Tengger Javanese

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by

Thomas J. Conners

Dissertation Director: Dianne Jonas

May, 2008

Abstract

TENGGER JAVANESE

Thomas J. Conners, Ph.D.

Yale University 2008

This dissertation provides a description and analysis of salient aspects of the grammar of Tengger Javanese. The data collected for and described in this work is the result of the author's year and half long fieldwork in the volcanic highlands of the Tengger Massif in East Java, Indonesia. The first part of the dissertation focuses on the history, sociolinguistic setting, and lexicon, phonology, and basic morphology of the dialect. The second and larger part of the dissertation describes and explores aspects of the morphosyntax and syntax of Tengger Javanese. The language is placed within both a synchronic perspective, though a comparison with other dialects of Javanese and other related languages, and also a diachronic perspective, through comparisons with Old and Middle Javanese. Issues relevant to Austronesian studies in general such as voice systems, word order, constituency, and applicatives are all addressed. The approach taken is empirical, with the primary goal of the dissertation to accurately explicate the complex linguistic patterns found within the Tengger dialect itself. Through the synchronic and diachronic comparisons, the Tengger dialect is placed typologically both within the Austronesian language family, but also comparatively with non-Austronesian languages.

BIOGRAPHICAL SKETCH

Thomas J. Conners received his BA in Linguistics from Yale University in 1995, writing his thesis on the syntax of Old Icelandic. He continued his studies in the Linguistics Department at Yale University, earning an MA and MPhil in 2000 and 2001 respectively, writing about the phonology and syntax of Indonesian.

In the summer of 2000, Tom took an intensive course and earned an advanced certification in Indonesian at the Universitas Sam Ratulangi, in Manado, Sulawesi. After several summers spent collecting data around the Indonesian archipelago, Tom lived in Ngadas village, East Java, Indonesia for a year and half in 2003-2004 on a Fulbright Scholarship, conducting intensive field work on Tengger Javanese, for his Ph.D. from Yale University.

Since February 2005, Tom has worked as a researcher for the Max Planck Institute for Evolutionary Anthropology, Jakarta Field Station, living full time in Jakarta, conducting a Javanese dialect mapping project.

This dissertation is fondly dedicated to the memory of my grandparents, Thomas J. and Eleanor P. Conners, who were here to see the beginning of it, and doubtless are proud of its completion.

ACKNOWLEDGEMENTS

I owe the deepest and warmest thanks to the Tengger people, and all the people I have met in Indonesia, who have welcomed me into their homes, shared their food, and spent endless hours cheerfully discussing their languages. The Tengger specifically and Indonesians in general are the warmest, most welcoming people, who invite you in and accept you as family. I hope in some way that this dissertation can repay in some small part the countless kindnesses which have been shown to me during all my years in Indonesia.

To my colleagues at the MPI EVA Jakarta Field Station I owe a great debt. I would never have finished this dissertation, or come to have the deeper understanding of Indonesian and Javanese that I do now without their help. To Antonia Sorriente, Betty Litamahuputty, Singgih, Dalan, Dini, Okki, Olly, Sarah, Tessa, Lani, Dian, Joni, Febi, Erni, Diana, Yanti, I thank you all.

In particular I need to mention both David Gil and Uri Tadmor, who not only gave me a job in Jakarta, but pushed me to finish my dissertation. If not for a series of conversations in a 'café' in Manokwari, Papua with them, I would never have completed this work.

The theoretical aspects of the dissertation owe an incalculable debt to many late night discussions in a blessedly warm apartment during the cold German winter of 2007-8 in Leipzig with David Gil. Without his expertise, encouragement, enthusiasm, and probing mind, this dissertation would never have become what it is.

I also benefited enormously from my conversations and long walks in Leipzig with Peter Cole, Gabriella Hermon, Jaklin Kornflit, and Andrej Malchukov.

My years spent at Yale were some of the best of my life. There are myriad people there who have helped me during various phases of my studies, and with the dissertation in particular. Specifically, I would like to thank my fellow graduate students, my many professors, and my friends, all from whom I have learned so much. Stanley Insler has been a constant friend and profound teacher.

To my dissertation advisor, Dianne Jonas I give my thanks.

I also owe my deepest thanks to J. Joseph Errington, who always managed to find time and provide exceptional insights into Javanese and Indonesian. I am humbled greatly by his vast knowledge of things Javanese, and I am grateful for his sharing some of that with me.

To my many other friends in the US and in Indonesia who have always supported me through all of my endeavors, I thank you all.

I would like to thank those who provided me with financial support over the years, including Yale University, the Council on Southeast Asia, the Graduate School, Fulbright, and Fulbright-Hayes

My family has been my greatest source of strength and encouragement throughout, Pop, Gran, Nora, Frank, Frankie, Tommy, and more recently Adam, I love you and thank you all.

The one person who, more than any other, gave me constant support, encouragement, advice; the only one without whom I would not be who I am today, and would never have achieved so much, to my mother, Chris, I owe the world. From

beginning to end, without fail, without doubt (of which I had plenty), she has always been there; she is my rock. Thanks mom, for everything!

TABLE OF CONTENTS

PART I: BACKGROUND

CHAPTER ONE: INTRODUCTION

1.1 Overview

This dissertation presents a description and aspects of the Tengger dialect of Javanese, spoken in the mountainous region of the Tengger massif in East Java, Indonesia, by approximately 30,000 speakers. The language of the Tengger is extremely threatened as this once isolated people have increased contact with lowlanders who speak the prestige variety of East Javanese. Further, there have been large scale conversions to Islam in recent years, which have weakened the connection of the Tengger to their unique code. While some past work has been conducted describing anthropological aspects Tengger life (Hefner 1985) and sociolinguistic aspects of their language (Smith-Hefner 1983), no significant work has ever documented, described, and analyzed formal aspects of Tengger language.

This dissertation provides a description and analysis of salient aspects of the grammar of Tengger Javanese. The data collected for and described in this work is the result of the author's year and half long field in the volcanic highlands of the Tengger Massif in East Java, Indonesia. The first part of the dissertation focuses on the history, sociolinguistic setting, and lexicon, phonology, and basic morphology of the dialect. The second and larger part of the dissertation describes and explores aspects of the morphosyntax and syntax of Tengger Javanese. The language is placed within both a synchronic perspective, though a comparison with other dialects of Javanese and other

related languages, and also through a diachronic perspective, through comparisons with Old and Middle Javanese. Issues relevant to Austronesian studies in general such as voice systems, word order, constituency, and applicatives are all addressed. The approach taken is empirical, with the primary goal of the dissertation to accurately explicate the complex linguistic patterns found within the Tengger dialect itself. Through the synchronic and diachronic comparisons, the Tengger dialect is placed typologically both within the Austronesian language family, but also comparatively with non-Austronesian languages.

So this is broadly a look at the Tengger language *in perspective*. It is not, nor does it attempt to be a grammar of the language. In some respect, it can be thought of as a 'best of', as I address those issues which have been particular crucial to Indonesian and Austronesian linguistics in general.

1.2 Organization of the Dissertation

The first two chapters contain historical information about the Tengger people and language, and about their broader geographic and linguistic setting. They also contain information about my fieldwork there. If one is primarily interested in the grammatical description, the reader is invited to skip directly to Chapter 3 section 2. If one is interested particularly in the description and aspects of Tengger syntax and morphosyntax, the reader is invited to skip directly to part II.

In chapter two I place the Tengger people and the Tengger language within their historical context, both from the earliest migrations of the Austronesians from the Asian

mainland and into island southeast Asia, and in their more recent history on the island of Java. I sketch elements of the basic grammar of the Tengger language, particular in terms of lexicon, phonology, and morphology in chapter 3. Also there I discuss the sociolinguistic setting of the dialect, and the unique position of Tengger among Javanese dialects in never having developed the elaborate speech level system that characterizes other varieties.

In the second half of the dissertation, I explore aspects of the syntax and morphosyntax of Tengger in much greater detail. Chapter 4 contains a discussion of word order and constituency in Tengger. I argue that Tengger has two stable word orders which are more frequent than others and which represent the shift from Middle Javanese VOS to a system prevalent in other Western Indonesian languages, SVO. Further, I argue that unlike in many colloquial varieties of Malay and Indonesian, Tengger makes a clear categorical distinction between substantives and predicates, though there are no unique syntactic categories such as preposition, adjective, etc. in Tengger. In chapter 3, I argue for the existence of a distinct NP in Tengger. There I also provide arguments for the existence of a VP, though it is unclear whether there are identifiable constituents other than an NP and VP in Tengger.

Chapter 5 presents the Tengger voice system, and provides synchronic comparisons with other dialects of Javanese and diachronic comparisons with Old and Middle Javanese. I place Tengger in a generalized typology of Austronesian voice, arguing that Tengger had until recently a symmetric voice system, but that recent contact with lowland dialects has introduced an asymmetric construction, and so Tengger currently has a mixed voice system.

Verbal morphosyntax is discussed in great detail in Chapter 6. Mood, aspect, and applicative markers are all taken into account, and a novel analysis is provided. I argue that the Tengger system, while simpler than that found in other dialects, actually represents the more original system, and has not undergone as process of simplification, as has been argued elsewhere. Further, I show that the Tengger paradigm has in fact added person distinctions where none had existed.

Finally, I provide a general conclusion in chapter 7.

1.3 Field Work and Data Collection

The primary data for the present dissertation come from a year and half of field work I conducted in East Java from June 2001 to September 2001 and again from June 2002 to September 2003. During this time, I spent over three months in the Tengger village of Ngadas and one month in the Tengger village of Ranu Pane. The Tengger region is situated around the active volcanoes of Mt. Bromo and Mt. Semeru, and falls into four distinct administrative districts or *Kebupaten*: Pasuruan, Malang, Lumajang, and Probolinggo. Tourists traveling to the caldera in Mt. Bromo almost always follow the paved road from Probolinggo up the northeast side of the mountain range. The southwest area of the region, encompassing both of the villages under investigation here, remains the most isolated and difficult to access.

I relied on two methods for the collection of primary data. Initially, consultants were given a survey that included very simple questions on a number of sociolinguistic issues such as language/dialect use, perceptions towards the dialects, and self-assessment

of fluency and literacy in the local dialect, in the lowlands dialect, and in high/low speech levels. These surveys also included modified and expanded Swadesh lists (see appendix I).

The bulk of the data used in the present dissertation was collected, with permission, from recordings of spontaneous conversation, and natural daily interaction of villagers. As any field researcher will attest, this method is the most effective way in which to obtain accurate, reliable data of actual patterns occurring naturally in a language. An elicitation method of data collection has significant drawbacks for any researcher, this is especially true for a non-native speaking outsider, and the problem is particularly acute given the current linguistic situation in Indonesia.

That being said, it is also necessary to conduct elicitations in order to examine in greater detail phenomena which might otherwise not occur in recordings. Wherever possible, I have taken examples from the spontaneous speech recordings, and only used the elicitations as a guide. For most of the recordings, the context is clear, straight forward, and not in need of further explication. However, when a particular example is not clear, I have provided the context of the utterance immediately following the English translation line, in square brackets.

Most Indonesians, regardless of their own ethnicity or native language, upon meeting a foreigner will attempt to assess if the visitor can speak any Indonesian (a few words is generally more than enough to elicit praise of one's fluency in the language!). Indonesian is the national language, and is used for education, law, government, in the media, and on TV, among other functions. Although it is not the native language of the majority of the population, latest estimates hold that there are 20-30 million native

speakers in a population 240 million (CIA Factbook); the vast majority of the population has a high degree of fluency in Indonesian. It is often the language of inter-ethnic group dialogue, and certainly the language with which most foreigners would be acquainted.

However, when one tries to engage in conversation in a local language there are several difficulties that emerge. The first is that Indonesian is very markedly the prestige language throughout the archipelago. There is great impetus then, for locals to show their education, urbaneness, and general prestige by using Indonesian as opposed to a local language or dialect when dealing with outsiders.

The problem is further exacerbated with the Javanese case. As noted in chapter 3, the very complex system of speech levels in Javanese arose over the course of several centuries; the language itself – through word choice, usage, etc. – directly reflects this culturally defined, exceedingly intricate pattern of social role and status of speech act participants. These patterns are exclusively defined by in-group participants. Non-Javanese, let alone foreign non-Indonesians do not traditionally figure into this cultural in-group definition. Status, and hence language choice among the Javanese is determined by family, history, profession, age, and position within the village. Since outsiders do not have a clearly defined role in these terms, and as Indonesian has spread recently, there is an overwhelming tendency among Javanese to use Indonesian when interacting with non-Javanese. It has also recently become a strategy when interacting with other Javanese with whom one is newly acquainted or not yet familiar. Language use and speech level choice among the Javanese in a period of high bilingualism is in fact a very complex and important question, but falls outside the general scope of the present discussion. That is, except to note the general difficulties in actually collecting natural speech data.

Given that the Javanese are very conscious and self-aware of their language and its usage and implications, this fact further complicates any method relying on elicitation as a primary source of data. After overcoming the obstacle of using Indonesian as opposed to Javanese, there is a general tendency among all Indonesians when asked either to clarify a sentence or if a particular construction is grammatical to respond with what is perceived as the 'correct' or 'textbook' answer instead of what is actually used. This risk with an elicitation method is certainly not restricted to Indonesian, it is however more prevalent than with many western language speakers (perhaps because of the wide gap between the 'standard' and colloquial registers in Indonesian and Javanese, and a number of other Indonesian languages). Western languages tend to have greater standardization across regional varieties than Indonesian or Javanese. Indonesian was first adopted as a national language in 1928[1] and as an official language in 1948. Since then, there have been a number of efforts to expand, promote, and standardize the language, including an annual 'language month'. This has to some extent been successful, but it has also often led to a greater divergence between the 'standard', educated, official language, and the quotidian vernacular. The divide is further heightened by the existence of a number of non-mutually-intelligible regional Malay dialects, and perhaps as many as 500 distinct other languages throughout the country, which each uniquely influence vernacular Indonesian in a particular region.

For Javanese, the language of the two Central Javanese court cities of Yogyakarta and Surakarta act as 'exemplary centers' (Errington 1986). These two varieties are considered the most *alus* or 'refined', and are thereby the most prestigious. They stand in opposition to the East Javanese dialect, referred to as a whole although there are any

[1] As expressed in the *Sumpah Pemuda* 'Youth Declaration.'

number of regional varieties, is referred to as *kasar* or 'crude, crass'. The speech levels in East Java are less intricate and less rigidly adhered to, especially in the coastal areas which are heavily influenced by Madurese. In fact, many East Javanese who travel to either of the 'exemplary centers' will often nowadays use Indonesian so as to avoid offending an interlocutor with an improper speech level or to avoid sounding parochial and uneducated. However, to a lesser degree than Indonesian, Javanese has also been subject to a period of government standardization. Beginning with the Dutch colonial government in the 19th century, attempts to document, describe, standardize, and teach Javanese took on a great role, as the colonial mentality held that natives, except the truly elite, should not be allowed to learn Dutch (Sneddon 2004). Colonial administrators therefore had to learn Javanese, and the two Central Javanese courtly dialects were elevated as the 'standard'.[2]

In my first attempts to collect data in Malang, East Java, I would have conversations for up to 10 or more minutes where I would speak only in Javanese and my interlocutor would respond only in Indonesian! Having eventually overcome that hurdle, when I would try to elicit responses from consultants, they would typically respond with some version of what they thought the Central Javanese version of a particular clause would be, leaving out all elements clearly regional. Eventually I employed the services of a local college student, who with some training became a very efficient recorder. Most of the data from Malang are from recordings completed by Mario, a native Malang resident, without the conspicuous and often biasing presence of a *londo* or 'Hollander—i.e. white man'.

[2] The history here, and ultimate selection of a 'standard' is actually much more complicated, but is not directly germane here.

The situation in the Tengger villages is somewhat different to that encountered in the lowlands. First, there is much less Indonesian-Javanese bilingualism among the Tengger, this is especially true of Tengger women and older villagers. Most men under 40 and a majority of children—most of whom attend at least a few years of school where Indonesian is taught—have a fairly high degree of fluency in Indonesian. This means that the Tengger are less likely that other Javanese to resort to Indonesian with outsiders. Second, for the Tengger the 'exemplary centers' are not the far off Central Javanese courtly cities, but rather the low vernacular of the surrounding lowlands. Those villagers who are comfortable with *krama* or high Javanese, actually speak what is referred to as *krama desa* or 'village krama'. *Krama desa* is a highly colloquial, regionally variant, form of the high 'standard' Javanese, generally spoken in rural areas; the term is usually used by urban central Javanese as a derogatory term. Given that the speech level system is centered and most clearly articulated around the old courts, with all of their associated social stratification, which is largely absent in a rural village setting, *krama desa* is an attempt by rural villagers to imitate the high courtly language. It very often mixes lexical items from different speech levels and incorporates regional and often Indonesian vocabulary, which is difficult to differentiate from the equally foreign *krama* vocabulary and therefore confuses elements of different speech levels. For example, it is not uncommon to hear a villager use a *krama inggil* (honorific and strictly 2nd or 3rd person referent) element in reference to himself or a *krama andhap* (humbling and strictly 1st person referent) element in reference to another person, both strictly prohibited by the rules of the speech level system itself (see chapter 3 for a very detailed discussion of the speech level system in Javanese and Tengger).

It is only the Tengger men who ever use *krama desa*, and then generally only those who serve or have served in some bureaucratic position within the village, and only when dealing with general bureaucratic affairs or when speaking to non-Tengger Javanese. So, although the difficulties associated with bilingualism and the speech level system found in the lowlands were not present among the Tengger, there was another unforeseen obstacle. Unlike the lowland Javanese, who are bound by a very complex cultural perception of 'hospitality'—very similar to that often discussed for Indo-Europeans, or that found in the sagas of the Icelanders—the Tengger tend to be rather xenophobic and suspicious of outsiders. So, whereas in the lowlands it was generally easy to gain access to consultants and participate actively, it took some time for the Tengger villagers to become comfortable with my presence in the village. In fact, I had to reassure them at a village meeting that I was not working for the CIA (why the CIA would be interested in isolated cabbage and potato farmers is beyond me!).

It was also impossible to train an assistant in the village to help collect the data. There were also no lowlanders sufficiently fluent in the dialect and familiar with the village (so as to avoid being given the cold shoulder by the villagers) who could act as assistants. So, after a period of about a month in the village, after I had more or less become a common enough sight and had spoken with most of the villagers, I was able to begin collecting samples of natural speech production.

This is not to say that no elicitation was conducted, as mentioned above. Working with several consultants in Malang and Ngadas, I spent a significant amount of time, especially near the end of the research project, completing surveys and elicitations related to a number of different, generally syntactic and morphosyntactic phenomena.

The languages described in this dissertation then are based on over 100 hours of field recordings made in Malang, Ngadas, and Ranu Pane, between June 2002 and September 2003, and subsequent clarification sessions. From these recordings, I have over 800 pages of usable transcription and translation. By usable I mean both that the audio quality of the recording was sufficient for accurate transcription, and that the dialogue recorded was of sufficient length to allow for context in translating the recording[3].

1.4 Methodology

I do not approach this project with any one particular theoretical framework to defend or rebuke. This is not to say that do not have a particular theoretical background in which I work and was schooled. This dissertation is more descriptive in nature, focusing on uncovering salient linguistic patterns attested in actual language, exploring the related and conflicting patterns in adjacent languages, and ideally pointing to some historical basis for the genesis of both the patterns themselves and explicating the divergences.

My own personal background in linguistics is within generative syntax and historical linguistics. I do share a number of basic assumptions about linguistic systems in general with from this tradition, and that will become clear throughout the dissertation. However, as with many linguists who stray outside of the core generative areas of English, Romance, Germanic languages and to some extent Japanese, much of the theory

[3] These recordings are maintained on both mini-disk and cd-rom, and they are fully transcribed and translated in Word© documents. I would like to make them available to any interested parties, who are invited to contact me directly.

seems less suitable for accurately describing the structure of the language, especially when confronted with empirical evidence from such divergent languages.

The description in this dissertation is driven by the facts of a particular language: Tengger. When I discuss a new phenomenon, I present the Tengger data first and foremost. When relevant, I also discuss the patterns found in the lowland dialects, the standard language and make comparison with other Indonesian languages, Austronesian languages, and beyond. When a particular diachronic issue is raised, I discuss data from Old and Middle Javanese as well. So, what will hopefully emerge is a picture of a language in perspective, both synchronic and diachronic. Where the data is particularly relevant to some other theoretical issue, be it formal, functional, typological or otherwise, only then do I make reference to any individual framework. I intend this study to be language driven, not theory driven.

CHAPTER TWO: HISTORICAL INTRODUCTION

2.1 Introduction

With over 235 million people spread out over 17, 000 islands representing some 500 plus ethnic groups, Indonesia is one of the most diverse countries on the planet. The small island of Java, roughly the size of Louisiana, is home to over half of Indonesia's 235 million people. It is the most densely populated area of comparable size in the world. There are three main languages spoken in Java. With about 13 million native speakers, Sundanese is the predominant language of West Java, excluding the capital area of Jakarta. In East Java and on the island of Madura, Madurese is the native tongue of roughly 8 million ethnic Madurese. However, with an estimated 80-90 million native speakers in Central and East Java, Javanese is by far the most widely spoken *bahasa daerah,* or 'local language' throughout the entire archipelago.

Javanese is the native language of some 90 million speakers spread across the globe, which places it among the 15 most widely spoken native languages in the world (number 11 according to the Ethnologue 2005). As a result of the Dutch colonial legacy, there are sizable populations of Javanese speakers in Suriname and the Netherlands. Within the Indonesian archipelago itself, through centuries of conquest of the outer islands, and more recently through government programs of *transmigrasi* or 'transmigration—the internal relocation of peoples from high density areas to sparsely populated areas', small pockets of Javanese speakers can be found on almost every island within Indonesia as well. Within Java itself, however, there is great variation among

dialects. Typically the dialects of the Yogya and Solo exemplify what is held to be the 'standard' language. However, there are numerous other dialects on the island that have often been overlooked by scholars. From Banten in West Java, to Banyumas in Central Java to Banyuwangi and Osing in East Java, many of these 'dialects' are not mutually intelligible and most have received little if any scholarly treatment.

Most scholars who study Javanese focus on the courtly 'dialects' of the Central Javanese cities of Yogyakarta and Surakarta, or Solo. After the fall of the Majapahit Empire, the last great Hindu-Buddhist kingdom of insular Southeast Asia, near the end of the 15th century, Yogyakarta and Surakarta arose as the preeminent polities in the region.

In this chapter, I will give a very broad history of the Javanese language and help to set the historical context for the particular dialects under investigation here, namely eastern Tengger and Malang Javanese. It is crucial to understand this context before continuing with a far more specific and detailed analysis of these dialects. Beyond the natural geography of Java, the rise and fall of various empires, the confluence of competing religions, and the contact with external peoples and languages are factors that have had the most significant influence on the development of these two dialects. In order to understand the differences and similarities between these two dialects, we need first to explore how and why the speech communities diverged, the historical pressures that came to bear in establishing the circumstances which led to the divergent evolution of the two languages.

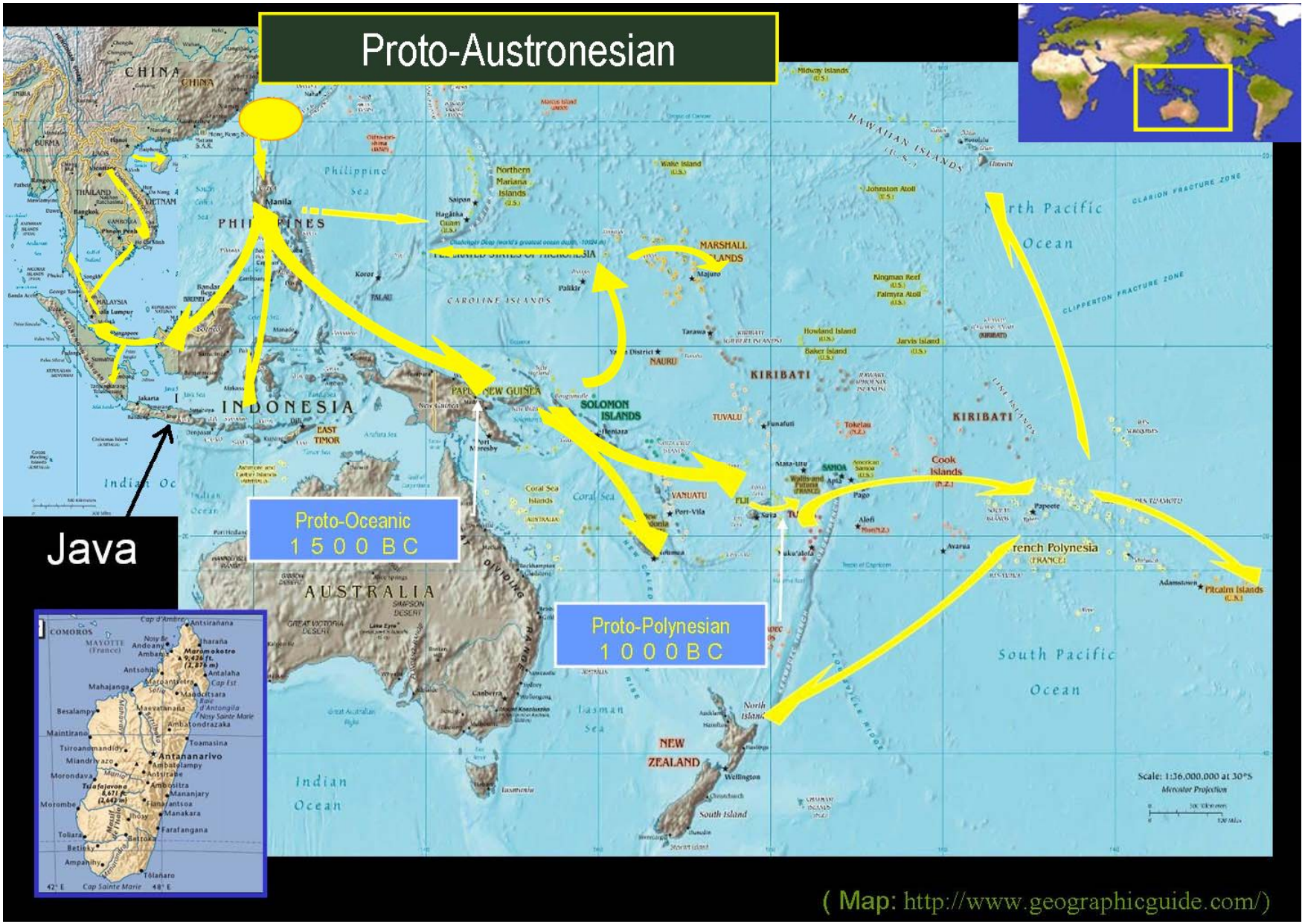
First, I discuss the broader context of the place of Javanese within the Austronesian language family.

2.2 Austronesian/Western-Malayo-Polynesian Foundations

The Austronesian language family includes some 1500 plus languages spread from Malagasy on the island of Madagascar off the east coast of Africa, to Hawaiian on the Hawaiian islands in the eastern Pacific ocean, to Maori in New Zealand in the far South Pacific, to the indigenous Formosan languages of Taiwan, off the East coast of China. It is generally accepted that the Austronesian homeland was either on the southeast coast of mainland China or on Taiwan itself. Based on linguistic, ecological, botanical, and archeological evidence, Bellwood (1997) takes the view that Taiwan was settled by colonists from the Chinese mainland sometime at the end of the fifth millennium BCE or the beginning of the fourth millennium BCE. For several centuries, or possibly up to a millennium, no further expansion took place. Then in the third millennium BCE the northernmost island of the Philippines, Luzon, was settled by colonists from Taiwan. This marks the most significant linguistic division within the Austronesian language family. According to Blust (1985), Lynch (1998), Nothofer (1996), Wolff (1996), and many others, there are four first-order sub-branches within Austronesian. More recent work has argued for even more first order branches within Austronesian, all of them located in Taiwan and representing Formosan languages.

Three of the four, Atayalic, Paiwanic, and Tsouic are located in Taiwan. The fourth first-order sub-branch, Proto-Malayo-Polynesian, or PMP, encompasses all of the Austronesian languages found outside Formosa. All of the languages of the Philippines, Indonesia, and Oceania, here a cover term used to incorporate Melanesia, Micronesia, and Polynesia, Peninsular Southeast Asia and Madagascar, are descended from a

common ancestor, PMP. Some form of PMP was spoken during the third millennium BCE on the northern Philippine island of Luzon.


Proto-Austronesian
Java
Proto-Oceanic
1 5 0 0 B C
Proto-Polynesian
1 0 0 0 B C
(Map: http://www.geographicguide.com/)

From Luzon, expansion continued in many directions into the southern Philippine islands, Borneo and Sulawesi. Bellwood claims that by at least 2000 BCE PMP had begun to break up (1997:241). Eastern Borneo, Sulawesi, the Malukus, Nusa Tenggara, Bali and Java were all settled during the period 2500 BCE to 1500 BCE. Further expansion continued to the east, with Melanesia being settled from 1500-1000 BCE, Micronesia from 500 BCE to 500 CE, and Polynesia from 300-1200 CE. To the west, Sumatra, western Borneo, and peninsular SEA were settled from 1500-500 BCE, with Madagascar following some 1000 years later.

Proto-Malayo-Polynesian has three subgroups: Western-Malayo-Polynesian, Central-Malayo-Polynesian and Eastern-Malayo-Polynesian. This latter group has two further primary subdivisions: South Halmahera-West New Guinea and Oceanic. Western-Malayo-Polynesian includes the modern languages of Sumatra, Borneo, Java, Madagascar, west Nusa Tenggara, and peninsular SEA. Nothofer (1973) is one of the first systematic attempts to further classify the languages of the Malayo-Javanic sub-branch of Western-Malayo-Polynesian. The primary languages within this subgroup are Malay, Javanese, Sundanese, Madurese, and Achenese, spread over the islands of Borneo, Sumatra and Java, and the Malay Peninsula. Proto-Malayo-Javanic is the most immediate proto language from which these languages stem. Recent work (Adelaar (2005), Tadmor (2007)) has argued, though, that Javanese in fact branched off earlier, and forms its own subgroup, with Malayic (Achenese, Chamic and Malayic languages), Sundanese, Madurese, Sasak, and languages of Sumbawa forming another group.

2.3 A Short History of Java, Tengger

The history of Javanese is generally broken down into three periods: Old, Middle, and Modern Javanese. This does not, however, represent a linear progression of the language but rather three distinct forms of the language that had different, but overlapping periods of use. In fact, there are surviving documents and inscriptions from the fifteenth century in each of Old, Middle, and Modern Javanese**.** Based on the criteria in Zoetmulder (1994), Modern Javanese are those forms of the language that have undergone contact with Arabic; they are the sundry varieties of Javanese spoken on the island after the initial contact with Islam. Old and Middle Javanese, by contrast, are pre-Islamic varieties of the language, which show no influence from Arabic. Old Javanese is that form of the language found in the *Kakawin* literature, and attested from the 9th century onward. This was the language of the old Central Javanese kingdoms, including the original Mataram. It is likely that this was already a somewhat fossilized literary language by the turn of the millennium. Middle Javanese is characterized by the *Kidung* style of literature. It was centered in East Java, and was most likely the spoken language of the Majapahit Kingdom. Simultaneously however, Old Javanese was also present at the Majapahit court, but apparently as a literary and courtly language, and not as the common spoken language.

Among Austronesian languages, Javanese has the oldest surviving manuscripts and the earliest flowering of a literary culture.[4] Many charters and stone inscriptions have been found throughout Java and in the surrounding islands. The oldest extant

[4] There are earlier inscriptions from the Malay speaking Sriwijaya Empire of Sumatra. However, these are limited in number, do not represent a literary tradition, and for the most part are composed in Sanskrit.

inscription in Old Javanese is the *Sukabumi Charter from* March 25th, 804 CE.[5] Most Old Javanese writings where preceded by a *Manggala.* This included an often complex and elaborate method of the marking the date of composition according to a number of different calendrical systems, which allows modern scholars to determine the date of composition with great accuracy. The oldest inscription from Java dates to 732 CE.[6] All inscriptions found in Java from that point on until the *Sukabumi Charter* are written in Sanskrit. The *Sukabumi Charter* marks the great departure from the practice of writing exclusively in Sanskrit and the switch to the local vernacular. Very few inscriptions in Sanskrit are found after 804 CE.

It is not clear, however, whether Old Javanese at this point represented a true vernacular language in common use among most Javanese as a courtly language, used by the royal family and their attendants, or a truly literary and scholarly language, much as Sanskrit was in India.

Along with the rest of SEA, Java was very heavily influenced by the Hindu-Buddhist culture from India. Contacts with India probably began in the first or second century CE. The length, depth, and significance of Indic influence on Java is attested to in the magnificent Buddhist temple of Borobodur and the Shivaite temple of Prambanan, along with hundreds of other smaller Hindu and Buddhist temples spread throughout the island. That influence is also seen through the literature of Old Javanese. From the ninth century CE to the 15th century CE, hundreds of literary works were composed in Old Javanese. Most are based on Indic sources of one sort or another, such as the Ramayana or Mahabharata. There are several different genres within Old Javanese poetry, including

[5] van Stein Callenfells, 1934.
[6] Krom 1931:123.

Parwa, the oldest, and the Kakawin. These are all based on Indic myths, epics and stories, and most are composed in using Indic verse conventions and Indic meter. Old Javanese was in continuous use as a literary language from the ninth century through the 20th century. However, the last texts composed in Old Javanese to be found on the island of Java itself date from the 15th century and the fall of the last great Hindu-Buddhist kingdom of Java, Majapahit. Thereafter, the language was maintained in both Bali and Lombok as a literary language.

Geography has played a rather significant role in influencing the settlement and development of Java. It is also a significant factor in the modern distribution of Javanese dialects, a topic we will return to later. Java has a spine of plateaus, mountains and active volcanoes running east-west through the center of the island. Between the mountains and volcanoes lay isolated plains that include some of the most productive and fertile arable lands in the world.[7] Although relatively little is known about the political organization of Java before the 11th century, it appears that this geography was conducive to small independent kingdoms and chiefdoms, with no major or overarching polity forming until the 13th century. The Majapahit Empire (1294-1478?) was based along the banks of the Brantas River in east Java, and grew out of the earlier Singosari Empire of east Java. There are a number of inscriptions in Old Javanese, the Old Javanese text *Nāgarakĕrtāgma,* written in 1365 (though found only in later manuscripts), and the Middle Javanese text *Pararaton* (surviving only in later Balinese redactions) which provide most of the information on the Majapahit empire.[8] At its zenith (1350-1389 CE),

[7] In fact, Java was the world's largest rice producer until the 19th century (Ricklefs 2002:20).
[8] Ricklefs (2001:20) notes that the Dutch scholar of Javanese, C.C. Berg, casts doubt on the historical accuracy of these sources, arguing that they should be seen rather as 'supernatural' documents, '…to be understood within the context of politico-religious myths which the authors of these records were

it controlled the whole of Java, Bali, parts of Sumatra and other smaller islands of the archipelago and had contacts with Vietnam, China, Cambodia, Siam, and Champa. Both Buddhism and Hinduism were practiced within Majapahit. In fact, it is here where the greatest mixing of Hindu, Buddhist, and traditional Javanese beliefs took place, and the syncretic practice later called Hindu-Javanese arose (see Hefner 1985 among many others).

During the 15th century, the inland based Majapahit began to lose control of the primary trading ports along the north coast of Java. Small Muslim principalities arose all along the north coast including Cirebon, Jepara, and most significantly in Demak. As their power grew, and their influence began to spread inland, Majapahit almost completely collapsed. It seems though that the fall of Majapahit in 1478 was not to a major Muslim principality, but rather some other smaller Hindu-Javanese state, or perhaps due to civil war within the empire itself. Majapahit did continue to exist, albeit in a much smaller and weaker state, until c. 1527, when it finally fell to Demak.[9]

Much of the Majapahit nobility fled to Blambangan, in the far eastern salient of Java, and to Bali. In the second half of the 16th century, the Muslim Mataram emerged as the preeminent power on Java. Based in the interior of Central Java (a shift from Majapahit in East Java), reached its peak in the early to mid-17th century, by which time the vast majority of people in Java had converted to (a syncretic form of) sufistic Islam. The Hindu-Javanese courts of Blambangan were an exception to this until they came under the influence of Mataram in 1640. After that time large scale conversion took

---

concerned to support.' Prof. Berg's arguments, however, are not supported by most scholars, and Ricklefs argues that some certainty about the history of Majapahit is certainly possible.

[9] Cribb 2000.

place in Blambangan. By the end of the 17th century, only small, isolated pockets of people still practiced the traditional beliefs of Hindu-Javanese.

It is important to note the strongly syncretic nature of most Javanese beliefs, including those influenced by Hinduism, Buddhism, and Islam. Java has a long history and tremendous capacity to assimilate and nativize external influences. This is seen in religion, politics, arts, law, culture, writing and language, among other aspects. The effect of four centuries of Islamic and Arabic influence is seen equally in Java's religion and language. It will become increasingly apparent later when we compare the lowland dialects of East Java, which are spoken by a largely Islamic population and therefore heavily influenced by Arabic, with the highland Tengger dialect, which to this day remains very strongly Hindu-Javanese.

The primary differentiation between Old and Middle Javanese on the one hand, and Modern Javanese on the other, has been characterized by the presence/lack of Arabic/Islamic influence.[10] Alternatively, Modern Javanese can be characterized by the lack of much Sanskrit morphological material present in Old and Middle Javanese. Old and Middle Javanese are those languages which existed in Java prior to the 15th/16th century conversion to Islam. However, as noted above, when the Majapahit Empire fell, much of the Javanese nobility fled to Bali. On Bali, Hindu-Javanese practices developed a distinctly Balinese flavor. But there can be no doubt that the basis for the Balinese Hinduism of today stems from the culture, language and beliefs of 15th century Hindu

[10] Zoetmulder.

Java. In fact, works in Old and Middle Javanese continued to be composed in Bali into the 20th century.[11]

Most of the literary works of Old and Middle Javanese were for centuries written on *lontar*, or palm leaves. Given the moist tropical conditions of Java and Bali, these manuscripts had at best a life-span of about 200 years. This means that texts had constantly to be transcribed again and again. Apart from the problem of accurate transcription and transmission—which I will address again below when discussing the *Sudamala*—this means that most copies of Old and Middle Javanese texts survive only from later redactions in Bali. Those that do exist from Java are often adulterated to better coincide with Islam.

Old Javanese then, was the pre-Islamic language of *Kakawin* poetry, which survives largely in copies of texts from Bali, and the prose inscriptions from Java itself. It is clear that for much of its 1100 year history, Old Javanese functioned primarily as a literary language, the domain of elite scholars, poets, and scribes associated with royal courts.[12] Middle Javanese, on the other hand, is thought to represent a common spoken language, during the Majapahit rule, in Central and East Java. Middle Javanese is characterized by a novel genre of poetry called *Kidung*, as well as prose works and. Where *Kakawin* poetry is based on Indic stories, legends and epics, and is composed in

[11] This practice has probably mostly died out in Bali today. However, in many cultural and religious ceremonies in contemporary Bali Old Javanese texts, epics, and prayers can still be heard. In fact, Bali is a primary center for the study of Old Javanese.

[12] Some scholars dispute this, claiming that literacy was rather widespread in Java before contact with Europeans. There are a number of reports from early European contacts with Indonesia stating that literacy was widespread among both men and women. This argument is further supported by the fact that inscriptions found on Java do not correspond to the boundaries of any known polity, and that there is a great deal of variety among inscriptions found within single states, such as Majapahit. Further, some claim the Indic-based scripts of pre-Islamic Indonesia are more suited for writing on palm or paper, as opposed to stone inscriptions, suggesting a rather wider audience than royal courts alone. Cribb, 2000:39.

Indic meters, *Kidung* poetry is of a more syncretic Hindu-Javanese content and, significantly is composed in native Javanese meters.

Apart from the differences in style, Old and Middle Javanese are quite distinct languages, probably not mutually intelligible. As noted above, the terms 'Old' and 'Middle' themselves here are misleading. They do not refer to a linear chronology in a single language. Rather, it seems that for some time Old Javanese coexisted with Middle Javanese, with the former being a petrified literary language and the latter either a quotidian courtly or common language. In fact, there are different manuscripts dating from the 15th century found on Java in Old, Middle and Modern Javanese.

As noted above, Modern Javanese is generally defined as the post-Islamic language ethnic Javanese. It is however a cover term used to describe a broad range of temporally, geographically and religiously distinct varieties. What is true about each of the three broad categories of Javanese entailed by the terms Old, Middle, and Modern, is that they are largely similar within themselves and diverge greatly in reference to the others. That is, although Modern Javanese refers to many different dialects, spoken across the island of Java (and on other islands and in other countries); these dialects share greater typological affinities to each other than either does to Old or Middle Javanese.

There are a certain number of poetic and prose texts from East Java, written in Middle Javanese, that are thought to represent the spoken language of that area some 500 years ago, which eventually developed into the modern eastern Javanese dialects. Among these is the *kidung Sudamala.* This text will form the basis for the historical comparison of the several dialects under consideration here. I will explore this in much greater detail in a later chapter.

Most scholarly works written on the Javanese language almost invariably concentrate on the Central Javanese dialects of the 'exemplary centers' of Yogyakarta and Surakarta. In fact, until the beginning of the 20th century, the courtly language of Surakarta was considered to the most *halus* or 'refined', and therefore the most prestigious. However, during the course of the last 100 years, the courtly language of Yogyakarta has also gained in prestige, to a point where the two are held in equal esteem today. Even among the many works written about these dialects, most have focused on the literary, sociolinguistic, ethnographic and anthropological aspects of the language. Very little work has been done on the grammar of Javanese itself. With the exception of several works by Cole & Hermon (2000, 2003, 2004, inter alia) and Davies (1999, 2003), Javanese has remained completely outside the purview of current theoretical linguistics

Modern Javanese, then, has two primary dialects, what we may call Central and Eastern. Central Javanese is typified by the dialects of Mataram's capitals of Surakarta and Yogyakarta, whereas the dialects of Surabaya and Malang are representative of Eastern Javanese. These too are just cover terms, as there are a number of typological differences within the Eastern and Central Javanese dialects themselves. The variation is mostly in the domain of lexicon; however there are some significant differences in phonology, morphology, and, to a lesser extent, syntax. Almost every city or region within Java is home to some unique linguistic features. However, there are several regions whose dialects stand out, primarily in the number of archaisms and retentions present. Of note are the Banten dialect of West Java, Banyumas in Central Java, and the Tengger and Osing, or Banyuwangi dialects of East Java.[13]

[13] The population numbers here are slightly out of date. The CIA World Factbook estimates there are 94 million Javanese speakers in Indonesia, as of 2007. www.cia.gov

Provinces	% of TP	Total Population	Number of Speakers
Banten	>5% 9.000.000	> 500.000	
Jakarta	n.a.	10.000.000	n.a.
West Java	15%	30.000.000	5.700.000
Central Java	97%	34.000.000	32.980.000
Yogyakarta	80%	3.500.000	2.800.000
East Java	78%	38.000.000	30.400.000
Lampung	80%	7.000.000	5.600.000
Total	a	139.000.000	77.980.000

Banten, which lies on the far west side of the island of Java, was settled by Javanese Muslims during the 16th century, coming from the Central Javanese Mataram Empire (by way of Cirebon). Banten Province, which only came into existence in 2000, is separated from Central Java by the Province of Western Java, which is the heartland of the Sundanese who predominantly speak Sundanese, which is a closely related but distinct language. The isolation of the Javanese settlers of Banten from other Javanese speakers and their intense contact situation with Sundanese as well as speakers of Betawi has caused this dialect to diverge rather strikingly from the 'standard' language of Central Java.

The Tengger and Osing dialects differ from Banten in that their speakers were Hindu-Javanese. As mentioned above, when the Majapahit Empire fell near the end of the 15th century, many Hindu-Javanese fled to the eastern salient (*Oesthoek)* of the island of Java and to Bali. Although the Kingdom of Blambangan didn't survive past the 17th century, there are a number of people who still practice traditional Hindu-Javanese customs in *Kebupaten Banyuwangi*, or Banyuwangi District, in the far east of Java. Being outside of the sphere of the last major Javanese empire, Mataram, and in contact with neighboring Balinese speakers, the Osing dialect also followed a unique path, and survives as a distinct dialect today.

The popular story of the original Tengger follows a similar track to the Osing dialect. When Majapahit was collapsing at the end of the 15th century, many people fled into the high mountains west of Malang, today the Tengger-Semeru-Bromo National Park. This mountain region is densely forested and very difficult to access. The Tengger people lived in relative isolation for a number of centuries. Several of the highest Tengger villages—which are the highest in Indonesia outside of Papua—still do not have paved roads. During my time in the Tengger region, two villages got electricity for the first time, and the bottom of the road leading up from Malang to Ngadas and Ranu Pane were paved with macadam for the first time, though the majority of the road remains a dusty, bumpy, steep, and arduous dirt 'road'.

However, it is clear that the Tengger have lived around Mt. Bromo at least as far back as the 14th century, and very probably much earlier. In 1880, a Javanese woman in a field in the district of Penanjangan in the Tengger area found a bronze plate (along with a bronze container for slaked lime used in betel chewing shaped like a phallus). The

inscription on this plate is the Charter of Walandit. It is a copy, made in 1405 CE, of an older charter ascribed to King Hayam Wuruk of Majapahit, who is referred to in the Charter by his posthumous name of *Wekas-ing-suka*, or 'Limit of Bliss'. According to the Middle Javanese text *Pararaton*, King Hayam Wuruk died in 1389 CE.

The Charter of Walandit also has a number one engraved on it. According to Brandes (1899), this probably indicates that this Charter is the first one of a series of plates containing the texts of charters referring to the Tengger districts. However, no other plates are known today.

Charter of Walandit

1381/1404 CE

Recto,

1. It shall be known by those who act as claimants of *titileman* (dues claimable at new moon) in Walandit, that Our Word is
2. in respect of that district of Walandit, Mamanggis-Lili, Jebing, *Kacaba*, inasmuch as they shall be exempt in respect to the
3. claimants of *titileman*, because of that district of Walandit is a *hili-hila* (sacred) district, of Spirits' servants, (worshipping) the honored holy
4. mountain Brahma (Mt. Bromo in the Tengger Massif), therefore now shall have an interdict all people living in the shade of the district of Walandit:
5. no *titileman* shall be claimed of them, let they not be examined, because the district if *hila-hila* (sacred). This Royal Seal

Verso,

1. when it is read shall be kept by those people of Walandit. Dated in the fifth month (*Margashirsa,* ~Nov/Dec), in the year '3. // in the Shaka-year
2. 1327, in the month *Asadha* (June/July), dated the ninth of the waning moon, *Pahing* (a day of the five-day week calendar), *Radite* (Sunday) in the week *Dungulan (Galungan,* the 11th week of the 30 week *wuku*-year), that was the time that
3. for the common families of Walandit was made a bronze plate of the Royal Seal, (containing) the Word of His Grace the Lord the Holy Wekasing-Suka (Limit of Bliss) because it is a *hila-hila* (sacred) district
4. of Spirits' servants (worshipping) the honored holy mountain Brahma. That was the reason that a bronze plate was ordered to be made for them by *kabayan* (beadle) Made, *buyut* (headman)…
5. …..

The area of Walandit is also mentioned in the Nagarakertagama, and de Casparis (1940) notes that the Charter of King Sendok of 939 CE also mentions Walandit as an important religious district. This means that the Tengger have been living in and around Mt. Bromo from at least the 14th century, and quite possibly as early as the 10th century CE.

The language of the Charter itself is interesting. The style and spelling of the Charter are the same as in a number of other charters of the same time that do not belong to the court sphere. This seems to have been a transitional period between Old and Middle Javanese. In its phonology, for example, the Old Javanese diphthongs /ya/ and

/wa/ have been collapsed to the Middle Javanese /e/ and /o/. However, the verbal suffix –*a* clearly has a future, irrealis reading, as in Old Javanese.

The Charter of Walandit then is most significant here because it confirms that the Tengger have lived in and around Mt. Bromo for at least 700 years. The Tengger still practice the rights and rituals that exempted them from taxes some seven centuries ago.

Given that the Tengger Massif is extremely isolated from the surrounding lowlands areas, and remains that way in some parts today, the divergence of their language comes as little surprise. And unlike the Banten and Osing dialects, there is no contact situation among the Tengger which has influenced the direction of change.

Although very little work has been done on Javanese dialectology, there are a number of native works concerning these dialects. These studies tend to address a single topic such as verb reduplication, or nominal morphology. No comprehensive study of any Javanese dialect other than Yogyakarta or Surakarta has yet been conducted. We turn next to a closer look at the language of the Tengger themselves. I present a short grammatical sketch of the lexicon, phonology, and morphology of Tengger. The syntax and morphosyntax will take up the bulk of the dissertation and is presented in part II.

CHAPTER THREE: THE TENGGER LANGUAGE

3.1 The Sociolinguistic Setting of Tengger

When they conducted their field work on the Tengger, some twenty-five plus years ago, the anthropologist/ethnographer Robert Hefner and his soon to be wife, sociolinguist Nancy Smith(-Hefner) report that there were fully twenty-eight villages and 40,000 people identified as Tengger. Fifty years ago, there were perhaps as many as forty villages which considered themselves (self-identified as) Tengger (S. G. Sutrisno, p.c.). Even in the twenty years before the groundbreaking work of Hefner and Smith-Hefner, there had been a significant amount of attrition among the Tengger. The last twenty-five to thirty years have been a time of rapid development and change throughout the whole of Java, including the Tengger region. According to Sutrisno, no more than 20 Tengger villages remain today, and those that do are under significant pressure from the outside.

First, though, we must briefly consider what it means to be Tengger. Most importantly, Tengger are Javanese. They speak a dialect of Javanese, practice an older, pre-Islamic Hindu-Javanese religion, and adhere to many of the cultural practices of lowland Javanese. Historically, they settled the upland region around the sacred Mts. Bromo and Semeru, in East Java, probably sometime before the 14th century. Their numbers almost certainly grew due to migration and flight of Hindu-Javanese when the last great Hindu-Javanese Empire of Majapahit fell to invading Muslim sultans from the

north coast during the end of the 15th century. Since that time, they have lived as a predominantly agrarian, settled, peaceful, Hindu-Javanese society.

The term Hindu-Javanese was first used to describe the religion of the Tengger by Hefner (1985). This term refers more broadly to the syncretic religious practices found throughout Java before the large scale conversion to Islam, which took place largely during the 13-15th centuries. The practice is a mix of indigenous Javanese traditions heavily influenced by Hinduism and Theravada Buddhism imported, via other Southeast Asian states, from India. The modern practices probably evolved over a millennium of contact between Javanese polities and Indian travelers, traders, and religious men.

With the advent of Islam in Java[14], and its rapid rise to political dominion, the Hindu-Javanese who did not convert fled east or took to the mountains and other remote areas. By the end of the mid-18th century, with the defeat of Hindu-Javanese Blambangan in the far east of Java, there were no longer any large or organized political entities that espoused Hindu-Javanism. The religion and practices did survive in neighboring Bali and in a few isolated regions within Java. Today, with 4 million people, Bali has the largest numbers of 'Hindus' in Indonesia; though in the 500 years since the collapse of organized Hindu-Javanism in Java proper, the Hinduism of the Balinese has taken on a very uniquely Balinese tone. Within Java itself, adherents of the *agama asli* or 'original religion', are found in several scattered pockets in Banyuwangi, on the far east end of Java, Banten in the west, and the Tengger of East Java. Other adherents of *kejawen* religion are to be found scattered throughout Java, though there are few who would identify themselves as such, particularly given the recent shift in political Islam,

[14] Today, around 90% of ethnic Javanese identify themselves as Muslim, though the degree of adherence and belief in Islamic traditions and principles varies greatly.

where there is a sort of social pressure to appear more Islamic as evidenced by the rise of the *jilbab*, or Islamic woman's scarf.

Of the remaining villages that identify themselves as Tengger, it is no longer the case that this implies a practice of Hindu-Javanism. The increase in trade, better transport, paving/stoning of road links, and more intense contact that has come in the past 25 years between the Tengger and the *wong ngaré* or 'lowlanders' has also brought with it cultural, religious, and linguistic influence. Large scale conversion to Islam, adoption of lowland cultural attitudes, and language drift/shift from the Tengger dialect to the lowland dialect represent significant challenges to the maintenance of a unique Tengger identity. Religious conversion, however, need not imply linguistic 'conversion' as well. The village of Ngadas serves as an excellent example.

Of the 1,601 villagers in Ngadas, there are 1110 Buddhists, 411 Muslims, 78 Hindus, and two Christians.[15] Ngadas is unique in having such a large population of Buddhists, which is generally associated with ethnic Chinese in the lowlands and throughout the archipelago. This is however, something of an historical accident. When Islam became the dominant religion in Java, people who did not practice Islam referred to themselves as *wong buddha* 'Buddhist'. This did not actually refer to any particular doctrine of Buddhism, but rather to the syncretic Hindu-Buddhist-Javanese religion that was dominant prior to Islam. This appellation seems to have been maintained, especially among the Tengger, for several centuries. With the advent of the modern Indonesian government, by law every citizen had to affiliate with one of the five recognized religions: Hinduism, Islam, Buddhism, Christianity, or Catholicism. By the 60's and

[15] Only 42 villagers are not native to Ngadas. Each émigré is married to a native; in each case the religion of the husband has been maintained, be it Buddhist or Muslim.

70's, most Tengger villages stopped using the term *wong buddha* and identified themselves as Hindu, given the historic relationship with the Hindu's of Bali.[16] However, Ngadas was the most remote of the Tengger villages, and in some ways the most parochial, and they continued to call themselves *wong buddha*. During the 1980's, after several bad harvests, the people of Ngadas searched for outside help, which they found from some Chinese Buddhist organizations in Surabaya. These groups provided financial support for the farmers contingent upon their association with official Buddhism and most of the villagers converted. However, in response to some very corrupt dealings on the part of several villagers, a group of several families requested assistance from Bali, and adopted official Hinduism. During all of this, with the increased contact with lowlanders, Islam was beginning to have a significant influence on many of the villagers, who eventually adopted official Islam.[17] So, today in Ngadas, there is a *sanggar* 'Buddhist temple', a *pura* 'Hindu temple', and a *mesjit* 'mosque'.[18]

This is all significant because, regardless of religion, the people of Ngadas have all, without exception, maintained their local dialect and this is even true among the younger generations. The ultimate effect of Islamization on the language, and with it the stronger connection to the lowlands, is unclear. Common Arabic phrases which entered SJ probably centuries ago, but which had not entered the Tengger lexicon have begun to emerge in Ngadas, but are restricted only to those villagers who have converted to Islam.

[16] And in many cases, they were actually given money from Bali. In this system, the greater the number of adherents of a particular religion, the larger the share of the national budget that goes toward that religion, thus promoting aggressive conversions.

[17] This trend in Ngadas has been rather recent, beginning only in the 90's, and gaining momentum over the past decade. Construction of a mosque in the village began in 2001, and was not yet complete when I left in August 2003.

[18] Two young men in their twenties were educated at a Christian school in the lowlands and have adopted official Christianity, though no church exists in the village.

The borrowing situation in Tengger is actually extremely complex, and I will take it up again in a later chapter.

Among Tengger males, it is not uncommon today to have adequate fluency in the Tengger dialect, the lowland dialect, Indonesian, and some *krama desa* or 'village krama'. However, the first language a young Tengger learns is the dialect, and the dialect is the language commonly used for all purposes within the village.

3.2 Speech Levels in Javanese and Tengger

Javanese displays a unique linguistic phenomenon termed *speech levels*.[19] In this system, word choice is intimately tied to the relative social position of speaker and interlocutor. In some respects this system is similar to the socially stratified honorific system found in other Asian languages such as Japanese and Korean, and to the polite-familiar distinction found in Romance pronominal systems. However, the speech levels in Javanese are not limited to honorifics or pronouns, and are spread through every category in the language.

It is also important to note that this is not a register system, with one language used for formal situations and another in casual settings. The choice of speech level depends entirely upon those engaged in speaking, although any formal public speaking calls for *krama*. It is also not a literary-spoken language distinction, as exists in a number of different languages, such as Bokmål and Nýnorsk Norwegian. Contemporary Javanese literature makes full use of the speech level system, generally using whichever level is appropriate to the context.

[19] The speech level system has been well documented. See Clynes (2000), Errington (1985, 1998), Geertz (1960), Poedjosoedarmo (1968, 1969), Uhlenbeck (1978), Wolff and Poedjosoedarmo (1982) among others for a more detailed discussion.

A register system can be termed symmetrical as all participants engaged in a speech event use the same register. Most languages in fact have register distinctions, though the difference between acrolectal and basilectal speech may be minor, as in English, or extremely significant, as in standard Indonesian and Jakartan Indonesian, for example. The speech level system is asymmetrical; participants in a speech event would only use the same level if they were or equal social status. Otherwise, the exchange would be between different speakers using different levels.

There are three basic levels in Javanese: *krama (basa), madyo,* and *ngoko. Krama* is the high, refined, polite language, *madyo* is the middle language, and *ngoko* is the familiar, intimate, everyday language. In addition to these base levels are two further sets of vocabulary: *krama inggil* and *krama andhap*. These latter forms are limited sets of vocabulary mostly relating to personal affects, actions, and requests and can be used with any of the three base levels. *Krama inggil* is used for words referring to very high status or honored people; they can never be used to refer to oneself. On the other hand, *krama andhap* is only used in reference to the speaker, in a humbling or self effacing manner.

1. a. Punika punapa inggih kagungan panjenengan? *Krama*

   b. Niki napa nggeh gadhahan sampeyan? *Madyo*

   c. Iki apa yo dhuwek- -mu? *Ngoko*

   this interr yes possession 2nd sg.

   'Is this yours?' (Errington 1998:37)

In this example, the *krama* would be used from a social inferior to a superior, a younger person to an older person, to a new acquaintance, or to anyone deserving of respect. The *ngoko* is only used from older persons to younger, from superiors to inferiors, or between extreme intimates. *Ngoko* is the first language that Javanese learn and it is the base language, unrefined. It is used for humor and in anger. Keeler (1984:xviii) notes that *madyo* 'is a less refined manner of speaking that krama, but it is also less familiar than ngoko. It is used in several types of situations: with and among persons of low status; with people one has known a long while but with whom one is not truly intimate; and to people with whom one is close but to whom one must still show respect.'

Not all words have an equivalent in each level. There are a number of words which are invariant across levels, such at *Ratu* 'king' below. There are also affixes which have different forms in the various levels. These words and affixes differ in form but not in sense as they are semantically identical. On the other hand, pragmatically they differ to show very intricate relationships between speakers.

2. Speech Levels in Javanese

Ngoko	*Madyo*	*Krama*	*Krama Andhap*	*Krama Inggil*	*English*
wong		tiyang			person
iki	niki	punika			this
kandha	sanjang	criyos	matur	ngedika	say
mata		mripat		paningal	eye
aku		kula	dalem		1st sg.
kowe		sampeyan		panjenengan	2nd sg.
ratu					king

The speech language system has been borrowed, generally in a somewhat reduced form, into the neighboring related languages of Madurese, Sundanese, Balinese and Sasak. It is found in no other languages. There is general agreement among scholars that the system emerged first in Javanese and was subsequently borrowed into the other languages at some point during the Majapahit or Mataram era when Java held sway (though not necessarily dominion) over its neighbors.[20] Clynes (1994) argues rather convincingly against the general assumption that the speech level system developed sometime during the 17th and 18th centuries CE, around the Central Javanese court areas of Yogyakarta and

[20] Zurbuchen (1987) argues that Balinese speech levels arose independently from the Javanese, and were based initially on a Sanskrit honorific register used in the Old Balinese period from around the ninth to the eleventh centuries CE. It is only during the Majapahit era, 12th-15th centuries when Java held political dominion over Bali, Lombok (Sasak), Madura, and West Java (Sundanese) and probably then during this time when the speech level system spread.

Surakarta.[21] He supports his claims with borrowing evidence from Balinese. Bali was under Java's dominion from as early as the eighth century, lasting through to the collapse of the Hindu-Javanese Majapahit kingdom in the early 15th century. The contact was intense, and Javanese had an enormous effect on Balinese. After the fall of the Majapahit empire, the Hindu Balinese no longer looked to the now Islamic Javanese for cultural or linguistic example (cf. Schulte 1986, in Clynes 1994). So, by the beginning of the 15th century, there were very few borrowings from Javanese entering Balinese. In a rather careful study, Clynes shows that the Javanese borrowings in Balinese show very clear evidence of a speech level system. This would indicate, then, that the speech level system was fully articulated by the beginning of the 15th century. His argument ends there, that the speech level system must have originated much earlier than generally assumed, by the beginning of the 15th century at the latest, but as early as the 8th-9th century when contact between Java and Bali first took place.

One remarkable feature about the Tengger dialect is its lack of a speech level system. (on the sociolinguistic implications of this see Smith-Hefner 1983). In chapter 2, I noted that the Tengger have been in their current position in Tengger-Bromo Massif since at least around the end of the 14th century. The Charter of Wlandit, which is dated 1405 CE, is a copy of an earlier charter of King Hayam Wuruk who died in 1389 CE. There are two main implications of a reading of this Charter. The first is that, contrary to the popular story which has them fleeing a the fall of the Majapahit Empire at the beginning of the 15th century, the Tengger have been identified as a distinct group living

[21] See Soebardi (1975), Gonda (1948), Mudjanto (1986) and Slametmulyana (1954) for evidence of a late emergence of the speech levels. However, much of the argumentation is circumstantial, and relies particularly on the absence of speech levels in the Old Javanese literature. We would not expect to find the level system in a purely literary language, which Zoetmulder (1983) argues is exactly the case.

in the relative isolation of the Tengger-Bromo highlands since at least the time of King Hayam Wuruk. Second, given the lack of speech levels present in Tengger and their separation from the lowland language, it can be posited that by this date the speech level system had not yet arisen within Javanese. Further, there is some linguistic evidence to support this hypothesis, such as the differences between the Tengger and lowland words for body parts.

Many of the Tengger words for body parts are not found in reference to humans in either Central or East Java. *Mata* 'eye', *congor* 'nose', and *endhas* 'head' for example are the Tengger items used to describe human and animal body parts. In the Central and East Javanese dialects the equivalents for human body parts are *mripat, irung,* and *sirah*, respectively. *Mata, congor,* and *endhas* are used exclusively in reference to animals (or in derogatory epithets). However, almost all dictionaries of standard Javanese will list *mata* as the *ngoko* or low word and *mripat* as the *krama* or high equivalent. Clearly, at some early point, *mata* was universally used in reference to a human or animal eye. However, after the development of the speech level system, when a more refined or prestigious equivalent became available, the status of *mata* was lowered to refer only to animals, with the *krama* form becoming the general word for 'eye' in reference to humans. The fact that the Tengger maintain these forms shows that they became isolated and that the dialect began to diverge before the speech level system arose. So where the Balinese evidence provides a 'no later than' date for the emergence of the speech level system, the Tengger evidence provides a 'no earlier than' date. We can therefore posit that the Javanese speech level system arose sometime after 1389 CE and before the beginning of the 15th century.

However, to say the Tengger do not display evidence of the speech level system is not entirely true. The speech level system is not native in Tengger, but a sort of parochial, rural version of the high *krama* is present. Used almost exclusively by Tengger males, and then generally by those who have or have had some administrative role in village affairs, there is some fluency in what is referred to a *krama desa* or 'village krama'. I describe *krama desa* in more detail in chapter 2. What is important to note is that even this often restricted use of *krama* is a recent occurrence, and it is a learned language, not native to Tengger. Interestingly, the phonological features which make the Tengger dialect appear so distinct from other dialects of Javanese are not present when a villager is speaking in *krama*, rather, the standard lowlands pronunciation is also incorporated.

The speech level system is also different in Malang, and in East Java in general, when compared to its full articulation in the Central Javanese court cities. Throughout East Java, speech levels are far less rigid or proscribed. This often leads Central Javanese to malign East Javanese as *kasar* or 'crude, brash, rough'. Of course, whether or not the system first evolved in Central or East Java is debatable, it is very clear that it found its greatest expression around the court cities of Central Java.

Of twenty males under 25 years old who were born and reared in Malang and whom I interviewed, one was comfortably fluent in *krama*. While each of them was a passive speaker, and most could easily have a conversation in krama, they did not feel comfortable in the use of *krama*. When I asked for *krama, madyo, krama inggil* and *krama andhap* equivalents of common *ngoko* words, out of 23 tokens which had either *krama inggil* or *krama andhap* counterparts, only one respondent could provide all 23

and most were unable to provide any.[22] The situation is exacerbated in the urban area in and around Malang, where television, radio, movies, advertisements—all in Indonesian—are all the more prevalent. Further, there is a popular sentiment that good Indonesian is necessary for success in today's Indonesia, and that the complex speech level system of Javanese is not as important. The speech level system in Javanese is clearly threatened.

3.3 A Grammatical Sketch of Tengger: Lexicon, Phonology, Morphology

In this section, I provide an overview of aspects of Tengger grammar. The overview of Tengger lexicon, phonology, and morphology will be necessarily brief as I aim to focus more on the syntax of the language. Where Tengger behaves similarly to SJ, I will refer the reader to other works where those aspects of the language are described in more detail. However, as noted earlier, most works on Javanese concentrate on literary and not grammatical phenomena in the language. Robson's (1992) *Javanese Grammar for Students* is the best and most thorough reference grammar currently available in English. In the introduction, Robson lists a number of earlier grammars in Dutch, including Kiliaan (1919), Prijohoetomo (1937), and Uhlenbeck (1978); as well as Sudaryanto (1991) in Indonesian. There are very few linguistic analyses of Javanese, though Uhlenbeck's writings constitute a major contribution (see bibliography for further references). There has been even less work done on Javanese dialects. Under the

[22] I did not use these terms when attempting to elicit said tokens, rather, I asked for more or less *halus* 'refined' words, or asked what speakers would use with their parents, grandparents, elders, etc. The terms used here to refer to the different speech levels, such as *krama andhap*, etc. are very formal, metalinguistic

Department of Education and Culture, the *Pusat Bahasa*, or Language Center[23] has produced a number of monograph sized studies of various aspects of regional dialects. However, in general these tend to be very brief descriptions which are not of a particularly high quality. It can safely be said that a systematic and analytical study of Javanese, particularly of Javanese dialects is currently lacking. I hope that this dissertation will make some small contribution in remedying that situation.

It should be remembered that this study is based on the dialect of two villages, namely Ngadas and Ranu Pane. As such, the language described herein, while largely representative of the Tengger dialect as a whole, will not necessarily be completely congruous with that found in other villages. Ngadas and Ranu Pane are the most isolated of the Tengger villages, with the least amount of contact with non-Tengger. At altitudes of 2700m and 3100m respectively, Ngadas and Ranu Pane are the highest incorporated villages in Indonesia outside Papua. It was only in the past year, during my stay there, that Ngadas finally was connected to the electrical power grid, and the bottom half of the road to the two villages was beginning to be paved as I left in August 2003. However, there are no plans yet to pave the top section!

Given this isolation, the language of Ngadas, and to a lesser extent—due to a higher influx of non-Tengger settlers—Ranu Pane, are the most conservative of the Tengger dialects. This is seen most clearly in its lexicon and phonology, although, as I will demonstrate below, there are a number of significant differences in syntax as well. As I aim to provide a synchronic comparison of the Tengger dialect with that of the surrounding lowlands, with some reference to the standard language of Central Java, and

[23] This is the board's name since 2000 prior to which it was called the Pusat Pembinaan dan Pengembangan Bahasa, or Center for Language Cultivation and Development.

a diachronic comparison of these dialects with Middle Javanese, and not a detailed study of Tengger dialects per se, I will refer below to the dialects of Ngadas and Ranu Pane simply as the Tengger dialect, given the caveats above, and Javanese will refer to the standard Central Javanese *ngoko*. Further, as other dialects of Javanese, especially the standard Central Javanese dialect have been described elsewhere, I will only make reference to non-Tengger dialects below where there is a salient difference. I would also like to reiterate that this is not a grammar of Tengger and is not intended as such. Such an undertaking would require an entire dissertation or a separate study dedicated solely to that purpose. I aim here to provide a basic grammatical sketch so as to explore in later chapters several of the significant variations that exist between Tengger and surrounding lowlands dialects, and their historical derivation.

3.3.1 The Tengger Lexicon

The most significant and apparent area where Tengger differs from other dialects of Javanese is in its lexicon. It is also the area where the greatest similarity between Tengger and Middle Javanese is seen. The long period of isolation and lack of contact with other languages, such as Arabic, Dutch, English, among others, and dialects is the primary reason for the large number of historical retentions in Tengger vocabulary. The following list provides just a few examples (see App I for a more thorough listing):

3. Tengger lexical retentions

Tengger	Javanese	English
a. (r)eyang	aku	1st person pronoun (male)
b. isun	aku	1st person pronoun (female)
c. paran	apa	what
d. kate	arep	want; fut. aux.
e. njajal[24]	cobo	try
f. bantah	omong	talk, say
g. manja	nandur	plant v.
h. ayo	mangga	please
i. amit	kula nuwun/nuwun sewu	excuse me
j. njare	kepiye/yo opo	how('s that)
k. neki	yen	if, when
l. karo	lan	and
m. ngapa/njare	kok/kena apa	why
n. he-eh	iya	yes
o. dhek[25]	neng	at

What is immediately remarkable from just this short list of unique Tengger lexical items is that they mostly belong to what are considered to be the most stable classes of words: pronouns, auxiliaries, interrogatives, prepositions and conjunctions. These are all very

[24] This is also used in Central Javanese, though with a more restricted use as a parenthetical utterance closer. I thank J. Errington for pointing this out.
[25] In all dialects of Javanese, this is commonly used to refer to a point in past time.
[25] *Kito* as a 1st person plural pronoun in SJ is a recent borrowing from Indonesian.

high frequency items that are resistant to borrowing and tend to remain more stable over time. It should be noted that pronouns across Javanese dialects are all highly unstable. Most dialects have only distinct first and second person pronouns, unmarked for number, case, or gender. This paucity of forms has allowed for the development of a wide variety of other lexical items to take on pronominal roles.

For a number of these items, the Tengger word is a retention from OJ/MJ. However, this is not always the case, and there are a number of examples where it is the standard form which has preserved the older meaning and the Tengger item which has undergone semantic shift. One example is T *mentas* 'to go up, ascend', from the root *entas*. In both OJ and SJ *m-entas* means 'to come out of the water'. The shift here is an obvious one, the Tengger are mountain dwellers, quite removed from any major body of water, and so the verb has come to mean 'to go up' from its original 'to come (up) out of the water'.

Another remarkable feature is seen in the pronominal system. Note that 3 (a&b) given above are both first person pronouns. SJ, OJ, and MJ do not mark case, number, or gender on pronouns.[26] However, a gender distinction has been introduced on the first person personal pronouns in Tengger. There is no other place in Tengger or SJ grammar where a gender distinction exists. Indonesian has introduced a gender distinction on a very few substantives borrowed from Sanskrit, but not on pronouns. *Ingsun, isun, sun, ningsun,*and *nisun* are all forms that emerge in the Kidung literature in MJ for the first person pronoun. However, the origin of *(r)eyang* is not clear. It seems not to have appeared in OJ or MJ. In ModJ it is the respectful form for 'grandparent', and its use in T could have developed from its function as an appellative for an older person, except

[26] *Kito* as a 1st person plural pronoun in SJ is a recent borrowing from Indonesian.

that T never developed the *krama*, or polite level of SJ (see section 2 on speech levels). On a sociological note, it is interesting that a gender distinction should arise among the Tengger who have, compared to lowland Javanese, a far more egalitarian society with much less clearly defined gender roles and a far greater sharing of basic tasks and duties (though Tengger women are still responsible for cooking!).

Probably since as early as the fourth or fifth century CE, Javanese have been in contact with non-Austronesian speaking people. They have readily borrowed both culturally and linguistically from these foreign influences. As noted in chapter 2, Sanskrit was the earliest and most significant external influence on the Javanese language, with some older manuscripts containing almost 50% Sanskrit vocabulary. The evidence of this is readily apparent in the Tengger lexicon as well. In the 14th and 15th centuries CE, Javanese began to borrow heavily from Arabic, as Islam spread throughout the island. What is remarkable is the lack of Arabic lexical material in Tengger. Some early European borrowings have entered the Tengger lexicon, though it is not clear when this occurred: From Portuguese Tengger has *sendhela* 'window'. (SJ *jendhelo).*

The absence of speech levels also accounts for a number of differences between SJ and Tengger. This is particularly apparent with words for body parts:

4. Words for body parts in Tengger

Tengger	Javanese	English
a. endhas	sirah	head
b. mata	mripat	eye
c. congor	irung	nose
d. cengkul	sikut	elbow
e. belung gambang	iga	rib
f. lambe	cangkem	mouth

With the exception of (d & e), all of the Tengger words also occur in SJ with the same meaning. However, in SJ, they are applied only to animals and not humans. *Endhas, mata,* and *lambe* all occur in OJ and MJ referring to humans. It is only with the development of the speech level system where these words acquired a 'low' reading, and became part of the *ngoko* or familiar speech level. However, even when speaking in *ngoko* today one will often use a *madyo* (middle) or *krama* (polite) word, when available, when directly referring to another person. Over time, then the original Javanese words for body parts have all been replaced with more 'polite' terms, leaving the original words used only for animals, or in a very insulting way. That the speech level system never arose in Tengger, means there was never any trigger for these words to have a reduced status, and so remain today used comfortably for human noses and heads.

3.3.2 A brief overview of Tengger phonology

Tengger phonology diverges very little from that of SJ. Other regional dialects of Javanese have developed quite different phoneme inventories, such as the presence of a series of palatalized stops in the Osing dialect of Banyuwangi. Tengger has the following phoneme inventory:

5. Tengger Consonant Inventory

p	m		
t	n		
ʈ		j	l, r
ʧ	ɳ		
k	ŋ	w	
ʔ			h

Each of the stops and affricates, except for the glottal stop, also appear in a pre-nasalized form, with a homo-organic nasal, occurring only in onset position word medially. It is not clear whether word initial stops appearing with a nasal are pre-nasalized or nasal + stop clusters.

Below I give the Tengger vowel inventory:

6. Tengger Vowel Inventory

i	I		u	ʊ
e	ɛ	ə	o	ɔ
			ɑ	ɒ

These paired vowels have a very regular distribution. The tense vowel occurs in open syllables and the lax vowel in closed syllables. Only schwa /ə/ (SJ *pepet)* has an unpredictable idiosyncratic distribution.

Tengger, like other varieties of Javanese, also has a series of murmured or breathy voiced vowels.[27] These occur following what is written with a voiced stop and historically developed when the voicing on consonants was reinterpreted as being a quality of the vowel and not the consonant (Thurgood 2004). Consequently, Javanese has lost all of its voiced stops and affricates. While the voiced-voiceless distinction is still maintained in the standard orthography of Javanese, this distinction no longer exists in the spoken language.[28] The fact that Tengger shares this feature with SJ, means that it must have been firmly established by the 10th or 11th century (see discussion of the rise of speech levels section 2), before the Tengger dialect began to diverge from the standard.

[27] What Ladefoged and Maddieson (1996) call 'slack voice' vowels

[28] A similar phenomena occurred in the history of the Chamic languages of peninsular SEA, and eventually developed in some instances into full tonal distinctions. The Chamic languages, though were heavily influenced by surrounding tone language families, such at Tai-Kadai, Tibeto-Burman, and Austroasiatic. Why the same initial stages are seen in Javanese, which has no contact with tone languages, is unclear. See Thurgood (1999).

The opposite holds true for vowel harmony. SJ has a little studied vowel-harmony system where /a/ in final open syllables becomes /o/ or /ɔ/, and this spreads to an /a/ in a preceding open syllable. For example, */dawa/* 'long' becomes [dowo]. However, suffixation blocks harmony: /*dawa-ne/* 'long-the, its length' is [dawane]. This harmony rule only operates on /a/ and on no other vowels. A pre-nasalized medial stop does not block harmony, but any other cluster will: /*bandha/* 'wealth, fortune' [bɔn̪ɖɔ], but */arta/* 'money K' [artɔ]. To my knowledge, Cohn (1999) is the only scholar to provide any systematic treatment of Javanese vowel harmony (though she never published those results). There is a great deal of idiosyncratic variation on how individual words are pronounced, using /o/ or /ɔ/. There is also a great deal of regional variation on what blocks harmony. For example, I have heard both /*kanca-ne/* 'friend-the' [kanʧa-ne] and [kɔnʧɔ-ne] from the same speaker.

What is significant for our purposes here is that Tengger does not display any vowel harmony. In fact, this is one of the most significant obstacles to mutual intelligibility with lowland Javanese speakers. Vowel harmony is almost certainly an innovation, with the Central Javanese dialects the earliest innovators.

Tengger does share with both SJ and EJ a system of vowel assimilation at morpheme and word boundaries referred with the Sanskrit term *sandhi*. As in Sanskrit, Old Javanese has both 'external' sandhi rules that govern the combination of vowels occurring across word boundaries, and 'internal' sandhi rules that govern the combination of vowels occurring across morpheme but not word boundaries. Although there are some distinctions between external and internal sandhi in Sanskrit, which heavily influenced

the Javanese system, the two vowel sandhi rules are completely parallel in OJ. Modern Javanese, on the other hand, only maintains internal sandhi; that is, no vowel assimilation takes places between consecutive vowels across a word boundary.

Old Javanese recorded a contrast between long and short vowels in its orthography, which was based on an early Pallavi script. It is not clear whether or not the language itself actually had a productive distinction between long and short vowels or if it was just a relic of a writing system originally developed for a language which did have that distinction. The sandhi rules for OJ are as follows:

7. Sandhi in Javanese

i. two identical vowels, whether long or short, merge to one long vowel

ii. an initial schwa of a suffix is lost before a final non-schwa vowel of a root

iii. /a/ + /u/ > /o/, e.g. *a* + *umah* > *omah* 'house'

iv. /a/ + /i/ > /e/, e.g. *impi* 'to dream' + *an* > *impen* 'a dream'

v. u, o, ɔ > w/_V (except schwa) e.g. *sinusu* + *an* > *sinuswan*

vi. i > y/_V (except schwa) e.g. *manguni* + *aken* > *mangunyaken*

vii. e > ay/_V e.g. *magawe* + *a* > *magawaya*

Zoetmulder (1992:3-4)

Whether or not this feature was actually a borrowing from Sanskrit is unclear. However, this type of sandhi does not appear in other closely related languages. It is possible that the phenomenon was influenced by the borrowed/adapted Pallavi script, which included rules to account for Sanskrit sandhi. That the two systems, that is the Sanskrit and Old Javanese vowel sandhi, from two genetically unrelated languages, should be largely

congruous is some evidence supporting the hypothesis that Javanese sandhi was a borrowed feature from Sanskrit. Further the fact that the sandhi operates on general principles of vowel gradations, which are present and productive in Sanskrit but completely lacking in the uninflected Javanese, further points to influence from Sanskrit.

As noted above, Modern Javanese does not have external sandhi. The rules are largely similar to those for Old Javanese with some exceptions. There are no long vowels in Modern Javanese. With this exception, rules (i-iv) above also hold for Modern Javanese. Rules (v-vii) no longer apply in Modern Javanese. The prefixes *ke-* and *ka-* deserve special note. *Ka-* is one form of the passive prefix which was very productive in Old Javanese, but occurs mostly in fossilized examples in Modern Javanese. This prefix is never subject to sandhi, except perhaps in *karep* 'a wish' from *k(a)-arep* 'to wish', where the vowel of the prefix and the initial vowel of the root have merged.[29] In contrast, we have *kaanan* 'state, condition' from *ka-ana-(a)n* 'to be', where the vowels remain independent. The *ke-* prefix forms a class of verbs referred to as 'accidental passives' or 'adversatives'. Here the schwa of the prefix is lost before any initial vowel in the root. The Tengger dialect follows the same pattern as SJ, and has no external sandhi.

3.3.3 An overview of Tengger Morphology

As Smith-Hefner notes, 'The rather slight differences in the verb morphology of the Tengger dialect are not a particularly salient feature of dialect speech.' (1983:142). The

[29] Though in the modern language, especially CJ, this is probably a fossilized form, as one also gets *di-karep-ke* 'wants', with the *k-* still attached to the root.

reader is referred there and to any of the grammars of standard Javanese given in the bibliography for a full accounting of Javanese morphology, especially Uhlenbeck (1978).

However, that being said, it should not be assumed that either verbal or nominal morphology are uniform across Java. Consider the following chart of some of the most frequently used verbal morphology:

8. Tengger Verbal Morphology

	Tengger	Central Java	East Java	Old Javanese
causative, benefactive	-(k)na/en	-(k)ake	-(k)nɔ	-akən
passive imperative	-(k)ən	-(k)ən	-(k)ən	-ən
passive imp. caus/bene	-na	-nɔ	-nɔ	-akna
active imp. caus/benef	-(k)ən	-nɔ	-nɔ	-akna
active imp./ subjunctive	-a	-ɔ	-ɔ	-a

Taking into account only these three dialects, we can see that there is a good deal of variation among them. What is initially remarkable about both the Tengger and the East Javanese dialects is the collapsing of forms, though not in a uniform way. In Tengger,

the passive causative/benefactive remains the distinct form, whereas in the East Javanese dialect, it is the passive imperative.

There has been much debate in the literature on Austronesian languages as how best to characterize the basic opposition present in the verbal system. Various authors have referred to a 'traditional' active-passive dichotomy, an actor-patient focus system, and a topic-comment system. The debate first emerged when researchers were trying to analyze the Tagalog (and other Philippine languages') verbal system, in which affixes added to the verbs refer to the semantic roles of noun phrases that are the 'topic' or 'subject' of the verb. This focus system allows reference to traditional 'subjects' and 'objects' but also has specific morphology to refer to 'locations' and 'instruments' as the grammatical or syntactic subject of a given verb.[30]

Old Javanese has a verbal system very similar to that of Tagalog and the other Philippine languages. Hunter (p.c) identifies three primary focus markers for Old Javanese. The infix *–um-* marks actor focus (what others have called 'active'), the infix *–in-* marks undergoer or goal focus (passive in other characterizations), and he points to

---

[30] Examples of this are discussed again in chapter 6, but the pattern is presented here as well.

a.	b-um-ili ng isda sa tindahan ang lalake	[Tagalog]
	VC-buy CORE fish OBL store man	
	'The man bought fish in the store.'	[agent subject]
b.	bi-bilh-in ng lalake sa tindahan ang isda.	
	IRR-buy-VC CORE man OBL store fish	
	'The man will buy fish in the store.'	[theme subject]
c.	bi-bilh-an ng lalake ng isda ang tindahan.	
	IRR-buy-VC CORE man CORE fish store	
	'The man will buy fish in the store.'	[locative subject]
d.	ipam-bi-bili ng lalake ng isda ang salapi	
	VC-IRR-buy CORE man CORE fish money	
	'The man will buy fish with the money.'	[instrumental subject]
e.	i-bi-bili ng lalake ng isda ang bata.	
	VC-IRR-buy CORE man CORE fish child	
	'The man will buy fish for the child.'	[benefactive subject]
		(Foley 1998:2)

the complex infix *–in-* plus suffix –an as analogous to Tagalog locative focus. Both of these infixes are present in Modern Javanese, though they are no longer productive and only found in fossilized examples on a few roots.[31]

In both Modern and Tengger Javanese, this three-way distinction has been reduced to a simple dual opposition. A homo-organic nasal assimilated in place of articulation added to the onset of a verb indicates actor focus or active clauses.[32]

9.	a.	delok	'to see'	n-delok 'see act.'
	b.	pati	'death'	m-(p)ati 'kill'
	c.	tulung	'help'	n-(t)ulung 'to help'

This pattern is observed both in lowland and Tengger Javanese. However there is some variation in the undergoer focus or passive systems.

There are several distinct forms of the passive in Javanese. There are two passive affixes which are inherited from OJ and are now archaic, occurring only in fossilized examples. The prefix *ka-* and the infix *–in-*, discussed above. For example, both occur with *timbang* 'weigh' in *ka-timbang, t-in-imbang* 'compared to, weighed against'. These are no longer productive.

[31] The non-productivity of these roots is seen very clearly in the example of *pinarak* 'please come in, sit down'. This is from the OJ root *parək* 'near' with the passive infix –in-, a polite way to invite one to draw near. None of my consultants was able to identify this as the root for *pinarak*, clearly indicating that the –in- infix is no longer productive. This was especially remarkable among the Tengger, who continue to use *parək* as the common word for 'near', in the lowlands this has been replaced with *cedhak*.

[32] In words beginning with a voiceless consonant, the nasal assimilates in POA to the subsequent consonant and then the voiceless consonant is lost. Voiced consonants are maintained. The underlying form of the prefix is /ŋ/ as this appears before vowels, liquids, and rhotics.

The common method for passivization uses the prefix *di-*, also present in SJ, Indonesian, and a number of other Austronesian languages.

10. a. Éyang katé n-andur-i bibit-é sésuk.
1.m will N-plant seed-É tomorrow
'I will plant the seeds tomorrow.'

b. Bibit-é katé di-tandur-i (Marsam) sésuk
seed-É will DI-plant-I Marsam tomorrow
'The seeds will be planted (by me) tomorrow.'

c. Bibit-é tak-tandur-i (*eyang) sésuk.
seed-É 1-plant-I tomorrow
'I intend to plant the seeds tomorrow.'

The syntax of passives in Javanese is very complex, and I will spend a good deal of time discussing this aspect of the language in a later chapter. For the time being, I would like simply to introduce the various constructions which have been analyzed as passives by different researchers and comment briefly on their morphological markings.

Example (10a) above gives a standard active, transitive sentence showing the order SVO (see chapter 4 for discussion of unmarked vs. marked order). Note that the verb is marked with the active prefix *N-*, a nasal which assimilates in place of articulation to the initial consonant of the root. The grammatical and logical subjects are congruent and appear in the specifier position, where the object appears in the post-verbal complement position.

Example (10b) shows the passive equivalent of (10a). The grammatical subject is the patient and appears in the pre-verbal position, the grammatical object is congruent with the logical subject, in this case the agent *Marsam*. The agent here can only occur immediately adjacent to the verb, it is ungrammatical in any other position unless further marked by a dative preposition *mbek* 'of, by'. The verb is marked by the passive prefix *di-*.

The *di-* prefix is actually part of a series of pronominal passive markers, as reported in most grammars:

11. Passive Pronominal Paradigm (standard Javanese)

	Ngoko	*Krama*
First Person	d/tak-	kula-
Second Person	ko(k)-	sampeyan-
Third Person	di-	dipun-

However, in my recordings of both Tengger and Malang Javanese, I have not come across one example of the second person passive prefix *kok-*. The case with the first person passive marker is confused by the existence of a propositive prefix with the same form *tak-*. The propositive occurs only in the first person and means something like 'I will; I intend; let me…'. However, the propositive can occur with an active or a passive verb, which is discussed in great detail chapter 6.

There are several different affixes that function as a passive in Old Javanese. Zoetmulder (1952, 1961) distinguished two types of passive affix for OJ based on the 'degree of involvement of the performer of the actions'.

In the *–in-* passive the action and its performer are treated as more important than in the *ka-* passive. In the *ka-* passive what is treated as central is the state brought about by the action. Hunter notes that the 'difference of agency and volitionality in the two Undergoer Focus types is brought out, says Zoetmulder, in differing constraints on the use of pronominal enclitics marking third person agents of *–in-* and *ka-* constructions.'

Old Javanese is very rich in derivational morphology. Given that nouns and adjectives do not inflect for case, number, or gender, and verbs are unmarked for tense, aspect, person, and number, there is no productive inflectional morphology other than that marking an active-passive distinction. Modern Javanese has a somewhat reduced amount of nominal and verbal morphology when compared to Old Javanese. In the most common colloquial spoken dialects there is a tendency to reduce the morphology even further. This may be an influence from the national language of Indonesian, which also shows significant morphological simplification in most of its colloquial dialects. The Tengger dialect is no exception in its move toward simplification of morphological markers. In the next section I demonstrate, however, that there is in fact less simplification, and the Tengger system represents an already simpler system. It is other dialects of Javanese, particularly the 'standard' dialects, which have in fact undergone significant complexification.

PART II: MORPHOSYNTAX/SYNTAX OF TENGGER

CHAPTER FOUR: WORD ORDER AND CONSTITUENCY

4.1 Introduction

In this chapter I explore issues relating to word order and constituency in Tengger. I focus on the order of clausal constituents, and do not address the order within the NP here. As with the other topics addressed in the dissertation, I place Tengger in both a synchronic perspective, offering comparisons with other dialects of Javanese as well as other Austronesian languages. Further, I provide a diachronic perspective for Tengger following its development from Old and Middle Javanese. I address several questions in this chapter. Does Tengger have an underlying or unmarked word order? Does Tengger make a distinction between arguments and adjuncts; does it allow bare obliques? Is Tengger configurational? Is there evidence for a verb phrase (VP) in Tengger, and what are the implications for the nature of grammatical and lexical categories?

These are significant questions. For some, I simply pose the questions and present the data. For others I suggest possible analyses. For most, however, I venture significant claims. First, I demonstrate that Tengger has two unmarked word orders. I argue that it is shifting from the VS(O) pattern found in Old Javanese, to the SVO pattern prevalent in most western Indonesian languages. However, I claim that the language is stable. While it is often difficult to find hard empirical evidence of constituency (it is a characteristic of free word order languages generally, but languages can allow free null elements and still show clear evidence of constituency) in a language that strongly prefers

null constituents, I do claim there is sufficient evidence to posit the existence of some clause internal structure in Tengger.

Tengger seems at first glance to be an almost completely free word order language, allowing constituents to appear in almost any linear order. While this is largely the case, it is not entirely true. A comparison with the standard language is useful. The question of underlying, or base word order in Standard Javanese is quite complicated. It is generally claimed that SJ demonstrates SVO word order; however, almost any combination of phrasal elements is possible. The general maxim from European languages, that the greater the morphological complexity in a language the freer the word order, and vice-versa, does not generally apply to Austronesian (and many Asian and even some European) languages, nor especially to Javanese and Tengger. For example, Javanese predicates are not marked for tense, number (except in the optative), aspect, or agreement, (they carry no phi-features); similarly substantives are unmarked for case, gender,[33] or number. Given this complete lack of inflectional morphology, we might expect that word order in Javanese would be quite rigid, based on assumptions from the study of European and other languages. This expectation is not borne out. In fact, in a simple transitive sentence with two arguments, each of the six logically possible word orders is perfectly grammatical in Tengger:

[33] The Tengger first person personal pronouns being the sole exception.

1. a. Supari tuku gedhang.

Supari buy banana

'Supari buys bananas.'

b. ?Supari gedhang tuku.

c. Gedhang Supari tuku.

d. Gedhang tuku Supari.

e. Tuku Supari Gedhang.

f. Tuku gedhang Supari.

Of the six examples above, only (1b) is slightly marginal, though it is still acceptable, and I do have examples of it in my recordings.[34] It should be noted here also that each of the above examples can have multiple intonation contours. The intonation here in no way disambiguates word order, nor does it mark thematic role. I discuss that further below.

Beyond a simple transitive sentence as in (1) above, Tengger also allows seemingly free order of other elements in a clause. To illustrate the scope of the issue, it is worth presenting the following examples, which are based on examples from Uhlenbeck (1994: 184-185)[35] but taken from my own fieldwork in Ranu Pane:

2. a. Dhek (waktu) iku / Marsam / cengkul-é / tokleg.

at time that Marsam elbow-ASS broken

'At that time Marsam's elbow was broken.'

b. Dhek iku / Marsam / tokleg /cengkul-e.

[34] The reason that (1b) is marginal has to do with the order of arguments, theme-agent-verb, which is the same structure as the passive *semu* construction, discussed extensively in chapter 5.
[35] I present the standard Javanese forms later in this section.

c. Dhek iku / cengkul-e / Marsam / tokleg.

d. Dhek iku / cengkul-e / tokleg / Marsam.

e. Dhek iku / tokleg / Marsam / cengkul-e.

f. Dhek iku / tokleg /cengkul-e / Marsam.

g. Marsam / dhek iku / cengkul-e / tokleg.

h. Marsam / dhek iku / tokleg / cengkul-e.

i. Marsam / cengkul-e / dhek iku / tokleg.

j. Marsam / cengkul-e / tokleg / dhek iku.

k. Marsam / tokleg / dhek iku / cengkul-e.

l. Marsam / tokleg / cengkul-e / dhek iku.

m. cengkul-e / dhek iku / Marsam / tokleg.

n. cengkul-e / dhek iku/ tokleg / Marsam.

o. cengkul-e / Marsam / dhek iku / tokleg.

p. cengkul-e / Marsam / tokleg / dhek iku.

q. cengkul-e / tokleg / dhek iku / Marsam.

r. cengkul-e / tokleg / Marsam / dhek iku.

s. tokleg / dhek iku / cengkul-e / Marsam.

t. tokleg / dhek iku / Marsam / cengkul-e.

u. tokleg / cengkul-e / Marsam / dhek iku.

v. tokleg / cengkul-e / dhek iku / Marsam.

w. tokleg / Marsam / dhek iku / cengkul-e.

x. tokleg / Marsam / cengkul-e / dhek iku.

This kind of variation is striking, and more inline perhaps with what we might expect either from a heavily inflected language such as Sanskrit or Latin, or from some Australian language which has been argued to have genuine free word order. Every logical ordering of phrases is grammatical. These examples are all variations of a simple predicate nominal sentence with a prepositional phrase. The Tengger examples are completely equivalent and, remarkably all are grammatical. To be sure these sentences are not all semantically and pragmatically equivalent, but those differences are subtle and not particularly germane to the issue of determining underlying word order here.[36] For many of the examples the intonation pattern is distinctive. In particular in (2c) *Marsam* is oddly placed between the substantive *cengkul-é* and the predicate *tokleg*; similarly the prepositional phrase in (2i) is equally odd intervening in that same position. Also, the placement of *Marsam* in (2v) reads as a tag or afterthought. Uhlenbeck (1994) notes that (2a & b) are most commonly "found in written Javanese. It is hard to discover a difference between them. Together with (2o) and (2q) they are the unmarked members of the set." (1994: 185-186). Though of course some of the attested patterns are more marginal than others.

Based on the above examples, and other data, Uhlenbeck claims that a phrase structure analysis of Javanese is inappropriate, or at least provides little insight into the structure of Javanese, and by extension closely related languages. Rather, he posits the 'sentence segment' as the syntactic unit most relevant and useful for study in Javanese. He is not alone in making this claim. In fact, it seems from the literature on Javanese, and the closely related languages of Balinese, Madurese, and Indonesian (there is almost

[36] It should be noted that in some of these examples, there are constituents which I will not argue for here. For example, (2a) here represents what could be analyzed as an external possessive construction, similar to those found in Hebrew. Similarly, (2c) is a normal genitive construction.

no literature addressing Sundanese and Sasak), there are two major competing claims. On the one hand, are researchers like Uhlenbeck (1976, 1996), Gil (1994, 1999, inter alia, for Riau Indonesian), Cumming (1991 on pragmatic preferences vs. syntactic constraints in Malay), who claim that these languages have minimal internal structure and that a constituency analysis is inappropriate and unproductive for languages such as these.[37] On the other hand are different approaches which either directly argue for or tacitly assume the existence of constituent structure in these languages, they are represented by Davies (1999 for Madurese and Javanese), Cole and Hermon (2002, 2003, 2007, etc. on Indonesian and Javanese), Chung (1976 on Indonesian), Voskuil (1996, 1999 on Indonesian), Wechsler and Arka (1998 on Balinese), Guilfoyle, Hung, and Travis (1992 on Indonesian and other Austronesian languages), among others. The question is begged, does the Tengger data lend itself more readily to a phrasal or non-phrasal analysis, that is, a constituent or non-constituent analysis?

Although not a totally free word order language, the above examples certainly suggest that Tengger might be a free phrasal order language, that is distinct phrases, as opposed to heads, can appear in any order and that there are no free split constituents. This would indicate that a phrasal analysis might most productively provide insight into the structure of Tengger. It is not the case, however, that Tengger is a totally free phrase order language, as there are restrictions on word as well as phrasal order in Tengger. After presenting a description of the basic facts, I discuss recent work by Davies on the order of constituents in Javanese and Madurese, and Uhlenbecks's analysis of standard Javanese. Davies (1999) argues that, contrary to appearances, Javanese and Madurese

[37] There has been a great deal of work on the syntax of Australian languages, showing that it is possible to give a syntactic analysis of such seemingly completely free word order patters.

are, in fact, strict word order languages, and that the great variety of surface variation falls out from the common use of dislocation constructions and null pronouns. After a description of the facts below, I will demonstrate that this claim does not extend to Tengger, and, by extension, probably does not account fully for the noticed patterns in standard Javanese either. In a derivational framework the problem is determining between free word order and free split constituents, and free phrase order derived by scrambling (and other movement) operations. I address those issues here. Finally, I will place Tengger typologically within the Austronesian languages based on recent work by Himmelmann.

4.1.1 A Note on Verbs and Null Elements in Tengger

What constitutes a clause in Tengger? What are the constituents? In order to answer these questions, the crucial element is the verb. Most verbs in Tengger are overtly marked for voice. The question of voice, or focus, or any of number of other analyses discussed in the literature, is complex and extremely important. The entire next chapter is devoted to voice. For the present discussion of word order and constituency, then, when possible, I make use of examples containing verbs that do not take overt voice marking, saving the analysis of the voice morphology for the subsequent chapter. There is a small class of verbs which do not take voice marking or any other verbal morphology when in a neutral sentence, that is, when not applicative, causative, imperative, etc, which are discussed in detail in chapter 6. This class includes verbs that seem to be inherently intransitive or transitive, and there is a single inherently ditransitive member of the class

as well *kirim* 'to send', though I give that special treatment below. Below I give a partial list of those verbs which do not take voice marking. Most verbs in this class are what might typically be called unaccusative or unergative. There are a few verbs that appear without voice marking which correspond to transitives in other languages. Also in this class are existential predicates:

3. Verbal classes in Tengger

existential predicates

ana 'there is, are/exist'	dadi 'to be, become'

unaccusatives

rutu 'fall'	tiba 'trip'	mulih 'go home'[38]
lair 'be born'	tangi 'wake up'	
lunga 'go'	teka 'come'	

unergatives

adus 'take a bath'	budhal 'leave'

transitives[39]

tuku 'buy'	éntuk/olèh 'get'
duwé 'have, own'	gawé 'make, do, use'

[38] This word belongs to a class of verbs which historically were formed with the active infix marker *–um-*, but which exist only in fossilized form in the modern language. Most words in this class show alternations with a nominal form and a verbal form with *N-*, such as *laku* 'walk, gait', *m-laku* 'to walk', and *ng-lako-ni* 'to do, carry out s.t.' *Mulih* however does not show any such alternations, and is very productive, so I include it here. I discuss the *M-* marked verbs in greater detail in chapter 5.

[39] The final two of this set sometimes appear with a prenasalized initial—*n-duwé, ng-gawé*. It appears to be a phonological property for some speakers. I have not been able to find any syntactic difference between those examples with and those without prenasalization, unless there is further suffixation, in which case the nasal is obligatory.

To these we can also add t*uru* 'sleep', which encompasses both the unaccustive 'to be sleeping' and the unergative 'to go to sleep'. In the discussion below on word order, I make use of these verbs as much as possible. It will become clear that many of these verbs are inherently neither transitive nor intransitive.

It will become increasingly clear that Tengger, as with many western Indonesian languages, allows for syntactically unexpressed constituents. In fact, it strongly prefers null constituents. This makes word order analysis extremely difficult but this is not always so. On the one hand, constructions and orders which may be perfectly grammatical may never come up in spontaneous speech recordings, no matter how many hundreds of hours – and there must be some discourse related reason for this. Alternatively, elicitations run the risk of presenting just such constructions which would never naturally occur in speech. The researcher must therefore rely on both methods, and generally proceed carefully. I base most of my work on naturalistic recordings. However, especially for the data in this chapter, I often check and try to extend that with elicitations.

4.2 Unmarked Word Order

It will be useful to be begin our discussion of unmarked word order in Tengger with some examples that as a first hypothesis might stand as strong candidates for representing unmarked word order. Based on the analysis of standard Javanese and many other closely related languages as SVO, we would expect that simple neutral sentences such as those below would represent the unmarked word order in Tengger:

4. a. Marsam budhal.

Marsam leave

'Marsam is leaving.'

b. Supari tuku gedhang.

Supari buy banana

'Supari bought bananas.'

Example (4a) shows a typical intransitive construction, with the actor argument preceding the verb. In (4b), we have a typical transitive construction. Here the agent argument precedes the verb and the patient argument follows. This might lead us as first to posit that Tengger has SV(O) or agent-verb-patient as its unmarked order. However, consider the following:

5. Budhal Marsam.

leave Marsam

'Marsam is leaving.'

The example in (5) above with the actor following the verb has exactly the same interpretation as (4a) above, where the actor precedes the verb, and can be uttered in exactly the same context. Remember too, that a simple transitive sentence allows for all of the six logically possible word orders, as shown in (1) above. Each of the examples above is grammatical. To be sure, some are highly contextualized, and many of them have distinctive intonation patterns. Nevertheless, each of the six examples in (1) above

can have multiple pragmatic structures, multiple theme-rheme orders. Intonation is most likely associated with these different pragmatic interpretations. The fact that there are multiple pragmatic interpretations for each of these six examples, demonstrates very clearly that there are not fixed meanings associated with any particular word order. Certainly there will be tendencies, certain pragmatic interpretations will correlate with certain intonation patterns, and these tendencies may be very strong. Nonetheless, they are not fixed patterns. For each of these examples, I have a number of instances of such word order in my recordings. In some cases, these are clear focus constructions, that is, some element in the utterance is given pragmatic salience, as indicated by intonation and possibly by linear order.

For the intransitive constructions, in elicitations the SV order found in (4a) above is preferred. However, in my recordings, there are actually a slightly higher percentage of examples where the actor follows the verb, or VS order (around 55% of examples with with an overt subject). There are almost as many examples of the actor preceding the verb. For the transitive constructions there is a preference for the orders in (1a) and (1f), that is SVO and VOS order. Note that in all of the above examples, the thematic roles of the participants do not change. The agent remains *Supari* and the patient remains *gedhang* 'banana' throughout. There is never any confusion as to bananas buying me. The participants in this type of example are semantically disambiguated. Note, though, that verbs with neutral participants, that is, those which cannot easily be determined semantically, also allow the same range of possible word orders. Generally, however, there is some voice marking on the verbs, and so I discuss those examples in greater detail in the next chapter. Consider the following here:

6. a. Kanca duwé ula.

friend have snake.

'My friend has a snake.'

*The snake has a friend.

b. Ula duwé kanca.

c. Ula kanca duwé

d. Kanca ula duwé

e. Duwé kanca ula.

f. Duwé ula kanca.

What is important here is that each of the above examples has available the interpretation where 'my friend' is the one who owns the snake. Under a different context, it could equally mean that the snake has a friend, both interpretations are 'available' and the context will determine which one is meant.[40]

Further, adjuncts—adverbs and prepositional phrases can be added to almost any position to both a transitive and an intransitive construction:

[40] Note that the construction in (6c) showing OSV word order has been analyzed by some for Javanese and other related languages as a passive construction. I discuss this in detail in the next chapter.

7. a. Wingi Marsam budhal.

yesterday Marsam leave

'Marsam left yesterday.'

b. Marsam wingi budhal.

c. Wingi budhal Marsam.

d. Marsam budhal wingi.

e. Budhal wingi Marsam.

f. Budhal Marsam wingi.

8. a. (wingi) Supari (wingi) tuku ??(wingi) gedhang (wingi)

yesterday Supari buy banana

'Supari bought bananas yesterday.'

b. (wingi) gedhang (wingi) tuku (wingi) Supari (wingi)

9. a. (Dhek pasar) Supari (dhek pasar) tuku??(dhek pasar)gedhang (dhek pasar)

at market Supari buy banana

'I buy bananas at the market.'

b. (Dhek pasar) gedhang (dhek pasar) tuku (dhek pasar) Supari (dhek pasar).

In both the second and third set of examples above, there are many more combinations of possible word orders with an adjunct adverb or PP which are grammatical. For the sake of space, I have only listed several. Note that with the intransitive verb, all combinations with a temporal adverbial are possible. This is also true of any number of other adjuncts.

The only inherently ditransitive verb to appear regularly without overt voice morphology is *kirim* 'send:'

10. a. Dukun kirim donga nang para leluhur.
priest send prayer to PLU ancestor
'The priest sends prayers to the ancestors.'
b. Dukun kirim para leluhur donga.

The above examples provide the pattern of a typical ditransitive double object construction.[41] In (10a) the agent precedes the verb, followed immediately by the theme. The goal here appears after the preposition *nang*.[42] In the subsequent example, the goal immediately follows the verb and there is no preposition. This is exactly parallel to many languages, including English:

11. a. The priest sends prayers to the ancestors.
b. The priest sends the ancestors prayers.

Unlike in English, however, other orders are also grammatical in Tengger. I do not provide an exhaustive list here. In the first set of examples I show possible word orders when the goal/recipient is introduced by a preposition. When the goal appears in a prepositional phrase, almost any word is possible:

[41] In the next chapter I discuss the word order and properties of ditransitives again in chapter 6.
[42] See below for discussion of the distinction between core and oblique arguments.

12. a. Nang para leluhur dukun kirim donga.

to PL ancestor priest send prayer

'The priest sends prayers to the ancestors.'

b. Nang para leluhur donga kirim dukun.

c. Nang para leluhur kirim donga dukun.

d. Donga nang para leluhur kirim dukun.

e. Donga nang para leluhur dukun kirim.

However, the goal does not need to appear in a prepositional phrase. However, in this case, there are restrictions on possible word orders:

13. a. Para leluhur dukun kirim donga.

PL ancestor priest send prayer

'To the ancestors the priest sends prayers.'

b. Para leluhur donga kirim dukun.

c. *Para leluhur kirim donga dukun.

d. *Donga para leluhur kirim dukun.

e. *Donga para leluhur dukun kirim.

What is immediately visible in the above examples is that the benefactive goal *para leluhur*, can appear in almost any position. Note however the order in the first example above. Here the benefactive goal appears in initial position, followed by the agent verb

patient. In this construction, as in the double object construction above, the benefactive does not obligatorily take a preposition.

In (13d) and (13e) above the patient appears in initial position. Here the goal must appear in a PP. Crucially, note the following:

14. Dukun kirim donga para leluhur.
priest send prayer PL ancestors
'The priest sends prayers to the ancestors.'

In this example the agent precedes the verb. However, note the order of constituents following the verb, the theme precedes an unmarked goal. We saw above that Tengger allows both the double object construction (13b), where the recipient goal immediately follows the verb. However, this example clearly demonstrates that there is no adjacency restriction, further, unlike in the 'typical' ditransitive construction where the recipient is in a PP, here there is no PP. This will factor very crucially into the discussion of applicatives in chapter (6). For the present, it underscores the variation in word order patterns attested in Tengger, extending that to inherent ditransitive verbs in addition to the intransitive and mono-transitives verbs already considered here.

If all of these orders are grammatical, does it even make sense to speak of an 'unmarked' order? First, to say that all attested word orders in Tengger are all grammatical is not to say that they are all equal. Secondly, there are a number of languages which allow various surface orders of elements but clearly still have an

unmarked order, perhaps Tengger is just such a language. I now turn to some of the available analyses in of Javanese word order in the literature.

4.3 Javanese and Tengger as Strict Word Order Languages

There have been very few analyses of the syntax of Javanese or dialects of Javanese. Of these, the works of two scholars stand out, and both are representative of their respective approaches. Uhlenbeck (1979, 1994, etc.), in a classical European structuralist tradition, provides the most thorough and detailed description and analysis of modern Javanese to date. He focuses on the standard dialects of Yogyakarta/Surakarta. Although the majority of his studies are on the morphology of Javanese, he also comments on the syntax and word order. As noted above, he claims that Javanese has almost completely free word order, based, among other evidence, on the examples below:[43]

15. a. Ing kala punika / Amat / manah-ipun / sakit.
at time that Amat heart-ASS sick
'At that time Amat was offended/heartsick.'

b. Ing kala punika / Amat / sakit /manah-ipun.

c. Ing kala punika / manah-ipun / Amat / sakit.

d. Ing kala punika / manah-ipun / sakit / Amat.

e. Ing kala punika / sakit / Amat / manah-ipun.

f. Ing kala punika / sakit / manah-ipun / Amat.

g. Amat / ing kala punika / manah-ipun / sakit.

[43] The Tengger equivalents of each of these were given near the beginning of this chapter.

h. Amat / ing kala punika / sakit / manah-ipun.

i. Amat / manah-ipun / ing kala punika / sakit.

j. Amat / manah-ipun / sakit / ing kala punika.

k. Amat / sakit / ing kala punika / manah-ipun.

l. Amat / sakit / manah-ipun / ing kala punika.

m. manah-ipun / ing kala punika / Amat / sakit.

n. manah-ipun / ing kala punika / sakit / Amat.

o. manah-ipun / Amat / ing kala punika / sakit.

p. manah-ipun / Amat / sakit / ing kala punika.

q. manah-ipun / sakit / ing kala punika / Amat.

r. manah-ipun / sakit / Amat / ing kala punika.

s. sakit / ing kala punika / manah-ipun / Amat.

t. sakit / ing kala punika / Amat / manah-ipun.

u. sakit / manah-ipun / Amat / ing kala punika.

v. sakit / manah-ipun / ing kala punika / Amat.

w. sakit / Amat / ing kala punika / manah-ipun.

x. sakit / Amat / manah-ipun / ing kala punika.

I have shown above that the same range of possible word order patterns are also available in Tengger, though some are clearly external possessor and standard genitive constructions.

Unfortunately the example that Uhlenbeck chose to demonstrate the extreme word order variation in Javanese contains a rather idiomatic phrase, ‘to be broken hearted’. I

have changed that in the Tengger data given above. Nonetheless, each of the constituents in the standard language can appear in any order, just as we saw with the Tengger.

That such a wide variety of word order patterns are possible in Javanese leads Uhlenbeck to claim that any sort of constituent, phrasal, or movement type analysis of Javanese is completely unproductive. Nonetheless, he did posit that there was an underlying word order. However, he claimed that the examples in (15: a, b, j, and l) above all represent the 'unmarked order', as he puts it. This in and of itself is telling, he is identifying four distinct word orders as unmarked for standard Javanese. These four constructions share a number of salient features. To begin with, the temporal adjunct in each of the four examples comes either initially or finally, crucially, it does not intervene between any of the other elements. This, though, is more of a comment on the addition of adjuncts than it is about unmarked word order. It is more natural for the adjuncts not to intervene between other material in a sentence, though it is still grammatical if it does. This is true for any number of languages. Another feature, however, is that Amat always precedes *sakit* and *manah-ipun*. Those two items, however, appear in any order. In the non-idiomatic reading (which is available in Tengger at least)[44], the 'heart' is the theme and 'sick' is the predicate, or S and V, and crucially here, either order is perfectly grammatical. Although he did claim that certain of the attested word orders were 'unmarked', the fact that Uhlenbeck himself identifies both SV and VS orders as unmarked suggests that minimally, there is no single underlying word order in Javanese.

[44] It is also striking that the idiomatic interpretation is still available no matter what the order of constituents is, even when there is intervening material.

A more recent analysis of Javanese syntax comes out of the generative tradition. In his analysis of Madurese and Javanese, Davies (1999) shows that on the surface, Madurese also seems to allow free word order[45]:

16. a. Hasan ngerem sorat dhaq Siti. [Madurese]
Hasan N-send letter to Siti
'Hasan sent the letter to Siti.'
b. Ngerem sorat dhaq Siti Hasan.
c. Ngerem sorat Hasan dhaq Siti.
d. Ngerem Hasan sorat dhaq Siti.
e. Dhaq Siti Hasan ngerem sorat.
f. Dhaq siti ngerem sorat Hasan.
g. Dhaq Siti ngerem Hasan sorat.

Davies 1999:155

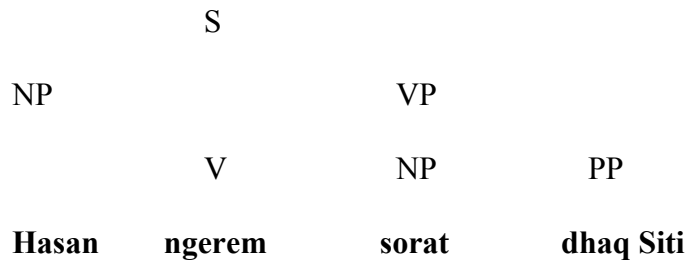
He further notes that 'word orders other than those shown in (16) are possible, but some of them, particularly those with the object occurring between the subject and the verb, are of relatively marginal acceptability (1999:155).'

Davies claims that both Madurese and Javanese are underlyingly SVO languages. Further, he refutes the claim that the above examples demonstrate that these are relatively free word order languages. He argues that the seeming freedom is a result of dislocation structures with co-indexed null pronouns in the place of the dislocated constituent, as has been proposed for some Australian languages.

[45] For his discussion, Davies uses Madurese examples, but claims that the same holds for Javanese. It is not made explicit which dialect of Javanese he used for his examples, but from the forms it is clear that they are from the standard Yogyakarta/Surakarta dialect.

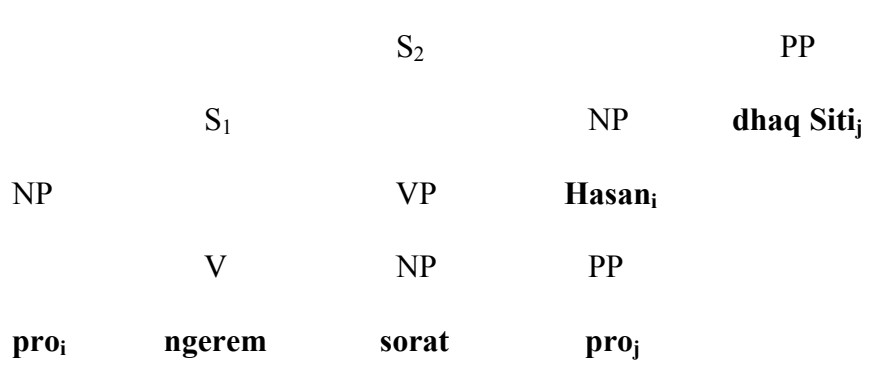
For the unmarked order in (16a) he proposes the following, very general structure:

17.



In order to derive, for example number (16c) above, he proposes that there are two distinct higher projections, S_2 and S_3 which are available positions for adjoining constituents. Further, he claims that the base structure given in (17) above therefore remains constant. The elements that are merged higher in the structure are replaced by co-indexed null pronouns in the base structure:

18.



While this analysis is parsimonious in maintaining a single underlying structure, it also adds a significant amount of unnecessary structure and movement, although it is one of the "standard" analyses of true (split constituent) free word order languages. To begin

with, it is not clear, within the generative framework itself, how the arguments here would receive case, though possibly the null pronominals occupy theta positions and are linked to adjoined XPs – thus they are linked to Case positions, particulary if the external argument raises to a Case position before being adjoined . Further, this analysis adds at least two distinct projections and postulates null elements that are otherwise completely unmotivated. There is nothing simple about this analysis except that it maintains a single underlying word order.

To motivate the existence of the null pronouns, Davies shows that Madurese has a number of what he calls dislocation structures, where constituents are moved rightward.[46] He also shows that in the place of the dislocated, or moved element, a co-indexed pronoun can appear.

19. a. Abaqeng$_i$ ngerem sorat dhaq Siti Hasan$_i$ [Madurese]
3.sg N-sent letter to siti Hasan
'He sent the letter to Siti, Hasan.'

b. Abaqeng ngerem sorat Hasan dhaq Siti.
3.sg N-send letter Hasan to Siti
'Hasan sent the letter to Siti.'

These are both dislocation structures, corresponding to (16 b & c) respectively above. Davies claims, essentially, that they are the same constructions, fully synonymous and with the same discourse function.

[46] We see this in simple right dislocated examples in Italian – *pro* ha telefonato Gianni.

He also notes here that many, if not most of the examples he uses are rather unnatural. Madurese, like Javanese and most other western Indonesian languages prefers null constituents. So it is difficult to get grammaticality judgments on examples such as these through elicitations, let alone to find them occurring in natural speech. Nevertheless, Davies then postulates that the only difference, for example, between (19b) and (19a) is that the latter has an overt pronoun in initial position. He claims that (19b) is also a dislocation structure, with a null pronoun in initial position.

20. a. Sorat jhuwa e-kerem biq Hasan dhaq Siti. [Madurese]
letter that di-sent by Hasan to Siti
'That letter was sent to Siti by Hasan.'

b. Sorat jhuwa e-kerem dhaq abaqeng biq Hasan dhaq Siti.
letter that di-send to 3.sg by Hasan to Siti
'That letter was sent to Siti by Hasan.'

While the Tengger version corresponding to the example in (20a) would be acceptable, the corresponding version of (20b) would not:

21. a. Layang iku di-kirim mbèk Hasan nang Siti. [Tengger]
letter that di-send by Hasan to Siti
'That letter was sent by Hasan to Siti.'

b. *Layang iku di-kirim nang dhéké mbèk Hasan nang Siti.
letter that di-send to 3 by Hasan to Siti
'That letter was sent to her by Hasan to Siti.'

If analyzed as part of a contiguous structure, then (20b) above is completely ungrammatical. The same holds true for the standard Javanese version, and I would also venture for the Madurese.[47] The problem is the analysis of these sentences as unitary structures.

Beyond positing several layers of otherwise unmotivated structure, and using data that Davies admits is unnatural, the main problem with his analysis has to do with what he claims are dislocation structures. This is the core of his argumentation, remember. He is claiming that the seeming free word order in Javanese and Madurese is actually the result of a number of sometimes cyclic dislocation constructions. There is a single underlying word order, and though all logically possible word orders for a simple transitive sentence, for example, are grammatical, they are all derived through dislocation. Javanese and Madurese then, have rampant and seemingly unrestricted dislocation.[48] The issue though, comes with is claims of what actually represents a dislocation construction. Consider the following:

[Madurese]

22. a. Sorat jhuwa e-kerem dhaq abaqeng biq Hasan dhaq Siti.
letter that di-send to 3.sg by Hasan to Siti
'That letter was sent by Hasan to Siti.'

[47] I have not collected any Madurese data, so I do not want to make any strong claims. However, given the issues I raise with Davies' Javanese data, of which I have collected much, I venture that the same issues apply to the Madurese. Nonetheless, I will confine my comments from here to the Javanese.
[48] Though such an analysis has also been made for Malagasy (underlying SVO, but with right/left adjoined dislocation structures (topic/focus etc.) by Potsdam (2007).

b. Burus jhuwa ngekeq abaqeng neng sakolaan dhaq anaq-eng Hasan.

dog that N-bite 3.sg as school to child-DEF Hasan

'The dog bit Hasan's child at school.'

As mentioned above, neither of the corresponding examples of these sentences is grammatical in either Javanese or Tengger, if analyzed as single utterances. The Tengger and Javanese are actually remarkably similar to the equivalent sentences in English:

23. a. *The letter was given to her by Hasan to Siti. [English]

b. *The dog bit him at school Hasan's child.

As they stand, both of these examples in English are ungrammatical. That is, if they are to be considered a single unit or unitary construction. However, such linear orders are quite possible in English. In the first example, if there is a pause either between *by* and *Hasan,* or between *Hasan* and *to*, then the sentences dramatically improve:

24. a. The letter was given to her by Hasan, to Siti.

b. The letter was given to her, by Hasan to Siti.

For me the first example is better, but both are acceptable. However, no one would claim that these examples represent a single structure. In (24a) *to Siti* is an afterthought, a

clarification, an echo. Similarly in (24b) the clarification is much more robust, *by Hasan to Siti* is added as an afterthought and not part of the original utterance. The same is true of the other English example in (23b) above. If there is a pause between *school* and *Hasan* then the sentence improves:

25. The dog bit him at school, Hasan's child.

It is obvious here that *Hasan's child* is an afterthought. It is not part of the same structure as the rest of the utterance, and hence does not represent a dislocation structure. The same is true of both the Tengger and the Javanese examples. When there is no pause, they are completely ungrammatical. However, when there is a pause, they are much improved. It should be noted, that this is not a matter of simple intonation contours. Davies does discuss the intonational patterns of the Madurese examples. To begin with, these do not necessarily apply to the Javanese. But more importantly, he is conflating structures that need to be measured independently. It is not a rising or falling contour that is crucial to render the English, Javanese and Tengger examples grammatical, it is a prosodic break sufficient to mark the following material as separate.

[Tengger]

26. a. Layang iku di-kirim nang dhéké, mbèk Hasan nang Siti.
letter that di-send to 3 by Hasan to Siti
'That letter was sent to her by Hasan to Siti.'

b. Kirik iku nyokot dhéké dhèk sekolah, anak-é Hasan.

dog that N-bite 3 at school child-é Hasan

'The dog bit him at school, Hasan's child.

Again, these examples are only grammatical with prosodic breaks, in exactly the same positions as the equivalent English examples above. They are afterthoughts, and as such, are distinct structures.

Davies is the only recent work to address word order in Javanese and Madurese. His claim that they are, contrary to appearances, strict word order languages posits unmotivated structure, relies on phonologically null elements, and crucially derives all of the non-base word orders via a series of dislocation constructions – as has been done for a number of languages with apparent discourse governed dislocation structures of this type. I have shown here that what he claims are dislocation structures, at least for Javanese and Tengger, are in fact distinct structures, as they are ungrammatical if considered as part of the same unit. This does not necessarily show that Javanese and Tengger are in fact free word order languages, but crucially, it adds no evidence that they are strict word order languages.

4.4 Constituency In Tengger

As mentioned above, I make no claim that Tengger is in fact a free word order language. At best I put forward the hypothesis that Tengger is a free phrasal order language – a scrambling type language. In order to assess this question, though, we must

first demonstrate that there are in fact, distinct constituents in Tengger. In the preceding section on grammar, I have demonstrated that there are in fact distinct NPs in Tengger. Here, specifically, we will be looking at whether or not there is a distinct verb phrase in Tengger, that is, is there an especially close syntactic bond between a verb and its complement(s). By complement here, I mean the verb and all non-subject arguments. (I will also briefly demonstrate that Tengger has a PP, though it comes very close to resembling a VP). This is by no means to be taken for granted, as it has been claimed for many languages with apparent free word order that there is no underlying structure. However, there is a huge amount of work on "free word order" languages (Australian, Hungarian and many others), addressing exactly this question.

The question of word order is tightly bound with the issue of the constituency, which in turn is also impacted by the nature of categories. I turn to that next.

4.4.1 A Note on Categories in Tengger

Before moving on to a discussion of constituency in Tengger, it is necessary to address the issue of categories, or word classes in TJ. As Evans (1985) notes, 'From Dionysius Thrax onwards it has been normal grammatical practice to define parts of speech through a combination of morphological, syntactic, and semantic criteria. This works well enough when all three coincide.' (84) However, there are many languages where these properties do not coincide, and it is not clear whether a particular lexeme is, for example, a verb, a noun, or both. Further, it has been common practice to assume that lexical categories which were originally developed to account for Greek and Latin grammar are

universal, and that all languages are neatly defined according to these same categories. Whorf (1945:1) noted the problem long ago, 'the very natural tendency to use terms derived from traditional grammar, like verb, noun, adjective, passive voice, in describing languages outside of Indo-European is fraught with grave possibilities of misunderstanding.' As noted in Malouf (2000:3), Sapir and Bloomfield both reject the notion or utility of parts of speech as grammatical primitives.

> A part of speech outside of the limitations of syntactic form is but a will o'the wisp. For this reason, no logical scheme of the parts of speech—their number, nature, and necessary confines—is of the slightest interest to the linguist. (Sapir 1921:118f)

> The term *parts of speech* is traditionally applied to the most inclusive and fundamental word-classes of a language, and then […] the syntactic form classers are described in terms of the parts of speech that appear in them. However, it is impossible to set up a fully consistent set of parts of speech, because the word-classes overlap and cross each other. (Bloomfield 1933:196.)

And yet, given these early observations and caveats, the generative tradition has always tacitly assumed that a certain limited number of syntactic categories were primitives and universal, 'for the most part, the structuralists' concerns have been ignored by generative grammarians.' (Malouf 2000:4) Malouf goes on to explore the nature of verbal gerunds

in the Head-driven Phrase Structure Grammar framework, examining the nature of 'transcategorial' or mixed category constructions, largely overlooked by other scholars.

Chomsky's X'-theory (1970) does make some attempt to define the nature of syntactic categories. He decomposes the major 'lexical categories' into terms of the two primitive features [+/- N] (substantive) and [+/- V] (predicative) to derive the following schema:

27. N(oun) = [+N, -V]
A(djective) = [+N, -V]
P(reposition) = [-N, -V]
V(erb) = [-N, +V]

This analysis avoids claiming that categories themselves are primitives, but rather the features associated with them are taken as primitives. Baker (2003:2) notes, though, that:

> 'this theory is widely recognized to have almost no content in practice. The feature system is not well integrated into the system as a whole, in that there are few or no principles that refer to these features or their values… All the features do is flag that there are (at least in English) four distinct lexical categories.'[49]

[49] The same criticism is leveled against Jakendoff's (1977) alternative account, using the features +/-subj and +/-obj instead of N/V, which provides for different natural classes of categories, but is no more productive.

In his extensive study, Baker contributes significantly to our understanding of lexical categories. For instance, in differentiating verbs on the one hand, and nouns and adjectives on the other, he notes 'that only verbs are true predicates, with the power to license a specifier, which they typically theta-mark. In contrast, nouns and adjectives need help from a functional category Pred in order to do this.' (2003:20). While this analysis works well with one I developed in the preceding chapter, also making use of a PredP for small clauses and other phenomena, it is striking that Baker's overall analysis provides the following, hardly radical, schema:

28. Noun is +N = 'has a referential index'
    Verb is +V = 'has a specifier'
    Adjective is –N, -V
    Preposition is part of a different system (functional)

Schacter (1985) claims that a number of Austronesian languages, among which Javanese can be included, make no distinction between nouns and verbs. Further, it is not clear if a distinction exists between verbs, adjectives, and prepositions (to be discussed shortly) either. While Baker noted such a potential problem, his analysis does little to resolve the dispute, as he ultimately concludes that there are still three universal, primitive lexical categories, with adpositions as a fourth, functional category.

While I do not here formulate a radically new approach to syntactic categories, I do wish to present the TJ data, and show how it bears on the question. Gil (2001:116)

makes the case for a single syntactic category in Riau Indonesian, a language closely related to TJ, with the following sentence and multiple glosses:

29. Ayam makan. [Riau Indonesian]
chicken eat

Gil argues that there is no categorial distinction between *ayam* 'chicken' and *makan* 'eat' in Riau Indonesian. This is based on distributional evidence. He claims that these items, as representatives of the categories of 'noun' and 'verb' respectively, have exactly the same distribution in Riau Indonesian. For example, they can both occur as one word utterances, take the same range of complements—of both categorial type, be modified by the same range of modals and functional words, quantification, in addition to a number of other shared properties. These same arguments can made time and again for any number of Indonesian languages, especially colloquial varieties of Indonesian.

In fact, it has been argued that several related languages in fact make no categorial distinctions between lexical elements (see, e.g. Gil 2001, on Riau Indonesian.) I find Gil's arguments for the lack of category distinctions within Riau Indonesian to be very convincing. In fact, I think the same is true at least of Jakartan Indonesian.[50] Consider the following examples:

30. a. Lagi makan. [Jakartan Indonesian]
PROG eat
'I'm eating.'

[50] From my own data.

b. Lagi ayam.

PROG chicken

'I'm (eating) chicken.'

31. a. Makan.

eat

'I'm eating.'

b. Ayam.[51]

chicken

'I'm eating chicken.'

32. a. Setiap ayam.

each chicken

'Each of the chickens.'

b. Setiap makan.

each eat

'Each time (I) eat.'

For each of the six examples above, multiple interpretations are possible. I have chosen one for each case to serve as an example translation. This by no means suggests that other translations are not possible. Strikingly, the above examples show three different constructions which generally allow either a verb or a noun to appear, to the exclusion of the other. While some of the interpretations would be context dependent, what is crucial here is the distributional facts. The first set of examples shows the progressive marker, a

[51] Remarkably, this can be uttered unsolicited. It is generally considered polite to announce that you are eating when someone walks in. On occasion, instead of saying *makan* 'I'm eating', one can state *what* one is eating, hence, *ayam.*

functional word, combining freely with both *makan*, as expected, and with *ayam*, as perhaps not expected. Similarly, in the subsequent set of examples, both words can appear independently as one word utterances. Finally, in the third set of examples above, we have a quantification construction. As the grammaticality of both examples attests, *ayam* can appear following *setiap* 'each' again perhaps as expected, but *makan* can also appear following *setiap*. This demonstrates that in Jakartan Indonesian, as in Riau Indonesian, there is no distinction between syntactic category types, there is no syntactic noun-verb distinction.[52] It is striking, but these can be true, independent utterances.

However, the same does not hold true of Tengger. Minimally, Tengger makes a categorical distinction between nouns, or substantives, and verbs, or predicates.[53]

There are two broad classes of modals in Tengger, which I discuss in detail shortly. One of the defining characteristics of each group is the complement type they select for. We will take here *katé* 'want, will' as representative of one class of modals and *gelem* 'want, will' as representative of the second class of modals. I should note that this distinction is not present in the standard language, which lacks *katé* altogether, but it does extend to the eastern Javanese lowlands dialects in addition to Tengger. The standard language uses *arep*, which behaves quite differently from *katé.*

The distribution of *katé* and *gelem* is the same. However, they do not take the same types of complements:

[52] A full discussion of Jakartan Indonesian is beyond the scope of this paper. Suffice it to say, there are, minimally, distinct functional categories in addition to the single lexical category. These are syntactic categories, I am not making any claims about semantic categories.

[53] I discuss the position of prepositions below. For the time being, I leave the question of adjectives open. I can find no compelling evidence to claim that adjectives are a distinct category from verbs in Tengger. This is interesting in light of the discussion of Baker above, where he groups nouns and adjectives together as opposed to verbs. The Tengger data would suggest a very different classification.

33. a. Isun katé nyambut gawé ora bisa.

b. Isun gelem nyambut gawé ora bisa.

1.f want connect make NEG can

'I wanted to work but I couldn't.'

34. a. …katé mentas nang nggon-é mbah-é kana.

b. …gelem mentas nang nggon-é mbah-é kana.

want go.up to place-é grandparent-é there

'I wanted to go up to my grandparent's place up there.'

35. a. *Nggo mangan éyang katé sega.

b. Nggo mangan éyang gelem sega.

for eat 1.m want rice

'To eat, I want rice!'

36. a. *Pak Inggi katé peraturan anyar.

b. Pak Inggi gelem peraturan anyar.

Pak Inggi want regulations new

'The (village) chief wants new regulations.'

In examples (33) and (34) above, *katé* and *gelem* are almost interchangeable. The difference between them is slight, and purely semantic. *Gelem* has more of a desiderative connotation, indicating a desire on the part of the agent or actor. In (33), the modals take the (compound) lexical verb *nyambut gawé* 'to work' as a complement; in (34) the lexical verb *mentas* 'to go up.' In every way here, the (a) examples are the same as the (b) examples.

In both (35) and (36), however, there is a difference. In both of these examples, the sentence with *katé* is ungrammatical where the sentence with *gelem* is grammatical. What is the difference? The word order is the same, the valency is the same. The difference is in the class of complement. In (35) the modal is followed by a substantive, *sega* 'rice' and only (35b) is grammatical. Similarly, in (36) the modal is followed by *peraturan* 'regulation,' an abstract substantive. Again, only (36b) is grammatical.

These two modals can take other types of complements:

37. a. Sapa katé nang Tumpang?
who wants to Tumpang
b. Sapa gelem nang Tumpang?
who wants to Tumpang
'Who wants to go to Tumpang.'

Here the modals are followed by a prepositional phrase—to be discussed in detail in the following subsection. What is important is that there is no distinction between what type of complement the two modals can take in terms of prepositions.

The asymmetry between these two (classes of) modals is strong evidence suggesting that Tengger in fact makes a grammatical distinction between at least two categories or words: nouns and verbs. This distinction holds for an entire class of words. To be specific, note the actual distinction. The modal *gelem* has no restrictions on what it can take as a complement. As such, it does not really provide any evidence for any kind of categorical distinctions in Tengger. In contrast to *gelem*, however, *katé* does place

restrictions on what it can take as a complement. A broadly delimited class of substantives, abstract and concrete, proper and common, or nouns, cannot serve as the complement of *katé*. Whether or not the set of nouns in Tengger is co-extensive with nouns in other languages, that is, whether there is a universal category of 'noun', is beyond the scope of the present discussion.

What the above data makes very clear, however, is that Tengger does have a syntactice noun/verb distinction. Remember example (29) above, where Gil argues that Riau Indonesian does not have a syntactic noun/verb distinction, based on distributional evidence. Similar claims have been made for other related languages such as Gil (1994) on Riau Indonesian, or Donohue (2004) for some languages of eastern Indonesia, and it may turn out to be a feature of some western Indonesian languages that they do not make a syntactic distinction between nouns and verbs. Here though, I have adduced evidence demonstrating that Tengger makes a significant distributional distinction that clearly delineates a syntactic category of noun, and by default verb (or non-noun).

4.5 Constituency and Categories: Prepositions in Tengger

Having argued for a clear categorial noun/verb distinction in Tengger, it makes sense to ask if there are other syntactic categories that operate in the language. Further, it will assist our discussion of constituency if we are able to identify any other distinct categories. Remember, that I argued for a distinct NP in chapter 3, now we will explore if Tengger has prepositions and prepositional phrases.

It is an empirical question to see if a language has different categories and phrasal projections. I have already argued for an NP in Tengger in chapter 3. In the preceding section, I argued for both an N and a V. In this section I will explore whether or not Tengger has a P and/or PP, before returning to the more complicated discussion of the VP.

It is well known that the semantic and syntactic distinction between verbs and prepositions is not uniform across languages. The mainland Southeast Asian languages for example, are often considered to not make a categorical distinction between verbs and prepositions, or at the very least not the same type of distinction that is clearly present in say, western European languages (see e.g. Clark 1978). In Tengger we have examples like the following:

38. a. Éyang mlaku teka Gubukklakah.

1.s walk from Gubukklakah

'I came from Gubukklakah.'

b. Éyang teka Gubukklakah.

1.m from Gubukklakah

'I came from Gubukklakah.

c. Éyang teka.

1.m from (this must be in context?)

'I came (from Gubukklakah).'

In (38a) above the verb *mlaku* is followed by what appears to be a preposition *teka*. However, in the second example, it appears as if *teka* itself is the verb. Finally, in the last example, we have only *teka*. Further, the complement of the preposition need not be concrete:

39. a. Éyang mlaku teka jam loro.
1.m walk from hour two
'I've been walking since two o'clock.'

b. Éyang teka jam loro.
1.m from hour two
'I came at two o'clock.'

The same set of examples can be reproduced for many 'prepositions' in Tengger. In fact, in stative clauses, there is a strong preference for just the 'preposition', though that is due to a general preference for a null copula.

That the locative goal need not be preceded by a preposition is repeated in any number of other verbal constructions. And the verbal behavior of *nang* is also the case with almost every other preposition in the language:

40. a. lunga: to go (out, away)
b. teka: to come (from)
c. anjlog: to jump down (on, onto)

41. a. mbèk: (to be) with

b. nang: (to go) to

c. nèng: (to be) at, in, on

d. dhèk: (to be) at, in, on[54]

e. teka: (to come) from

f. karo: (to be) with, by

g. nganti: (to go) until

It is interesting to note that historically, many Tengger pronouns actually either contain verbal morphology, or they are the result of the coalescence of a verb with another element. For example, *karo* 'with' is derived from the base word *ro* 'two' with the passive marker *ka-*, thus *ka-ro* 'to be together, to pair.'[55][56] *Nèng* 'in, at, on' is a general stative locative. Historically, it derives from *ana ing* 'to be at', which with normal Javanese sandhi rules yields *anèng.* The initial /a/ was subsequently lost. Further was can see in nasalization and prenasalization in *nganti, nang,* and *mbèk*. *Teka* can stand as both a full verb and a preposition, in the standard language. Tengger has no independent form for the preposition, but the standard language also has *saka* 'from', so we can get examples like:

42. a. Aku wis teka saka Semarang. [Standard Javanese]

1 pft come from Semarang

'I already arrived from Semarang.'

[54] This is a regional variant, common in Tengger and EJ, unknown elsewhere.
[55] This base form of **ro* 'two' never appears independently. However the normal cardinal number, *loro,* is derived from the reduplicated base **ro-ro* which then underwent dissimilation.
[56] There is also a yearly festival celebrating the dualities of existence called *Karo*.

b. Aku wis mangkat teka Semarang.

1 pft M-leave from Semarang

'I already left from Semarang.'

Here we can see the clear verbal usage of *teka* in (42a) where it is preceded by an auxiliary and followed by preposition. In (42b) we can see where *teka* behaves more like a preposition. The case of *teka* however, may be different from the other 'prepositions', and there may in fact be two distinct lexemes. I discuss these and other uses of *teka* in the section on grammaticalization below.

The other prepositions can also appear after an auxiliary, however, they cannot stand alone or in final position:

43. a. Isun gek nang omah Joko. [Tengger]

1.f PROG to house Joko

'I'm on my way to Joko's house.'

b. Gek m-laku.

PROG M-step

'He's on the way.'

c. *Gek nang.

PROG to

'He's on the way.'

So although this class of prepositions can appear without a lexical verb, and they can appear after modals and in the position of verbal elements, they do not have the same distribution as the prepositions discussed above. On a cline of properties which are identified as either verbal or prepositional then, they are clearly more verbal than their counterparts in, say, European languages. However, with the exception of *teka*, they probably do not go quite as far as toward full verbal behavior as some of the mainland Southeast Asian languages. One way to capture this, is to differentiate between semantic and syntactic categories. Semantically, then, there seems to be very little difference between prepositions and verbs in Tengger. And while they seem to occupy many of the same syntactic positions (43a-c) above, it is clear that they have a different syntactic distribution over all as evidenced by the ungrammaticality of (43c).

Further, having demonstrated that there are in fact distinct prepositions in Tengger, however 'verbal' they may be, there is also clear evidence that Tengger has a distinct prepositional phrase:

44. a. Éyang nang Tumpang.
1.m to Tumpang
'I'm going to Tumpang.'

b. Nang Tumpang éyang.

c. *Éyang Tumpang nang.

d. *Nang éyang Tumpang.

e. *Éyang nang wingi Tumpang.
1.m to yesterday Tumpang

There is clear adjacency here between the preposition and its complement. Further, the whole phrase can appear in initial or final position, but nothing may intervene. This is true even of temporal adverbs which otherwise have a very free distribution (44e).

Further, it is possible to coordinate PPs:

45. a. Engko sik, brambang pré-né mbèk sapi-né mbèk pitik-é iku
later first onion leek-é with cow-é with chicken-é that
mau kenèk di-gawé tuku paran?
just.now can di-make buy what
'Wait a sec, the leeks and the cow and those chickens can be used to buy what?

b. Banyu-né trus nang tengen nang kiwa.
water-é continue to right to left
'The water just continued to flow to the right and to the left.'

The first example above shows coordination of PPs introduced by the preposition *mbèk* 'with'. Further, note that the NPs is marked with the associative marker *–é*. The demonstrative pronoun here, followed by the temporal adverb combine to mean something like 'those just now (mentioned)'. Together, they are modifying each of the distinct PPs, so the leeks, cow, and the chicken, all of which had just been discussed. In the second example is simple coordination of PPs, here headed by *nang*.

The word order and coordination facts are more than sufficient to posit the existence of a PP in Tengger. Further, I have demonstrated that unlike the clause in general, the order of constituents within the PP is much more rigid. There is indeed much more to said about the structure of the PP in Tengger, and especially interesting would be a more detailed comparison of the most 'verbal' like aspects of prepositions, and to what degree the structure is similar to that of other predicates. That, however, is outside of the domain of the present study.

4.6 Is There a Verb Phrase in Tengger?

Before moving on to an analysis of the various attested word order patterns in Tengger, it is necessary to address further issues of constituency in Tengger. First, to address the most extreme possibility, briefly, Tengger is a configurational language. It has been argued for a number of Austronesian languages that they are, in fact, non-configurational. Baker (2001) has recently proposed a typology of non-configurationality which argues for three distinct types of non-configurationality, and in fact argues that languages can display certain properties of non-configurationality while remaining configurational languages. The primary tests for differentiating among the three types come from the binding of anaphora, and Principle C of binding theory, and weak crossover effects. In languages where anaphoric relations are sensitive to word order and not to grammatical functions, exemplified by Japanese, Baker argues for a movement analysis of non-configurationality. Baker then makes a distinction between pure head marking languages (Nichols 1986, 1992) such as Mohawk, for which a dislocation analysis is provided, and

dependent marking languages with rich case morphology, such as Walpiri, for which a secondary predication analysis is given. Tengger would seem to fall into the second category. However, as we have shown above, a dislocation analysis does not explain all of the attested word orders demonstrated for Tengger.

It is very difficult to come up with productive tests for constituency in a language that strongly prefers null constituents. Much of the normal arsenal of tests focuses on replacements, ellipsis, and substitutions. In Tengger, of course, a null element is preferred in all cases. Further, movement tests are unhelpful for a language that allows such loose word order. This leaves us with relatively few tests to apply.

We have also already noted particular intonation contours that occur naturally in Tengger. Perhaps prosodic domains provide the best initial indication of constituents in Tengger. Remember that the elements in a bare transitive verb construction can appear in any order:

46. a. Marsam tuku gedhang.
Marsam buy banana
'Marsam bought bananas.'

b. Tuku gedhang marsam.

c. Gedhang tuku marsam.

d. Tuku marsam gedhang.

e. Gedhang marsam tuku.

f. ?Marsam gedhang tuku.

Each of the examples in (46) can appear with different intonation patterns, and they can appear with multiple intonation patterns—however, there are some strong preferences which can be ascertained. In (46a) there are no breaks, the entire utterance is one prosodic unit. Similarly, (46b) can be said without any pause or break, and can constitute a single prosodic unit. It can also be said with a pause following *gedhang*, separating the subject in final position. In (46c) there is a clear break after *gedhang*, marking *tuku marsam* as a single prosodic unit. Similarly in (46d) *tuku marsam* form a single prosodic unit, and there is a break before *gedhang*. Again in (46e) the subject and verb constitute a single prosodic unit, with a break following *gedhang*. Although only marginal, for the present discussion I gloss over that and treat (46f) equally with the other patterns. Here, the subject *marsam* is followed by a break, and *gedhang tuku* forms a prosodic unit.

So what conclusions can we draw from the prosodic evidence? To begin with, the prosodic evidence points to two different potentially unmarked word orders, those found in (46a & b) above, SVO and SOV. These are the only two in which the verb and both its arguments can form a single prosodic unit. However, (46b) also points to the predicate forming a constituent with its complement *tuku gedhang*. This is supported by the prosodic grouping in (46f) as well. However, (46c, d, and e) all point to the predicate and the agent forming a constituent *marsam tuku*—leaving aside for the moment that the ordering within the proposed constituents is variable. What we are looking at here is simply how the prosody of the utterances groups the various elements together, and thereby suggests constituency. So, what are we to make of this data? Two of the examples point to the verb forming a constituent with the theme, and three examples

actually suggest that the agent and the verb form a constituent, a highly unusual state of affairs to say the least!. Note that there is no evidence of the two arguments forming a constituent.

It is possible to argue that one or even more of the above examples represents a topicalization construction. That would fail however, to account for the difference between, say (46c) and (46e), where the patient is topicalized, but the order of other elements is reversed. I have already laid out arguments for why an unconstrained, dislocation analysis is also not sufficient to account for the Tengger data. It may just be that, as with replacement and substitution tests, prosody is not a reliable gauge for determining constituency in Tengger. That is, the phonological domain here does not necessarily reflect the constituency of the syntactic domain.

Himmelmann (2005) notes that there is strong evidence however, to suggest that most western Austronesian languages have a VP containing the predicate and all non-subject core arguments. Apart from clitics and a restricted class of adverbs, no element can intervene between the predicate and its complement/non-subject core arguments. So in most western Austronesian languages the subject and any adjuncts cannot appear adjacent to the predicate separating it from its complement. He gives some examples from Totoli to demonstrate the point (citing also Donohue (1999:151) for Tukang Besi and Artawa (2001:15) for Balinese):

47. a. gaukan [$_{VP}$ no-gutu ponguman itu] [Totoli]
king *AV.RLS-make story* *DIST*
'(Yesterday) the king made this announcement…'

b. *nogutu **gaukan** ponguman itu

c. ***ponguman itu** nogutu gaukan

In the second example above, placing the subject argument between the predicate and its complement results in ungrammaticality. Similarly, the sentence is also unacceptable if the two arguments switch position (without a concomitant change in verbal morphology). Himmelmann does point out that if the theme argument appears in initial position, *ponguman itu, gaukan nogutu*, then the sentence is grammatical. However, he argues that this is clearly a topicalization construction, and the theme argument constitutes its own prosodic unit here.

However, although these arguments hold for most western Austronesian languages, the situation is different in Tengger:

48. a. Ratu (ng)-gawé pengumuman iku. [Tengger]
king N-make announcement [IND] that
'The king made that announcement.

b. (Ng)-gawé ratu pengumuman iku.

c. Pengumuman iku gawé ratu.

Unlike the Totoli examples, each of their equivalents in Tengger is also grammatical. Note specifically (48b) where the agent intervenes between the predicate and its complement. The nasal marker on the verb here is optional, that is an idiosyncrasy of this particular verb, and also in Tengger. For (48c), we have the reversed order of elements.

Semantically and pragmatically it is clear here what the thematic roles are, the king is clearly the agent. These are clearly not determined syntactically. In this example, it is more natural to have the verb without the nasal prefix. This is a complex issue to which I devote the entire next chapter. For the time being, it suffices to point out that those criteria for determining the existence of a VP in other western Austronesian languages shed no light on the Tengger case.

As with the discussion of PPs above, coordination is one of the few constituency tests that yields some results for Tengger. Coordination, it should be noted however, is one of the least prescriptive constituency tests, in that it often identifies potentially overlapping constituents. [57] However, given the difficulty of most of the other standard diagnostics, it is very much worth while to examine the coordination facts.

Unlike the standard language which makes use of *lan* 'and', Tengger uses a preposition macro-functional element *karo* for positive conjunction and in most coordination constructions. I discussed the prepositional uses of *karo* in the preceding section, but the range of uses of this element is rather wide. It can be a conjunction, a coordinator, or mark comitative, instrumental, agentive, dative, ablative, among many other functions [58]. There are some other elements which can serve to coordinate structures, such as *trus* found in the first example below.

In Tengger, a verb-complement unit can be coordinated:

[57] Some authors advocate an idea of 'flexible constituency' (Ades & Steedman 1982, Dowty 1988, Pickering & Barry 1993, Pesetsky 1995).
[58] A full analysis of *karo* is well beyond the scope of the present study. However, it is interesting to note that is has many similarities with the use of *sama* in many colloquial dialects of Indonesian, especially Jakartan Indonesia. See also Gil (2002) on *sama* in Riau Indonesian.

49. a. Marsam ng-gawé omah dhéwé trus ngé'i karo anak-é.
Marsam N-make house self continue N-give with child-é
'Marsam made the house himself and gave it to his daughter.'

b. Anak-é Supari gelem-é ng-rampung-na sekolah karo ng-golek
child-é Supari want-é N-finish-na school with N-look.for
kerja nèng ngaré.
work at lowlands
'Supari's child wants to finish school and look for work in the lowlands.

The first example above is not unambiguously a coordination construction. Although there are two verb-complement units in sequence, and they share the same agent, expressed only once, *Marsam*, it is possible that *trus* here is in fact a complementizer subordinating the second clause to the first.

The second example, however, is a clear coordination construction. Under some analyses of standard Javanese and related languages, these constructions are actually nominalizations, with the *–é* marker on *gelem* marking a nominalization. In this case, the sentence could be translated as 'Supari's child's wish is to…' I addressed the question of nominalizations briefly in the general grammar section, and it is not particularly germane to the present point. Regardless of the analysis of the preceding element as a N head or an auxiliary (I or T) head, the point remains that there are clearly two VP like structures coordinated after them. Further, from background knowledge, all participants here know that Pak Supari's child is currently going to school in the lowlands, and looking for work

there. So the VP like structures also coordinate with the following PP. This example does not mean that Pak Supari's child is going to school in the village and then wants to look for work in the lowlands.

So a verb and a patient argument (49a) or theme argument (49b) can be coordinated. This suggests that the verb and its complement argument form a constituency. Consider also the following:

50. a. Biasa, isun masak tapi bapak karo arèk-arèk mangan sega sik.
usual 1.f cook but father with RED-kid eat rice first
'It's typical, I cook (the rice) but dad and the kids eat the rice first.'

It should be noted that this is not a typical sentence pattern. I have found several similar examples though in my recordings. Here, an agent argument and verb sequence is repeated twice. The single patient, *sega* 'rice', is clearly the patient of both verbs, *masak* 'cook' and *mangan* 'eat'. I was talking in the kitchen with a woman who was in the process of cooking, rice and side dishes, and she was, not so much complaining as pointing out how things usually went, she cooked, they ate. It should be noted here the *sega* 'rice' is often used in Tengger, and in many Indonesian languages, to refer to food in general, and any kind of substantive eating—as opposed to snaking. So, the sentence can be paraphrased as 'I cook and they eat the food.' In this example, the agent and verb are clearly coordinated. Thus, if we were to posit the existence of a VP in Tengger based on this type of construction, we would have to say that the VP contains the verb and agent argument, to the exclusion of the patient argument. This is highly unlikely.

Rather, it seems that coordination also turns out to be a less than reliable diagnostic for VP constituency in Tengger.

So where does this leave us? Movement tests in a language that allows seemingly free word order are unproductive. Similarly, replacement, substitution and ellipsis tests in a language that strongly prefers null constituents are equally unproductive. Above, I presented data from intonation patterns and coordination. However, these diagnostics alternately identified the subject and verb or the verb and object as a constituent. Neither of these tests pointed unambiguously to the existence of a VP in Tengger. It seems most parsimonious to me to not posit the existence of something without sufficient empirical evidence. VP constituency is a standard test for configurationality, so it is a good reason to ask if Tengger does as well, however, it is not a valid reason for positing its existence in Tengger without other unambiguous evidence. I can find no such evidence at this time.

4.7 Restrictions on Word Order

4.7.1 Modals In Tengger

In the grammatical overview above I discussed word order and word order restrictions within NP in Tengger. Here I discuss certain restrictions on word order in clauses. Tengger has a series of modal verbs which in general can combine with other verbs and in some cases can stand on their own.

51: Tengger Modals [the standard equivalents are given in brackets]

a.	wis	‘already, perfective’	[wis]
b.	gèk	‘progressive’	[lagi]
c.	gelem	‘want’	[gelem]
d.	katé	‘will, want’	[arep]
e.	(d/g)(u)rung	‘not yet’	[durung]
f.	kathik	‘progressive’	—
g.	isih	‘still’	[isih]
h.	(b)isa	‘can’	[isa]
i.	kudu	‘must’	[kudu]
j.	mesthi	‘must’	[mesthi]

This list is near exhaustive for Tengger. There are a few other modals that are very rarely heard, and may well be due to interference from the lowland dialects. In general, the modals appear adjacent to the verb or word they are modifying:

52. a. Balé desa urung di-jembar-né…
hall village not.yet di-broad-é
’The village meeting hasn’t been expanded yet.’

b. Kathik nduwé rencana kudu teka Malang jam pira?
PROG n-have plan must come Malang time how.much
‘Do you have a plan for what time you should arrive in Malang?’

In the above examples, the modal appears immediately preceding the verbal element. In the second example there are two modals and two verbs, in each case the modal immediately precedes the verbal element. The modal can also follow the verbal element quite naturally:

53. a. Iku budhal gèk sore mau.
that leave PROG afternoon earlier
'He was just leaving earlier this afternoon.'

b. Lho gelem-é lunga isih.
prt want-é to.go still
'Wow, I still want to go!'

In both of the examples above, the modal appears following the verbal element. Note in these cases there is still adjacency.

Many of the modal elements can also combine with non-verbal elements, and again in general they appear adjacent to the word they are modifying:

54. a. Nèk wis cukup ngono, kuwi mulih dhisik.
if pft enough like.that that go.home. first
'If we're all finished here, then I'll just head home now.'

b. Bawang tropong gèk rong éwu, rong éwu munggah.
garlic binoculars PROG two thousand two thousand go.up
'K.o. large scallion are 2,000 (rupiah/kilo), 2,000 and up!'

In the first example above, the perfective modal *wis* combines with an adjective, *cukup* resulting in a meaning 'already enough.' In the second example, there is no verbal element other than the progressive modal. It is possible that this is simply a null copular construction, or just that the verb itself is phonologically null. The progressive marker is indicating simply a present and continuing truth, for the time being at least, the current cost of *tropong* scallions.

Multiple modals can also appear in a sequence. In many related languages, there are restrictions on order in which certain modals can appear or which can appear in series. Most of the restrictions are semantic, so it is unlikely to find perfective *wis* combined with *urung* 'not yet', or *isih* 'still.' These are only semantic restrictions, and not syntactic restrictions, and still some surprising combinations are possible (55c).

55. a. Basa Inggeris gèk urung di-terjemah-en.
language English PROG not.yet di-translate-na
'There's still not anyone to translate the English.'

b. Dadi isih urung cetha.
so still not.yet clear
'So it's still not yet clear.'

c. Masyarakat wis gèk m-butuh-na dhukun manèh.
society pft PROG N-need-na priest again
'The people are already needing a new priest again.'

In each of the above examples there are two modals preceding the element they modify. The order of the modals is a matter of scope, with the outer modal taking scope over the subsequent modal.

In each of the preceding examples, the modal appears immediately adjacent to the word it is modifying, though it can either precede or follow that element. However, there is no strict adjacency between modal and verb (or other modified element):

56. a. Anu gèk kathik iku lho nerus-na iku...
um PROG PROG iku prt N-continue-na that
'Um, still doing that continuing with that thing.'

b. Nèk isih urung sésuk-ésuk teka Pak Tris…
if still not.yet tomorrow-morning come Pak Tris
'If Pak Tris still hasn't arrived by tomorrow morning…'

Interestingly both of these examples contain multiple modals. In both cases, neither of the modals is adjacent to the verb, though they are in each case adjacent to each other. These examples, demonstrate that there is no adjacency restriction which holds between modals and verbs.

Several of the modals also often occur in final position. However, like a number of other particles, when in final position they frequently appear with the suffix *–an*. Most commonly this is true of *wis, isih,* and *urung*, as well as the interrogative particle *kok*, the negatives *ora* and *dudu*, the affirmative particle *ya*, and the agreement particle *rak*.

57. a. Nèng Amrik ya ana wong cilik?

at America aff there.is person small

'Are there peasants in America?'

b. Masa gèk nyangking koran ya-an.

no.way PROG N-carry.in.hand newspaper aff-an

'No way was he carrying a newspaper in his hand right!'

c. Nèk sing kéné ki Inggris-é isih urung ana sing bisa.

if REL here this English-é still not.yet there.is REL can

'Here there still aren't any (people) who can speak English yet.'

d. Dhèk Wanakitri during isih-an.

at Wanakatri not.yet still-an

'(I) still have not yet ever been to Wanakirti.'

Note that in examples (57a) and (57c) the particle and modal appear internally, and have no suffix. However, when they appear sentence finally, they almost always have the suffix *–an*, as in (57b) and (57d) above. The meaning appears unchanged by the addition of the suffix. Indeed, it may well function as a place marker. It could also be that the suffix licenses the modal in a position in which it is otherwise ungrammatical, except that we do have plenty of examples of these particles and modals in final position without the *–an* suffix. I leave this particular question open here.

Not all modals can appear word finally, nor can they all appear as one word utterances. In fact, these properties delineate two distinct groups of modals. *Kudu, gèk,*

kathik, and *katé* cannot appear word finally, nor can they appear as one word utterances.[59] The most restricted distribution of the modals is also one of the most frequent, *katé*, I discuss this in more detail in section (4) on categories. Each of the other modals can appear as one word utterances, contextualized of course.

58. a. Isih!
still
'I still am (sleeping).'

b. Wis.
already
'I'm done (eating).'

For the class of modals which cannot appear as a one word utterance, the only restriction on their distribution is that they not appear alone or sentence finally. Otherwise, there is no adjacency restriction and they can modify any other class of words:

59. a. Nèk wis dadi pe-tinggi desa Ngadas, jiwa-né kudu jiwa nasional.
if pft become pe-high village Ngadas soul-é must soul national
'If you become the village chief of Ngadas, you must have a nationalistic spirit.'

[59] *Kudu* 'must' can appear as a one word utterance if it suffixed with *–né*, meaning 'it should be, it should have been'. Without the suffix, however, it is ungrammatical.

b. Kudu ya sa'botol dadi cukup.

must aff one-bottle become enough

'One bottle must be enough.'

c. *Ya sa'botol dadi cukup kudu.

d. Ya sa'botol dadi cukup kudu-né.

In the first example above, the modal modifies a nominal, *jiwa* 'soul.' In the second example, the modal appears in initial position followed by an affirmative particle. Here the modal is modifying the verb *dadi* 'become', but there is clearly no adjacency restriction with the verbal element. The contrast between the last two examples shows clearly that this class of modals cannot appear in final position. Specifically for *kudu*, it is rescued if it has the *–é* marker on it.

We can also mention three other words that are often classed together with modals here, and which share a number of properties with the other modals, the negative markers. Tengger has two distinct indicative negative markers, *ora/ga* and *dudu*. Historically, *ora* negated verbal or predicative constructions. This is also the form found in many other Javanese dialects. In Tengger, it is most often realized at /raʔ/, though /ora/ is also common. In the Malang and Surabaya dialects in the lowlands, *ora* has basically been replaced with some form of *ga*, which is a borrowing from Indonesian. This is most often realized as /gaʔ/, though occasionally /ŋgaʔ/ or even /ga/ are also heard. This form has been borrowed recently into Tengger. The two negatives are in complete free variation, there is no difference whatsoever between the two. The most common form in Tengger is still /raʔ/.

Functionally, in the standard language, *ora* is described as a verbal negative. It is parallel to Indonesian *tidak* and related forms. This contrast with another negative, *dudu*, which has an invariable form. For the standard language, this is described as a nominal negative, or in some cases as an emphatic negative. It is parallel to *bukan* in Indonesian and related dialects.

60. a. Ora isa merga-né gak n-duwé ijasah, rak n-duwé pe-paran.
NEG can because-é NEG N-have certificate NEG N-have RED-what
'I can't because I don't have a diploma, I don't have whatever is needed.'

b. Iku suguh-an, iku dudu tonton-an ha ha ha
that food.offered.to.guest-an that NEGN watch-an ha ha ha
'That's food to be eaten not a spectacle (just to be looked at!)'
[encouraging guests to eat and not just look at the food]

The first example above makes use of three negatives, all negating verbal elements. Note three distinct forms, /ora/, /gaʔ/ and /raʔ/ are used in the same sentence, with exactly the same meaning and function. In the second example we find a prototypical usage of *dudu*. Note that there are no expressed verbal elements at all in the sentence, and the negative is used to negate one of a pair of nominals.

Most frequently, when used utterance finally, *ora* indicates a question[60], though this is not always the case. One of the negative forms, /raʔ/ is homophonous with another particle which is used to invite agreement (similarly to *kan* in Indonesian). When this particle appears in final position it almost always has *–an* suffixed to it, otherwise it appears unrestricted and without any further marking. The two can usually be differentiated by context and through intonation.[61]

61. a. Nèk sega demen ora?
if rice like NEG
'Do you like rice or not?'[62]

b Truk sayur-an wis teka rung?
truck vegetables-an pft arrive not.yet
'Has the vegetable truck already come or not?'

c. Ana sing percaya ana sing gak.
there.is REL believe there.is REL NEG
'There are those who believe and those who do not.'

The first two examples above are very clearly questions. Here the negative comes in final position, must as it does in the equivalent English construction. However, as shown in (61c), the negative can also appear in final position when not an interrogative as well.

[60] This is true also of *urung* 'not.yet' discussed above.
[61] This particle can not co-occur with any other interrogative particles, so a sequence of *kok rak ...* is clearly the negative and not the particle, meaning 'why don't…?'
[62] People were of course very concerned that I should like rice!

While *dudu* can appear as a one word utterance, otherwise it cannot appear finally:

62. a. Sing iki? Dudu![63]

REL this NEGN

'(do you want) this one?' 'No (not that one).'

b. *Sing isun gelem iku dudu.

REL 1.f want that NEGN

'The one I want is not that one.'

c. Sing isun gelem dudu iku.

The first example clearly shows that *dudu* can stand as a one word utterance. However, as we can see in (62b), it cannot otherwise appear word finally. When the negative appears preceding the element it is negating, however, as in (62c) the sentence is grammatical.

The difference between (62b) and (62c) suggests that there is a word order restriction here, and indeed there is. The nominal negative *dudu* must immediately precede the element that it modifies. Consider the following:

[63] There are many examples of *dudu* appearing as the final element in two word utterances:

a.	Isun dudu.	b.	iku dudu.
	1.f NEGN		*that NEGN*
	'Not me.'		'Not that one.'

These would seem to contradict my claim that *dudu* cannot appear word finally. However, in examples such as these, there is a very significant pause between the two words, and they are better considered at two distinct utterances. 'Me? No.' is a better translation.

63. a. Dudu di-ternak-na, teka alas, teka liar.

NEGN di-livestock-na come forest come wild

'It's not raised, it comes from the forest, from the wild.'

[talking about pet snakes]

b. Nèk ula dudu halal?

if snakes NEGN permissible

'But snakes aren't halal right?'

[talking about eating snakes]

What is going on here? The nominal negative is modifying a verb in the first example and an adjective in the second example above. The first example demonstrates another use of *dudu*, which is as an emphatic negative. Here the speaker was stressing the fact that the snakes were not raised, but caught wild. In the second example, however, the negative modifies an adjective. There is no emphatic element here, it is a simple interrogative question. This example would be ungrammatical in the standard language, however, in Tengger it is perfectly acceptable. It seems as if Tengger is undergoing a loosing on the distributional restrictions on *dudu*. Perhaps this is due to the emphatic use, where *dudu* can modify non-nominal elements, which has been reanalyzed as a more general negative. On the other hand, *ora* (in all it's variations), is restricted to any non-nominal, verbs, adjectives, prepositions. I have no examples of *ora* negating a nominal in my entire corpus. So that pattern is robust.[64]

For the sake of completeness, I mention here also the negative imperative marker, *aja* 'don't'. The negative imperative marker can appear as a one word utterance, though

[64] Many colloquial dialects of Indonesian also allow a wider distribution of *bukan*. So the pattern in Tengger may be due to influence from Indonesian, or the two may be affecting each other.

when combined it almost always immediately precedes what it is prohibiting, which is not restricted and can be any nominal or predicative element. In a sense it can combine with other modals, however, it always precedes them as well:

64. a. Aja dhek kéné/Tomo/larang!
NEGIMP at here/Tomo/expensive
'Don't (do it) here!/ (pick) Tomo / (make it) expensive!'

b. Tapi aja wis pak, mbésuk nang Amrik.
but NEGIMP pft sir tomorrow to America
'Don't be done/go so soon, tomorrow you're leaving for America.'

In example (64a) above the negative imperative marker is followed by a preposition, proper name, and finally an adjective. The modal appears preceding the thing prohibited. In the second example the negative imperative appears modifying another modal, in this case the perfect marker *wis*, with no expressed lexical verb. The sense here is 'don't let it be over/done'. Note in this case that the negative imperative must precede any other modal.

There are a few limited cases where another element may precede the negative imperative:

65. a. Tomo, aja.
Tomo NEGIMP
'Don't Tomo.'
'Don't (pick/use/invite) Tomo.'

Under the first interpretation, Tomo is simply a vocative, and the prohibition is general. Under the second interpretation the speaker mentions the thing prohibited, then pauses, then adds the negative imperative. There is clearly a large pause, and the two are most likely separate constructions.

Some of the basic properties of Tengger modals are given in the table below:

66. Tengger Modals

Modal	One word utterance	appears finally	appears finally with *–an*	Open complement type	combines with other modals
wis ‘perfective’	+	+	+	+	+
(b)isa ‘can’	+	+	-	+	+
mesthi ‘must’	+	+	-	+	+
gelem ‘want’	+	+	-	+	+
isih ‘still’	+	+	+	+	+
kudu ‘must’	-	-	-	-/+	+
katé ‘will, want’	-	-	-	-	+
gèk ‘PROG’	-	-	-	-/+	+
kathik ‘PROG’	-	-	-	+	+
(g)(u)rung ‘not.yet’	+	+	-	+	+
ora ‘NEGV’	+	+	+	-	+
dudu ‘NEGN’	+	+	-	+	-
aja	+	+	-	+	+

It is very clear from the above table that there are two large classes of modals in Tengger. While almost all modals are able to combine with other modals, many have restrictions on what can serve as a complement. I discussed the difference between *katé* and *gelem* above in the section on categories and between *dudu* and *ora* in this section. One class of modals contains *wis, mesti, bisa, isih,* and *gelem*. This class is characterized by being able to occur as a one word utterance, appearing utterance finally, and having no restrictions on which words can serve as complements. The second class consists of *katé, kathik, kudu,* and *gèk*. Items in this class cannot appear as a one word utterance, cannot appear in final position, and are restricted in what can serve as a complement. Note that for *kudu* and *gèk* I have placed both a plus and a minus sign. Consider the following:

67. a. Gèk mangan paran?
PROG eat what
'What are you eating?'

b. Gèk sega.
PROG rice
'(I)'m (eat-)ing rice.'

Although not common, it is perfectly grammatical for the progressive marker *gèk* to combine with a noun. The verb in this example is null, but it is clearly there semantically. *Kudu* behaves in a similar way. This contrasts to the use of *katé* discussed above, which uniformly cannot take nominal complements.

4.8 Summary

The discussion of modals in the previous section makes it clear that there are some restrictions on word order in Tengger. In fact, we have seen a number of cases of restrictions on possible word orders in Tengger. Most of them, however, are lexical. What is remarkable is the extremely wide variety of orders which are possible.

4.9 Diachrony of Tengger Word Order: Word Order in Old and Middle Javanese

Proto-Austronesian had VSO word order, and many languages of the western Austronesian group, including the Philippine and Formosan languages maintain VSO order. The languages of western Indonesia, however, have largely developed SVO order.[65] All of the languages which are closely related to Javanese including Balinese, Sasak, Sundanese, Madurese, the Malayic languages, the Batak and Dayak languages are generally, though not uniformly as I have noted, reported to share SVO order. Given this it could be assumed that the change from VSO to SVO occurred early on in the development of a proto-language immediately above all these, probably Proto-Western-Malayo-Polynesian and is not the result of unrelated isolated innovation, convergence, areal spread, or contact with non-Austronesian languages. This is not the case however. Many languages of eastern Indonesia, as well as the other subgroups below Western-

[65] Of course not all members of this set are SVO. Malagasy, which is closely related to Malay is VOS. Most languages of Eastern Indonesia, Oceania, and Micronesia maintain VSO order.

Malayo-Polynesian maintain VSO word order. More crucially, as Old Javanese is reported to have VSO order (Hunter 1999, Zoetmulder 1974, Ogloblin 2005).[66]

68. a. Ahyun ta sira patya dang hyang Drona. [Old Javanese]
want prt 3sg death of HON Drona
'He wants for the death of Drona.'
b. Wruh ta-aku bhakti-nta.
know prt-1sg devotion-2sg
'I know (of) your devotion.'

These examples clearly show VSO word order, and note that the verb in each case also appears without any further morphology. Like modern Javanese, Old Javanese had complex verbal morphology, marking voice, valency, semantic relations of arguments, mood and other grammatical information. However, as with Tengger, there is also a class of transitive verbs which can appear without any morphology. It is not clear how transitive (68b) is here.

Like the Philippine languages, Old Javanese also had verbal morphology which marked the semantic role of the participants. I discuss this in the next chapter on voice and focus. Importantly here, however, is that word order in Old Javanese also was not fixed:

[66] Unfortunately there has not been enough scholarly attention given to Middle Javanese, and so it is unclear to me whether Middle Javanese was fully VSO or in the initial stages of developing SVO order.

69. a. Sira ta k-um-on i nghulun tamolah-a ngke pancatirtha.

3sg prt orde-ACTr to 1st stay at pancatirtha

'He ordered us to stay at Pancatirtha.'

b. Aku k-um-on-aken i kite won ri wwang ngkana.

1sg order-ACT-CAUS to 2nd guarded by people there

'I already ordered you to be protected by those people there.'

Both of these examples have active voice. The agent precedes the verb, which precedes the patient. Here, the Old Javanese examples display SVO order.

Poedjosoedarmo (2002) traces the history of word order from Old Javanese to Modern (standard) Javanese. She claims that the dominant word order in OJ was VS(X), almost without exception. I have given rather run of the mill examples above however, with SVO order. She claims that OJ sentences '…having SV order, the verb phrase following the noun constitutes a dependent clause in all cases. (2002:325)'. It is hard to reconcile such an analysis with the examples given above.[67] Each of the two examples above has a dependent and an independent clause. In (69a) the subject, *sira* precedes the active verb *k-um-on*, this is clearly the matrix verb. It takes the entire embedded clause, *i nghulun tamolah-a ngke pancatirtha*, as its complement, or object. The order within the dependent clause is also SVO. Under an analysis where a verb following a subject NP represents a separate clause, the example in (69a) above would have to contain four different clauses. This is clearly not the case.

[67] Of course, given that the examples are taken from Old Javanese, it is hard to determine if they are pragmatically marked in some way.

CHAPTER FIVE: VOICE AND THE CLAUSE IN TENGGER

5.1 Introduction

Perhaps more than any other feature the voice or focus system is the most salient, characteristic, and contentious issue in Austronesian linguistics. It has been the target of an enormous amount of recent work, ever since Schacter (1976) questioned the applicability of the notion 'subject' in Philippine languages. Kroeger (1993), Foley (1998), Arka and Manning (1998), Wouk and Ross (2002), Arka and Ross (2005), Himmelman and Adelaar (2005) *inter alia* are just some of the recent works and entire volumes that have been published dedicated solely to issues of voice and focus in Austronesian languages.

The characterization of the basic voice opposition within Indonesian languages has also been the subject of much recent work. Standard Indonesian, a language closely related to Javanese, has been analyzed as having an active and two distinct passive constructions (Chung (1976), Hamid (1992), Alsagoff (1992), Guilfoyle et. al. (1992), Sneddon (1996), *inter alia.*) Problems with this analysis have been pointed out recently by Saddy (1991, 1992), Cole and Hermon (1998), Gil (2000), and Tadmor (1996). Work on Indonesian and Indonesian dialects is far ahead of work on Javanese, and there is little consensus there. Here, I wish mostly to present the data from Tengger and show how it may bear on the issue.

In this chapter, I present the data from the voice system in Tengger. I place it within its historical and geographical perspective. Further, I place it typologically in

relation to recent work on symmetric versus non-symmetric voice systems. Perhaps more than in other chapters, I will look at some more abstract theoretical issues here, including some recent work in the generative tradition and issues of ergativity, to show how the Tengger data bear on such questions.

Primarily though, the chapter will be descriptive, presenting data on a range of different voice constructions, word order variation, and aspects of the behavior of ditransitives. I will also explore questions of subjecthood, core vs. oblique arguments and discuss crucially how they behave under different syntactic tests such as reflexive binding and quantifier float.

5.2 Voice Markers in Tengger

There are two large classes of verbs in Tengger: those that take voice morphology and those which appear in bare form. those that appear in a neutral clause without any overt morphological markers are exemplified by *lunga* 'go' and *tuku* 'buy'. I discussed these at length in the previous chapter. Compared to standard Indonesian and many Malay varieties, this class of verbs is particularly large in Tengger, and other Javanese dialects. It is further remarkable for containing a large number of transitive verbs.

Most verbs in Tengger, nonetheless, generally appear with some overt voice marker in an otherwise neutral clause. There is a basic opposition between active and passive voice. A note on the terminology used here is necessary. There is a tremendous variety of different terminology used to describe the same phenomena in western Austronesian languages, active/passive voice, actor/patient focus, topic/comment, etc.

These all refer to the morphosyntactic systems employed to identify the semantic relation between the verb and its arguments. In the literature, there have emerged two basic systems, with any number of variations thereon: the Philippine-like systems and the Indonesian-like systems. Later on I will discuss the Philippine-like focus system in greater detail, as Old Javanese followed that pattern. I will further place the Tengger within a typology of different voice systems. For the time being, I adopt the term voice, and the associated distinctions of active and passive for the description of Tengger. However, this is not to imply that the Tengger voice system is completely analogous to other more familiar voice systems. Ultimately, I argue for the voice markers in Tengger as *weak generalized voice markers*, with the active markers associated loosely with an agent/actor/experiencer argument and the passive markers associated with a patient/theme argument, on a semantic level, and crucially not on a syntactic level of expression.

5.2.1 Active Voice in Tengger

There are two different markers of active voice in Tengger: a homo-organic nasal prefix that can be applied to almost any root, and a bilabial nasal that appears on a very restricted class of words. I discuss each in turn.[68]

[68] Kats (1939:243) (quoted in Smith-Hefner 1983) reports another active prefix in Tengger: *ang-*:

Kak isuk bud, ang-katé nyang ngendi rika?

PRT morning ?? ANG-want to where 2.formal

'It's so early, where are you going?'

5.2.1.1 The nasal prefix N-

To make a simple active sentence, a nasal prefix, marked *N*-, is added to almost any root. The nasal assimilates in place of articulation to the initial consonant of the root. If the root begins with a voiceless obstruent, it is lost after affecting assimilation. Roots beginning with vowels and liquids take a velar nasal. Roots beginning with nasals are not (distinguishably) prefixed. The pattern is the same as most other dialects of Javanese. It differs from standard Indonesian, though is similar to many colloquial varieties of Indonesian, especially Jakartan Indonesian. The pattern is given below:

This prefix comes from Old Javanese *mang*-, neither of which were attested when I conducted my fieldwork.
Smith-Hefner also reports rare occurrences of the prefix *mer*-, I have found no examples of this, and none of my consultants knew the form.

1. Nasal Prefixes in Tengger

Root Initial	Prefix	Is Replaced With	Example
p		m	pikir > mikir ‘think’
b	m		balik > mbalik ‘return’
t		n	takon > nakon ‘question’
d	n		duwé > nduwé ‘have, own’
th		n	thuthuk > nuthuk ‘hit, pound’
dh	n		dhudhuk > ndhudhuk ‘sit’
s		ny /ɲ/	semprot > nyemprot ‘spray’
c		ny /ɲ/	colong > nyolong ‘steal’
j	n		jajal > njajal ‘try
k		ng /ŋ/	kopi > ngopi ‘drink coffee’
g	ng /ŋ/		gawa > nggawa ‘carry’
w		m	waca > maca ‘read’
l	ng /ŋ/		langi > nglangi ‘swim’
r	ng /ŋ/		ricik > ngricik(i) ‘clean’
N	-	-	nyanyi > nyanyi ‘sing’
V	ng /ŋ/		isi > ngisi ‘fill’

There are some other minor peculiarities to the system. For example, roots with initial /w/ are sometimes idiosyncratically replaced with /ng/ instead of /m/. In keeping with the phonotactic preference for disyllabic words, monosyllabic roots most often take *nge-*, for example *dol* > *ngedol* 'sell.'

The nasal prefix can be added to form intransitive, transitive, and ditransitive words:

2. a. Soal-é bumi **muter**, molak-malik. < N-puter
matter-É earth N-turn back.and.forth
'The thing is, the earth rotates, back and forth.'

b. Mbiyèn, éyang ya **ng-golèk** anu iku…**ng-golèk** kobra…
past 1.m AFF N-look.for anu that N-look.for cobra
'In the past, I searched for, um, what's it… I looked for cobras (to sell).'

c. Nèk-é mas Tom **n-jaluk** awak-é dhéwé basa kéné, sanggup-a
if-É mas Tom N-ask.for body-É self language here capable-INTRG
Mas Tom **m-bel-ajar**-i laré kéné?
Mas Tom N-ber-teach-I child here
'But if you (mas Tom) ask us for the local language, will you be able to teach (your language) to the children here?'

In (2a) above the unaccusative verb *muter* appears with prenasalization. The voiceless obstruent initial of the root /p/ has been lost. The actor here precedes the verb. In (2b), there is a transitive verb, *ng-golek* 'to look for.' The agent *éyang* precedes the verb and the patient *kobra* follows the verb. The last example shows an inherent ditransitive verb, *jaluk* 'to ask for.' In this case the verb is marked with a nasal prefix and the agent precedes the verb. This is a double object construction, and so the goal/source immediately follows the verb, and the theme is in final position.[69]

---

[69] Example (2c) is an extremely interesting sentence, on a number of levels, and warrants a short footnote. To begin with, this utterance shows interference from three distinct languages/dialects. The primary language of the utterance is clearly Tengger as is clear from the phonology, but note the following items:

i.	awak-é dhéwé 'us'	lowland dialects of East Java
ii.	laré 'child'	Krama Inggil, very high speech level, Central Java
iii.	mbelajari 'teach'	Indonesian, with Tengger morphology

Tengger pronouns are not marked for number, as with most Javanese dialects. However, the eastern lowland dialects have innovated a complex form. The speaker has borrowed this form. As already noted, the most salient sociolinguistic feature of Tengger is the total lack of the elaborate speech level system found in the lowlands, and particularly in Central Java. The form *laré* here is extremely formal and respectful, and not part of the local dialect. The most fascinating form, though, is *mbelajari*. Javanese, including Tengger, has asymmetrical lexemes for teach/learn, as in English, *mulang* 'teach' and *sinau* 'learn/study.' However, Indonesian has symmetrical forms based on the root *ajar*. The 'teach' verb is formed with the active/passive prefixes *meN-/DI-*, so you get *meng-ajar* in the standard language, or simply *ng-ajar* in most colloquial varieties. The 'learn/study' verb is formed with the generally intransitive marking prefix *ber-,* with an idiosyncratic form *bel-* in this root, *bel-ajar* 'study.' The speaker here, however, has taken *belajar* as a monomorphemic root, and attached the Tengger active nasal prefix *N-* together with the applicative *–I*, making a causative form 'cause to learn' i.e. 'teach,' and perhaps also in order to transitivize an unergative base, and so we get *m-belajar-i*, with the meaning 'teach.'

As with the unmarked verbs presented in the previous chapter, other word orders are possible even with overt voice morphology.

3. a. bèn **muter** sik montor-é.
   let N-turn first motorcycle-É
   'Let the motorcycle turn around first.'

   b. akhir-é énak-é, cepet-é **ng-golek** dhèkné.
   end-É delicious-É fast-É N-look.for 3
   'In the end he just went after comfort and ease (speed).'

In (3a) above we have the same verb as in (2a) *muter*, with exactly the same meaning. Here, however, the verb appears before the actor, followed by temporal adverb adjunct, the actor is in final position. In (3b) we have a transitive verb appearing with the nasal prefix, as in (2b) above. Here, though, the agent is in final position, and the double theme arguments are preverbal.

Other word orders are possible, but these will suffice for the time being. What is important to note, is that the variation discussed in the last chapter for the class of verbs that regularly appears without any voice markers, is also noticed when the verb has the active nasal prefix.

Note also the form *nèk-é*, which shows the complementizer with the *–É* affix. It is rare cross-linguistically for complementizers to be morphologically complex, though there are languages with agreeing complementizers.

I have also noted elsewhere many western Indonesian languages have a strong preference for the passive voice. Tengger shows a similar preference. Within the corpus, there also appears to be a strong preference for bare verb stems, as discussed in the last chapter. This, however, is due to the fact that the class of unmarked verbs in Tengger is rather large (especially compared to Indonesian varieties), contains a number of transitive verbs, and generally members of this class are high frequency words.

The nasal prefix can be attached to a root of almost any class to create a verb:

4. a. Sapa sing **nyambel**? < N-sambel
   who REL N-chili.sauce
   'Who wants (to use) chili sauce?'

   b. **Minggir** sik ya!
   N-side first AFF
   'Move to the side ya!'

   c. **Ngéné**! < N-kéné
   N-here
   'Like this!'

In the first two examples above, nominal bases are prenasalized. The resulting meaning in such cases is varied, it can mean to use, eat, drink, make, or anything associated with the object. The default in most cases would be whatever action was most commonly associated with the noun. The second example demonstrates yet another use. When combined with locations or directions, the nasal prefix generally indicates movement. So

in the example above, when the nominal 'side' is appears with a nasal prefix, it means 'to move to the side.' Yet another use is demonstrated by (4b). Here a locative pronoun appears with the nasal prefix. This is a set form, there is no passive counterpart. It demonstrates a very idiosyncratic use, and in this case the resulting word means 'like this', 'in this way.'

The nasal prefix can be combined with other roots as well. In the last chapter I discussed prepositions, many of which belie their verbal origins with prenasalization or another verbal marker. Adjectives and adverbs can also take prenasalization, however they most commonly then also take an applicative ending. I discuss such examples in chapter 6.

5.2.1.2 The active prefix M-

There is another active prefix in Tengger, with an invariant form *M-*. Historically, this marker comes from the Old Javanese infix *–um-*, which indicated active voice in Old Javanese. The *–um-* infix, when attached to roots that began with a vowel, was reduced to */m/* and prefixed to the root. In Tengger, phonologically, the *M-* prefix is limited to roots beginning with vowels or sonorants. The class of words which appear with the *M-* prefix is extremely limited.

5. a. Kadhang nèk wong **m-lebu** ra katok.
   sometime if person M-enter NEG visible
   'Sometimes if someone comes in we don't see them.'

b. Lha sing gelem **m-laku** ming éyang thok ki….

PRT REL want M-step only 1.m only this

'The only one who wants to walk is me.'

c. Ya **metu** ndhèk gaga kana. < M - wetu

AFF M-go.out at field there

'Ya it lets out by the field there.'

d. Sapi-né wis **mati**. < M –pati

cow-É PFCT M-die

'The cow died.'

In each of the examples above, there is an intransitive verb with a single argument and this form is most often described as exclusively forming intransitive verbs. However, we saw above that *N-* can also be used to form intransitive verbs. Unlike *N-*, however verbs formed with *M-* have no corresponding passive construction, unless the verb undergoes further modification:

6. a. * di-lebu

DI-enter

b. * di-laku

DI-step

c. *di-wetu

DI-go.out

d. *di-pati

Di-die

In fact, roots which can take the *M-* prefix, can all also take the *N-* prefix. However, when they appear with the *N-* prefix, the root is almost always further modified with an applicative suffix, and the semantics of the verb are altered accordingly (as well at the phonology, for which see chapter 6).

7. a. ng-lebon-i 'enter s.t. into s.t.'

b. ng-lakon-i 'do, perform s.t.'

c. metok-na 'make come out'

d. maten-i 'kill'

I discuss the nature of applicatives in great detail in chapter 6. For the current discussion, it should suffice to note the difference between the verbs in (5) and (7). In (5), the verbs are all intransitive, whereas in (7) above, they are all transitive, with an additional locative (7a), transitive (7b), or causative (7c, d) semantic interpretation.

It might seem then, that *M-* marks this class of verbs as intransitive, and *N-* marks transitive. Remember we have already seen that on other verbs and especially on non-verbal roots, *N-* can mark intransitive constructions. So it is only the particular distribution of the two active markers on these particular roots that we are examining.

Consider the following:

8. a. Aku sih **m-lebu** wong mbiyèn…

1 prt M-enter person before

'I once brought someone in the past…'

b. Rekam-an-é sapa sing **m-laku**.

record-an-É who REL M-step

'Who make the recording?'

In both of the above examples, the verb is transitive, yet it is marked with the *M-* prefix, and has not other verbal morphology, and, crucially, no applicative marker. In (8a) the agent *aku* appears preverbally and the patient *wong* appears immediately following the verb. The examples in (8b) is a cleft, but clearly transitive.

So evidently, the *M-* marker is not restricted to indicating intransitive verbs, nor is the *N-* nasal prefix required when a verb is transitive. However, the *M-* prefix can never appear with an applicative marker:

9. a. *m-lebon-i

b. *m-lakon-i

For roots that begin with /w/ or /p/ there is no formal distinction between the invariant *M-* prefix, and the homo-organic nasal *–N*.

So, how are we to make sense of all this? In Tengger there is a lexically determined class of roots which take the *M-* active marker only when there is no further verbal modification. Although the verbs of this class are frequently intransitive, this is

not always the case. Therefore, it can only be claimed that *M-* marks active voice. This actually is unsurprising diachronically, as I noted above the prefix is derived from the Old Javanese active infix *–um-*. When suffixed, all roots take the nasal *N-* prefix. I discuss further the basic opposition between active and passive markers below. First, I turn to the passive markers in Tengger.

5.2.2 Passive Voice in Tengger

There are four distinct passive voice forms Tengger. Two of these forms are rare, and found mostly in fossilized examples. The constructions with *ka-* and the infix *–in-* appear in a very limited set of archaic forms, and they are no longer productive in Tengger, as in standard Javanese. An example of each is given below:

10. a. Aku kan **ka-wit** biyèn kaé ana ndhèk Ranu,
1.sg PRT KA-begin formerly that there.is at Ranu
'Right, I began (at first) like that, when I was at Ranu.'

b. Nèk pound **t-in-imbang** karo kilo, loro titik loro kilo, sa'pound.
if pound weigh –in- with kilogram two point two kilogram one pound
'If pounds are compared with kilos, 2.2 kilos are one pound.'

The first example above shows the *ka-* passive, and is notable for two features. First, this form is not frequent, if occurring at all, in other Javanese dialects, though it was

productive in Old Javanese, at least in the form *ka-wwit-an* 'began, was begun' (Zoetmulder 1982:234); as such it represents another lexical retention in Tengger. I have not found many other examples of the *ka-* passive in my Tengger corpus, and consultants do not feel that it is productive, so although this form stands as an example which is maintained in Tengger as opposed to other dialects, it is not evidence for the productivity of this construction in the modern language.[70] The second notable characteristic of this form is that it appears with an agentive first person personal pronoun *aku.* This is odd both because it is not the expected and usual Tengger pronoun *éyang*, showing clear signs of lexical interference, but also, and more importantly, because in this 'passive' construction, the agent should either be postverbal or absent all together. Given that this is the only example of this token in the corpus, it could simply be a speech error. However, and foreshadowing the discussion to come, there seems to be more at work here as Tengger seems to be undergoing a significant syntactic shift in relation to the valence of passive forms, with passive morphology appearing with active interpretation and typically active syntax. As the *ka-* passive is non-productive and of such limited scope in the modern language, I do not base any of the arguments made for this change on this evidence, but conversely, the analysis provided below for the other passive constructions will equally be applicable here.

The infix *–in-* is similarly now non-productive in Tengger. In Old Javanese, this was the most common way to mark passive voice, or patient focus voice, in contrast to the active *–um-* infix discussed above. Compared to Malay varieties, there are more occurrences of fossilized *–in-* in Javanese varieties, and to many speakers it is still

[70] The form is still found in the literary language of Central Java, and also sometimes in one of the higher speech levels.

analyzable. Note in the example above, the two expressed arguments carry the same thematic relation, theme. The agent is not expressed. The statement is equational, to a degree. One thing is being compared to another, but in no particular order, and with neither being given default status.

These two forms are archaic, literary, and no longer productive in the modern language. Passive voice in the modern language is generally marked with the prefix *di-*, which I turn to presently.

5.2.2.1 The passive markers DI-, TAK-

The primary method for passive formation in Tengger is equivalent to what Robson called the first passive (1992:87), formed with the proclitics *tak-, kok-,* and the passive prefix *di-*. This form is inflected for person, with *tak-* first person, *kok-* second person, and *di-* neutral, or third person. The passive is formed by adding either one of the proclitics or the prefix to the bare verb stem. There are many examples in the corpus of both the forms with *tak-* and *di-*, however I have found no tokens with *kok-*. There are several possible explanations for this. First, the second person passive is by far the least common in both standard Javanese and in Indonesian, so we might expect the same in Tengger. While some consultants are familiar with the form, it does seem to be very rarely used in Tengger. The clitic is a dependent, shortened form of the second person personal pronoun *kowé*, which is used in the standard language in central Java in the *ngoko* speech style. However, second person pronouns are highly variant across Javanese dialects, with *koen* being used in Malang, *samang* in Lumajang, and *sira* in Tengger.

Hence, the form *kok-* is unfamiliar to Tengger speakers not acquainted with the central Javanese dialects.[71]

Another reason this form is so infrequent is due to a generalized taboo avoidance of second, and to a lesser extent, first person forms, throughout Javanese and a large number of other Austronesian languages, as discussed above in Chapter 3. Several strategies are employed, including substituting proper names and demonstrative pronouns for second person, and again to a lesser extent, first person pronouns. It is also possible to utilize third person forms with second person referents or use impersonal demonstratives with second person referents. In fact, there are a number of examples, discussed below, where it is clear that a second or first person referent is implied, though the form that is used is the third person *di-* form. In Tengger, it seems that the second strategy, that of substituting demonstrative pronouns is the preferred one, as discussed above.

The use of the first person passive is also relatively rare in Tengger. This may point to a historical explanation for the lack of the just the second person passive. Cross-linguistically it would be rather odd for a language to have first and third person passive markers, with no corresponding second person. Smith-Hefner reports that, while generally absent, some uses of *tak-* were acceptable to Tengger speakers when she conducted her field work some 30 years ago (1983:155). A possible historical explanation could be that the Tengger dialect split from the other Javanese before the innovation of passive forms inflected for person, and so never developed the first and second person forms. Alternatively, it could have split after they were developed and then lost the forms, reducing the passive to a single *di-* form. The occurrence of the first

[71] The form of the passive proclitic itself is variable across dialects. In the north coast central Javanese dialects, for example, the form is *mbok-*.

person forms in the current day Tengger could then represent a borrowing or recent interference from the lowland dialects. This would not be wholly surprising. The propositive construction, which I discuss in the following chapter, makes use a homophonous proclitic, though it has very distinct properties. Smith-Hefner also reports that it was not present in Tengger speech. However, there are many clear examples of it my recordings, and it demonstrates clear recent interference from the lowlands. It is then possible that both of these homophonous forms are becoming more common in Tengger. Further research will be needed, however, to determine the exact historical development of the passive markers in Tengger.

Below, I give several examples of both the first person passive with the *tak-* proclitic, and the neutral passive marked with the *di-* prefix. In standard Javanese, Indonesian, and related languages, these two constructions are treated as distinct types of passive, a 'standard' passive marked with *di-* and a passive with *semu*, with the agent of the verb appearing preverbally. I will discuss these issues here, but introduce the forms in order, then discuss them together.

11. a. biyèn nèk mules-mules, weteng mules-mules **di-kèi** Norit iku.
past if RED.cramp stomach RED.cramp DI-give Norit that
'Before if one had stomach cramps, one would be given Norit [medicine].'

b. iku areng **di-gerus**, iku di-céthak dadi pil.
that charcoal DI-grind that DI-print become tablet
'That charcoal was ground, and made into tablets.'

c. [japa-né] **di-apal-en** mbek Aman teka buku-ku…

incantation DI-memorize-EN by Aman from book-1

'The incantation was memorized by Aman from my book(s).'

If we examine (11c), it shows the form of what might be a canonical passive in Javanese, Tengger, and many other languages across the world: the logical object, or patient, *japa-né*, appears in pre-verbal position and functions as the syntactic subject. The logical subject has appears here as an adjunct in the by-phrase *mbek* (this form being unique to Tengger and surrounding lowland dialects), and there is characteristic passive morphology, *di-,* on the verb, and in this particular example there is another adjunct PP, here carrying a source thematic role. The (a&b) examples in (11) also show the underlying patient appearing in preverbal position, and there is again the characteristic passive morphology on the verb. The agent has been demoted, and as is common cross-linguistically, or does not appear at all, as it is no longer a licensed argument of the now unaccusative verb. (11a) shows the passive in an embedded clause, (11b) in a coordinated construction, and (11c) shows main clause passivization. In each case the patient appears in preverbal position.

Unlike in many European languages, the agent of a passive *DI-* construction, can also appear immediately following the verb, and not in an adjunct by-phrase. In this construction, the agent is necessarily limited to the third person:

12. a. Gaga wis di-panja **réwang-é**.

field PFT DI-plant friend-É

'The fields have already been planted by my husband.'

b. *Gaga wis di-panja isun.

field PRT DI-plant 1.f

'I already planted the fields.'

c. Dadi di-bagi **wong** loro ngono.

so DI-divide person two like.that

'So the two of them split it up.'

[talking about the younger and older village shamans, who split the two areas of the village for holy days]

d. *Dadi di-bagi **sira** ngono.

so DI-divide 2 like.that

'So it was split by you.'

Examples (12b & d) show that the agent is restricted to third person in the *DI-* passive. In (12a & c) the agent appears immediately adjacent to the verb. When an agent is expressed, and not in an adjunct phrase, there is generally strict adjacency with the verb. However, this is also not always the case, see (13b) below where the agent precedes the verb. It should come as no surprise at this point that other word orders are possible:

13. a. Trus di-obong menyan-é nèng omah-é dhéwé-dhéwé.

then DI-burn incense-É at house-É RED.self

'Then incense is burned in each house (by themselves).'

b. Tas mari rapat kaé, Pak Tom di-critak-na karo Pak Danu-né.

just after meeting that Pak Tom DI-story-NA with Pak Danu-É

'Just after the meeting Pak Tom told (the story to) Pak Danu.'

[It was clear that I was doing the telling.]

c. Biasa-né ya di-tanggap jaran kepang iku.

usual-É AFF DI-hire.to.perform horse woven that

'Normally *Jaran Kepang* dancers are hired to perform.'

In each of the above examples the verb appears with the passive marker *di-*. In (13a) the patient appears in immediate postverbal position, not in preverbal or subject position, as we saw in (11) above. In (13b) the order is agent-DI-verb-recipient. This particular exchange happened right after the major bi-annual village meeting, during which I was introduced by the village chief (fortunately having arrived only a week or so prior to the meeting). As I returned home, my host mother asked if had told Pak Danu why I was in the village, as he had suspiciously asked of my host mother earlier. My host father, who had been at the meeting with me, then responded with this sentence. It was very clear that I was the agent, the one doing the telling. The verb here means something like 'to tell a story', and so the theme argument is inherent in the verb. The order of constituents here is crucial, though, in demonstrating a very important fact, namely that word order does not seem to necessarily correlate with the verbal morphology that appears on a verb. In (13), the verb is marked with the passive marker *DI-*, and yet the agent appears preverbally and the recipient argument appears postverbally.

Similarly, in (13c), it is clear that the *jaran kepang* or 'horse dancers', is not the agent, yet it appears immediately following the verb. I will resume the discussion of this construction shortly, but now I introduce the first person passive.

The first person passive behaves quite differently from the *DI-* passives in a number of respects:[72]

14. a. Enggak, ng-gawé teka gembrèng tak-thuthuk – i, dadi rak kaya jepit.
NEG N – make from oil.can 1-hit-I so PRT like pincers
'No, it's made from an oil can which I shaped, so that it's like pincers.'

b. Banyu tawar tak-semprot-en, anu-né lugur kabèh iku wis.
water fresh 1-spray-EN PRT-É fall all that AUX
'I sprayed it with fresh water, all of the whatsits fell off.' [pests]

c. Pak dhukun malah tak-ancam.
Pak. Priest even 1-threaten
'I even warned the priest/shaman.' [on being late to a ceremony][73]

[72] Note that I have provided active translations for each of the examples in (14). The passive in Tengger often has the 'force' of an active, and consultants will often give active translations into languages like English. However, passive translations are also available. Tengger, Javanese, and many Indonesian languages predominantly make use of passive constructions, as opposed to languages like English, which stylistically prefer active.

[73] As mentioned above, Smith-Hefner reports that this construction is rare or almost non-existent. It is important then to distinguish this from the propositive, which makes use of a homophonous proclitic *tak-*. The propositive, or desiderative, has a near future, intentional interpretation, 'let me…' or 'I intend to…' Each of the examples here, however, clearly has a past tense, perfective reading, which is incompatible

Unlike the passive in *di-*, the agent in (14) appears as a proclitic attached directly to the verb. The word order in all three examples above then is patient-agent-verb, with the agent necessarily expressed, again, unlike the *di-* passive. It is not clear that the clitic pronouns associated with this construction can actually be considered characteristic passive morphology, as they are clearly not part of the inflectional paradigm, though they are both phonologically and syntactically dependent. The syntactic dependence is shown by the strict adjacency which obtains between the verb and the clitic:

15. a. Gaga iku wis tak-semprot.
    field that PFT 1-spray
    'I already sprayed that field.'
    'That field was already sprayed by me'
    b. *Gaga iku tak wis semprot.
    c. *Gaga iku wis tak nyemprot < N-semprot
16. a. Gaga iku wis di-semprot.
    field that PFT DI-semprot
    'That field was already sprayed.'
    b. *Gaga iku di wis semprot
    c. *Gaga iku wis di nyemprot

with the propositive. In addition, formally, the propositive is most often—though not always—followed by verbs which take the nasal marker *N-*, which does not appear on any of the verbs in (14). So these are all clearly instances of the first person passive.

The above examples show that there is strict adjacency between both *tak-* and *di-*, as the (b) examples are ungrammatical when a modal intervenes. Note also the (c) examples, that the constructions are ungrammatical if the verb takes the active *N-* marker. This data suggests that both passive markers are clitics, and not prefixes. This is perhaps surprising for the *di-* marker. The active nasal marker is clearly a prefix, as demonstrated by it's close phonological bond with the root, triggering assimilation and sometimes deletion. This is not the case with *di-*. We have already discussed the complex sandhi rules that govern affixation in Tengger, both word initially and word finally, and we will see below that the *ke-* passive is clearly a prefix as it undergoes these phonological changes. *Di-*, and *tak-* on the other hand, does not undergo any phonological modification ever, their form is invariant. Further, unlike the active nasal prefix, *di-* and *tak-* do not ever enter into the base form in reduplicated words:

17. a. Aja nyampur-nyampur < N-campur
NEGIMP RED.N-mix
'Don't mix it all together!'

b. Trus tak-éling-éling-é gambar-é Pak Aman.
then RED.1-remember-É picture-É Pak Aman
'Then I was thinking about (remembering) the picture of Pak Aman.'

c. Iki ora di-gawé-gawé anu ngono.
this NEG DI-make um like.that
'This isn't made like that.'

d. di-asuh

di-rear

e. *dèsuh. isèni < isi + an + i ‘fill s.t.’

The first example above shows that the nasal marker is clearly a prefix, when attached to a reduplicated root the prefix also undergoes reduplication. This is not the case for *tak-* and *di-* as can been in (17b & c), they do not undergo reduplication. Further, they undergo no sandhi. For *tak-* the final phoneme of the clitic is a glottal stop, which never undergoes any sandhi. For *di-* however, two vowels at a morpheme boundary do undergo coalescence, as can be seen in *isèni*. Under similar phonological conditions, a final /i/ followed by an initial /a/ at an internal morpheme boundary, the *di-* form does not undergo any phonological modification. So we can safely claim that both of the passive markers are clitics and not prefixes, unlike the active markers.

A significant different between the two passives is that for the *tak-* passive, the agent is encoded in the same morphological unit that marks the form as passive, it is first person, as opposed to the *di-* form which does not encode any features of the agent. The first person passive will never appear with any further specification of the agent, that is, no full agent pronouns can precede the verb or appear in an adjunct by-phrase (also unlike the propositive construction).

18. a. *Gaga iku wis éyang tak-semprot.

b. *Gaga iku wis tak-semprot **karo éyang**.

field that PFT 1 1-spray with 1

The first person passive is clearly ungrammatical with any other expressed agent.

Is this a *pasif semu* as described for Indonesian? A major difference between the Javanese passives and the Indonesian passives is the form of the agent. In all Javanese dialects, the agent is restricted to the proclitic pronouns *tak-* and *kok-* (or the dialectal variation thereof). In Indonesian and Indonesian dialects, there is a set of reduced proclitic pronouns, but full pronouns, and pronoun substitutes such as proper names with first or second person referent can also appear in the agent position immediately preceding the verb.

Beginning in (1976), Chung argued that similar constructions in standard Indonesian were not simply a topicalization construction, as might appear at first glace, but was rather a genuine passive. A number of generally subject related properties, such as control and binding, identified the theme argument as the grammatical subject and not just a topicalized object. This analysis has been taken up by a number of subsequent researchers, as noted above. In the *pasif semu*, or P2 as discussed in Cole et. al. (2006), the verb appears in bare stem form, the theme appears pre-verbally and the actor is expressed also pre-verbally, crucially unlike more familiar European passives where the actor is either not expressed or expressed in an adjunct by-phrase. In their (2006) paper, Cole et. al. claim that adult Jakartan Indonesian does have the P2 construction, as demonstrated through word order in relative clause constructions, where the agent shows strict adjacency to the verb.

Though there are a number of *tak-* passives in the Tengger corpus, none of them occur in relative clauses or in any complex clause. Further, speakers tend not readily to accept the *tak-* passive as a dialect feature, though again it does occur in spontaneous production. However, it is not the case that all of the uses of *tak-* in the corpus are simply instances of code-switching with the lowland dialect and Tengger:

19. Tak-donga-en kana dadi guru gedé.
1-pray-NA there become teacher big
'I prayed that she would become a senior teacher.'

In this example there are a number of unique dialect features, which do not occur in the surrounding lowland dialects. Phonologically, /kana/ does not show the vowel mutations of the lowlands. The morpheme *–en* is the specifically dialectal form of the applicative II (discussed in chapter 6). Hence it is not wholesale code-switching between Tengger and the lowland dialect. However, it is clearly interference, representing more intense recent contact with the lowlands. It is unclear to me to what degree this construction has been fully borrowed into Tengger, though it is clearly moving in that direction.

Another significant difference between the two passives is semantic. For the *di-* passives in (11), I have given translations with passive verbs in English. For all of the examples in (14) above, I have given translations with active verbs. It is generally clear from the pragmatic context that these sentences have an active reading, and almost all consultants agreed that they have an active force. Example (14a) is a particularly interesting case in this respect. In this multi-clausal sentence, the matrix clause, *ng-gawé*

teka gambreng 'it's made from an oil can' has both a null object and null subject. Further, it has typically active morphology, initial nasalization, on the verb *ng-gawé.* The agent is unexpressed, but was a farmer, the speaker of the utterance, as is the patient, a metal tong-like tool involved in fertilizing the crops. However, the first embedded clause is passive, again with the syntactic subject—also the topic of the whole sentence and unexpressed object of the matrix clause—remaining covert. The force of the active verb in the matrix clause is clearly morphologically active, though semantically a passive English translation is best. Alternatively, the translation could be '(I) made (it) from an oil can, which I shaped so that it became like pincers.' This translation makes it obvious that the embedded clause is in fact a headless relative clause. The semantic ambiguity of the first person passive in Tengger, that it has active force but traditionally passive morphology, is mirrored in both the standard and lowland language and is generally true in Indonesian and Indonesian dialects.

This semantic ambiguity is not limited in Tengger to the first person passive, however. There are abundant examples from the corpus that are marked morphologically passive, but semantically are clearly active. Further, the word order in such examples, is also not what we have seen above, in the examples below, the patient follows the verb. There seems to be a clash, between morphological and syntactic marking which is similar to some of the literature on morphological versus syntactic ergativity.

20. a. Nèk sa'lokasi di-tandur-i patang macem – limang macem …
if one location DI-plant-I four variety five variety
'If in one location four or five varieties are planted…'
'If in one location, (I) plant four or five varieties…'

b. Saiki iku répot-é nèng kéné, umpama-né di-gawé sa'tèngki,
now that busy-É at here example-É DI-make one-tank.for.liquid
'Now that's difficult here, suppose one tank [of fertilizer] is used…'
'Now that's difficult here, suppose (I) use one tank…'

c. engko di-obat sa'macem…sing mati sa'macem, sing liya-né sih urip.
later DI-medicine one variety REL dead one variety REL other-É still life
'Later one variety is fertilized, the one that dies is that variety, the others still live.'
'Later (I) fertilize one variety…'

These three examples all have clear morphological passive marking on the verb, *di-*. As with a canonical passive, the agent is suppressed or demoted, and has no overt realization except for the marking on the verb. However, as noted earlier, it is a common strategy to use this 'third person' passive marker *di-* even when the agent involved is first or second person. It is clear from the second translations I have provided for each of the tokens,

that the agent involved is indeed the first person speaker. That a first person referent is intended for the agent is particularly clear in (20a), however, the (b & c) examples could also be translated with 'one' as they are more abstract in their discussion.

So, although these examples have special morphology marking the verb as passive, and the agent is unexpressed, the NP/DP object has not raised to syntactic subject position, and remains in-situ. Traditional generative analyses of passives in many European and other languages hold that a passive verb, in addition to showing characteristic morphology and demoting the subject, loses its ability to assign/check accusative case to its complement, which must then raise to a higher A-position, such as T, in order to check the accusative case feature and be licensed.[74] Descriptively, the typical method of passivization in SVO languages has the underlying object and syntactic subject appearing in preverbal position[75] According to either view, then, the above examples from Tengger pose a problem. It will be instructive to look at several non-canonical passives in other languages.

It should be further noted that each of the above examples involve DP/NPs which are marked for number or have a numeral modifier. In some sense, then, these are 'heavy NPs', which may be 'shifting' to the right, similar to English NP shift. However, there is another class of passives which do not involve movement, the so called 'impersonal passives', found in languages like French and Italian:

[74] Though this is not always the case. Some languages allow constructions such as *Yesterday have been sunk many ships* where the underlying object does not raise from its base generated position.
In such ca

[75] This is a simplification. For some languages definite DPs/pronouns must be raised, but indefinites do not have to be, eg. English *There have been three people arrested.*

21. a. Es wurde bis spat in die Nacht getrunken. (German)

it was till late in the Night drunk

'Drinking went on till late in the night.' (Jaeggli 1986:22b)

b. In de zomer wordt er hier vaak gezwommen. (Dutch)

in the summer it is here frequently swum

'In the summer it is swum here frequently.' (Perlmutter 1978:68)

The German and Dutch impersonals, unlike other passives (e.g. all English passives), do not correspond to active transitive verbs, but intransitive unergative verbs (unergatives are argued to be transitive under some accounts). However, these are unlike the Tengger forms above, all of which have corresponding active verbs, *to plant, to use,* and *to fertilize*. Some of the romance languages, however, also have impersonal passives that have corresponding active sentences:

22. a. Le fué entregado un libro a Maria por Pedro. (Spanish)

to-her was handed a book to Maria by Pedro

'Maria was handed a book by Pedro.' (Jaeggli 1986:13)

b. Il a étè tué un homme. (French)

there has been killed a man

'A man has been killed.' (Belletti 1988:10)

c. É stato messo un libro sul tavolo. (Italian)

has been put a book on the table

'A book has been put on the table.' (Belletti 1988:18a)[76]

In the above impersonal passives in Romance, the agent is either not expressed or in an adjunct by-phrase. In each case, the object appears in post-verbal position, which can be taken to be in-situ; that is, in its base generated position. However, as Baltin (2001:229) notes, 'Spanish and Italian allow subjects to be postposed, and French allows stylistic inversion. Therefore, one might ask whether the objects are actually in-situ, or are in the postposed construction.' The same question applies to the Tengger examples as well. Tengger also frequently allows subjects to appear in post-verbal position, as we have seen before:

23. a. Gawok **éyang iki**!
astounded 1m this
'I was astounded!'

b. Se-taun neng kéné, pokok-é gelis apal wis **sampéyan**.
one-year at here point-the fast memorize already 2sg.[madya]
'One year here, the point is you have already learned quickly!'

c. Kok ora melu sembahyang **iku**?
INTER NEG follow ritual.prayers 3sg
'Why doesn't he come to ritual prayers?'

[76] The German, Dutch, French, Italian, and Spanish examples are all cited from Baltin (2001:229).

Under a rather traditional generative type analysis, one would claim that in each of these examples, the subject—in bold—remains in its usual position in SpecAgrS (specIP, specTP), while the entire predicate (PredP/VP) has raised to the focus position below CP.[77] These examples differ from the in-situ passives above in that the former all have an intonation contour which rises at the end of the fronted predicate and then falls sharply on the post-verbal subject. There is no such intonational break found in the latter. Further, in (23c) there is an overt object which precedes the subject, and I have found no such instances in the corpus with the in-situ passive construction.

So neither the impersonal passive analysis nor the postposed (i.e. Pred raising) subject analysis is sufficient to account for the Tengger data. The situation is further complicated when we consider examples like the following:

24. a. sa'tèngki di-gawé macem papat … kan srampangan.
one tank.for.liquid DI-use variety four PRT haphazard
'One tank uses four types [of chemicals for fertilizer], it's haphazard [the mixing].'

b. Terus sing sapi-né iku di-gawé peng-hasil-an sing sipat-é taun-an
continueREL cow-É that DI-use pe-result-an REL character-É year-an
'And the cow uses the yield/products which should last a year.'

c. Iku di-peken kembang-é.
that DI-pick flower-ART

[77] It is unclear from 23a whether just the head of the phrase or the entire phrase itself has raised to the higher projection, as both would be possible.

‘He picks flowers.’

‘Flowers are picked by him.’

Here again there is passive morphology on each of the verbs, but unlike the in-situ passives above in (20), each of the above examples has both an overt object and an expressed agent. The agent in these three sentences, however, is not post-verbal or in an adjunct by-phrase, as would be expected when the verb is marked with *di-*, but rather it appears in preverbal, subject position. Further, the force of the verb in each case is very clearly active, especially when clarified by the context. For example, in (24b) the speaker was discussing why he does not raise cows, due to the exorbitant expense involved. In (24c) it is obvious that the flowers are not doing the picking, but rather the preverbal demonstrative cum personal pronoun.

While it is true that for most of the examples in the corpus which have passive morphology and a post-verbal object/patient with an expressed agent, the agent is in pre-verbal position and it is marked as definite or specific, this is not always the case. Soemarmo (1970), among others, argues that indefinite, non-specific subjects are not possible in both Indonesian and Javanese. However, this is not the case. In fact, Davies (1999) shows that it is also not the case in Madurese. There is a clear preference for the subject to be marked definite or specific, however, this it is far from obligatory:

25. a. Kirik-é nyokot kucing. Kirik ny-(c)okot kucing.

dog-É N-bite cat

‘The dog bit a/the cat.’ ‘A dog bit a/the cat.’

b. Kucing-é di-cokot kirik. Kucing di-cokot kirik.

cat-É DI-bite dog

'The cat was bitten by a/the dog.' 'A cat was bitten by a/the dog.'

As in Indonesian and Madurese, definiteness or specificity can be marked through use of demonstrative pronouns, quantifiers, or the *–É* marker (Indonesian *–nya* and Madurese *–na*). The examples in (25) make it very clear that there is no definiteness or specificity requirement on subjects in Tengger.

In many of the examples of this construction, the object is also definite or specific, though not always, as seen in (24b) above, *peng-hasil-an,* and not *peng-hasil-an-é.*[78] In a standard *di-* passive in related languages such as Indonesian, there is strict adjacency between the verb and the post-verbal agent, unless it appears in an adjunct by-phrase. However, this is not the case in Tengger, compare the previous examples with:

26. a. Kucing-nya di-gigit anjing. [Standard Indonesian]

b. *Kucing-nya di-gigit kemarin anjing.

c. Kucing-nya di gigit kemarin oleh anjing.

dog-NYA DI-bite yesterday by dog

'The cat was bitten yesterday by the dog.

27. a. Kucing-é di-cokot kirik [Tengger]

b. Kucing-é di-cokot wingi kirik.

c. Kucing-é di-cokot wingi mbèk kirik.

[78] Peng-hasil-an itself is a borrowing from Indonesian.

cat-É DI-bite yesterday by dog

'The cat was bitten by a dog yesterday.'

Clearly in Tengger the same adjacency restriction does not hold, as all three examples in (27) are grammatical, including (27b) where a temporal adverb intervenes between the verb and the agent, and the agent is not in an adjunct by phrase. This is further evidence, then that we are not dealing with subject postposing.

What exactly are we dealing with here? The in-situ passives show a clear case feature violation, and the in-situ passives with expressed pre-verbal agents violate standard accounts of passives cross-linguistically in that the agent is not demoted. It is interesting to note that active readings are available for most passive constructions, it is not clear, however, if this is true for all *di-* constructions. What then is the function of the *di-* marker? Before moving on to an analysis of *di-,* I present the other productive passive construction in Tengger.

5.2.2.2 The Accidental Passive, Ke-(an)

There is another productive construction in Tengger, similar to that in closely related languages, called the *accidental* passive. A verbal root is prefixed with *ke-* if the root begins with a non-sonorant. If the root begins with a /r/, /l/, or /w/, then *k-* is prefixed (occasionally *ku-* appears before /w/). For vowel initial roots, again, just *k-* is generally prefixed, a schwa being lost according to the sandhi rules. That the marker is subject to sandhi rules indicates that it is a true prefix and not a clitic, unlike the other

passive markers. Forms often occur with simultaneous suffixation of *–an*, this indicates that the corresponding active form has the applicative I marker, and so forms in *ke-an* are also not uncommon.

28. a. Katé nyambut gawé ora bisa merga **kodanan** [k(e) - udan – an],
want N-take make NEG can because KE-rain-AN
'I wanted to work but I couldn't because I got rained on.'

b. Wah, tas-é mas Tom ke-gawa Marsam.
PRT bag-É Mas Tom KE-carry Marsam
'Oh no, your bag was inadvertently taken by Marsam.'

c. Ya mboh, éyang ki ke-blasuk!
AFF don't.know 1.m this KE-lost
'I don't know, I got lost!'

These constructions are variously referred to as accidental or adversative passives. They are similar to English 'get' passives, 'get robed', 'get pick pocketed' etc. However, the marker can be combined with any verbal root to indicate non-volitionality. There is often no agent or actor per se, as in (28a) above. When there is a specific agent or actor it is often left unexpressed, though in (28b) we see *Marsam* expressed in post-verbal position.

These constructions are generally considered passive because the patient appears pre-verbally, and often there is no expressed agent or actor. The agent/actor can also be expressed in an adjunct by-phrase:

29. Wah tas-é mas Tom ke-gawa karo Marsam.
PRT bag-É Mas Tom KE-carry with Marsam
'On no, Marsam (accidentally) carried your bag off!'

However, word order in these constructions is also not fixed:

30. a. Wah, ke-gawa tas-é mas Tom (karo Marsam).
PRT KE-carry bag-É Mas Tom with Marsam
'Oh no, your bag got carried away (by Marsam).'

b. Ke-blasuk éyang iki!
KE-lost 1.m this
'I'm lost!'

In these examples, the patient appears post-verbally. Again, as is usual, no agent or actor is expressed. This construction is very productive in Tengger, and other Javanese dialects. It has also been borrowed into Indonesian, and most Indonesian dialects. In the native form in Indonesian, *ter-* is prefixed to the verb. In a number of colloquial dialects of Indonesian, the form in *ke-* has become much more common. In all cases, what sets this construction apart is the pragmatic implication of non-volitionality, accident, or adversative. As the form and the meaning are the same as in other languages, where descriptions already exist, I will not comment further here on the accidental passive.

5.3 Voice in Standard Javanese

Voice in the standard language and in other dialects of Javanese is relatively similar to what has been described here for Tengger. The major differences are to be found in the passive forms. The active markers *M-* and *N-* function almost exactly the same throughout the different Javanese dialects (at least those with which I am familiar). Certain lexemes may vary, and the set of verbs which regularly appear with no voice morphology may not be exactly co-extensive. One major difference with what I have presented above and what is to be found in some grammars and descriptions are differences in possible word orders, and perhaps in the definiteness restrictions on subjects. Some grammars and descriptions describe the standard language as having more rigid word order than I have shown Tengger to have. However, data that I have collected on a wide variety of different dialects, ranging over the whole of the Javanese speaking area, confirm that such descriptions are prescriptive. Indeed, word order across Javanese dialects is extremely free, though to cover that topic in any satisfying depth is well beyond the scope of this study. The overall behavior of the active markers across Javanese dialects is sufficiently analogous not to have to present it any further here. The differences in the passive, however, warrant further discussion.

5.3.1 The Passive in Standard Javanese

Robson (2002) describes four distinct passive constructions in standard Javanese.[79] He notes that the passive is more frequently used in Javanese than in English. We have seen a similar range of four distinct passives in Tengger. Three of the four are almost completely parallel in Tengger and in the standard language, so I cover these briefly first.

In Robson's 'second' passive, the prefix *ka-* is added directly to a bare root. Its use is highly formal, and so appears more in *krama* and in the literary language. It does not appear to be particularly productive in the modern language, and most examples are from fossilized forms. The *ka-* passive can only appear with third person agents, and it never undergoes sandhi—though in speech it can be reduced to [kə] when before a consonant only.

The 'third' passive makes use of the infix *–in-*, developed directly from OJ. In the modern language it is no longer productive and when it appears in a few fossilized items it has a distinctly archaic flavor. When a root begins with a vowel, *ing-* is prefixed to the bare root. Again, it can only appear with a third person agent.

Unlike the preceding two types of passive, the 'fourth' passive is again very productive in both standard Javanese and Tengger. This form is characterized by the prefix *ke-*, 'which is added to root-words beginning with any consonant except *r, l,* or w. Before *r* and *l* we find just *k* and before *w, ku-*. When the root-word begins with a vowel, *k-* is prefixed and usually the vowel does not change.' (Robson 2002:76) The *ke-* passive has the force of an accidental or inadvertent act, and often there is no explicit agent. When the corresponding active appears with the locative/iterative suffix *–i*, then the

[79] There are of course both *ngoko* and *krama* equivalents. However, given the lack of a high speech style in Tengger, I only discuss the *ngoko* forms.

passive appears with *–an* suffixed. Some examples of each of these three types of passive are given below:

31. Standard Javanese Passive

	Active	Passive
ka- passive	a. ng-golong 'to group'	ka-golong 'grouped with'
	b. ng-ecap 'to print'	ka-ecap 'printed'
-in- passive	c. ng-gawe 'to make'	g-in-awe 'made'
	d. ng-utus 'to send'	ing-utus 'sent'
ke-(-an) passive	e. ng-gawa 'to carry, take'	ke-gawa 'carried away accidentally'
	f. m-banjir-i 'to flood (s.t.)'	ke-banjir-an 'to get flooded'

In addition to the three forms described here, there is also what Robson calls the 'first' passive. The 'first' passive is the most frequent voice marker in standard Javanese. It has distinct forms for first, second, and third person, though there is no distinct form for singular or plural; it inflects for person not for number. This type of passive is marked by the use of dependent clitic personal pronouns for the first and second person: *tak-, kok-*[80]*,* respectively.[81] These pronoun proclitics encode for a first or second person agent, unlike the form for the third person. For the third person, the passive prefix *di-* is added to the bare verb stem; that is, there is no nasalization. In a construction with *di-* an agent NP may appear immediately following the verb, or in an adjunct by-phrase,

[7]The first person can also appear written as *dak-*, though this is exclusively restricted to literature. The second person can be written either as *kok-* or *ko'*, though both are pronounced [koʔ].

[81] I provided arguments above for the analogous forms in Tengger being clitics and not prefixes. The same arguments hold here.

preceded by the prepositions *ing* or *dening*, though Robson notes that this is rather formal. A first or second person agent is strictly ungrammatical used in conjunction with *di-*.

The primary difference between the standard language and Tengger is that the standard language has the full paradigm of person marked passives, as follows:

32. a. Buku iku tak-jupuk. [Standard Javanese]
book that 1-take
'That book was taken by me.'

b. Buku iku kok-jupuk.
book that 2-take
'That book was taken by you.'

(32a) shows the form of the first person passive and (32b) the form of the second person passive. In both examples, the patient appears in initial position. The agent, crucially, appears pre-verbally, and there is strict adjacency between the verb and the agent. The patient can appear either pre- or post-verbally, or can remain unexpressed:

33. a. Wis tak-jupuk buku iku.
PFT 1-take book that
'I already took the book.'

b. Wingi wis tak-jupuk.
yesterday PFT 1-take

'I took (it) yesterday.'

The modal *wis* cannot intervene between the agent and verb, as seen in (33a). In (33b), we see that the patient does not need to be overtly expressed. *DI-* passives behave quite differently:

34. a. Buku iku di-jupuk Pak Marsam.
book that DI-take Pak Marsam
'That book was taken by Pak Marsam.'

b. Buku iku di-jupuk wingi *(dening) Marsam/*aku/*kowe.
book that pass-take yesterday by M 1sg 2sg
'That book was taken yesterday by Marsam.'

Here we can see that the agent can either immediately follow the passive verb or appear in an adjunct by-phrase. It also shows that a first or second person agent is clearly ungrammatical in this construction.

As with active verbs, the preverbal NP is generally, though not obligatorily specific or definite, as demonstrated by the following examples:

35. a. Kucing di-cokot asu.
cat DI-bite dog
'A/the cat was bitten by a/the dog.'

b. Kucing-é di-cokot asu.

cat-É DI-bite dog

'The cat was bitten by a/the dog.'

So, as in Tengger, there is no definiteness or specificity requirement on the arguments of a passive construction in standard Javanese, though there is a preference.

What is distinct from Tengger is the presence of the full personal inflection for the passive markers in the standard language. I have argued above that this type of passive construction is most likely an innovation that occurred after the Tengger had split from other groups of Javanese. What we see of the first and second person passive in Tengger then, is the result of very recent interference from the lowland dialects. It is not clear to me whether these constructions first developed earlier in the East Javanese courts or later in the Central Javanese courts. That the form is an innovation is clear as it did not exist in Old Javanese, to which we turn now.

5.4 Voice in Old/Middle Javanese: Diachrony of Tengger Voice

Different scholars have described the voice system of Old Javanese as a focus system (see Hunter (1999), Poedjosoedarmo (2002) *inter alia*), more akin to the focus system of the Philippine languages than to other western Indonesian languages. This however seems to conflate the active/passive distinction marked by verbal prefixes and infixes in Old Javanese, with the applicative, locative, and benefactive distinction marked by verbal suffixation. I address these verbal suffixes in the next chapter on verbal morphology. Here, I concentrate on the basic voice distinction between passive and active.

Active verbs in Old Javanese were marked either with the infix *–um-* or the prefix *(m)aN-*. There are several different affixes that function as a passive in Old Javanese. Zoetmulder (1952, 1961) distinguished two types of passive affix for OJ based on the 'degree of involvement' of the performer of the actions:

In the *–in-* passive the action and its performer are treated as more important than in the *ka-* passive. In the *ka-* passive what is treated as central is the state brought about by the action.

Hunter notes that the 'difference of agency and volitionality in the two Undergoer Focus types is brought out in differing constraints on the use of pronominal enclitics marking third person agents of *–in-* and *ka-* constructions (1999).' I have already noted above that the *ka*-passive survived almost into the present day language. It is no longer productive however. Similarly, the *–in-* passive remains only in fossilized examples, though it was the primary passive marker in Old Javanese:

36. a. Um-ahas sira ri-ng tapowana. [Old Javanese]
UM-visit 3 at penance.grove
'He visited penance-groves.'

b. Tan dadi ri sisya mangan drawya ning guru.
NEG fitting to student N-eat that.owned by teacher
'It is not fitting for a student to eat what which belongs to his teacher.'

37. a. T-in-amuy de-nira sang mahamuni.
IN-guest by-3.obl title great.sage

'He was treated as a guest by the great sages.'

b. Ikang naga p-in-angan-an ing apuy…

that dragon -IN-eat in fire

'The dragon which was eaten by fire.'

((a) examples Hunter 1999)

((b) examples Zoetmulder & Poedjawijatna 1954)

The first example above shows a typical, unmarked active, (extended) intransitive verb in Old Javanese. The verb is in initial position, marked with the infix *–um-*, (prefixed to roots with initial vowels). The second sentence shows a typical transitive verb, marked here with prenasalization, the word order here is SVO, which is uncommon in Old Javanese. The examples in (37) show the Old Javanese passive construction. In both sentences the verb is marked with the infix *–in-*. In (37a), the patient is not expressed, but the actor appears in an adjunct by-phrase. Similarly in (37b) the agent, *apuy* appears in an adjunct by-phrase.

The two active prefixes found in Tengger were both inherited directly from Old Javanese: *aN-* became the *N-* prenasalization marker, and the OJ infix *–um-* became the *M-* active marker. What we do not find in Old Javanese, however, are the passive markers *tak-* and *di-*. The first case is perhaps more straight forward. OJ had a first person pronoun *kita*, and *tak-* is simply the reduced, dependent form, with a final glottal stop. The marker *Di-* is more complex. Wolff (1996), reconstructs *di-* as a reflex of the infix *–in-*, for Malay becoming *ni-* and finally *di-*. Ross (2004), however claims that *di-* is an innovation in Proto-Malayic. He claims that the form is not based on the old PAN passive infix, but rather on the agentive preposition *di-*, and that it has spread through

contact to other languages, including Javanese and Tengger. Another proposal is similar to that, that it is derived from the third person pronoun *dia*, which has been grammaticalized. This nicely accounts for the third person restriction on agents in Indonesian; however, many other varieties of Malay allow first and second person agents with *di-* passives. I leave open here the question of the exact origin of the *di-* passive in Javanese, whether it was borrowed from Malay or inherited from Old Javanese.

For the passive *semu* (passive constructions with the order theme-proclitic agent-bare V), however, there is something more to be said here. The passive *semu* in Javanese, characterized by a bare verb stem with agent pro-clitic, was clearly not present in Old Javanese. Ross (2002) reconstructs back to both Proto-Malayo-Polynesian and Proto-Austronesian bare verbs as a-temporal, non-indicative actor voice. We will see in the next chapter that Tengger has an optative, marked by a suffix, not a verbal prefix, which encompasses both imperative and irrealis moods. These forms in Tengger are clearly inherited from PAN. Wolff (1996) claims that the bare verb stem (with a proclitics agent) in a passive voice is an early innovation in Indonesian, but he notes that it is a reflection of the PMP a-temporal patient, location, and circumstantial voice forms due to the loss of the original suffixes. If the passive *semu* is in fact an innovation in Malay/Indonesian, then it has been clearly borrowed into standard Javanese. This actually goes a long way to explaining why Tengger does not have the construction productively, except through very recent contact phenomena. The passive *semu* must have been borrowed into Javanese after the period of Old/Middle Javanese, as the construction is not found there, but after the Tengger had already split off from the main

lowland populations. It spread widely throughout the Javanese speaking area, though not completely:

38. a. La ng-embil epring, tak-sogok mau pring [Banten Javanese]
EXCL N-take bamboo 1-prod earlier bamboo
'Then I took a small piece of bamboo and I poked into it.
Bamboo.'

b. Anu sog-an tag-lèdhèg-i ng-guyu-guyu thog [Banyumas Javanese]
um often-AN 1-tempt-I N-RED.laugh only
'…um I often flirt (with him) but he just laughs and laughs.'

These examples show that the construction has spread to many dialects of Javanese, including some very isolated dialects such at Banten, spoken on the far west coast of Java and surrounded by speakers of Sundanese.[82] In many dialects the second person form is rarely if ever encountered.

5.5 Typology of Voice: Symmetric Voice Systems

Much recent work on Austronesian voice, particularly typological studies such as those of Foley (1998), Kroeger (1998), Arka (2002), Himmelmann (2005), *inter alia*, though also more formal studies such as Cole, Hermon, and Yanti (2008), have attempted to

[82] It appears as if Osing Javanese, spoken on the far east coast of Java, like Tengger, does not have the passive *semu*. This might indicate that the construction reached its eastern limit in the Tengger area, however the construction is well attested in the Lumajang and Jember dialects, which are to the east of the Tengger area. However, more research is needed to determine the full geographic spread of the construction.

define and place particular languages and language groups on a scale of symmetrical versus asymmetrical voice systems. A symmetrical voice system is argued to characterize the Philippine-type languages, many Formosan languages, and has even been reconstructed for Proto-Austronesian. There is a cline that exists, with symmetrical systems gradually eroding through the western Indonesian languages, and not existing at all in eastern Indonesian and Oceanic languages. From this point of view, it will be very interesting to see exactly where Javanese sits, as it is on the boundary, between those languages displaying symmetrical or asymmetrical voice systems. It has been argued that some languages, such as standard Indonesian, are actually mixed languages, having both symmetrical and asymmetrical voice constructions. Before presenting the Tengger data, I will briefly describe what is meant by a symmetrical voice system.

Different scholars have slightly different definitions for what constitutes a symmetrical system. In general, a symmetrical voice system is characterized by the following primary properties (Cole et.al 2008):

39. a. All forms of the verb are marked for voice.
    b. In 'passive-like' constructions agents are not demoted to adjuncts, but rather are arguments of the predicate
    c. There are a variety of passive-like voices, promoting nominals exhibiting different grammatical and semantic relations to surface subjecthood.

These can be illustrated very clearly with examples from Tagalog:[83]

40. a. b-um-ili ng isda sa tindahan ang lalake [Tagalog]
VC-buy CORE fish OBL store man
'The man bought fish in the store.' [agent subject]

b. bi-bilh-in ng lalake sa tindahan ang isda.
IRR-buy-VC CORE man OBL store fish
'The man will buy fish in the store.' [theme subject]

c. bi-bilh-an ng lalake ng isda ang tindahan.
IRR-buy-VC CORE man CORE fish store
'The man will buy fish in the store.' [locative subject]

d. ipam-bi-bili ng lalake ng isda ang salapi
VC-IRR-buy CORE man CORE fish money
'The man will buy fish with the money.' [instrumental subject]

e. i-bi-bili ng lalake ng isda ang bata.
VC-IRR-buy CORE man CORE fish child
'The man will buy fish for the child.' [benefactive subject]

(Foley 1998:2)

The verbs in (40) above appear with a number of different voice affixes, prefixes, suffixes, infixes, reduplication, and applicatives. Either independently or in conjunction

[83] I use Foley's glosses here. VC is a voice affix. CORE/OBL are nominal markers indicating whether they are core or oblique arguments. IRR is irrealis mood (reduplication in all the examples here).

with another affix, they function to indicate the semantic role of the noun marked with *ang* in each sentence. The *ang* nominal is the syntactic and surface subject of the sentence, and in each of these examples appears in final position, yielding VOS order – the unmarked order in Tagalog. That *ang* identifies the 'subject' of the clause is not uncontroversial, however, it is an issue orthogonal to the current discussion.

In (40a) above, the agent is marked with *ang* (marking it as the subject), and so the verb takes the *–um-* infix. For a theme subject, *-in* is suffixed to the verb; for a locative *–an* and so on. Note crucially the difference between (40a) and (40b). In the agent focus voice, or active, the agent is marked with *ang* and the theme is marked as a core argument, with *ng*. For the patient/theme focus counterpart in (40b), the theme is marked as the subject with *ang*, and the verb has 'passive' or theme focus morphology. Note, however, that the agent is also a core argument, marked with the same marker that marked the theme in the active construction *ng*. The agent does not appear in an adjunct phrase, as in asymmetrical voice systems such as English:

41. a. The man bought fish at the store. [English]
    b. Fish were bought at the store by the man.
    c. *Fish were bought the man at the store.

So, we can see that Tagalog satisfies the basic requirements of a symmetrical voice system, as given in (39) above.

Some of the Austronesian languages, of course, display properties of both symmetrical and asymmetrical voice systems, and are thus labels mixed languages:

Standard Indonesian, standard Javanese, and, as Cole et.al. (2008) argue, most colloquial dialects of Malay—to a differing degree—are all mixed languages. There are three distinct voice constructions, referred to by various names, here I use active (a) examples, passive (b) examples, and passive *semu* (or object voice, second passive, P2, etc.) for (c) examples[84]:

42. a. Isteri-ku me-lihat orang itu. [Standard Indonesian]
wife-1.sg meN-see person that
'My wife saw that person.'

b. Orang itu di-lihat isteri-ku.
person that DI-see wife-1.sg
'That person was seen by my wife.'

c. Orang itu saya-lihat.
person that 1.sg-see
'I saw that person.'

43. a. Bojo-ku n-delok wong kaé. [Standard Javanese]
wife-1.sg N-see person that
'My wife saw that person.'

b. Wong kaé di-delok bojo-ku.
person that DI-see wife-1.sg
'That person was seen by my wife.'

[84] There are some very significant differences between the passive semu in standard Indonesian and Javanese. For standard Javanese, the agent is restricted to the proclitics pronouns *tak-* and *kok-*. Nothing else can appear immediately preceding the verb in this construction. In Indonesian, proclitics pronouns, both phonologically reduced and full forms, proper names, and other NPs can all appear in the agent position immediately preceding the verb.

c. Wong kaé tak-delok.

person that 1-see

'I saw that person.'

The (a) and (c) examples above show the 'symmetric' alternation in Indonesian and Javanese. It has been fairly well established since Chung (1976) that the patient in the passive *semu* (c) acts as the surface or syntactic subject of the clause. More recently, the agent has been analyzed as an argument (Guilfoyle et.al. (1992), Cole et.al (2008)) based on its surface position preceding the verb, its obligatory appearance, and crucially, on its ability to antecede reflexives, this, however, is not the case in Standard Javanese:

44. a. Diri-mu mesti kau serah-kan ke polisi. [Standard Indonesian]

self-2.sg must 2.sg give.in-KAN to police

'You must surrender yourself to the police.'

(Arka and Manning 1998)

b. *Dhéwé kudu kok-serah-aké nang polisi. [Standard Javanese]

self must 2-give.in-AKÉ to police

'You must surrender yourself to the police.'

The difference, however, is not due to the status of the agent, but to differences between the reflexives themselves. The Indonesian reflexive is inflected for person, though not for number, whereas the Javanese reflexive is not inflected for person but does have

distinct forms for singular *dhéwé* and plural *dhéwé-dhéwé*. The two forms also have a different distribution. I discuss reflexives more shortly.

What is important here is that the agent in a passive *semu* is in fact an argument. It is not 'demoted' to an optional adjunct as in English and other asymmetrical voice languages. On first glance, it would appear as if the agent in examples (42b) and (43b) above also demonstrate symmetrical behavior. As I have shown above for Tengger, this is also not the case for Indonesian or Javanese:

45. a. Orang itu di-lihat kemarin (olèh isteri-ku). [Standard Indonesian]
person that DI-see yesterday by wife-1.sg
'That person was seen yesterday by my wife.'

b. Wong kaé di-delok wingi (dèning bojo-ku).[Standard Javanese]
person that DI-see yesterday by wife-1
'That person was seen yesterday by my wife.'

In both standard Indonesian and Javanese, the agent of a passive is optional. Further, it can appear either immediately following the verb, as in (42) and (43) above, or in an adjunct by-phrase, as in (45). So, is the agent of a passive an argument? Certainly (42) and (43) would appear to suggest that it is an argument, as there is strict adjacency with verb. Cole et.al (2008) argue contra Donohue (2006) that the agent in a passive is always an adjunct, regardless of whether or not it appears with a preposition. Certainly its optionality would suggest that it is not a true argument, but further evidence, based on

Arka and Manning (1998) comes from the fact that the agent in *di-* passives cannot act as the antecedent of a reflexive:

46. a. *Diri-nya tidak di-perhati-kan Amir. [Standard Indonesian]
self-NYA NEG DI-attend.to-KAN Amir
'Himself was not taken care of by Amir.'

b. *Diri-nya selalu di-utama-kan Amir.
self-nya always DI-priority-KAN Amir
'Himself was always given priority by Amir.'

The ungrammaticality of these examples shows that the agent cannot serve as the antecedent for the reflexive, and hence is an adjunct and not an argument, even though it is not 'demoted' in these examples.

It is difficult here to make direct comparison with the Javanese. The parallel form to the Indonesian reflexive *diri-nya* in Javanese is *dhéwék-é*. However, this form has become part of the pronominal paradigm in the standard language, and acts as a third person pronoun. Remember from the discussion in chapter 3 that Tengger has no real third person pronoun. In fact, in many Javanese dialects there is no distinct form, with demonstratives and locative adverbs often serving as personal pronouns. For this particular case, then, the equivalent Javanese examples would be grammatical, but not strictly anaphoric:

47. a. Dhéwék-é ora di-delok Amir.

self-É NEG DI-see Amir

'He (himself) wasn't seen by Amir.'

b. ??Dhéwé ora di-delok Amir.

Self NEG DI-see Amir

The above example is perfectly grammatical in Javanese. However, if the true reflexive is used *dhéwé*, without the *-É* ending, then the sentence becomes marginal at best (47b).

Of course, it is not only descriptively convenient to distinguish these types of voice systems. A number of other grammatical properties are associated with one type of system or the other. Himmelman (2005) provides a typological distinction between symmetrical voice system languages on the one hand, and preposed possessor languages on the other, as two basic types of languages found in the western Austronesian area. There are a bundle of different properties that are associated with each. To begin with, there is a negative correlation between symmetric voice and possessor preposing, that is languages with preposed possessors tend either not to show any grammatical voice alternations, or manifest clearly asymmetrical voice alternations.

There are also syntactic correlations. For example, in many symmetrical voice languages, there are severe restrictions on extraction processes, such as WH question

formation and relativization. In Tagalog, only the NP marked with *ang*, the syntactic subject, can be questioned or relativized[85]:

48. a. Isda=ng i-b-in-igay ng=lalake sa=bata [Tagalog]
fish=LNK OV-PERF-give GEN=man DAT=child
'The fish which was given to the child by the man.'

b. bata=ng b-in-igy-an ng=lalake ng=isda
child-LNK PERF-give-DV GEN=man GEN=fish
'The child which was given fish by the man.'

c. *Isda=ng nag-bigay ang=lalake sa=bata
fish=LKN AV-PERF-give NOM=man DAT=child
'The fish which the man gave to the child.'

d. *isda=ng b-in-igy-an ng=lalake ang=bata
fish=LNK PERF-give-DV GEN=man NOM=child
'The fish which the child was given by the man.'

(Kroeger 1993:23-24)

5.6 Summary and Conclusion

[85] Again here I use the same glossing as in the original, in this case Kroeger. LNK = linker; PERF = perfect; OV = object voice; AV = actor voice

In this chapter we looked at the voice morphology in Tengger. At first glance, it seems to behave like a fairly unremarkable system, much like that described for other Indonesian languages. To begin with, Tengger does not have the complex voice marking system found in Tagalog and the other Philippine languages, where the verbal voice marker encodes the semantic role of the subject: agent, theme, locative, benefactive, and instrumental. The general pattern in Tengger is for active verbs to appear with prenasalization, with the agent preceding and the theme or patient following the verb (2). Similarly, for the passive construction the marker *di-* the patient or theme appears in preverbal position and the agent appears either immediately postverbally or in a demoted by phrase (11 and 12). However, I also noted that other word orders are possible, even with overt voice morphology. So we have examples like (5) for active and (13) for passive where we have the opposite of the expected order. This verbal morphology then seems no to directly affect the syntax, that is the linear order of constituents. Clearly some pragmatic function is at work here, how else are we to account for the following:

49. a. Pak Tom di-critak-na karo Pak Danu-né.
Pak Tom DI-story-NA with Pak Danu-E.
'Tom told the story to Pak Danu.

b. Pak Tom nyritak-na karo Pak Danu-né. < N-critak
Pak Tom N-story-NA with Pak Danu-E.
'Tom told the story to Pak Danu.'

In (49a) the predicate appears with the passive marker *Di-* and in (49b) it appears with the active marker *N-*, yet both sentences have the same linear order and roughly the same meaning. Clearly these voice markers do not affect the syntax.

So, the crucial question we must answer is what then is the function of these 'voice markers'? I would propose that these are semantic markers, and not syntactic markers. They overtly mark the presence of either an agent/actor argument in the case of the *N-* prefix, and a patient/theme argument in the case of the *DI-* prefix. The argument need not be overtly expressed, and indeed, in Tengger it is most often omitted. Further, it is not necessary, though it is sufficient, to have the voice marker in order to license the argument. In many cases they are optional. This helps to explain why often both a passive and an active interpretation are available.[86]

I flesh out this proposal in the next chapter on morphosyntax, where the evidence from applicatives provides further evidence.

[86] In many ways, then, these markers are like plural reduplication in Tengger. It is optional, and generally left unexpressed, with the plural reading available from the context.

CHAPTER SIX: MORPHOSYNTAX

6.1 Introduction

Having considered the voice system in Tengger in the previous chapter, in this chapter I explore other aspects of Tengger morphosyntax: mood, the propositive, applicatives and the interaction of these elements. In many respects, the Tengger system is relatively strait forward, when viewed comparatively both diachronically and synchronically. As with most western Indonesian languages, Tengger has lost the focus system common in the Philippine languages, and also present in Old Javanese. However, it has not developed the elaborate applicative system which characterizes many other dialects of Javanese. I argue that the Tengger system represents an intermediary stage in the development from a focus system marked through infixation, found in Old Javanese, and the complex system of applicative suffixes characterizing standard Javanese and other dialects. Further, I argue that, unlike applicatives in other languages, Tengger applicatives do not affect argument structure, and contra Son and Cole (2008) that there is no direct mapping from the argument structure to the overt syntax. Rather, I adopt a modified version of Gil's (2002) analysis of Riau Indonesian generalized voice markers, arguing as I have for the voice markers in the previous chapter that the applicatives in Tengger operate on the semantic frame of a verb, but they do not affect the overt syntax. I place the Tengger within a typology of generalized voice markers, showing that they are weak, indicating the existence of a particular argument, but not necessarily implying any correlate properties.

1.0 Mood in Tengger

Given the complete lack of inflectional morphology in Tengger, including tense, number, aspect, and gender, it would appear at first glance that the morphosyntax of predicates is rather simple. This is, however, not the case. Beyond the verbal prefixes discussed in the previous section, Tengger has a very complex system of suffixal morphology which interacts with base words to affect the type and number of complements, mood, transitivity, ordering of constituents (in a very limited way as I show below), and focus and topic relations in a given clause.

Before moving into a discussion of applicatives in Tengger, we need to take a look at how mood is marked. I present here a very different typology of mood than has been presented for Tengger (Smith-Hefner 1983) or for standard Javanese (e.g. Robson 1992, Poedjosoedarmo 1996, Uhlenbeck 1970, inter alia). First, I claim that there are only three distinct moods, and not four. Second, I argue that there is a person distinction in the imperative/subjunctive. There is also a complex interaction among voice, mood, and applicatives that we will discuss shortly.

The Tengger predicate distinguishes between three different moods: indicative, propositive[87], and imperative/subjunctive, most with unique forms for active or passive. Except for the propositive, mood is marked by the addition of elements to the right edge of the root, as demonstrated in the table below , using the verb *jukuk* 'take' as an example:

[87] I discuss the propositive in more detail in a subsequent subsection. It is marked through prefixation, and not suffixation as are the other moods. It will become more relevant when we compare the applicative system of Tengger with that of other Javanese dialects.

40. Mood in Tengger

	Active	Passive I	Passive II	No voice
Indicative	n-jukuk	di-jukuk	ke-jukuk	
Optative	n-jukuk-a	2p: jukuk-en 3p: di-jukuk-a	Ø	iku-a, kéné-a
Propositive	(isun) tak n-jukuk	Ø	Ø	

Neutral or indicative verbs are marked by their lack of suffixation, in both the active and the passive forms.

The next mood is what I call here optative, and I consider the paradigm above to all fall under this unified category. That is, as opposed to the indicative, which has just one form for the active and each of the passives, the optative has one form for the active, and two distinct forms for the passive I. There is yet another form which is neither active nor passive. The optative does not appear with the passive II (see below). The optative encompasses a large variety of meanings, though generally they can be characterized as irrealis and imperative. It can also express a desire for how things should be, or should have been. The optative can also be used to express a polite command. For the active, *-a* is suffixed to the pre-nasalized root:

2. a. N-jukuk-a pethès-é teka wuwung-é!

N-take-a monkey-e from roof-e

'Get that monkey off the roof!' (Someone should take the monkey off the roof!)

b. Ng-ombé-a kopi-né, pak, ng-ombé-a.

N-drink-a coffee-é, VOC, N-drink-a

'Drink your coffee, drink your coffee.'

In the first example above, we can see that the speaker is issuing either a general command, or a strong wish that a certain state of affairs would come about, i.e. the monkey being off the roof. In the second example, a polite command is given as way of an invitation to drink. In fact, one of the most common phrases heard in Tengger villages is *mampir-a* 'stop by, drop in', which makes use of optative in the form of a polite command.

If we look at the imperative meaning of the optative here, then Tengger is rather rare from a cross-linguistic view. In most of the world's languages, the imperative form is often the smallest verbal form, or at least reduced vis-à-vis the indicative (Xrakovsky and Volodin 1986). Here, to the contrary, the imperative is formed, essentially: *indicative* + *a*, that is, the indicative *plus* some other ending. Typologically, then this sort of imperative construction is, at least, rare.

In the passive, the optative is marked for person. This is a striking feature of Tengger, for no other dialect of Javanese, and almost no other Indonesian languages have

distinct grammatical markers for person (except the passive clitic pronouns). In the second person passive optative, *-en* is suffixed to the bare root:

3. a. Pethès-é jukuk-en teka wuwung-é
   monkey-e take-en from roof-e
   'That monkey should be taken off the roof (by you)!'
   b. Maringono japa-en entas-entas ya …
   afterwards magical.incantation RED-quick aff
   'After that (you) say an incantation [i.e. pray] then quickly

It is important to note here that the second person passive optative has a bare root whereas the active has the nasal prefix. This is exactly the opposite of Indonesian imperative, and that of other related languages where the active imperative takes a bare root and the passive imperative has passive morphology:

4. a. Ambil-lah monyet dari atap! (Indonesian)
   take-prt monkey from roof
   'Get that monkey off the roof!'
   b. Silahkan, di-ambil-(lah).
   please di-take
   'Please take (some/it).'

This distinction plays an important role in the interplay of mood and applicatives that I discuss shortly.

Both of the examples above relating to getting the monkey off the roof, in (3:a and 4:a) have the force of an active imperative, that is, they are both clearly commands directed at an individual or individuals. The formal distinction here encodes a politeness effect, common in Indonesian languages, where an agentless passive is used as a polite form of a command. There is also a distinction in the nature of the optative markers, as we shall see below.

In the third person passive optative the predicate appears with passive morphology prefixed, and takes the optative marker –a:

5. Gaga-né di-panja-a karo kenthang.
field-e di-plant-a with potato
'That field should be planted with potato.'

The force of the optative here is not restricted to simple irrealis. It also carries with it a hortative sense, of what 'should be' or 'should be done'. It is distinct from a simple imperative in not directly commanding anyone to do anything. The above example is from a story a farmer was recounting. When prices for garlic were high, he planted garlic—as did many others. His neighbor planted potatoes. When harvest time came, there was a glut of garlic, and dearth of potatoes. He was reprimanding himself, in a way, he should have planted potatoes!

It is also interesting to note in regard to this, the formal similarities between the active optative and the third person passive optative. They are both marked with *–a*. This would suggest that they are an exclusive set. However, it is very clear from the semantics, that the second person passive optative also belongs to this set. There is also a further member of this set.

The optative is unique among verbal morphology in Tengger in being able to take almost word as its root, even functional words:

6. a. Yo turu-a kéné-a nak ya
aff sleep-a here-a TRU-child aff
'Just sleep here kids.'[88]

b. tuku iku-a cap lenga wangi
buy that-a brand oil fragrant
'Just buy that one, wangi oil brand'

In the first example above we have two instances of the optative. On the first it marks the verb *turu* 'sleep' and the demonstrative pronoun *kéné* 'here'. The verb *turu* here is active, it's part of the small class of verbs which does not take the nasal prefix. But more interesting here is that the optative can also be added to a demonstrative pronoun. Note also that there is no prosodic break between *turu* and *kéné* here, that is, this is a single prosodic utterance. So it is not that case that these are two independent imperatives, *Sleep! Here!* Rather they must be taken together. The demonstrative in that case seems

[88] Note the mixed use of the lowland *yo* and Tengger *ya*.

to have more of a desiderative meaning, or even a hortatory, suggesting, but not commanding the kids to just sleep *there*.

Tengger has an additional strategy to mark commands. Here the verb has no affixation but is rather marked through intonation. This is actually the most common form of imperative in Tengger.

7. Nduk, turu kéné, turu ng-arep.
   VOC.f sleep here sleep N-front
   'Sleep here nduk, sleep out front (with us)!'

One final and unique use of the *–a* ending in Tengger is as a general interrogative. This is not a verbal ending however, though it is homophonous with the verbal *–a*. In this usage, it generally appears on the final element in the utterance, and is accompanied by a rising intonation at the end:

8. a. Turu kéné-a?
      sleep here-a
      'Are you going to sleep here?'
   b. Dhèk kapan réné-a?
      past when come-here-a
      'When did you get here?'

In its interrogative use, the *–a* marker can attach to any root. The root appears without any morphology.

There is no optative marker that goes with the passive II. Remember that the passive II marks accidental or adversative events, real events and often past events. As such, it is semantically incompatible with the optative which indicates irrealis, events that should be, or commands, which are not yet complete.

6.2.1 The Propositive

The propositive (desiderative) is a construction perhaps unique to Javanese among Indonesian languages. Although its formation is different from the other moods considered here, it still has the semantic force of a mood. The propositive only functions in the first person (and as such is like a cohortative as opposed to a jussive mood). Remember though that while number is not morphologically marked, even in the pronominal system, the propositive almost always has a singular meaning. For the propositive, the proclitic *tak*[89] is placed before the predicate, which has obligatory prenasalization[90].

9. a. (éyang) tak ng-adus-i sik ya.
1.m prop N-bathe-i first aff
'I'm going to wash him first.' [speaking of a child]

[89] In older and often in the written language the form is *dak*. In the spoken language, the initial consonant is invariably voiceless.
[90] Except on those roots which never take prenasalization.

b. panci-né (isun) tak n-jukuk sik.

pot-é 1.f prop N-take first

'Let me get the pot first.'

The propositive adds a meaning of intentionality or immediate plans, as in (9:a) above. It can also indicate a subtle request, as in 'let me be the one who…' as seen in (10:b). As mentioned above, it is restricted to the first person, and has an active reading, as indicated by the nasal prefix. The first person agent pronoun is optional, though generally omitted. We can tell that the *tak-* marker is a clitic by its ungrammaticality when it appears not immediately preceding the predicate:

11. a. *tak éyang ng-adus-i sik ya.

b. *panci-né tak isun n-jukuk sik.

As we will see shortly, the propositive can also take applicative suffixes, and the different constructions interact in complex ways.

There is also a passive propositive construction in Standard Javanese. In this case, there are a unique set of verbal suffixes, in order to mark a construction as propositive, which would otherwise be formally indistinguishable from the first person passive construction.

1. Standard Javanese Propositive

	Active	Passive
Standard Javanese	(aku) tak n-jupuk	tak-jupuk-é
Tengger	(éyang) tak n-jukuk	Ø

However, Tengger does not make use of these affixes. So, if there were in fact a passive propositive it would be formally indistinguishable from the first person passive. Further, as discussed above, the force of the first and second person passive constructions is often reported as active. So the two constructions would be both formally and semantically almost indistinguishable.

Smith-Hefner reports that even the active propositive is generally absent from the Tengger repertoire. Given the above discussion it is clear that there is only one formal distinction between the active propositive and the first person indicative passive constructions. The propositive verb must have the active voice marker *N*- whereas the passive has the bare stem. This is another instance of very recent interference from the lowland dialects.

6.3 Mood in Standard and Lowland Javanese

Unlike Tengger, many dialects of Javanese distinguish clearly between four distinct moods: indicative, propositive, imperative, and subjunctive. Similarly to Tengger, there is no marking

Mood	Voice	Prefix	Ending

beyond the voice markers on the indicative. The propositive as we noted above, is probably a recent borrowing into Tengger, therefore is has the same form as in the standard and other lowland dialects for the exception of the agent pronoun, if present.

The major difference between the standard language and Tengger is in the imperative and subjunctive paradigms. My analysis has essentially collapsed these two into a single paradigm. That's not exactly it either. They haven't 'been collapsed' at all. In fact, it seems more likely that the imperative and subjunctive in the standard language are really part of the same paradigm. Consider first the active subjunctive and the active imperative are indistinguishable formally, both take the nasal prefix and both take the ending *–a*. In the passive, there are two formal differences between the imperative and the passive. The passive imperative takes a bare root suffixed with *–en*, the passive subjunctive takes passive verbal morphology and the marker *–a*.

13. Mood in Standard Javanese

Mood	Voice	Prefix/Proclitic	Suffix/Enclitic
Indicative	Active	N-/M-	Ø
	Passive	tak-/kok-/di	
Propositive	Active	(aku) tak N-	Ø
	Passive	tak-	-é
Imperative	Active	N-	-a
	Passive	Ø	-en
Subjunctive	Active	N-	-a
	Passive	tak-/kok-/di-	-a

In his grammar, Robson points out that the subjunctive is often found only in subordinate clauses. This may be the distinguishing factor between the active imperative and the subjunctive. Robson notes that the subjunctive is 'to be found in non-literary as well as literary sources ... (it) is an important part of Javanese grammar' (1992:131). The subjunctive may well appear in non-literary sources, but it is not often heard in daily speech, if at all. And it certainly has a very formal air to it.

6.4 Mood in Old Javanese

Old Javanese also had an irrealis mood. There were two distinct markers, but they did not encode voice distinctions or person distinctions, as with the Tengger and other modern Javanese dialects. Rather there was a distinction between transitive and intransitive (incl. reciprocal and some non-volitional transitives).[91]

14. Old Javanese Irrealis

	transitive	intransitive
OJ irrealis	-a	-en

15. a. ganjar-an anut-a guna-nira [Old Javanese]

reward-an aN-follow-a quality-3

'The reward should be in accord with his skill.'

[91] It is possible that the distinction encoded by the two irrealis markers in Old Javanese could be characterized in general terms of volitionality. However that would require much more work, and a better understanding of Old Javanese than I currently possess.

b. dur-an hana dasa-sila ka-wruh-en-mw-en

distant-an exist ten-virtues ka-know-en-2-en

'It's hard to believe that you would know of the ten virtuous practices!'

[Hunter 1999:23]

In Old Javanese, the force of the irrealis mood seems much closer to what we are perhaps more familiar with (in that the term irrealis originated to describe the mood in classical western languages). It marks an unreal, unaccomplished event, or states how someone feels something should, would, could be or have been.

Diachronically, it is clear that the markers in Tengger as well as the standard language derive directly from Old Javanese. However, there is a major semantic shift in the function of these markers. Where Old Javanese has different forms depending on the transitivity of the root, the standard language and Tengger both have forms encoding voice distinctions. Tengger goes even further to encode distinct forms for second person. We saw above where Tengger is unique among almost all Austronesian languages in encoding a gender distinction in first person pronouns.

The imperative in Old Javanese had two forms as well. Bare verb roots indicated the imperative, so that *inum* 'Drink it!' and *linggih* 'Sit down!' Typologically of course this is a more common type of imperative as compared to the Tengger form. For predicates with a nominal base, the prefixes *pa-/paN-* indicated the imperative (as opposed to the indicative in *ma-/maN-*): *pa-weh* 'give (it)!' and *pang-renge* 'Hear! (listen)' These examples come from Ogloblin (2005).

Historically then the Tengger optative markers come from Old Javanese imperative markers. There was a semantic shift in the forms, from encoding valency to encoding voice and person. Finally, there seems to have been a total semantic collapse of the Old Javanese imperative and subjunctive. However, it is also possible that the Old Javanese subjunctive was purely a literary mood, similar to the passé simple in French, or forms of Bokmal Norwegian. I leave that question open here.

6.5 Applicatives

In addition to the voicing contrasts marked through prefixation, and mood marked through suffixation, there is a series of predicative applicative suffixes in Tengger. Cross-linguistically applicative endings serve to encode particular relationships between verbs and arguments, promote oblique arguments, and affect the valency of a verb. They can relate an argument to an event or encode the relationship of two arguments. By affecting the valency of a verb, they can introduce new arguments, or license NPs in the verb phrase.[92] While most languages have only a single applicative marker, some languages do have multiple markers which generally correspond to different semantic categories, benefactive, locative, instrumental, etc.

Similarly to many other Austronesian, and especially Indonesian languages, applicatives are used to add or indicate the role of arguments and adjuncts in the sentence. Like Standard Indonesian, but unlike most colloquial varieties of Indonesian, Tengger has two distinct applicative affixes, but given that they each display a wide

[92] For standard Indonesian, Hopper and Thomson (1980) call *–kan* a valency increasing morpheme. Arka (1992), Sie (1989) and Postman (2002), among others, claim it is a transitivizer.

variety of meanings, and in some cases even overlap, I simply refer to them as applicative I and applicative II. The Tengger applicatives also interact with voice and mood affixes, sometimes in a very complex way.

In this section, I present the applicative predicate paradigm that operates in Tengger. First, I provide a description of the facts, before moving into an analysis and comparison of the Tengger applicatives both diachronically with Old Javanese, and then synchronically, with the patterns found in the standard language, and some other relevant dialects.

6.5.1 Applicative I

Tengger has two classes of applicative markers applicative I and applicative II. These applicatives occasionally have distinct forms depending on the voice and mood of the sentence. The applicative I suffix has a wide range of semantic meanings. Frequently, it marks location, and is thus sometimes referred to as a locative suffix. In some ways it is thus analogous to a locative applicative in e.g. some Bantu languages (see Baker 1988, Marantz 1993, Pylkkänen 2002, inter alia).

12. Applicative I in Tengger

	Active	Passive I	Passive II
Tengger	-i	-i	-an

In its neutral form in indicative clauses, the suffix –i is added directly to the root of the active or passive predicate[93], which bears normal verbal morphology (i.e. *N-* for active and *di-* for passive):

17. a. Dhek wingi éyang manja-ni gaga karo kenthang.
prt yesterday 1.sg.m N-plant-i field with potatoes
'I planted the field with potatoes yesterday.'

b. Dhek wingi éyang manja kenthang dhek gaga.
prt yesterday 1.sg.m N-plant potatoes prt field
'I planted potatoes in the field yesterday.'

These two examples provide a clear contrast. In (17:b) the predicate appears with out the *–i* suffix. In this case, it is followed directly by the patient argument, and the locative adjunct is introduced by *dhek*. In (17:a) where the predicate has the *–i* suffix, it is followed immediately by the locative argument. In this example, the patient is an adjunct, introduced with *karo*. Note then the total effect of the *–i* suffix here, the argument structure of the predicate was not changed in number, but in semantic role. This usage of applicative I corresponds very directly to locative applicatives in other languages. According to standard analyses of applicatives, in this type of construction the locative goal is shifted from an adjunct to an argument position. In relational grammar terms, the oblique is promoted to a core argument.

This is however, an insufficient characterization of the applicative I. Describing the applicative I in the standard language, Robson (1992:57) characterizes verbs with the

[93] If the root ends in a vowel, the ending will be *–ni* with concomitant vowel sandhi.

–i suffix simply as having 'the addition of an extra dimension of meaning.' By which he means the difference between an unmarked verb and one with the applicative I marker is often subtle, frequently idiosyncratic, and occasionally imperceptible or even stylistic. However, beyond the locative function described above, there are some patterns that can be discerned.

The applicative I, in addition to licensing locative adjuncts, can also seem to affect the basic argument structure of a predicate, making an intransitive root transitive, or even a transitive root ditransitive. Consider the following:

18. a. Montor-é mesti m-undur sik!
motorbike-é must M-reverse first
'The motorbike has to pull back first.'

b. Lho, dhéké mesti ngundur-i sapi.
prt 3 must N-reverse-i cow
'He's got to get his cows out of the way!'

In (18:a) above there is only one external argument *montor* 'motorbike'. Unlike languages such as English, unergative verbs in Tengger can be transitive. In order to make unergative verbs like withdraw, pull back, enter, talk, etc. transitive, the most common strategy is to use the applicative I marker. So from the unergative *mundur* 'reverse', we get the transitive *ngunduri* 'reverse s.t.' Note also, though, that the prefix has changed, and the transitive verb also takes the *N*- nasal prefix.

There are also non-argument affecting functions of the applicative I as well, which have a purely semantic effect upon the predicate. For example applicative I can simply mark iterativity of an action or plurality of a patient. For example, *mangan* 'to eat', *mangan-i* 'to eat again and again; to eat many things'.

There are also cases where multiple functions are acting simultaneously:

19. a. Éyang madhah-i pelem-é nang tas plastik.
1.sg.m N-container-i mango-é to bag plastic
'I put the mangoes into a plastic bag.'

b. Apik-é pelem iki di-wadhah-i tas plastik.
good-é mango this di-container-i bag plastic
'It would be best for this mango to be put in a plastic bag.'

In the first example (19:a), in somewhat poetic language, the substantive *container* appears as a predicate with nasal prefixation. However, in order to make it transitive, it must appear with the *–i* suffix. In this case, an overt directional, *nang* is required to introduce the locative adjunct. However, in the second example, the predicate is followed directly by the locative goal. In this case, the *–i* suffix is licensing the locative goal, which entails also marking transitivity.

Yet another function of the applicative I is causative. This is not as common as the locative or transitivizing uses of the applicative, and it is also not typologically common for an applicative ending to mark causative predicates. Most of the roots that

take causative *–i* do not also appear with the applicative II, though some do. So we have examples such as:

20. a. mati ‘die’ matè-ni ’kill’
    b. ilang ‘gone, disappeared’ ng-ilang-i ‘remove s.t., make s.t. disappear’
    c. wedi ‘scared’ medè-ni ‘frighten’

So far, I have only presented examples of the applicative I with active voice morphology on the verb. The forms and functions of the applicative I are the same for the passive I as they are for the active. For the passive II, however, the form of the applicative ending becomes *–an* and not *–i*. This is the same as in the lowland and standard language. There is also vowel variation that occurs between both the root and suffix and prefix. So we get the following examples:

21. a. éling ‘remember’ k-éling-an ‘remember s.t.’ *k-éling-i
    b. lali ‘forget’ ke-lalè-n ’forget s.t.’ *ke-lalè-ni

There are other, very idiosyncratic uses of the applicative I, but as this is not a grammar, we need not be exhaustive here. What has been presented above is sufficient for our discussion and analysis below.

6.5.2 Applicative II

The applicative II also has a wide range of semantic meanings, though again some broad patterns can be discerned. However, in Tengger the form of the applicative II is interesting. Smith-Hefner reports the form as being *–en* and notes that this is a particular feature of the dialect, however she adds:

> The saliency of verbal morphology as an identifying feature for both standard and dialect speakers is very low. Many lowlands immigrants to the Tengger area never adjust their speech completely to the dialect system of verb morphology.
>
> (Smith-Hefner 1983:158)

Since her fieldwork was completed almost 30 years ago, there have been significant changes in the language. Today, there are two competing forms of the applicative II, *-en* and *–na*. The latter form is the form common in the lowlands around the Tengger area. Although there has not been a significant in-migration to the village of Ngadas, there has been to Ranu Pané. Further, as mentioned in the chapter on history, there are clearly greater contacts between all Tengger villages and lowland areas now, compared to 30 years ago. It is clear that there is a significant amount of grammatical interference in the speech of Tengger. It is also possible that in many ways the language is undergoing dialect shift or possibly drift. In my recordings from Ranu Pané, there is a clear preference for the lowland *–na* marker. Interestingly though, it has a different pronunciation that in the lowlands due to the lack vowel raising in Tengger. So where the lowland dialects have /*-nɔ*/ the Tengger have /*-na*/. In Ngadas both forms are

common. It is unclear whether or not this represents a stable situation, or whether some shift is taking place. But it is clear from Smith-Hefner that these patterns were never identified as salient dialect speech, so it is not surprising to see a good deal of interference from the lowland dialects here. In what follows, I will use the *–na* form, for the sole reason of keeping it separate from the passive imperative ending which is also *–en*.

22. Applicative II in Tengger

	Active	Passive I	Passive II
Tengger	-en/-na	-en/-na	Ø

The most frequent use of the applicative II is as a benefactive marker[94,95]:

23. a. …padha karo tuku rokok nem éwu kéné ya?

same with buy cigarettes six thousand here ya

'It's the same cigarettes as we buy for 6,000 Rp. here ya?'

c. Dhéké nukok-na isun ya paran?[96]

3 N-buy-na 1.f aff what

'What did he buy for me?'

[94] Beyond those listed here, there are a number of idiosyncratic uses of *–na*, e.g. *rungu* 'hear' *ng-rungok-na* 'listen'.

.

[96] Note the form of the verb here. For verbs that end in a vowel, a final glottal stop is inserted before the suffix, and the preceding vowel is lowered. Also, for roots ending in /n/, the /n/ becomes /?/.

b. Dhéké tuku paran ya nggo isun?

3 buy what aff for 1.f

'What did he buy for me?'

The first example above shows the transitive verb *tuku* 'buy' without any verbal morphology. Here the theme object comes directly after unmarked verb. In the second example, the verb appears with the applicative II marker, which automatically triggers the appearance of the *N-* prefix. Here we can clearly see the benefactive meaning of (23:b) which has the *–na* suffix, as opposed to the bare root in (23:a). Note also that in (23:b) for the bare verb stem to take an applicative it obligatorily takes either active or passive I verbal morphology. It is ungrammatical with an indicative reading to appear with any applicative ending if not also simultaneously overtly marked for voice.

Compare (23:b) with (23:c). In the latter, the verb again appears with no morphology. Here the benefactive goal is in an optional adjunct PP phrase and the theme argument is adjacent to the verb. In a typical analysis of applicatives, the change from (23:c) to (23:b) the benefactive goal adjunct is promoted and becomes an argument of the verb. Note however that there is not strict adjacency in the benefactive construction between the two objects, the *wh-* patient 'what' here appears after the affirmative marker *ya*.

Benefactive applicatives in many languages appear obligatorily adjacent to the verb which carries the applicative marker (see Bresnan & Moshi,1993 on Bantu). However, not only is there no strict adjacency between the two objects, it is possible for the benefactive goal to be missing completely:

24. a. Dadi godhong-é jati iku éyang ng-ungkab-na…

so leaf-é teak that 1.m N-open-na

'So I opened the teak leaf *for him*…'

[talking about assisting his son]

In the above example, the verb appears with the applicative II marker. The patient argument appears in initial focus position. However, there is no expressed benefactive goal. From the meaning of the sentence, it is clear that a benefactive interpretation is intended. But what is the status of the benefactive goal? As mentioned Tengger allows, in many cases prefers phonologically (and possibly syntactically) null constituents. Semantically, it is clear that there is a benefactive goal here, and the occurrence of the applicative II marker on the verb confirms this.

It is also possible for the verb to appear with the benefactive marker and for the goal argument to still appear in an adjunct PP phrase:

25. ndhèk kéné wis akèh anggrèk kembang, iku di-pèk-en nggo Pak Tris

prep here PFT many orchid flower that di-gather-en for Pak Tris

'There are lots of flowering orchids here, we picked them for Pak Tris.'

[Pak Tris worked with the local conservation board]

The verb here is marked with the applicative II ending, and it clearly has a benefactive interpretation. However, the benefactive goal here appears in an adjunct PP phrase. This is unexpected if the function of the applicative ending is to license an additional

argument of the verb or to promote the oblique to a core position. So, what exactly is the function of the applicative II? Before I begin to answer that, we must present the other uses of the applicative II marker.

In addition to the benefactive, there is also a causative use of the applicative II. Cross linguistically, causatives tend to behave differently from applicatives, and have unique morphology. In Tengger, however, both applicative I and II can have a causative reading.

26. a. Soal-é di-penting-na iku, upacara Karo.
matter-é di-important-na that, ritual Karo
'The thing is that the Karo festival is given priority.'
[discussing the agenda for a village meeting]

b. Maringono ng-lebok-na nang omah
after N-enter-na to house
'Afterwards I put (it) in the house.'

In these set of examples, the causative meaning of the *–na* suffix is clear. From the root *penting* 'important', the applicative ending gives the causative meaning 'to prioritize, to make important.' The predicate has voice morphology in both examples. In (26a) the verb takes the passive I marker; however the patient appears linearly after the verb both the demonstrative *iku* and then the co-indexed *upacara Karo*.[97] In (26b) the base verb is *lebu* 'enter'. This is an interesting case. It would appear at first glance

[97] In a derivational analysis, we might claim that the verb has raised to a focus position, which must be below CP which is filled by *soalé*, and must be TP or FocP.

as if the applicative marker here actually serves simultaneously to transitivize the intransitive root, and to add a causative interpretation. This is not the case, however.

In example (26:b) above the root is prefixed with the active nasal marker, in contrast to the intransitive *M-* suffix. The nasal marker here is marking the predicate as active. The *–na* suffix gives the verb a causative interpretation. I argue that the applicative suffix does not affect the syntactic structure, that is, the argument structure of the verb. Rather, the applicative here only affects the semantic structure, in adding a sense of causative (or benefactive for other examples). It would appear then, to be the collocation of both voice morphology and the applicative marker which renders the intransitive verb transitive. This is not a circumfix, per se—although genuine circumfixes do occur and are productive in the language, but rather the co-occurrence of two verbal affixes to overtly mark an emergent effect. This is not the whole story however.

Further evidence for claiming that the applicative markers do affect the syntactic structure of verbs comes from the following examples:

27. a. Aku sih m-lebu wong mbiyèn…

1 prt M-enter person before

'I once brought someone in the past…'

b. …ng-gawa petinggi rong botol, trus éyang ng-gawa sa'botol.

N-carry chief two bottle then 1.m N-carry one-bottle

'He brought the chief two bottles, and I brought one.'

In (27:a) above we have the intransitive form of the verb *lebu* 'enter'. However, in this sentence it is clearly transitive, as it is followed by a patient argument *wong* 'person'. If the applicative markers encoded transitivity, that is, if its function was to license an additional argument of the verb, we would expect the verb to be *ng-lebok-na*, with the active nasal prefix and causative/transitive *–na* suffix. However, this is not what we get. Further, if it were simply the *N-* nasal prefix and not the *M-* prefix which encoded transitivity, we would expect minimally to get *ng-lebu* in (27:a). Again, this is not what we find.

Similarly, in (27:b) the verb appears in its unmarked transitive form, with regular nasal prefixation. However, the meaning of the sentence is benefactive. Compare (27:b) above with the following:

28. a. Kudu ng-gawèk-na Marsam ondhé-ondhé.
must N-make-NA Marsam k.o.rice.cake
'I must make ondhé-ondhé for Marsam.'

b. …sing gawé cathet nggo pak Tinggi.
rel make notes for Mr. Chief
'(he's the one) who takes notes for the chief.'

In (28:a) we see the verb *gawé* 'make, do, use' with nasal prefixation and the benefactive goal has appears immediately following the verb. However, in (28:b), where there is no applicative ending (and no nasal prefix, which we discuss above), the patient or theme

argument immediately follows the verb, and the benefactive goal is reduced to an oblique, following the preposition *nggo* 'for'.

So, how can we make sense of (28: a & b) above? In both of these cases, the verb appears with no applicative ending. Example (28:a) 'should be' an intransitive, and yet there is a very clear patient argument immediately following the verb. Similarly, (28:b) 'should be' a simple transitive, yet it is clearly a ditransitive, with a benefactive goal/recipient and patient object, neither licensed by a preposition.

In one sense, it is clear that applicatives in Tengger are not necessary to, or rather do not alter the syntactic structure the verb to which they are attached. This is also supported by the form of the applicative in passive II constructions.

When the active equivalent of a passive II verb is marked by the applicative II, the passive II is unmarked, that is, it takes no suffix:

29. a. Kenthang-é kudu di-pendhem-na njero…
potato-é must di-underground-na N-inside
'Potatoes must be put in deep.'

b. Kabèh sing ke-pendhem iku gampang… soal-é
all REL ke-underground that easy matter-é
'All the ones put underground are easy… you see…'
[talking about how the ones not yet planted/in the ground can get easily hurt or damaged]

Here we have a non-verbal root. In order for it to become verbal, it must take active or passive I verbal morphology and obligatorily the applicative II: **mendhem* and **di-pendhem* are both ungrammatical. However with the passive II the applicative is not only not obligatory, it is ungrammatical: **ke-pendhem-na*. Neither of the examples has an expressed agent, and the second example appears in a relative clause. The *ke-pendhem* element in the second example is clearly verbal and passive, it is marked with the passive II marker. However, its underlying root is not verbal. In the active and the passive I, the applicative is obligatory. Again, this suggests that the applicative marker is not actually affecting the argument structure of the verb. It is not making the substantive transitive or verbal, in that sense that *ke-pendhem* is still verbal although is does not have the applicative ending. So, what is the function of the applicative endings, both I and II?

Before we move on to try to answer that question, we need to discuss how the applicatives interact with both voice and mood markers in Tengger. We will also need to consider how these applicatives operate in related dialects of Javanese. But before that, a brief comparison of the two applicatives will be useful.

6.5.3 Applicatives Compared

Although there are some broad semantic categories into which the two applicatives fall, I have also raised a question as to what their real functions are. It will be useful now to briefly compare the two to show that they are indeed distinct in their usage, beyond the number of idiosyncratic usages, several of which have already been mentioned.

It is possible for the *–i* and *–na* endings both to indicate some type of causality. However, when a word can appear with either ending, there is still a sometimes very subtle distinction to be made. Consider the following[98]:

30. a. Nah, kursi-né iku sing di-parèk-i.
prt chair-é that rel di-close-i
'Well, that chair should be the closer one.'
[asked which of several chairs should be moved closer]

a. Cèrèt-é mesti di-parèk-na.
earthen.pot must di-close-na
'The pot must be moved closer (to the fire).'

The contrast between the two is indeed subtle. Applicative I indicates causality in relative terms, that is, the chair must be moved closer *than* the other chairs. Applicative II on the other hand, indicates causality in absolute terms, the pot must be moved closer, full stop.

It is not common among the world's languages to have two distinct applicative suffixes with broad and generalized semantic functions. Another language which does is Standard Indonesian, which has the *–i* and *–kan* endings which are roughly analogous to applicative I and applicative II in Tengger. However, in many of the colloquial varieties of Indonesian, these are reduced to a single form. For example, in Jakartan Indonesian, all of the various roles described above for the two applicatives in Tengger are subsumed

[98] As these contrasts are subtle, and I did not record any naturalistic examples of them, this is elicited data, based on the descriptions in Robson (1992:61), and altered appropriately for the dialect.

by the marker *–in.* In this light, it is interesting to note that I have comparatively few examples of applicative I in my recordings, and also heard it infrequently while conducting fieldwork. This is contrasted with the frequency of the applicative II. It appears as if Tengger is gradually moving toward a single applicative ending. There is further, even recent historical evidence to support this claim, which I will discuss shortly.

Further evidence that the language is undergoing shift comes from the following examples. The following comes a story a farmer was telling about spraying his crops with insecticide. He uses the same verb in three consecutive utterances, each with the same meaning. Note carefully the form of the verb:

31. a. Dadi iku tak-semprot-i ng-anggo wangi-wangi-an waé.
so that 1pass-spray-i N-use RED.fragrance only
'So I just sprayed it with the smelly one (insecticide).'

b. éyang ésuk tangi, iku tak-semprot kosong-an ng-anggo banyu biasa
1.m morning wake.up that 1pass.spray empty-an N-use water normal
'I woke the next morning and I sprayed it with plain water.'

c. Ng-anggo banyu tawar iku tak-semprot-en, anu-né, lugur kabèh iku wis
N-use water normal that 1pass-spray-en whatchmacallit-é fall.off all that prt

'I sprayed it with plain water and all the whatchmacallits just fell right off.' [talking of pests on his crops][99]

In all three examples, he is talking about spraying his fields. The fields are here represented with the demonstrative pronoun *iku* 'that'. In each of the tree examples, then, the patient or theme argument here precedes the verb. All three examples are in the first person passive, marked with the *tak-* enclitic. This indicates the first person agent. Each of the three examples have an instrumental adjunct phrase introduced by the verb/preposition *nganggo* 'use'. In the first two examples, the instrumental is in final position after the verb. In the last example it appears in initial position for focus. An alternative translation for the last example could be 'It was with just plain water that I sprayed it…' which captures more the focus of the instrumental. The second example above has a medial adverb *kosong-an*, taking narrow scope of the preceding verb. The only other differences among the three example is the verbal morphology. In the first example, the verb takes the applicative I ending. There is no ending in the second example. And the third example takes the applicative II ending. Yet all three verbs are seemingly synonymous. In argument structure, thematic structure, and even linear surface order, these utterances are exactly parallel. This raises several very serious questions. Here I wish to address two questions, the first is how are the two applicatives distinct? These examples suggest either that the distinction which had existed between them is weakening and they are converging to some degree, or that the distinction between them was not as significant as once thought. Given the arguments above

[99] Note that this is one of the rare examples of the first person passive in Tengger. This particular speaker frequently traveled to the lowland marker, where he had a lot of interaction with lowland speakers.

comparing the causative aspects of the two applicatives, where they demonstrate a very clear, if subtle, distinction, it appears that the present examples are just further evidence that Tengger is in the process of collapsing the usage of the two applicatives into a single applicative, as has happened to many other languages in the region.

Of course the major question that this small data set raises is what the function of the applicatives is in general.

6.6 Verbal Suffixes: Clitic or Affix?

Based on evidence from reduplication, vowel mutations, and a particular ludling, I have argued above that the active verbal markers *N-* and *M-*, and the passive marker *ke-* are affixes whereas the passive markers (and the propositive) *éyang-, sira-,* and *di-* are in fact clitics. Where do the applicatives and other verbal endings fall into this classification?

Unfortunately the evidence for the applicatives is not as strong as for the other affixes. Nonetheless, it is clear from the vowel mutation effects that are caused by the applicative endings on the root that there is a tight bond between the root and the applicative. This clearly points to applicatives as affixes and not clitics. Consider the following:

32.	a.	m-lebu ‘enter’	ng-lebon-i ‘enter s.t.’	ng-lebok-na ’put in’
	b.	mati ‘die’	matèn-i ‘kill’	matèk-na ’murder’

The mutations caused by the addition of the applicative are not limited to vowel mutations. In both (32:a & b) there is an epenthetic /n/ in the second forms and an epenthetic /k/ or glottal in the third forms. All of these forms in addition undergo vowel mutation.

There is also interesting evidence that the applicative endings are also affixes in the lowland dialects. Though this does not necessarily imply that the same is true Tengger, if they share a common origin, which they do, it is suggestive. Malang Javanese (as well as the standard and many other dialects), has vowel harmony which operates on /a/ in open syllables (as described above), and spreads to a preceding /a/ in an open syllable. In these cases, the /a/ is raised and rounded to /ɔ/. So we have examples like:

[Malang Javanese]

33. a. mata /mɔtɔ/ 'eye' b. mangga /mɔŋgɔ/ 'please'

c. tamba /tambɔ/ 'medicine' d. tata /tɔtɔ/ 'arrangement, order'

However, when an applicative is added to root that ends in /a/ the vowel harmony is blocked:

34. a. namban-i /nambani/ 'treat' [Malang Jav]

b. natak-na /nataʔnɔ/ 'put in order'

That vowel harmony is blocked demonstrates that there is a very close prosodic bond between the applicative and the root.

Finally there is evidence from reduplication. Here however, the evidence points to the applicatives as clitics. Again here there are at least two distinct forms of reduplication. Lexical reduplication can be either where the reduplicated lexeme has an idiosyncratic meaning vis-à-vis its non-reduplicated form, or an idiosyncratic reduplicated lexeme that does not stand in opposition to a non-reduplicated form. However here I am talking about grammatical reduplication, which has a wide scope of semantic meanings, but is more systematic and in general marks some grammatical function such as plural, scope, iterativity, intensity, etc. In almost all cases where a verb has applicative ending and is reduplicated the applicative is not reduplicated:

35. a. kudu di-aran-aran-i. *kudu di-aran-i-aran-i
must di-RED-name-i
'You've got to give it a name.' [of a newborn]

b. kok di-larang-larang-na?
prt di-RED-expensive-na
'Why are you making it so expensive?'

On first glance, it would appear that these examples point to applicatives as clitics, in that they are not sufficiently close to the root to undergo reduplication. However, it is also possible that reduplication simply precedes affixation. This could be true in the

language in general, or specifically relating to the reduplication of verbs.[100] However, we have already seen that the nasal prefix *N-* does in fact get reduplicated, *ng-omong-ng-omong* 'by the way' from the root *omong* 'speak'. I leave the question open here.

With the evidence from lowland dialects where the applicative blocks vowel raising and harmony, coupled with the evidence from vowel mutation and epenthesis that are caused by the addition of applicative endings, it seems safe to conclude that applicative endings are affixes.

The case with the other verbal endings, however, is not easy. The active imperative and the subjunctive ending are homophonous, and they have the same behavior in terms of clitic or affixal properties. I will therefore treat them together here.

The optative *–a* marker is unique among the verbal endings considered here in not causing vowel mutation or epenthesis when attached to a root that ends in vowel, but rather it is added directly to the unaltered root:

36. a. Lé, teka-a saiki!
boy come-a now
'Come on home now son.'

b. …ora gelem mati-a!
neg want N-die-a
'I don't want for him to die.'

[100] If the applicatives are genuine affixes, as I claim here, then the facts from nominal reduplication would suggest that there is a particular morphological order of processes which is distinct for nouns and for verbs. In nominal reduplications with circumfixes, which are clearly affixal in nature, the circumfix is reduplicated with the root.

The optative is also unique in being able to take almost any word as its root. Again, in these cases, no vowel mutation or epenthesis takes place:

37. a Kéné-a!
here-a
'Come here!'

b. Iku-a!
that-a
'(buy) that one!'

In these examples a directional pronoun and a demonstrative pronoun serve as the base for the imperative/subjunctive. It is interesting to note, that these bases cannot appear with any other verbal morphology, there is no form **di-kéné-na* or *ng-iku-i*.

In the lowland dialects with vowel raising/rounding and vowel harmony, the imperative/subjunctive marker does not block either process. So we get examples like *teka-a /təkɔ-ɔ/* 'come on' and *mata-a /mɔtɔ-ɔ/* 'look!' where there is both vowel raising/rounding and subsequent harmony.

Normal patterns of vowel harmony only affect the final syllable and the penultimate syllable. There seems to be a defined foot over which harmony spreads. Gil has made similar claims about a core prosodic foot in Riau Indonesian (p.c., 2007 Manokwari paper). In those few words with three syllables, each with a coda-less /a/ in final position, and no blocking consonant clusters intervening, harmony only affects the last two syllables. For example *katara* 'evident' is pronounced /katɔrɔ/ and not */kɔtɔrɔ/

as we might expect if vowel harmony were not limited by a prosodic boundary, here a foot. However, as we saw above, when the imperative/subjunctive is added to disyllabic roots with final and penultimate open /a/, such as *mata* 'eye', vowel harmony still operates on the penultimate: /mɔtɔ-ɔ/ and not */matɔ-ɔ/. This is very strong evidence suggesting that the imperative/subjunctive ending does not enter into a close prosodic relationship with the root.

Interestingly, the imperative/subjunctive would also appear to be subject to vowel raising/rounding here. This would point to the ending perhaps being an affix and not a clitic, in that the normal phonological processes which act on the word level are also here affecting the ending. However, it is more likely that the underlying form of the ending itself is actually -ɔ in the lowland dialects. As no further suffixation or cliticization is allowed after the use of the imperative/subjunctive, the form of the ending is invariably -ɔ. This contrasts with the other cases under discussion here, where there is a clear alternation between a/ɔ depending on the phonological climate of the word. One argument against the underlying form of the imperative/subjunctive as -ɔ in some lowland dialects is the uniform distribution of vowel raising/rounding and imperative/subjunctive -ɔ in dialects. That is, every dialect which has the phonological phenomenon of vowel raising/rounding also has the form of the imperative/subjunctive as –ɔ. If the actual underlying form were different, we might expect to find a dialect which lacks vowel raising/rounding, or alternatively, vice versa. To my knowledge, there is no dialect of Javanese which has such a combination. I leave the question open

then, as to how strong this particular argument is to treat the imperative/subjunctive as a clitic and not an affix. There is yet further evidence to be considered.

In reduplicated items taking the imperative/subjunctive, the *–a* marker appears only once at the end of the reduplicated element; it does not get reduplicated itself:

38. a. Gundhul-gundhul-a!
RED-bald-a
'Just shave it all off!'

c. Bosen-bosen-a ngko lek wis mari lunga…
RED-bored-a later if PFCT finish go
'Even though we're bored we'll go after it's over.'

In the above examples, the imperative/subjunctive ending attaches to the end of the reduplicated word, and does not get reduplicated.

Here we have adduced quite significant evidence demonstrating that the imperative/subjunctive marker does not form a close prosodic bond with its root: lack of vowel mutation, epenthesis, not blocking vowel raising or harmony, and reduplication facts. The question then is begged, is this a clitic or free particle?

That there is some prosodic bond between the imperative and its root is clear from the fact that the clitic must immediately follow the root, there can be no intervening material. However, given that the clitic can attach to almost any root, it is sometimes difficult to show. The clitic cannot in general attach to functional words, like *ngko* in the example below:

39. a. *bosen-bosen lek-a wis mari lunga…

RED-bored if-a PFCT finish go

'Even though we're bored we'll go when it's over.'

Of course with a different meaning, the marker could appear on another word in the sentence. However, it is clearly ungrammatical attached to functional element, and for the given reading it must be attached to *bosen-bosen*. So the imperative/subjunctive ending has very clear clitic properties.

6.7 Applicatives: Voice and Mood

The interaction of the applicative markers with the various voice and mood alternations in Tengger is at first glace straightforward. I have already discussed the interaction of the applicative markers with the passive II. As the passive II also does not occur in the other moods, I won't present it further here.

For the indicative, the applicative suffix is added directly to the root, which bears appropriate voice morphology, with appropriate phonological changes as noted above:

66. Tengger Indicative

	neutral	Applicative I	Applicative II
Active	n-jukuk	n-jukuk-i	n-jukuk-na
Passive I	di-jukuk	di-jukuk-i	di-jukuk-na
Passive II	ke-jukuk	ke-jukuk-i	ke-jukuk

The propositive only has an active form. The applicative affixes attach directly to the end of the root, which appears with a nasal prefix and the propositive clitic:

67. Tengger Propositive

	neutral	Applicative I	Applicative II
Active	(éyang) tak n-jukuk	(éyang) tak n-jukuk-i	(éyang) tak n-jukuk-na

There is a more complex process at work in the optative. When the applicative endings are added, the mood marker disappears. In other dialects of Javanese, there are unique forms the applicatives in different moods, as we shall see, but this is not the case in Tengger. This does not, however, mean that none of the forms are marked for mood. Consider first how they are formed. In the active forms of the optative, the applicative

suffixes are added directly to the root which bears the active *N-* prefix. For the passive, there are different forms:

68. Tengger Optative

		neutral	Applicative I	Applicative II
Active		n-jukuk-a	n-jukuk-i	n-jukuk-na
Passive I	2nd person	jukuk-en	jukuk-i	jukuk-na
	3rd person	di-jukuk-a	Ø	Ø
	unmarked	Ø	di-jukuk-i	di-jukuk-na

So, formally, in the active, the optative verb with either applicative is indistinguishable from an indicative.

The second person passive I, however, is marked in all cases by the lack of voice morphology, compare *jukuk-i* with *di-jukuk-i*. The rest of the paradigm is indistinguishable from the indicative.

There are again a number of interesting questions to be asked here. Why are there so many homophonous forms in the Tengger paradigm? In one way, I have tried to present the data here as clearly as possible. I have also tried to analyze the data on its own merit, and not merely compared against that of the standard language. This analysis has in fact kept to the smallest number of actually occurring distinct forms. In doing so, I have argued for only three distinct moods in Tengger, with a person distinction in the Optative. I also claim that there is no passive propositive, as it would be completely

indistinguishable from the first person passive. That there is a unique optative is beyond question, and it would be more parsimonious to claim simply that the optative cannot co-occur with applicative endings. However, the second person passive optative is in fact marked as optative by its lack of voice morphology, and it does appear with applicatives. Perhaps it is just that the active and the rest of the passive optative paradigm do not co-occur with applicatives. It will also depend on what the actual function of the applicative endings is, a discussion we are saving for just a bit later.

According to my analysis then, the Tengger verbal paradigm is much less complex than either that of the lowland or the Standard language. It will be useful to see how the verbal system operates in these other varieties.

6.2 Standard Javanese Verbal Paradigm

In the following description of the standard Javanese verbal system I rely largely on the work of Poedjosoedarmo (1979) and the many studies of E.M. Uhlenbeck, published collectively as *Studies in Javanese Morphology* (1978).[101]

Uhlenbeck provides the following classification of the Javanese verbal system, with minor modifications here (R = root; N = nasal prefix[102]):

[101] That only two proper analyses of the Javanese verbal system exist, and that both were written over a quarter of a century ago is a telling mark of how marginalized many non-Western languages are in recent theoretical linguistics, especially as Javanese has some 90 million native speakers, making it the 12th largest language in the world, by such a measure.
[102] See section 3.2 on phonology.

41. Standard Javanese Verbal Paradigm, á la Uhlenbeck:

Neutral Form	Applicative I	Applicative II
1. N-R	N-R-i	N-R-ake
2. tak-N-R	tak-N-R-i	tak-N-R-ake
3. ke-R	ke-R-an	
4. di-R	di-R-i	di-R-ake
5. ka-R	ka-R-an	ka-R-ake
6. –in-R	-in-R-an	-in-R-ake
7. tak-R	tak-R-i	tak-R-ake
8. tak-R-e	tak-R-ane	tak-R-ne
9. kok-R	kok-R-i	kok-R-ake
10. R-en	R-ana	R-na
11. R-R-an		

Uhlenbeck (1978:119)

The three elements in each numbered row stand as a unified set, and, as Uhlenbeck notes 'the division in three vertical columns is justified by the existence of a continuous proportionality between the forms in one horizontal row.' For example, from the root *adus* 'bath', from row 1 the relation between *ng-adus, ng-adus-i,* and *ng-adus-ake* is parallel to the relation between *adus-en, adus-ana,* and *adus-na* from row 10.

A further opposition exists between the six elements in rows 1 and 2 and the rest of the chart. The former elements indicate action from the point of view of the actor, that is they are actor-focus (or active) whereas the remaining elements are 'not positively

considered from the standpoint of the actor.'; they are patient or theme focused (passives in many analyses).

From lines three and five above, we see that there are two distinct *k-* passives in Standard Javanese, whereas Tengger has one. In addition to the accidental/adversative passive marked with *ke-* just as in Tengger, Standard Javanese also has a passive marked with an invariable *ka-*. This form does not show any vowel mutation, even when prefixed to a root that begins with a vowel. While there are some fossilized examples of this form in Tengger, it is not productive, and it is not clear if it ever was productive. The *ka-* passive generally appears with no expressed agent, and it is found more often in *Krama* than in *Ngoko*, hence it is marked as formal. This is one reason why it may never have been productive in Tengger, though it did also appear in Old Javanese.

While Uhlenbeck's analysis of the standard Javanese verb is the most thorough, it is also a bit confusing. It does however capture an important point. There is a three way contrast in the verbal paradigm in Standard Javanese, as noted by the three columns in Uhlenbeck's analysis. Based on the descriptions of the verbal paradigms of Standard Javanese given in Robson (1992), Uhlenbeck (1978), Horne (1961), Keeling (1984) we can give the following, perhaps clearer visual representation:

44. Standard Javanese Verbal Paradigm

Mood	Voice	Prefix	Neutral	Applicative I	Applicative II
Indicative	Active	N-	Ø	-i	-aké
	Passive I	tak-/kok- /di-			
	Passive II	ke-		-an	Ø
Imperative	Active	N-	-a	-ana	-na
	Passive I	Ø	-en		
Propositive	Active	(aku) tak N-	Ø	-i	-aké
	Passive I	tak-	-é	-ané	-né
Subjunctive	Active	N-	-a	-ana	-na
	Passive I	tak-/kok- /di-	-en	-na	

Soebroto et.al. (1991) point out the decompositional elements of the verbal affixal system are still clear. According to them, the applicative I ending becomes *–an* in non-indicative moods. Similarly, the applicative II ending becomes *–n* in non-indicative moods. To these, are added the modal endings, *-a* for irrealis and *–é* for the propositive. This analysis, however, does not hold for the active propositive, as only the passive is uniformly marked with *–é*.

The most recent brief description of Standard Javanese is Ogloblin (2005). He also analyzes the standard language as having a person distinction, with the first and second persons opposed to the third person forms. I have also analyzed the Tengger forms as making a person distinction. He provides the following chart:

45. Mood Forms of Transitive Verbs (Non-Formal) [Standard Javanese]

Voice	Indicative	Irrealis	Imperative	Propositive
Active	n-jupuk	n-jupuk-a		tak n-jupuk
	n-jupuk-i	n-jupuk-an-a		tak n-jupuk-i
	n-jupuk-aké	n-jupuk-n-a		tak n-jupuk-aké
Passive ½	ø-jupuk	ø-jupuk-a	jupuk-en	ø-jupuk-é
	ø-jupuk-i	ø-jupuk-an-a	jupuk-an-a	ø-jupuk-an-é
	ø-jupuk-aké	ø-jupuk-n-a	jupuk-n-a	ø-jupuk-n-é
Passive 3	di-jupuk	di-jupuk-a		
	di-jupuk-i	di-jupuk-an-a		
	di-jupuk-aké	di-jupuk-n-a		

(Ogloblin 2005:600)

The chart nicely captures the decompositional nature of the applicative and mood suffixes. Note however, where it does not work fully for the propositive, as noted above. Note also that the base imperative form is not marked here with the *–a* ending, but the applicative forms are.

Ogloblin has also here treated as a single class the active irrealis/imperative that other authors have treated as separate categories. Here, the imperative has only a unique (presumably second person—Ogloblin does not specify) imperative form. Compare this above with the analysis presented for Tengger, where I also argue that the irrealis/imperative are a single category, what I call optative, though with a distinct second person form.

What is noteworthy when comparing the two systems, whether that provided by Uhlenbeck or that proposed by Ogloblin, is how much more fully articulated the verbal paradigm is in the Standard language as opposed to Tengger. Counting only the active and passive I forms, under Uhlenbeck's analysis the standard Javanese verb has 15 distinct forms, 27 for Ogloblin, and 18 for the 'composite' analysis presented in chart (45) above. In contrast, the Tengger verb only distinguishes between 10 different verbal forms. To begin with, the range of proposed forms for the standard languages is truly astounding, from 15 to 27. If we move outside of the Yogyakarta/Surakarta dialect area (in any direction!), the number of distinct forms also drops dramatically. Also, the difference between even the lowest number there and the Tengger system, 15 to ten, is a major divergence. The divergence on both parts, that is among the various analyses of the standard language on the one hand, and between the standard language and all other dialects of Javanese comes mostly from a distinct spoken versus literary, formal vs. informal differences. Given that Tengger is only spoken, and that there are no formal vs. informal distinctions overtly encoded in the language, it stands to reason that it would have fewer forms. However, even in Yogyakarta/Surakarta, most speakers do not use this many forms in daily conversation. Many of the forms given for the standard

language are literary, and to some extent part of the spoken language of those still involved with the courts and courtly culture. None of the scholars listed above mentions this fact, and though it clearly requires further research, the basis is sound. There is a bias in many studies of Javanese not only to look at only the Yogyakarta/Surakarta dialects—aberrant as they are from most dialects of Javanese—but also to look only at the most 'refined' speech. There is a tendency to overlook the fact that a Javanese, like almost all languages, has a register distinction. Of course, this is separate from the elaborate speech level system that so uniquely characterizes Javanese. Minimally though, the speech level system is an asymmetrical code, where speaker and interlocutor often use differing forms, depending upon social status. A register distinction, however, is symmetrical. It is code that varies according to the context, ritual, courtly, legal, academic, etc. and not according to the participants involved. Although I have not yet conducted in depth research in the Yogyakarta/Surakarta area, I would hypothesize that one factor explaining the huge variation in the verbal paradigm as described by different authors is in fact related to this register system. I would further venture that the most common, quotidian, and widespread speech to be found would have a system much closer to that of Tengger and not the pattern described for the standard language here. Unfortunately, we must leave the question open here and await future research, however, to substantiate the claim at least diachronically, it will be instructive to look at applicatives in Old Javanese.

6.3 Applicatives and Causatives in Old Javanese

Compared to the modern Javanese dialects, Old Javanese made more use of prefixes and infixes, rather than suffixes. To mark lexemes transitive, Old Javanese used the prefixes *pa-, paha-,* and *paka-* (have, use as), the latter two here are clearly compositional. There was a designated causative prefix *pi-*. Nonetheless, Old Javanese did have applicative constructions that made use of suffixes. According to Hunter (1999) there are two distinct applicative suffixes: *-i* and *–aken* both of which should be familiar to us by now. The *–i* ending has been maintained in both Tengger and almost all other Javanese dialects. The functions of the applicative I in Old Javanese are very similar to those described above for Tengger including applicative, extensive, and causative, among others. If attached to a root ending in a vowel, the form became *–ani*. There was also a distinct passive form *–an.*

The second applicative ending in Old Javanese *–aken* became *–aké* in the standard language. It is likely that the Tengger form in *–en* is a shortening of this form, with the initial syllable of the suffix lost. The form common in lowland East Javanese dialects, as well as in many others including the Banyumas and Pesisir Lor dialects, *-na* is a little trickier to reconstruct. Remember above the compositionality of the applicative paradigm in standard Javanese, the *–n-* element represents the applicative II. This is perhaps derived from the final /n/ of OJ *–aken.* The /a/ element may perhaps be the irrealis marker. If this is the case, then *–na* originally marked the applicative II in imperative and irrealis verbs, and was subsequently generalized to the indicative.

In Old Javanese the applicative II also had a wide range of uses, but especially salient are as benefactive, transitivizer, and causative.

A simple example will suffice here:

[Old Javanese]

46. a mang-adeg ta sira t-um-inggal-aken pa-turw-an-(n)ira
AF2-stand top pro3 AF2-leave-behind-Tr1 AF1nom-sleep-LF-pro3

'He stood up, leaving behind his sleeping place.'

b. in-adeg-aken ta Sang Hyang Teas
GF1-stand-Tr1 top title title border.stone

'...the sacred border stone was set up.'

[Hunter 1999:19]

In (46a) the verb is intransitive meaning 'to stand up'. With the applicative II in (46:b), here marking the verb as transitive, the meaning becomes 'to set/stand s.t. up.'

The two applicatives were very clearly distinct, both semantically and syntactically. For ditransitive verbs like *tarima* 'receive' *-aken* promotes the theme argument. In these cases, the *–i/-an* applicative I targets the recipient, promoting it to a focus position (remember that OJ had a focus system similar to that of the modern Philippine languages, and unlike modern Javanese). The contrast is seen in (47:a) and (47:b) below:

[Old Javanese]

47. a. Sang Ambika t-in-arima-(a)ken i Sang Citranggada
pro3f-lig Ambika GF1-receive-Tr1 dat pro3f-lig Citranggada
'Lady Ambika was given (in marriage) to Lord Citranggada.'

b. ri kala ni-ng wiwaha sedeng
dat time obl-lig marriage mod-prog
rahadyan sanghulun t-in-arima-(a)n Sang Dewayani
my noble lord" (you) GF1-receive-LF title Dewayani
'...at the time of the wedding, when you, my Lord, were given to Lady Dewaya.'

[Hunter 1999:20]

Both of the verbs in the above examples are passive (goal focus in Hunter's terminology)—hence the applicative I appears in the passive form *–an*. The contrast between the two is clear. In (47:a) the theme is focused, and in (47:b) it is the recipient that is focused. Note that focus here must be referring to morphological focus, syntactically the sentences are roughly equivalent.

To these two forms, Ogloblin adds an applicative/benefactive circumfix *pa(N)-aken*, which he notes also has a causative function, so we have forms such as *–gi:ta* 'song' [clearly borrowed from Sanskrit] > *pang-gi:ta-aken* 'to sing for s.o.'. The applicative marker here serves both to transitive a substantive base and to give it a benefactive interpretation. However, it is not clear that these two roles are in fact the result of the circumfix. Again here, the marker *paN-aken* is clearly compositional.

There are other examples where the only the *pa-* prefix or the *–aken* suffix serve to transitivize a non-transitive root, or to add causative or benefactive semantics:

[Old Javanese]

48.	a.	a-hayu	'good'	pa-hayu	'put in order'
	b.	wijil	'go out'	wijil-aken	'produce'

Both of the above examples are transitive, both have a causative interpretation, 'to make good' i.e. 'to put in order', and 'to make go out' i.e. 'produce something.' It is unclear to me in what way Ogloblin intends the circumfix as a unitary applicative marker. I have argued elsewhere (Conners 2002) that Modern Indonesian and Javanese have genuine circumfixes, which represent unitary morphological processes, so I am amenable to the suggestion that Old Javanese also had productive circumfixes. However, given the clear decompositionality of the independent elements of this particular circumfix in these particular constructions, I am not convinced that it does represent a unitary morphological element.

There are two typological notes to stress here. Cross-linguistically, those languages with applicatives very strongly tend to have a single applicative affix. Further, the causative is almost universally distinct from the applicative. Standard and Old Javanese, and Tengger have two very distinct and productive applicative constructions. Further, in addition to the normal range of applicative semantics—benefactive, transitive, locative, etc.—both allow for a causative interpretation. Across the western Austronesian languages, most have a distinct causative formation involving the prefix

pa, as in Old Javanese.[103] With the exception of the Philippine-type languages, most western Austronesian languages also distinguish two distinct applicative affixes: a locative formed with *–i* (or a cognate) and another with the form *–akan* (or a cognate). This latter applicative, applicative II in our discussion, often has a wider range of semantic roles. Himmelmann points out that :

> Although the morphosyntax is quite different, there is no doubt that the conveyance and locative voice alternation in the Philippine-type languages have much in common semantically with applicative alternations in other languages. There is furthermore a formal similarity in that the most widely attested applicative formative *–i* is also widely attested as (subjunctive) locative voice suffix.
>
> (2005:170)

This makes sense in the Javanese case. As the language changed from Philippine-type language, to which Old Javanese was closer, the focus prefix system was lost and replaced with an applicative suffixal system. Tengger then, is not typologically odd in having two distinct applicative constructions, at least not when compared with other western Austronesian languages.

Still compared cross-linguistically it is remarkable. Further, even when compared with other Austronesian languages, the fact that Tengger applicatives also have causative semantics is exceptionally rare cross-linguistically. Why should the Javanese applicatives behave so aberrantly typologically? We discuss that presently.

[103] There are exceptions, including many Malay varieties. Though many Indonesian varieties do make use of a causative *per-* prefix, though generally in combination with some other marker.

6.4 Summary

So far in this chapter, I have presented the non-voice marking verbal morphology in Tengger. Synchronically I have compared it with both the lowland East Javanese dialect and the Central Javanese standard dialect. Further, I have presented the facts from Old Javanese to allow for diachronic comparison.

Tengger marks three distinct moods, indicative, propositive, and optative, the last marked also for person. Except for the propositive, mood is marked by suffixation or encliticization. Further, Tengger has two unique applicatives, I & II. In some ways, these seem to behave like 'normal' applicatives from e.g. the Bantu languages, they mark transitivity, license locative goals, and promote benefactive recipients., affecting both the semantic meaning of a sentence, and the syntactic structure in changing both the linear order of constituents but also the underlying argument structure. However, I have also presented examples where verbs were transitivized, locative goals were licensed (without a preposition), and benefactive recipients where promoted, all without the accompanying verb having any (non-voice) morphology. Further, I have presented examples of verbs with applicative endings, yet having no expressed arguments. This all questions the role of the applicative function. I have hinted above at the solution to these problems, and I present more thoroughly the argument below. Basically, I claim that the applicative construction does not have any direct effect of the syntactic structure of the verb to which it is attached. It does not affect the argument structure of the verb. Rather, it is an

(optional) overt marker of a semantic relationship between the verb and its arguments or between arguments themselves.

6.5 Analysis

To my knowledge, there are no analyses of the applicative constructions in any dialect of Javanese. To be sure, there are several descriptions, some of which I've presented above. In the recent literature there have been several studies of the *–kan* marker in Indonesian and related dialects. The *–kan* marker is analogous to the applicative II in Tengger.[104] I present in some depth two of those analyses, as they contain valuable insights into the Tengger case, Gil (2002) discussing Riau Indonesian, and Son and Cole (2008) on standard Indonesian.

Son and Cole (2008) provide an event-based, semantic analysis of acrolectal standard Indonesian. Ultimately, they claim that *–kan* is a morphological reflex of the RESULT head, which yields causative interpretations. Further, they show that there is 'a direct mapping between semantic decomposition of predicates and the (morpho-)syntax.' There are two immediately interesting things to note here. First, both Gil and Son & Cole zero in on the semantics of the marker in order to provide a unified analysis. For Gil, it is 'end-point' theme, for Son & Cole RESULT, but is it clear that the two are analogous. Second, the syntactic claim in Son and Cole (2008) is weaker than that in Cole and Son (2004), namely they do not claim in the later paper that the *–kan* marker licenses an argument in the argument structure that is not licenses inherently by the verb. I briefly review their analysis here.

[104] It is cognate with the form in standard Javanese *–aké*.

Son and Cole (2008) summarize four distinct uses of *–kan* in standard Indonesian: causative, benefactive, goal—PP, and inherently ditransitive constructions:[105]

[Indonesian]

49: a. Cangkir-nya pecah. b. Janet me-mecah-*(kan) cangkir-nya.
mug-3 break Janet meN-break-kan mug-3
'The mug broke.' 'Janet broke her cup.'

50. a. Tika me-manggang-kan roti itu (untuk Erik).
Tika meN-bake-kan bread that for Erik
'Tika baked the bread for Erik.'

b. Tika me-manggang-*(kan) Erik roti itu.
Tika meN-bake-kan Erik bread that
'Tika baked Erik the bread.'

51. a. Dia meng-ikat tali itu.
3.sg meN-tie rope that
'He tied the rope.'

b. Dia meng-ikat-(kan) tali itu ke anjing.
3.sg meN-tie-kan rope that to dog
'He tied the rope to the dog. (Sneddon 1996)

[105] Unless otherwise noted, the following examples are taken from Son and Cole (2008), with only minor adjustments.

52. a. Dia me-nyerah-*(kan) pekerjaan itu kepada saya.

3.sg meN-entrust-kan job that to 1.sg

'He entrusted the job to me.' (Dardjowidjojo 1967)

b. Pekerjaan itu di-serah-kan kepada saya (oleh dia).

job that di-entrust-kan to 1.sg by 3.sg

'The job was entrusted to me (by him).'

c. *Saya di-serah-kan pekerjaan itu (oleh dia).

1.sg di-entrust-kan job that (by 3.sg)

'I was entrusted the job (by him).'

Son and Cole point out that in the first two set of examples above, the causative and the benefactive, the difference between the (a) examples and the (b) examples involves changes in the actual argument structure of the verb. (49:a) is an inchoative verb (adjectival predicate) with only one argument. However, in (49:b), with the *–kan* affix, it becomes causative and both an agent and a theme argument are licensed. Similarly in (50:a) the benefactive recipient appears in an optional adjunct PP *untuk Erik*. However when *–kan* is added to the verb, the benefactive recipient is promoted to an argument and appears linearly immediately following the verb, which now has three overt arguments.

So the first set of examples here would support past analyses by e.g. Hopper and Thomson (1980), Arka (1992), Sie (1989), Postman (2002) which claim that *–kan* is either a transitivizing or valency changing morpheme. That is it interacts syntactically with verb to which is attaches to add an extra argument to its argument structure. Son

and Cole observe, however, that this is inconsistent with the facts from goal-PP and inherently ditransitive constructions, as given in the second set of examples above.

For example, in (51:a) above, the verb has two overt arguments, an agent subject and a theme object. When *–kan* is added to the verb, there are still only two arguments, the same agent subject and theme object. Here a locative goal appears as an adjunct in a prepositional phrase. The underlying argument structure of the verb *ikat* has not been altered. Note also, that the *–kan* itself is optional here, the omission of which neither renders the sentence ungrammatical nor alters the interpretation.

Similarly, in inherently ditransitive verbs the suffix does affect the underlying argument structure of the verb. Further, it can be seen above that theme argument in the primary object, and not goal argument, as the theme can be passivized but the goal cannot (52:b vs. 52:c).

So of the four cases of *-kan* that they analyze, there are two where it seems to be affecting the underlying argument structure of the verb, and two where it clearly does not affect the argument structure. So far, this is very similar to the case for Tengger presented above. In fact, we have seen that Tengger has each of these types of uses of the applicative II.

They unify the various uses of *–kan* claiming that it is not accidental homophony, but rather the four constructions all share the same event structure, and the *–kan* marker is the 'morphological reflex of the aspectual component, i.e.., a result-state, that is shared by those constructions (2008:30).' They rely on a event decomposition analysis to explain both the unity of the structure and the diversity of its uses, following much recent

work arguing for a direct mapping the event structure of lexical/logical semantics into syntax.

They claim there that 'a now widely held position in the literature on verbal meaning is that the lexical semantic representation of verbs involves complex event structure, with semantic primitives like DO, CAUSE, and BECOME; events are not unanalyzable atomic units, but are decomposable into subparts of events.' Following this line of argument, events are broken further into inner and outer events, where inner events are associated with telicity and change of state, and outer events are associated with agency and causation. Under such an analysis, the semantic structure of the verb 'break', decomposed into inner and outer events, means something like 'x causes y to change into a state of becoming broken.'

53. [x CAUSE [y BECOME [$_{AP}$ broken]]]

\ / \ /

Outer Inner Event

However, given that Indonesian lacks unaccusative causatives, crucial evidence demonstrating the existence of a CAUSE (projection) separate from Voice is lacking, the *–kan* causatives are not projected from CAUSE. Rather, they propose that *–kan* is an instantiation of a Result Phrase projection which is generated very low in the VP.

Son and Cole then discuss in some detail two specific constructions where the addition of *–kan* in standard Indonesian does not affect the argument structure of the verb: goal PP (54) and inherent ditransitive (55) constructions:

[Indonesian]

54. a. Dia me-nempel gambar itu.
3.sg meN-patch picture that
'He glued the picture.'

b. Dia me-nempel-(kan) gambar itu ke tembok.
3.sg meN-patch-KAN picture that to wall
'He stuck pictures on the wall.' (Sneddon 1996)

55. a. John mem-beri-*(kan) surat itu kepada Peter.
John meN-give-KAN letter that to Peter
'John gave a letter to Peter.'

b. John mem-beri-*(kan) Peter surat itu.
John meN-give-KAN Peter letter that
'John gave Peter that letter.'

In (54:a) the goal of the action is not expressed. However, as it clear from (54:b) it can be expressed optionally in a PP adjunct phrase. In (54:b), in fact, the applicative itself is not obligatory, the agent and patient arguments appear without the addition of the applicative ending, and the goal is still optional, and in a PP. The effect of the applicative here is purely semantic, there is no overt affect on the syntax.

Similarly in (55) is an inherent ditransitive verb. In (55:a) the theme argument appears immediately following the verb and the recipient *Peter* is in an adjunct PP. In (55b) there is a double object construction, and the recipient is in an argument position

immediately following the verb. Crucially, here the applicative is obligatory, whether the recipient is in an argument position or appears in an adjunct PP. The applicative does not affect the overt syntax of the clause.

In both (54) and (55) the effect of the applicative is semantic. For Son & Cole, that translates into an additional syntactic projection as well, the Result Phrase. They claim that '…the four aforementioned constructions…all involve the same aspectual components of a causing event and a result state, which correspond to separate verbal projections in the syntax (2008:45).' They further propose 'that the verbal head projecting a result-state-encoding constituent, i.e., a Result head, is explicitly expressed by the suffix *–kan* in Indonesian (45).' Crucial evidence for this comes from sentences with unexpressed constituents:

56. a. Roti itu di-panggan-kan. [Indonesian]
bread that DI-bake-KAN
'The bread was baked ***for someone.***'

b. Rumah-rumah-an itu di-buat-kan.
RED-house-an that DI-make-KAN
The toy house was made ***for someone.***'

The verbs in both of these examples appear with the applicative ending. What is crucial here, although no goal or recipient is expressed, the only possible interpretation for these examples is benefactive. These both have a null benefactive interpretation. For Son and Cole, then, the benefactive reading indicates the existence of the Result Phrase, and the –

kan is the overt expression of the head of that phrase. For them, this is crucial evidence of the interaction between semantics and syntax, in fact they posit that this demonstrates that the syntax must have access at some level of representation directly to the argument structure of the verb, that is, the semantic frame of the verb.

Some of the facts from Tengger, while not completely congruous with those from standard Indonesian, show similar patterns:

57. a. Lawang-é di-inep-na.
door-É DI-open-NA
'The door was opened ***for someone.***'

b. Iwak-é di-tukok-na.
fish-É DI-buy-NA
'The fish was bought ***for someone***.'

In both of the above examples, there is no expressed benefactive recipient. However, both of the sentences have a benefactive reading available. For the Indonesian, the benefactive interpretation is reported as being obligatory in such examples. However, in Tengger, the benefactive is available, and may even be the preferred reading, but it is not the only reading. These examples would suggest that Tengger is like Indonesian, and that even when the recipient argument is not syntactically expressed, it is still semantically present. However, the facts are not all the same. In (57) above we saw that a verb in Tengger that appears with the applicative marker, but with no expressed recipient, can still have a semantically benefactive interpretation. Crucially, the reverse

is also true in Tengger, that is, we have examples where we might expect to find an obligatory occurrence of an applicative marker, and yet none is expressed.

58. a. Maringono isun m-lebu.
afterwards 1.f M-enter
'After that I went in.'

b. Aku sih m-lebu wong mbiyèn…
1 prt M-enter person before
'I once brought someone in the past…'

M-lebu 'enter' in (58:a) is a prototypical intransitive verb. However, in (58:b) we have the same form of the verb, yet here it is clearly transitive 'bring'. If the applicative II in Tengger were an obligatory transitivizer, as Postman (2002) proposes for Indonesian –*kan*, we would expect to see it on the verb in (58:b), or rather we would expect (58:b) to be ungrammatical. Yet it is clearly grammatical.

This is also the same in Jakartan Indonesian:

59. a. Aku masuk nanti siang aja ya. [Jakartan Indonesian]
1.sg enter later noon only aff
'I'm going to just come in later this afternoon ok.

b. Mesti masuk-(in) motor di mana?
must enter-in motorcycle loc where
'Where should I put my motorcycle?'

As noted above, there is only one productive applicative marker in Jakartan Indonesian, *-in*, which takes all of the semantic roles of *–i* and *–kan* in the standard language. In (59:b) above the *–in* is optional, and does not change the grammaticality or the interpretation of the sentence. Clearly we can see that the facts from Jakartan Indonesian are the same as the Tengger, and both stand opposed to the standard language. *Masuk* is seemingly an unergative verb, yet in both of the (b) examples in the two sets of examples above, it has an overt object patient.

In Tengger, it is not only limited to the transitivizing function of the applicatives. We find similar examples with a benefactive interpretation, yet no expressed applicative marker:

60. a. …sing ng-gawa ilmu-né wong Tengger
REL N-carry science-é person Tengger
'…who carried the knowledge of the Tengger people.'

b. …ng-gawa petinggi rong botol, trus éyang ng-gawa sa'botol.
N-carry chief two bottle then 1.m N-carry one-bottle
'(He) brought the chief two bottles, and I brought one.'

The first example above is a simple transitive verb, with an agent subject and a theme object. There are only two arguments, and the verb has no suffixation, a typical monotransitive.[106] The second example here is doubly instructive. First, there are two instantiations of the same verb and both are clearly benefactive. The verb in the first

[106] Note here that *wong Tengger* is a genitive, not a distinct argument.

clause is followed immediately by the benefactive goal. In the second clause, however, the benefactive goal is not expressed. However, it is still clear that the second verb has an inherent benefactive reading, his bottle was also brought *for the chief*.' Crucially, there is no benefactive marking on either verb. This is especially surprising in the first clause, where the benefactive goal is overtly expressed. So the licensing of the benefactive goal here, and the transitive patient in the preceding examples, would appear to be unrelated to the applicative markers.

The above examples at the very least make it clear that the applicatives are not necessary to alter the argument structure of a verb. That is, to add an overt argument it is not necessary to affix an applicative ending in Tengger. The next question then, is the addition of an applicative sufficient to license a additional argument in the argument structure of a verb? Is there any relationship between the applicative endings and argument structure, and overt syntax?

Gil (2002) looks at similar constructions in Riau Indonesian, and provides a unified semantic analysis of the *–kan* marker in Riau Indonesian. In the previous chapter on voice, I have gone into detail concerning Gil's analysis of generalized voice markers in Riau Indonesian, (2000). He essentially claims there that the verbal affixes *di-* and *N-* in Riau Indonesian are generalized voice markers, indicating simply the existence of a patient or agent/experiencer argument in the argument structure of a verb. Crucially, he argues that these are semantic markers, and that there is no necessary syntactic affect from the markers.

For the applicative, there is nothing in print other than a handout from a 2002 conference in Canberra. I will therefore go into some detail to reconstruct the

argumentation there.[107] Riau Indonesian makes more frequent use of the applicative –*kan* than any other:

61. a. Bikin kopi untuk saya. [Riau Indonesian]
make coffee for 1.sg
'Make coffee for me.'

b. Bikin-kan saya kopi.
make-KAN 1.sg coffee
'Make me coffee.'

The above alternation at first glance looks to be the same that we have seen for standard Indonesian above. In (61a) the theme argument follows the unmarked verb, and the benefactive goal is in an adjunct PP. When the verb is marked with –*kan* in (61b), however, the benefactive goal appears in an argument position immediately following the verb. This is a typical double object like construction.

According to Gil, the unitary function of –*kan* can be expressed as follows :

> When applying to an expression X, the enclitic –*kan* asserts that the derived expression X-*kan* has an argument with the thematic role of end-point (EP).

There are three distinct subtypes of –*kan* in Riau Indonesian:

[107] Gil (2002) also shows that –*kan* in Riau Indonesian is interchangeable with –*in*. Further, he demonstrates that –*kan*, in that dialect, is clearly a clitic. These are interesting, but not crucial for the discussion here, so I do not reconstruct those arguments here.

i. marks argument already present in argument structure as ‘end-point’

ii. introduces a new argument into the argument structure of a verb, and marks an already present argument as ‘end-point’

iii. introduces a new argument into the argument structure of a verb and marks it as endpoint

Each of the three subtypes is demonstrated respectively below:

62. a. Rudy kasi sama aku.

Rudy give accompany 1.sg

’Rudy gave it to me.’

b. Kita kasi-kan sama dia lagi?

1.pl. give-KAN accompany 3 again

’Would they give it to him again?’

[Discussing what would happen if passport were lost]

63. a. Kamar-ku bersih.

room-1.sg clean

‘My room is clean.’

b. Ini kita bersih-kan besok.

this 1.pl clean-KAN tomorrow

’We’ll clean this up tomorrow.’

64. a. Dapit beli ikan.

Dapit buy fish.

'Dapit bought some fish.'

b. Beli-kan yang disko.

buy-KAN REL disco

'Buy me some disco ones.'

In (62) the verb itself, *kasi* is inherently ditransitive. In (62a) the verb appears unmarked, and the recipient follows the verb introduced by the macrofunctional element *sama*. In (62b) the verb has the applicative ending, and the recipient still appears following *sama*. In this case, the addition of the applicative enclitic simply serves to mark one of the already existing arguments as 'end-point', in this case *dia*, the recipient. In some sense, this use of *–kan* is redundant. The syntax of the utterance is clearly unaffected. Further, the argument structure, that is, the semantic frame of the predicate is similarly unaffected. The marker simply serves to emphasize a particular thematic relation.

The second use of *–kan* described in (ii) and demonstrated in (63) above, is essentially a causative use. An inchoative verb, like *bersih* 'clean', is made into a causative through the addition of *–kan*. Thus, we see in (63) addition of *–kan* yields a meaning of 'cause to be clean.' In this type, an argument is added to the semantic representation, a cause, which in (63a) is also the agent *kita* 'we.' The theme, *ini* 'this' is marked as 'end-point', the result of the action. So, an argument has been added to the semantic frame of the verb, and an existing argument has been marked as 'end-point.'

The examples in (64) show the third function of *–kan* in Riau Indonesian, which is basically the benefactive. Here a new argument is introduced to the semantic frame of the verb, a benefactive goal or recipient. This new argument is marked as the 'end-point' thematic role. Note here that the beneficiary is not expressed syntactically. Semantically, it is there nonetheless, as is clear from the interpretation. This is the same as that described above for standard Indonesian (56) and Tengger (57).

Crucially, here, Gil argues that *–kan* is a purely semantic marker. It has no affect on the clause structure. We already saw in (62b) above where the verb is marked by *–kan* and yet the recipient immediately following the verb is still introduced by *sama*. Similarly in (63b) the theme precedes the agent in a causative construction. Perhaps more striking is the following examples:

65. a. Bayar-kan taksi saya?[108]
pay-KAN taxi 1.sg
'Will you pay the taxi for me?'
[Getting out of the taxi]

b. Si Arip tak ada di-beli-kan aku, e.
PERS Arip NEG exist DI-buy-KAN 1.sg ADVERS
'Arip didn't buy me anything.'

In (65a) the benefactive goal follows the theme argument. There is no preposition here, and clearly no 'promotion'. Similarly, in (65b), notice that the verb has patient voice.

[108] Note that this is not a genitive construction, as both were getting out of the taxi, so a plural pronoun would have been used.

Nonetheless, the agent comes in preverbal position, and the benefactive immediately follows the verb.

We can now place *–kan* within the paradigm of weak generalized voice markers discussed in the previous chapter on voice:

69. Paradigm of Weak Generalized Verb markers

Form	*di-*	*N-*	*-na/-en*
Thematic Role	patient	agent	end point
Properties	existence	existence	existence

Again, as discussed in the previous chapter, voice markers in some languages encode very specific properties. This is true of say the passive in English, which is associated with a number of strong subject properties. In Tagalog, some subject properties are associated with the passive, though not as many as with the English passive. The claim that is made for Riau Indonesian, and that I am adopting here for Tengger, is that the only property associated with the voice markers is existence. That is, the presence of a voice marker, now including the applicative markers, simply asserts that a particular argument exists. It may not even be expressed, as we have seen a number of examples for Tengger. It does not imply that that argument has any other properties associated with it.

In another paper, Cole and Son (2004), take a syntactic approach to analyzing standard Indonesian *–kan*. They provide a unified account of *–kan* claiming that it is a

derivational morpheme 'affecting the argument structure of the verb to which it is affixed' (339). This is clearly not the case in Tengger. Further, they claim there that 'the role of *–kan* is to indicate the syntactic licensing of an argument in the argument structure that is not licensed syntactically by the base verb.' I have already shown that this type of analysis cannot account for the Tengger data. I have also shown above that this type of analysis also fails to account for the facts from Jakartan Indonesian.

We have already discussed in detail that Tengger like many western Indonesian languages, has a preference to not overtly express constituents. I have not made the claim, as is often assumed in much current syntactic theory, that these are simply phonologically null elements. That assumes that unexpressed elements are syntactically present, and simply fail to project into the final phonological representation (spell-out, PF, or other relevant level). Instead, I have argued that such constituents are present in the semantic frame of the verb, but that in fact they do not find expression in the syntax at all.

Similarly, this is exactly what I am claiming for the applicative markers in Tengger. In fact, they behave exactly as the applicative marker found in Riau Indonesian. They are optionally expressed markers of semantic relations present in a clause. Again, they have no direct affect on the overt syntax of a clause. Contrast the following two short conversations:

67. a. Lho, picis-é semana bisa tuku sapi, bisa tuku sepédha.

prt money-é that.much can buy cow can buy bike

'Well with that much money you could buy a cow, you could buy a bike.'

b. Nèk sapi tuku pira?

if cow buy how.much

'How much would it cost if you/one bought a cow?'

b. Ya emboh, tuku pira.

aff neg.know buy how.much

'Ya, I don't know how much you could buy it for.

68. a. Nah, pisan di-ungkab-na.

prt first di-open-na

'So, first you open it.'

c. Maringono di-lebok-na.

afterward di-enter-na

'Then you stuff it.' [*lit.* then you put it (the rice)in (the leaf).]

d. Wis mari, di-bungkus-na manèh.

prft finish di-wrap-na again

'When you're done, you wrap it again.'

In the above examples, those elements which are null in the Tengger version, are underlined in the English translation. In the first set of examples, only the patient is expressed in (67a and b). Note also that the patient precedes the verb in (68b). In the

second set of examples, there is not a single overt argument. Here, the speaker was showing me how to make *lontong* 'rice cooked in banana leaves' for a special occasion.

What is important here can be exemplified with the verb in (68b), which will be familiar by now, but it applies to all three constructions. In (68b) it is clear that not only is an agent argument not expressed, the one doing the stuffing, but necessarily so is a patient argument.[109] There are two null arguments. It is impossible to get an intransitive reading here. The *–na* implies, necessarily that there is a patient argument, even though it may not be expressed. This is crucial.

So, I have demonstrated clearly that the addition of an applicative ending does not affect the argument structure of a verb. It does not necessarily add an additional argument to a verb, or 'promote' a benefactive oblique. As can be seen in cases like (67) above, the presence of the applicative does necessarily imply that there is an additional argument. I have argued that the applicative endings in Tengger are optional overt markers of particular semantic relations that obtain either between a verb and its arguments or between arguments themselves. The applicative markers are not obligatory, but if they do appear, they necessarily encode one of these semantic relations. The relations that obtain, have been called 'end point' by Gil or 'RESULT' by Son and Cole, the actual term used is not of consequence here. Semantically however, both Son and Cole's 'RESULT' and Gil's 'End Point' identify the same semantic relationship between the verb and its arguments that is encoded by the applicative marker. This analysis actually provides partial evidence for Cole and Son's (2004) account of *–kan* in standard Indonesian, which they showed 'indicates the syntactic licensing of an argument in the argument structure that is not licensed syntactically by the base verb.' Indeed,

[109] A locative goal here would appear in a PP, *nang godhong* 'in the leaf'.

what I am claiming is that for Tengger, as well as many colloquial dialects of Indonesian, specifically here I have presented the facts for Jakartan Indonesian, the applicative is optional. When it does appear, it is merely the overt morphosyntactic realization of a semantic relation. In some cases, this relation allows overt changes in the syntax, i.e. the benefactive and the causative. However, there are many examples where there is no affect at all on the syntax of the clause, i.e. goal PP, inherent ditransitives, where the applicative is signaling the existence of the relation itself, and nothing more.

6.6 Conclusion

In this chapter, I have described in detail the main features of Tengger morphosyntax, other than the voice system. Tengger has a series of mood and applicative markers. I have shown that the mood markers in Tengger behave rather differently from those in both Old and standard Javanese. The propositive in Tengger is a recent borrowing from the lowlands. The optative marker in Tengger, encompasses both imperative and subjunctive functions, and inflects for second and third person. The Tengger system, however, has not undergone significant simplification. Rather, the standard language has undergone significant complexification.

There are two distinct applicative markers in Tengger. What I have called the applicative II, is undergoing interference from the lowland dialects, and there are now two competing forms in free variation, the original *–en* and the lowland *–na*. Further, I have argued that both of these applicatives are optionally overt semantic markers, encoding thematic relationships between a verb and its arguments.

CHAPTER SEVEN: CONCLUSION

In this dissertation I have presented aspects of the Tengger dialect of Javanese. The goal has been to present the facts of the language free from any a priori theoretical considerations. That being said, I have also addressed theoretical issues both when the theory was germane to the data at hand, and also when the data had significance for a particular theory.

What has emerged, I hope, is a picture of the Tengger language in perspective. In the first half of the dissertation, that perspective began with a brief discussion of Tengger lexicon, phonology, and morphology, before moving on to a presentation of the sociolinguistic setting of the Tengger language. Uniquely among Javanese dialects, Tengger does not have the highly elaborate speech level system. Sociolinguistically, the language is far more egalitarian. In the lexicon, phonology and morphology are seen many archaic features that have remained unchanged from Old and Middle Javanese. There have also been a number of significant innovations, though comparatively the 'standard' dialects of Yogya and Solo were clearly shown to be the most innovative, in terms of the speech level system, vowel harmony, and in the verbal paradigm.

In the second half of the dissertation, I considered aspects of the morphosyntax and syntax of Tengger. Tengger allows a striking variety of possible word orders, made all the more striking by the fact that Tengger has no agreement, case, tense, aspect, number, or person (with a few minor exceptions) overtly marked. It would be possible to provide a derivational analysis for many of the demonstrated word orders, however one would have to invoke such a range of movement opertations, VP fronting, V fronting, I fronting, object shift, heavy NP shift, scrambling, DO topicalization, IO topicalization,

and various focus constructions, among others. To do this does not help us to explain the facts of Tengger itself. Instead, I have argued that Tengger has multiple base generated word orders that result from the historical shift from VOS to SVO. Further, I have shown that the situation in Tengger is stable. I have not claimed that Tengger is a totally free word order language, but rather have shown that Tengger clearly has NP and VP constituents.

When scholars address Austronesian languages, the voice system often emerges as one of the most interesting and characteristic aspects. The Tengger voice system is no different. I have shown that due to recent influence from the lowlands, the Tengger voice system, previously a typologically symmetrical system has developed aspects of a mixed system, as in many western Indonesian languages. Initially the active and passive voice markers appear to license the presence of arguments in the argument structure of verbs, and demand a strict linear order of arguments. However, I have shown that this is clearly not case, so we have examples such as:

1. a. (Mas Tom) di-golek (Marsam).
   Mas T DI-look.for Marsam
   b. (Mas Tom) ng-golek (Marsam).
   Mas T N-look.for Marsam
   'Mas Tom is looking for Marsam.'

In the first example above the verb is marked with passive verbal morphology and in the second example the verb is marked with active verbal morphology. However, the linear

order of constituents is exactly the same, and the interpretation is also exactly the same. Further, either of the arguments in either example could be unexpressed. I have argued that just such evidence demonstrates that the voice morphology in Tengger does not license arguments of the predicate. Rather, the active and passive voice markers in Tengger mark the presence of a particular semantic role in the argument structure of a predicate. Note that the argument itself can be unexpressed. Thus, the occurrence of *DI-* simply indicates that there is a theme or patient present in the argument structure, and the nasal prefix *N-* simply indicates the presence of an agent or actor.

A similar analysis was then extended to account for the applicative suffixes in Tengger. They were placed within a typology of generalized voice markers, which are option, overt markers of the presence of particular thematic roles in the argument structure of a predicate. Under this analysis then, we can capture the difference between examples where the applicative markers seem to directly license the presence of an overt argument, and examples where the argument is present, but the applicative marker fails to appear.

BIBLIOGRAPHY

1978. *Upacara Kasada dan Beberapa Adat Istiadat Masyarakat Tengger.* Jakarta: DepDikBud.

Adams, Karen L., and Thomas J. Hudak. 1992. *Papers from the 2nd Annual Meeting of the Southeast Asian Linguistics Society.* Tempe: ASU.

Adelaar, K. Alexander. 2005. *Salako or Bademea.* Wiesbaden: Harrossowitz Verlag.

Adelaar, K. Alexander and Nikolaus P. Himmelmann (eds.). 2005. *The Austronesian languages of Asia and Madagascar.* London: Routledge.

Adger, D. 1995. *Agreement and the DP in Celtic Languages.* Ms. (Doctoral Dissertation).

Alieva, Natalia. 1978. Observations on Typological Evolution in Indonesian Languages. In Wurm, S.A. and Lois Carrington eds*., Second International Conference on Austronesian Linguistics: Proceedings*: 1.327-334.

Alves, Mark, Ed. 1993. *Papers from the 3rd annual Meeting of the Southeast Asian Linguistics Society.* Tempe: ASU.

Anderson, Henning, Ed. 2003. *Language Contacts in Prehistory.* Amsterdam: John Benjamins Publishing.

Anderson, Stephen. 1992. *A-Morphous Morphology.* Cambridge: Cambridge University Press.

Arka, I Wayan. 1998. *From Morphosyntax to pragmatics in Balinese: A lexical-functional approach.* Ph.D. dissertation. University of Sydney.

Arka, I Wayan and Malcolm Ross. 2005. *Aspects of Austronesian voice systems.* Canberra: Pacific Linguistics.

Arifin, Syamsul, et.al. 1990. *Tipe-tipe Klausa Bahasa Jawa*. Jakarta: PPPB.

Artawa, K. 1994. Ergativity and Balinese Syntax. Ph.D. dissertation. La Trobe University, Bundoora, Australia.

Baker, Mark. 1988. *Incorporation*. Chicago: Chicago University Press.

Baker, Mark. 2002. *Lexical Categories: Verbs, Nouns, and Adjectives.* Cambridge: Cambridge University Press.

Baker, Mark. 2008. *The Syntax of Agreement and Concord.* Cambridge: Cambridge University Press.

Baltin, Mark, and Chris Collins, eds. 2001. *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell.

Bayer, Joseph. 1999. Comments on Cardinaletti & Starke. In van Riemsdijk ed.

Bellwood, Peter. 1985. *Prehistory of the Indo-Malaysian archipelago.* Sydney: Academic Press.

Bellwood, Peter, James J. Fox, and Darrell Tryon, Eds. 1995. *The Austronesians: Historical and comparative perspectives.* Canberra: ANU

Berman, Laine. 1998. *Speaking through the Silence.* Oxford: Oxford U. Press.

Bittner, Maria, and Ken Hale. 1996. The Structural Determination of Case and Agreement. In *Linguistic Inquiry* 27.1:1-68.

Blake, Barry J., and Kate Burridge, Eds. 2003. *Historical Linguistics 2001.* Amsterdam: John Benjamins Publishing.

Blust, Robert. 1985. *Telaah Komparatif Bahasa Nusantara Barat.* Jakarta: Djambatan.

Bresnan, Joan. 2000. The Emergence of the Unmarked Pronoun. Ms. Stanford University.

Brown, Keith, and Jim Miller, Eds. 1999. *Concise Encyclopedia of Grammatical Categories.* Amsterdam: Elsevier.

Cardinaletti, Anna, & Michal Starke. 1994. The Typology of Structural Deficiency. Ms. University of Venice, University of Geneva.

Cardinaletti, Anna, & Michal Starke. 1996. Deficient Pronouns: A View From Germanic. In H. Thrainsson et. al. eds*., Studies in Comparative Germanic Syntax II*: 21-65. Amsterdam: Kluwer Academic Press.

Cardinaletti, Anna, & Michal Starke. 1999. The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns. In van Riemsdijk ed.

Cardinaletti, Anna. 1994. On the Internal Structure of Pronominal DPs. In *The Linguistic Review* 11:195-219.

Carnie, Andrew, and Eithne Guilfoyle, Eds. 2000. *The Syntax of Verb Initial Languages.* Oxford: Oxford Press.

Chelliah, Shobhana L., and Willem J. de Reuse, Eds. 1995. *Papers from the 5th Annual Meeting of the Southeast Asian Linguistics Society.* Tempe: ASU.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge: MIT Press.

Chomsky, Noam. 1998. *Minimalist Inquiries: The Framework.* MIT Occasional papers in linguistics: 15.

Chomsky, Noam. 1999. *Derivation by Phase.* MIT Working Papers in Linguistics: 18.

Chung, Sandra. 1976. On the Subject of Two Passives in Indonesian. In C. Li ed., *Subject and Topic*: 57-98. New York: Academic Press.

Clynes, A. 1995. Topics in the Phonology and Morphosyntax of Balinese. Ph.D. dissertation, ANU, Canberra.

Cmerjrkova, Svetla, and Frantisek Sticha, Eds. 1994. *The Syntax of Sentence and Text.* Amsterdam: John Benjamins Publishing.

Cole, Peter, and Gabriella Hermon. 1998. The Typology of Wh-Movement: Wh-questions in Malay. In *Syntax* 1.33:221-258.

Cole, Peter, and Gabriella Hermon. 2000. Partial Wh-Movement: Evidence from Malay. In Lutz, Gereon, and von Stechow.

Cole, Peter, and Gabrielle Hermon. 2003. The typology of Malay reflexives. In *Lingua:* 115: 627-644.

Cole, Peter, and Min-jeong Son. 2004. The Argument Structure of Verbs with the Suffix *–kan* in Indonesian. In *Oceanic Linguistics* 43.2: 340-364.

Cole, Peter, and Min-jeong Son. To appear. An Event-Based Account of *–kan* Constructions in Standard Indonesian. In *Language.*

Conners, Thomas J. 2002. Circumfixation: An Unnoticed Complication for Indonesian Stress. Ms. Yale University.

Corver, Norbert, and Denis Delfitto. 1994. Feature Asymmetry and the Nature of Pronoun Movement.

Croft, William. 1990. *Typology and Universals.* Cambridge: Cambridge U. Press.

Cumming, Susanna. 1986. Word order change in Malay. In *FOCAL I: Papers from the Fourth International Conference on Austronesian Linguistics.* Paul Geraghty, et.al eds.

Cumming, Susanna. 1991. *Functional change: The case of Malay constituent order.* Berlin: Mouton de Gruyter.

Davies, William D. 1999. Madurese and Javanese as Strict Word-Order Languages. In *Oceanic Linguistics* 38.1: 152-167.

Davies, William D. 2003. Extreme Locality in Madurese Wh-Questions. In *Syntax* 6.3: 237-259.

Davies, William D. 2005. The richness of Madurese voice. In Arka and Ross eds.

De Guzman, Videa. 1978. *Syntactic Variation of Tagalog Verbs.* Honolulu: University Press of Hawaii.

Diesing, Molly. 1999. Comments on Cardinaletti & Starke. In van Riemsdijk ed.

Dobrovie-Sorin, Carmen. 1999. Comments on Cardinaletti & Starke. In van Riemsdijk ed.

Donohue, Mark. 1999. *A Grammar of Tukang Besi.* Berlin: Mouton de Gruyter.

Duranti, Alessandro. 1997. *Linguistic Anthropology.* Cambridge: Cambridge U. Press.

Dutton, Tom, and Darrell Tryon, Eds. 1994. *Language Contact and Change in the Austronesian World.* Berlin, New York: Mouton de Gruyter.

Dyen, Isidore. 1953. *The Proto-Malayo-Polynesian Laryngeals.* Baltimore: LSA.

Dyen, Isidore. 1965. *The Lexicostatistical Classification of the Austronesian Languages.* Baltimore: Waverly Press.

Effendi, S. 1981. *Beberapa Masalah Sintaksis Bahasa Jawa.* Jakarta: PPPB.

Epstein, Samuel David, and Norbert Horstein, eds. 1999. *Working Minimalism.* Cambridge: MIT Press.

Errington, J. Joseph. 1985. *Language and social change in Java: linguistic reflexes of modernization in a traditional royal polity*. Athens, Ohio: Ohio University Press.

Errington, J. Joseph. 1988. *Structure and style in Javanese: a semiotic view of linguistic etiquette.* Philadelphia: University of Pennsylvania Press.

Errington, J. Joseph. 1998. *Shifting Languages: Interaction and Identity in Javanese Indonesia.* New York: Cambridge University Press.

Foley, William Auguste. 1976. *Comparative Syntax in Austronesian.* Ph.D. Dissertation, University of California, Berkeley.

Foley, William Auguste. 1998. Symmetrical Voice Systems and Precategoriality in Philippine Languages. Sydney: University of Sydney.

Geraghty, Paul, L. Carrington, and S.A. Wurm. 1986. *Focal I: Papers from the Fourth International Conference on Austronesian Linguistics*. Pacific Linguistics Series #93: The Australian National University.

Gil, David. 1994. The Structure of Riau Indonesian. In *Nordic Journal of Linguistics* 17:179-200.

Gil, David. 2001. Creoles, Complexity, and Riau Indonesian. In *Linguistic Typology* 5:325-371.

Gil, David. 2001b. Escaping Eurocentrism: Fieldwork as a process of Unlearning. In *Linguistic Fieldwork*. P. Newman and M. Ratliff eds. Cambridge: Cambridge University Press.

Gil, David. 2002. The Prefixes *di-* and *N-* in Malay/Indonesian dialects. In Wouk and Ross eds.

Gil, David. 2004. Riau Indonesian *sama,* Explorations in Macrofunctionality. In *Coordinating Constructions*, M. Haspelmath ed. Amsterdam: John Benjamins.

Gil, David. 2006. The acquisition of voice morphology in Jakarta Indonesian. In *The Acquisition of Verbs and Their Grammar.* N. Gagarina and I. Gulzow eds. Netherlands: Springer.

Goodenough, Ward H., Ed. 1996. *Prehistoric Settlement of the Pacific.* Philadelphia: American Philosophical Society.

Grenoble, Lenore, and Lindsay Whaley, Eds. 1998. *Endangered Languages.* Cambridge: Cambridge U. Press.

Grohman, Kleanthes Kostas. 1996. The Left Periphery of West Germanic Embedded Clauses. In *University of Maryland Working Papers in Linguistics* 4:21-47.

Guilfoyle, Etienne, H. Hung & L. Travis. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. In *Natural Language and Linguistic Theory:* 10.375-414.

Haegeman, Liliane. 1999. Comments on Cardinaletti & Starke. In van Riemsdijk ed.

Hale, Kenneth, and Samuel Jay Keyser, eds. 1993. *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger.* Cambridge: MIT Press.

Hale, Kenneth, and Samuel Jay Keyser. 1993. On Argument Structure and the Lexical Expression of Syntactic Relations. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger,* Kenneth Hale and Samuel Jay Keyser eds. Cambridge: MIT Press.

Halim, Amran, L. Carrington, and S.A. Wurm. 1981. *Papers from the Third International Conference on Austronesian Linguistics vols. 2, 4*. Pacific Linguistics Series #77: The Australian National University.

Hawkins, John A. 1983. *Word Order Universals.* New York: Academic Press.

Hefner, Robert W. 1985. Hindu Javanese: Tengger Tradition and Islam. Princeton: Princeton University Press.

Hestvik, Arild. 1992. LF Movement of Pronouns and Antisubject Orientation. In *Linguistic Inquiry* 23.4:557-594.

Himmelmann, Niklaus. 2005. The Austronesian languages of Asia and Madagascar: typological characteristics. In Adelaar and Himmelmann eds.

Holmberg, Anders. 1999. Comments on Cardinaletti & Starke. In van Riemsdijk ed.

Hopper, Paul J., and Elizabeth Closs Traugott. 1993. *Grammaticalization.* Cambridge: Cambridge U. Press.

Iskandarwassid, Et.Al. 1985. *Struktur Bahasa Jawa Dialek Banten.* Jakarta: PPPB.

Janse, Mark, and Sijmen Tol, Eds. 2003. *Language Death and Language Maintenance*. Amsterdam: John Benjamins Publishing.

Janse, Mark, Ed. 1998. *Productivity and Creativity.* Berlin: Mouton de Gruyter.

Jeffers, Robert, and Ilse Lehiste. 1979. *Principles and Methods for Historical Linguistics.* Cambridge: MIT Press.

Kana, Marit. 1983. The Syntax and Semantics of Verbal Prefixes in Bahasa Indonesia. In Halim, A. et. al. eds*., Papers from the Third International Conference on Austronesian Linguistics*: 4.21-33.

Kaswanti Purwo, Bambang. 1984. *Deiksis Dalam Bahasa Indonesia*. Jakarta: PN Balai Pustaka.

Kayne, Richard S. 1989. Null Subjects and Clitic Climbing. In *The Null Subject Parameter* O. Jaeggli and K. Safir eds. Dordrecht: Kluwer.

Koopman, Hilda. 1999. The Internal and External Distribution of Pronominal DPs. In K. Johnson and I. Roberts eds. *Beyond Principles and Parameters.* Dordrecht: Kluwer.

Kosadi Hidayat, et.al. 1992. *Kamus Bahasa Cirebon-Indonesia.* Jakarta:PPPB.

Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog.* Stanford: CSLI Publications.

Lasnik, Howard. 1999. *Minimalist Analysis.* Oxford: Blackwell.

Lass, Roger. 1997. *Historical Linguistics and Language Change.* Cambridge: Cambridge U. Press.

Lehiste, Ilse. 1988. *Lectures on Language Contact.* Cambridge: MIT Press.

Lightfoot, David. 1979. *Principles of Diachronic Syntax.* Cambridge: Cambridge U. Press.

Longobardi, Giuseppe. 1994. Reference and Proper Names: A Theory of N-Movement in Syntax and Logical Form. In *Linguistic Inquiry* 25.4:609-665.

Lüem, Barbara. 1988. *Wir sind wie der Berg lächelnd aber stark.* Ph.D. Dissertation: Universität Basel.

Lutz, Uli, Gereon Muller, and Arnim von Stechow Eds. 2000. *Wh-Scope Marking*. Amsterdam: John Benjamins Publishing.

Lynch, John. Fa'afo Pat, Eds. 1993. *Oceanic Studies: Proceedings of FICOL.* Canberra: Pacific Linguistics, ANU.

Malouf, Robert P. 2000. *Mixed Categories in the Hierarchical Lexicon.* Stanford: CSLI.

Martin, Roger, David Michaels, Juan Uriagereka, eds. 2000. *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik.* Cambridge: MIT Press.

McGinn, Richard, ed. 1988. *Studies in Austronesian Linguistics*. Athens, Ohio: Ohio University Monographs in International Studies #76.

Meinunger, Andre. 2000. *Syntactic Aspects of Topic ad Comment.* Amsterdam: John Benjamins Publishing.

Miller, D. Gary. 2002. *Nonfinite Structures in Theory and Change.* Oxford: Oxford U. Press.

Mühlhäusler, Peter, and Rom Harre. 1990. *Pronouns and People.* Basil: Blackwell.

Musgrave, Simon. 2001. Non-subject arguments in Indonesian. Ph.D. dissertation. Melbourne University.

Myhill, John. 1988. Nominal Agent Incorporation in Indonesian. *In Journal of Linguistics*: 24.111-136.

Nardiati, Sri, et.al. 1995. *Sistem Morfemis Nomina Bahasa Jawa-Indonesia.* Jakarta: PPPB.

Naylor, Paz Buenaventura. 1979. *Austronesian Studies: Papers from the 2nd Eastern Conference on Austronesian Languages*. Michigan Papers on South and Southeast Asia.

Nothofer, Bernd. 1975. *The reconstruction of Proto-Malayo-Javanese*. 's-Gravenhage: Martinus Nijhoff.

Nothofer, Bernd. 1996. *Reconstruction, Classification, Description: Festschrift in Honor of Isidore Dyen.* Hamburg: Abera.

Ogloblin, Alexander K. 2005. Javanese. In Adelaar and Himmelmann eds.

Padmapuspita. ND. *Candi Sukuh dan Kidung Sudamala.* Jakarta: DepDikBud.

Patmadiwiria, Munadi. 1977. *Kamus Dialek Jawa Banten-Indonesia.* Jakarta: PPPB.

Paul, Ileana, V. Phillips, and L. Travis. 2000. *Formal Issues in Austronesian Linguistics*. Dordrecht: Kluwer Academic Publishers.

Paul, Ileana, Vivianne Phillips, and Lisa Travis, Eds. 2000. *Formal Issues in Austronesian Linguistics.* Dordrecht/Boston/London: Kluwer Academic Publishers.

Pesetsky, David. 1995. *Zero Syntax*. Cambridge: MIT Press.

Pigeaud, Theodore G. Th. 1976. *Islamic States in Java 1500-1700.* The Hague: Martinus Nijhoff.

Poedjosoedarmo, Soepomo, dkk. 1979. *Morfologi Bahasa Jawa.* Jakarta: PPPB.

Postman, Whitney. 2002. Thematic role assignment in Indonesian: A case study in agrammatic aphasia. Ph.D. dissertation, Cornell University.

Prawiroatmojo, S. 1981. *Bausastra Jawa-Indonesia.* Jakarta: Gunung Agung.

Pylkkanen, Liina. 2002. *Introducing arguments.* Ph.D. dissertation, MIT.

Ritter, Elizabeth. 1995. On the Syntactic Category of Pronouns and Agreement. In *NLLT* 13:405-443.

Roberts, Ian, and Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization.* Cambridge: Cambridge U. Press.

Robson, Stuart. 2002. *Javanese Grammar for Students.* Clayton: Monash Asia Inst.

Robson, Stuart, and Singgih Wibisono. 2002. *Javanese English Dictionary.* Hong Kong: Periplus.

Ross, Malcolm D. 2002. The history and transitivity of Western Austronesian voice and voice marking. In Wouk and Ross eds.

Runner, Jeffrey T. 1994. A Specific Role for AGR. Ms. University of Massachusetts, Amherst.

Saddy, Douglas. 1991. Wh Scope Mechanism in Bahasa Indonesia. In Lisa Cheng and Hamida Demiradash (eds), More Papers on Wh-Movement. *MITWPL* 15:183-218.

Saddy, Douglas. 1992. A versus A-bar movement and Wh-fronting in Bahasa Indonesian. Ms., University of Queensland.

Sapir, Edward. 1921. *Language: An introduction to the study of speech.* New York: Harcourt, Brace, and Co.

Schachter, Paul. 1976. The subject in Philippine languages. In Li ed.

Sirk, Ulo. 1996. On the History of Transitive Verb Suffixes in the Languages of Western Indonesia. In Steinhauer, H. ed., *Papers in Austronesian Linguistics*: 3.191-205.

Smith-Hefner, Nancy. 1983. Language and social identity: Speaking Javanese in Tengger. Ph.D. dissertation, University of Michigan.

Sneddon, James Neil. 1996. *Indonesian: A Comprehensive Grammar*. London, New York: Routledge.

Soedjiatno, Et.Al. 1984. *Kata Tugas Bahasa Jawa Dialek Jawa Timur.* Jakarta: PPPB.

Soedhito et. Al. 1984. *Struktur bahasa Jawa dialek Tengger.* Jakarta: PPPB.

Soemarmo, Marmo. 1970. Subject-predicate, focus-presupposition, and topic-comment in Bahasa Indonesia and Javanese. Ph.D. dissertation, UCLA.

Soetoko, Et.Al. 1981. *Geografi Dialek Banyuwangi.* Jakarta: PPPB.

Soetoko, Et.Al. 1984. *Geografi Dialek Bahasa Jawa di Kabupaten Surabaya.* Jakarta: PPPB.

Sornikola, Rosanna, et.al., Eds. 2000. *Stability, Variation, and Change of Word-Order Patterns Over Time.* Amsterdam: John Benjamins Publishing.

Steinhauer, H. 1991. *Papers in Austronesian Linguistics No. 1*. Pacific Linguistics Series #81: The Australian National University.

Steinhauer, Hein, Ed. 1988. *Papers in Western Austronesian Linguistics No. 3.* Canberra: ANU.

Sudaryono, et.al. 1988. *Geografi Dialek Bahasa Jawa Kudus.* Jakarta: PPPB.

Sugono, Dendy. 1985. *Verba Transitif Dialek Osing: Analisis Tagememik.* Jakarta: PPPB.

Tadmor, Uri. 2003. Final /a/ mutation: a borrowed areal feature in western Austronesian languages. In *Issues in Austronesian Historical Phonology*, John Lynch ed. Canberra: Pacific Linguistics.

Tadmor, Uri. 2007. Kontroversi asal-usul Bahasa Melayu-Indonesia. In *PELBBA 18:* Jakarta

Tasmowski, Liliane, and S. Paul Verluyten. 1987. Control Mechanisms of Anaphora. In *Journal of Semantics* 4:341-370.

Thurgood, Ela. 2004. Phonation Types in Javanese. In *Oceanic Linguistics 43.2.*

Uhlenbeck, E.M. 1964. *The Languages of Java and Madura.* The Hague: Martinus Nijhoff.

Uhlenbeck, E.M. 1975. Sentence segment and word group: Basic concepts of Javanese syntax. In *NUSA* 1, ed. J.W.M. Verhaar, 6-10. Jakarta: NUSA.

Uhlenbeck, E.M. 1983. *Javanese Linguistics.* Dordrecht: Foris Publications.

Uriagereka, Juan. 1995. Aspects of the Syntax of Clitic Placement in Western Romance. In *Linguistic Inquiry* 26.1:79-123.

Uriagereka, Juan. 1999. Comments on Cardinaletti & Starke. In van Riemsdijk ed.

Van Gelderen, Elly. 1993. *The Rise of Functional Categories.* Amsterdam: John Benjamins Publishing.

Van Riemsdijk, Henk. 1999. *Clitics in the Languages of Europe.* Berlin, New York: Mouton de Gruyter.

Voskuil, Jan. 2000. Indonesian Voice and A-bar Movement. In I. Paul, V. Phillips, L. Travis eds., *Formal Issues in Austronesian:* 195-212. Dordrecht: Kluwer Academic Publishers.

Wechsler, S. and I W. Arka. 1998. Syntactic Ergativity in Balinese: an Argument Structure Based Theory. In *NLLT:* 16:387-441.

Widyaprakosa, Simanhadi. 1994. *Masyarakat Tengger.* Yogyakarta: Kansius.

Wiesman, Ursula, ed. 1986. *Pronominal Systems.* Tubinen: GNV.

Wilkinson, R.J. N.D. *A Malay-English Dictionary.* Tokyo: Daitoa Syuppan Kabusiki Kaisya.

Winford, Donald. 2003. *Introduction to Contact Linguistics.* Oxford: Blackwell Publishing.

Wolff, John U. 1989. *Kamus Indonesia Inggris*, Edisi Ketiga. Jakarta: PT Gramedia. (Originally compiled by Echols, John M. and Hassan Shadily. 1961, revised 1963.)

Wolff, John U. 1996. The development of the passive verb with pronominal prefix in western Austronesian languages. Ms. Cornell University.

Wolff, John U. and Soepomo Poedjosoedarmo. 1982. *Communicative Codes in Central Java.* Ithaca: Southeast Asia Program.

Wouk, Fay, and Malcolm Ross, Eds. 2002. *The History and Typology of Western Austronesian Voice Systems.* Canberra: Pacific Linguistics, ANU.

Wurm, S.A. and Lois Carrington, eds. 1978. *Second International Conference on Austronesian Linguistics: Proceedings*. Pacific Linguistics Series #61: The Australian National University.

Zoetmulder, P.J. 1982. *Old Javanese-English Dictionary.* Leiden: 'S-Gravenhage—Martinus Nijhoff.

Zoetmulder, P.J. 1983. *Kalangwan.* 2nd ed. Djambatan: Yogyakarta.

Zribi-Hertz, Anne, and Liliane Mbolatianavalona. 1999. Toward a Modular Theory of Linguistic Deficiency: Evidence from Malagasy Personal Pronouns. In *NLLT*

Zurbuchen, Mary S. 1976. *Introduction to Old Javanese Language and Literature.* Ann Arbor: Center for South and Southeast Asian Studies, UM

APPENDIX: WORD LISTS

Javanese Word List Comparison

English	Yogyakarta	Tengger/Ngadas	Malang	Old/Middle Javanese
1st per. pron.masc	aku	éyang/réyang*	aku	aku, kami, i(ng)sun -ku/nghulun
1st per. pron.fem	aku	isun	aku	*ibid*
2nd per. pron.	kowé*	sira	koèn*	kanyu/kita –mu/-nta *sira '3rd person'*
2nd per. pron. (distant)	sampéyan#/ penjenèngan#	rika#	sampéyan#	*ibid, ri+ika* 'there' *sampe* 'disgusting, lowly'; *jeneng* 'function, mastery' *panyjenengan* 'stop, permanent position'
3rd per. pron.	dhéwèké#	sira, dhéwèké#, iku, ika, dhéwèkné#, sebutan keluarga	arèké, dhéké#	sira, -nya/-nira *demen* 'to like' *dhewek (dawak)* 'self'
head	sirah	endhas	sirah, ndhas	eṇdhas; śirah (skt)
nose	irung	congor*	éróng	(h)irung
hand	tangan	tangan	tangan	tangan
elbow	sikut#	céngkul*/ sikut#	sékót#	siku; *sikut* 'edge, border, size'
(finger/toe) nail	kuku	kuku	kuku	kuku
finger	driji*	driji*	driji*	
foot	sikil*	sikil*	sékél*	
ankle	polok*	bènthol, pelok*, ugel-ugel*	tungkak#	*tungkak* 'heel'
knee	dhengkul#	dhéngkul#	dhengkól#	*dengkul* 'squeeze, fold, lap over'
calf	kéntol*/kémpol*	kèntol*, bèntis#, wèntis: 'shin' wèntis, bèntis	pedèt*	we(n)tis; *bentis* 'about the stomach ?'
skin	kulit	kulit	kólét	kulit
stomach	weteng	weteng	weteng	weteng ('rahim')
waist, loins	boyok*	béyèk*, pangéan*	bankèkan*	
bone	balung	belung	balóng	b/walung

liver	ati	ati	ati	(h)ati
heart	jangtung*	jantung*, ati, montong	jantóng*	
breast/milk	susu	susu	susu	susu
shoulder	bau	bau, pundhak	pun	pundhak, *bahu* (skt) ‘arm’
chest, breast	dhódhó	dhadha	dhodho	dhadha
rib, flank, side	igó	belung gambang, iga	igo	(h)iga
neck	gulu	gulu	gulu	gulu
blood	getih*	getih*, darah	getéh*	
hair	rambut	rambut	rambót	rambut
cheek	pipi	pipi	pipi	pipi
mouth	cangkem	lam bé#, cangkem	cangkem, tutók	cangkem, *lambé* ‘lips’
tongue	ilat	ilat	ilat	(h)ilat
ear	kuping	kuping	kópéng	ku/oping
eye	mripat*	mata	mripat*, moto	mata,
wash hair	kramas*	wuwung#, kramas	kramas*	*wuwung* ‘roof cross-beam’,
wash clothes	umbah-umbah	masoh, umbah-umbah, ombah, wumbah, ngumbah, memasoh	umbah-umbah	kumbah; (w)ase/uh =masoh/inasehan
wash plates	isah-isah	esoh-esoh	korah-korah*	(w)ase/uh(?); mang-eső ‘scrub’
wash face	raup	raup	raóp	rahup
try	nyobo*	njajal, hakéta jajal	nyobak*	jajal
meet	ketemu	ketemu campuh, kepethuk*	temu	ka-temu; campuh/r ‘go outside of one’s caste’
play	dolan	main*, dedolan, dolan, moto, dolanan	dulinan*, dólén	adodolan
warm by the fire	api-api	gegeni	njarang*	*api-api* ‘make excuses’
talk/chat	omong*	bantah*, omong*, guneman	omong-omong*	
teach	mulang	mulang, ngajari, muruk, mlajari	ngajar	ajar; *wulang* ‘rope’
plant	nandur	manja*, nandur	nandór	tandur;

crush/smash	diblok/dheplok	mbebeg, ajur (menghilangkan kulit): ngécros	ngejór	
hoe	gebuk	molah, macul	macól	pacul; *um-olah* 'busy with, moving w'
ascend, go up	munggah	munggah, mènèk* (wit), mentas#	mundak, munggah	m-unggah, *m-entas* 'get out of water'
light a fire	cethik*	dèdèn, gawé#	ngórópno	hurip *dadi-en?, gawé* 'do, make, cause'
to weed	bubut-bubut	ngrumat**	matón	bubut, cabud,
glean	ngasak	mèpèl	nglómpókno	
free corn from cob	mipili	ng//kocèki*, ngeculèn, mbosèt*, (lepas dari stalk): mecok	mipél	amipili (pipil),
to light	madhangi/ sumed	nguripèn, murup, nyocok, mbarak	ngórópno	hurip, pandhang
place flowers on a grave	nyekar	nyekar, mburéi, ndokok, nyeleh kembang	nyekar	sekar/anekar, kembang/ angembang
sew	njait	(mesin) njait, (tangan) nondomi	njaét	dom, *jahit* 'decorated cloth'
see	ndèlok/ndèlèng	masèn, nguasèn	ndelok	delő, deleng (in later texts)
sift	nampèni/ngayak	ngiriki#	ngayak, nyaring	ayak *(i)rik* 'shrill sound'
walk	mlaku	mlaku	mlaku	laku
run	mlayu	mlayu	mlayu	layu, alayu, palayu 'larinya'
come, arrive	tekó	teka	teko	teka
return	mbali	mbalik	mbalék, móléh	balik
bathe	adus*	adus*	adós*	
know	wruh/ngerti	wruh	ngerti	wruh, *arthi* (skt) request, meaning'
think	mikir*	mikir*	pékér*	i/emut 'remember, be aware'
breathe	ambeg(an)	ambekan	ambekan	ambekan n. mambekan v.
smell	ngambu	ngambèt, mambu	ambu	ambe/ő, angambung

kiss	ngambung	ngambung	ambóng	angambung, inambung
laugh	ngguyu	ngguyu	ngguyu	aguyu
cry	nangis	nangis	nangis	tangis
vomit	mu(n)tah	mlèkok*, mutah	mutah	wutah/mutah
spit	ngidu	ngidu, idu-idu	ngidu	(h)idu
eat	mangan	mangan	mangan	pangan/amangan
cook rice	adang*	adang*, nglisik*	adang*	liwet, angliwet
drink	ombé*	ngombé*	ngómbé*	
bite	nyokot//cokot	nyokot	cókót	cokot, anoko
hear	krungu	ngrungoken	krungu	rungu, rungő
lying down	lujih/turon	léyé-léyé, leluyé, léyéan, luyéan	turon	*lume, lumay, alumeh* 'no energy' luya 'exhausted'
sleep	turu	turu	turu	turu
dream	ngimpi	ngipi	ngimpi	ipi, angipi
sit	lungguh	lungguh	lónggóh	lungguh, linggih
stand up	ngadeg	ngadeg	ngadeg	adeg
wake up	tangi	tangi	tangi	tang(h)i
say	kandha#/ngomong*	bantah*	ngomong*	*kandha* 'part. esp of poem, story'
ask	takon	takon	takon	takwan, atakwan
answer	nyauri*/sumaur*	nyauri/*, sumaur*, njawab*, balas takon	njawab*	
lie, deceive	ngapusi/mbujuki	nggoro'i, moro'i	goroh, ngapusi	apus (rare), goroh
hunt		gladhak*, nggundang*	mbebedag	buru, amuru; bedhag, ambedhag
fight	gelut	gelut, tukar, pepathalon, gégèr, keputan	gelót	gelut
push	surung	nyurung, medel, jrengen	sóróng	surung, asurung
pull	tarik	mbatek, matek, mbaduk	tarék	tarik, anarik; watek, amatek
steal	nyolong	nyolong	colong	colong, anyolong
give	mènèhi/ké'i	ngènéi, ngewé'i	kèk*	?weh, w-in-eh-i
take	njupuk*	njukuk*, njupuk*	jópók*	
kill	matèni	matèni	matèni	pati, amati
live, life	urip	urip	órép	hurip
die, dead	mati	mati, matèk	mati	pati

		(untuk kewan)		
cut	ngethok//kethok *	ngethok*, negor	ketók*	tugel, anugel; tegor, ategor (tree); *potong* ‘release horns (animal)’
work	kerja/nduwé gawé	nyambut gawé, kerja	kerjo	*karya (skt)*
grow, sprout	thukul*	cukul*, urip	thukól*	
swell, puff up	aboh*	aboh*	ngembang, mekar	
hold, grasp	cekel	nyekel, ngandhèli, nggacèki	cekel	cekel, cumekel
sieze	tangkep	nyekel, nangkep	rebót	*tangkep* ‘fight, argue, oppose’
buy	nuku/tuku	tuku	tuku	tuku
sell	dol	dodol, adol	dol	dwal, dol, adol
borrow	silih	nyilih	síléh	*silih* ‘each, other, exchange, change’
open	mbuka(k)	mbukak, ngengaken/(lawang)	bukak	*buka* ‘beginning, openning’
close	tutup	nutup, nginep# (lawang)	tótóp	tutup; *inep* ‘sleep over’
throw, toss	mbalang	ngguwang, ngucale#n, nganthem	uncal	balang, ambalang; *ucal* ‘tweezers, pinch; button’
fall	tibó	rutu*, tiba (trip)	rótóh*	tiba;
fly	mabur	miber, mabur	miber	iber, (u)miber;
fly (of a bird)	miber	miber	miber	
flow	mili	mili	kényót	hili, ili "aliran air"
burn	ngobong//obong	mbakar (makanan), obong (barang)	kobong	ababakar(sst), binakar (mayat)
urinate	uyuh	nguyuh	ngóyóh	uyuh, ěyěh"air kencing"
ride (a bus)	numpak	numpak, munggah, nunggang	nyópér	
get down/off	medhun	mudhun, mudhuk, misor	módhón	
dog	asu	kirik	asu	asu
-puppy				kirik

monkey	kethèk	pethès, kethèk	ketèk	
cat	kucing	kucing	kócéng	kucing
bird	manuk	manuk	manók	manuk
-baby bird				
chicken	pitik	pitik	péték	ayam, pitik
-baby chicken				pitik
egg	èndhog	endhog	ndhok	ĕndog
feather, fur	wulu	wulu/, ulu/	wulu	wulu, wulu-wulu"orang dari status sosial rendah"
wing	sewiwi	sewiwi/, siwi	swiwi	hĕlar, ĕlar, lar/ ahelar "bersayap"
rat, mouse	tikus	tikus	tékós	tikus,
tail	buntut	buntut	bóntót	buntut
snake	uló	ula	ulo	ulā
venom	wisa	upas, wisa, racun ula, semburé ula	upas	bisa, wişa, upas, racun, aŋracun/rinacun "meracuni"
leech	lintah	lintah, pacet, kelelet	lintah	lintah, pacĕt
hair lice	tumó	tuma	tumo	tuma
chicken lice	gurem	gurem, reki, tumané pitik	gurem	
nest	sarang/susoh*/ kurungan*/sang kar*	susuh, tarang, mimik (di tumpang)	susóh	susuhan/sĕsĕhan"sarang", susuh"little water snail", sarang "special tree"
mosquito	lemut/jingklong	nyamuk, lemut, merutu	nyamók	
spider	sawang #/ kólómónggó*	meluntu	kolomonggo	reŋit "fly",
cockroach	coro	coro, melinjé, godhèk, mblinjing, èmbès	córó	sawaŋ, labah: labuh "start"
fish	iwak	iwak, gerih (ikan asin)	iwak	
turtle	bulus	--	kuro	
water buffalo	kebo	kebo	kebo	iwak, gĕrih "salty"
land	lemah	lemah, lemah gaga	lemah	bulus "kind of weapon for stab",

				peñu
rock, stone	watu	watu	watu	kĕbo
sand	wedhi	wedhi, pasir/	wedhi	lĕmah
water	banyu	banyu	banyu	watu
sea	segóró	segara	laót	pasir "sea/beach",
lake	t(e)logo/ranu (lit)	telaga, segara, segara cilik	tlogo	bañu
river	kali (cilik)/bengawan	kali, sumber, banyu	kali	garĕ" status society"
forest	alas	alas	alashok	ranu,
mud	endhut	lédhok, bledhok, èndhut (utk halus)	blètok	kali, kali "dewi/ ruh jahat: batari durga, kali "jaman besi", sumber "sumber
mountain	gunung	gunung	gonong	alas, alas "mat"
hill	gunung cilik	pusung, pudhuk, gunung	--	ĕndhut, blĕdhog "kind of firecracker",
sun	srèngèngé	sengéngé	srengéngé	
moon	re(m)bulan	bulan	mbulan	pusuŋ, gumuk
star	lintang	lintang	lintang, bintang	
cloud	mèndhung	mendhung	awan, mendóng	wulan
fog, mist, haze	pedhut	pedhut, mendhung, kabut	kabót	lintang = wintang, lintang/halintang/a lintang "pass away"
rain	udan	udan	udan	1.awăn "tall, big", 2. awăn "street", 3. awăn "fish from sea", mendhung "black cloud, dark, water cloud", lamuk "fog", megha
thunder	gludhug	gludhug, guntur, longsor, bledhèk	glódók	pĕdhut
lightning	thathit	kilapan, bledhèk	kilat	udan = hudan
wind	angin	angin	angén	gludhug = gl**ĕdhĕg, guntur "lava flood" ,**
smoke	beluk *reg*	kukus, beluk	belók	kilat*
bark	kulité wit	gelumbré	uwét ?	1. angin 2. angin "status society" *
tree	wit	wit, kayu	wit	asĕp "dupa",

				kukus, abeluk " berteriak",
field	tegal	gaga	sawah	uwat "medicament"
dust	awu	bledhuk	blédhók	wit = wwit "origin, foundation, tree"
sky	langit	langit	langét	kĕbon=kĕbwan "kebun/perkebunan", tĕgal, sawah, gagā "sawah tanpa irigasi"
wood	kayu	kayu	kayu	blĕdog = bĕlĕdog " cara membuat alat musik", (h)awu, lĕbū
uncooked rice	beras	beras	beras	*langit
rice plant, still on ear	pari	pari	pari	kayu
meat	daging	iwak, daging	daging	bĕras lih bras
contents	isi	isi	isi	pari
fat	gajih/lemak	gajih	gajéh	daging
oil	lèngó	lenga	lengo	(h)isi
branch	pang	pang	pang	gajih (animal),
leaf	godhong	godhong	godong	lĕnga "wijen oil"
root	oyod	oyod, cengkar	oyot	1. pang, 2. pang lih mapang "
flower	kembang	kembang	kembang	godhong, dhadhang "fish dedang"
fruit	woh	woh	woh	ayat "plan to"
-fruiting tree				kĕmbang, kĕmbĕng "pool"
grass	suket	suket	suket	who:wwah (special for pinang)
fire	geni	geni	geni	wit = wwit "origin, foundation, tree"
ember	wowo*	wawa, barek geni, murup areng	areng	sukĕt
ash	awu	awu	awu	gĕni
island	pulo	--	--	arang"rare", wowo lih wwawa "gibbon", wangwa lih wangba,

				hawu,ewu: iwu "thousand"
fiancé(e)	tunangan	pacangan, gendhakan	?	pulo
female servant	bidoyo*/batur	buru	babu	
friend	kóncó	kanca, réwang	konco	lamar: anglamar:linamar, ikět "roof", pacang "candidat", bakalan "candidat"
child	laré/bocah	anak	anak	babu"mother", batur, buru "nunter/hunting", rewang:rowang "friend"
hearth, fireplace	tungku	pawon	pawonan	batur
sleeping platform	ambèn/paturon	ambèn, lincak, paturon, balih, kerdin	dipan	anak
-for corpse				
pillow	bantal	bantal	bantal	amběn lih hamběn "balut", paturon,
person, people	wong	wong	wong	
man	lanang	lanang	lanang	bantal, kampil "bag/pocket"
woman	wedok/wadon	wedok	wedok	wong, manusa

spouse	bojo	bojo	bójó	rabi, laki
younger sibling	adhi	adhik, adhèk	adék	
older brother	kangmas	kakang	mas	kakang, mās "gold"
older sister	mbakyu	iyah	mbak	
mother	ibu/mbok/biyang	mbok, biyung, mak	ibuk	ibu
father	bapak	bapak	bapak	bapa
grandmother	mbah (putri)/éyang	yungwèk, biyung, ninèk, wèk	mbah	wek=weh "particle", nini
little boy	nak	(tho)lé	tholé	nak, lek "month"
little girl	nak	ndhuk	ndók, gendók	nak "children", nduk lih duk "time(when)",

				lare=rare "chil(m/w)
house	omah	umah	omah	umah, griya=gṛha
roof	payon	wuwung(an), larap (ceiling): pyan, ampé-ampé	genthèng	payon=payu=pay wan(papaywan), wuwungan, larap "kerlap-kerlip",
door	lawang	lawang	lawang	lawang
wall	tèmbok	tèmbok, gedhong	tèmbok	tembok(?) lih tombok "pemberian tambahan (barter), kenaikan taruhan, taruhan", gedhong "bulid"
wall of bamboo	gedhèg	pager, gedhèg	gedhèk	gĕdhek: at hr dibaca godhek "bercambang", pagĕr "wall"
window	jèndhèló	sèndhéla	cendèlo	
stair, step	andha	andha	undak-undakan	andha, undhakan "horse"
neighbor	tónggó	tangga, dulur(an)	tonggo	(ta)tangga, dulur "friend"
kitchen	pawon	pawon	pawon	
name	jeneng	jenèng	jeneng	jĕnĕng "stand, berfungsi, penguasaan, pendirian tetap", aran=haran=ngara n,
needle	dom	dom	dom	dom
grime, filth	bolot	karak, bolot, kothé, lega, blédhok	bolot	bolot(ĕn) "dirty", karak(a) "crust", lĕga "free, satisfied, free of worry"
nail	paku	paku	paku	paku (= pakö "penetapan/pengu kuhan")
sick	lóró	lara	loro	lara, gĕring= gring
upset stomach	mules//pules	lara weteng	mules	lara weteng
headache	ngelu	lara èndhas	ngelu, mumet	lara ĕndhas= tĕndhas, puyĕng, ngĕlu

knife	lading	lading	ladèng	lading
pregnant	meteng//weteng	meteng	meteng	(a)mětěng
near	cedhak	parek, cedhek	cedhek	parěk, pěrěk=prěk=pědhěk,
far	(a)doh	adoh	adoh	
often	kerep	gelek, kerep	seréng	kěrěp "meet, mass", gělěk "fast, soon"
rarely	arang/soksok	arang-arang	jarang, kolo-kolo	arang
difficult	angèl	éwoh, sara, angèl	angèl	gati "street, tindakan", ngel=angel= ang(h)el "tired"
easy	gampang	gampang	gampang	gampang,
naughty	cerèwèt	memel, nakal, ngamak, mbongol	nakal	
native	asli	asli	asli	
left	kiwó	kiwa	kiwo	kiwa
right	tengen	tengen	tengen	těngěn
clean	resik	ricik	sesik	rěsik
dirty	kothor/reged	rusuh, uleg, reged, kothé, jembruk	rusuh	(a)rěrěd, rusuh "difficult", uleg=ulug "no perfect"
back, rear	mburi	mburi	mburi	buri=wuri, gěgěr
sharp	landhep	landhep, lincing	landhep	landhěp
dull, blunt	kethol*	kethul, papak	gak masah	kěthul,1. papak= pakpak "merayu, memuji" 2. papak "general"
broken	tugel	putung (belung), tokleg (kayu), pudul, tugel	pecah	tugěl, putung
putrid, rotten (smell)	bosok	bosok, mburuh, mambu, amis	tengik, banger	ambu=ambo "pengawal", "water", amis, bangěr
hot	panas	panas	panas	panas, běnter "bright, blinded"
cold	adhem/atis	adhem	adhem	adhěm lih dhěm "lonely, silent, calm"

dry	garing	garing, asat	garéng	gĕring "ill"
wet	telees	teles	teles	udan= hudan "rain", tĕlĕs, kĕpus
heavy	abot	abot	abot	
light	èntheng	ènthèng	ènthèng	
black	ireng	irèng	ireng	(h)irĕng
white	putih	putih	pótéh	putih
red	abang	abang	agang	(a)bang
yellow	kuning	kuning	kónéng	kuning
green	ijo	ijo	ijo	(h)ijo
little, small	cilik	cilik	cilik	cili(k)
big, large	gedhé	gedhé	gedhé	gĕdhe
short	cèndhèk	cendhek èndhèk	ndhèk	
long	dowo	dawa duwer	dowo	dawā, jangkung "kind of egret"
lazy, disinclined	wegah/males	wegah, males, kesèt, mbling	males	wĕgah=wugah "tall, very big",
thin	tipis	tipis	tipis	tipis=(an)ipis
thick	kandel	kandel	tebel, kandhel	(a)kanděl
narrow	ciyut	cubit, cubèk, subèk	ciut	sĕsĕk, sĕsök "fuul"
wide, spacious	ómbó	amba, jembar	ombo	amba, sĕmbar=sĕmbur "sprinkling", wera
shy, ashamed	isin	isin	isin	isin
old	tuwó	tuwèk	tuwèk	tuwa= tuha
young	nom	nom	nom	(a)nom, nom=nwam "kaum muda"
just (arrived)	mentas	èntas/tas	tas	ĕmbe=ĕmbih "crying", mĕntas "out from water", tĕmbe "first time"
long (arrived)	suwé	suwé	suwé	suwe lih sowe
new (things)	anyar	anyar	anyar	
old (things)	lawas	lawas, èlèk	lawas	lawas,suwe
good	apik/bècik	apik	apik	bĕcik, bagus lih wagus "handsome"
bitter	pait	pait	pait	pa(h)it
sweet	legi	legi	legi	lĕgi, manis
bad, evil	óló/èlèk	kereng, gethot, nyerèkèl, èlèk	jahat	jahat"shattered", (h)ala
slippery	lunyu	lunyu, alit	lunyu	lunyu, alit "soft"

full	kebak	akèh, bek (banyu)	kebek	kěběk, (a)keh=kweh "amount", běk
wound	tatu	jatu	tatu	tatu, jatu " latex", catu "measurement, portion"
salt	uyah	uyah	uyah	uyah
money	dhuit	picis	duwék	picis lih pisis
woman's head scarf	kudhung	kethu, gelung	kódóng	kudhung, kandhung"pregnant", kěthu, gělung "hair bun"
earthenware pot	pengaron	cèrèt, ngaron, omplong	pot	aru= haru "belanga for cooking rice",
please	mónggo	ayo, bacut/	tólóng	ayo lih haywa "do not",
excuse me	kula nuwun/nuwun sewu	amit, wis ya/	amét	
thank you	matur nuwun	matur suwun, nedan nrima, kesuwun	suwón	matur "salah satu istri raja',
how(‘s that?)	yo ópó/kepiyé	njaré	yok opo	
what	ópó	paran	opo	apa, paran
if	yèn	nèki/nèké	lèk	yen=yan, ning/gening "from/-nya", nek "climb", lek "month", kadhung "short"
because	mergi/sebapé	polahé, sebapé, merkané	sokalé	polah "behavior", paran "how"
very	banget	temen(an)	nemen	těměn, bangět "speed, current, flow", pisan "first",
snacks	camilan/nyamikan	panganan, rimikan, camilan, manisan	jajan	jajan "jenis ikan",
food	panganan	panganan	panganan	pangan
night	bengi/wengi	bengi	bengi	wěngi
yesterday	wingi	wingi	wingi	wingi
two days ago	wingènané	wingènané	wingènané	wingi

tomorrow	sésuk	sésuk	sésók	
two days from now	sésuk-èmbèn	èmbèn	mbèn	ĕmben
year	taun	taun	taón	tun= ton "imperatif"
when	suk/dhèk kapan	kapan	kapan	kapan
before, first, earlier	ndhisik	biyèn, (di)kit, dhisik	sakdurungé	tuli "fish", tuli "deaf", durung "not yet"
afterwards, then	(ba)njur	maringono, mari, singkari, sawisé terus	sakwisé	wis "allgone, finished", bañjur "continue", mari "maybe", tĕrus lih trus "sharp, to stick"
day	dinó	dina	dino	awan "1.when sun is very tall, way", dina
where	nèng èndhi	nèng èndhi	ndhi	ĕndi=ndi, nĕng=hĕnĕng "clearness", ning=ni "particle conjuction"
who	sópó	sapa	sopo	s(y)apa
I myself	aku dhéwé	éyang/isun dhéwé	aku déwé	dhewek, aku, i(ng)sun, kita "2sg"
different/other	liyó	liya, sici/1	liyo	
all	kabèh	kabèh	kabèh	kabeh
and	lan	karo, lan	karó	karo=karwa "ant", kambe=amben "bed",ambi
why	kok/kénó ópó	ngapa, njare	kenopo	kĕna(a)pa, (ng)apa, kĕna(ng)apa,
not yet	durung	urung	dóróng	durung, urung=wurung "fail"
not (nominal)	dudu	dudu	gak ono	dudū "be different", duduk lih duk "a stab", gak ana "gak:particle onomatop, ana=hana : there's"

not	ora	ora(k)	gak	ora, gak "particle onomatop"
yes	(i)yo	he-eh	iyo	(h)iya "seriously"
how many	piró	pira	piro	pira, pirĕ= amirĕ "sheer off"
one	siji	siji	siji	siji, (h)iji "name of bird like manyar"
two	loro	loro	loro	kalih
three	telu	telu	telu	tĕlu
four	papat	papat	papat	papat lih pat
five	limó	lima	limo	lima, limo "orange"
21	selikur	selikur	selékór	
25	selawé	selawé	selawé	(sĕ) lawe "thread"
50	seket	seket	sèket	sekĕt
60	sewidak	sewidak	sewidak	(sa)widak
there	kónó	kana	kono	kana, riku "arrange", rike "here"
here	kèné	kèné	kéné	(ng)kene
there aren't	ora ónó	nana	gak ono	laka "red", langka "functionary", ora ana, nanā "broken"
suddenly	moro-moro		moro-moro	(a)dadak
road	dalan	èmbong	dhalan	dalan, gili "isle",
turn	mbelok/menggok		mènggok	putĕr
swim	nglangi		nglangi	inggĕk=minggĕk "rotate"
guts	usus		osos	jĕro-an "inside", usus
afraid	wedi	wedi	wedhi	wĕdi
laugh	ngguyu	ngguyu	ngguyu	guyu
chew	mamah		mamah	kĕmah
suck	nyedhot		mót	(h)isĕp
yawn	angop		mbengok	
rope	tali	tali	nalèni	tali, kambang "float", tambang
tie	nalèni		nalèni	ikĕt, tali, jirĕt=jirĕk"kind of tree"
shoot	nèmbak/mbedhil		tèmbak	(ti)nembak, (a)mbĕdhil
stab			jojoh, cubles (needle)	

hit, beat	ngantem		antem	(h)antěm "bed smell", (ng)anu "someone"
scratch	nggaruk		garók	(a)kukur
split	misah, mbagi		suwèk	sigar, kěthok, (a)misah,
choose	miléh		péléh	(a)pilih
squeeze			peres	padhět "solid"
dig up	ngedug		kedók, dódók	(ang)dhudhuk, (a)ngědhuk,
earthworm	cacing	cacing	cacéng	cacing
blow	nyebul		seból	damu
correct, true	bener	bener	bener	běněr
hide	ndhelik		singitan	umpět=upět ", singit "incline"
climb	munggah	mèntas, mènèk (wit)	pénék	munggah, (a)menek, měntas "out from water",
at	nèng	ndhèk	ndék	něng=hěněng=ěně ng "quiet", ning=ni "conjuction particle"
in, inside	njero	njera	ndék jeró	jěro
above	nduwur	nduwur	ndék nduwór	dhuhur
below	ngisor	ngisor	ndék mburi	isor=iswar
this	iki	iki	iki	iki
that	kuwi, iku	iku	iku	iku
count	ngitung	ngitung	étóng	(a)ngitung, initung,
want/will	arep	katé	katé, gelem	gělěm, pan-apan "because", (h)arěp "front"
come in	mlebu	mlebu	mlebu	luměbu

* not in Zoetmulder
# differs from OJ meaning in Zoetmulder
% different phonological form