What Madurese Can Tell Us About Other Languages and What Other Languages Can Tell Us About Madurese

William D. Davies
Department of Linguistics at the University of Lowa

Abstract: There are many reasons to study a particular language. One might study a language to learn to speak that language in order to be able to communicate with native speakers or conduct business. One can study a language to learn about the culture of the people or understand their literature and poetry. One can study the grammar of a language to understand the structure of that language and compare it to the structure of other languages It is this last reason that motivates the following discussion.

Key words: Causative constructions, cross-linguistic comparison, Innateness Hypothesis, Javanese, Madurese, morphology, prolepsis, raising, syntax, universal grammar, word order.

A common question addressed to me is "Why do you study Madurese? Why don't you study Javanese?" The question of why to study Madurese is essentially the question of why to study any and all languages. It has been said that language is the window into the human mind. And it is clear that of all animal communication systems, language is by far the most highly articulated and most complex. It is a uniquely human characteristic. To understand the structure of language may provide a key to understanding the human mind. Much of modern linguistics is guided by what has come to be known as 'the logical problem of language acquisition' (Hornstein & Lightfoot, 1981). This refers to the fact that despite being given very little direct instruction, all children who are born without any physical or mental defects learn their first language remarkably quickly. By the time they are about five years old, they have mastered a large vocabulary and the majority of the morphological and syntactic structure of their native language. They

learn the language through observation and trial and error, but the error is relatively limited compared to adults attempting to master a second or additional language. This remarkable fact has led modern linguistics to adopt a hypothesis developed by Noam Chomsky in the 1950s (e.g., Chomsky 1959) commonly referred to as the Innateness Hypothesis. According to the Innateness Hypothesis humans are born with certain capacities for language and that is why mere exposure to meaningful language enables a child to learn such a complicated system so quickly.

The Innateness Hypothesis includes the premise that there are universal organizing principles that all languages obey. While languages can differ from one another considerably, there are universal aspects to the structure of language--structural aspects that all languages share. Within the modern traditional, this is referred to as 'Universal Grammar' or UG. The question facing linguists operating within this paradigm is how to gain insight into the principles guiding the process of language acquisition. Put differently: What is included in Universal Grammar? The general methodology is to compare the structure of a variety of human languages. By seeing how languages are alike and how they are different, we can extract generalizations about languages that provide us with the information we need to develop principles that help us understand restrictions on the form that languages can take. That is, we can develop principles that show how language is organized, which takes us a step closer to discovering how language works in the brain.

Some generalizations regarding language seem fairly obvious. To take a phonological example, there appears to be no human language that includes the sound associated with belching as part of the meaningful inventory of sounds that make up the words, phrases, and sentences of a language. On the syntactic side, it appears that reversing the order of every word in a sentence is not a process than any human language employs to form a question. Thus, no language has yet been discovered in which "Book the read Siti" would mean "Did Siti read the book?" While these may seem like ridiculous examples, it is important to note that there is no *a priori* reason that these should not be part of human language. It is simply the case that they are not. They are outside the bounds of the principles that constrain what a possible human language is.

CROSS-LINGUISTIC COMPARISON

At the crux of developing our Universal Grammar is cross-linguistic comparison. That is, we must understand the structure of as many languages as possible in order to form and test our hypotheses. Obviously, the reason that we know that no language includes belching sounds in its phonological inventory and that no language includes complete sentence reversal in question formation is because we have examined many languages. If we just consider the Indonesian language, we know that it is true that these are not part of the grammar of Indonesian. However, there is no guarantee that our conclusion might simply be an accident of basing this conclusion on consideration of this one particular language. Only by examining a large number of languages can we safely hypothesize that this is true of human language in general.

To take a potentially more interesting example, it is generally assumed that we can describe the basic word order of a simple declarative sentence in terms of the notions 'subject', 'object', and 'verb'. Considering a small group of languages, we can see that there are many that have the same basic word order. So, in English, Indonesian, French, and Madurese, we find that the basic word order for a simple declarative transitive sentence is subject-verb-object (SVO).

(1) English   :   Ali sees Hasan.
     Indonesian :   Ali melihat Hasan.
     French   :   Ali voit Hasan.
     Madurese   :   Ali neggu Hasan.

However, as is well known, not all languages have the same basic word order as the four in (1). Choctaw, an American Indian language spoken in the states of Oklahoma and Mississippi, like many other languages, has the basic word order of subject-object-verb (SOV), the same basic word order as Japanese.

(2) Choctaw   :   Aliyat Hasan pisatok.
     Japanese :   Ali wa Hasan o mimasita.

Still other languages have other word orders. So, Malagasy, an Austronesian language spoken on Madagascar, has a basic

verb-object-subject (VOS) order, and the Philippine language Cebuano has basic verb-subject-object (VSO) order.

(3) Malagasy   : Nahavory ny ankizy ny mpampiantra.
                      gathered the children the teachers
                      'The teachers gathered the children.'
     Cebuano   : Nagpalit si Linda ug dulsi.
                      bought Linda candy
                      'Linda bought candy.'

Therefore, in examples (1-3) we see one way in which languages are similar: the basic word order of a transitive sentences can be characterized in terms of subject, verb and object. We also see a way in which languages can differ: it is not the case that in all languages subject, verb and object are ordered the same way.

There are other generalizations that can be extracted regarding the word order of human languages. Through careful and painstaking examination of a vast corpus of languages, the American linguist Joseph Greenberg discovered a universal tendency: languages in which the object follows the verb almost exclusively have prepositions rather than postpositions, while languages in which the object precedes the verb are almost exclusively postpositional (Greenberg, 1966). Thus, despite the fact that basic word order in English, Indonesian, French, and Madurese is subject-verb-object, in Malagasy verb-object-subject and in Cebuano verb-subject-object, all of these languages have prepositions, as in (4), where the preposition is given in boldface.

(4) English   : Ali wrote a letter **to** Hasan.
     Indonesian : Ali menulis surat **kepada** Hasan.
     French   : Ali a ecrit une lettre **à** Hasan.
     Madurese   : Ali noles sorat **dha**' Hasan.
     Malagasy   : Nividy mofo **ho an**'ny ankizy aho.
                      bought bread for the children I
                      'I bought bread for the children.'
     Cebuano   : Nagpalit si Linda ug dulsi **para** Lito.
                      Bought Linda candy for Lito
                      'Linda bought candy for Lito.'

Greenberg's principle predicts that Choctaw, one of our languages in which the object precedes the verb, should have postpositions, not

prepositions; and this, in fact, is the case. Japanese behaves likewise. (In the Japanese example, *wa* is a topic marker on the subject and *o* marks the object of the sentence.)

(5) Choctaw : Ali Hasan **anshaka** hikiyatok.
  Ali Hasan behind stood
  'Ali stood behind Hasan.'
  Japanese : Rakoto wa kono senzai **de** ifuku o sentakusita.
  Rakoto TOPIC this detergent with clothes OBJ washed
  'Rakoto washed clothes with this detergent.'

Again, once pointed out, these are fairly obvious generalizations, accessible through fairly casual but careful observation of a large number of languages. However, unlike the previous belching and sentence reversal generalizations, they tell us something interesting about the structure of human languages. Yet other types of generalizations require truly detailed descriptions of particular languages.

INTRA-LINGUISTIC VARIATION AND CROSS-LINGUISTIC GENERALIZATION

At times, the type of cross-linguistic generalizations sought within the modern paradigm can inform variation within a single language. Madurese causatives offer a striking example of this. Every language has a means for expressing ideas such as "Ali made the baby cry", "Bambang made Siti sad", "Siti made Marlena slice the bread", "Mother made the children put the toys in a box", and so on. In Madurese, in many cases the prefix *pa-* is used to form causatives. (The prefix occurs as *ma-* in the actor voice in the following examples.) So, we find sentences such as[1]

(6) Ali *ma*nanges baji'.
  A AV.CS.cry baby

---

[1] The following abbreviations are used in the morphemic glosses of the examples:

AV = actor voice	NOM = nominalization
COMP = complementizer	OV = object voice
CS = causative	PERF = perfective
DEF = definite	RED = reduplication
UT = future	

'Ali made the baby cry.'
(7) Bambang *ma*sossa Siti.
     B AV.CS.sad S
     'Bambang made Siti sad.'

These are sentences that all speakers agree on. But differences arise when the situations include more participants. So, we get different information about the next sentences.

(8) Siti *ma*ngerra' Marlena rote rowa.
     S   AV.CS.slice M bread that
     'Siti made Marlena slice the bread.'
(9) Siti *ma*nyaba' ana'eng en-maenan e dhalem kothak.
     S   AV.CS..put child.DEF RED-play.NOM at inside box
     'Siti made her children put the toys in a box.'

Speakers split on the acceptability of sentences such as (8) and (9). Some speakers accept these as well-formed sentences, while others reject them, preferring the periphrastic causatives in (10) and (11).[2]

(10) Siti nyoro Marlena ngerra' rote rowa.
     S AV.order M AV.slice bread that
     'Siti ordered Marlena to slice the bread.'
(11) Siti nyoro ana'eng nyaba' en-maenan e dhalem kothak.
     S AV.order child.DEF AV.put RED-play.NOM at inside box
     'Siti order her children to put the toys in a box.'

Finally, some speakers consulted accept (8) as grammatical but reject (9), preferring (11).

These causative data are very stable, the speakers unwavering in their judgments. However, rather than simply being inconsistent or perhaps even appearing to be chaotic, the causative data are actually quite interesting from the standpoint of cross-linguistic comparison, and what we know about other languages can help explain this intra-linguistic variation.

The base predicates involved in these data differ from one another. Predicates such as *nanges* 'cry' and *sossa* 'be sad' take one participant,

---

[2] Speakers may also accept predicates other than *soro* as the 'causative' predicate, such as *gabay* 'make'.

they are intransitive or one-place predicates. The verb *ngerra'* 'slice' is a transitive or two-place predicate, taking both a subject and an object. The verb *nyaba'* 'put' is a 3-place predicate, requiring a subject, and object, and a location. Thus, the data can be distinguished in terms of the transitivity of the base predicate. All speakers accept using the causative prefix with intransitive predicates, as in (6) and (7). Then there are differences. Some speakers allow the prefix only with intransitive predicates, while others allow it with intransitive and transitive predicates only. Still still others use the prefix completely productively and accept (11), where the base is a 3-place predicate. This results in a hierarchy of cases.

(12) Bases used with causative *pa-*
intransitive (1-place) > transitive (2-place) > ditransitive (3-place)

Different speakers divide the hierarchy in different ways, which can be illustrated as follows (where '|' indicates the point at which the hierarchy is divided):

(13) Case A
accepts (6) & (7) but not (9) or (10)
intransitive (1-place) | > transitive (2-place) > ditransitive (3-place)

(14) Case B
accepts (6), (7) & (8) but not (10)
intransitive (1-place) > transitive (2-place) | > ditransitive (3-place)

(15) Case C
accepts (6), (7), (8) & (9)
intransitive (1-place) > transitive (2-place) > ditransitive (3-place)

Thus, there is a regularity in the data.

But not only can we discover regularity in the data and incorporate that into a description of Madurese grammar, these data are interesting from a cross-linguistic standpoint because there are languages that operate in the same way represented by each of the cases in (13-15). Many languages have the kind of causative affix found in Madurese. And there are languages that split up the hierarchy in (12) precisely the same way that different speakers of Madurese do. For example, in Seri, an American Indian language of Mexico, the causative morpheme can only

be added to intransitive predicates. Thus, Seri is just like Case A. Choctaw is just like Case B: the causative suffix can be added to intransitive and transitive bases but not to ditransitive bases. Finally, the Philippine language Tagalog is just like Case C. The causative prefix, which like Madurese happens to be *pa*, can be added to any verbal base. Thus, what at first looks like intractable data in Madurese, actually reflect a larger regularity and can provide interesting facts that help build significant linguistic generalizations.

INTRA-FAMILY COMPARISON

Detailed description of Madurese can also provide interesting comparisons within the family of closely related languages, in this case Javanic languages (Nothofer, 1975) such as Balinese, Indonesian, Javanese, and Sundanese. Here I will concentrate on comparison with Javanese. There are a number of obvious parallels between Madurese and Javanese, and clearly there has been mutual grammatical influence over the centuries, given the close proximity and interaction among the speakers and the shared linguistic ancestry. This is obvious through cursory examination and lexical comparison; a large number of lexical items are shared between the two languages, particularly in the higher speech levels. And Davies (1999) provides evidence for similarity in word order configurations.

However, perhaps the most conspicuous feature of Austronesian languages is the morphosyntactic system for identifying the semantic role of the most prominent argument in a given clause. This argument has been referred to as "subject", "topic", and "focus" of the clause, what Starosta (1986) succinctly characterizes as the "perpetual centre of the sentence". Since this is not the issue of interest here, I will simply refer to it as the subject. The literature on Javanic languages generally recognizes two basic functional sets of verbal prefixes. The first is variously referred to as "active", "actor focus", "actor voice", and others. The second has been referred to as "passive", "nonactor focus", "object focus", "object voice", and others. For a number of reasons that have been spelled out by Bintoro 1980, Cumming 1986, Naylor 1978, Thomas 1980, and others, there are good reasons not to use the terms "active" and "passive". So, I will follow current use widespread in the Austronesian

literature and refer to these as actor voice (AV) and object voice (OV) in the discussion here.

The Madurese system can profitably be compared to systems in the other Javanic languages because of the great similarity of the systems. But the differences between the languages in regard to the verbal morphology is instructive as well. Here I will compare certain aspects of the Madurese system with that of Javanese, as described in Bintoro 1980.

In Madurese, actor voice occurs on syntactically transitive verbs and some intransitive verbs when the agent or actor of the clause is the subject. The morphological manifestation of actor voice is either a nasal consonant or the prefix *a-*. The choice of the nasal or *a-* prefix appears to be largely a lexical idiosyncrasy (although its use is more widespread in the eastern variety of the language). As is true of the actor voice prefix in Javanese, the nasal consonant assimilates to the place of articulation of the initial consonant of the verb root. Additionally, all root-initial obstruents are deleted. Unlike Javanese, in which the deletion occurs only with voiceless obstruents, all three series of consonants (voiceless, aspirated, and voiced) delete in the presence of the nasal actor voice prefix, as is obvious in the examples in (16).

(16) Verb Root   Actor Voice

    enom 'drink'   *ngenom*
    rosak 'ruin'   *ngrosak*
    baca 'read'   *maca*
    toles 'write'   *noles*
    kera 'think'   *ngera*
    bundhu' 'wrap'   *mundhu'*[3]
    semprot 'spray'   *nyemprot*

The prefix *a-* marks actor voice for a variety of syntactically transitive verbs. While Stevens (1968) reports that *a-* is used predominantly with roots with initial aspirated and voiced consonants, there seems to be a great deal of dialectal and individual variation. The roots in (17) exemplify some of those that generally take the *a-* form.

(17) Verb Root   Actor Voice

---

[3] Although not obvious from the orthography, the initial consonant of *bundhu'* 'wrap' is a voiceless aspirated bilabial stop. This example is from Stevens 1968.

berri' 'give'	*aberri'*
temmo 'meet'	*atemmo*
gabay 'make'	*agabay*
jaling 'look at'	*ajaling*
sasa 'wash'	*asasa*

There are also verb roots that take either prefix, as in (18).

(18) Verb Root   Actor Voice

kerem 'send'	*ngerem* or *akerem*
bukteagi 'prove'	*mokteagi* or *abukteagi*[4]
bukka' 'open'	*mokka'* or *abukka'*

As stated above, in addition to the difference between Madurese and Javanese with respect to the widespread use of *a-* as a marker of actor voice, the manifestation of the nasal prefix differs somewhat. In Javanese, the nasal prefix assimilates to the place of articulation of obstruents, but only replaces voiceless obstruents such as *p, t, th, c, k,* and *s*. With the voiced varieties, the initial consonant of the verb root does not delete. This is illustrated in (19), which come from Bintoro 1980.

(19) Verb Root   Actor Voice

pedhot 'cut'	*medhot*
bayar 'pay'	*mbayar*
tumpes 'exterminate'	*numpes*
deleng 'look at'	*ndeleng*
thuthuk 'strike'	*nuthuk*
dhidhik 'educate'	*ndhidhik*
cokot 'bite'	*nyokot*
jupuk 'take'	*njupuk*
sawang 'look at'	*nyawang*
kirim 'send'	*ngirim*
gered 'pull'	*ngered*

---

[4] As a general rule, high vowels follow voiced obstruents. Thus, there is a vowel alternation apparent in *mokteagi~abukteagi* and *mokka'~abukka'* However, there is also speaker variation, and for some the high vowel of the root perseverates in the actor voice form resulting in *mukteagi* and *mukka'*.

Aside from these morphological differences, the use of actor voice in Madurese and Javanese are quite similar. However, an important difference is that in Madurese the use of the actor voice is obligatory with syntactically transitive verbs. In Javanese, the use of the actor voice prefix alternates with the absence of the prefix depending on whether or not the object is specific or not. This is evident in (20), where the variant with the actor voice morphology takes a specific object (20a) and the prefixless variant takes a generic (or nonspecific) object (20b) (from Bintoro 1980).

(20) a.. Jono nuku sepeda-ku.   (*Jono tuku sepeda-ku.)
      J AV.buy bicycle-my
      'Jono bought my bicycle.'
   b. Jono tuku sepeda.   (*Jono nuku sepeda.)
      J buy   bicycle
      'Jono bought a bicycle.'

As these examples show, use of the unprefixed form with a specific object is ungrammatical (20a) and use of the prefixed form with a nonspecific object is deemed ungrammatical (20b). In the Madurese counterparts, the verb must take the actor voice prefix regardless of the specificity of the object. When it does not, the sentence is ungrammatical.

(21) a. Ali melle tangsapedha.   (*Ali belli tang sapedha.)
      A AV. Buy my   bicycle
      'Ali bought my bicycle.'
   b. Ali melle sapedha.   (*Ali belli sapedha.)
      A AV.buy bicycle
      'Ali bought a bicycle.'

Thus, comparing the use of actor voice provides an indication of some of the subtle ways Madurese and Javanese differ from one another, despite the strong similarities in their syntactic and morphological systems.

Interesting and subtle differences also emerge when comparing the object voice morphology and the verbal suffixes in the two languages. However, limitations of space and the focus of the present discussion necessitate that such investigations be left to another forum.

WHAT MADURESE MIGHT TELL US ABOUT OTHER LANGUAGES

Finally, Madurese may be able to provide insights into the proper analysis of constructions in other languages. One area in which this is true is with respect to what is referred to in the literature as Raising-to-Object. Raising-to-Object is illustrated by the two English sentences in (22).

(22) a. Hasan believes that Ali bought a new car.
b. Hasan believe Ali to have bought a new car.

The sentence in (22b) is the Raising-to-Object structure, and this construction has a long and storied history in generative grammar. In early generative treatments, such as Rosenbaum 1967 and Postal 1974, (22b) was derived from a structure such as (22a) by "raising" *Ali*, the subject of the complement clause (*that Ali bought a new car*), to object position in the matrix clause. Chomsky (1973) argued that this is the wrong analysis and that *Ali* in (22b) is the subject of an embedded infinitival clause. This analysis persisted in the writings of Chomsky and his close associates until a theoretical shift in the early 1990's, at which time the Raising-to-Object analysis was reinstated (Lasnik & Saito, 1991).

Of relevance to Javanic languages is the fact that Balinese, Indonesian, and Javanese sentences such as those in (23-25) respectively have been analyzed as instances of Raising-to-Object or its equivalent in other theoretical frameworks.

(23) Balinese (Wechsler & Arka 1998)
Polisi tawang=a lakar ngangkep Wayan.
Police OV.know=3 FUT AV.arrest W
'He knew the police would arrest Wayan.'

(24) Indonesian (Chung 1976)
Mereka anggap buku itu sudah dibaca oleh Ali.
They believe book that PERF OV.read byA
'They believe this book to have been read by Ali.'

(25) Javanese (Davies 1990)
Amir ngarepna Musa arep menyang dina Selasa.
A   AV.expect M will to day Tuesday

'Amir expected Musa to go Tuesday.'

The Madurese counterpart is a sentence such as (26).

(26) Hasan ngera Ali melle montor anyar.
  H AV.think A AV.buy car new
  'Hasan thought Ali bought a new car.'

There is good reason to believe that the grammar of Madurese does not contain Raising-to-Object, and that sentences such as (26) and those that might be analyzed as Raising-to-Object are in fact a different construction. The proper analysis of Madurese may shed light on the proper analysis of other Javanic languages as well as languages outside the immediate language family.

"Raising" nonsubjects

The first piece of evidence is that the "raised" NP need not be a subject. In (26) *Ali* does play the role of subject in the embedded clause. However, unlike Raising-to-Object in English and most other languages for which it has been proposed, the "raised" element can bear any of a number of grammatical functions. In (27) *Hasan* is the possessor of the object of the embedded clause, *ana'* 'child', and in (28) *Hasan* is the possessor of the subject of the embedded clause.

(27) Siti ngera Hasan ja' dokter juwa mareksa ana'eng.
  S AV.think H COMP doctor that AV.examine child.DEF
  'Siti thinks that the doctor examined Hasan's child.'

(28) Marlena abala'agi Hasan ja' embi'eng ngekke' Ali.
  M AV.say.AGI H COMP goat.DEF AV.bite A
  'Marlena said that Hasan's goat bit Ali.'

More literal translations of (27) and (28) are 'Siti thinks about Hasan that the doctor examined his child' and 'Marlena said about Hasan that his goat bit Ali', respectively. And in (29) the embedded direct object has "raised".[5]

(29) Siti ngera Hasan ja' dokter juwa mareksa.
  S AV.think H COMP doctor that AV.examine
  'Siti thinks that the doctor examined Hasan.'

---

[5]There is speaker variability with respect to the grammaticality of (28). As shown shortly, all speakers will accept the sentence with a pronoun coreferential with *Hasan* in the embedded object position.

While there have been proposals that nonsubjects can raise to object in a few languages Niuean (Seiter 1983) and Kipsigis (Jake & Odden 1979)), these cases are exceedingly rare and perhaps bear closer scrutiny. Additionally, there appear to be no proposals for possessors raising directly from embedded clauses to matrix clauses.

Coreferential pronouns

Second, unlike familiar instances of Raising-to-Object, a pronoun coreferential with the "raised" NP can occur in the base position in the embedded clause. In (30-33), the raised element is coindexed with the pronominal *aba'eng* in the complement clause.

(30) Hasan ngera Siti$_I$ ja' aba'eng$_I$ melle montor.
    H AV.think S COMP she AV.buy car
    'Hasan thinks Siti bought a car.'
    lit: 'Hasan thinks about Siti that she bought a new car.'

(31) Siti ngera Hasan$_I$ ja' dokter juwa mareksa aba'eng$_i$.
    S AV.think H COMP doctor that AV.examine he
    'Siti thinks that the doctor examined Hasan.'
    lit: 'Siti thinks about Hasan that the doctor examined him.'

(32) Siti ngera Hasan$_I$ ja' dokter juwa mareksa ana'eng aba'eng$_i$.
    S AV.think H COMP doctor that AV.examine child.DEF he
    'Siti thinks that the doctor examined Hasan's child.'
    lit: 'Siti thinks about Hasan that the doctor examined his child.'

(33) Marlena abala'agi Hasan$_I$ ja' embi'eng aba'eng$_i$. ngekke' Ali.
    M AV.say.AGI H COMP goat.DEF he AV.bite A
    'Marlena said that Hasan's goat bit Ali.'
    lit: 'Marlena said about Hasan that his goat bit Ali.'

The sentences in (30-33) would be instances of copy raising, as proposed for Greek by Joseph (1976). Just as with the raising of nonsubjects, copying raising is quite rare in the world's languages and the data leading to such proposals may well be open to alternative analyses. Additionally, it has been argued extensively that copy raising affects only complement subjects (Moore 1998, Potsdam & Runner 2001). Thus, a raising analysis of the Madurese data would be unusual on a number of counts.

Idioms

Casting further doubt on the raising analysis is the behavior of idioms in this construction. It has been widely observed that parts of idiomatic expressions that are involved in Raising-to-Object structures can retain their idiomatic meaning. For example, the English expression in (34) has both a literal and an idiomatic interpretation.

(34) The cat is out of the bag.

In (34) *the cat* can refer either to an actual animal that is no longer contained in a bag or it can refer to a secret in the idiomatic interpretation "The secret is known". In the Raising-to-Object structure in (35), *the cat* can retain its idiomatic meaning of *the secret*, even when it is the subject of a passive form (35b). This is not true in (36), where *the cat* can only have its literal sense; and (36) is a very odd sentence even then, since cats are very difficult to persuade of much of anything.

(35) a. Siti believes the cat to be out of the bag.
　　 b. The cat is believed by Siti to be out of the bag.
(36) a. ?Siti persuaded the cat to be out of the bag.
　　 b. ?The cat was persuaded by Siti to be out of the bag.

The distinction between the behavior of (35) and (36) with respect to idioms is one of the principle ways of distinguishing between structures that are Raising-to-Object from structures that are not. This is because even when a part of the idiom has moved out of the clause containing it, the idiomatic meaning can be retained.
This is not true of the Madurese construction. Consider the idiom in (37)

(37) Nase' la daddi tajjin.
　　 rice PERF　become porridge
　　 'It is too late to do anything about it.'
　　 lit. 'The rice has become porridge.'

As (38) shows, when wholly contained in the complement clause, this string can be interpreted idiomatically. It is clear that it is wholly contained in the complement clause since it follows the complementizer *ja'*.

(38) Siti ngera bari' ja' nase' la daddi tajjin.
　　 S AV.think yesterday COMP rice PERF become porridge

'Siti thought yesterday that it is too late to do anything about it.'
lit: 'Siti thought yesterday that the rice had became porridge.'

However, when *nase'* 'rice' appears in the matrix clause, the idiomatic meaning is no longer available.

(39) Siti ngera nase' bari' ja' la daddi tajjin.
    S AV.think rice yesterday COMP PERF become porridge
    'Siti thought about the rice yesterday that it became porridge.'

In (39), *nase'* precedes both the matrix adverb *bari'* 'yesterday' and the complementizer *ja'*, showing that it is a constituent of the matrix clause here. The sentence in (39) does not allow the idiomatic reading but admits only the literal interpretation. Thus, idiom chunks operate in a way inconsistent with the predictions of a raising analysis. The pair of sentences in (40) shows the same effect with the idiomatic expression *ajam atellor e beras* "S/he has it easy' (literally: 'A chicken laid an egg in the rice').

(40) a. Siti namtoagi ja' ajam atellor e beras.
      S AV.certain COMP chicken AV.egg at rice
      'Siti is certain that he/she has it easy.'
      lit: 'Siti is certain that the chicken laid an egg in the rice.'
   b. Siti namtoagi ajam ja' atellor e beras.
      S AV.certain chicken COMP AV.egg at rice
      'Siti is certain about the chicken that it laid an egg in the rice.'

In (40a), the referent of the idiom is established through discourse context. However, when *ajam* 'chicken' occurs in the matrix clause in (40b), it loses any idiomatic meaning.

"Raising" triggers

Another argument against the Raising-to-Object analysis is the fact that the construction in Madurese is very productive. While Raising-to-Object in English and other languages is only possible with a relatively small number of predicates, the Madurese construction includes the majority of verbs that take sentential complements. Thus, alongside example with the predicates *kera* 'think', *bala* 'say', and others already illustrated, the structure is possible with *yaken* 'sure', *loppa* 'forget',

lapor 'report', *kabaragi* 'spread the news', *terrangngagi* 'explain', *janji* 'promise', *tao* 'know', *koto'* 'whisper', *bukteyagi* 'prove', and virtually any other predicate that can take a clausal complement.

Prepositional objects

Finally, the embedded NP that occurs in the matrix in the Madurese construction can occur as a prepositional object in the matrix clause. In each case illustrated thus far, it is possible for the NP to occur as the object of *parkara* or *halla* 'about' or in some instances *ka* 'to'. Examples are given in (41-43).

(41) Siti ngera parkara Hasan$_I$ ja' dokter juwa mareksa aba'eng$_i$.
S AV.think about H COMP doctor that AV.examine he
'Siti thinks about Hasan that the doctor examined him.'

(42) Marlena abala parkara Hasan ja' embi'eng ngekke' Ali.
M AV.say about H COMP goat.DEF AV.bite A
'Marlena said about Hasan that his goat bit Ali.'

(43) Bambang yaken parkara Ita$_I$ ja' Ali aberri' buku dha' aba'eng$_i$.
B sure about I COMP A AV.give book to she
'Bambang is sure about Ita that Ali gave her a book.'

In each case here, the matrix prepositional object is coreferential with an embedded NP.

An analysis

A Raising-to-Object analysis of the Madurese construction is contraindicated by the fact that (i) the targets of raising are not exclusively subjects, (ii) embedded coreferent pronouns are possible in all cases, (iii) idiomatic meanings are not preserved, (iv) there is no apparent lexical restriction on matrix predicates, and (v) the raised element may occur as a prepositional object in the matrix clause. The Madurese construction closely parallels English construction in (44).

(44) John assumed about Mary$_i$ that Ted had warned her$_i$.

The English and Madurese constructions, which Higgins (1981) refers to as 'prolepsis', both typically require the matrix argument be coindexed with an embedded argument. The appearance that the

Madurese construction might be raising is due to the possibility of zero pronouns in the language, an option unavailable in English.

Madurese freely admits zero pronominals in the appropriate discourse context, especially as subjects or possessors. The option is also available in direct object position, although there is some speaker variability in the acceptability of this. So when the embedded NP in prolepsis is a subject or possessor it looks very much like raising, but in fact the structures with and without overt pronouns are precisely the same, as in the similarity between (45) and (46).

(45) Siti ngera Hasan$_i$ [ja' dokter juwa mareksa ana'eng *pro$_i$*]
(46) Siti ngera Hasan$_i$ [ja' dokter juwa mareksa ana'eng aba'eng$_i$]
    'Siti thinks that the doctor examined Hasan's child.'
    lit: 'Siti thinks about Hasan that the doctor examined his child.'

In (45) a zero pronoun possessor (symbolized here as *pro*) is coindexed with the matrix NP *Hasan* and in (46) the overt pronminal *aba'eng* is coindexed with *Hasan*.

The fact that the matrix NP can occur as a prepositional object is the key to the full analysis of the Madurese. The base form of the structure includes matrix PPs. Base-generation of the NP in a PP not only accounts for the obvious cases in which these NPs occur as prepositional objects, but also accounts for verbal morphology. In Madurese the suffixes *-agi* and *-e* can occur in certain environments when a prepositional object surfaces as a bare NP. This is illustrated for *-agi* in (47) and *-e* in (48).

(47) a. Ita melle buku kaangguy Bambang.
       I AV.buy book for B
       'Ita bought a book for Bambang.'
    b. Ita melleagi Bambang buku.
       I AV.buy.AGI B book
       'Ita bought Bambang a book.'
(48) a. Bambang ngerem paket dha' Siti.
       B AV.send package to S
       'Hasan sent a package to Siti.'
    b. Bambang ngereme Siti paket.
       B AV.send.E S package
       'Hasan sent Siti a package.'

In (47a) the beneficiary *Bambang* occurs as the object of *kaangguy* 'for'. When it occurs as the bare NP in (47b), the verb must be suffixed with *-agi*. In (48b), the prepositional object of (48a), *Siti*, occurs as a bare NP and the verb obligatorily includes the suffix *-e*.

These same suffixes occur with some verbs when the matrix NP in the complex clauses occurs as a bare NP. This is obvious in the pairs in (49) and (50).

(49) a. Marlena abala parkara Hasan ja' embi'eng ngekke' Ali.
     M AV.say about H COMP goat.DEF AV.bite A
     'Marlena said about Hasan that his goat bit Ali.'

   b. Marlena abala'**agi** Hasan ja' embi'eng ngekke' Ali.
     M AV.say.AGI H COMP goat.DEF AV.bite A
     'Marlena said about Hasan that his goat bit Ali.'

(50) a. Sengko' loppa parkara Ita ja' Hasasn ngerem paket dha' aba'eng.
     I forget about I COMP H AV.send package to she
     'I forgot about Ita that Hasan send a package to her.'

   b. Sengko' ngloppa**e** Ita ja' Hasasn ngerem paket dha' aba'eng.
     I AV.forget.E I COMP H AV.send package to she
     'I forgot about Ita that Hasan send a package to her.'

With the predicate *bala* 'say', when the matrix element is a bare NP, the suffix *-agi* occurs on the verb (49b). With predicate *loppa* 'forget', *-e* is suffixed when the matrix element is a bare NP (50b). The use of these suffixes parallels their appearance in structures in (47b) and (48b).

There is thus strong evidence that the Madurese analogue of the construction that has been analyzed as Raising-to-Object for other Javanic languages is in fact not a raising construction at all and is most appropriately analyzed as a case of prolepsis. It is entirely possible that closer examination of this type of structure in both closely related and unrelated languages may provide insights into the appropriate analyses of corresponding constructions in these languages.

CONCLUSION

The foregoing discussion is intended to illustrate the various ways that in-depth investigation and analysis of any particular language that a

community of speakers uses can shed light on the investigation of the properties of language in general and the characteristics of other natural languages. It is through the detailed description and examination of as many of the world's languages as possible that we can gain insight into one of the most fundamental but ultimately most distinct behaviors of human beings.

ACKNOWLEDGEMENTS

REFERENCES

Bintoro. 1980. Javanese transitive verbs: a tagmemic analysis. *NUSA Linguistic Studies in Indonesia and Languages in Indonesia 8.* Jakarta: NUSA.

Chomsky, N. 1959. Review of B.F. Skinner: verbal behavior. *Language, 35,* 26-58.

Chomsky, N. 1973. Conditions on transformations. In S. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle.* New York: Holt, Rinehart and Winston.

Chung, S. 1976. On the subject of two passives in Indonesian. In C. Li (Ed.), *Subject and Topic* (pp. 57-98). New York: Academic Press.

Cumming, S. 1986. Word order change in Malay. In P. Geraghty, L. Carrington, & S.A. Wurm (Eds.), *FOCAL I: Papers from the Fourth International Conference on Austronesian Linguistics* (pp. 97-111). Canberra: Pacific Linguistics.

Davies, W. D. 1990. Javanese evidence for subject-to-object raising. In K. Dziwirek, P. Farrell, & E. Mejias-Bikandi (Eds.), *Grammatical Relations: A cross-Theoretical Perspective* (pp. 95-108). Stanford: CSLI.

Davies, William D. 1999. Madurese and Javanese as strict word-order languages. *Oceanic Linguistics, 31,* 152-167.

Davies, W. D. (To appear). The richness of Madurese voice. In I W. Arka & M. Ross (Eds.), *Aspects of Austronesian Voice Systems*. Canberra: Pacific Linguistics.

Greenberg, J. 1966. *Language universals, with special reference to feature-hierarchies.* The Hague: Mouton.

Higgins, R. 1981. Proleptic objects and verbs of perception in Zacapoaxtla Nahuat. *Texas Linguistics Forum, 18*, 69-88.

Hornstein, N., & Lightfoot, D. (Eds.). 1981. *The logical problem of language acquisition*. New York: Longman.

Jake, J., & Odden D. 1979. Raising in kipsigis. *Studies in the Linguistic Sciences, 9,* 131-155.

Joseph, B. 1976. Raising in modern Greek: a copying process? *Harvard Studies in Syntax and Semantics*, Volume II.

Lasnik, H., & Saito, M. 1991. *Papers from the 27th regional meeting of the Chicago linguistic society*. University of Chicago Department of Linguistics.

Moore, J. 1998. Turkish copy-raising and A-chain locality. *Natural Language and Linguistic Theory 16*, 148-189.

Naylor, P. B. 1978. Toward focus in Austronesian. In S.A. Wurm & L. Carrington (Eds.), *Second International Conference on Austronesian Linguistics: Proceedings* (pp. 395-442). Canberra: Pacific Linguistics.

Nothofer, B. 1975. *The reconstruction of proto-Malayo-Javanic*. The Hague: Mouton.

Postal, P. M. 1974. *On raising*. Cambridge, MA: MIT Press.

Potsdam, E., & Runner, J. T. 2001. Richard returns: copy raising and its implications. *Papers from the Blah 37th Regional Meeting of the Chicago Linguistic Society*. University of Chicago Department of Linguistics.

Rosenbaum, P. S. 1967. *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.

Starosta, S. 1986. Focus as recentralisation. In P. Geraghty, L. Carrington, & S.A. Wurm (Eds.), *FOCAL 1: Papers from the Fourth International Conference on Austronesian Linguistics* (pp. 73-95). Canberra: Pacific Linguistics.

Stevens, A. 1968. *Madurese phonology and morphology.* American Oriental Society.

Seiter, W. 1983. Subject-direct object raising in Niuean. In D.M. Perlmutter (Ed.), *Studies in Relational Grammar 1* (pp. 317-359). Chicago: University of Chicago Press.

Thomas, M. R. 1980. Verb affixes and focus in Bahasa Indonesia. In P.B. Naylor (Ed.), *Austronesian Studies* (pp. 63-70). Ann Arbor: University of Michigan.

Wechsler, S. & Arka., I. W. 1998. Syntactic ergativity in Balinese: an argument structure based theory. *Natural Language and Linguistic Theory, 16,* 387-441.