

ON THE NUMBER OF VOICES IN MADURESE*

Helen JEOUNG
University of Pennsylvania
jeoungh@sas.upenn.edu

Abstract

Two voices have previously been documented in Madurese: an active voice, and a non-active voice with *e-* prefixed verbs. In this paper I examine the non-active voice, which has variously been called passive voice or object voice, and demonstrate that it is a canonical passive voice. Furthermore, I document the existence of a distinct third voice, an object voice similar to that of other languages of the area. The object voice is used in the polite register, but not in the familiar register, so the two-voice system in the familiar register differs from the three-voice system in the polite register. The registers are also differentiated by nominal extraction patterns: objects may extract in polite speech, but not in familiar speech. In view of these contrasts, I argue that the familiar and polite registers each operate with distinct sets of morphosyntactic rules, or distinct grammars.

Keywords: Madurese, voice, register, Austronesian

ISO 639-3 codes: mad

1 Introduction

This paper examines the voice system of Madurese, a Malayo-Polynesian language spoken primarily in western Indonesia. Two distinct voices are identified in previous literature on Madurese (Davies 2010, Kiliaan 1897, Stevens 1968), illustrated in the familiar register in (1) and (2).^{1, 2}

- (1) *Ali ng-ato-eh ana'-eng.* (Familiar)
Ali AV-call-Appl child-Def
'Ali called his child.'
- (2) *Ana'-eng e-kato-eh bi' Ali.*
child-Def PV-call-Appl by Ali
'The child was called by Ali.'

Previous authors agree that clauses such as (1) are active. However, clauses with *e-* prefixed verbs such as (2) have been described as either passive or object voice. The first goal of this paper is to address this issue: I demonstrate that the morphosyntactic properties of (2) are characteristic of a canonical passive voice rather than object voice. Thus (1) and (2) show an active-passive voice opposition in the familiar register.

I next show that the voice system in the polite register differs from that of the familiar register. In addition to active and passive voices, in the polite register there exists a previously undocumented object voice, which is illustrated in (3).

* Thanks to my consultants, Isya Mahfud, Maimuna and M. Khoirun Najib. Thanks also to Julie Legate for much helpful input, and to the audiences at 13-ICAL and AFLA 22, as well as two anonymous reviewers whose comments have improved this work overall. All errors remain my own.

¹ Abbreviations: Appl=applicative; AV=active voice; Caus=causative; Def=definite; Irr=irrealis; OV=object voice; Perf=perfective; PV=passive voice; Redup=reduplicant; Rel=relative morpheme.

² The orthography suggested by my consultants is used throughout the paper, as Madurese orthography is not completely standardized. Data cited from other sources appear with the original orthography.

- (3) *Potra-epon ampon ramah tembhal-ih.* (Polite)
 son-Def Perf father OV.call-Appl
 ‘Father called his son.’

This object voice is not marked with the verbal prefix *e-*, but rather with a null prefix (\emptyset -). Object voice is employed in the polite register, but not in the familiar register. Object voice was identified as a distinct type of voice configuration in Chung 1976 for Indonesian, which is related to Madurese. Prior to that time, research on languages of Indonesia had identified only active and passive voices, which were well attested, e.g. in European languages. Subsequent research has identified object voice in a number of related languages (see e.g. Arka 2002, 2003, 2008, Arka and Manning 1998, Cole et al 2006, 2008, Guilfoyle et al 1992, Legate 2014, McKinnon et al 2011, Sneddon 2006, Sneddon et al 2012, Yanti 2010). For Madurese, the object voice may have escaped notice in more recent work such as Davies’ (2010) grammar because the data is drawn primarily from the familiar speech level, which does not appear to employ the object voice.³

A third goal of this paper is to demonstrate that the polite and familiar registers also differ in their patterns of object extraction: while object extraction is possible in the polite register, it is illicit in the familiar register. One consequence of these differences is that each register has its own grammatical system; a Madurese speaker who controls both the familiar and polite registers employs two distinct grammars.

This paper is organized as follows. In section 2, I demonstrate that Madurese *e-* verbs are a canonical passive voice in both registers. In section 3, the polite register is shown to have a distinct third voice, the object voice. The discussion in section 4 extends to another grammatical difference between the registers: the pattern of nominal extraction. In section 5 I propose an analysis that accounts for the grammatical differences between the registers, and argue that two distinct grammars operate within Madurese. Conclusions are given in section 6.

2 A canonical passive in Madurese

2.1 Background

Madurese has several colloquial varieties, and their geographic distribution is characteristic of a dialect continuum (Lewis et al 2015). The Madurese consultants who provided data for this paper speak a western variety that is sometimes called the Bangkalan dialect because it is associated with the Bangkalan regency on Madura island.⁴

Madurese has at least three registers or speech levels (Davies 2010, Pawitra 2008, Safioedin 1977, Stevens 1968), which are named for the terms ‘no-yes’ in each register, or for the characterization of the speech level.

- | | | |
|------------------------|------------------------------|-----------------|
| • <i>Enjha’-iya</i> | <i>kasar</i> ‘rough, coarse’ | Familiar speech |
| • <i>Engghi-enten</i> | <i>tengnga’an</i> ‘middle’ | Middle speech |
| • <i>Engghi-bunten</i> | <i>alos</i> ‘smooth, polite’ | Polite speech |

Throughout this paper, the terms ‘register’ and ‘speech level’ are used interchangeably. Familiar register is generally used when speaking to friends and children, while polite register is used when speaking to, or of, superiors. The middle register is less productive, with fewer lexical items, and often middle speech can be mixed with either familiar or polite speech. For the purposes of this paper, in order to draw a clear contrast between registers, the middle speech level is not considered here.

Madurese morphosyntax shares many characteristics with other Indonesian languages of the area. In particular, voice is morphologically marked as verbal affixes. Active voice (AV) is marked either with the

³ Both Davies 2010 and this paper are based on the western dialect. It is possible that in other dialects, there may be variation in the number of voices in each register.

⁴ After the primary data had been collected for this paper, judgments were also checked with speakers from Jember and Sampang.

prefix *a-* or a homorganic nasal prefix *N-* as in (1).⁵ I now turn to the question of what voice is encoded by *e-* verbs.

2.2 *E- verbs are a canonical passive*

The voice that is encoded by *e-* verbs has been called both passive voice and object voice in the literature. Kiliaan 1897 and Stevens 1968, both of which document Madurese morphology and phonology, identify *e-* verbs as passives. On the other hand, clauses with *e-* verbs are described in Davies 2010 as object voice. In this section I show that the first view must be correct, and that Madurese *e-* verbs pattern as a canonical passive.

Many Austronesian languages have more than one non-active voice, and in the Austronesian literature non-active voices are sometimes collectively referred to as ‘passives’. I reserve the term ‘passive’ for a canonical (European-style) passive voice rather than a range of non-active voices. Cross-linguistically, canonical passives share several characteristics (Haspelmath 2001): an argument other than the Agent is raised or promoted to the position of grammatical subject; typically, this is a VP-internal object. The Agent argument is optional in a passive clause, and may be implicit. When the Agent is pronounced however, it appears as an oblique argument.⁶ In languages of the area (Indonesia), another characteristic trait is an alternation between passive and active morphology on the verb. For example, in both Indonesian and Balinese, active and passive affixes are distinct, and passive morphology on the verb is obligatory in passive clauses (Arka 2003, Cole et al 2008, Sneddon et al 2012).

Madurese *e-* clauses in both familiar and polite speech display characteristics of a canonical passive voice. First, the Patient in an *e-* clause occurs as grammatical subject. In (4a) the Patient *nase’eng Fatima* occupies the pre-verbal subject position; likewise the Patient *potraepon* in (5a).

- (4a) *Nase’-eng Fatima e-kakan (bi’ ale’)*. (Familiar)
 rice-Def Fatima PV-eat by younger.sibling
 ‘Fatima’s rice was eaten by little brother.’
- (4b) **E-kakan nase’-eng Fatima bi’ ale’*.
 PV-eat rice-Def Fatima by younger.sibling
- (4c) **Nase’ e-kakan (bi’ ale’)*.
 rice PV-eat by younger.sibling
- (4d) *Nase’eng Fatima se e-kakan*
 rice-Def Fatima Rel PV-eat
 ‘Fatima’s rice that was eaten’
- (5a) *Potra-epon e-tembhal-ih (sareng ramah)*. (Polite)
 son-Def PV-call-Appl by father
 ‘The son was called by father.’
- (5b) **E-tembhal-ih potra-epon sareng ramah*.
 PV-call-Appl son-Def by father
- (5c) **Potra e-tembhal-ih (sareng ramah)*.
 son PV-call-Appl by father

⁵ The distribution of the two active prefixes is discussed in Davies (2010), but is not relevant to the discussion here.

⁶ I use Agent as a general thematic category that includes Experiencers.

- (5d) *Potra-epon se e-tembhal-ih*
 son-Def Rel PV-call-Appl
 ‘the son that was called’

The subject position must be filled in Madurese, so it is ungrammatical for the Patient to remain in its post-verbal thematic position in (4b), (5b). Like all subjects in Madurese, the preverbal Patient in *e-* clauses must be definite or specific; note the ungrammaticality of (4c), (5c). Another piece of evidence that the Patient is a grammatical subject in (4a), (5a) is that it occurs without the relative morpheme *se*, which obligatorily occurs after extracted nominals in (4d), (5d).

As expected of a canonical passive, an overt Agent is optional in (4a) and (5a). When an Agent does occur, it is introduced by a preposition (familiar *bi’/moso*; polite *sareng*). This by-phrase may occur in different surface positions in the clause, as expected for a PP adjunct.

- (6a) *(Bi’ embi’ rowa) jhuko’ jiyah e-kakan.* (Familiar)
 by goat that fish that PV-eat
 ‘(By the goat) the fish was eaten.’

- (6b) *Jhuko’ jiyah e-kakan bari’ (bi’ embi’ rowa).*
 fish that PV-eat yesterday by goat that
 ‘The fish was eaten yesterday (by the goat).’

- (6c) *Jhuko’ jiyah e-kakan (bi’ embi’ rowa) bari’.*
 fish that PV-eat by goat that yesterday
 ‘The fish was eaten (by the goat) yesterday.’

The forms that introduce an Agent in a passive clause (*bi’*, *moso*, *sareng*) have other prepositional uses as well. For instance, they can introduce an instrument PP and also may have a comitative usage (examples 7 and 8 are from Davies 2010:229, 27 and 25). The Agent occurring with *e-* clauses is thus an oblique argument, embedded within a PP by-phrase.

- (7) *Ali ma-becce’ sapedha motor-ra bi’ obing.* (Familiar)
 Ali AV.CS-good motorcycle-DEF with screwdriver
 ‘Ali fixed his motorcycle with a screwdriver.’

- (8) *Mangkana engko’ keng wa’-dhuwa’-an bi’ ba’eng.*
 therefore I only RED-two-AN with you
 ‘Therefore, I will only be two with you.’

Finally, similar to other areal languages, the verbal prefix *e-* is obligatory in all passive clauses. The *e-* prefix in the passive clauses (4a), (5a) alternates with nasal prefixes in the active counterparts (9), (10).

- (9) *Ale’ ng-akan nase’-eng Fatima.* (Familiar)
 younger.sibling AV-eat rice-Def Fatima
 ‘Little brother ate Fatima’s rice.’

- (10) *Ramah n-embhal-ih potra-epon.* (Polite)
 father AV-call-Appl son-Def
 ‘Father called his son.’

In brief, Madurese *e-* verbs display the characteristics of a canonical passive, a finding that supports Kiliaan 1897 and Stevens 1968. I assume that *e-* is a functional morpheme that encodes passive voice on the verb.

3 Object voice in Madurese

In this section I establish that in addition to active and passive voice, Madurese has a distinct third voice: an object voice akin to that in other Indonesian languages such as Acehnese, Balinese, Indonesian and Javanese.

3.1 Terminology for passive voice and object voice

I use ‘passive voice’ and ‘object voice’ to distinguish the two non-active clauses in Madurese, following e.g. Cole et al 2008 and Legate 2014. However, the terminology varies in the literature describing languages of the area. For passive and object voices, respectively, Chung 1976 uses ‘canonical passive’ and ‘object preposing passive’; Dardjowidjojo 1978 and Sneddon et al 2012 use ‘Passive type I’ and ‘Passive type II’; Aldridge 2008 uses ‘passive’ and ‘pronominal passive’; Arka and Manning 2008 use ‘passive voice’ and ‘objective voice’. The Indonesian term *pasif semu* ‘pseudo-passive’ also refers to object voice, whereas the term *obyek fokus* ‘focused object’ is sometimes used inconsistently in the literature for various constructions in which an object DP appears at the left edge of the clause (including object voice clauses, active clauses with an extracted object and topic-comment clauses).

Despite the varying terminology, object voice across these languages shares characteristics that distinguish it from both active and passive voices. The primary criteria I utilize are morphosyntactic properties; the analysis presented here is essentially a syntactic one, following previous literature on object voice in related languages (Arka 2002, 2003, 2008, Arka and Manning 1998, Cole et al 2006, Cole et al 2008, Guilfoyle et al 1992, Legate 2014, McKinnon et al 2011, Sneddon 2006, Sneddon et al 2012, Yanti 2010). I leave open the question of which voice is more ‘basic’ in terms of frequency or discourse function.

3.2 Object voice in the Polite register

In this section I demonstrate that the Polite register of Madurese has object voice. First, (11-13) illustrates both AV and PV clauses in the Polite register.

(11a) *Ramah ampon n-embhal-ih potra-epon.* (Polite)
 father Perf AV-call-Appl son-Def
 ‘Father called his son.’

(11b) *Potra-epon ampon e-tembhal-ih sareng ramah.*
 son-Def Perf PV-call-Appl by father
 ‘The son was called by father.’

(12a) *Kaulah kodhuh ng-obang-eh sepeda motor.*
 1sg must AV-money-Appl motorbike
 ‘I must purchase a motorbike.’

(12b) *Sepeda motor-epon kodhuh e-obang-eh.*
 motorbike-Def must PV-money-Appl
 ‘The motorcycle must be purchased.’

(13a) *Red-mored panekah ng-omba rasoghan-epon.*
 Redup-student that AV-wash shirt-Def
 ‘The students washed their shirts.’

(13b) *Rasoghan-epon e-komba sareng red-mored.*
 shirt-Def PV-wash by Redup-student
 ‘The shirts were washed by the students.’

The same verbs occur in the Object voice (OV) clauses in (14-16).⁷

⁷ OV glosses are rendered with active verbs, as there is no equivalent for object voice in English.

- (14) *Potra-epon ampon ramah tembhal-ih.* (Polite)
 son-Def Perf father OV.call-Appl
 ‘Father called his son.’
- (15) *Sepeda motor-epon kodhuh kaulah obang-eh.*
 motorbike-Def must 1sg OV.money-Appl
 ‘I must purchase a motorbike.’
- (16) *Rasoghan-epon ampon ebhuh komba.*
 shirt-Def Perf mother OV.wash
 ‘Mother washed the shirt.’

In OV, the thematic Patient occurs in the pre-verbal position of grammatical subjects, as it does in the passive voice. However, a number of morpho-syntactic features distinguish object voice from passive voice. Unlike PV, the verb in object voice must be ‘bare,’ which refers specifically to the lack of a phonologically overt voice prefix in (14-16). Bare verbs may have other affixes (e.g. the applicative suffixes in 14 and 15), but all voice prefixes are illicit; for example, the OV clauses in (14-16) are rendered ungrammatical if they have active or passive voice prefixes as in (17-19).

- (17) **Potraepon ampon ramah n-embhal-ih.* (Polite)
 son-Def Perf father AV-call-Appl
- (18) **Sepeda motor-epon kodhuh kaulah e-obang-eh.*
 motorbike-Def must 1sg PV-money-Appl
- (19) **Rasoghan-epon ampon ebhuh ng-omba.*
 shirt-Def Perf mother AV-wash

In addition to bare verbal morphology, another difference between the object voice and the the passive voice is the Agent argument. As previously discussed, in a canonical passive clause, the Agent is an oblique argument embedded in a PP. By contrast, the OV Agent cannot occur with the preposition *sareng* ‘by’.

- (20) *Potra-epon ampon (*sareng) ramah tembhal-ih* (Polite)
 son-Def Perf by father OV.call-Appl
 ‘Father called his son.’

That the Agent *ramah* cannot be introduced by a preposition in (20) indicates that the Agent argument is a bare DP, rather than a PP.⁸ Furthermore, the OV Agent must always be overt, unlike the Agent in a canonical passive, which is optionally expressed/implicit. Compare (14) with the ungrammatical sentence in (21), in which the Agent is unexpressed.

- (21) **Potra-epon ampon tembhal-ih.* (Polite)
 son-Def Perf OV.call-Appl
 ‘(Father) called his son.’

The OV Agent therefore must be present in the syntactic structure, unlike the optional Agent in a passive. The obligatory nature of the Agent, and its status as a bare DP, are both expected if the Agent is an external argument of the verb, rather than an adjunct.

⁸ The prepositions *bi*, *moso* and *sareng* may be phonologically null under restricted conditions (see 3.4). However, the distribution of null P is such that wherever null P is licensed, its overt counterpart is also licit.

The ungrammaticality of (21) also rules out *pro*-drop of the OV Agent. Note that Madurese allows *pro*-drop of subjects when the dropped argument is already understood (Davies 2010). Both the AV clause in (22) and the PV clause in (23) allow null subjects if previously mentioned in the discourse.

- (22) *Teros n-embhal-ih potra-epon.* (Polite)
 then AV-call-Appl son-Def
 ‘Then (he) called his son.’
- (23) *Teros e-tembhal-ih sareng ramah.*
 then PV-call-Appl by father
 ‘Then (he) was called by father.’

In fact, the subject of an OV clause (the Patient argument) may also be dropped as in (24), given the appropriate discourse context.

- (24) ... *teros ampon ramah tembhal-ih.* (Polite)
 then Perf father OV.call-Appl
 ‘Then Father called (him).’

Thus OV is consistent with the generalization that Madurese allows subjects to *pro*-drop; however, the OV Agent may never be dropped. The OV Agent, then, does not behave as a grammatical subject in Madurese.

Rather, the OV Agent always occurs to the right of the subject, and to the right of auxiliaries and modals as well. The Agent must also be immediately to the left of the verb. Compare the relative word orders in AV (25) and OV (26).

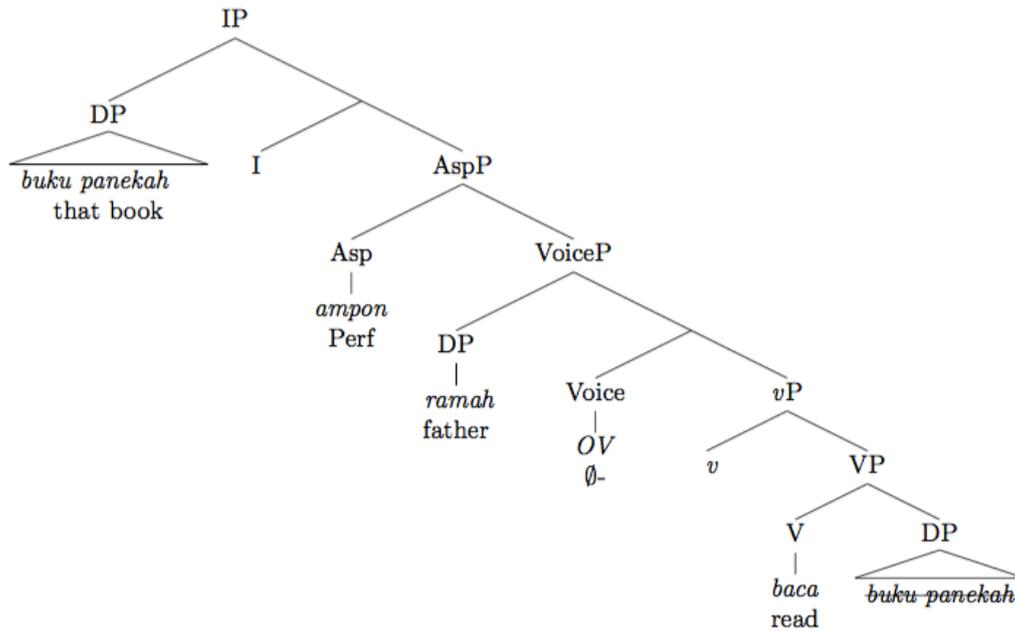
- (25) *Ramah ampon m-acah buku panekah.* (Polite)
 father Perf AV-read book that
 ‘Father already read that book.’
- (26) *Buku panekah ampon ramah bacah.*
 book that Perf father OV.read
 ‘Father already read that book.’
- (27) **Buku panekah ramah ampon bacah.*
 book that father Perf OV.read

Modals and aspectual morphemes such as *ampon* occur to the right of the Agent in an AV clause (25), but must occur to the left of the Agent in an OV clause (26). In object voice, no element may intervene between the Agent and the verb; (27) shows that it is ungrammatical for the auxiliary to occur in this position. Cole et al 2008 utilize this relative ordering in Indonesian (i.e. Agent-Auxiliary-Verb vs. Auxiliary-Agent-Verb), to differentiate between AV and OV, respectively.

The linear position of the OV Agent supports an analysis in which the OV Agent remains in its thematic position (as previously argued by Cole et al 2008, Guilfoyle et al 1992, Legate 2014, Yanti 2010). See Figure 1 for the structure of an object voice clause in Madurese. I assume that VoiceP is the functional projection that introduces an external argument in its specifier in AV and OV (following Harley 2013, Legate 2014, Pylkkänen 2002).⁹ The functional head Voice⁰ is spelled out as a voice prefix on the verb: this head has a phonological exponent in AV and PV, but is phonologically null in object voice.

⁹ The precise configuration between Voice and *v* is not at issue here, but rather the main point is that the head of the projection hosting the external argument in its specifier is realized as a different prefixes in active, passive or object voice. Recall that a ‘bare’ OV verb does not bear an overt voice prefix, but it may bear other prefixes, such as a causative prefix hosted in *v*. Example (i) shows the co-occurrence of the prefix hosted in Voice and the prefix hosted in *v*: the voice prefix *e-* occurs with the causative prefix *pa-*.

Figure 1: Structure of object voice (OV) clause



Unlike AV clauses in which the external argument raises to the derived subject position, SpecIP, the external argument in an OV clause remains in Spec,VoiceP (Guilfoyle et al 1992). We have already seen that it is the Patient that occurs as subject in object voice. The structure above thus derives the surface word order Patient-(Aspect/Modal)-Agent-bare verb.¹⁰

From its position in Spec,VoiceP the Agent in object voice c-commands VP internal arguments. In other languages with object voice, authors have applied various diagnostics as evidence of this c-command relation. For instance, Arka and Manning (2008) use binding evidence to show that in Indonesian OV, when a reflexive is raised to subject, the Agent binds the reflexive in its base position. Unfortunately this diagnostic cannot be applied to Madurese, since reflexives are not licit in subject position.¹¹ Legate 2014 also presents arguments to establish that in Acehnese, the OV Agent c-commands arguments introduced low in the VP. Here I only briefly mention several reasons why similar diagnostics are unavailable in Madurese. First, binding tests using personal pronouns are not available because the polite register lacks third person pronouns altogether (Davies 2010). Possessive DPs do not contain a third person possessive pronoun (i.e. a variable), only definite marking. The OV Agent cannot be controlled PRO, but this is not surprising since it

(i) Motor-ra e-pa-becce' bari'. (Familiar)
 motorbike-Def PV-Caus-good yesterday
 'The motorbike was fixed yesterday.'

¹⁰ Object voice appears to violate locality: the Patient is raised to subject over an argument that is structurally closer, the thematic Agent. Several analyses might explain this phenomenon, although the Madurese data at hand does not adjudicate between them. For example, if the Agent receives inert (inherent) Case as described in McGinnis 1998, the Agent is ineligible for A-movement, and thus a lower DP is raised to subject. (Interestingly, McGinnis notes that Agents do not appear to receive inert case; Madurese would be an exception to this generalization.) A different solution is offered by Legate (2014), who adopts a leapfrogging analysis (a la Bobaljik 1995) for Acehnese object voice: the head Voice, which selects an external argument in its specifier, also attracts the Patient DP to create an additional specifier of VoiceP; the Patient DP is then visible for further movement. For other approaches, see Aldridge 2008; Levin 2015.

¹¹ Davies 2008 provides examples of apparent reflexives occurring as subjects. However, the relevant judgments are not shared by my consultants; I find instead that these forms are not reflexives, but rather personal pronouns modified by an intensifier.

is not a subject and must be overt. Furthermore, since OV Agents cannot be modified (see below), tests involving quantifier scope and quantifier float are also ruled out as diagnostics.

I take OV word order (Patient–auxiliary/modal–Agent–verb) and the strict adjacency between the verb and Agent as indications that the OV Agent remains in Spec,VoiceP. Several properties of object voice clauses follow naturally from this. Other elements cannot intervene between the Agent and the verb because they lack a structural position to do so. The OV Agent and the verb appear to form a phonological unit, possibly as a result of the Agent’s obligatory adjacency to the verb. In speech, it is reported to be unnatural, or even impossible, to pause between the OV Agent and the verb. Similar morphophonological effects are reported in other languages with OV as well. Finally, because the OV Agent must be pronounced in its position in SpecVoiceP, the Agent is said to be ‘immobile’ (Legate 2014); it cannot undergo movement. For example, wh-movement such as relativization (28) or topicalization (29) is ungrammatical.

- (28) **Ramah se potra-epon ampon tembhal-ih* (Polite)
 father Rel son-Def Perf OV.call-Appl
 ‘The father who called his son’¹²
- (29) **Ramah, potra-epon ampon tembhal-ih*
 father son-Def Perf OV.call-Appl
 ‘Father, (he) called his son’

In Madurese, object voice Agents are selected from a closed class of DPs, including polite personal pronouns (see 15) and a small set of kinship terms such as *ramah* ‘father,’ *ebhu* ‘mother,’ *ale* ‘younger sibling’ etc.^{13, 14} Other DPs are disallowed, including common nouns, names and modified nominals.¹⁵ Compare the acceptability of full DPs such as ‘Mr. Tono’ and ‘the teacher’ in active voice (30a, 31a) with the ungrammaticality of the equivalent object voice clauses (30b, 31b).

- (30a) *Pak Tono n-embhal-ih potra-epon.* (Polite)
 Mr Tono AV-call-Appl son-Def
 ‘Mr Tono called his son.’
- (30b) **Potra-epon ampon Pak Tono tembhal-ih.*
 son-Def Perf Mr Tono OV.call-Appl
- (31a) *Ghuruh panekah m-acah buku.*
 teacher that AV-read book
 ‘The teacher read a book.’
- (32b) **Buku panekah ampon ghuruh panekah bacah.*
 book that Perf teacher that OV.read

Furthermore, wh-words do not belong to the set of possible OV Agents. Madurese has three strategies for wh questions: in-situ, partial movement and clefted/moved wh (Davies 2003, 2010). However,

¹² The surface word order of (28) is possible as a pseudo-cleft with extracted object, i.e. ‘It is Father who his son called’. See section 4.1 for discussion regarding the ambiguity between OV clauses and AV clauses with extracted object.

¹³ Not all languages restrict the class of Agents in object voice, e.g. Acehnese (Legate 2014), colloquial Jakarta Indonesian (Sneddon 2006), some dialects of Jambi Malay (Yanti 2010).

¹⁴ Unlike other languages of the area, kinship terms cannot be used pronominally (for 2nd person address).

¹⁵ There is some interspeaker variation regarding the acceptability of some DPs as OV Agent. Simple polite pronouns are accepted by all speakers, and common nouns and complex DPs are judged ungrammatical. However, the acceptability of kinship terms and titles may vary.

questioning the Agent in-situ (immediately to the left of the verb) is disallowed. In object voice, the selectional requirements thus limit the Agent to definite/specific DPs.

- (33) **Potra-epon ampon paserah tembhal-ih?* (cf. 14)
 son-Def Perf who OV.call-Appl
 ‘Who called his son?’

Given the restrictions on the OV Agent, questioning the OV Agent with a moved wh word is expected to be ungrammatical, which is the case.

- (34) **Paserah se potra-epon ampon tembhal-ih?* (Polite)
 who Rel son-Def Perf OV.call-Appl
 ‘Who called his son?’

Questions such as (34) are ill-formed on two counts: as a consequence of the immobility of the Agent, as well as the selectional requirements on the Agent in object voice. Interestingly, the inability to form either wh-in situ questions or clefted questions has the effect that in Madurese, Agent questions are not possible at all in OV.

Finally, the pragmatics associated with object voice are still somewhat unexplained in Madurese and other Indonesian languages. It is not clear why speakers employ object voice rather than active or passive voice, and what semantic or pragmatic differences arise. Discourse and information structural factors appear to be relevant; since Madurese subjects must be definite (Davies 2010), and the set of possible OV Agents is also definite, the Agent and Patient are both be old information, rather than new information. I suggest that in AV and PV clauses, the argument that occurs as grammatical subject is given prominence, whereas in an OV clause, both the Patient and the Agent retain equal prominence (neither is ‘demoted’). Since neither argument is given more salience than the other, both the Patient and the Agent share ‘aboutness’ in an OV clause. However, more research is needed in order to better understand the use of object voice in these languages.

3.3 No object voice in the Familiar register

Having demonstrated that the polite register of Madurese has object voice, I next find that the familiar register of Madurese does not have object voice. The expected OV word order and null voice marking are given in (35-37) with an acceptable OV Agent:

- (35) **Ana’eng la engko’ kato-eh.* (Familiar)
 child-Def Perf 1sg OV.call-Appl
 ‘I called his child.’
- (36) **Buku jiyah ale’ bacah.*
 book that younger.sibling OV.read
 ‘Little brother read that book.’
- (37) **Jhuko’ jiyah bhakal hedah kakan-ah.*
 fish that will 2sg OV.eat-Irr
 ‘You will eat that fish.’

The ungrammaticality of these sentences appears to arise from both the word order as well as the bare verb form, which are not possible at all in familiar speech.¹⁶ Other combinations of pronominal Agents and bare verbs are also unacceptable. I conclude that in the familiar register, object voice is not possible.

¹⁶ A few verbs such as *mole* ‘go home’ are exceptions in that they do not occur with voice prefixes.

3.4 Postverbal Agents in *e*- clauses without (overt) P

I now turn to an ambiguous type of clause in which an *e*- prefixed verb is immediately followed by an Agent (without a preposition) as in (38a) and (39a):

(38a) *Jhuko' jiyah e-kakan (embi' rowa).* (Familiar)
 fish that PV-eat goat that
 'The fish was eaten by the goat.'

(38b) *Jhuko' jiyah e-kakan (bi' embi' rowa).*
 fish that PV-eat by goat that
 'The fish was eaten by the goat.'

(39a) *Potra-epon e-tembhal-ih (Pak Tono).* (Polite)
 son-Def PV-call-Appl Mr Tono
 'The son was called by Mr Tono.'

(39b) *Potra-epon e-tembhal-ih (sareng Pak Tono).*
 son-Def PV-call-Appl by Mr Tono
 'The son was called by Mr Tono.'

Note that (38a) and (39a) are the same as (38b) and (39b), except that no preposition occurs before the Agent. I have previously shown that *e*- prefixed verbs are canonical passives, and I argue that the examples above are no exception. Several pieces of evidence indicate that (38a) and (39a) are canonical passive clauses with a phonologically null P. In other words, the postverbal Agent is embedded in a PP whether or not the P is pronounced. Davies 2010 also provides further examples of optionally pronounced prepositions in Madurese:¹⁷

(40a) *Maleng rowa e-tangkep polisi.* (Davies 2010:256-7, examples 20, 22, 24a,b)
 thief that PV-catch police
 'The thief was caught by police.'

(40b) *Maleng rowa e-tangkep so polisi.*
 thief that PV-catch by police
 'The thief was caught by police.'

(41a) *So polisi maleng rowa e-tangkep.*
 by police thief that PV-catch
 'By police the thief was caught.'

(41b) **Polisi maleng rowa e-tangkep.*

Immediately following the passive verb *etangkep*, the familiar preposition (*mo*)*so* may be null (40a) or pronounced (40b). However, if the Agent and verb are separated, the preposition must be overt (41a, 41b). Thus the immediate post-verbal position of the PP licenses an optionally null P (familiar *bi'*, (*mo*)*so* or polite *sareng*); in other positions, the P must be overt. In other words, (38a) and (38b) are structurally identical; likewise (39a, 39b) and also (40a, 40b). As expected under this analysis, no semantic difference arises between the pairs of sentences, as their syntactic structure is identical. The Agent is also optionally expressed in all cases, as expected for a passive clause.

Let us consider an alternate analysis for (38a, 39a, 40a) in which the Agents *embi' rowa* and *Pak Tono* are not embedded in a PP, but are verbal arguments. The only position this argument occurs is the immediate

¹⁷ Recall that in Davies 2010, passive *e*- verbs are labeled object voice. To avoid confusion, I have replaced the gloss of OV with PV wherever the *e*- prefix occurs. I also use passive voice in the glosses.

right of the verb (we have seen that in other positions, it must occur with P). This is equivalent to an object voice analysis: the Patient occurs as subject, and the Agent is an external argument that obligatorily occurs adjacent to the verb.¹⁸ This analysis, however, raises several questions. Davies (1999) shows that Madurese has fairly rigid SVO word order: external arguments are generated to the left of the verb in Madurese in both active voice and object voice. VP fronting is possible in active clauses, but results in Verb-Object-Agent word order, rather than the Object-Verb-Agent order in (38a, 39a and 40a). Although this is like an object voice analysis, the postverbal Agents here are not subject to the selectional restrictions of object voice: the Agent can be a common noun modified by a demonstrative in (38a, 40a) or a complex name in (39a).

On this latter point, one reviewer suggests that the post-verbal Agent is an external argument that is base generated to the right of the verb; in object voice, a subset of these nominals can then undergo clitic movement to a preverbal position, accounting for the selectional restrictions on the OV Agent. An obvious difficulty with this suggestion is that the morphology on the verb differs in these two situations: this approach requires that the post-verbal external argument occurs with the passive prefix *e-*, while the moved clitic arguments always occur with bare verbs. This requires *e-* to be a syncretized form, encoding both a canonical passive voice *in addition to* another voice that is clearly not passive, since it selects an external argument. It is also difficult to explain why the postverbal Agent is optional, but the (moved) preverbal Agent is obligatory. Finally, this type of clitic movement is otherwise absent from the language: pronouns (as well as names) do not display such movement. If clitic movement applies to *ramah* ‘father’ and *ebhu* ‘mother,’ which are licit as OV Agents, these nominals should be equivalent to pronouns; but there is no evidence for this. Unlike other related languages of the area, kinship terms cannot be employed pronominally (for second-person address) in Madurese.

Returning to the present analysis, null P is also attested in areal languages. Indonesian exhibits a pattern similar to Madurese: immediately following a passive verb, the preposition *oleh* ‘by’ may be unpronounced. Various authors adopt this view (Macdonald and Dardjowidjojo 1967, Dardjowidjojo 1978, Jeoung and Biggs (to appear), Sneddon et al 2012). Arka and Manning’s (2008) analysis concludes that the postverbal Agent without an overt P is an oblique argument in Indonesian, rather than a core argument. Jeoung and Biggs (to appear) call this phenomenon P-drop, arguing that in Indonesian, when PP modifies VoiceP and P introduces an Initiator, linear adjacency between the Initiator and passive verb licenses a variable deletion rule in the morphophonology.

To summarize this section, the Agent immediately following a passive *e*-verb is embedded in a PP by-phrase, even when no overt P introduces the Agent, and all *e*- clauses are canonical passives. I repeat previous examples here, using notation that indicates the P may be optionally pronounced, but the structures are identical.

- (42) *Jhuko’ jiyah e-kakan (bi’/moso) embi’ rowa.* (Familiar)
 fish that PV-eat by goat that
 ‘The fish was eaten by the goat.’
- (43) *Potra-epon e-tembhal-ih (sareng) ramah.* (Polite)
 son-Def PV-call-Appl by father
 ‘The son was called by father.’
- (44) *Maleng rowa e-tangkep (so) polisi.* (Familiar)
 thief that PV-catch by police
 ‘The thief was caught by police.’

4 Differences in nominal extraction

In section 3 I demonstrated that the voice system of polite Madurese is different than that of familiar Madurese, as polite speech employs three distinct voices while familiar speech only employs two. A natural question that follows is whether the registers display any other syntactic differences besides the availability of object voice. In this section I demonstrate that the two registers also deviate in another aspect of the

¹⁸ This general view is taken in Aldridge 2008, Arka and Manning 1998 and Nomoto and Kartini 2014, although these approaches also differ from one another.

grammar: the pattern of nominal extraction. Just as the familiar register is more restricted with regard to the number of voices it employs, I find that DP extraction is more restricted in familiar speech than in polite speech.

4.1 Both subject and object extraction in the polite register

In nominal extraction, the extracted argument occurs in initial position, separated from the rest of the clause by the relative morpheme *se* (strikethrough indicates the position of nominals before extraction or A'-movement).

- (45) Ramah se ~~ramah~~ ampon m-acah buku panekah. (Polite)
 father Rel Perf AV-read book that
 'It was Father who read that book.'

(45) can be interpreted as either a pseudo-cleft, 'It was Father who read that book' or a relative, 'the father that read that book'. Pseudo-clefts and relatives have the same surface form (but are disambiguated in speech by intonation and context). I assume that in (45), a null operator undergoes movement within a headless relative clause (following analyses of pseudo-clefts in related languages, e.g. Aldridge 2013, Massam 2003, Paul 2001, Pearson 2001). However, for ease of exposition in the discussion below, I refer to DP or argument movement, rather than 'operator movement'.

In the polite register, the grammatical subject may be extracted in active, passive and object voices:¹⁹

- (46) Paserah se ~~paserah~~ ampon m-acah buku panekah? (Polite)
 who Rel Perf AV-read book that
 'Who read that book?'

- (47) Ponapah se ~~ponapah~~ e-bacah?
 what Rel PV-read
 'What was read?'

- (48) Buku panekah se ~~buku panekah~~ ampon ramah bacah.
 buku that Rel Perf father OV-read
 'It was that book that father read.'

Extraction of a subject does not change the verbal morphology, which obligatorily reflects the voice of the clause in (46-48). In particular, note that the AV prefix is retained in (45) and (46), a fact that will contrast with object extraction in the discussion below.

Since a thematic Patient obligatorily raises to grammatical subject in both PV and OV (see 4b, 5b) direct extraction from object position can only be examined in AV clauses. In contrast to subject extraction, object extraction is ungrammatical when the AV prefix occurs on the verb.

¹⁹ One reviewer suggests that the Patient argument can be directly extracted from its VP-internal position in PV and OV, rather than obligatorily raising to subject position first (thus making Madurese similar to Philippine languages such as Tagalog). However, evidence from possessor movement shows that extraction is from subject position.

(i) Sapah se buku-nah e-bacah? (Familiar)
 who Rel book-Def PV-read
 'Who was it that (their) book was read?'

(ii) *Sapah se e-bacah buku-nah?
 who Rel PV-read book-Def

When the wh-possessor *sapah* is extracted from the passive clause (i), the possessum *bukunah* (the Patient) is stranded in subject position; it cannot remain in object position (ii). The Patient thus becomes a subject before A' extraction of the possessor. Similar facts hold in the Polite register; see Jeoung 2016.

(49) *Buku panekah se ramah ampon m-acah ~~buku panekah~~. (Polite)
 book that Rel father Perf AV-read
 ‘It was that book that Father read.’

(50) *Ponapah se ramah ampon m-acah ~~ponapah~~?
 what.Pol Rel father Perf AV-read
 ‘What did father read?’

However, if the verb does not have AV voice morphology, then the same clauses allow object extraction. Compare (49) and (50) with (51) and (52), which differ only in voice morphology:

(51) Buku panekah se ramah ampon bacah ~~buku panekah~~. (Polite)
 book that Rel father Perf AV.read
 ‘It was that book that Father read.’

(52) Ponapah se ramah ampon bacah ~~ponapah~~?
 what Rel father Perf AV.read
 ‘What did father read?’

Therefore, both subjects and objects may extract in polite Madurese, but object extraction requires that the verb be bare, i.e. without an AV prefix.

Note that (51) and (52) closely resemble OV clauses with bare verbs. Recall however that the surface word order Aux-Agent-Verb is diagnostic of OV (see 48), while Agent-Aux-Verb must be AV (51, 52). In principle, without the presence of a modal or auxiliary (i.e. *ampon*), OV clauses with extracted Patient/subject (48) and AV clauses with extracted Patient/object (51) cannot be differentiated. (But see section 5 regarding a strong preference for pronouncing a modal or aspectual morpheme in Madurese object voice.)

4.2 Subject extraction only in the familiar register

Subject extraction is licit in the familiar register, in both active and passive voices:

(53) Daud se ~~Daud~~ la m-acah buku jiyah. (Familiar)
 Daud Rel Perf AV-read book that
 ‘It was David who read that book.’

(54) Sapah se ~~sapah~~ ng-ato-eh ana’-eng?
 who Rel AV-call-Appl child-Def
 ‘Who called his child?’

(55) Apah se ~~apah~~ e-bacah bi’ Daud?
 what Rel PV-read by Daud
 ‘What was read by David?’

Unlike the polite register however, object extraction is ungrammatical in familiar speech, either with an active prefix (56) or bare verb (57).

(56) *Buku jiyah se Daud la m-acah ~~buku jiyah~~. (Familiar)
 book that Rel Daud Perf AV-read

(57) *Buku jiyah se Daud la bacah ~~buku jiyah~~.
 book that Rel Daud Perf AV.read

Note that the impossibility of object extraction with the bare verb in (57) contrasts directly with the licit object extraction in the Polite register (51), which has an identical structure.

Finally, the verb *bacah* ‘read’ is purposefully used throughout the examples because this particular root may occur in both polite and familiar registers. This removes the possibility that the availability of object voice, and furthermore the availability of object extraction, might be dependent on the properties of the verbal root. I have shown that *bacah* occurs in Polite object voice and Polite object extraction; yet the same structures with *bacah* in the Familiar speech are ill-formed. I conclude that it is not particular lexical roots which determine the difference between Polite and Familiar registers. In the following section, I discuss the source of the divergence between the registers.

5 Two grammars within Madurese

Many languages of Indonesia (e.g. Balinese, Javanese, Sundanese, and to a lesser extent, Indonesian) utilize two or more speech levels, which require (at least) a choice of lexical items appropriate to each register. For example, personal pronouns may have both a polite and familiar form. This is true of Madurese as well; many lexical items have a polite variant and a familiar variant, including pronouns, titles, determiners, prepositions and *wh*-words. Some nouns and verbs also have two forms that correspond to register. Yet the discussion above has shown that the speech levels in Madurese cannot be distinguished merely by replacing lexical items from the appropriate speech level.

The registers display at least two productive syntactic differences: first, the polite register has three morphosyntactically distinct voices, while the familiar register has only two; second, the polite register allows subject and object extraction, while the familiar register allows only subject extraction. Below, I propose that these differences arise from two distinct loci of variation in the grammar: the number of voices is determined by the number of available Voice heads in each register, while the extraction patterns are determined by the feature [D] on Voice.

5.1 Sources of variation between the registers

In the analysis below, two syntactic properties (voice and object extraction) arise from one source: the functional Voice heads available in each register. Literature on Austronesian languages suggests a strong connection between Voice and the movement of DP arguments (Aldridge 2008, Cole et al 2008, Legate 2012, 2014, Rackowski and Richards 2005, *inter alia*). Madurese is like other Indonesian-type languages that require null Voice marking when objects are extracted from AV clauses; the voice system is implicated when DPs move out of VoiceP.

First, consider polite Madurese. The data support the availability of three different Voice heads in polite speech. Merging the Voice head into the syntactic structure derives the configuration of its arguments. For instance, the passive Voice head does not select an external argument; the active Voice head selects an external argument in its specifier; the object Voice head selects an external argument from a restricted set of DPs. These Voice heads are instantiated as prefixes in the verbal morphology. The active Voice head is realized as the prefix *a-* or *N-*; the passive Voice head is realized as *e-*; the OV Voice head is phonologically null. Familiar Madurese, on the other hand, has a reduced inventory of only two functional Voice heads: active and passive. The OV Voice head is unavailable in the familiar grammar, so object voice clauses, with concomitant null voice morphology (i.e. bare verbs) are not derived in the familiar grammar.

While the number of available Voice heads accounts for the differences in the number of voices on each register, it does not account for the difference in nominal extraction patterns. Object extraction from AV clauses is available in polite speech, but not in familiar speech. Since both polite and familiar registers have AV, it cannot be the number of Voice heads that derives this difference. Rather, I propose that it is an edge feature on the active Voice head that is responsible for the difference in extraction patterns.

I assume that object DPs undergo successive cyclic movement, and that in Madurese, VoiceP is the relevant domain for movement (i.e. a phase). Cyclic or phase-based movement is discussed in much previous literature (see Chomsky 1986, 2000; as well as Aldridge 2008, Cole et al 2008, Legate 2008, Rackowski and Richards 2005, van Urk and Richards 2015 *inter alia* for analogous movement in related languages. The details of these proposals vary, but they share the general idea that VP internal arguments must first pass through the edge of vP/VoiceP before further movement.) Since voice marking is implicated by object movement in Madurese, I propose that DP extraction depends on the presence or absence of an edge feature [D] on an active Voice head. The [D] feature regulates DP movement from within VoiceP: if the feature is

present on the active Voice head, an internal object may move to the edge (specifier) of VoiceP, from where it is visible for further movement; but absence of the [D] feature means that the object cannot raise to the edge of VoiceP, and object extraction is impossible.

On this view, the variation in object extraction between the registers falls out from variation in the feature [D] on the active Voice head. In the Polite register, the active Voice head may have a [D] feature that attracts the object to the edge of VoiceP, from which position it is available for further movement to a high position in the clause. In contrast, the active Voice head in the Familiar register never bears the [D] feature, so objects cannot raise to the edge of the phase and cannot be extracted in Familiar speech.

Several implications follow. The active Voice head shares certain properties across the registers (i.e. selecting an external argument); it is only the featural makeup of this functional head that differs. The active Voice head in both registers can enter the derivation without [D]. In the Polite register only, the active Voice head may bear [D], in which case object extraction occurs. The [D] feature also has a phonological reflex: Voice⁰_[D] is phonologically null, while Voice⁰ without [D] is pronounced as *a-* or *N-*.

Another theoretical implication that arises from this discussion is the source of variation between the two registers. Madurese speakers who control both registers must allow for syntactic differences in each register which appear to affect the structure of entire clauses as well as various types of DP movement. However, the view presented here places the locus of the syntactic differences in the functional elements in each grammar, that is, the number of Voice heads and the features present on the active Voice head.

In Madurese, the register used by the speaker determines which inventory of functional items is available for a derivation. Using one inventory over another is essentially a theory of competing grammars (see Kroch 1989, 1994) within one language. That is to say, polite Madurese has one grammar, with an inventory of lexical items (including functional heads); the Familiar register has another grammar with different lexical items (notably, one less functional Voice head). In each grammar (register), the (un)availability of an OV Voice head determines whether object voice clauses can be derived.

Variation between the two grammars is also attributable to more fine-grained differences in the lexicon, i.e. features on functional heads. In the polite register, the AV Voice head is compatible with [D], allowing object extraction. For the same speaker, in familiar speech, an AV Voice head may not bear the [D] feature. The presence of these two distinct grammars, then, falls out from the number of functional elements (i.e. Voice heads) or micro-parametric differences on those functional elements (i.e. presence or absence of an edge feature) (the Chomsky-Borer conjecture; cited in Baker 2008).

5.2 Madurese voice in diachronic perspective

Madurese register differences present an interesting question, which is how the two registers have historically come to have distinct grammars, especially with respect to voice.

The discussion that follows must be speculative, since definitive evidence is lacking for a diachronic account. However, I note that previous work in the development of Indonesian-type languages has laid some initial groundwork. Indonesian-type languages are thought to retain some characteristics of an older symmetrical or Philippine-type voice system (see Himmelmann 2005, Wouk and Ross 2002 for discussion on the typology of Philippine-type and Indonesian-type voice.) In a symmetrical voice system, several different arguments can be made the prominent argument in a clause (including arguments that are otherwise thought to be oblique or non-core, such as instruments and locations); morphology on the verb indicates which argument is prominent.

Cole et al (2008) describe AV and OV as a symmetrical voice system ‘in miniature,’ while the canonical passive (PV) in Indonesian is thought to be a more recent addition in the historical development of the language. Cole et al investigate several colloquial varieties of Indonesian and Malay, concluding that the loss of a fully symmetrical voice system has resulted in a number of different voice systems across Indonesian/Malay languages. (See also Arka 2002, 2008, Arka and Wouk 2014, Yanti 2010 for several languages displaying varying degrees of symmetrical voice across Indonesia). Interestingly, the correlation between voice and object extraction is further supported in cross-linguistic comparisons. While object extraction with AV is not possible in a symmetrical-type voice system, it is possible in various languages that have developed a ‘mixed’ Indonesian-type system (Cole and Hermon 2005, Cole et al 2008, Saddy 1991, Voskuil 1996).

Madurese provides a unique situation in which to observe changes in voice: instead of comparing closely related language varieties, we have the opportunity to observe one language that displays two

synchronically distinct voice systems.²⁰ If, on parallel to Indonesian, Madurese AV and OV instantiate an (older) Philippine-type voice system, then PV is a later development. If this is correct, then the object voice might have been once employed in both familiar and polite registers, but was lost in the familiar register. This would mean that the polite register is the more conservative one, maintaining both AV and OV along with the more recent addition of PV.

5.3 Madurese OV is not vigorous

In this section I make note of some unusual aspects of Madurese OV, which taken together suggest that object voice is not as vigorous as AV and PV. To begin, Madurese OV clauses are not compatible with negation. (14) is repeated here as (58).

- (58) *Potra-epon ampon ramah tembhal-ih.* (Polite)
 son-Def Perf father OV.call-Appl
 ‘Father called his son.’

All AV and PV clauses may be negated, but an OV clause is not compatible with negation:

- (59) **Potra-epon {lo’ / ghilo’} ramah tembhal-ih.*
 son-Def Neg not.yet father OV.call-Appl
 ‘Father {did not/ has not yet} called his son.’

Note that OV clauses in Indonesian and other languages are compatible with negation, so this is a surprising restriction on OV in Madurese.

Furthermore, in Madurese AV or PV clauses, an aspectual morpheme is optional. The AV clause in (60) can receive a perfective interpretation whether or not *ampon* occurs.

- (60) *Ramah (ampon) n-embhal-ih potra-epon.* (Polite)
 father Perf AV-call-Appl son-Def
 ‘Father called his son.’

However, without a modal or aspectual morpheme occurring before the Agent, OV clauses are degraded.

- (61) ??*Potra-epon ramah tembhal-ih.* (Polite)
 son-Def father OV.call-Appl
 ‘Father called his son.’

When considering (61), Madurese consultants strongly prefer to pronounce the morpheme *ampon* as in (58). Since *ampon* is not obligatory in the other voices, and adds little to the semantics, it is remarkable that it is strongly preferred in OV. Similar morphemes are not required in OV in other languages, such as Indonesian.

- (62) *Buku itu (udah) aku baca.* (Indonesian OV)
 book that Perf I OV.read
 ‘I read that book.’

(62) is well-formed with or without the perfective morpheme *udah*, which is not required for a perfective reading in Indonesian, and is not obligatory in any voice. In Indonesian, the AV morpheme is optionally pronounced in many clauses: this means that (62) is ambiguous between an OV clause and an AV clause with topicalized object and unpronounced AV prefix. This is not the case with Madurese however. (61) is ill-formed as an AV clause with extracted object because an extracted object must occur with the relative

²⁰ All speakers may not control the polite register; however, the discussion here considers those speakers whose native control of Madurese includes both registers.

morpheme *se*. Furthermore, AV prefixes are not optional in Madurese, so (61) should unambiguously be interpreted as an OV clause. However, my consultants judge the clause marginal without *ampon*; the aspectual morpheme appears necessary to signal that the clause is OV.

Without corpus data showing that younger speakers of Madurese use OV much less frequently than older speakers, or other diachronic evidence, it is an open question whether OV is in danger of being lost in the Polite register of Madurese. I submit that it appears that the use of OV is restricted compared to the other voices. OV is restricted to polite speech only; OV clauses cannot be negated; and aspectual morphemes are needed to signal that the clause is OV. Recall from section 3.2 that Agent questions are also not possible in OV. These facts indicate that in Madurese, object voice is not as vigorous as AV and PV.

6 Conclusion

I have demonstrated that Madurese has three morphosyntactically defined voices: an active voice, a canonical passive voice and a previously undocumented object voice. All three voices are available in the polite register, but the familiar register does not have object voice, resulting in only two voices. The grammatical difference between the registers also extends to nominal extraction: the polite register allows both subjects and objects to extract, but the familiar register allows only subjects to extract. Thus the registers of Madurese are differentiated not only by lexical items, but in the syntax as well: two distinct grammars operate in the polite and familiar registers. Syntactic differences are formalized as variation in the functional head Voice and its ability to have a [D] feature that raises objects to the edge of VoiceP. These findings contribute to an understanding of synchronic variation across two registers of the same language.

References

- Aldridge, Edith. 2008. Phase-based account of extraction in Indonesian. *Lingua* 118:10, 1440-1469.
- Arka, I Wayan. 2002. *Voice systems in the Austronesian languages of Nusantara: Typology, symmetry and undergoer orientation*. Accessed Dec 2, 2015 from <http://hdl.handle.net/1885/41058>.
- Arka, I Wayan. 2003. *Balinese Morphosyntax: A Lexical-functional Approach*. The Australian National University.
- Arka, I Wayan. 2008. Voice and being core: Evidence from (eastern) Indonesian languages. in *Voice and Grammatical Relations in Austronesian Languages*, eds. P.K. Austin and S. Musgrave. CSLI Publications.
- Arka, I Wayan and Christopher Manning. 1998. On the three subjects in Indonesian: Evidence from binding. On-line Proceedings of the LFG98 Conference, csli-publications.stanford.edu/LFG/3/lfg98.html.
- Arka, I Wayan and Christopher Manning. 2008. Voice and grammatical relations in Indonesian: A new perspective. *Voice and Grammatical Relations in Austronesian Languages*, P. K. Austin and S. Musgrave, eds., 45-69. CSLI Publications.
- Arka, I Wayan, and Fay Wouk. 2014. Voice-related constructions in the Austronesian languages of Flores. *Argument Realisations and Related Constructions in Austronesian Languages: Papers from 12-ICAL*, W. Arka, N. Indrawati, eds., 313-333. Asia-Pacific Linguistics.
- Baker, Mark. 2008. *The Syntax of Agreement and Concord* (Vol 115). Cambridge University Press.
- Bobaljik, Jonathan. 1995. *The syntax of Verbal Inflection*. Ph.D dissertation, MIT.
- Borer, Hagit, 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Foris Publ.
- Chomsky, Noam. 1986. *Barriers* (Vol. 13). MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program* (Vol. 28). MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. *Step By Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, Roger Martin, David Michaels, and Juan Uriagereka, eds., 89-155. MIT Press.
- Chung, Sandra. 1976. On the subject of two passives in Indonesian. *Subject and Topic*, C.N. Li, ed. Academic Press.
- Cole, Peter and Gabriella Hermon. 2005. Subject and non-subject relativization in Indonesian. *Journal of East Asian Linguistics* 14:1, 59-88.

- Cole, Peter, Gabriella Hermon and Yassir Tjung. 2006. Is there pasif semu in Jakarta Indonesian? *Oceanic Linguistics*, 64-90.
- Cole, Peter, Gabriella Hermon and Yanti. 2008. Voice in Malay/Indonesian. *Lingua* 118:1500-1553.
- Dardjowidjojo, Soenjono. 1978. *Sentence Patterns of Indonesian*. University of Hawaii Press.
- Davies, William D. 1999. Madurese and Javanese as strict word-order languages. *Oceanic Linguistics*, 152-167.
- Davies, William D. 2003. Extreme Locality in Madurese Wh-Questions. *Syntax* 6:3, 237-259.
- Davies, William D. 2008. Madurese reflexives with reference to Malay/Indonesian. *Lingua* 118: 1603-1616.
- Davies, William D. 2010. *A Grammar of Madurese*. Walter de Gruyter.
- Harley, Heidi. 2013a. External arguments and the Mirror Principle: On the distinctness of Voice and v. *Lingua* 125, 34-57.
- Guilfoyle, Eithne, Henrietta Hung and Lisa Travis. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10, 375-414.
- Haspelmath, Martin, ed. 2001. *Language Typology and Language Universals*. Walter de Gruyter.
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. *The Austronesian languages of Asia and Madagascar*, K.A. Adelaar and N.P. Himmelman, eds., 110-181. Routledge.
- Jeoung, Helen. 2016. Possessor extraction in Indonesian-type languages. *Proceedings of the Linguistic Society of America* 1, 35:1-15.
- Jeoung, Helen and Alison Biggs. To appear. Variants of Indonesian Prepositions as Intra-speaker Variability at PF. *Penn Working Papers in Linguistics*.
- Killiaan, H. N. 1897. *Madoereesche Spraakkunst*. Batavia: Landsdrukkerij.
- Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese*. MIT Press.
- Legate, Julie Anne. 2012. Subjects in Acehnese and the nature of the passive. *Language* 88:3, 495-525.
- Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2015. *Ethnologue: Languages of the World, 18th edition*. SIL International. Online version: <http://www.ethnologue.com>.
- Levin, Theodore F. 2015. *Licensing without Case*. Ph.D dissertation, MIT.
- Macdonald, R. Ross and Soenjono Dardjowidjojo. 1967. *Indonesian Reference Grammar*. Georgetown University Press.
- McGinnis, Martha. 1998. Locality and inert Case. In *Proceedings of NELS* 28, 267-281.
- McKinnon, Timothy, Peter Cole and Gabriella Hermon. 2011. Object agreement and 'pro-drop' in Kerinci Malay. *Language* 87:4, 715-750.
- Nomoto, Hiroki and Kartini Abd Wahab. 2014. Person restriction on passive agents in Malay: Information structure and syntax. *NUSA: Linguistic Studies of Languages in and around Indonesia* 57: 31-50.
- Pawitra, Adrian. 2008. *Kamus Lengkap Bahasa Madura-Indonesia*. Jakarta: Dian Rakyat.
- Pylkkänen, Liina. 2002. *Introducing Arguments*. PhD dissertation, MIT.
- Pylkkänen, Liina. 2008. *Introducing Arguments* (Vol. 49). MIT press.
- Rackowski, Andrea and Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36:4, 565-599.
- Saddy, Douglas. 1991. WH scope mechanisms in Bahasa Indonesia. *More papers on WH Movement*, L. Cheng and H. Demirdache, eds., 183-218. *MIT Working Papers in Linguistics* 15.
- Safioedin, Asis, ed. 1977. *Kamus Bahasa Madura-Indonesia*. Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan.
- Sneddon, James Neil. 2006. *Colloquial Jakartan Indonesian* (Vol. 581). Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- Sneddon, James Neil, K. Alexander Adelaar, Dwi N. Djenar, and Michael Ewing. 2012. *Indonesian: A Comprehensive Grammar*. Routledge.

- Stevens, Alan. 1968. *Madurese Phonology and Morphology*. New Haven, Connecticut: American Oriental Society.
- Voskuil, Jan E. 1996. *Comparative Morphology: Verb Taxonomy in Indonesian, Tagalog and Dutch*, Vol. 21. Holland Academic Graphics.
- Wouk, Fay and Malcolm Ross. 2002. *The History and Typology of Western Austronesian Voice Systems*. Pacific Linguistics, Research School of Pacific and Asian Studies, the Australian National University.
- Yanti. 2010. *A Reference Grammar of Jambi Malay*. Ph.D dissertation, University of Delaware.

Reviewed: Received 14 March 2016, revised text accepted 3 January 2017, published 1 February 2017
Editors: Editor-In-Chief Dr Mark Alves | Managing Eds. Dr Paul Sidwell, Dr Nathan Hill, Dr Sigrid Lew