

DE GRUYTER  
MOUTON

*William D. Davies*

# A GRAMMAR OF MADURESE

# A Grammar of Madurese

# Mouton Grammar Library

## 50

Editors

Georg Bossong

Bernard Comrie

Matthew Dryer

De Gruyter Mouton

# A Grammar of Madurese

by

William D. Davies

De Gruyter Mouton

ISBN 978-3-11-022443-6  
e-ISBN 978-3-11-022444-3  
ISSN 0933-7636

*Library of Congress Cataloging-in-Publication Data*

Davies, William D., 1954–  
A grammar of Madurese / by William D. Davies.  
p. cm. – (Mouton grammar library; 50)  
Includes bibliographical references and index.  
ISBN 978-3-11-022443-6 (alk. paper)  
1. Madurese language – Grammar. I. Title.  
PL5352.D385 2010  
499'.22345–dc22

2010028789

*Bibliographic information published by the Deutsche Nationalbibliothek*

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;  
detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

© 2010 Walter de Gruyter GmbH & Co. KG, 10785 Berlin/New York

Printing: Hubert & Co. GmbH & Co. KG, Göttingen  
∞ Printed on acid-free paper

Printed in Germany

[www.degruyter.com](http://www.degruyter.com)

for Patty



## Acknowledgments

The influence of many people is manifested in numerous ways in the pages that follow.

Eschewing time-honored tradition, I would first and foremost like to recognize the inestimable contribution of my family. Patty, Billy, and Kate provided vast quantities of moral and physical support. They showed great patience in the face of my sometimes inexplicable passion for this subject, actually encouraging it. They endured the absence of husband, father, coach, companion, playmate, crossword puzzle chum for a couple months each year while I was off gathering data; at least they did not complain about it very strenuously. Thanks guys, you really are the best.

Surachman Dimiyati introduced me to the Madurese language, first as my research assistant, later my colleague in many of these endeavors, but always as my friend, opening numerous paths both physical and otherwise for the past 15 years.

Many speakers of Madurese endured my virtually unending string of strange queries, my aberrant pronunciation, and my occasional confusion with great equanimity. I am most grateful for the enthusiasm they showed for the project. All were wonderful, but deserving of special mention are Masharudin, Moh. Hasan Sasra, Ruskawi Sjah, and Sugiyono, whom were generous with their time and energy and willing to set me straight on more than one occasion. Thanks to Sri Mulyati, Sarwo Santuso, Mauliyati, Mohammad Ridwan, Laila Qodriyah Ervan, and A. Haitul Umam for their important contributions to my understanding of their language. And thanks to ‘the boys’ at the Hotel Ningrat in Bangkalan; they not only contributed linguistic and personal information through our attempts to understand one another but also made my stays on Madura all the richer (not to mention once helping me out of a jam).

I’d like to thank those folks who read a chapter and sometimes much, much more as the work progressed. Jill Beckman, Stan Dubinsky, Ivan Ivanov, Kumyoung Lee, Ileana Paul, Cathie Ringen, and Roumyana Slabakova all made important suggestions, many of which found their way into this volume. In addition to reading the entire manuscript, Kumyoung provided much needed editorial assistance and frequently rescued me from my lack of mastery of the word processing software, and special thanks to Eri Kurniawan and Stan for some technical stuff. Tom Frakes whipped up some nice maps on very short notice. Bernard Comrie, who apparently drew the short straw, served as series editor for the book, providing invaluable feedback through his careful reading of the manuscript. Now adhering to time-honored tradition, I reluctantly release all of

these fine folks from any responsibility for errors, shortcomings, and misapprehensions in what follows, all of which I claim as my own.

Doing this kind of work can require considerable financing. The National Science Foundation (through grant number SBR 98-09044 to the University of Iowa), the Fulbright Foundation through the Fulbright-Hays program, and the United States-Indonesia Society provided much needed support. Closer to home, the University of Iowa has been extremely generous. Funding was received from the Arts and Humanities Initiative, the May Brodbeck Fellowship in the Humanities, the Obermann Center for Advanced Studies, the Stanley International Programs-Obermann Center Research Fellowship, the Social Sciences Funding Program, the College of Liberal Arts Summer Fellowship for Senior Faculty, and the Center for Asian and Pacific Studies.

My sincere thanks to all.

# Contents

Acknowledgments	vii
Abbreviations	xvi
<b>Chapter 1 Introduction</b>	1
1. Madura and its people	1
2. The language	4
3. Previous studies	6
4. Methodology and framework	8
4.1 Framework	9
<b>Chapter 2 Phonology</b>	11
1. Basic inventory and description of sounds	11
1.1 Consonants	11
1.2 Vowels	18
2. Sound correspondences between Madurese and Indonesian (and Javanese)	20
2.1 Consonants	20
2.2 Vowels	23
3. Syllable structure and phonotactic constraints	25
4. Vowel inventory and distribution	29
4.1 Vowel harmony	29
4.2 Other vowel processes	38
5. Phonological processes	40
5.1 Epenthesis	40
5.2 Gemination	43
5.3 Deletion	45
6. Morphophonemic processes	46
7. Stress	51
8. Orthography	52
<b>Chapter 3 Lexical categories</b>	61
1. Open class categories	62
1.1 The category noun	62
1.2 The category verb	67
1.3 The (sub)category adjective	70
1.3.1 Adjective-verb similarities	70

1.3.2	Potential adjective-verb differences	72
1.3.3	A subcategory adjective	74
1.4	Precategorial roots	76
2.	Closed class categories	78
2.1	Adverbs: straddling the open/closed dichotomy	78
2.2	Auxiliaries	80
2.3	Prepositions	81
2.4	Pronouns	82
2.5	Demonstratives	84
2.6	Numerals and quantifiers	85
2.7	Interrogatives	88
2.8	Conjunctions	89
2.9	Particles and interjections	90
2.10	Clitics	92
<b>Chapter 4</b>	<b>Morphology</b>	97
1.	Affixation	97
1.1	Verbal morphology	99
1.1.1	Actor voice <i>ng-</i> , /ŋ/	99
1.1.2	Actor voice <i>a-</i> , /a/	101
1.1.3	Object voice <i>e-</i> , /ɛ/	102
1.1.4	Involitive <i>ta-</i> , /ta/	103
1.1.5	Verbal <i>ka-</i> , /ka/	103
1.1.6	Adversative <i>ka-...-an</i> , /ka...an/	104
1.1.7	Causative <i>pa-</i> , /pa/	105
1.1.8	Irrealis <i>-a</i> , /a/	105
1.1.9	Iterative <i>-an</i> , /an/	105
1.1.10	Durative <i>-e</i> , /ɛ/	106
1.1.11	'Locative' <i>-e</i> , /ɛ/	106
1.1.12	'Benefactive' <i>agi-</i> /ak <sup>h</sup> ɛ/	107
1.2	Nominal morphology	109
1.2.1	Definite <i>-na</i> , /na/	109
1.2.2	Result nominals with <i>-an</i> , /an/	112
1.2.3	Nominal <i>ka-...-an</i> , /ka...an/	112
1.2.4	Agentive and instrumental <i>pa-</i> /pa/ and <i>pang-</i> /paŋ/	113
1.2.5	Locative/agentive <i>pa-...-an</i> , /pa...an/	114
1.2.6	<i>pang-...-an</i> , /paŋ...an/	114
1.2.7	<i>pa-</i> with process nominals	115
1.2.8	<i>sa-</i> , /a/	115
1.2.9	<i>sa-</i> in nominalizations	116
1.3	Derived adjectives	117
1.4	Derived adverbs	118

1.5 Infixes	119
2. Co-occurrence of affixes	120
3. Extensions	125
4. Compounds	127
<b>Chapter 5 Reduplication</b>	129
1. Reduplication types	129
2. Reduplication with other affixes	131
3. Functions of reduplication	137
3.1 Reduplication of verbs	138
3.1.1 Multiple actions	138
3.1.2 Imperfect aspect	139
3.1.3 Emphasis	140
3.1.4 Reduplication of adjectives	142
3.2 Reduplication of nouns	143
3.2.1 Temporal expressions	144
3.2.2 Additional functions	144
3.3 Indefinite expressions	146
3.4 Adverbs	146
3.5 Numbers	147
3.6 Compounds	147
<b>Chapter 6 Clause types</b>	149
1. Basic word order	149
2. Non-verbal clauses	151
3. Existential clauses	154
4. Verbal clauses	157
4.1 Intransitive clauses	157
4.2 Semantically-transitive statives	159
4.3 Transitive clauses	162
4.4 Ditransitive clauses	163
5. Exclamatory clauses	164
6. Hortative clauses	165
7. Imperative clauses	167
8. Comparative and superlative clauses	170
9. Topic-comment structure	175
10. Subjects, definiteness and specificity	177
<b>Chapter 7 Nominals and noun phrases</b>	181
1. Nominal types	181
2. Nominal morphology	181
2.1 Nominal inflection	182

2.2 Nominal derivation	183
3. Pronouns	184
3.1 Distribution of pronouns	189
4. Demonstratives	189
5. Possession	193
6. Quantifiers	194
7. Numbers	199
8. Measure phrases	205
9. Adjectival modification	207
10. Prepositional phrase modifiers	210
11. Nominalizations and clausal NPs	211
12. Constituent order in NPs	214
<b>Chapter 8 Prepositions and prepositional phrases</b>	<b>225</b>
1. Basic prepositions	225
1.1 <i>e</i> and <i>neng</i> ‘at’	225
1.2 <i>dha'</i> and <i>ka</i> ‘to’	227
1.3 <i>dhari</i> ‘from’	228
1.4 <i>bi'</i> and <i>moso</i> ‘with’	229
1.5 <i>sampe'</i> and <i>gan</i> ‘until’	231
1.6 <i>akantha</i> ‘like’	234
2. Derived prepositions	235
2.1 <i>kaangguy</i> ‘for’	235
2.2 <i>ngangu</i> instrumental ‘with’	236
2.3 <i>kalaban</i> and <i>banding</i> ‘than’	236
2.4 <i>katembang</i> and <i>moso</i> ‘with’	239
2.5 <i>lebat</i> ‘through’	240
2.6 <i>parkara</i> and <i>hal</i> ‘concerning’, ‘in regard to’	240
2.7 <i>sabab/bab</i> ‘about’	240
2.8 <i>sabellunna</i> ‘before’, <i>sa'ellana</i> ‘after’, <i>samarena</i> ‘after’	241
3. Complex locative expressions	241
4. Nominal morphology	245
4.1 Verb-preposition combinations	245
4.2 Prepositional marking of direct objects	247
<b>Chapter 9 Verb phrases and verbal marking</b>	<b>249</b>
1. Voice	249
1.1 Actor voice	250
1.1.1 Actor voice with intransitive roots	253
1.1.2 The unaccusative/unergative split	255
1.2 Object voice	256
1.3 Further manifestations of voice	259

2. Irrealis mood	260
3. Aspectual <i>-an, -e</i> , reduplication	264
4. Auxiliaries and adverbial tense, mood and aspect marking	267
5. Negation	273
6. Additional voice constructions	275
6.1 Resultative/abilitive <i>ka-</i>	275
6.2 Involitive	277
6.3 Adversative	279
<b>Chapter 10 Modifications to clause structure</b>	<b>283</b>
1. ‘Locative’ <i>-e</i>	283
1.1 Ditransitives	283
1.2 ‘Give’	286
1.3 Verbs of communication	289
1.4 Verbs of cognitive state	291
1.5 Dynamic intransitive verbs	295
1.6 Stative intransitive verbs	297
2. The suffix <i>-agi</i>	299
2.1 Benefactives	299
2.2 Verbs of communication	304
2.3 <i>-agi</i> specifying an endpoint	307
2.4 <i>-agi</i> with instruments	308
2.5 Intransitive predicates and causativization	311
2.6 Polite imperatives	313
2.7 Miscellaneous instances	314
2.8 The meaning/function of <i>-agi</i> and Indonesian <i>-kan</i>	315
3. Causatives	315
3.1 Causatives of stative intransitive verbs	317
3.2 Causatives of dynamic intransitive verbs	318
3.3 Causatives of syntactically transitive verbs	321
3.4 Patterns of acceptability	323
3.5 Structural variations	325
4. The role of <i>-e</i> and <i>-agi</i>	327
<b>Chapter 11 Complex sentences</b>	<b>335</b>
1. Coordination	335
1.1 Subclausal coordination	338
2. Complementation	341
3. Sentential subjects	346
4. Control	350
5. Aspectual predicates	354
6. Prolepsis	357

7. Relative clauses	365
8. Complements of nominals	373
9. Clefts	374
9.1 The presuppositional clause	376
9.2 Predicative nature of the clefted element	381
9.3 Clefting adjuncts	383
10. Direct speech	385
<b>Chapter 12 Adverbs and adverbial clauses</b>	<b>389</b>
1. Types of adverbial expressions	389
1.1 Temporal expressions	389
1.2 Frequency adverbs	392
1.3 Epistemic adverbials	393
1.4 Manner adverbials	394
1.5 Degree modifiers	396
2. Derivation of adverbs	397
3. Adverbial clauses	398
3.1 Purpose clauses	398
3.2 Reason clauses	400
3.3 Conditional clauses	401
3.4 Concessive clauses	404
3.5 Temporal clauses	405
3.5.1 <i>sa-</i> adverbials	411
3.6 Parataxis	412
3.7 Correlative structures	413
4. Adjuncts in complement clauses	413
<b>Chapter 13 Anaphora</b>	<b>417</b>
1. Reflexives	417
1.1 The uses of <i>dibi'</i>	418
1.2 The distribution of <i>aba'na dibi'</i>	421
1.3 Reflexive interpretation without <i>dibi'</i>	423
1.4 NP + <i>dibi'</i> reflexives	425
2. Reciprocals	425
2.1 The components of reciprocals	427
2.2 A special property of <i>saleng</i> , <i>patang</i> and <i>gante</i>	429
2.3 Stative predicates	430
3. Null pronouns	431
4. Cross-clausal null anaphora	433
5. Ellipsis and gapping	436

<b>Chapter 14 Questions</b>	439
1. Yes/no questions	439
2. Alternative questions	442
3. Tag questions	443
4. Constituent questions	444
4.1 In situ questions	445
4.2 Clefted questions	448
4.3 Fronted questions	449
4.3.1 Adverbial questions	449
4.3.2 Prepositional questions	450
4.3.3 <i>barampa</i> ‘how much/how many’	451
5. Long-distance questions	453
5.1 Long-distance in situ questions	453
5.2 ‘Partial’ movement questions	455
5.3 Long-distance ‘movement’	455
5.4 Evidence against long-distance movement	456
5.5 Embedded <i>wh</i> -adjuncts and movement	459
5.6 Long-distance questions and island effects	461
6. Embedded questions	463
7. The question particle <i>ba</i> ’	465
<b>Chapter 15 Speech levels</b>	469
1. The vocabularies	470
2. Use of the vocabularies	473
2.1 Kasar, tengnga'an and alos	473
2.2 Alos tenggi and alos mandhap	475
2.3 Identifying distinct speech levels	479
3. Textual examples	482
4. Implementation of the speech levels	487
<b>Chapter 16 Texts</b>	491
1. Parembi'na Joko Tole	491
2. Radin Sagara	508
3. Siti Lalumba	533
4. Kerrapan Sape	557
<b>References</b>	569
<b>Index</b>	581

## Abbreviations

AD	adversative
AGI	<i>-agi</i> morpheme
AUX	auxiliary
AV	actor voice
BEN	beneficiary
COMP	complementizer
CONJ	conjunction
CS	causative
DEF	definite
DUR	durative
E	<i>-e</i> (locative) morpheme
EMPH	emphatic
EXCL	exclamatory/exclamation
HORT	hortative
IMP	imperative
IN	involitive
IRR	irrealis
IT	iterative
KA	<i>-ka</i> (abilitive/result) morpheme
LOC	locative
NEG	negative
NOM	nominalization
ORD	ordinal
OV	object voice
PERF	perfective
PL	plural
Q	question particle
RED	reduplication
REL	relative marker

# Chapter 1

## Introduction

### 1. Madura and its people

The Madurese people are originally indigenous to the island of Madura, located in the Java Sea just north of East Java (see map 1, following page), and Madura is part of the province of Jawa Timur (East Java). Due to poor soil conditions and a low annual rainfall, with an area of 4250 km<sup>2</sup>, the island was unable to sustain a population of more than 3-4 million or so. Thus, there is a long history of Madurese emigrating to other areas of Indonesia, particularly East Java, where agricultural conditions and other economic opportunities were markedly superior, resulting in large, permanent populations in such cities as Besuki, Situbundo, Probolinggo, Jember, Surabaya, and elsewhere. As of 2000, roughly 3.5 million Madurese live on Madura itself, meaning that there is a sizable population of Madurese-speaking people living in East Java and other areas of Indonesia, including significant populations on the islands of Kalimantan and

When suffixed to certain nouns and adjectives, the result is a derived adverb, as in *Madurese-speaking*. When used as a noun, it refers to the people of Madura. The Madurese language is a member of the Austronesian language family, and is one of the most widely spoken languages in Indonesia (the national language, Bahasa Indonesia, being the largest).<sup>2</sup> The Madurese people are originally indigenous to the island of Madura, located in the Java Sea just north of East Java (see map 1, following page), and Madura is part of the province of Jawa Timur (East Java). Due to poor soil conditions and a low annual rainfall, with an area of 4250 km<sup>2</sup>, the island was unable to

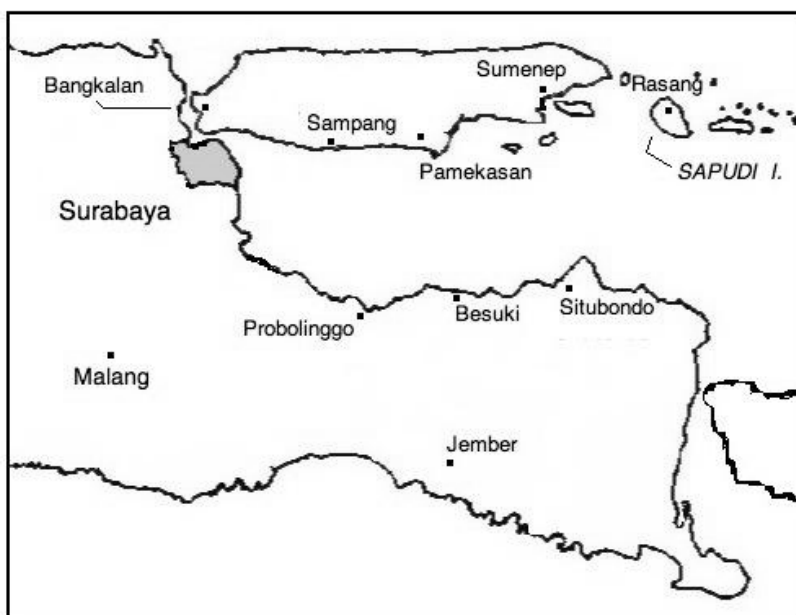
---

<sup>1</sup> Suryadinata, Arfin, and Ananta (2003) contend that there are roughly 200,000 more Malays than Madurese, arriving at this figure by combining 13 distinct categories of Malays recognized in the official census (Badan Pusat Statistik 2000).

<sup>2</sup> Other reports put the number of Madurese considerably higher. *Ethnologue* (Gordon 2005) gives a figure of 13.7 million, while Nothofer (2006) cites 13 million, and Pawitra (2009) 15 million. Others provide yet different numbers. Suryadinata, Arfin, and Ananta (2003) note that the percentage increase in Madurese population from the 1930 census to the 2000 census falls well below that of other groups and of the country as a whole, speculating that some may not identify themselves as Madurese for various reasons.

Sumatra. As is true of many ethnic groups in Indonesia, considerable Madurese enclaves are also found in Jakarta, Bandung, and other large cities in Central and West Java.

According to Husson (1997:80), early Madurese migration to East Java began in the 13th century, serfs that were sent “to colonize agricultural lands and strengthen the Majapahit kingdom”. Additionally, their reputation as fierce combatants led to the voluntary enlistment or forced conscription of soldiers,



Map 1: East Java

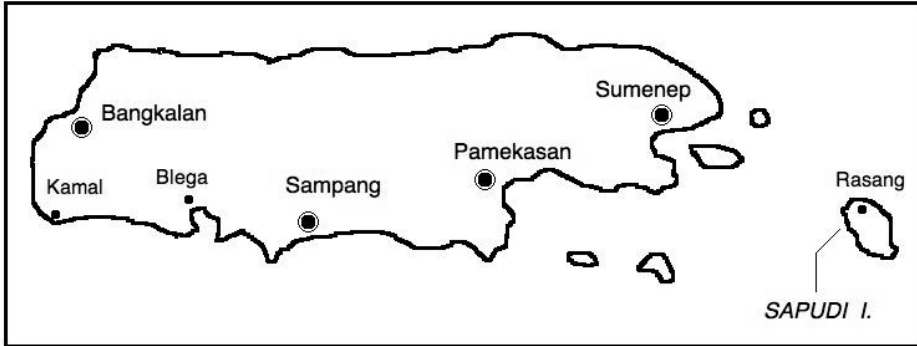
accounting for other waves of Madurese leaving the island during the period from the 16<sup>th</sup> to 19<sup>th</sup> centuries and eventually settling on Java. Added to this were the shifting alliances of the different kingdoms or sultanates on Madura with the Mataram empire and the Dutch colonialists.<sup>3</sup> There were also patterns of migration to East Java in search of work and a source of income due to limited opportunities on Madura (Husson 1997:86). More recently, attempts by the Republic of Indonesia government to redistribute the national population to ease crowding and fuel economic development through a program of transmigration account for the establishment of significant Madurese populations on the islands of Kalimantan and Sumatra (in Riau Province) where the Madurese have sought additional economic opportunities. There is thus a long history of the migration from Madura.

However, most previous works on the Madurese language and people

---

<sup>3</sup>See Husson (1997) for details of the political situation.

have focused on Madura Island and the 67 small nearby islands including Bawean, Sapudi, Kangean, and others. Madura proper is comprised of the regencies (*kabupaten*) of Bangkalan, Sampang, Pamekasan, and Sumenep, moving eastward on the island (see map 2). Although Sumenep is considered the cultural capital, each regency had its own political institutions and various local influences.



Map 2: Madura

As stated above, Madura itself presents a somewhat inhospitable environment. Compared to the fertile soil and abundant rainfall characteristic of Java, Sumatra, and Kalimantan, Madura receives far less rainfall and has poor chalky soil. This results in extremely unfavorable conditions for agriculture and an inability to sustain a large population. Maize was traditionally the principal crop, supplemented by beans and some other vegetable crops. Raising livestock has been critical to the subsistence economy, particularly cattle and goats. (The Madurese are well-known for the tradition of *kerapan sape* ‘bull racing’, and the annual event draws spectators to the island from a wide region. A description of the tradition is included in the texts in Chapter 16.) Given the conditions present on the island, a tradition of fishing and seafaring developed, and many Madurese became traders. Those Madurese who migrated to Java and other parts of Indonesia (to escape these conditions on Madura) largely worked as laborers and food vendors.

The tradition of trading created close ties with the *pasisir* kingdoms on the northern coast of Java, and these ties were important in the rapid introduction of Islam to Madura in the early 16<sup>th</sup> century. This was hastened further by a political system which made the more egalitarian religion of Islam particularly appealing to the large proportion of the population that did not belong to the nobility in the courts of the three principal kingdoms of Sumenep, Pamekasan, and Bangkalan. According to Mansurnoor (1990), as Islam spread, the religious leaders, the *kyai*, increasingly assumed positions of standing and some power as the kings, the *rato*, recognized the wisdom of officially embracing the popular

religious leaders as a way of maintaining their power. Subsequently, when the *rato* were stripped of their power by the Dutch in the late 1800's, the *kyai* filled the leadership vacuum and became a central focus of authority in the various villages and larger governmental units, thus taking on an important role in the governance system. As a result, the Madurese people are almost entirely Muslim, with an admixture of local customs, and a strong commitment to Islam is at the core of the Madurese identity and their lives. Traditionally, daily life in the villages in the countryside and enclaves in larger cities revolved around an Islamic center, under the leadership of an *ulama* or *kyai*, and the associated boarding school, *pesantren*, played a pivotal role in society, particularly to the education of young Madurese men, and does so in many Madurese communities throughout Jawa Timur today. Of course, with rapid transitions in modern Indonesian society, any sweeping generalization is difficult to maintain.

However, sweeping generalizations and stereotypes have for centuries plagued the Madurese and play a critical role in understanding their position in Indonesian society. The hard and uncertain livelihood and difficult conditions on Madura resulted in a tough-minded temperament and independent character, a character infused with fealty to family, friends, and leaders. A tendency toward straightforward, direct expression has generally been viewed as less refined and polite by some other groups, particularly on the main island of Java—this despite the fact that the Madurese language shares with Javanese, Sundanese, Balinese, and some other languages of the area the intricate system of politeness vocabularies and rules for their use commonly referred to as ‘speech levels’ (described in Chapter 15). But this reputation for straightforward expression, combined with the tradition of *carok* (a system of ritualized violent vengeance operating outside of the recognized legal system<sup>4</sup>) has created a situation succinctly summed up by de Jonge (1995:7), “There are few ethnic groups in the Indonesian archipelago about whom as many negative and equivocal stereotypes exist as about the Madurese...it is striking that the current image of the *suku Madura* [‘Madurese tribe’] in Indonesia today hardly differs from the one present during the colonial period.” As described below, this has contributed to what has been until relatively recently as relatively small Western literature on the Madurese people and their language.

## 2. The language

Madurese is a member of the Malayo-Sumbawan subgroup of the Western Malayo-Polynesian branch of the Austronesian language family (Adelaar 2005; Gray, Drummond and Greenhill 2009). Western Malayo-Polynesian in-

---

<sup>4</sup>See Smith 1997, van Dijk 2002, and Wiyata 2002 for descriptions and discussion of *carok*.

cludes some 500+ languages of western Indonesia, the Philippines, and the Malagasy language of Madagascar. On the basis of lexical and phonological evidence, it has been determined that Madurese is most closely related to Sundanese, Balinese, Malay/Indonesian, Sasak, Sumbawa, and Chadic, which is evident its morphology and syntax as well.

This work might perhaps be better titled *A Grammar of the Language of Bangkalan, Madura* because the Madurese described here is predominately the Madurese spoken in Western Madura, more specifically that in the regency of Bangkalan. In this, this work differs from most that predate it. The Eastern dialect is the recognized standard, the Madurese spoken in the regency of Sumenep, where the principle palace of the Majapahit era and earlier 'dynasties' is located. The attitude regarding the prestige dialect is widely held, even among the people of other dialect areas. However, as is true of any language with a significant number of speakers spread over a large geographical area, identifying a homogeneous standard can actually be somewhat controversial, or, at least, misleading. The fact is that the majority of Madurese live outside the island of Madura and speak distinct forms of the language. All, however, are speakers of Madurese. The situation is not unlike that found with most languages, and closely parallels the case of Javanese: Central Javanese, the language of the royal court in Solo and Yogyakarta, is taken to be the standard and differs in a number of respects from the Javanese spoken elsewhere on the island.

Speakers of the Western dialect tend to characterize themselves as very straightforward and direct, saying what is on their minds without equivocation (in a way reflecting what some outsiders say of the Madurese in general). They impressionistically characterize their speech as sonically more clipped, spoken at a high volume. Eastern Madurese are considered more soft spoken and refined, given more to circumlocution. But these are gross generalizations, given here in the terms used by the Madurese to describe themselves. As this is not intended to be a dialect study, no effort is made to evaluate or interpret these characterizations. In fact, while some lexical differences such as those cited above are identified in what follows, no concerted effort is made to exhaustively document the differences between the Bangkalan and Sumenep varieties. Such description is properly the province of a concerted dialect study. Some dialectal differences in pronunciation have been documented recently in Sutoko et al. 1998.

As might be expected given the close proximity of Bangkalan to Surabaya (30 minutes by ferry and as of 2009 connected by the Suramadu Bridge), there is somewhat more influence of Javanese on Western Madurese than in points further east on the island (but likely no more than the Madurese of the 3 million plus speakers living in East Java).

### 3. Previous studies

The amount of work on the Madurese language does not approach the scope of linguistic research done on other languages of the region such as Balinese, Javanese, Malay/Indonesian, and Sundanese. As alluded to above, the reason for this may be prevailing attitudes toward the Madurese people, attitudes which extend to views of the language as well. For example, one observer remarked, in comparison to Javanese and Sundanese “Madurese sounds much less modulated than one of those other two. Jerkier and rougher, more peasant-like and more a language for rough sailors...or pirates.” (Brandt Buys 1926:369) Such characterizations are not difficult to find in other writing. There has, nonetheless, been a substantial amount of work on the language, in Dutch, English, and Indonesian.

H.N. Kiliaan, a Dutch civil servant, did the most extensive work during the Dutch colonial rule of Indonesia. Most important are (i) his grammatical description of Madurese, *Madoereesche spraakkunst*, commissioned by the Dutch government and published in 1897 and reprinted in 1911, and (ii) a Madurese-Dutch dictionary, *Madoereesche-Nederlandsch woordenboek*, published in two parts in 1904/05. By one assessment, “After Kiliaan no linguists have done any extensive work on Madurese” (Uhlenbeck 1964:176). However, there was work that followed Kiliaan’s as well as some that predated it. The first published work was that of Vreede (1874/76), who did 4 volumes including a grammar sketch, wordlists, some literary work, among others. Elzevier Stokmans and Marinissen (1880) produced a short practical guide for learning Madurese, as did Penninga and Hendriks (1930).

The Western literature following the Dutch era has, like the work in Dutch, primarily focused on the phonology and morphology of Madurese. Foremost among these is the body of work of Stevens. The most extensive among these, *Madurese Phonology and Morphology* (1968), provide a detailed phonological description and an extensive catalogue of morphemes together with their phonological effects. In other articles Stevens describes the system of speech levels (1965), Madurese reflexes of proto-Malayo-Polynesian (1966), juncture rules in phonology (1980), and the unique system of reduplication (1985, 1994). Cohn examines Madurese vowel harmony in a series of papers (1993a,b; Cohn and Lockwood 1994). Two aspects of Madurese phonology have received quite a bit of attention in other generative literature: reduplication (Wilbur 1973; Marantz 1982; Steriade 1988; Silverman 2002) and vowel harmony (Trigo 1987, 1991 and Anderson 1991). Hamm 2001 and Maddieson (In press) include extensive discussion of gemination in Madurese. The most extensive work on the syntax and morphosyntax of Madurese is that of Davies

(1999, 2000, 2003, 2005a,b,c, 2008).<sup>5</sup>

Third, there is a growing Indonesian descriptive literature. The first description published in Indonesian that is available is Asmoro's (1950) 26-page *Paramasastra Madura*. Since 1978, a number of descriptions especially of the phonology and morphology have appeared, though somewhat less work on sentence structure. The largest body of work has been carried out under the auspices of Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan ('the Center for Training and Development of Language, Department of Education and Culture'), headquartered in Jakarta: Zainudin, Soegianto, Kusuma, and Barijati (1978) surveys the phonemic and morphological inventories; Moehnilabib, Wahab, Prijambada, Huda, and Ghazali (1979) survey the morphological inventory and basic sentence structure; Pratista, Harjono, Matorahardjo, Soekemi, and Hatib (1984) provide a detailed inventory of reduplication forms and their meanings; Soegianto, Soetoko, Soekarto, Soetarto, and Kustiati (1986) again survey the morphological inventory and basic clause structure. Other works have been sponsored by other Indonesian government agencies: Oka et al (1986) provide a somewhat abridged version of the information in Zainudin et al; Budi, Satyaputra; Koentjahjo, Adriana, and Marhaen (1986/1987) survey the phonemic inventory; Wibisono, Sofyan, Suyanto, Soegianto, and Suparmin (2001) provide an extensive description of negative formation. Safioedin (1977) compiled the first published Madurese-Indonesian dictionary, *Kamus Bahasa Madura-Indonesia*, again under the auspices of the Pusat Pembinaan dan Pengembangan Bahasa, Departemen Pendidikan dan Kebudayaan. A more recent small dictionary, *Kamus Bahasa Madura*, compiled by Kutwa, Dradjid, Basar, Sadik, Muakmam, Sastro, Hartono, and Efendy (2007), represents the first published dictionary that includes definitions in Madurese in addition to Indonesian; this was published through the Dinas Pendidikan dan Kebudayaan ('Department of Education and Culture') of the Kabupaten of Pamekasan. Additionally, the agency sponsored two dialect studies: Soegianto, Soetoko, Soekarto and Soetarto's (1986) *Pemetaan Bahasa Madura di Pulau Madura* and Sutoko, Soegianto, Surani, Sariono, and Suyanto's (1998) *Geografi Dialek Bahasa Madura*. Most recently, a new dictionary has been compiled by Adrian Pawitra (Pawitra 2009). At nearly 750 pages, *Kamus Lengkap Bahasa Madura Indonesia* is a monumental work which includes pronunciations in phonetic transcription for each entry as well as etymological information, definitions in Indonesian, Madurese synonyms, and some common expressions. I regret that the dictionary was published after the manuscript of this volume was completed and thus could not have been more extensively consulted.

---

<sup>5</sup> There is one further Western monograph on Madurese grammatical structure, Ogoblin 1986. Regrettably, this work, written in Russian, has not been accessible to me, and thus the current work does not benefit from Ogoblin's discoveries and insights.

#### 4. Methodology and framework

Compiling a description of a language that has already been documented naturally entails slightly different methodology than that used when describing a language for which little or no documentation exists. And that is reflected to some degree in the procedures followed in this work. The data comprising the ‘corpus’ reported on here were gathered over a 10 year plus period from a variety of sources:

1. direct elicitation (both in Indonesia and Iowa),
2. recorded narratives and conversations (provided by my primary language consultants),
3. existing accounts of the language (as described in the section on previous literature),
4. dictionaries (published and unpublished) and
5. narratives and stories published in Madurese (e.g. *Campaka* (Imron 1979), *Santre Gudhigan* (Asmara 1984)).

The first three categories represent the sources from which the vast majority of the data were gathered. Existing accounts of the language (as well as insights from work on Javanese) provided initial leads and later specific structures to follow up on. Such data were checked and rechecked with language consultants in Bangkalan and Iowa, however, to verify their currency in the language described here. Where there was divergence it has been noted. It is clear from citations contained in the chapters that follow, which secondary sources proved the most valuable. Recorded narratives and conversations not only provided important examples of variability in structures in ‘natural speech’ as opposed to prescriptive standards but more importantly in introducing new structures and puzzles to investigate. Such naturally occurring data were also checked with language consultants in order to filter out sentences that might be considered ‘production errors’ that occur with some frequency whenever we speak.

All of them were extremely gracious and patient in the face of sometimes persistent and annoying questioning, and all deserve my special thanks, as do the extensive staff of the Hotel Ningrat in Bangkalan, who contributed data through casual conversation and sometimes through direct questioning. Deserving of special mention, however, is the late Pak Masharudin, more affectionately referred to as Pak Ending by many who met him. For 35 years or more, Masharudin compiled a dictionary of Madurese, painstakingly written by hand. He was generous enough to allow me to copy his labor in 2000, the first year I met him. There are currently no plans to publish this work; therefore, here I cite the work as Masharudin 2000. While by his own admission an incomplete work, Masharudin’s dictionary includes some forms not found in the excellent pub-

lished dictionaries of Safioedin (1977), Kutwa et al. (2007), and Pawitra (2009), thus providing important data in the present work and a valuable resource to those fortunate enough to have access to it.

#### 4.1. Framework

This volume is intended to be a largely descriptive work. While I have attempted to provide a comprehensive description of the phonology and morphology of Madurese, this is the first work I am aware of that focuses quite so heavily on syntax. No particular theoretical framework has been adopted here. However, the basic perspective makes assumptions common to generative linguistics. This is clearly reflected in the extensive coverage of certain topics as well as aspects of the organization of this work. Nevertheless, the goal being a general description, many opportunities for theoretical discussion or in-depth consideration of certain domains of data have been passed over. That said, at times it strikes me as appropriate to discuss the data in more theoretical terms and to bring standard generative argumentation to bear on the data in order to explicate a particular analysis. These aspects, too, will be obvious to the reader.



# **Chapter 2**

## **Phonology**

This chapter presents the basic sound system of Madurese. Section 1 outlines the segmental inventory and the phonetic realization of consonants and vowels. Section 2 describes some of the regular sound correspondences that are found between Madurese and its two closest relatives, Indonesian and Javanese. Section 3 describes some phonotactic constraints on syllable structure. Section 4 again takes up the vowel inventory, specifically providing evidence for the phonemic inventory, describing the distribution of vowels and the conditioning environment for the system of alternating vowels. Section 5 presents the major phonological processes and section 6 the morpho-phonological processes. Section 7 includes a brief note on stress. Section 8 discusses orthography and presents the spelling system that will be used in the chapters that follow.

### **1. Basic inventory and description of sounds**

#### 1.1. Consonants

The basic inventory of Madurese consonants is given in the chart on the following page.

Notable in the consonantal inventory is the set of stops in Madurese. Leaving aside the glottal stop, there are five places of articulation for oral stops. Of the languages most closely related to Madurese, this is true only of Javanese, which has stops at the same places of articulation. It is not true of Indonesian, Balinese, or Sundanese. Additionally, there is a three-way contrast in phonation at each of the five places of articulation. The most-closely related languages, including Javanese, have only a two-way contrast in the series of stops, either voiceless unaspirated and voiced or voiceless unaspirated and breathy, a type not present in Madurese.

	labial	dental <sup>1</sup> / alveolar	retroflex <sup>2</sup>	palatal	velar	glottal
stop:						
voiceless	p	t	ʈ	c	k	ʔ
vl aspirated	p <sup>h</sup>	t <sup>h</sup>	ʈ <sup>h</sup>	c <sup>h</sup>	k <sup>h</sup>	
voiced	b	d	ɖ	ɟ	g	
nasal	m	n		ɲ	ŋ	
fricative	(f)	s				(h)
liquid		l,r				
glide	(w)			(j)	(w)	

Notable in the consonantal inventory is the set of stops in Madurese. Leaving aside the glottal stop, there are five places of articulation for oral stops. Of the languages most closely related to Madurese, this is true only of Javanese, which has stops at the same places of articulation. It is not true of Indonesian, Balinese, or Sundanese. Additionally, there is a three-way contrast in phonation at each of the five places of articulation. The most-closely related languages, including Javanese, have only a two-way contrast in the series of stops, either voiceless unaspirated and voiced or voiceless unaspirated and breathy, a type not present in Madurese.

Following is a description of the consonant phonemes.

/p/ voiceless bilabial stop  
 [p] in syllable-initial position  
 [paman] ‘uncle’  
 [ɖɤpaʔ] ‘arrive’

<sup>1</sup> Although the oral stops have dental articulation, the dental diacritic will not be included here as no confusion should arise from its omission.

<sup>2</sup> Rather than using the IPA symbols for the retroflex consonants, the . diacritic signals retroflex articulation, which is consistent with the orthographic tradition as well as previous studies (e.g. Stevens 1968, Moehnilabib et al. 1979, Budi et al. 1986/87, and Cohn & Lockwood 1994).

[p̚] in syllable-final position

[taŋkəp̚] ‘catch’

[aɖɖəp̚] ‘front’

/p<sup>h</sup>/ voiceless aspirated bilabial stop

[p<sup>h</sup>] in syllable-initial position, does not occur in syllable-final position

[p<sup>h</sup>urus] ‘dog’

[tamp<sup>h</sup>ɣŋ] ‘mine’

/b/ voiced bilabial stop

[b] in syllable-initial position, does not occur in syllable-final position

[bɣtə] ‘rock’

[t<sup>h</sup>ibiʔ] ‘alone’

/t/ voiceless dental stop

[t] in syllable-initial position

[tanəŋ] ‘hand’

[atɛ] ‘heart/liver’

[t̚] in syllable-final position

[sərat̚] ‘letter’

[pəʔlət̚] ‘pencil’

/t<sup>h</sup>/ voiceless aspirated dental stop

[t<sup>h</sup>] in syllable-initial position, does not occur in syllable-final position

[t<sup>h</sup>isa] ‘town’

[t<sup>h</sup>ut<sup>h</sup>iŋ] ‘accuse’

/d/ voiced dental stop

[d] in syllable-initial position, does not occur in syllable-final position

[dibɣsa] ‘mature’

[duŋgu] ‘dull-witted’

/t̚/ voiceless alveolar retroflex stop

[t̚] in syllable-initial position, does not occur in syllable-final position  
(except geminates)

[t̚wat̚] ‘scream’

[kəʔak̚] ‘box’

- /tʰ/ voiceless aspirated alveolar retroflex stop  
 [tʰ] in syllable-initial position, does not occur in syllable-final position  
 (except geminates)  
 [tʰimɛn] ‘too much’  
 [pʰuʈʰu] ‘capital’
- /d/ voiced alveolar retroflex stop  
 [d] in syllable-initial position, does not occur in syllable-final position  
 (except geminates)  
 [dʒʔ] ‘to’  
 [bɤdʒ] ‘exist’
- /ç/ voiceless palatal stop (lightly affricated)  
 [ç] in syllable-initial position, does not occur in syllable-final position  
 (except geminates)  
 [cɔkpʰ] ‘enough’  
 [bɤca] ‘read’
- /cʰ/ voiceless aspirated palatal stop (lightly affricated)  
 [cʰ] in syllable-initial position, does not occur in syllable-final position  
 (except geminates)  
 [cʰukɔʔ] ‘fish’  
 [kʰɤcʰɤ] ‘elephant’
- /ʃ/ voiced palatal stop (lightly affricated)  
 [ʃ] in syllable-initial position, does not occur in syllable-final position  
 (except geminates)  
 [ʃɤgɔ] ‘wake up’  
 [parcaʃɤ] ‘believe’
- /k/ voiceless velar stop  
 [k] in syllable-initial position  
 [kɛrɛm] ‘send’  
 [pakɔ] ‘nail’
- [kʰ] in syllable-final position  
 [kɔʔakʰ] ‘box’  
 [candʒkʰ] ‘hold’

- /k<sup>h</sup>/ voiceless aspirated velar stop  
 [k<sup>h</sup>] in syllable-initial position, does not occur in syllable-final position  
 [k<sup>h</sup>ʁbɔŋ] ‘make’  
 [p<sup>h</sup>ik<sup>h</sup>ʁl] ‘rob’
- /g/ voiced velar stop  
 [g] in syllable-initial position, does not occur in syllable-final position  
 (except geminates)  
 [gʁʒi] ‘salary’  
 [ʒʁgɔ] ‘wake up’
- /ʔ/ glottal stop<sup>3</sup>  
 [ʔ] occurs only in syllable-final position  
 [anaʔ] ‘child’  
 [raʔjatʔ] ‘people’
- /m/ voiced bilabial nasal  
 [m] in syllable-initial and syllable-final position  
 [mɛra] ‘red’  
 [maɛn] ‘play’  
 [kɛrɛm] ‘send’  
 [ɔkɔm] ‘punish’
- /n/ voiced alveolar nasal  
 [n] in syllable-initial and syllable-final position  
 [nəŋ] ‘at’  
 [nasɛʔ] ‘rice’  
 [kakan] ‘eat’  
 [k<sup>h</sup>ʁn] ‘until’

---

<sup>3</sup> Word-final [ʔ] and [kʔ] are virtually auditorily indistinguishable even to native speakers (leading to frequent inconsistency in spelling). Their phonemic identity is revealed when a suffix is added to the stem, however, resulting in [ʔ] for phonemic /ʔ/ and [k<sup>h</sup>] for /k/, as in:

- (i) sabʁʔ ‘put’      sabʁʔa ‘put’ (irrealis)  
 candʁk ‘hold’      candʁk<sup>h</sup>ʁ ‘hold’ (irrealis)

The surface form of the irrealis morpheme /a/ results from the vowel raising process described in section 4.1. Aspiration is discussed in section 6.

- /ɲ/   voiced palatal nasal  
 [ɲ] in syllable-initial position, does not occur in syllable-final position  
 (except geminates)  
 [ɲaman] ‘delicious’  
 [tɲa] ‘ask’
- /ŋ/   voiced velar nasal  
 [ŋ] in syllable-initial and syllable-final position  
 [ŋara] ‘possible’  
 [taŋɛs] ‘cry’  
 [sənnəŋ] ‘happy’  
 [sɛŋkɔʔ] ‘I’
- /s/   voiceless alveolar fricative  
 [s] in syllable-initial and syllable-final position  
 [səkɛʔ] ‘sick’  
 [səmpɛʔ] ‘until’  
 [alas] ‘forest’  
 [maskɛ] ‘although’
- /r/   voiced alveolar trill  
 [r] in syllable-initial and syllable-final position  
 [raʃɾ] ‘large’  
 [sɔratɾ] ‘letter’  
 [kəp<sup>h</sup>ɾ] ‘news’  
 [kərtas] ‘paper’
- /l/   voiced alveolar lateral approximant  
 [l] in syllable-initial and syllable-final position  
 [lɛʔɛr] ‘neck’  
 [alas] ‘forest’  
 [bɔntal] ‘pillow’  
 [ʃunɛl] ‘skilled’

The glides [j] and [w] have a special status. [j] is phonemic only in word-final position of native Madurese words, such as [sɔŋaj] ‘river’, [kaʔaŋk<sup>h</sup>uj] ‘for’, and others. In other native words, [j] is not phonemic. An epenthesis process described in section 5.1 is largely responsible for its occurrence in intervocalic environments. Word-initial [j] occurs in some loanwords.

- /j/ palatal semi-vowel  
 [j] in syllable-initial and syllable-final position  
 [nɛjatʰ] ‘intend’  
 [rɔpɛja] ‘rupiah’  
 [tamɔj] ‘guest’

Word-internal [w] is also primarily a result of epenthesis. Pawitra (2009) lists 36 w-initial entries in his dictionary (Safioden gives just 19), the majority of which are borrowings.<sup>4</sup> [w] does not occur in syllable-final position.

- /w/ labio-velar semi-vowel (19 w-initial words in dictionary)  
 [w] in syllable-initial position (largely non-phonemic)  
 [wəjipʰ] ‘obliged’ (Arabic)  
 [wawancara] ‘interview’ (Indonesia *wawancara*)  
 [wɔrtəl] ‘carrot’ (Dutch *wortel*)

The phones [h] and [f] occur in borrowings. [h] occurs primarily in initial position in words that are largely of Arabic origin. There are 67 entries for *h* in Pawitra 2009 and 38 in Safioedin 1977.

- /h/ glottal fricative  
 [h] in syllable-initial position  
 [haɖirin] ‘audience’ (Arabic)  
 [hɔrmatʰ] ‘honor’ (Indonesian *hormat*)  
 [hɔtəl] ‘hotel’

[f] occurs in loanwords from a variety of languages. Pawitra list 29 for *f* in his dictionary and Safioedin 17 in his.

- /f/ labiodental fricative (17 f-initial words in dictionary)  
 [foto] ‘photograph’  
 [fana] ‘transitory’ (Arabic)  
 [fardu] ‘duty’ (Arabic)

---

<sup>4</sup>While phonemic /w/ in loanwords is not limited to word-initial position (e.g. [wawancara] ‘interview’ from Indonesian), the relatively small number of dictionary entries is indicative of the circumscribed nature of its use. This is true of other ‘non-native’ phonemes described below.

## 1.2. Vowels

Identifying the inventory of vowels is a more complicated story. By Stevens' estimate, roughly 95% of the Madurese lexical items in his corpus employ the eight-vowel inventory given in the following chart:

	front	central	back
high	<b>i</b>	<b>ɨ</b>	<b>u</b>
		ɻ	
mid	<b>ɛ</b>	<b>ə</b>	<b>ɔ</b>
low		<b>a</b>	

These can be considered the 'primary' surface vowels of Madurese. Due to a system of paired alternating vowels, the vowels that are contrastive can be identified as those that are in bold. This system is detailed in section 4. Also, the precise phonetic realization of some of the vowels is subject to some variation among speakers and perhaps across dialects. Some of this variation is discussed in subsequent sections as well. The primary vowels are as follows.

- /i/ high front unrounded vowel  
 [i] in open and closed syllables  
 [bijasa] 'usual'  
 [paði] 'rice plant'  
 [giʔ] 'yet'  
 [ɪ] in some closed syllables (subject to inter- and intra-speaker variation)  
 [bɪntaŋ] 'star'  
 [k<sup>h</sup>ɪppɔŋ] 'confused'
- /ɛ/ mid front lax unrounded vowel  
 [ɛ] in closed and open syllables  
 [kɛnɛʔ] 'small'  
 [ɛŋaʔ] 'remember'  
 [sakteɛ] 'magic power'  
 [e] in some open syllables (subject to inter- and intra-speaker variation)  
 [se] 'relative particle'  
 [reja] 'that'

- /u/ high back rounded vowel  
 [u] in open and closed syllables  
 [k<sup>h</sup>unɔŋ] ‘mountain’  
 [c<sup>h</sup>ukɔʔ] ‘fish’  
 [p<sup>h</sup>urus] ‘dog’  
 [ɔ] in some closed syllables (subject to inter- and intra-speaker variation)  
 [bɔnt<sup>h</sup>ɔʔ] ‘tail’  
 [k<sup>h</sup>ɔttɛ] ‘uncle’
- /ɔ/ mid back lax rounded vowel  
 [ɔ] in closed and open syllables  
 [cɔkɔp] ‘enough’  
 [c<sup>h</sup>ukɔʔ] ‘fish’  
 [bɔktɔ] ‘time’  
 [o] in some open syllables (subject to inter- and intra-speaker variation)  
 [pɔlɛ] ‘again’  
 [toko] ‘store’
- /a/ low central vowel  
 [æŋ] ‘water’  
 [kapala] ‘head/chief’  
 [bɔca] ‘read’
- /ə/ mid central vowel  
 [ənnəm] ‘six’  
 [nəssəl] ‘regret’

As indicated here, there is variation in the realization of some vowels depending on whether they occur in open or closed syllables. This is taken up in section 4.2.

Although they have a widespread distribution, the high central vowels [i] and [ɿ] are arguably not phonemic. Rather, they are allophones of [ə] and [a], respectively. As described in section 2.4, the distinction between [ə] and [i] is not universally recognized in the literature; however, acoustic analysis does support the difference (Cohn & Lockwood 1994, Bortscheller 2007). Additionally, the vowel transcribed with the symbol [ɿ] is a mid-close central unrounded vowel most nearly equivalent to the vowel represented by [ə] in standard IPA transcription.<sup>5</sup> The symbol [ɿ] is used here to ensure consistency with the con-

---

<sup>5</sup>My thanks to Jill Beckman for help interpreting the acoustic measurements. Pawitra (2009) uses [ɐ], a low central vowel, for pronunciations in his dictionary.

vention adopted in the literature on Madurese (e.g. Stevens 1985, Cohn 1993a,b, Cohn & Lockwood 1994).<sup>6</sup> Examples of their use include:

- [i]      [billi] ‘buy’  
          [ginnaʔ] ‘complete’  
[ɤ]      [bɤca] ‘read’  
          [sanɤaʔɤ] ‘intentional’

The inventory of vowels and their distribution is discussed in extensive detail in section 4.

## 2. Sound correspondences between Madurese and Indonesian (and Javanese)

There are some regular phonemic sound correspondences between Madurese and the languages most closely related to it, Indonesian and Javanese. As will be discussed in Chapter 15, in the higher speech levels the influence of Javanese is unmistakable. At the lower level of speech, *kasar* ‘rough’, and with *biyasa* ‘usual’ vocabulary used in all speech situations, there is greater similarity between Madurese and Indonesian, although similarity between Madurese and Javanese is widespread here as well.

### 2.1. Consonants

The similarities and differences among consonants in Madurese, Indonesian, and Javanese cognates are most striking at the phonemic level, so in what follows, underlying representations are given. The Madurese phonetic form also appears. The processes by which these forms are derived are presented in later sections.

Aside from loanwords, Madurese has no phonemic /w/; where Indonesian and Javanese cognates have /w/, /b/ occurs.<sup>7</sup>

---

<sup>6</sup>[ɤ] is used to transcribe the Sundanese equivalent of this vowel as well, which is sometimes described as a mid-back unrounded vowel, but which like the Madurese vowel is a mid-close central unrounded vowel by acoustic measurement.

<sup>7</sup>Even in many borrowings, /w/ is realized as /b/. E.g., *birit* [birit] ‘pray’, Indonesian *wirid* from Arabic.

(1)	Madurese		Indonesian		Javanese	
	/batu/	[bʁtɔ]	/watu/		/watu/	‘stone’
	/baktɔ/	[bʁktɔ]	/waktu/		/waktu/	‘time’
	/c <sup>h</sup> aba/	[c <sup>h</sup> ʁbʁ]	/dʒawa/		/jawa/	‘Java’
	/labang/	[labʁŋ]	(/pintu/)		/lawang/	‘door’
	/t <sup>h</sup> ɛbɛʔ/	[t <sup>h</sup> iʁiʔ]	(/diri/)		/dɛwe/	‘alone/self’
	/baca/	[bʁca]	/baca/		/waca/	‘read’

Stevens (1966) cites Kiliaan (1897) as identifying the Madurese-Javanese correspondence, as well as a correspondence between Madurese /p<sup>h</sup>/ and Javanese /b/. This is true of Indonesian as well.<sup>8</sup>

(2)	Madurese		Indonesian		Javanese	
	/p <sup>h</sup> aʔɛk/	[p <sup>h</sup> ʁʔɛk]	/batik/		/batik/	‘batik’
	/p <sup>h</sup> ak <sup>h</sup> ɔs/	[p <sup>h</sup> ʁk <sup>h</sup> us]	/bagus/		/bagus/	‘good’
	/p <sup>h</sup> ɔt <sup>h</sup> ɔ/	[p <sup>h</sup> uʔ <sup>h</sup> u]	/bodoh/		/boʔo/	‘stupid’
	/ɛp <sup>h</sup> ɔʔ/	[ɛp <sup>h</sup> uʔ]	/ibu/		/ibu/	‘mother’
	/kəp <sup>h</sup> ɔj/	[kəp <sup>h</sup> uj]	/kərbaw/		/kərbaw/	‘water buffalo’

The phoneme /h/ is rare in Madurese and occurs principally in loanwords of Arabic origin, and some recent loans from Indonesian. For the most part, aside from Arabic loans, there is no /h/ in Madurese cognates of Indonesian words containing /h/. (Under the influence of Indonesian, some cognates are sometimes pronounced with [h], e.g., [asɛl] ~ [hasɛl] ‘result’. However, one might interpret this as a speaker simply using the Indonesian word, as speakers frequently will pepper their speech with Indonesian or Javanese.)

(3)	Madurese		Indonesian		Javanese	
	/taɔ/	[taɔ]	/tahu/		/tau/	‘know’
	/c <sup>h</sup> aɛ/	[c <sup>h</sup> ʁi]	/dʒahe/		/jae/	‘ginger’
	/atɛ/	[atɛ]	/hati/		/ati/	‘heart/liver’
	/asɛl/	[asɛl]	/hasil/		/asil/	‘result’
	/mɛra/	[mɛra]	/merah/		(/abaŋ/)	‘red’
	/lɛɔllɛ/	[lɛɔllɛ]	/oleholeh/		/oleholeh/	‘souvenir’ <sup>9</sup>

In many lexical items, glottal stop occurs where the Indonesian cognate

<sup>8</sup>Stevens (1966) notes the possibility of widespread borrowing from Javanese may be a plausible explanation for the /p<sup>h</sup>/ ~ [b] correspondence as Javanese *b* is breathy and voiceless and so might be perceived as aspirated by Madurese speakers.

<sup>9</sup>These are reduplicated forms.

has a syllable-final voiceless stop.<sup>10</sup> However, this is a sporadic phenomenon; in many cases, the cognates in the two languages have the same voiceless stop. Examples of both phenomena are given below.

(4)	Madurese		Indonesian	Javanese	
a.	/ataʔ/	[ataʔ]	/atap/	(/wuwuŋ/)	‘roof (made of palm fronds)’
	/gənaʔ/	[gənaʔ]	/gənap/	/gənəp/	‘complete’
	/dʁddʁʔ/	[dʁddʁʔ]	/dadap/	/dadap/	‘Erythrima tree’
	/taŋkəp/	[taŋkəpʰ]	/taŋkap/	/taŋkəp/	‘catch’
	/atʰəp/	[atʰəpʰ]	/hadap/	/adəp/	‘front’
b.	/sakeʔ/	[sakeʔ]	/sakit /	(/lɔrɔ/)	‘ill’
	/əmpaʔ/	[əmpaʔ]	/əmpat /	/papat/	‘four’
	/c <sup>h</sup> aeʔ/	[c <sup>h</sup> ɛiʔ]	/dʒahit /	/jait/	‘sew’
	/cəpət/	[cəppətʰ]	/tʃəpat /	/cəpət/	‘fast’
	/pəʔlɔt/	[pəʔlɔtʰ]	/potlot /	/pətɫɔt/	‘pencil’
c.	/anaʔ/	[anaʔ]	/anak/	/anak/	‘child’
	/rəkɔʔ/	[rəkɔʔ]	/rokok/	/rəkək/	‘tobacco’
	/bapaʔ/	[bapaʔ]	/bapak/	/bapak/	‘father’
	/p <sup>h</sup> atək/	[p <sup>h</sup> ətəkʰ]	/batik/	/batik/	‘batik’
	/kɔtak/	[kɔtakʰ]	/kotak/	/kɔtak/	‘box’

All of the underlying forms of the Indonesian and Javanese items have /k/ in final position, which by rule is realized as [ʔ]. In the Madurese, /ʔ/ and /k/ represent a phonemic contrast. So, as with [p] and [t], some word-final phonemic /k/ in cognates correspond to underlying /k/ in Madurese while others correspond to underlying /ʔ/.

The palatal glide /j/, while present in many words, is largely non-phonemic, its presence due to an epenthesis rule discussed in section 2.5.1. Where the palatal glide occurs in Indonesian cognates, a voiced palatal stop /ɟ/ is generally found in Madurese.

<sup>10</sup> Stevens (1966) notes that Proto-Malayo-Polynesian word-final \*p, \*t, and \*k developed into Madurese /ʔ/. It is possible that words that do have final /p/, /t/, and /k/ entered the language through borrowing from Malay and/or Javanese or at least were influenced by them.

(5)	Madurese	Indonesian	Javanese		
	/parcaja/	[parcaɣ]	/pərtʃaja/	/pərcaja/	‘believe’
	/jakən/ <sup>11</sup>	[ɣkən]	/jakin/	/jakin/	‘sure’
	/jatəm/	[ɣtəm]	/jatim/	/jatim/	‘orphan’
	/c <sup>h</sup> aɣa/	[c <sup>h</sup> ɣɣ]	/dʒaja/	/ɣaja/	‘victory’
	/sɔrbaɣa/	[sɔrbɣɣ]	/surabaja/	/surabaja/	‘Surabaya’

In some cognates, Madurese has a voiceless aspirated /t<sup>h</sup>/ or voiced /d/ retroflex where one finds /d/ in Indonesian. As spelled out by Stevens (1966), this is attributable to the fact that Proto-Malayo-Polynesian \**d* was realized as /d/ in Indonesian but either /t<sup>h</sup>/ or /d/ in Madurese.

(6)	Madurese	Indonesian	Javanese		
	/dapaʔ/	[dɣpaʔ]	/dapat/	(/əleh/)	‘arrive’
	/datəŋ/	[dɣtəŋ]	/datan/	(/teka/)	‘come’
	/sapɛda/	[sapɛdɣ]	/səpeda/	/səpeda/	‘bicycle’
	/t <sup>h</sup> ampəl/	[t <sup>h</sup> ɣmpəl]	/dəmpul/	/dəmpul/	‘putty’
	/t <sup>h</sup> əkən/	[t <sup>h</sup> ukən]	/dukun/	/dukun/	‘traditional doctor’

## 2.2. Vowels

Many vowel correspondences between Madurese and Indonesian are quite regular because the quality of the Madurese vowel is conditioned by a rule of vowel raising in which high vowels occur after aspirated and voiced stops and mid and low vowels elsewhere. (The process and the limited exceptions are discussed more fully in section 4). Because the surface form is more revealing in the case of vowels, phonetic representations of Madurese and Indonesian are compared here. These correspondences include instances in which Madurese has a mid vowel because of the preceding conditioning environment. So, in some cases the high vowel of the Indonesian corresponds to a mid vowel in the Madurese cognate.

(7)	Madurese	Indonesian	
	[kəceŋ]	[kutʃiŋ]	‘cat’
	[sapɛ]	[sapi]	‘cow’
	[kɛrɛm]	[kirim]	‘send’
	[tələs]	[tulis]	‘write’

<sup>11</sup>Many speakers currently pronounce this as [jakən], likely influenced by Indonesian or Javanese.

At other times, precisely the opposite obtains—cognates in which Madurese has a high vowel where Indonesian has a mid vowel.

(8)	Madurese	Indonesian	
	[t <sup>h</sup> isa]	[desa]	‘village’
	[k <sup>h</sup> itək]	[getek]	‘raft’
	[ʃiʃir]	[dʒedʒer]	‘row’
	[p <sup>h</sup> uʔ <sup>h</sup> u]	[bodoh]	‘stupid’

There are however, ample examples of the same vowel surfacing in both languages.

(9)	Madurese	Indonesian	
	[bukkaʔ]	[buka]	‘open’
	[diriʔ]	[diri]	‘stand’
	[kəntan]	[kontan]	‘cash’
	[tənəŋ]	[tenoŋ]	‘a type of basket’

For the most part, non-high vowels (ɛ, ə, a, ɔ) do not occur following voiceless aspirated and voiced stops and high vowels (i, ɤ, i, u) occur only in the environment after voiceless aspirated and voiced stops. This regularity accounts for the vowel correspondences illustrated in (6-7). For example, consider the Madurese and Indonesian words for ‘raft’ in (8), [k<sup>h</sup>itək] and [getek]. The vowels in the second syllable in both are mid front vowels. However, the mid front vowel of the first syllable of the Indonesian word is a high front vowel in Madurese, owing to the fact that it follows a voiceless aspirated stop [k<sup>h</sup>]. This pattern recurs in the other data in (7-8). The system of alternating vowels in Madurese is discussed in detail in section 4.

There are two other notable vowel correspondences that occur sporadically. In one case, Indonesian /ə/ occurs as /a/ in Madurese.<sup>12</sup> This occurs only in open syllables and is due in part to the fact that Madurese /ə/ occurs only in closed syllables, as discussed in section 3.

(10)	Madurese	Indonesian	
	[ka]	[kə]	‘to’
	[careta]	[tʃərita]	‘story’
	[pareksa]	[pəriksa]	‘examine’

<sup>12</sup>These correspondences again largely reflect the development of Proto-Malayo-Polynesian vowels in the two languages. See Stevens 1966.

The opposite correspondence obtains in some cognates. Indonesian /a/ is /ə/ in Madurese. Due to the restriction on schwa, this occurs only in closed syllables.

(11) Madurese	Indonesian	
[sənnəŋ]	[sənəŋ]	‘happy’
[maləm]	[malam]	‘night’
[kərrəp]	[kərap]	‘often’
[taŋkəp]	[tangkap]	‘catch’

The words ‘happy’ and ‘often’ illustrate an alternative to the /ə/ ~ /a/ correspondence. Rather than a phonemic /a/ where Indonesian has schwa in an open syllable, the onset of the following syllable is geminated, creating the proper environment for Madurese /ə/.

### 3. Syllable structure and phonotactic constraints

The majority of Madurese roots are disyllabic, made up of CV and CVC syllables.<sup>13</sup> Syllables can, however, take any of the following forms.

#### (12) Licit syllables

V	a.bɤ.lɤ ‘say’, ɛ.pɔ.kɔl ‘hit’, a.ɛŋ ‘water’
CV	bɤ.ca ‘read’, bɪl.li ‘buy’, sɔ.ratʰ ‘letter’
VC	ə.m.paʔ ‘four’, ɛ.n.tar ‘go’, ka.ɔs ‘t-shirt’
CVC	kɛ.rɛm ‘send’, pʰu.rus ‘dog’, bɤkʰ.tɔ ‘time’, tʰu.ŋɛŋ ‘story’
CCV	pra.ɔ ‘boat’, gla.nɔn ‘excuse me’
CCVC	ɛs.trɛ ‘wife’, trɛs.na ‘love’, klam.pʰi ‘clothes’

The clusters in CCV(C) syllables come largely but not exclusively from two sources: forms that have undergone vowel deletion and borrowings. As is shown below, [praɔ], [glaɔn], [trɛsna] and [klam<sup>h</sup>i] are derived from /paraɔ/, /galaɔn/, /taresna/ and /kalamp<sup>h</sup>i/, respectively. [ɛs.trɛ] is borrowed from Indonesian.

Preference for disyllabic roots manifests itself in the large number of words with initial əC<sub>1</sub>C<sub>1</sub> which correspond to Indonesian monosyllabic roots with initial C<sub>1</sub>. Some of these are borrowings from Dutch and English. A few of these are given in (13).

<sup>13</sup>There are many words of three syllables and more, but the majority of these are forms with one or more affixes.

(13)	Madurese	Indonesian	Javanese		
	/əccɛt/	[əccɛtʰ]	/tʃat/	/cɛt/	‘paint’
	/əcc <sup>h</sup> am/	[əcc <sup>h</sup> am]	/dʒam/	/ʒam/	‘hour’
	/əllak/	[əllakʰ]	/lak/	/lak/	‘sealing wax’
	/ənnɔr/	[ənnɔr]	/nur/	/nur/	‘light’
	/əssɛl/	[əssɛl]	/sel/	/sɛl/	‘cell’
	/əttɔŋ/	[əttɔŋ]	/toŋ/	/tuŋ-tuŋ/	‘tun’

It is frequently the case, however, that in speech these forms are shortened and only the final syllable is pronounced.

As illustrated in section 1, all consonants other than [ʔ] can occur in syllable-initial position. Leaving aside cases of gemination, the set of consonants that can occur in syllable-final position is much more restricted, and includes p, t, k, m, n, ŋ, s, r, l, j, and ʔ. An example of each in word-final position is given in (14).

(14)	[cəkɔpʰ]	‘enough’
	[sɔratʰ]	‘letter’
	[kɔʔakʰ]	‘box’
	[kɛnɛʔ]	‘small’
	[kɛrɛm]	‘send’
	[kɔran]	‘newspaper’
	[nɔŋ]	‘at’
	[panas]	‘hot’
	[ɛntar]	‘go’
	[tɔkɔl]	‘hammer’
	[k <sup>h</sup> ɪbɪj]	‘make’

It should be noted that underlying /j/ is limited to syllable-final/word-final position. All [j]-onsets are arguably the product of an epenthesis process (section 5.1). Additionally, [ʔ] can only occur in syllable-final position. Syllabification must respect this. So any sequence of XVʔVY is syllabified as XVʔ.VY, as in the examples (15).

(15)	[sa.kɔ.laʔ.an]	‘school’
	[ɲabɪʔ.i]	‘put on’

Also, it is not uncommon for speakers of the western dialect to pronounce an [h] on words ending in vowels, e.g. [bɪtɔh] ‘rock’, [k<sup>h</sup>urɪh] ‘teacher’, [apah]

‘what’, and so on, particularly when occurring before a significant pause. This is, however, unrelated to Indonesian words with a final /h/. Pawitra (2009) includes many examples in his dictionary.

As touched on in section 2, schwa can occur only in closed syllables, and this includes the higher counterpart [i̯]; as shown in section 4, the restriction of [ə] and [i̯] to closed syllables can induce gemination. Examples of [ə] and [i̯] include,

- (16) [dʌtəŋ] ‘come’  
 [əmpaʔ] ‘four’  
 [maŋəssəl] ‘regret’  
 [səlɭəʔ] ‘ring’  
 [biccəʔ] ‘good’  
 [birriʔ] ‘give’  
 [pəkʰi̯l] ‘angry’

Generally speaking, nasals must be homorganic with a following stop. Dental nasals occur adjacent to both dental and retroflex stops.

- (17) [əmpaʔ] ‘four’                      [tampʰɤŋ] ‘mine’                      [ambu] ‘stop’  
 [mantə] ‘son-in-law’                      [kantʰɤs] ‘run aground’                      [səndu] ‘difference’  
 [kaləntəŋ] ‘handcuffs’                      [manʰi] ‘bathe’                      [tandəs] ‘fast’  
 [kaŋca] ‘friend’                      [əŋcʰɤk] ‘pedal’                      [gɪŋɤŋ] ‘lively’  
 [səŋkəʔ] ‘I’                      [aŋkʰuɟ] ‘use’                      [aŋɤʔ] ‘conceited’

There are a small number of exceptions to this; for the most part, these are instances involving the velar nasal, such as [taŋtʰaŋ] ‘dance’, [bɤruŋbuŋaŋ] ‘sheath’, [bɤrɤŋtaʔ] ‘infatuated’, [paləŋpəŋ] ‘deaf’, [taŋtʰɤŋ] ‘unequal’, [karatəŋtəŋ] ‘without family’, [cʰuruŋcʰɤŋ] ‘a type of tall grass’.<sup>14</sup> In root words that are not borrowed, [l], [r], and [s] are preceded almost exclusively by the velar nasal, there being only a handful of instances in which they are preceded by [m]

<sup>14</sup>Some of the instances which are clearly not reduplicated roots are being regularized. So ‘dance’ often surfaces as [tantʰaŋ], which is given as the entry in Safioedin’s (1977) dictionary. However, Masharudin’s (2000) dictionary gives the entry as [taŋtʰaŋ], and Stevens (1968) cites this form as well. (Pawitra (2009) includes entries for both forms.) These are not instances of reduplication, as the onset of the first syllable is unaspirated [t] and the onset of the second is aspirated [tʰ].

or [n].<sup>15</sup> Some examples follow.

- (18) [aŋlɔ] ‘brazier’, [buŋlɔn] ‘type of tree’, [cɔŋlet] ‘type of lamp’,  
 [saŋlɛʔ] ‘kick’ [sɔŋlɔt] ‘peel’  
 [p<sup>h</sup>ŋsa] ‘like’, [t<sup>h</sup>uŋsəŋ] ‘nasal sound’, [k<sup>h</sup>ŋɕɛ] ‘grind’,  
 [saŋsara] ‘difficult’  
 [bɔlŋrat] ‘scratched’, [caŋrəp] ‘submerged’, [cɔŋrɔt] ‘decrease’,  
 [tɔŋrap] ‘sprout all at once’  
 [pamrɛ] ‘reward’  
 [jumla] ‘sum’  
 [lɛnra] ‘fence’

Phonetic vowel clusters are relatively limited. Sequences of identical vowels are disrupted by glottal stop epenthesis (section 5.1) and [j]-epenthesis between the object voice prefix [ɛ] and the locative prepositional clitic [ɛ] ‘at’ (sections 5 & 6). And sequences of non-like vowels where the first vowel is either front or back are disrupted by glide epenthesis (section 5). There are no əV sequences, as ə only occurs in closed syllables. Finally, [ai] and [ɤə] are, to the best of my knowledge, unattested. Thus, the only surface vowel clusters that are found are [aɛ], [aɔ], [ɤi] and [ɤu], as in

- (19) [paɛʔ] ‘bitter’                      [taɛ] ‘feces’  
 [taɔn] ‘year’                              [paɔ] ‘mango’  
 [bɤi] ‘else’                                [c<sup>h</sup>ɤi] ‘ginger’  
 [ɔɤun] ‘leaf’                               [c<sup>h</sup>ɤu] ‘far’

---

<sup>15</sup> There are cases of what appear to be clusters of [ml], [mr], [ms], [nr] and [ns]. However, they are quite limited (perhaps fewer than 20) and they clearly arise from reduplication of a monosyllabic root to form the preferred disyllabic root. Four examples are:

- (i) [lɔmlɔm] ‘calm’  
 [rɔmrɔm] ‘submerged in water’  
 [ranran] ‘yank’  
 [sɔnsɔn] ‘smoke from incense’

#### 4. Vowel inventory and distribution

While there are occasional discrepancies in the consonantal inventory cited in various works, there is more variation with vowels. For one, determining the precise inventory of Madurese vowels depends on whether one considers the phonetic level or the phonemic level; and even then there has been disagreement in the literature. From the standpoint of phonetic realization, there is some agreement that there are 12 vowel qualities that manifest themselves in surface forms. These are given in the chart in the following chart:<sup>16</sup>

	front	central	back
high	i (ɪ)	ɻ ɨ	u (ʊ)
mid	e, ε	ə	o, ɔ
low		a	

Of these half are either allophones that can be derived through regular phonological processes or constraints or auditory distinctions that may be due to other factors. These include ɪ, ʊ, e, o, ɨ, and ɻ. According to Stevens (1968), [ɪ] and [ʊ] result from a rule that lowers [i] and [u] in closed syllables, though this may actually reflect vowel length (see Cohn & Lockwood 1994 and discussion below). [e] and [o] are the output of a rule that raises [ɛ] and [ɔ]. These processes are discussed further in section 4.2. [ɨ] and [ɻ] result from the Vowel Harmony process (referred to as Vowel Tensing in Stevens 1994 and Vowel Raising in Cohn 1993a) discussed in the next section.

##### 4.1. Vowel harmony

Vowel harmony is perhaps the most striking feature of Madurese phonology. It was first detailed in the work of Kiliaan (1897, 1904) and more recently in that of Stevens (1968, 1980, 1985, and 1994).<sup>17</sup> Vowel harmony is contained in the system of paired alternating vowels that characterize the vast majority of lexical

<sup>16</sup>Oka et al. (1988/1989) represent 13 different vowels. No one else reports this many.

<sup>17</sup>The system has also generated interest in the theoretical literature in the work of Cohn (1993a,b), Trigo (1987, 1991), and Anderson (1991).

items. Stevens (1968) estimates that 95% of all Madurese words make use of these alternating vowels. Each pair consists of a high vowel and a non-high vowel that match in backness and whose distribution is determined by the phonological environment in which it occurs. The pairs are given in (20).

## (20) Alternating vowels

ε ~ i

ɔ ~ u

a ~ ʌ

ə ~ ɪ

The conditioning environment is the preceding consonant: high vowels follow the voiced and voiceless aspirated stops; the low counterparts occur elsewhere, although there are systematic exceptions for [l], [r], [ʔ], and [s] under specific circumstances discussed below. The regular case is illustrated first with the ε ~ i pair.

## (21) ε ~ i complementarity

ε		i	
p	pəkker 'think'	p <sup>h</sup>	p <sup>h</sup> ik <sup>h</sup> ɪl 'rob'
t	tɛŋk <sup>h</sup> i 'tall'	b	bilɻ 'when'
t̚	t̚ɛt̚ɛŋɛl 'make anxious'	t <sup>h</sup>	t <sup>h</sup> ittɛl 'shake'
c	cɛʔ 'very'	d	dibɻsa 'mature'
k	kɛnɛʔ 'small'	t̚ <sup>h</sup>	t̚ <sup>h</sup> iccaʔ 'lame'
m	mɛmpɛ 'dream'	d̚	d̚ijɻ 'here'
n	nɻmɔr 'dry season'	c <sup>h</sup>	c <sup>h</sup> irɻʔ 'knot'
ŋ	ŋɛjɔr 'coconut'	ʃ	ʃibut <sup>7</sup> 'corpulent'
ŋ	ŋɛd̚d̚ɻ 'sleep soundly'	k <sup>h</sup>	k <sup>h</sup> ɪʔ 'yet'
l	lɛwat <sup>7</sup> 'pass'	g	gigit <sup>7</sup> 'teeth'
r	rɛbɻ 'lap'		
s	sɛŋkɔʔ 'I'		
#___	ɛŋa <sup>7</sup> 'remember'		

Consider first just the bilabial stops: [ɛ] follows [p] in [pəkker] 'think', but following [p<sup>h</sup>] and [b], the high variant [i] occurs, as in [p<sup>h</sup>ik<sup>h</sup>ɪl] 'rob' and [bilɻ] 'when'. The same is true for the other four series of stops. [ɛ] occurs following all of the nasals as well as [l], [r], and [s], when the last three are in word-initial position. [ɛ] also occurs in word-initial position. The overwhelming majority of lexical items show this pattern. (Exceptions are discussed below.) While illustration with word-initial consonants is highlighted here, the pattern obtains

word internally as well. In [tɛŋk<sup>hi</sup>] ‘tall’, the non-high [ɛ] follows the voiceless unaspirated stop [t̚], but the high vowel [i] follows the voiceless aspirated stop. This is also illustrated by [t<sup>hi</sup>ittɛl] ‘shake’, in which [i] follows the voiceless aspirated dental stop and [ɛ] follows the voiceless unaspirated dental stop. The rest of the data conform to this regularity.

The pattern is illustrated for the other three pairs in (22-24).

(22) ɔ ~ u complementarity

ɔ		u	
p	pəkɔl ‘hit’	p <sup>h</sup>	p <sup>h</sup> urus ‘dog’
t	tɔɖus ‘embarrassed’	b	bujɣ ‘salt’
t̚	t̚ɔʔ ‘only’	t <sup>h</sup>	t <sup>h</sup> umen ‘after’
c	cɔŋŋɔʔ ‘look for’	d	dunggu ‘dull-witted’
k	kɔcɛŋ ‘cat’	t̚ <sup>h</sup>	t̚ <sup>h</sup> uʔ <sup>h</sup> ɣʔ ‘widower’
m	mɔsem ‘season’	ɖ	ɖuccɔl ‘loose’
n	nɔna ‘young girl’	c <sup>h</sup>	c <sup>h</sup> ukɔʔ ‘fish’
ɲ	ɲɔɲɔr ‘inflammable’	ʝ	ʝuʝuʔ ‘great-great grandparents’
ŋ	ŋɔɖɣ ‘young’	k <sup>h</sup>	k <sup>h</sup> unɔŋ ‘mountain’
l	lɔncɔʔ ‘jump’	g	gudunʝ ‘stupid’
r	rɔma ‘house’		
s	sɔratʔ ‘letter’		
#___	ɔp <sup>hi</sup> ŋ ‘screwdriver’		

(23) a ~ ɣ complementarity

a		ɣ	
p	pacɔl ‘hoe’	p <sup>h</sup>	p <sup>h</sup> ɣk <sup>h</sup> us ‘good’
t	tanɔŋ ‘hand’	b	bɣʔɣriʔ ‘yesterday’
t̚	t̚aratʔ ‘scream’	t <sup>h</sup>	t <sup>h</sup> ɣtt <sup>hi</sup> ‘become’
c	candɪ ‘temple’	d	dɣnɔ ‘dull-witted’
k	kap <sup>h</sup> ɣr ‘news’	t̚ <sup>h</sup>	t̚ <sup>h</sup> ɣsar ‘base/foundation’
m	mandɪ ‘bathe’	ɖ	ɖɣtɔŋ ‘come’
n	nakal ‘naughty’	c <sup>h</sup>	c <sup>h</sup> ɣbɣ ‘Java’
ɲ	ɲaman ‘pleasant’	ʝ	ʝɣkɛn ‘sure’
ŋ	ŋara ‘possible’	k <sup>h</sup>	k <sup>h</sup> ɣkk <sup>h</sup> ɣr ‘fall’
l	labɣŋ ‘door’	g	gɣrɣŋ ‘beautiful’
r	ratt <sup>hi</sup> n ‘pretty’		
s	sampɛʔ ‘until’		
#___	aɛŋ ‘water’		

## (24) ə ~ ɨ complementarity

ə		ɨ	
p	pək <sup>h</sup> ɨl ‘angry’	p <sup>h</sup>	p <sup>h</sup> ɨkkəm ‘grip’
t	təbbɨl ‘thick’	b	biccεɛʔ ‘good’
t̚	t̚ɨŋtəŋ ‘robust’	t <sup>h</sup>	t <sup>h</sup> ɨrrɨp̚ ‘belch’
c	cəppət̚ ‘fast’	d	dɨrpək̚ ‘stout’
k	kəŋŋaŋ ‘full’	t̚ <sup>h</sup>	t̚ <sup>h</sup> ɨkkəp̚ ‘embrace’
m	məlla ‘lukewarm’	ɖ	ɖɨrrɨŋ ‘roar’
n	nəpp <sup>h</sup> ɨʔ ‘flat’	c <sup>h</sup>	c <sup>h</sup> ɨllɨŋ ‘see’
ɲ	ɲəccɛŋ ‘too short’	ʃ	ʃɨrruk̚ ‘citrus fruit’
ŋ	ŋəpp <sup>h</sup> ɨt̚ ‘face down’	k <sup>h</sup>	kɨlɨ ‘crazy’
l	ləmbʔ ‘soft’	g	gɨmmɔʔ ‘fat’
r	rəpp <sup>h</sup> ɨ ‘grass’		
s	sənnəŋ ‘happy’		
#___	əndɨʔ ‘willing’		

The alternations are most dramatically illustrated in the pairs of root/nasal actor voice forms of verbs. The actor voice nasal prefix, which is /ŋ/ in underlying form, replaces the initial consonant of the verb root with a homorganic nasal. This is illustrated first for roots with bilabial stops in initial position.

(25) underlying root	surface root	actor voice	
/penta/	[penta]	[mənta]	‘ask’
/betəŋ/	[bitəŋ]	[mətəŋ]	‘count’

The data in (24) demonstrate the Vowel Raising process in deriving surface root forms and actor voice forms from the underlying forms /penta/ ‘ask’ and /betəŋ/ ‘count’. The underlying forms contain only non-high vowels. In the surface root forms, [penta] and [bitəŋ],<sup>18</sup> the non-high vowel [ɛ] follows [p] in the first syllable of the first word and the high vowel [i] follows [b] in the first syllable of the second word, following the application of Vowel Raising. In the actor voice, the initial bilabial stop of each root is replaced with the bilabial nasal [m]. For the root [penta], the vowel of the first syllable remains the same. However, for the root [bitəŋ], the vowel of the first syllable surfaces as [ɛ] following the general principle that high vowels do not occur immediately following nasals.

<sup>18</sup>Surface root forms occur in the citation form of the word (as in a dictionary entry) as well as in the object voice form, which consists of the object voice prefix ε- and the root. For the two verbs in (25) the object forms would be [ɛpenta] and [ɛbitəŋ].

The forms in (26) further illustrate the pattern with surface root and actor voice forms, demonstrating its regularity.

(26)	underlying	surface root	actor voice	
	/kɛrɛm/	[kɛrɛm]	[ŋɛrɛm]	‘send’
	/k <sup>h</sup> ɛp <sup>h</sup> a/	[k <sup>h</sup> ip <sup>h</sup> ɿ]	[ŋɛp <sup>h</sup> ɿ]	‘carry’
	/pareksa/	[pareksa]	[mareksa]	‘examine’
	/baca/	[bɿca]	[maca]	‘read’
	/tamən/	[tamən]	[namən]	‘plant’
	/ɖapaʔ/	[ɖɿpaʔ]	[napaʔɛ]	‘arrive’
	/tɔləs/	[tɔləs]	[nɔləs]	‘write’
	/t <sup>h</sup> ɔtt <sup>h</sup> ɔʔ/	[t <sup>h</sup> utt <sup>h</sup> uʔ]	[nɔtt <sup>h</sup> uʔ]	‘point’
	/bøllɛ/	[billi]	[møllɛ]	‘buy’

An alternative to the nasal actor voice prefix is the actor prefix *a-*, which occurs with specific transitive roots (predominantly roots with initial aspirated or voiced stops).<sup>19</sup> Occasionally these roots take the *ŋ-* prefix (especially in the Western dialect), and when they do, the vowel optionally follows the vowel height generalization. For example, with the root [p<sup>h</sup>uktɛ] ‘prove’, the standard actor voice form is [ap<sup>h</sup>uktɛ]. However, some speakers also use the *ŋ-* prefix. In this case, the form [mɔktɛ] is predicted, as the non-high variant of the back vowel should follow the nasal [m], and speakers occasionally produce this form. However, speakers also produce [muktɛ], which employs the nasal prefix but maintains the high vowel of the root. Other examples are:

(27)	root	surface root	<i>a-</i> prefix	<i>ŋ-</i> prefix
	/bɔkkaʔ/ ‘open’	[bukkaʔ]	[abukkaʔ]	[mɔkkaʔ] ~ [mukkaʔ]
	/k <sup>h</sup> ɔrɛŋ/ ‘fry’	[k <sup>h</sup> uriŋ]	[ak <sup>h</sup> uriŋ]	[ŋɔrɛŋ] ~ [ŋuriŋ]
	/t <sup>h</sup> ɛŋk <sup>h</sup> al/ ‘remain’	[t <sup>h</sup> iŋk <sup>h</sup> ɿl]	[at <sup>h</sup> iŋk <sup>h</sup> ɿl]	[nɛŋk <sup>h</sup> ɿl] ~ [niŋk <sup>h</sup> ɿl]

Clearly, the pairs of vowels represent allophonic alternation, and only one from each pair is a true phoneme. Given the wider distribution of the non-high vowels, /ɛ/, /ɔ/, /a/ and /ə/ can reasonably be assumed to be the underlying vowels, a conclusion reached by Stevens (1968, 1980) and Cohn (1993a,b) as well. The high variants are derived by a regular process variously referred to as Vowel Tensing (Stevens 1994) or Vowel Raising (Cohn 1993). As outlined above, descriptively, vowels are tensed or raised following voiced and aspirated stops. Trigo (1991) and Cohn (1993a) have proposed formal phonological ana-

<sup>19</sup>The *a-* prefix is also used with the majority of active intransitive verbs. See Chapter 9.

lyses in which the conditioning environment depends on a feature that they refer to as [lowered larynx] which is manifested in the voiced and voiceless aspirated consonants.

Given the conditioning environment, the non-high vowels are expected following [l], [r], [ʔ] and [s]. And that is what the data show when considering only word-initial [l], [r], and [s]. (Recall that [ʔ] does not occur syllable initially.) However, there are data in the previous sets that show instances of high vowels following [l] and [r]. Consider the data in (28), which also include forms with [ʔ].

(28)	non-high V	high V
	[tɔləs] ‘write’	[bilɣ] ‘when’
	[møllɛ] ‘buy’	[billi] ‘buy’
	[kɛrɛm] ‘send’	[p <sup>h</sup> urus] ‘dog’
	[kərrɛs] ‘knife’	[t <sup>h</sup> irɣɣp] ‘belch’
	[napaʔɛ] ‘arrive at’	[nabɣʔi] ‘put on’

The data in (28) indicate that the vowels that follow [l], [r], and [ʔ] match the height of the vowels that immediately precede them. This is true for [l] and [r] even when geminated. These segments are thus transparent for the determination of the quality of the vowels that follow them. That non-high vowels follow [l] and [r] when these consonants occur word-initially can be attributed to the fact that there is no word-internal conditioning vowel that precedes [l] and [r] in this environment and so the default non-high vowel occurs. The glides [w] and [j] show the same effect. However, when internal to a word, it can be argued that these glides are epenthesized (see section 2.5.1). If so, then the rule determining vowel height can operate prior to the epenthesis, in which case it is unnecessary to consider these elements ‘transparent’ in the same sense as [l], [r], and [ʔ].

Finally, [s] behaves differently yet. Internal to a morpheme, [s] is followed by non-high vowels. This was illustrated for word-initial [s] above, and is also clear in the following:

(29)	[bisa]	‘can’
	[t <sup>h</sup> ɣsar]	‘base/foundation’
	[nɛsɛr]	‘pity/love’
	[pasar]	‘market’

Regardless of whether or not the vowel of the preceding syllable is high, such as [i] or [ɣ], or non-high, such as [ɛ] or [a], the following vowel is non-high. However, at a morpheme boundary, the height of the vowel following [s] depends on

the height of the vowel in the preceding syllable. This is illustrated in the data in (30).<sup>20</sup>

(30)	tələs + -ε	→	[tələsɛ]	‘write’ (durative)
	p <sup>h</sup> uŋkəs + -an	→	[p <sup>h</sup> uŋkəsən]	‘package’ (wrapped thing)
	gibɾs + -ε	→	[gibɾsi]	‘wag’ (durative)
	bɾlis + -an	→	[bɾlisən]	‘returned thing’

In the durative form of ‘write’ [tələsɛ], the durative morpheme occurs as the non-high [ɛ] and the vowel of the root-final syllable is non-high. However, in the durative form of ‘wag’, [gibɾsi], the durative morpheme occurs as the high vowel [i], and in this instance the vowel of the root-final syllable is high [ɾ]. Similarly, the nominal suffix *-an* alternates between [an] and [ən] when suffixed to a root ending in [s], again depending on the vowel of the final syllable of the root. In (28), when the root-final syllable contains the non-high vowel [ɔ], the suffix takes the non-high form [an], [p<sup>h</sup>uŋkəsən] ‘package’. Conversely, in [bɾlisən] ‘returned thing’, the vowel of the final syllable of the root is high [i] and the suffix takes the high vowel form [ən].

Returning to the question of the phonemic inventory, there are a number of lexical items that contain what Stevens (1968) refers to as non-alternating vowels. These vowels do not follow the general pattern, including cases of high vowels that follow segments other than voiced and voiceless aspirated stops (or [l], [r], [ʔ], and [s] under the prescribed conditions). Most of these lexical items are borrowings, predominantly from Arabic, Dutch, English, and Indonesian. Nonetheless these forms have been fully absorbed into the language. Examples include:

---

<sup>20</sup> In the discussion that follows, phonological processes are applied to intermediate forms, forms which have already undergone the process of Vowel Raising (VR). Therefore neither phonemic nor phonetic bracketing is used. This is a matter of both clarity and brevity. For example, the derivation from underlying representation (UR) to surface representation (SR) is as follows:

(i)	UR		VRaising		Affixation		VRaising/ Phonetic Realization
	/tələs/	→	---		tələs + -ε	→	[tələsɛ] ‘write’
	/gɛbas/	→	gibɾs		gibɾs + -ε	→	[gibɾsi] ‘wave’

Dispensing with the step in which Vowel Raising takes place reduces unnecessary clutter, thus making the phonological process of interest more perspicuous.

(31) Lexical items with non-alternating vowels<sup>21</sup>

[ban] ‘tire’  
 [baŋ] ‘bank’  
 [bɔla] ‘ball’  
 [bɔrgɔl] ‘manacle’  
 [buku] ‘book’  
 [gɔbis] ‘cabbage’  
 [kibik] ‘cubic’  
 [kipas] ‘fan’  
 [kɔpi] ‘coffee’  
 [kuwi] ‘cake’  
 [sidaʔ] ‘choke’  
 [sipil] ‘civilian’  
 [surgap] ‘conceited’  
 [susu] ‘milk’  
 [tipəs] ‘typhus’

The list includes cases in which high vowels follow voiceless consonants, as in [buku] ‘book’, [kɔpi] ‘coffee’, [kibik] ‘cubic’, [susu] ‘milk’, and [tipəs] ‘typhus’. There are also instances in which a non-high vowel follows a voiced stop, such as [ban] ‘tire’, [baŋ] ‘bank’, [bɔla] ‘ball’, and [gɔbis] ‘cabbage’. The only high vowels that are found in such lexical items are [u] and [i]; [ɨ] and [ɻ] are never found outside the conditioning environment for vowel alternation.

There is another important difference among [u], [i], [ɨ], and [ɻ]: [u] and [i] occur word-initially in a small number of lexical items (32), whereas [ɨ] and [ɻ] never do.

(32) [irigɻsi] ‘irrigation’  
 [iseŋ] ‘troublesome’  
 [istimɛwa] ‘special’  
 [uɲijɻn] ‘examination’  
 [usul] ‘character’  
 [uwɻʔ] ‘steam’

---

<sup>21</sup>Often, the forms that have been borrowed from Indonesian have variants with non-alternating and alternating vowels. Examples include:

- (i) [susu] ~ [sɔsɔ] ‘milk’  
 [tulɛn] ~ [tɔlɛn] ‘pure’  
 [tulus] ~ [tɔlɔs] ‘reliable’

In order to account for these instances of [i] and [u] outside the conditioning environment, they must be posited in underlying representations. It is never necessary to posit [ī] and [ɤ] in underlying representations of lexical items. As the occurrence of [ī] and [ɤ] can always be predicted, there is no need to include them in the phonemic inventory of the language.

Having established the vowel harmony (or raising or tensing) process and the phonemic inventory as the six vowels /ɛ/, /ɔ/, /a/, /ə/, /i/, and /u/, it is clear this should be reflected in the underlying representations of lexical items. Thus, we have the following:

(33) underlying	surface
/k <sup>h</sup> ɛp <sup>h</sup> a/	[k <sup>h</sup> i <sup>h</sup> p <sup>h</sup> ɤ] ‘carry’
/t <sup>h</sup> ɛŋk <sup>h</sup> al/	[t <sup>h</sup> i <sup>h</sup> ŋk <sup>h</sup> ɤl] ‘remain’
/pəkk <sup>h</sup> əl/	[pəkk <sup>h</sup> i <sup>h</sup> l] ‘angry’
/c <sup>h</sup> ɔkɔʔ/	[c <sup>h</sup> ukɔʔ] ‘fish’
/k <sup>h</sup> ɔrɛŋ/	[k <sup>h</sup> urɛŋ] ‘fry’

All of the high vowels in the surface forms are derived via the harmony rule. In words with non-alternating vowels, the surface and underlying forms are identical. However, in order to focus discussion and make forms maximally transparent, I include the surface form of the vowels in all roots.

A final issue regarding the vowel inventory is the status of [ɨ] as a surface vowel. In works such as Stevens (1980, 1994) and Cohn & Lockwood (1994), eight vowels are posited as being the principle vowels in Madurese, the eight described above as participating in the vowel alternation. However, other sources differ on the number of primary vowels. (I use the term ‘primary vowels’ here as the authors tend not to identify the inventories as being phonemic or phonetic.) Budi et al. (1986/1987) recognize seven vowels, as do Safioedin (1977) and Zainudin (1978). While there seems to be some conflation of [ə], [ɤ] and [ī] in Budi et al., Safioedin (1977), Zainudin (1978), and Pawitra (2009) posit [ī], [ɛ], [u], [ɔ], [a], [ə], and [ɤ] (although transcription systems vary somewhat). Moehnilabib (1979) recognizes six vowels, corresponding to the phonemic inventory posited here. Notably absent in all but Stevens and Cohn & Lockwood is any distinction between [ə] and [ī]. And, in fact, no practical orthography ever devised for Madurese has ever recognized such a distinction (see section 8). Those that make use of symbols other than *a*, *e*, *i*, *o*, *u* identify six or seven graphemes but never include a distinct symbol for [ɨ].

What, then, is the status of [ī]? It might be recognized for theoretical reasons, i.e. to regularize the system of alternating vowels. However, absent other evidence, ensuring paradigm uniformity for the alternating vowel system seems an insufficient justification. There does, however, appear to be some instrumental evidence for distinguishing [ə] and [ī]. Cohn & Lockwood (1994) report that

the analysis of recordings of one male and one female speaker from East Madura indicate the vowel space (based on first and second formants) occupied by [i] is distinct from that of [ə]. At the same time, the difference in the vowel space that they report for [ə] and [i] is much less dramatic than the differences for the other pairs, and indeed there was some overlap in their distribution. The space occupied by each of the vowels in the other pairs was essentially completely distinct from its partner. Cohn & Lockwood found little difference in the average F2s of the pair of vowels but found the difference between the average F1s to be comparable for [i]/[ɛ], [u]/[ɔ] and [i]/[ə] (the difference in the F1 averages for [ɤ]/[a] were dramatically higher). In attempting to replicate these results with a single male speaker from Jember, East Java, Bortscheller (2007) reports a somewhat more dramatic difference in the vowel space occupied by [ə] and [i]. He also reports a slightly greater difference between the average F1s for [ə] and [i] than found by Cohn & Lockwood for their male speaker. But again, the difference in average F1 is comparable to that found with the other pairs of vowels. While not a matter of great urgency, given the instrumental evidence (albeit slim), it seems that the distinction should be recognized.

#### 4.2. Other vowel processes

Finally, there are three low-level phonetic processes that have been proposed for some apparent tense/lax allophonic variation and a vowel nasalization process. The first is a rule that alters the quality of a high front or high back vowel in closed syllables, accounting for perceived instances of [ɪ] and [ʊ]. Stevens (1968) cites this rule, and the reflexes [ɪ] and [ʊ] are included in Safioedin (1977) and Oka et al. (1988/89). Potential examples include:

- (34) [bintaŋ] ~ [bɪntaŋ] ‘star’  
 [k<sup>h</sup>ippɔŋ] ~ [k<sup>h</sup>ɪppɔŋ] ‘confused’  
 [buntɔʔ] ~ [bɪntɔʔ] ‘tail’

Cohn & Lockwood (1994) report that for the two speakers they studied, high front and high back vowels were significantly shorter in duration in closed syllables than in open syllables. They reported no significant differences in the F1 and F2 values in the same environments. I have been able to confirm these results to some degree through acoustic analysis of data collected from four speakers.<sup>22</sup> There were no significant differences in the F1 and F2 values in open and closed syllables, but there was some difference in vowel length. These

---

<sup>22</sup> Three speakers were recorded on Madura and one speaker originally from Situbundo (East Java) was recorded in Iowa City. Thanks to Michael Bortscheller, Craig Dresser, and Jeff Press for assistance with the acoustic analysis.

results indicate that speakers vary with respect to the strength and prevalence of this process, the distinction more robust with some speakers and nearly non-existent with others. However, this same type of laxing in closed syllables has been reported some closely-related languages (Indonesian (Lapoliwa 1981), Karo Batak (Woollams 1996)).

Similar to the tense/lax alternation is a reported raising and/or tensing process affecting non-high front and back vowels in open syllables. The reflexes of this get reported as [e] and [o] at times, e.g. Oka et al. 1988/89, and at others as a slightly raised  $\epsilon$ , [ $\epsilon^{\wedge}$ ], or  $\text{ɔ}$ , [ $\text{ɔ}^{\wedge}$ ] (Stevens 1968). Potential examples include:

- (35) [s $\epsilon$ ] ~ [s $\epsilon^{\wedge}$ ] ~ [se] ‘relative particle’  
 [reja] ~ [r $\epsilon^{\wedge}$ ja] ~ [reja] ‘that’  
 [p $\text{ɔ}\epsilon$ ] ~ [p $\text{ɔ}^{\wedge}$ l $\epsilon^{\wedge}$ ] ~ [pole] ‘again’  
 [rat $\text{ɔ}$ ] ~ [rat $\text{ɔ}^{\wedge}$ ] ~ [rato] ‘king’

Again, there is a great deal of speaker variation. Acoustic analysis does not confirm a reliable difference in open and closed syllables. Additionally, some lexical items seem to be more susceptible to this process than others, for example, the word *toko* ‘store’ is usually pronounced [toko], whereas *nyoro* ‘command’ is usually pronounced closer to [n $\text{ɔ}\text{r}\text{ɔ}$ ].<sup>23</sup>

Cohn and Lockwood (1994) report a different tense/lax alternation involving [ $\epsilon$ ]/[e] and [ $\text{ɔ}$ ]/[o]. They report systematic tensing and raising of front and back non-high vowels following nasals, citing the data in (36), which confirms Stevens’ (1968) observation (and reflects the nasalization rule to be discussed presently).

- (36) [p $\epsilon$ ka] ‘pinch’                      [m $\epsilon$ kətʰ] ‘trap’  
 [p $\text{ɔ}$ ka] ‘break’                         [m $\text{o}$ ka] ‘break (actor voice)’

*Nasalization.* Another minor phonetic realization rule affecting vowels is nasalization. Vowels immediately following nasals are lightly nasalized, as are vowels and glides adjacent to nasalized vowels. Glottal stop is transparent to vowel nasalization.

- (37) maca                                 →        [m $\tilde{a}$ ca] ‘read’  
 $\epsilon\text{n}\text{ɔ}\text{m}$                                     →        [ $\epsilon\text{n}\tilde{\text{o}}\text{m}$ ] ‘drink’

<sup>23</sup>In fact, Stevens (1968:37) reports this in some instances as well, e.g. [soto] ‘a kind of soup’, but not others, e.g. [s $\text{ɔ}\text{s}\text{ɔ}$ ] ‘milk’.

nɛat	→	nɛjat	→	[nɛ̃ʃjātʰ]	'intend'
ɛŋaʔ	+ ε		→	[ɛŋãʔɛ̃]	'remember'

As it is unimportant to points which follow, nasalization is not marked on vowels.

## 5. Phonological processes

There are a number of process that derive surface structures from underlying phonemic forms.

### 5.1. Epenthesis

There are three principal processes that insert consonants between two contiguous vowels.

*Glottal insertion.* A glottal stop is epenthesized between 2 identical vowels, either at a morpheme boundary or root internally, which has the effect in (38).

$$(38) \emptyset \rightarrow ? / V_1 \_ V_1$$

(39)	sakɔla	+ an	→	[sakɔlaʔan]	'school'
	mate	+ ε	→	[mateʔɛ]	'kill'

In the word for 'school', [sakɔlaʔan], [ʔ] is inserted between the root-final [a] of [sakɔla] and the suffix-initial [a] of [an]. In [mateʔɛ] 'kill', [ʔ] is inserted between the root-final [ɛ] of [mate] and the suffix [ɛ]. Other examples include:

(40)	maca	+ a	→	[macaʔa]	'will read'
	abɤlɤ	+ ak <sup>hi</sup>	→	[abɤlɤʔɤk <sup>hi</sup> ] <sup>24</sup>	'tell about'
	birse	+ ε	→	[birseʔɛ]	'clean'

---

<sup>24</sup>With the suffix *-ak<sup>hi</sup>*, one actually finds variation between [abɤlɤʔɤk<sup>hi</sup>] and [abɤlɤʔak<sup>hi</sup>], in which the initial vowel of the suffix does not undergo the harmony rule. Stevens (1994) attributes this optionality to the fact that the form *ak<sup>hi</sup>* can be more closely bound to the root as a suffix, resulting in [abɤlɤʔɤk<sup>hi</sup>] (in this case), or more loosely bound, almost as an independent word, resulting in [abɤlɤʔak<sup>hi</sup>].

Note that in the form [abʌlʌʔak<sup>hi</sup>] ‘tell about’, [ʔ] is inserted not between two identical surface vowels but between two vowels of one of the alternating pairs, [ʌ] and [a]. This provides additional evidence that in underlying representation the paired vowels are identical. Glottal insertion also occurs root internally, as illustrated by:

(41)	lɛɛr	→	[lɛʔɛr]	‘neck’
	tɔɔt	→	[tɔʔɔtʰ]	‘knee’
	bʌʌ	→	[bʌʔʌ]	‘flood’
	naas	→	[naʔas]	‘misfortune’

Root-internal [ʔ] is completely predictable, occurring only between two identical vowels. Additionally, the Indonesian cognates of these words are virtually identical, merely lacking the glottal stop and the precise quality of vowels.

*j-epenthesis.* The lone exception to the glottal insertion process is a morphophonemic rule occurring in the environment in which the ‘passive’ prefix  $\varepsilon$ - precedes a stem-initial  $\varepsilon$  or the prepositional element  $\varepsilon$  ‘at’ cliticizes to a stem-initial  $\varepsilon$ . In this environment, the palatal glide [j] is epenthesized rather than the expected [ʔ].<sup>25</sup>

(42)  $\emptyset \rightarrow j/ \varepsilon + \_ \varepsilon$

(43)	$\varepsilon + \varepsilon t^{hi}n\varepsilon$	→	[ɛjɛt <sup>hi</sup> nɛ]	‘be allowed’
	$\varepsilon + \varepsilon nɔm$	→	[ɛjɛnɔm]	‘be drunk’
	$\varepsilon + \varepsilon s\varepsilonʔ\varepsilon$	→	[ɛjɛs\varepsilonʔ\varepsilon]	‘be filled’
	$\varepsilon + \varepsilon lɔŋ$	→	[ɛjɛlɔŋ]	‘in (the) nose’

*Glide epenthesis.* As stated in section 1, glides in Madurese are rarely phonemic. There are a small number of loan words in which word-initial /j/ and /w/ are found and there are some native words with final /j/. (See section 1.1.) The vast majority of surface glides arise from a process of glide epenthesis that inserts [j] or [w] between contiguous vowels that have different specifications for backness. This occurs word internally or before a suffix, the glide conditioned by the backness of the first vowel. [j] is inserted after a front vowel, and [w] is inserted after a back vowel.

<sup>25</sup>At times, speakers override this j-epenthesis, inserting [ʔ] instead, as in

(i)  $\varepsilon + \varepsilon c\varepsilon t \rightarrow [\varepsilonʔ\varepsilon c\varepsilon tʰ]$

- (44) mǎlle + ak<sup>hi</sup> → [mǎllejak<sup>hi</sup>] ‘buy for’ (actor voice)  
 ŋa + taɔ + ak<sup>hi</sup> → [ŋataɔwak<sup>hi</sup>] ‘convince’ (actor voice)

In the form [mǎllejak<sup>hi</sup>] ‘buy for’, [j] is inserted between the stem-final [ɛ] of [mǎlle] ‘buy’ and the initial [a] of the benefactive suffix [ak<sup>hi</sup>]. However in [ŋataɔwak<sup>hi</sup>] ‘convince’ [w] is inserted between the stem and [ak<sup>hi</sup>] because the stem-final vowel is back and round [ɔ]. Other examples include:

- (45) libɔli + an → [libɔlijɔn] ‘several times’  
 t<sup>h</sup>ɔtt<sup>hi</sup> + a → [t<sup>h</sup>ɔtt<sup>hi</sup>jɔ] ‘will become’  
 ɔlle + a → [ɔlleja] ‘will get’  
 ka + ratɔ + an → [karatɔwan] ‘palace/kingdom’  
 ɛka + taɔ + ɛ → [ɛkataɔwɛ] ‘is known’  
 abɔntɔ + a → [abɔntɔwa] ‘will help’

The occurrence of root-internal glides is also completely predictable and can be derived through glide epenthesis. Examples include:

- (46) neat → [nejat<sup>h</sup>] ‘intend’  
 sɛər → [sɛjər] ‘fall asleep’  
 ŋɛər → [ŋɛjər] ‘coconut’  
 bɔriɔŋ → [bɔrijɔŋ] ‘feel unwell’  
 diəm → [dijəm] ‘calm’  
 t<sup>h</sup>iuk → [t<sup>h</sup>ijuk] ‘commotion’  
 tɔa → [tɔwa] ‘old’  
 sɔər → [sɔwər] ‘notch’  
 duɔʔ → [duwɔʔ] ‘two’  
 bui → [buwi] ‘fetter’

It is telling that for those roots with an Indonesian cognate, the Indonesian word lacks the glide but in all other respects is virtually identical to the Madurese: *niat* ‘intend’, *diam* ‘calm’, *bui* ‘fetter’, *dua* ‘two’, *tua* ‘old’.

## 5.2. Gemination

There are a number of gemination processes in Madurese, but by far the most productive and frequent of these is gemination that ensures that the syllable structure constraint on [ə] is satisfied. Recall that [ə] can only occur in closed syllables. When a vowel-initial suffix is added to a consonant-final root, resyllabification takes place so that the final consonant becomes the onset of the fol-

lowing syllable. (45) illustrates.

(47) ηε.ρεμ ‘send’ + a ‘IRR’ → [ηε.ρε.μα] ‘send (IRR)’

In (47), [m] forms the coda of the second syllable of the root. But, in the derived form [ηερεμα], [m] is the onset of third syllable. Importantly, [ε] is the vowel of the second syllable, and [ε] can occur in an open syllable, and so the form [ηερεμα] is well formed. However, when the root-final syllable contains [ə], the same process would result in a form that violates the phonotactic constraint on [ə], as illustrated in (48).

(48) μα.πəs.σəl ‘regret’ + a ‘IRR’ → \* [μα.πəs.σə.λα]

Gemination of the stem-final [l] of [μαπəσσəl] preserves the environment for [ə] and ensures well-formedness, as in

(49) μα.πəs.σəl ‘regret’ + a ‘IRR’ → [μα.πəs.σəl.λα] ‘regret (IRR)’  
Additional examples include:

(50) p<sup>h</sup>usən ‘bored’ + ε ‘LOC’ → [p<sup>h</sup>usənnε] ‘bored with’  
pəkk<sup>h</sup>il ‘angry’ + ε ‘LOC’ → [pəkk<sup>h</sup>illi] ‘angry at’  
ka + ənnəŋ ‘stay’ + an ‘NOM’ → [kənnəŋŋan] ‘place’  
naŋkəp ‘catch’ + a ‘IRR’ → [naŋkəpp<sup>h</sup>ɣ] ‘catch (IRR)’<sup>26</sup>  
muɟ<sup>h</sup>ik ‘strike’ + ε ‘DUR’ → [muɟ<sup>h</sup>ikk<sup>h</sup>i] ‘strike repeatedly’

This contrasts with cases in which the vowel in the final syllable is neither [ə] nor [i]. In these cases there is no gemination of the consonant.

(51) tələs ‘write’ + ε ‘DUR’ → [tələsε] ‘write (DUR)’  
ηερεμ ‘send’ + ε ‘LOC’ → [ηερεμε] ‘send to’  
p<sup>h</sup>uŋkəs ‘wrap’ + an ‘NOM’ → [p<sup>h</sup>uŋkəsən] ‘package’  
ɟɰwɰp ‘answer’ + an ‘NOM’ → [ɟɰwɰp<sup>h</sup>ɰn] ‘answer’  
nətəp ‘close’ + a ‘IRR’ → [nətəp<sup>h</sup>ɣ] ‘close (IRR)’

Note that it is entirely possible that like ?-insertion and glide-epenthesis, this type of gemination may occur word internally as well. For example, the root

<sup>26</sup>The forms [naŋkəpp<sup>h</sup>ɣ] ‘will catch’ and [muɟ<sup>h</sup>ikk<sup>h</sup>i] ‘strike repeatedly’ undergo the process of aspiration (discussed in section 6) in addition to gemination. This additionally causes the raising of the vowel of the suffix.

[sənnəŋ] ‘happy’ includes geminate *n*. The cognate in Indonesian is [sənəŋ], with a single [n]. Indonesian does not contain the same type of syllable-structure constraint on [ə]. So, it is conceivable that the underlying form in Madurese is /sənəŋ/, and the gemination process takes place to create the appropriate syllable structure for [ɽ]. This is, in fact, Stevens’ (1968:57) approach, where, for example, an intermediate form for the actor form for ‘buy’ is *mələ* and gemination occurs in the derivation of the surface form [məllə].<sup>27</sup>

Stevens (1968) cites two additional gemination processes, both of which are obligatory. In the first, which is considerably more productive than the second, a stop is geminated when it is adjacent to a liquid root internally. The process derives a form such as [pöttɾɛ] ‘princess, daughter’ from the underlying form /pötɾɛ/. The underlying form proposed by Stevens is the same as the underlying form and phonetic realization of the Indonesian cognate *putri*. In the spirit of this analysis, this gemination process operates in the derivation of the following:

(52)  $C_1 \rightarrow C_1 C_1 / \_ \{l/r\}$   
[STOP]

(53) juk <sup>h</sup> lɔŋg	→	[juk <sup>h</sup> lɔŋg] ‘large hole, ravine’
katrəl	→	[kattɾəl] ‘pulley’
əbrɔl	→	[əbbɾɔl] ‘waste’
sadrika	→	[saddrika] ‘flat-iron’
seplak	→	[sepplak] ‘slap’
səklat	→	[səkklat <sup>ɽ</sup> ] ‘brown’

The second of the two minor gemination processes is quite restricted, affecting only *s* when adjacent to *r* word internally. Stevens cites the forms in (54) as examples.

(54) asɾɛŋ	→	[assɾɛŋ] ‘frequent’
pasra	→	[passra] ‘hand over’

### 5.3. Deletion

There is an optional deletion process which affects the vowel of the first syllable of a word. When the first syllable is the third syllable from the end of the

<sup>27</sup>As is the case with vowel quality, I include geminated consonants in all forms rather than listing any underlying forms with single consonants that are subsequently geminated in order to focus discussion and make forms maximally transparent.