

Selecting for and Selecting Despite: A Javanese case study

by

Cora Lesure

B.A. Honours Linguistics, McGill University, 2015

Submitted to the Department of Linguistics and Philosophy
in partial fulfillment of the requirements for the degree of

DOCTORATE OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2023

© 2023 Cora Lesure. All rights reserved.

The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Authored by: Cora Lesure
Department of Linguistics
September 8, 2023

Certified by: Adam Albright
Professor of Linguistics, Thesis Supervisor

Certified by: Norvin Richards
Professor of Linguistics, Thesis Supervisor

Certified by: Shigeru Miyagawa
Professor of Linguistics, Thesis Supervisor

Accepted by: Danny Fox
Anshen-Chomsky Professor of Language & Thought
Department Head, Department of Linguistics and Philosophy

Selecting for and Selecting Despite: A Javanese case study

by

Cora Lesure

Submitted to the Department of Linguistics and Philosophy
on September 8, 2023 in partial fulfillment of the requirements for the degree of

DOCTORATE OF PHILOSOPHY

ABSTRACT

This is an investigation of the argument structure of Javanese (Austronesian, Indonesia) which focuses on the distribution of four core derivational morphemes: the Actor Voice prefix, and the suffixes -ake, -i, and -an. The project is based on original consultant work conducted with a speaker of the Central dialect of Javanese. The work establishes language internal diagnostics for various aspects of a stem's lexical semantics and lexical category and then utilizes these criteria to analyze a wide variety of morphological derivatives, both verbal and nominal. The resulting analysis is able to predict the distribution of derivational morphemes and the nature of their resulting derivatives to a higher degree than what was previously understood to be possible.

Thesis supervisor: Adam Albright

Title: Professor of Linguistics

Thesis supervisor: Norvin Richards

Title: Professor of Linguistics

Thesis supervisor: Shigeru Miyagawa

Title: Professor of Linguistics

Acknowledgments

Every year of my graduate study was marked with some form of profound change either on a personal or global level. From my own experiences with chronic illness starting in 2017, through the global COVID-19 pandemic starting in 2020, and into my transition to becoming a liveaboard sailor from 2021 and ranging into the present, the times have been in a word "unprecedented". Subsequently, it is difficult to adequately express my gratitude for all the people and systems that supported me throughout this eventful period.

I am incredibly grateful for the continued support of my advisor Adam Albright, not only in guiding my exploration through the difficult questions posed by derivational morphology, but for providing a sense of community and a spirit of stick-to-itiveness through the most uncertain moments of the past few years. The morning check-in meetings you organized starting in the Spring of 2020 truly made a world of difference in the quality of my day to day life, and serve as a prime example of your dedication to your students.

Similarly I am very grateful for Norvin Richards, for always being present as a sounding board for sticky bits of data and continually encouraging the spirit of curiosity tempered by intentionality inherent to being a good fieldworker. Thank you also to Shigeru Miyagawa for his ability to provide straightforward insight into very big questions, and for somehow always finding the time to engage meaningfully with my ideas, even when they relate to such far flung topics as the acoustic nature of caves.

I am thankful for all the ways in which the entire department, despite wildly variable circumstances and an altogether unorthodox trajectory, continued to welcome and support me. I am grateful for my fellow students, my office mates, and particularly all the members of String Adjacent (the unofficial knitting group of the MIT linguistics department that I founded in 2019), who were willing to sacrifice some of their valuable time and energy to come together and knit.

Beyond academic support I am infinitely grateful for the family and friends that saw me through this experience, particularly my parents and my dearest friend Nicole. Through countless Facetime calls and last minute weekend visits home, your continued encouragement and friendship helped me maintain my sense of purpose as well as my sanity.

Finally, I feel incredibly fortunate for being able to work with Ismartilah Drummond, my Javanese language consultant. Even through switching to virtual meetings in the wake of the pandemic, she has continued to meet my incessant questions with patience, curiosity, and invaluable insight. This work would not have been possible without her and I look forward to continuing to learn more about this endlessly interesting language.

Contents

Title page	1
Abstract	3
Acknowledgments	5
1 Introduction	9
1.1 Structure of the dissertation	9
2 The Empirical Problem	11
2.1 Task 1: The Range of Possible Derivatives	13
2.1.1 Verbalizing Strategies	13
2.1.2 Nominalizing Strategies	19
2.2 Towards Task 2: Understanding what is actually <i>derived</i>	21
3 Understanding the Base	23
3.1 Disentangling Intransitives	23
3.1.1 <i>lagi</i> : ‘right now’	24
3.1.2 <i>isih</i> : being and doing <i>still</i>	29
3.1.3 <i>rada</i> : halfway and kind of doing/being	33
3.1.4 <i>meh</i> : ‘almost’	36
3.1.5 <i>ke-√-en</i> : ‘too X’	38
3.1.6 Interim Summary	39
3.2 Diagnosing Lexical Category	40
3.2.1 Reduplicating full bare stems	40
3.2.2 Lexical Category and the Content of Roots	46
4 Deriving Verbs	48
4.1 The Actor Voice Prefix as a Derivational Morpheme	49
4.1.1 Intransitive Bases	52
4.1.2 Transitive Bases	64
4.1.3 Nominal Bases	67
4.1.4 Interim Summary	70
4.2 The suffixes <i>-i</i> and <i>-ake</i>	71
4.2.1 Explaining the distribution of <i>-ake</i>	71
4.2.2 Explaining the distribution of <i>-i</i>	77

4.3	What about valence decreasing?	80
5	Deriving Nouns	82
5.1	The suffix -an	82
5.1.1	Verbal (and adjectival) Bases	83
5.1.2	Nominal Bases	87
5.1.3	A note about other nominalizing strategies	89
6	Concluding Remarks	91

Chapter 1

Introduction

Javanese possesses a small but highly flexible inventory of derivational morphemes. These morphemes are flexible in both their apparent range of functions and in terms of what types of bases they can attach to. Uhlenbeck remarked on this feature stating that there are “fairly subtle ways in which a few recurrent morphological devices may combine in order to characterize a variety of categories” (Uhlenbeck 1978 p.159). The present investigation seeks to understand the relationship between a given base and its resulting range of potential derivatives, a relationship which can become highly obscured in a system with these features.

In terms of methodology, the data presented largely represents the grammar of a single speaker, as collected through original consultant work conducted by me between March 2020 and August 2023. Some additional insight is provided by a Javanese language teacher who I have been working with independently since November of 2022. These sources are differentiated where relevant, and any additional data from the literature is also marked as such. Javanese is a widely spoken but severely understudied language of Indonesia, it is a member of the Malayo-Polynesian branch of the Austronesian language family.

My consultant speaks the Central dialect which is considered to be the closest dialect to “Standard” Javanese, and she is originally from the city of Surakarta (Solo), but currently resides in Massachusetts. She is also fluent in Indonesian. All data is representative of the *Ngoko* register of Javanese. This is the most casual register. The other two main registers, *Krama Madya* and *Krama Inggil* have almost entirely different lexical inventories and it remains to be seen the extent to which the grammatical generalizations found here transfer to these registers.

1.1 Structure of the dissertation

In Chapter 2 I will begin by presenting an overview of the empirical landscape and defining the core goals of the investigation, in the form of three tasks:

1. Determine the characteristics of the stems we find: What’s possible?

2. Delineate the contributions made by the morphemes present on those possible stems (eg. categorizing, argument introducing, etc): What changed?
3. Given this relationship, determine the syntactic position and selectional properties of the morpheme in question: What is a potential candidate to undergo this change?

This chapter will highlight the range of derivatives we see, addressing Task 1 from the list above, and will summarize the key problems these data present. Chapter 3 will begin to address Task 2 through the development of Javanese internal diagnostics for properties of lexical semantics and lexical category, defining the nature of the base to further derivation. Chapter 4 will then examine overt verbalizing and valence increasing morphology, and analyses will be proposed for the Actor Voice prefix, and the suffixes *-ake* and *-i*. Chapter 5 will examine the nominalizing suffix *-an* and provide a starting point for future work on other processes of nominalization. Chapter 6 concludes.

Chapter 2

The Empirical Problem

In Javanese both things like *sapu* ‘broom’ which are arguably canonical nouns, and things like *tiba* ‘fall’ which are arguably canonical verbs, can co-occur with the same broad range of derivational morphology:

- | | |
|---|--|
| (1) <i>sapu</i> ‘broom’ | (2) <i>tiba</i> ‘fall’ |
| a. Esti duwe sapu
Esti have broom
‘Esti has a broom’ | a. Esti tiba
Esti fall
‘Esti falls’ |
| b. Esti duwe sapu-sapu
Esti have RED-broom
‘Esti has many brooms’ | b. Esti tiba-tiba (terus)
Esti RED-fall (conti-
nous)
‘Esti falls (repeatedly)’ |
| c. Iki sapo-nan-e Esti
This broom-an-3.poss Esti
‘This is Esti’s sweeping’ | c. Banyu-ne tiba-n
water-DEF fall-an
‘Rain water’ (lit: the
fallen water) |
| d. Esti ny-apu
Esti AV-broom
‘Esti sweeps’ | d. Esti n-iba
Esti AV-fall
‘Esti falls’ (agentive) |
| e. Esti ny-apo-ake omah-e kanggo
Tuti
Esti AV-broom-ake house-DEF for
Tuti
‘Esti sweeps the house for Tuti’ | e. Esti n-iba-ake buku
Esti AV-fall-ake book
‘Esti drops a book’ |
| f. Esti ny-apo-ni jobin-e
Esti AV-broom-i floor-DEF
‘Esti sweeps the floor’ | f. Esti n-iba-ni Budi
Esti AV-fall-i Budi
‘Esti falls on Budi’ |

It is a defining feature of the Javanese derivational system, that the morphemes involved in these derivational processes appear to occur on stems of all kinds. Beyond this, the relationship between the starting stem and the resulting derivative is not surface apparent. The same combination of morphemes can appear to derive an intransitive in one case as in (2d)

above, and a transitive in another as in (3b):

- (3) a. Gelas-e pecah
glass-DEF break
'The glass breaks'
b. Budi m-[p]ecah gelas-e
Budi AV-break glass-e
'Budi breaks the glass'

Alternatively, the appearance of a morpheme that in some places derives a ditransitive as in (4b), could appear to cause no change to the valence of a given stem as in (5):

- (4) a. Budi ny-[s]ilih buku
Budi AV-borrow book
'Budi borrows a book'
b. Budi ny-[s]ilih-ake Esti book
Budi AV-borrow-ake Esti book
'Budi lends Esti a book'
- (5) Budi m-[p]ecah-(ake) gelas-e
Budi AV-break-ake glass-DEF
'Budi breaks the glass'

In general, when looking at an inventory of derivational morphemes and how they relate to the broader lexicon, implicational relationships should arise between lexical items and their derivatives. Though there are bound to be gaps and idiosyncrasies, the hope is that the observed behavior of a derived stem will follow more or less predictably from facts about the base, and the syntactic position of the additional functional elements that make it up. Furthermore, the selectional properties of those functional elements should serve to limit the set of possible derived stems in predictable ways. This way of characterizing word building, presents an investigation of derivational morphology for a given language with the following core tasks:

1. Determine the characteristics of the stems we find: What's possible?
2. Delineate the contributions made by the morphemes present on those possible stems (eg. categorizing, argument introducing, etc): What changed?
3. Given this relationship, determine the syntactic position and selectional properties of the morpheme in question: What is a potential candidate to undergo this change?

The purpose of this Chapter is to provide an overview of the empirical space, or the range of possible derived stems (Task 1). Given the complexity inherent to the system, this will involve some observation of what has changed (Task 2), insofar as we will observe cases within a given derivational paradigm where the presence or absence of an overt morpheme

impacts the valence of the resulting stem. As will be discussed further below, an understanding of the nature of bare stems in this language is not a trivial task, and so the focus of this discussion will center on the observable characteristics of resulting derived stems and eschew discussion of how that stem got to be that way. The latter topic is discussed in detail in subsequent chapters.

I will highlight the key problems posed by the empirical facts, namely those posed by the range of possible derivatives we can observe. Then I will discuss how these problems relate to the broader goals encompassed in Tasks 2 and 3. These goals being to understand the contribution made by each morpheme and subsequently the structural position and selectional properties of those morphemes. I will be focusing on stems affixed with derivational morphemes which interact with processes of categorizing and valence altering, to the exclusion of morphology that only ever modifies the meaning of a given stem without affecting category such as superlative or excessive morphology. To start, I will divide the inventory of stems into two broad groups, verbalizing and nominalizing. Again, these labels refer to the result of affixation and do not necessarily reflect the starting status of a given stem. The goal here is to delimit the range of possible stems and subsequently the observable range of morphological patterns, so that we can then move on to determining how those stems were actually arrived at in a meaningful way.

2.1 Task 1: The Range of Possible Derivatives

2.1.1 Verbalizing Strategies

There are at least five verbalizing strategies present in Javanese:

1. zero derivation : $\sqrt{\quad}$ - \emptyset
2. voice prefixation: ng-/di- $\sqrt{\quad}$
3. voice prefixation and suffixation with *-ake* : ng-/di- $\sqrt{\quad}$ -ake
4. voice prefixation and suffixation with *-i* : ng-/di- $\sqrt{\quad}$ -i
5. full reduplication : RED- $\sqrt{\quad}$

Zero derivation is included above in order to indicate the fact that verbs of all valences can occur grammatically as bare stems. In some cases, stems which are grammatical as verbs when bare, can optionally host verbalizing morphology *without* a change in valence¹. I have used the Actor Voice prefix in the following examples for simplicity and because it occurs with verbs of all valences. Given the nature of the Object Voice this prefix does not occur with intransitives due to the fact that it requires an object. The OV is largely

¹It is not to say that there is no potential change in meaning, but I am putting aside this nuance for the moment. It will be discussed in detail in Chapter 4.

interchangeable with the AV in transitive and ditransitive contexts however. The intricacies of its distribution are discussed further in Chapter 4, for our present purpose it is sufficient to track the availability of voice prefixation on the whole.

- (6) *intransitive*
 Esti n-[t]iba/tiba
 Esti AV-fall/fall
 ‘Esti fell/falls’
- (7) *transitive*
 Esti n-[t]uku/tuku buku
 Esti AV-buy/buy book
 ‘Esti bought/buys a book’
- (8) *ditransitive*
 Esti ng-[k]irim/kirim Budi surat
 Esti AV-send/send Budi letter
 ‘Esti sent/sends Budi a letter’

This optionality is not always present, in that some bare stems do not permit prefixation with voice morphology alone:

- (9) *intransitive*
 Banyu-ne umob/*ng-umob
 water-DEF boil
 ‘The water boiled/boils’
- (10) *transitive*
 Aku tilik/*n-[t]ilik ibu-ku
 1.sing visit mother-1.poss
 ‘I visited/visit my mother’

Furthermore, we can observe examples of verbs surfacing with obligatory voice prefixation as well. In these cases the bare stems are not grammatical as verbs on their own:

- (11) *intransitive* Budi ng-guyu/*guyu
 Budi AV-laugh
 ‘Budi laughed/laughs’
- (12) *transitive* Budi n-[t]ulis/*tulis surat
 Budi AV-write letter
 ‘Budi wrote/writes a letter’
- (13) *ditransitive* Budi ng-adol/*adol Tono buku
 Budi AV-sell Tono book
 ‘Budi sells Tono a book’

For some intransitive stems which have grammatical bare forms, the addition of voice prefixation appears to correlate with an increase in the valence of the stem by 1:

- (14) *intransitive* → *transitive*
- a. gelas-e pecah
glass-DEF break
'The glass broke/breaks'
 - b. Tuti m-[p]ecah gelas
Tuti AV-break glass
'Tuti broke/breaks the glass'

So to briefly summarize what we have seen so far, zero derived stems are possible for all valences. For such stems, the addition of voice prefixation can have no effect on valence, or it can be ungrammatical. In the case of some intransitives it correlates with the derivation of a transitive. We never see this prefix appear to derive a ditransitive from a transitive. We have also observed that for some stems which cannot occur grammatically when bare, voice prefixation can yield a grammatical verb. This is further summarized in the table below :

	-∅	AV/OV prefix
√	*	intransitive, transitive, ditransitive
	intransitive	intransitive, transitive, *
	transitive	transitive, *
	ditransitive	ditransitive

The table above leaves open the starting status of a given root, and indicates what sorts of verbs we can observe as bare stems. It then indicates the range of possible stems we can observe with overt voice prefixation for each type of bare stem observed, including stems which are ungrammatical when bare.

The remaining three verbalizing strategies all result in stems which are at least transitive. The presence of one transitivizing pattern in a verb's paradigm does not necessarily rule out the presence of another:

- (15) a. Banyu-ne umob/*ng-umob
water-DEF boil
'The water boiled/boils'
- b. Parta ng-umob-ake/-i banyu
Parta AV-boil-ake/-i water
'Parta boiled/boils water'
- (16) a. Omah-e resik
house-DEF clean
'The house was/is clean'
- b. Aku resik-resik omah-ku
1.sing RED-clean house-1.poss
'I cleaned/clean my house'
- c. Aku ng-resik-*(i) omah-ku

- 1.sing AV-clean-i house-1.poss
 ‘I cleaned/clean my house’
- (17) a. Buku-ne rusak
 book-DEF ruin
 ‘The book was/is ruined’
- b. Esti ng-rusak-(ake) buku-ne
 Esti AV-ruin-ake book-DEF
 ‘Esti ruined/ruins the book’

In (15) above suffixation with either *-ake* or *-i* yields a grammatical transitive. Similarly in (16) both reduplication and suffixation with *-i* are available. Finally, either the voice prefix alone or the voice prefix plus the *-ake* suffix are acceptable transitives in (17).

The suffixes *-ake* and *-i* are also present in the derivation of ditransitives. This occurs most commonly in conjunction with stems that themselves can surface grammatically as transitives as in (18), but not always as can be seen in (19) below. In the latter example the verb meaning ‘to give’ requires suffixation with either suffix to surface grammatically:

- (18) a. Parta ny-[s]ilih buku
 Parta AV-borrow book
 ‘Parta borrowed/borrows a book’
- b. Parta ny-[s]ilih-ake Esti buku
 Parta AV-borrow-ake Esti book
 ‘Parta lent/lends Esti a book’
- (19) Budi m-[w]eneh-*(ake)/*(i) Esti oleh-oleh
 Budi AV-give-ake/i Esti gift
 ‘Budi gave/gives Esti a gift’

Finally, the suffixes *-ake* and *-i* can occur optionally on some stems which otherwise can surface grammatically as ditransitive:

- (20) Esti ng-[k]irim-(ake/i) Budi surat
 Esti AV-send-ake/i Budi letter
 ‘Esti sent/sends Budi a letter’

The ditransitive examples above are all presented with a double object or applicative word order. Moreover, they can all be construed as *low* or change of possession type applicatives (Pylkkänen 2002), specifically in the direction of subject to recipient. It is generally the case that this word order is only possible for my speaker² in such change of possession cases. Despite this syntactic restriction, the *-ake* suffix is still associated with the relationship of benefaction and this relationship can still be instantiated when a syntactic applicative or double object configuration is not possible as in (21b) below:

- (21) a. *Parta m-[w]aca-ake Budi buku

²High applicatives are reported to be a syntactic possibility elsewhere in the literature, see Nurhayani (2014)

- Parta AV-read-ake Budi book
 Intended: Parta reads Budi a book
- b. Parta m-[w]aca-ake buku-ne kanggo Budi
 Parta AV-read-ake book-DEF for Budi
 ‘Parta reads the book for Budi’ (benefactive)

The benefactive nature of the reading in (21b) can be observed under pro-drop wherein the presence of the suffix correlates with the indication that there is some unnamed benefactee. The idea for this test comes from Cole and Son (2004) for Indonesian *-kan*:

- (22) a. Parta m-[w]aca buku
 Parta AV-read book
 ‘Parta reads a book’
- b. Parta m-[w]aca-ake buku
 Parta AV-read-ake book
 ‘Parta reads a book for someone’ (benefactive)

I am using the term *ditransitive* here to encompass both true syntactic applicatives, and cases where the second argument is syntactically introduced by other means but still thematically considered a core member of the verb’s argument structure as in (22b) above.

To briefly summarize these additional data. We have now seen 4 strategies used to derive transitive stems from intransitive stems, the addition of the voice prefix alone, the addition of the voice prefix plus the suffix *-ake* or *-i*, and full reduplication. We have also seen that the suffixes *-ake* and *-i* can derive ditransitives, most typically from existing transitive stems, however sometimes they can appear optionally or in cases where a grammatical transitive counterpart does not exist. Again these facts are further illustrated by use of a table. There are cases where suffixation with *-ake* or *-i* is not available, however the table indicates what the resulting valence would be in a case where these morphemes are permitted to surface.

	-∅	AV prefix	-ake/-i	RED
√	*	intransitive	transitive	-
		transitive	ditransitive	-
	intransitive	intransitive	transitive	transitive
		transitive	transitive	-
		*	transitive	-
	transitive	transitive	ditransitive	-
	ditransitive	ditransitive	ditransitive	-

As with the previous table, the starting status of the root is left open here. The columns which indicate prefixation or suffixation are organized to show the implicational relationships between stem types, so for example if a stem exists as a transitive when bare, that same stem when prefixed with only voice morphology is never going to surface as ditransitive. For space I have only indicated the suffixes themselves in the 4th column, but it should be noted that these suffixes never surface on their own. They only appear in conjunction with

a voice prefix. The bulk of the column indicating full reduplication of the bare stem is not used here. This is not because reduplication is impossible in these cases, but because the presence of reduplication indicates pluractionality or plurality here, and when reduplication is serving this function it never changes the category or valence of the base.

Problems presented by this data

Based off of these data, two core problems emerge:

1. What determines when a given strategy is unavailable? : the gaps
2. When a given strategy is available, what determines what the output will be? : the redundancies

Problem 1 can be illustrated by examining examples like (9) in conjunction with those like in (15), each are reproduced below for ease:

- (23)
- a. Banyu-ne umob/*ng-umob
water-DEF boil
'The water boiled/boils'
 - b. Esti ng-umob-*(ake/i) banyu-ne
Esti AV-boil-ake water-DEF
'Esti boiled/boils the water'

In (23) above, the intransitive form of *umob* 'boil' must be bare, and the transitive form must contain both a voice prefix and one of the suffixes. This is to the exclusion of any grammatical *ngumob* form. This is not an outlier either, there appears to be an entire class of stems which exhibit this pattern. It is a non-trivial question as to whether or not these are all accidental gaps or indicative of something deeper.

Problem 2 is two-fold. One aspect of it can be illustrated by the fact that within the paradigm of a given verb, multiple strategies to achieve a stem of the same valence exist (i.e. redundancies), as was seen in examples (15-17). This involves further complications however if we examine a case of this more closely. It is evident that the suffixes *-ake* and *-i* are capable of deriving ditransitives from transitives containing a voice prefix. However, they do not do so in every case, as was seen in (17b), an equivalent of which is present in (24) below:

- (24)
- a. Esti m-[p]ecah gelas-e
Esti AV-break glass-DEF
'Esti broke/breaks the glass'
 - b. Esti m-[p]ecah-ake gelas-e
Esti AV-break-ake glass-DEF
'Esti broke/breaks the glass'

- c. *Esti m-[p]ecah-ake Budi gelas
 Esti AV-break-ake Budi glass
 Intended: Esti breaks Budi a glass

We can see from (24b-c) that the *-ake* derivative in this case *must* be transitive, as opposed to ditransitive, despite the existence of the form in (24a). We know from the examples in (21-22) that even when a syntactic ditransitive is not available due to the lack of a possession relation, the benefactive relationship can still be instantiated when a verb is suffixed with *-ake*. However, the example in (24b) is analogous to our pro-drop test from above, and it is clear even with ample situational context provided that this simply does not indicate the presence of a benefactee.

So, despite the fact that a ditransitive version of *mecahake* is in theory *possible* due to the broader availability of benefactive ditransitives and the existence of the form in (24a), it is not available. Instead we have two transitive forms and no morphologically derived ditransitive form, applicative or otherwise. So it appears as though, in cases such as these transitivity is preferred despite its apparent redundancy and di-transitivity is ruled out.

From this we can see not only that redundancy is possible, but it is preferred in certain contexts. Moreover, from the lack of ambiguity in (24b), homophony is also dispreferred. So Problem 2 involves not only the availability of redundancy but the determining of what factors necessitate it to the exclusion of other conceivable outcomes.

2.1.2 Nominalizing Strategies

Nominalization, like verbalization, encompasses multiple strategies. The four core strategies are summarized below:

1. zero derivation: $\sqrt{-\emptyset}$
2. suffixation with *-an*: $\sqrt{-an}$
3. possession: $\sqrt{-ku/-mu/-(n)e}$
4. reduplication plus suffixation with *-an*: RED- $\sqrt{-an}$

Zero derivation is again included to indicate that many nouns freely occur as bare stems:

- (25) *kucing* ‘cat’
 Esti duwe kucing
 Esti have cat
 ‘Esti has a cat’

The suffix *-an* can combine with a variety of stems, including ones which appear to be already nominal when bare. This suffix never co-occurs with voice morphology or either of the

verbalizing suffixes. Here are some examples indicative of the range of possible meanings we observe when this suffix is present:

- (26) *sepur* ‘train’ → *sepur-an* ‘a train ride’
- (27) *langit* ‘sky’ → *langit-an* ‘ceiling’
- (28) *pangan* ‘to eat’ → *pangan-an* ‘food’
- (29) *luwe* ‘hungry’ → *luwe-nan* ‘perpetually hungry person’

The third process occurs specifically with verbal stems, as is evidenced by the presence of voice morphology. The result is something gerund like:

- (30) *tandur* ‘to plant’
N-[t]andur-e Parta cepet
AV-plant-3.poss Parta quick
‘Parta’s planting is quick’

The process of reduplication plus the suffix *-an* occurs with some nouns to derive a toy or imitation version of the base noun:

- (31) *gunung* ‘mountain’ → *gunung-gunung-an* ‘toy/miniature mountain’ (like something used in puppetry)
- (32) *sepur* ‘train’ → *sepur-sepur-an* ‘toy train’

So, to briefly summarize what we have seen, two strategies appear to correlate with the derivation of nouns from pre-existing nouns. These are suffixation with *-an* as well as reduplication plus suffixation with *-an*. Two strategies appear to occur with non-nominal stems as well, *-an* suffixation again and the process of suffixation with possessive morphology. The latter of these strategies appears to create a stem that is gerund-like.

Problems presented by the data

It is already evident from the data in (26-29) that nominalizing morphology can occur with a wide variety of stems and the range of potential meanings is quite broad and difficult to predict. This is true for all the processes, apart from maybe possession which appears to be limited to verbal stems. This in and of itself poses a challenge. This picture is further complicated by the fact that these strategies do not always yield nouns³:

- (33) *sapu* ‘broom’ → *sapon-sapon-an* ‘sweeping at each other’ (reciprocal)

In the example above, we might expect the reduplicated and *an* suffixed form of *sapu* ‘broom’

³You might have been wondering why this process was included, since I said earlier I wouldn’t be looking at morphology that doesn’t ever correlate with a change of category or valence. This is why!

to yield a toy or imitation broom. However, this is a reciprocal verb. We can observe that indeed, the process of reduplication plus *-an* can produce reciprocal verbs from otherwise verbal bases as in (34):

(34) *kirim* ‘to send’ → *kirim-kirim-an* ‘send things to each other’ (recipricol)

It is true that there exists a verbal form of *sapu* ‘broom’ that we saw in (1), reproduced below. However, the availability of this form is not enough to explain the fact in (33):

(35) *Esti ny-[s]apu*
Esti AV-broom
 ‘Esti sweeps’

(36) *Esti ny-[s]epur*
Esti AV-train
 ‘Esti rides the train’

As we can see in (36) above, *sepur* ‘train’ also has a verbal form and yet it did yield a toy/imitation type nominal derivative when reduplicated and suffixed with *-an* in (32). So the presence of reciprocal sweep remains puzzling. As we saw with transitive vs. ditransitive derivatives above, there is a choice point where multiple derivational options appear possible given the broader paradigm a stem is contained within. In this case with *sapu* ‘broom’, the presence of verbal *nyapu* ‘sweep’ takes precedent. But why, and how?

2.2 Towards Task 2: Understanding what is actually *derived*

Below is a summary of all the morphological patterns we have seen so far and their respective ranges of resulting meanings:

Morphemes Present	Derivative Type
√-∅	intransitive (including states and adjectives) , transitive, ditransitive, nominal
AV/OV-√	intransitive, transitive, ditransitive
AV/OV-√-ake	transitive, ditransitive (benefactive)
AV/OV-√-i	transitive, ditransitive (locative)
RED-√	transitive, plural/pluractional w/o category or valence change
√-an	nominal (heterogenous meaning)
(AV/OV)-√-POSS	nominal (gerund-like)
RED-√-an	nominal (toy/imitation), reciprocal

This is in essence the answer to Task 1, it is the range of what is *possible*. The goal of Task 2 is reproduced below:

- Delineate the contributions made by the morphemes present on those possible stems (eg. categorizing, argument introducing, etc): What changed?

Though in the column on the right above we can see what the output of a given derivation roughly equates to, it is still not clear what exactly has *changed*. To put it another way, given the data we have seen, it remains unclear what the presence or absence of a given morpheme actually corresponds to syntactically within the context of a given derivation: what is actually *derived*. In a language like Javanese, with a small but highly flexible inventory of derivational affixes and very little other morphology, determining the contribution made by an individual morpheme (Task 2) poses a challenge.

One element of this that we have already touched on, is the prevalence of grammatical bare stems:

- *kucing* ‘cat’ : noun
- *biru* ‘blue’ : adjective (?)
- *alun* ‘slow’ : adverb (?)
- *tiba* ‘fall’ : intransitive verb
- *tuku* ‘buy’ : transitive verb
- *kirim* ‘send’ : ditransitive verb

Though it is not the case that every Javanese stem can surface grammatically with no overt morphology, bare stems can be observed for all stem types including non-verbal as well as verbal predicates of every valence. Moreover, we have seen that stems of all kinds appear compatible with the full inventory of derivational morphemes, with mixed results.

As is alluded to with the use of question marks above, the attribution of a lexical category to any bare stem is not trivial given the lack of inflectional morphology in this language. This is particularly true in the domain of intransitives, which broadly includes things that are at least intuitively more verbal like *tiba* ‘fall’, as well as properties and states which are less clearly categorizable out of hand, like *biru* ‘blue’ and *luwe* ‘hungry’. So when it comes to delimiting the contribution made by a given base, as the input to a morphological derivation, developing language internal diagnostics for lexical category (and the underlying composition of bare stems more broadly) is the first obstacle. Before we can confidently ascribe a syntactic contribution to a particular morpheme (the goal of Task 2 above) we have to clarify the derivational starting point. The development of language internal diagnostics for lexical category and relevant properties of lexical semantics, is the topic of Chapter 3.

Chapter 3

Understanding the Base

In the following sections I will provide a system for distinguishing nouns from intransitive predicates. I will also divide the broader class of intransitives into aspectual classes using Javanese internal Aktionsart diagnostics. This will ultimately help tease apart modifiers such as adjectives and adverbs from intransitive verbs, and will serve to further differentiate sub-types of intransitive verbs. These language internal diagnostics will serve to shed light on several patterns of reduplication which have been previously acknowledged in the literature, but not in a way that highlights the importance of lexical category or stem class to the extent provided here. Still, the larger aim of this chapter remains to elucidate the nature of the input to the derivational processes of verbalization, valence increasing, and nominalization. The hope is that a deepened understanding of the input will help us delineate the contribution made by each piece of the derivation more meaningfully.

3.1 Disentangling Intransitives

Ogloblin (2005, p.599) observes that in Javanese "[intransitive verbals] include adjectives with some derivational possibilities unusual in verbs...In general, however, adjectives and intransitive verbs are very close in syntactic behavior". Given the sparsity of inflectional morphology, the lack of tense, and the lack of an overt copula in Javanese, differentiating intransitive verbs from adjectival and adverbial predicates is no trivial task:

- (37) Esti tiba
Esti fall
'Esti fell/is fallen?'
- (38) Bocah-e kesel
child-DEF tired
'The child is tired'
- (39) Buku-ne biru
book-DEF blue
'The book is blue'

As we know from the starting observations guiding this investigation, we cannot infer this information straightforwardly from the distribution of subsequent derivational morphology either. For example, the same morpheme *-ake* appears able to transitivize things like *tiba* ‘fall’, *kesel* ‘tired’, and *biru* ‘blue’, more or less indiscriminately. However, just as we have seen that the derivational picture becomes complicated once one looks across whole derivational paradigms, differences among these intransitive predicates do clearly arise once their broader behavior across contexts is considered.

The following subsections illustrate this through an examination of Aktionsart which attempts to situate these predicates into a system of aspectual classes. This is inspired by the aspectual classes outlined in such seminal works as Vendler (1967) and Dowty (1979), but adapted to serve the Javanese case. A full semantic analysis of all the auxiliary particles mentioned in this section is beyond the scope of the present investigation, but the empirical facts presented regarding how they interact with various stems, and the subsequent categorization schema developed will serve the present purpose of deepening our understanding of the input to morphological derivation. This schema along with the reduplication facts in section (3.2), will also serve to illustrate that despite the aforementioned Javanese specific factors that often blur the lines between lexical categories, systematic differences of this kind do indeed exist.

3.1.1 *lagi* : ‘right now’

The particle *lagi*¹ is typically considered to be a marker of progressive aspect (Horne 1961, Uhlenbeck 1978, Robson 1992 etc.), and in elicitation it is often translated by my consultant with the phrase ‘right now’. When looking at simple and clearly active/eventive clauses, this seems straightforwardly borne out:

- (40) Esti lagi m-[p]layu
 Esti *lagi* AV-run
 ‘Esti is running’
- (41) Esti lagi m-[p]angan pelem-e
 Esti *lagi* AV-eat mango-DEF
 ‘Esti is eating the mango’

In (40) above we can see an example of *lagi* with the intransitive *mlayu* ‘to run’, and in (41) *lagi* is co-occurring with the transitive *mangan* ‘to eat’. In the literature on Aktionsart

¹Following the summary in Vander Klok (2012, p.35), across dialects the following other forms are found:

- *gek* Tengger Javanese (Connors 2008), Peranakan Javanese (Cole et al. 2008), Paciran Javanese (Vander Klok 2012)
- *lage* Tegal Javanese (Suwadji 1981)
- *lagek* Paciran Javanese (Vander Klok 2012)

(Vendler 1957 et seq.), these verbs (in these usages) are traditionally classed as activities and accomplishments respectively. Their compatibility with the progressive is attributed to both predicates being durative as well as dynamic. Durative refers to duration or extension in time, while dynamic refers to the presence of some sort of change. Both properties must hold, which sets these types of verbs apart from states and achievements. States are durative but they are not dynamic, while achievements are dynamic but lack duration.

	Durative	Dynamic
<i>activity</i>	✓	✓
<i>accomplishment</i>	✓	✓
<i>state</i>	✓	X
<i>achievement</i>	X	✓

If we take these basic properties as a starting point, it becomes clear that *lagi* does not behave how we would expect. Notably this particle is compatible with predicates that denote states:

(42) Budi lagi luwe
 Budi *lagi* hungry
 ‘Budi is hungry (right now)’

(43) Bocah-e lagi gedhe
 child-DEf *lagi* big
 ‘The child is big’ (at stage of childhood where they are a relatively big kid)

The examples above are durative but clearly not dynamic, at least not in the sense of dynamic found above that referred to a change of some kind being underway. In (42) it is not the case that Budi is becoming hungry, and likewise in (43) it is not the case that the child is currently growing or becoming larger. The presence of *lagi* here does not indicate that a change is taking place or in progress. These data indicate that the subject is currently in a given state, namely the state of hunger or the state of being a larger child. We can however contrast the example in (43) with (44-45) below:

(44) #Buku-ne lagi gedhe
 book-DEF *lagi* big
 int: ‘The book is big’

(45) #Wit-e lagi gedhe
 tree-DEf *lagi* big
 ‘The tree is big’

The particle *lagi* is only possible with properties like *gedhe* ‘big’ when they are stage level (in the sense of Carlson 1977), or indicative of a temporary/transient state of being. This means that an object with a canonically fixed size is incompatible with *lagi*. Similarly even objects that grow are impossible as in (45) given that once a tree is big it is always a big tree. What makes *lagi* work with ‘child’ in (43) is the fact that a child will pass through the phase of being a big kid, and move on to other stages irrespective of their physical size. A

tree should be compatible with *lagi* and *cilik* ‘small’ because this is (barring intervention) a transient state of being for a tree, and indeed this is what we find:

- (46) Wit-e lagi cilik
 tree-DEF *lagi* small
 ‘The tree is small’ (right now)

What differentiates the individual level vs. stage level uses of these predicates is in essence the possibility that the property in question will no longer hold at some point in the future. It appears as though, despite the fact that in the felicitous cases with *lagi*, no change is asserted to be underway or in progress, some change needs to at least be possible. This notion holds when we look at predicates which denote the result of a change of state like those below. In these cases *lagi* asserts that the result of the change of state holds ‘right now’, not that the change in question is currently happening. Subsequently what differentiates the felicitous cases from the infelicitous cases below, is whether or not the result of the change of state is reversible:

- (47) a. Buku-ne (#lagi) rusak
 book-DEF *lagi* ruin
 ‘The book is ruined’ (* with lagi)
 b. Mobil-e lagi rusak
 car-DEF *lagi* dysfunctional
 ‘The car is broken down/not running (right now)’

In (47a) when *rusak* is used in conjunction with a non-mechanical object like a book the meaning is something like ‘ruined’, or irreparably damaged, unreadable. This result state, like with the tree above and *gedhe*, is irreversible and akin to an individual level property. On the other hand in (47b), the interpretation given to *rusak* when occurring with a complex mechanical object such as a car is one in which the vehicle is in a state of disrepair rendering it dysfunctional, but it is understood that this is a reversible state. The car is fixable and subsequently *lagi* is available.

To summarize, in these cases of ruining and disrepair, *lagi* does not indicate that some process is taking place, it indicates that a state is holding at a given time. Moreover it indicates that the state in question will no longer hold at a point in time in the future. This taken with the individual vs. stage level facts given above, calls for a revision to our original notion of *dynamic*. We have seen that *lagi* is compatible with stative predicates, either those denoting properties or result states, so long as the state is *subject to change*:

- Revised definition of ‘dynamic’ applicable to states : A predicate is considered dynamic if it denotes a state which is subject to change. It should be conceivable that the state denoted by the predicate will no longer hold at some point in the future.

As we have noted, it appears as though the bare stem *rusak* necessarily indicates the end state of ruining or disrepair as opposed to the process of becoming ruined or dysfunctional.

Other strategies are used to refer to a point in the process of the ruining, which will be discussed in later subsections. We can contrast the pattern above with a stem that does not encode an end state, *tiba* ‘to fall’ :

- (48) #Esti lagi tiba
 Esti *lagi* fall
 Int: ‘Esti is falling’ / ‘Esti is fallen’

In (48) above the predicate *tiba* does not provide a durative but transient end state for *lagi* to apply to such as ‘fallen’. This is not an activity either such as *mlayu* ‘to run’, but instead appears more like an achievement predicate which is considered dynamic but not durative. There is no interval to pick out to coincide with the notion of ‘right now’. Even when presented with scenarios of prolonged falling, speakers cannot use *lagi* to describe the interval during which an object is in the air. For example when describing the scene in Alice in Wonderland by Lewis Carroll in which Alice is falling down the rabbit hole, to describe this in Javanese my consultant used the future oriented temporal auxiliary *meh* ‘almost/about to’ to describe the fall.



“meh tiba,
 meh tiba,
 meh tiba...”

This type of predicate can be made durative, and subsequently compatible with *lagi*, through pluralization of the subject or the verb itself:

- (49) a. Udan-e lagi tiba
 rain-DEF *lagi* fall
 ‘The rain is falling’
 b. Wong-wong-e lagi tiba
 people-RED-DEF *lagi* fall
 ‘The people are falling’ (intermittently)
 c. Esti lagi tiba-tiba
 Esti *lagi* fall-RED
 ‘Esti is falling over and over’

In (49a-c) above the rain and the plural people can be said to be falling with *lagi* because this indicates many individual events of falling spanning an extended period of time. The particle *lagi* can be used to refer to the entire interval containing the separate instances of falling, which fulfills the need for duration. Similarly when Esti is falling pluractionally, with the reduplicated form of the verb *tiba-tiba*, the interval of these repeated falls can satisfy the need for duration. Reduplication will be discussed in more detail in the following section, for the present moment it is enough to observe that pluralization serves to make *lagi* available with this class of predicates where it otherwise was not. We can contrast this with result state predicates which are not reversible, or which fail to meet our revised dynamic requirement:

- (50) a. #Gelas-e lagi pecah glass-DEF *lagi* break
 Int: The glass is breaking
 b. #Gelas-gelas-e lagi pecah
 glass-RED-DEF *lagi* break
 Int: The glasses are breaking
 c. #Gelas-e lagi pecah-pecah
 glass-DEF *lagi* break-RED
 Int: The glasses are breaking (iterative?)

As can be seen in (50a-c), pluralizing the subject or attempting to pluralize the verb through reduplication fails to fix these. This follows from the notion that what matters here is the dynamic status of the end state of breaking, which remains unchanged.

The table below summarizes what we have seen so far. The auxiliary *lagi*, which is typically considered a marker of progressive aspect, in practice can serve as a means of diagnosing whether or not a predicate is durative. Beyond this, if we draw the initial distinction among intransitives between activity predicates like ‘run’ and state denoting predicates, *lagi* can be further used to probe whether or not the property or result state denoted by that predicate is inherent/permanent or subject to change. We have also observed that *lagi* does not refer to a change-of-state in progress or a point in which such a process is underway, but instead can be used to indicate that the result of that change or process holds ‘right now’.

		durative	dynamic	<i>lagi</i>
<i>activity</i>		✓	✓	✓
<i>accomplishment</i>		✓	✓	✓
<i>property state</i>	individual level	✓	X	X
	stage level	✓	✓	✓
<i>result state</i>	irreversible	✓	X	X
	reversible	✓	✓	✓
<i>achievement</i>		X	✓	X

There is an added layer of complexity with respect to the *lagi* facts which should be addressed. Namely, there is a set of predicates which on the surface we might expect to

behave as *gedhe* did above, but which are always unavailable with *lagi*:

- (51) a. #Langit-e lagi biru
sky-DEF *lagi* blue
Int: The sky is blue (right now)
b. #Omah-e lagi biru
house-DEF *lagi* blue
Int: The house is blue (right now)
- (52) #Tas-e lagi abot
Bag-DEF *lagi* heavy
Int: The bag is heavy (right now)
- (53) #Lempung-e lagi bunder
clay-DEF *lagi* round
Int: ‘The clay is round (right now)’

We can see from (51) that even when being blue is a mutable property, in that the sky is not *always* blue, we cannot use *lagi*. Similarly we cannot use *lagi* to indicate that the bag in (52) is only heavy right now, given a context where it is full of many books for example, nor can we use it to indicate that the clay in (53) is currently round, even in a context where a potter is planning to make it a different shape. So, given these facts it would appear that these stems which largely denote physical properties are simply impossible with *lagi* even if we construe a context in which the state of having that property should be temporary.

This makes the previous data with *gedhe* somewhat puzzling. Though upon reflection, the stage level version of *gedhe* is only available with the particular context of childhood, and so I think it is possible that the meaning in that instance is more complex than simply denoting physical size. Similarly *cilik* was possible in the context of a tree, which is again something that changes size of its own accord. We might be learning that when the property denoted by the predicate is mutable in a way that is *internally caused*, in the sense of Levin and Rappaport Hovav (1995), this can be possible with *lagi*, making it more akin to a result state. By contrast, something like *bunder* ‘round’ in the clay example, though mutable does not have this quality and is subsequently not possible with *lagi*. It would be worth exploring if *lagi bunder* could be attributable to something which can make itself round, such as an armadillo for example. All of this is to say that it is not enough to have a conceivable stage level interpretation for predicates denoting physical properties, and where this appears possible I would argue there is actually an internally caused change-of-state element involved. This distinction will be elaborated on in the following subsection where *isih* ‘still’ is discussed.

3.1.2 *isih*: being and doing *still*

The particle *isih* denotes that a property *still* holds at a relevant time. The requirements for the presence of *isih* are very similar to the requirements for the presence of *lagi*, namely being durative and dynamic, though with some key differences. Still, we can see that *isih* is available in many of the same places where we saw *lagi*:

- (54) a. Esti isih m-[p]layu
 Esti still AV-run
 ‘Esti is still running’
 b. Esti isih m-[p]angan pelem-e
 Esti still AV-eat mango-DEF
 ‘Esti is still eating the mango’
 c. Esti isih luwe
 Esti still hungry
 ‘Esti is still hungry’

One important point of difference involves properties which are treated as individual level for the sake of *lagi* but can co-occur with *isih* all the same. These are specifically the predicates denoting physical attributes which are mutable, discussed at the end of the previous section:

- (55) a. Langit-e isih biru
 sky-DEF *still* blue
 ‘The sky is still blue’
 b. Omah-e isih biru
 house-DEF *still* blue
 ‘The house is still blue’
- (56) Tas-e isih abot
 Bag-DEF *still* heavy
 The bag is still heavy
- (57) Lempung-e isih bunder
 clay-DEF *still* round
 ‘The clay is still round’

Its important to point out that it is not the case that *isih* can co-occur with any individual level predicate:

- (58) #Budi isih dhuwur
 Budi still tall
 Int: ‘Budi is still tall’

It appears as though *isih* as with *lagi*, is still sensitive to some notion of being dynamic. Moreover, *isih* is also ruled out with result states which denote an irreversible change, as *lagi* was:

- (59) #Pitih-ne isih m-[p]ati
 chicken-DEF still AV-death
 Int: ‘This chicken is still dead’
- (60) Lampu-ne isih m-[p]ati
 lamp-DEF still AV-death
 ‘The lamp is still off’

When the verb *mati* ‘to die’ is used in conjunction with animate subjects it refers to the irreversible result state associated with natural death. However, when used with a subject such as a lamp, this verb refers to the lamp simply being in the state of being switched off. This shows again that beyond being sensitive to duration, *isih* cannot be used if the result state is not transient. Still given the facts presented regarding individual level properties such as color and shape, *isih* is less sensitive than *lagi*.

If a property is mutable, this potential for dynamic change is enough to allow the presence of *isih*, but it is not enough to allow the presence of *lagi*. If a result state is irreversible, both particles are ruled out. Furthermore, we can observe the following three way contrast with the predicate *campur* ‘mixed’:

- (61) a. Jus-e #lagi/#isih campur
 juice-DEF *lagi/still* mixed
 Int: ‘The juices are mixed (right now)’ / ‘The juices are mixed still’
- b. Gabah-e lan beras-e #lagi/✓isih campur
 unhulled.rice-DEF and hulled.rice-DEF *lagi/still* campur
 Int: ‘The unhulled and hulled rice are mixed (right now) ✓’ / ‘The unhulled and hulled rice are still mixed’
- c. Murid-murid-e lagi/isih campur
 student-RED-DEF lagi/isih mixed
 ‘The students are mixed (right now)’ / ‘The students are still mixed’

When mixing ingredients that could never conceivably be separated again, both *lagi* and *isih* are ruled out as predicted. When the mixing involves objects which could conceivably be sorted or separated again, *isih* becomes available in both cases as in (61b-c). What is noteworthy here is that *lagi* only becomes available in (61c) when the subject of *campur* is a group of students. This aligns with the observation made with respect to *gedhe* and *cilik* above. When the subject of the predicate in question is in some way responsible (through agency or otherwise) for the state described, *lagi* becomes available where it otherwise was not. As we saw with predicates which denote physical characteristics, when the subject is animate or capable of undergoing the relevant physical change of its own accord, *lagi* is available. This appears to hold for predicates which denote physical position or orientation as well. It appears as though beyond being stage level, predicates of this kind require some degree of internal causation in order to permit the presence of *lagi*.

Additionally there are cases where *lagi* is possible, but the relevant interval it picks out is so short that *isih* is ruled out:

- (62) Esti (lagi) teka
 Esti *lagi* arrive
 ‘Esti is arrived’ (recently arrived)
- (63) #Esti isih teka

Esti still arrive
Int: ‘Esti is still arrived’ ??

It is impossible to use *isih* in (63) because though the state of having arrived is durative enough to permit the presence of *lagi*, *isih* requires the continuation of a state previously held and the state of arrival provides too short of a window for this to be available. So *isih* has the added requirement beyond being durative, a state must be *extendable*, meaning it must extend far enough into the future from the point in which the state was entered to provide a window of time for *isih* to apply to.

As we may expect, *isih* is impossible in conjunction with achievements because these are not durative. Still, as with *lagi* above, pluralizing the subject or verb can make *isih* available in some contexts where it previously was not, namely with achievements (like *tiba* ‘fall’), as well as with things where the state fails to be extendable like *teka* ‘arrive’ above. The repair is given in (64) below. In these cases *isih* indicates that the state denoted by the predicate ‘keeps’ occurring, so for a given interval it continues to be true that people iteratively or intermittently reach the state of having arrived. It, as with *lagi*, does not mean that plural people collectively continue to be in the state of having arrived, as this would not repair the problem.

- (64) Wong-wong-e isih teka
people-RED-DEF keep arrive
‘The people keep arriving’

The need for predicates to be durative seems to follow intuitively from the notion of ‘still’ in that this implies a continuation of some kind, either of a state or an action. The requirement for the state to be dynamic or transient, is less obvious. It seems as though beyond meaning ‘still’ *isih* also carries with it the notion of ‘for now’. For example, something like ‘isih mati’ when talking about a living thing is subsequently odd because there is an implication that at some point the subject will no longer be dead.

Taking all of this into consideration a revised data summary is provided in the table below:

			lagi	isih
Activity	Dynamic	✓	OK	OK
	Durative	✓		
Accomplishment	Dynamic	✓	OK	OK
	Durative	✓		
Achievment	Dynamic	✓	*	*
	Durative	X		
Property State	Dynamic	X: individual	*	*
		✓: stage	OK	OK
		✓: mutable	*	OK
	Durative	✓	-	-
Result State	Dynamic	X: irreversible	*	*
		✓: reversible	OK	OK
	Durative	✓: extendable	OK	OK
		✓: non-extendable	OK	*

I have added a line for ‘mutable’ property states to indicate items like *biru* which are always okay with *isih*. It should be understood that these stems when in contexts where the property is internally caused, pattern with result states. Subsequently the availability of *lagi* and *isih* becomes contingent on whether or not that result state is transient/reversible or not.

3.1.3 *rada*: halfway and kind of doing/being

The measure/degree term *rada* is typically glossed as ‘somewhat’ or ‘quite’ (Robson and Wibosono 2002). However with respect to my consultant it more accurately denotes either a quantitative or qualitative halfway point with respect to some relevant scale. For this reason I will be glossing this term as ‘half’ for ease. Examples of each type are given in (65-68) below:

- (65) Esti rada wareg
 Esti half full
 ‘Esti is half full’ (halfway to being full)
- (66) Banyu-ne rada umob
 water-DEF half boil
 ‘The water is half boiled’ (halfway to being boiled)
- (67) Esti rada luwe
 Esti half hungry
 ‘Esti is half hungry’ (kind of hungry, kind of not)
- (68) Bakal-e rada biru
 fabric-DEF half blue
 ‘The fabric is half blue’ (kind of blue, kind of not)

While the particles *lagi* and *isih* required the predicates they co-occurred with to be both

dynamic and durative, *rada* does not require a predicate to be dynamic. This means that *rada* can combine with individual level predicates as well as irreversible changes of state:

- (69) Esti rada dhuwur
 Esti half tall
 ‘Esti is half tall’ (kind of tall, kind of not)
- (70) Kertas-e rada suwek
 paper-DEF half torn
 ‘The paper is half torn’ (torn halfway/partially)

What *rada* does require is the predicate in question to be either gradable or incremental. In (71) below we can see that *mati* in its reversible usage as ‘off’ is a state that is not gradable, either something is or it is not off, and subsequently *rada* is infelicitous here. We can contrast this with the ‘dead’ usage of *mati* which is a result state that can be arrived at incrementally. Similarly, the change of state involved to give the result state *kandhas* ‘sunken’ is not incremental enough to permit the presence of *rada*. This cannot be used to indicate that something is halfway or partially sunken.

- (71)
- a. #Lampu-ne rada m-[p]ati
 lamp-DEF half AV-death
 Int: ‘The lamp is half off’ (partially off?)
- b. ✓ Pitih-e rada m-[p]ati
 chicken-DEF half AV-death
 ‘The chicken is half dead’ (kind of dead/kind of not)
- (72) #Watu-ne rada kandhas
 stone-DEF half sunken
 Int: ‘The stone is half sunk’ (sunken halfway/partially)

Given this requirement for gradability or incrementalness, *rada* is also ruled out with achievement predicates. An achievement lacks any duration and subsequently cannot be referred to in any meaningful way as ‘half’ complete. In (73) below, Esti either fell or she did not fall, there is no ‘kind of’ falling or ‘half’ falling:

- (73) Esti (#rada) tiba
 Esti half fall
 Int: ‘Esti half/kind of fell’

So far we have seen that for *rada* to be possible in combination with a given predicate, that predicate must be durative. Beyond this, the predicate must be in some way gradable or incremental. Gradability most readily aligns with property states like *luwe* ‘hungry’ or *adhem* ‘cold’, whereas incrementality applies to result states like *suwek* ‘torn’. Achievement predicates are ruled out given that they lack duration.

Turning to activities and accomplishments, these are both durative and subsequently we might expect such predicates to be compatible with *rada*. Activity predicates like *mlayu* ‘run’ can be construed as gradable with respect to a canonical running event, and so to ‘half run’ would be to run in some non-canonical way or to ‘sort of’ run. This is akin to the Evaluative usage of ‘half’ presented in Bochnak (2010). Similarly, accomplishments like *mangan peleme* ‘eat the mango’ could be construed as incremental in that ‘half’ could refer to quantitative progress towards whatever goal is indicated by the verb and its relation to the direct object (in this case eating the mango). This is Bochnak’s Eventive usage of ‘half’. However, in Javanese *rada* is impossible in these contexts:

- (74) Esti (#rada) m-[p]layu
 Esti half AV-run
 Int: Esti half/sort of runs/ran
- (75) Esti (#rada) m-[p]angan pelem-e
 Esti half AV-eat mango-DEF
 Int: Esti half ate the mango (ate half of it)

Though *rada* is ruled out with *mlayu* above, it is possible with *nguyu* ‘to laugh’:

- (76) Esti rada ng-guyu
 Esti half AV-laugh
 ‘Esti kind of laughed’ (laughs shyly, a little bit)

So it is possible that what *rada* is actually sensitive to is agentivity². This is further supported by the fact that it is possible to use *rada* with psych predicates like *pengin* ‘to desire’ and *bosen* ‘to tire of’:

- (77) a. Esti rada bosen m-[p]angan pelem
 Esti half tired.of AV-eat mango
 ‘Esti is half (kind of) tired of eating mango’
- b. Esti rada pengen m-[p]angan pelem
 Esti half desire AV-eat mango
 ‘Esti half (kind of) wants to eat mango’

Again taking into account these additional data, a summary of what we have seen so far is given in the following table. Given the incompatibility of *rada* with activities, accomplishments, and achievements these have been omitted for visual clarity:

²This is further supported by the fact that you can causativize *nguyu* ‘to laugh’ but not *nangis* ‘to cry’, and these facts pattern with *rada* in that as seen above it is available with the former, but it is not available with the latter: *Esti rada nangis Int: ‘Esti is kind of crying’. As will be discussed in detail in Chapter 4, morphological causatives of this type are only possible with non-agentive predicates.

			lagi	isih	rada
Property State	Non-dynamic	individual	*	*	-
	Dynamic	stage	OK	OK	-
		mutable	*	OK	-
	Durative	gradable	-	-	OK
non-gradable		-	-	*	
Result State	Non-dynamic	irreversible	*	*	-
	Dynamic	reversible	OK	OK	-
	Durative	extendable	OK	OK	-
		non-extendable	OK	*	-
		incremental	-	-	OK
		non-incremental	-	-	*

This table is designed to demonstrate how the availability of *lagi* and *isih* are closely bound to the property of ‘dynamic’, with the added complexity of extendability in time for *isih*. On the other hand, the availability of *rada* is dependent on the availability of a gradability or incrementality scale. It is also important to remember that *rada* is impacted in some way by agency.

3.1.4 *meh*: ‘almost’

The particle *meh* in Javanese has been labeled as a future marking auxiliary, glossed ‘will’, in the literature (Cole et al 2008, Vander Klok 2012). These discussions specifically pertain to the Peranakan dialect of Javanese. To my knowledge it has not been formally described for any other dialect. The major Javanese-English dictionaries as well as some pedagogical materials translate *meh* as ‘almost’ or ‘very nearly’ (Horne 1974, 1961 ; Robson and Wobisono 2002). The latter description most closely aligns with the usage of *meh* employed by my consultant so I will gloss this auxiliary as ‘almost’ for ease.

Dowty (1979) observed that in English, ‘almost’ gives rise to ambiguity when used in conjunction with accomplishments but not activities. The phrase ‘John almost walked’ can only mean that John did not walk, while ‘John almost painted the picture’ can either mean he never painted or he nearly completed it. This same duality of function can be observed with Javanese *meh*, in that there is a usage that indicates that something was about to occur or could have occurred but did not, as well as a usage that indicates something being close to culmination or completion. However, this is where the parallel to English ends. The usages of *meh* do not correlate with the activity/accomplishment distinction and the presence of *meh* does not give rise to ambiguity, either one meaning is possible or the other:

- (78) *activity*
 Budi meh mlaku
 Budi almost walk
 ‘Budi is/was about to walk’ (no walking has taken place)
- (79) *accomplishment*
 Budi meh m-[p]angan pelem

Budi almost eat mango

‘Budi is/was about to eat a mango’ (no eating has taken place)

(80) *achievement*

a. Esti meh wahing

Esti almost sneeze

‘Esti is about to sneeze’ (visual cue that the sneeze is coming)

b. Esti meh tiba

Esti almost fall

‘Esti is about to fall’ (visual evidence of ice in front of her)

In (78) above, with the activity *mlaku* ‘walk’, it is also the case in Javanese that no walking has taken place. Still, given the lack of any temporal information included in this example the result is ambiguous between a situation where Budi is on the precipice of walking or a situation where he was going to walk but something intervened (analogous to English example above with overt past tense). Either way, as in English, Budi did not walk. Unlike in English, the accomplishment in (79) does not give rise to any ambiguity. This too only receives an interpretation where no eating has taken place.

Achievements also pattern with activities and accomplishments, though with one added caveat. Achievements are generally non-agentive as in the falling example in (80), subsequently there needs to be some contextually salient reason why we expect Esti to fall incipiently. This could be that we see she is not paying attention, or see that there is ice in the vicinity. For activities and accomplishments, the choice of *meh* over another future oriented temporal auxiliary like *arep* ‘will, going to’, should be motivated by the context as well. However, the sort of evidence required is more flexible. For example, if you know that Budi tends to walk in the evenings around five and it is currently 4:30, then you can say (78).

The split in the meaning of *meh* arises when we look at the class of states:

(81) *property state*

Budi meh luwe

Budi almost hungry

‘Budi is/was about to be hungry’ (not hungry yet)

(82) *result state*

Budi meh wareg

Budi almost full

‘Budi is/was almost full’ (already mostly full, just not completely)

States that do not involve some form of causation or change-of-state, like *luwe* ‘hungry’ in (81), pattern with activities, accomplishments, and achievements in that *meh* refers to a point where the state has not been entered but will be entered shortly. States that do involve some sort of result of causation or a change-of-state such as *wareg* ‘full’, by contrast are already in progress when *meh* is present. This is where *meh* gives rise to the ‘almost completed’ meaning.

This distinction is sharpened when we consider the amount of progress denoted by *meh* for a given predicate vs. that of *rada*. We know that where available, *rada* indicates some mid point on either a quantity or a quality scale. So in cases like *luwe*, we would expect *Budi rada luwe* to indicate a situation where Budi is more hungry than *Budi meh luwe*, which as stated above indicates that he is not hungry yet. By contrast *Budi rada wareg* should indicate that Budi is **less** full than in the case of *Budi meh wareg*. The former should align with a situation where Budi is kind of or half full, where the latter should indicate that he is very nearly completely full. This is borne out.

meh	rada	luwe ‘hungry’
	rada meh	wareg ‘full’

So despite both of these states being in some way gradable as indicated by their compatibility with *rada*, only the result state which has some defined point of culmination or maximum, is compatible with the ‘almost’ usage of *meh*.

There are some additional limitations on the presence of *meh* that arise due to the fact that it is future oriented. For example *meh* is impossible with something like ‘unripe’, because there is no point at which a piece of produce can be on the verge of being unripe, regardless of what type of state this is there is no previous point of any kind that *meh* can refer to. Similarly, in the cases where *meh* is asserting that a certain state will be held in the near future, this is difficult or odd when discussing things like emotional and mental states that are arguably difficult to predict in this way. In cases where *meh* gets an ‘almost completed/reached the end state’ interpretation, this is difficult in cases where that end state is difficult to pinpoint or involves something going out of existence.

3.1.5 *ke-√-en* : ‘too X’

The prefix *ke-* has been most commonly described in conjunction with the adversative or accidental passive (Nurhayani 2015, Robson 2014), and with or without the suffix *-an*. In those cases it correlates with valence decreasing, and this process is discussed in detail in Chapter 4. When this prefix occurs with the suffix *-en*, on the other hand, it forms an excessive as in (83) below:

- (83) *panas* ‘hot’
ke-panas-en
ke-hot-en
‘too hot’

This process never occurs with stems which are arguably nominal when bare such as *kucing* ‘cat’, **kekucingen*, and it appears to only occur on non-nominal predicates which are intransitive. Of the types of intransitives we have observed so far, this circumfix is compatible with most property states (like *panas* above), to the exclusion of those denoting emotions, as well as a small number of result states:

- (84) *sedhik* ‘sad’
 *ke-sedhik-en
 ke-sad-en
 int: ‘too sad’
- (85) *bosok* ‘rotten’
 ke-bosok-en
 ke-rotten-en
 ‘too rotten’

In general it appears as though this process is limited to result states which can be internally caused, so things like *bosok* ‘rotten’, as well as *mateng* ‘ripe’, *ajer* ‘melted’/‘liquid’, *garing* ‘dry’. This is to the exclusion of things like *umob* ‘boiling’ **kemoben*, or *suwek* ‘torn’ **kesuweken*. This process is also unavailable with achievements like *tiba* ‘fall’, and activities like *mlayu* ‘run’.

3.1.6 Interim Summary

So far we have seen 3 different auxiliary particles, a measure term, and a circumfix which together diagnose various lexical semantic properties for intransitive predicates. Those properties are:

- Dynamic
 - Individual vs. Stage
 - Irreversible vs. Reversible
- Durative
 - Gradable vs. Non-gradable
 - Incremental vs. Non-incremental
 - Extendable vs. Non-extendable
- Internal Causation
- External Causation
- Agentivity

We have diagnosed predicates as intransitive through their lack of compatibility with object voice morphology. Things which denote canonical activities such as *mlayu* ‘run’ are more readily labeled as verbal. These stems often require overt Actor Voice morphology, and are clearly agentive. Beyond stems such as these, the lines begin to blur. The inventory of diagnostic criteria summarized above allows us to divide this blurrier space in concrete ways. In the next section I will zoom out to discuss the broader task of lexical category distinctions, after laying a foundation through a discussion of reduplication, we will return to the notion of how to divide the space of intransitive predicates in lexical terms, insofar as this makes sense for Javanese and the task(s) at hand.

3.2 Diagnosing Lexical Category

3.2.1 Reduplicating full bare stems

The previous sections have touched on the process of full bare stem reduplication in passing. This process is fairly heterogenous but overarching generalizations do exist which can be helpful with respect to further dividing up the lexical space. First we can observe that bare nouns typically form plural nouns when reduplicated, and this can come with an implication of variety but this is not obligatory³:

(86) *buku* ‘book’ → *buku-buku* ‘books’

(87) *wit* ‘tree’ → *wit-wit* ‘trees’, ‘forest/wood’

This is a very pervasive process with nouns though there are cases where in addition to the plural noun meaning, a noun will have an idiomatic verbal usage:

(88) *omah* ‘house’

a. *omah-omah* ‘houses’

b. *omah-omah* ‘to be married’

Despite these idiosyncracies, full reduplication remains a good diagnostic for noun-hood. This process yields a plural even in cases where the noun in question is produced through suffixation:

(89) *Esti duwe* [pangan-an]-pangan-an

Esti have [RED]-eat-an

‘*Esti* has (various) foods’

Turning to the broad class of intransitives, full reduplication results in three potential meanings⁴:

1. distributive: the property denoted by the predicate holds for every member of a plural set of nouns denoted by the subject
2. continous: the state denoted by the predicate holds continuously for a given span of time
3. pluractional: the change of state or action denoted by the predicate occurs repeatedly

The distributive usage occurs with individual level property states, the class of ‘mutable’ property states which largely denote physical appearance/orientation, and a limited number

³As was discussed in Chapter 2, reduplication plus the *-an* suffix can also be observed in conjunction with a ‘variety’ meaning.

⁴The data associated with these meanings are discussed in pedagogical grammars such as Robson (2014) and overview articles like Wijana (2021) but they are not organized in this way.

of stage level property states, particularly those denoting mental or emotional states of being. This meaning does not occur with result state predicates, activities, or achievements.

- (90) Wong-e dhuwur-dhuwur
 people-DEF RED-tall
 ‘The people are tall’/‘The tall people’
- (91) Buku-ne biru-biru
 book-DEF RED-tall
 ‘The books are blue’ / ‘The blue books’
- (92) Bocah-e kesel-kesel
 child-DEF RED-tired
 ‘The children are tired’ /‘The tired children’

If we think of reduplication introducing some notion of plurality, it is clear that with predicates such as these that plurality is taking scope over the entire phrase. The subjects above are necessarily interpreted as plural here, despite the fact that the nouns themselves are unmarked for plurality and in theory ambiguous for number.

We saw in the previous section that many of the ‘mutable predicates’ have result state type usages which allow them to co-occur with particles like *lagi* and *isih*. However, this capability is not reflected in their behavior under reduplication. A predicate like *gedhe* which had an internally caused result state type usage when used with *bocah* ‘child’, still becomes distributive under reduplication:

- (93) Bocah-e gedhe-gedhe
 child-DEF RED-big
 ‘The children are big’ / ‘The big children’

This contrasts with the behavior of predicates which always denote a result state⁵ as in (94), as well as predicates which denote a stage level physical property such as *luwe* ‘hungry’ in (95):

- (94) Kertas-e suwek-suwek (terus)
 paper-DEF RED-torn (continuously)
 ‘The paper is torn (over and over)’ /‘The papers are torn (over and over)’
- (95) Bocah-e luwe-luwe (terus)

⁵There is an idiosyncratic set of result state/change of state predicates that receive a different sort of meaning under reduplication akin to ‘doing some X’:

- (i) *resik* ‘clean’
 Budi resik-resik
 Budi clean-RED
 ‘Budi does some cleaning/cleans’

This is described by Robson (2014) as a generic version of an otherwise transitive verb, so *resik-resik* above stands in contrast with *ngresiki* ‘to clean (something)’.

child-DEF hungry-RED (continuously)
 ‘The child continues to be hungry/is continuously hungry’

Though you could coerce a plural reading of *bocah* in (95) above through additional context, given that the plurality of the subject is simply underspecified, the point here is that the presence of reduplication on *luwe* does not force plurality of the subject on its own as it did in (94). Furthermore the adverb *terus* ‘continuously’ is possible here, if not preferred to further clarify the purpose of the presence of reduplication.

The plurality of the subject is similarly ambiguous in the case of *kertase* ‘the paper’, but for a different reason. This example denotes a pluractional event of paper being torn. It is unclear whether or not those repeated tears are occurring iteratively to a single sheet of paper, or iterative distinct sheets of paper. This ambiguity is contingent on the lexical semantics of the change-of-state involved, and whether or not it necessarily impacts the entire subject or not. Tearing could impact only part of a piece of paper, leaving some of that paper able to be torn in another iteration of tearing. This contrasts breaking for example, the same glass cannot be broken more than once, breaking necessarily impacts the entire glass, and subsequently when we reduplicate *pecah*, the breaking event described necessitates iterative and distinct glasses:

(96) Gelas-e pecah-pecah (terus)
 glass-DEF RED-break (continuously)
 ‘The glasses break (repeatedly)’

We can note how this behavior mirrors the property of iterativity and subsequently the availability/unavailability of *rada* with these predicate:

(97) a. Kertas-e rada suwek
 paper-DEF half torn
 ‘The paper is half/kind of torn’
 b. #Gelas-e rada pecah
 glass-DEF half broken Int: ‘The glass is half/kind of broken’

It should be noted that both of these result states described thus far are irreversible. When it comes to reversible result states, the behavior under reduplication is more idiosyncratic. Some reversible result states are simply bad with all three options such as *wareg* ‘full’ and *umob* ‘be boiling’⁶. Others are ambiguous between an iterative reading like that above and a continuous reading like what was seen with stage level properties like ‘hungry’ :

(98) Es-e ajer-ajer (terus)
 ice-DEF RED-melt (continuously)

⁶Sometimes my consultant could allow a continuous reading with *umob-umob* under the specific circumstance of leaving the stove on so the water keeps boiling, but it still didn’t seem great.

- ‘The ice melts (continuously)’ / ‘The ice melts (over and over)’
 (99) Suket-e garing-garing (terus)
 grass-DEF RED-dry (continuously)
 ‘The grass dries (continuously)’ / ‘The grass dries out (over and over)’⁷

More work is needed to understand the connection, but it is worth noting that the cases where reversible result states are compatible with reduplication appear to align with the cases where the circumfix *ke-en* was available. These are predicates which could be construed as internally caused, at least in some usages.

Both property states and result states when reduplicated can be used in a construction to indicate that the state denoted by the predicate has not been reached yet (in the case of a result) or entered yet (in the case of a property):

- (100) Esti m-[p]ecah gelas-e ning ora pecah-pecah
 Esti AV-break glass-DEF but NEG break-RED
 ‘Esti broke the glass but it didn’t break’
 (101) Esti ng-umob-ake banyu-ne ning durung umob-umob
 Esti AV-boil-CAUS water-DEF but not.yet boil-RED
 ‘Esti boiled the water but it is not boiling/boiled yet’
 (102) Jam loro awan ning Budi durung luwe-luwe
 hour two afternoon but Budi not.yet RED-hungry
 ‘Its two in the afternoon but Budi is not hungry yet’

Achievement predicates necessarily become iterative as we saw previously with *tiba-tiba*. This is in contrast with activity predicates which can be either continuous or iterative:

- (103) Esti tiba-tiba (terus)
 Esti fall-RED continuously
 ‘Esti falls (repeatedly)’
 (104) Esti m-[p]layu-mlayu (terus)
 Esti AV-run-RED continuously
 ‘Esti runs (continuously)/(repeatedly)’

Some common activities have conventionalized usages, such as *mlaku* ‘to walk’ can mean ‘to hike/walk for recreation’ when reduplicated. There are also several processes of reduplication which involve vowel changes and add other nuanced meanings such as *mloka-mlaku* ‘to walk around, meander’, but a full description of these is beyond the scope of the present investigation.

So far we have observed that reduplication can differentiate nouns from intransitives, and

⁷This stem also has a distributive usage: *klambine garing-garing* ‘all the clothes are dry’. I would argue that this is a case of homophony, and that there is an adjectival form of *garing* that denotes the physical property of being dry.

as with the Aktionsart facts this process can help to subdivide the space of intransitives, though along somewhat different lines. Reduplication lumps individual level and mutable property states together, as well as some stage level properties which denote mental and emotional states. These on the whole encompass predicates which are most aligned with the lexical category of adjectives such as colors, shapes, sizes, as well as things like ‘happy’ and ‘tired’. For this group of predicates, the plurality instantiated by reduplication takes scope over the whole phrase, including the noun being modified, making that noun necessarily plural. I will tentatively label this class adjectival for the time being.

Reduplication treats stage level properties like *lara* ‘sick’ and *luwe* ‘hungry’ differently than these adjective-like predicates above. In this case, the subject is not necessarily pluralized and the adverb *terus* ‘continuously’ is preferred to clarify that the state being described is holding for a prolonged period of time. In cases where the subject is plural, there is still only one continuous occurrence of the state applying to the plural subject, there is no iterativity.

Result states which occur as the result of an irreversible change-of-state such as tearing or breaking, become iterative or pluractional under reduplication. The scope of the plurality is determined by the lexical semantics of the predicate in question. If the predicate is incremental, the plurality of the subject is ambiguous so the scope of plurality can be either high (including the subject) or low (excluding the subject). If the predicate is not incremental the subject is necessarily plural.

Reversible result states are heterogenous in their behavior. This at least tentatively appears to correlate with internal vs. external causation. Externally caused reversible result states seem incompatible with this type of reduplication. Internally caused reversible result states are ambiguous in their behavior, allowing for either continuous or iterative interpretations.

Achievements are necessarily iterative, which follows from the fact that they lack duration. Activities, like the internally caused reversible result states, are ambiguous between continuous and iterative interpretations. All of this is summarized in the table below:

Stem Class	Result of Reduplication
nominal	plural noun
adjectival	distributive
stage level state	continuous
achievement	iterative
activity	continuous, iterative
irreversible result	iterative
reversible result (externally caused)	*
reversible result (internally caused)	continuous, iterative

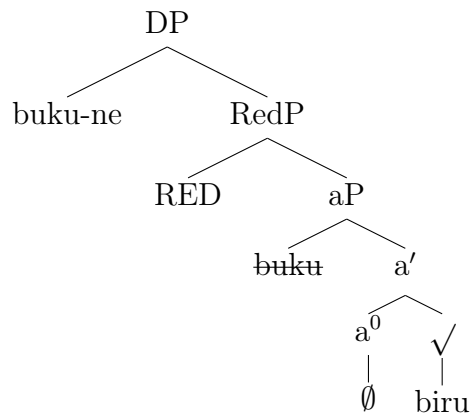
Given the types of reduplication we have observed I am going to divide this process into three core types associated with (and subsequently diagnosing) the three main lexical

categories:

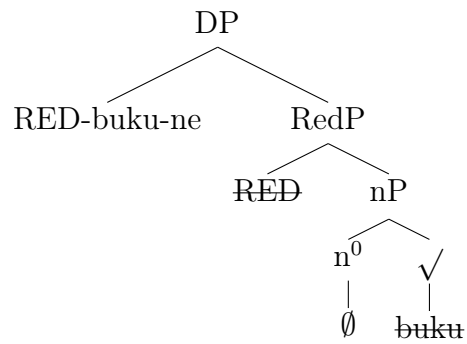
1. Plural: occurs with nominal bases
2. Distributive : occurs with adjectival bases
3. Pluractional : occurs with verbal bases, can instantiate iterativity or continuity depending on the semantics of the verb

Furthermore I am going to propose the following basic structures for reduplication:

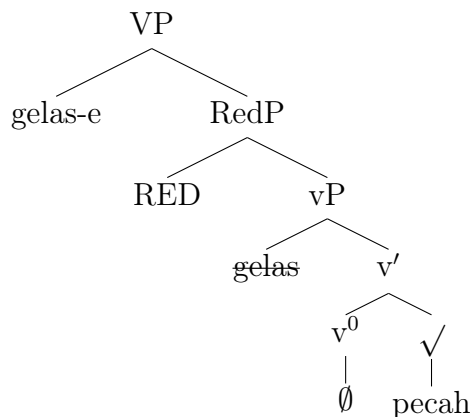
(105) The distributive reading:



(106) The plural noun reading:



(107) The pluractional reading:



The distributive reading is defined by the fact that there is no notion of pluractionality, and the fact that whatever noun is being modified by the predicate is necessarily plural itself, and all nouns in that set possess the relevant property. Reduplication I subsequently argue must be taking scope over the noun being modified by the adjectival stem. This structure is fundamentally different from the other two, and motivation for considering adjectives a defined lexical category of their own despite any overlap we may observe between adjectives and verbs with respect to the morphology these stems can host. This distinction will also prove useful in explaining the distribution of that morphology in Chapter 4.

3.2.2 Lexical Category and the Content of Roots

Working within a framework such as Distributive Morphology (Halle and Marantz 1993) in which all word building is syntactic, the lexical categories we have established above are necessarily arrived at via categorizing heads. I have, by virtue of positing these lexical categories, argued in favor of the classic trio of categorizing heads namely: *n*, *a*, *v*. Though this framework asserts that roots are category neutral, that does not mean they are content free. Before concluding our investigation of the base to the derivation, it is worth exploring the question of root content for Javanese.

Alexiadou and Lohndal (2017) in their chapter about the division of labor between roots and functional structure, establish a typology which places languages on a scale from ‘empty’ roots to ‘contentful’ roots:

‘empty’	→	‘contentful’
Hebrew	Greek	English

They place languages such as Hebrew, which Arad (2003) gives the property of "multiple contextualized meanings", at the ‘empty’ end of the scale. This is because roots appear to be highly flexible not only in their ability to combine with the rich inventory of functional structure, but in their range of potential meanings as well. Roots in general appear to be relatively unspecified for meaning and subsequently content-less. On the other end of the scale is a language like English, which has relatively little overt functional structure, ample zero-derivation, and roots which appear to have highly specific/inflexible meanings. Greek they argue based off of certain derivational processes including the creation of participles and verbalization, falls somewhere in the middle.

As we have now seen in greater detail, Javanese like English (though to a lesser extent) exhibits a great degree of zero-derivation. Javanese, unlike English and more so like Greek, possesses an abundance of overt verbalizing morphology and many agentive verbs require this to be overt at least to some extent. Still, as was noted in Chapter 2, stems of all lexical categories can combine with the full inventory of derivational morphemes present in Javanese. However, it is worth noting that zero-derivation does not generally give rise to homophony. We do not see bare stems of the same phonological form belonging to multiple lexical categories, at least this appears to be very uncommon. Moreover, even given the

flexibility of the morphology to apply to lexical categories of different types and the other empirical challenges raised in Chapter 2, the actual meaning of a resulting derivative remains largely semantically transparent and predictable, as will be made evident in Chapters 4 and 5. Given these facts about Javanese, I would propose that it falls closer to the ‘contentful’ range of the typological scale:

‘empty’	→	‘contentful’
Hebrew	Greek	<i>Javanese</i> English

I will carry forward this notion that Javanese roots are not empty into the discussion of verbalizing morphology in Chapter 4 and ultimately propose a four way distinction among roots based off of the three root classes proposed for English in Harley (2005), but modified for the Javanese case. These classes will be:

- $\sqrt{property}$
- $\sqrt{experience}$
- \sqrt{event}
- \sqrt{state}

Harley originally had a class of *state* roots which I have divided into *property* and *experience* roots. This division will help us predict the distribution of the Actor Voice prefix, which is the focus of the next section.

Chapter 4

Deriving Verbs

Javanese, as has been stated, employs very little inflectional morphology and contains only a small inventory of derivational morphemes. These derivational morphemes carry out a wide variety of functions however, leading to the formation of complex and rich derivational paradigms which ultimately form the backbone of the Javanese lexicon. In our investigation of this system, the present chapter turns to the derivation of verbs. First we will be looking at verbs prefixed only by voice morphology, then section (4.2) will expand this discussion to the inventory of valence increasing suffixes.

I will first provide an overview of existing analyses of the Javanese voice system. I will argue that the existing characterization of the actor voice (AV) prefix, which takes it as indicative of the presence of an agent in the highest specifier position of v and a lack of object shift, is insufficient. This approach does not capture the full scope of this prefix's distribution. I will propose that this morpheme should be considered a hybrid, belonging to the voice system in yes signaling a lack of object shift where relevant, but also belonging to the derivational system in its role as both a verbalizer and argument introducer. Moreover, the heterogenous nature of the derivational behavior I will argue is the result of homophony. There are two versions of this prefix which occur in complementary distribution:

1. Voice⁰ : selects for an intransitive vP, introduces the position for an external argument
2. v^0 : selects for a non-verbal stem, verbalizes and introduces an unaccusative subject

This will be done through an in depth empirical discussion of the various contexts in which the AV prefix appears. This will be followed by a discussion of where it does not appear, and how its distribution interfaces with other derivational processes as well as other voice morphology.

Generally in Javanese, the same morphemes are observed when deriving verbs from other verbs, as in the unaccusative/transitive alternation, or verbs from non-verbal predicates such as nouns and adjectives. Though they do not do so indiscriminately, given the complex nature of these distributions, the discussion of both processes will overlap. Based on the diagnostic criteria established in Chapter 3 for lexical category and predicate class, the following processes will be discussed :

1. $V \rightarrow V(+)$: Argument introduction
2. $A \rightarrow V$: Verbalization and argument introduction
3. $N \rightarrow V$: Verbalization and argument introduction

4.1 The Actor Voice Prefix as a Derivational Morpheme

In a study of Javanese verbal morphology, arguably the most fundamental overt morpheme is the voice prefix. In addition to its function within the voice system, this prefix correlates with a variety of verbalizing and argument structure altering processes and is obligatory in the presence of additional valence increasing morphology. It has been argued that the voice system in Javanese is dyadic (Uhlenbeck 1978, Poedjosoedarmo 1986, Keeler 1984, Robson 1992). The two parts in opposition have received a variety of names in the literature but I will use the terms Actor and Object here¹.

I will first focus on the actor voice (AV) prefix because it plays the most prominent role in verbalization. The object voice (OV) prefix does not occur in intransitive contexts, but mirrors the AV distribution in transitive and ditransitive contexts. The accidental passive will be briefly discussed in section (4.3). The AV prefix consists of a nasal consonant which assimilates in place of articulation to the first consonant of the root. No morphology is permitted to intervene between this prefix and the root. The assimilation correspondences are given in the table at the top of the following page, reproduced from Robson and Wibosono (2002), these are supplemented by the rules for prefixation given in Connors (2008). These rules refer to Standard Javanese.

It has been observed that this prefix frequently occurs in *agentive* contexts, specifically in transitives, though Robson (2002) notes that the AV prefix is also found with a number of intransitives as well. Given this distribution it has been argued that the prefix signals the introduction of an agent and the syntactic position of that agent. This is the analysis put forth in Sato (2015). Sato's analysis is an extension of the proposal presented in Cole et al (2008) for voice in Indonesian, which itself is based on Rackowski and Richards (2005)'s treatment of voice in Tagalog.

¹Recent studies (Ogloblin 2005: adopted in Krausse 2017, Vander Kloek and Evans 2022) argue that there are three voices: Actor, Patient, and Passive. However all that differentiates patient and passive is whether or not the agent is a participant (1/2 person) or in the third person, I gloss over this distinction at this time. Ogloblin presents data wherein there is an overt pre-verbal subject in cases of patient voice, but this data is not replicated in the dialect of my consultant.

	Initial letter of root	AV prefix form
	V	ng-
initial letter deletes	p t th c k s w	m- n- n- ny- ng- ny- or n- m-
no deletion	b d dh j g r l	m- n- n- n- ng- ng- ng-
deletion or degemination	n, m, ng	n, m, ng

Sato argues that the AV prefix in Javanese signals that the Agent of the verb is in the highest specifier position of *v*, and that no object shift has taken place. This analysis, he states, also conforms with the proposal given in Connors (2008) which argues that the AV prefix marks the presence of an agent or actor argument. In sum, the existing literature argues that the presence of this prefix is indicative of two things² :

1. The presence of a syntactic agent
2. The location of that agent is the (highest) specifier position of *v*

This characterization of the AV prefix is compatible with Robson’s observation mentioned above. Furthermore, it is indeed the case that examples of the AV prefix are overwhelmingly agentive and transitive, and it is true that the instances of intransitives hosting this prefix are *generally* agentive as well:

(108) Esti n-[t]uku buku-ne
 Esti AV-buy book-DEF
 ‘Esti buys/bought the book’

(109) Esti m-[p]layu
 Esti AV-run
 ‘Esti runs/ran’

In (108) and (109) above we can observe the prefix attaching to the root *tuku* which forms transitive verb ‘to buy’ as well as *playu* forming the agentive unergative ‘to run’. However, this picture is complicated by the fact that the AV prefix is not obligatory in order to form the verb ‘to buy’, as in (110):

²And by extension, also that object shift has not occurred

- (110) Esti tuku buku-ne
 Esti buy book-DEF
 ‘Esti buys/bought the book’

If the AV prefix is not necessary in the presence of the agent in (110) then its availability in (108) requires further explanation. We could argue that the presence of an agent merely makes the prefix available. In other words, the prefix (at least in some cases) correlates with the presence of an agent but is not actually obligatory in this context. This would necessarily mean that the prefix itself is not serving as an agent introducing head, but just appearing (optionally) as a reflex of a particular syntactic configuration. It is important to note that the AV prefix is not optional in (109). So argument introducing head or no, the availability or absence of optionality remains unexplained.

This situation is further complicated by the fact that this prefix can occur on non-agentive/theme unergatives as well as unaccusatives (111-112):

- (111) Esti ng-[g]uyu
 Esti AV-laugh
 ‘Esti laughs/laughed’
- (112) Esti m-[p]ati
 Esti AV-die
 ‘Esti died’

It could be argued that in the case of the theme unergative in (111), the subject is still an external argument which is necessarily originating in the specifier position of *v*, and despite that subject not being strongly *agentive* per se this is enough to warrant the AV prefix. However, the data in (112) if ‘to die’ is in fact an unaccusative (as is cross-linguistically typical), remains mysterious. This data would suggest that the AV prefix is signaling only the movement of the unaccusative subject to the specifier position of *v*. In this case there is no external argument present, agentive or otherwise. Moreover, we know from looking at other (probable) unaccusatives in this language that the AV prefix is not available just any old time a DP occupies this position:

- (113) Banyu-ne umob/*ngumob
 water-DEF boil
 ‘The water boils/boiled’

Putting aside the issue of language internal unaccusativity diagnostics for the time being, what is evident is that characterizing the AV prefix as indicative of the presence of an agent is insufficient. Given these data, the picture we are now left with is one in which the AV prefix has been reduced to a morpheme that sometimes indicates that something is in the specifier position of *v*, without speaking to the type of thing in that position, and without being clearly responsible for the movement itself or the introduction of any argument. This is the result given the data presented above, if we attempt a homogenous description of the prefix based entirely on what we know about the resulting stem.

This approach follows from the treatment of the AV prefix as strictly a voice-agreement morpheme. However, as the following sections will illustrate, this prefix is tightly bound up with the process of verbalization and the broader system of valence altering morphology in Javanese. This prefix exists in complementary distribution with the object prefix series *tak-/kok-/di-* and should be considered a part of the voice system. At the same time it exists in complementary distribution with the nominalizing suffix *-an*. Taking this into consideration along with the full observed distribution of the AV prefix across paradigms, I argue that it is a member of the derivational system as well.

Specifically, I will argue that what has been described as the AV prefix in the literature actually corresponds to two homophonous morphemes, which exist in complementary distribution. One of these serves to introduce an external argument in the context of a stem that is already verbal, and the other verbalizes a non-verbal stem and derives an unaccusative verb. In order to tease apart the collection of phenomena associated with the overt form of this prefix, it is necessary to contextualize it from above, examining the ultimate output of the derivation as we have done here, as well as from below. Namely, beyond taking into account the *result* of prefixation, we must take into account the full scope of possible inputs to prefixation.

4.1.1 Intransitive Bases

In this section we will look at the behavior of the AV prefix when occurring on all bases which fall within the broad umbrella of intransitive predicates, indicating how these facts relate to our more nuanced categorizing criteria from Chapter 3 as becomes relevant.

We will see throughout this section that 4 main patterns arise when looking at AV prefixation on intransitive bases, these are summarized in the following table:

Bare Stem	with AV prefix
intransitive	intransitive
intransitive	transitive
intransitive	*
*	intransitive

In cases where the bare stem is grammatical as an intransitive, the addition of the AV prefix has three potential results. Either the result is ungrammatical, there is no apparent change in valence and the stem remains intransitive, or an argument is added and the resulting stem is transitive. There are also cases where the bare stem is ungrammatical, and a grammatical intransitive is formed as a result of pre-fixation. We will examine each of these patterns in greater detail, beginning with cases which most align with the expected behavior of the AV prefix, namely the derived transitives.

Derived transitives from intransitive bases

Two main predicate types are compatible with transitivity under prefixation: adjectives denoting mutable properties as well as result states compatible with external causation³. Examples of a transitives derived from predicates denoting mutable properties are given in (114-115), and examples of those derived from result states are in (116-117)

- (114) Esti m-bunder lempung-e
Esti Av-round clay-DEF
'Esti rounds the clay'
- (115) Esti ny-[c]ampur jus-e
Esti AV-mixed juice-DEF
'Esti mixes the juice'
- (116) Esti m-[p]ecah gelas-e
Esti AV-broken glass-DEF
'Esti breaks the glass'
- (117) Esti m-bosok tomat-e
Esti AV-rots tomato-DEF
'Esti rots the tomatoes'

In Javanese the notion of 'rotting' is compatible with both internal and external causation, the availability of external causation is all that is required to make (117) above available. The subset of adjectives we see co-occurring with this process include mainly those denoting mutable properties of physical appearance or orientation, including color and shape. We don't see this occurring with those denoting emotional states or states that denote internally felt experiences such as 'tired' or 'cold'⁴. The adjectives that are available here are specifically ones which could readily be construed as the result of an externally caused event. This is supported by the fact that these predicates, though not typically compatible with *lagi*, were compatible with *isih*. The presence of *isih* is sensitive to a notion of being dynamic or susceptible to change. In general what appears to unite the set of predicates compatible with transitivity, is the availability of external causation.

Based on these data we can conclude that the AV correlates with the derivation of transitives from roots which are associated with both intransitive verbs as well as adjectives. So the challenge for our analysis then becomes determining specifically which adjectives and intransitive verbs form derivatives of this kind, because as we will see in the follow-

³These align with the observations made by Levin and Rappaport Hovav (1995) regarding the transitive alternation in English. They observed that this alternation occurred specifically with unaccusatives that are compatible with external causation. Additionally they observed that adjectives which are capable of forming de-adjectival verbs necessarily have a stage level usage, something first observed by Dowty (1979). I am not arguing that these transitives should be considered 'causative' in a meaningful way, just that it is worth noting that the lexical semantics of the stems involved do line up with what is observed elsewhere.

⁴Javanese has different words for cold which is internally experienced/inherent to an object and cold which is externally experienced such as when you touch something and it 'feels cold'. The word for externally felt cold is *anyep*, this does transitivity with the AV, though it also requires reduplication to *nganyep-anyep* to differentiate it from the would-be-homophonous *nganyep* 'to fast/eat only bland food'.

ing section, intransitive verbs and adjectives also correlate with the derivation of agentive unergatives and unaccusatives. I am ultimately going to argue that this distinction can be achieved by appealing to the notion introduced in Chapter 3 regarding roots having content. I will propose an inventory of root classes which will instantiate the observed distribution of derivatives. The stems contained in this section will be derived from what I will call *property* roots, which include roots that categorize as property denoting adjectives as well as result state verbs, both being compatible as was said above with external causation.

Derived intransitives from intransitive bases

Understanding what is happening when the AV prefix results in an intransitive stem requires a nuanced understanding of what properties are in fact changing when the prefix appears, as well as what rules out the availability of a transitive here. With respect to the first of these tasks, it appears as though prefixation with the AV does not increase the valence of the stem when occurring with a subset of adjectival predicates. The subset of adjectival predicates possible here form the complement set to the adjectival predicates possible in the subsection above. These denote internal experiences such as ‘cold’ and ‘tired’ as well as certain properties of physical position/orientation that were ruled out above such as ‘short’ and ‘tangled’. In general these are the predicates that blur the line between adjective and state. Things that fall into the camp of internally caused result states are also possible here such as ‘arrive’, as well as achievements such as ‘fall’.

Despite failing to correlate with an increase in the valence of the stem, the addition of the AV prefix does correlate with an observable change in every case with these stems. This prefix appears to either derive an externally caused result state as in (118) with *bundhet* ‘tangled’. This can be observed in the availability of *lagi* in (118b) only. Alternatively, it appears to add an agent in (119-121). This could be construed as adding internal causation, but the resulting predicates do not behave like result states with respect to *meh*⁵. For this reason I think it is more appropriate to refer to this as agency and to categorize the resulting stems as activities. This is further supported by the fact that the resulting stems are not compatible with *rada* which was stated in Chapter 3 to correlate with the presence of an agent⁶.

(118) adjective → result state (+external causation)

- a. #Tali-ne lagi bundhet
 rope-DEF lagi tangled
 Int: The rope is tangled (right now)
- b. Tali-ne lagi m-bundhet
 rope-DEF lagi AV-tangeled
 ‘The rope is tangled right now’

⁵*meh* gives an “almost complete” reading with result states, but an “about to start” reading with activities.

⁶This was shown in the availability for *rada* to modify things like *mati* ‘dead’ but not things like *mlayu* ‘run’ or *mangan* ‘eat’.

(119) adjective/state →activity (+agent)

- a. Es-e kambang
ice-DEF afloat
'The ice is afloat/floats'
- b. #Es-e lagi kambang
ice-DEF lagi afloat
Int: the ice is afloat (right now)
- c. Esti lagi ng-[k]ambang
Esti lagi AV-afloat
'Esti is floating (right now)'

- (120)
- a. Sup-e adhem
soup-DEF cold
'The soup is cold'
 - b. *Sup-e ng-adhem
soup-DEF AV-cold
Int: the soup cooled
 - c. Esti ng-adhem
Esti AV-cold
'Esti cooled (herself off)'

- (121)
- a. Esti lara
Esti sick
'Esti is sick'
 - b. Esti ng-lara
Esti AV-sick
'Esti pretends to be sick'

In (14) above, the meaning attributed to the agentive AV containing stem involves pretending to have the quality denoted by the predicate. This is one way the semantics of the verb can be bent to accommodate the introduction of an agent. Similarly in (15) below with the achievement *tiba*, the AV derivative specifically means to fall 'on purpose'. We get a similar result when we prefix *teka* 'to arrive' with the AV. The meaning is something like showing up unannounced.

(122) achievement →activity (+agent)

- a. #Esti lagi tiba
Esti lagi fall
Int: Esti is falling right now
- b. Esti lagi n-[t]iba
Esti lagi AV-fall
'Esti is falling right now (on purpose)'

In general what we can observe here is that there is a split between AV prefixed forms which

are agentless as in the unaccusative *mbundhet* ‘get tangled’, and those which contain agents such as *niba* ‘fall on purpose’. I want to argue that the difference is predicated on whether or not the AV is prefixing directly to an adjectival base or to a verbal base. In the former case the AV correlates only with verbalization and derives an unaccusative verb. In the latter case the AV introduces an external argument, promoting the internal argument of the stem, and deriving an agentive unergative.

What we must now account for is the fact that transitivity is unavailable to these stems. To do this I am going to return to the notion of contentful roots from Chapter 3. I am going to appeal to an analysis of root content in English proposed by Harley (2005), with one key modification made to fit the Javanese facts. Harley proposed a three way distinction among roots for English:

1. $\sqrt{\text{State}}$
2. $\sqrt{\text{Event}}$
3. $\sqrt{\text{Thing}}$

All of the derivatives we have discussed so far have either been derived from adjectival predicates or non-agentive verbal predicates, either result state roots or actions which are arguably unaccusative such as ‘fall’. I am going to argue that all of these stems are built from roots of the State type, but I am going to further divide this class to capture the Javanese facts. In Javanese, instead of a class of State roots there are :

- Property roots: includes adjectives and result state predicates which do not denote internally felt/undergone experiences
- Experience roots: includes adjectives and result state predicates which denote internally felt/undergone experiences

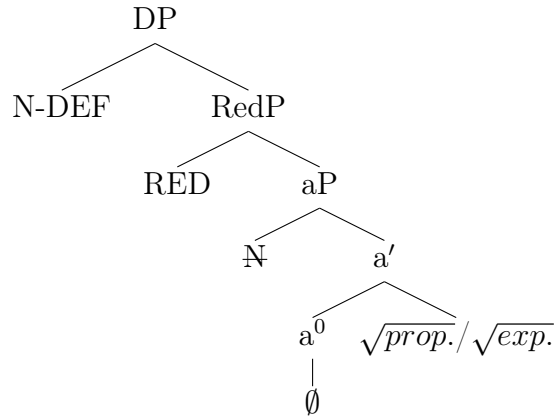
As stated above, both *property* and *experience* roots can form zero-derived adjectives as well as verbs:



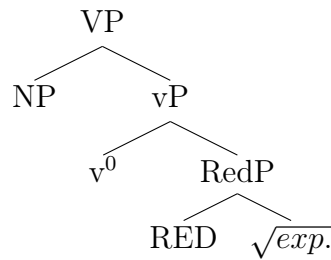
This is reflected in the behavior of these stems under reduplication. All zero-derived adjectives are compatible with the distributive reading. All zero-derived verbs are compatible with the pluractional verbal reading. It is then an added nuance that reduplication can directly attach to some experience roots, giving the continuous state reading in cases like

luwe-luwe ‘continuously hungry’:

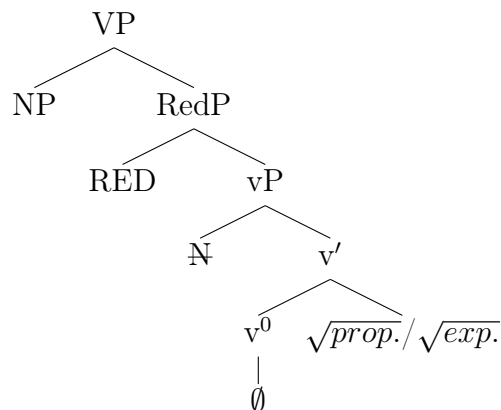
(123) Revised tree for distributive reduplication:



(124) Revised tree for continuous state reduplication:



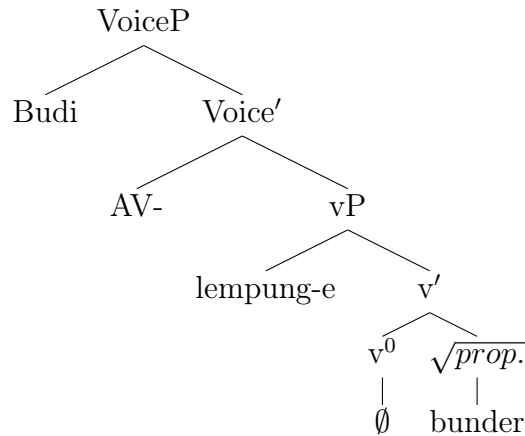
(125) Revised tree for pluractional reduplication:



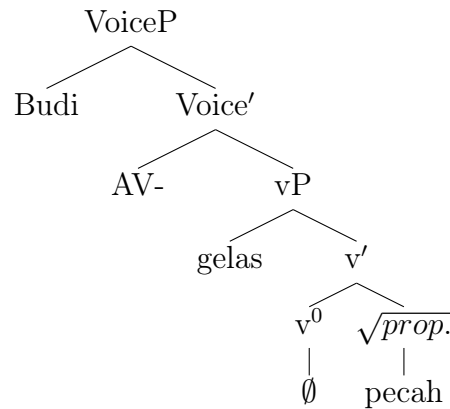
Moreover, this division allows us to readily account for the split in behavior under pre-fixation with the AV. Only property roots which are able to be directly categorized with *v* form transitives of the form seen in the previous section. This includes all externally caused result state predicates, as well as only those adjectival predicates which are compatible with external causation. The version of the AV in this case, given that it is attaching to some-

thing already verbal, necessarily introduces an external argument. This allows for the base generation of an agentive subject, and derives a transitive verb.

(126) Budi mbunder lempunge:

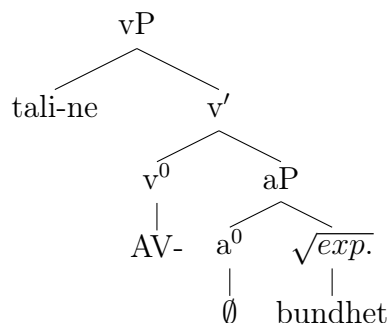


(127) Budi mecah gelase:

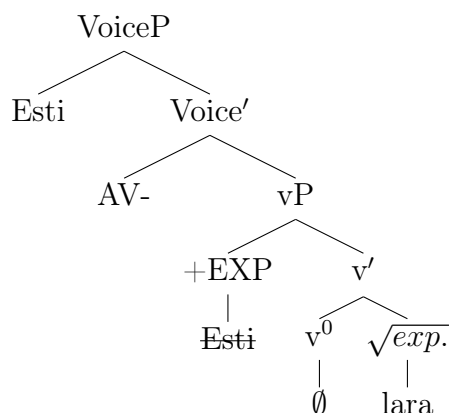


On the other hand, in the case of intransitive AV derivatives, these are derived from experience roots. Unaccusatives are derived from adjectival experience predicates, as is seen in example (118) with *bundhet* ‘tangled’. Unergatives are derived from verbal experience predicates, including states like *adhem* ‘cold’ and achievements like *tiba* ‘fall’. The variant of the AV that is available here again introduces an external argument because it is attaching to something that is already categorized as verbal. However, the internal argument which I will classify as having the property of being an ‘experiencer’, is able to act as the external argument and there is no need to base generate an additional argument. This is why these derivatives remain intransitive. This also aligns with a general principle that I will advocate for over the course of this chapter that favors derivatives that contain the fewest number of arguments possible, or which are structurally the smallest. We could propose that the feature of experience roots that allows for the unaccusative subject to then become an external argument is the fact that these stems are semantically associated with internal causation.

(128) tali-ne mbundhet



(129) Esti nglara



So to briefly summarize what we have seen so far, the division of *property* and *experience* roots subdivides the inventories of bare adjectives and unaccusative verbs. This division allows for us to account for the distribution of unaccusative, unergative, and transitive stems derived with AV morphology. I have argued that the AV has two versions, one which attaches to bases that have been categorized as verbal to which it adds an external argument, and one which attaches to bases that are non-verbal which it then verbalizes deriving an unaccusative. All property roots can verbalize directly and subsequently are compatible with the external argument introducing AV, moreover these stems are compatible with external causation and the external argument is base generated in the specifier of Voice, deriving a transitive verb. Experience roots which form zero derived verbs, also are compatible with the AV prefix which introduces an external argument position, but the subjects of these verbs are able to then be used as the external arguments of the AV derivatives, keeping the stem intransitive and forming an unergative. I proposed that this was related to the types of unaccusative subjects we see associated with these roots, *experiencers*, and their association with the possibility of internal causation. Of the roots associated with zero derived adjectives, only experience root adjectives resulted in unaccusative derivatives under prefixation with the AV. I argued that this was because these stems were necessarily already categorized with a a^0 and that the version of the AV available when attaching to a non-verbal base, verbalizes

that base but does not introduce an external argument, resulting in an unaccusative.

One question raised by this analysis is: What rules out instances of the AV affixing directly to property adjectives and deriving unaccusatives like ‘be round’ or ‘be blue’. Is this only available to adjectives derived from experience roots? I am going to propose that for a given speaker, there is a general principle that discourages homophony. Given this, within the paradigm of a property root that is compatible with adjectival categorization and with external causation, either the AV is going to derive a transitive or it will derive an unaccusative. The grammar will choose which of these options to privilege and rule the other out. This analysis predicts that, if we see inter-speaker variation we should see it in some ways but not others. Namely, we may see speakers where adjectival property roots AND adjectival experience roots form unaccusatives. Or we will see cases like the one given above where adjectival property roots verbalize and form transitives and adjectival experience roots form unaccusatives. Crucially we should not see cases where adjectival experience roots form transitives. So far this appears to be borne out. The data given above represents the grammar of my consultant, however p.c from my Javanese language instructor indicates the predicted alternative:

- (130) biru IV
Tengan-e Esti m-biru
hand-3.poss Esti AV-blue
‘Esti’s hand turned blue’
- (131) bunder TV
Esti m-bunder lempung
Esti AV-round clay-DEF
‘Esti rounds the clay ’
- (132) *TV bundhet
*Esti m-bundhet tali-ne
Esti AV-tangle rope-DEF
intended: ‘Esti tangles the rope’

It is worth noting that the status of *bundhet* ‘tangled’ as an experience root and zero derived adjective is externally supported by its compatibility with both the distributive reading and the continuous state reading under reduplication. The continuous state reading is only available to *experience* roots, and it does not result in pluractionality due to the fact that reduplication is affixing directly to the root.

So to summarize what we have seen so far, the AV prefix when selecting for a verbal base occupies Voice and introduces an external argument position. The external argument is either base generated high, or the unaccusative subject associated with an *experience* root becomes the external argument. This distinction I have argued is due to a property of the unaccusative subject of specifically *experience* roots, in order to act on the observation that it is specifically AV prefixed stems built from experience roots which form agentive unergatives.

The property vs. experience root distinction which I have proposed also helps us explain

the distribution of verbalizing AV and external argument introducing AV. This is because property roots, for my speaker, can categorize directly with v^0 , and subsequently form transitives, while experience roots can categorize with a^0 . When an aP is formed, the variant of the AV present is necessarily verbalizing and does not introduce an external argument, instead it derives an unaccusative verb as in *mbundhet* ‘get tangled’.

Cases of ungrammaticality under prefixation

So far we have addressed the problems inherent to delimiting when the AV prefix correlates with verbalizing or both verbalizing and the introduction of an external argument. Now we turn to the issue of its absence. We have seen above that this prefix can derive an unaccusative verb from an adjectival stem. It can also derive unergative verbs from verbal stems built from experience roots, and transitive verbs from verbal stems built from property roots. Now we must rule out the following:

- (133) adjective? : *amba* ‘wide’ *ngamba
- (134) experience verb? : *luwe* ‘hungry’ *ngluwe
- (135) property verb? : *garing* ‘dry’ *nggaring
- (136) unergative verb? : *lungguh* ‘sit’ *nglungguh

Given what we have seen the AV prefix do, it is expected that its presence would be ruled out in certain contexts, on either selectional or semantic grounds. This is because in order for the AV to appear on a verbal stem, it necessarily introduces an external argument, so the stem being prefixed must be compatible with external argument introduction. If the semantics of the stem do not permit this, or the stem through other morphological means already contains an external argument, we would expect to see the AV prefix disallowed. Similarly, in order to be a verbalizer and derive an unaccusative, this variant of the AV prefix requires a non-verbal stem, so a stem built from a root that selects for a little a^0 (or a little n head, but we will return to denominal verbs later), but one which contains the lexico-semantic properties necessary to make it compatible with subsequent non-agentive verbalization. If either of these criteria go unmet, we would again expect the AV prefix to be ruled out.

So, when we look at an example like *luwe* ‘hungry’, this stem is a zero derived verb built from an experience root. Subsequently, the AV would necessarily promote the experiencer subject to the external argument position. It might simply be the case that the semantics of ‘hunger’ rule out an agentive unergative version of this stem, though it is unclear what actually rules out a meaning like ‘pretend to be hungry’ given that we know such a meaning is possible with things like *lara* ‘be sick’.

In the cases of *amba* ‘wide’ and *garing* ‘dry’, arguably property roots, we must again appeal to some fact about lexical semantics or the compatibility of these stems with verbalization. The verbalization angle is ruled out for *garing* as this stem gets a pluractional reading under reduplication. We could argue that *amba* is simply incompatible with the direct verbalization necessary to make it compatible with the version of the AV that introduces an external argument, ruling out a transitive built from the AV. We would then have

to rely on the fact that this is a property root, and not an experience root, to say that the unaccusative is ruled out, as in the grammar of my speaker only adjectival stems built from experience roots are compatible with verbalization with the AV. As with *luwe*, the most we can say is that the basic requirements must not be met for the AV to form a transitive, an agentive unergative, or an unaccusative verb for these stems. We can speculate which factors must be missing on a stem by stem basis, but beyond this the true cause of the unavailability remains mysterious.

These stems can all form transitives with the addition of the *-ake* or *-i* suffix, which will be discussed in detail in section (4.2). What is relevant to the present discussion is that, the analysis of these suffixes that will be proposed does necessitate compatibility with an unaccusative vP structure, somewhere in the derivation. So this complicates the explanation of specifically *amba*, which is our strongest contender for an “it simply can’t verbalize that way” story. Though it could be the case that the nature of verbalization contained within an AV only structure, vs. an AV and *-ake/-i* structure is inherently different in some way. I am going to preliminarily say that this is the case, and further say that AV only transitives require the stem being transitivized to be compatible with external causation, something potentially ruled out for *amba*. This requirement does not hold for transitives built with an additional suffix, because the suffixes attribute properties to the unaccusative subjects that they contain, for example that of *causee*. This remedies the need for the stem to be already compatible with external causation on its own, and transitives with stems like *amba* become possible under suffixation.

Now we can expand this discussion to stems which are arguably agentive verbs when bare as in (137) below:

- (137) *lungguh* ‘sit’ Parta *lungguh*
 ‘Parta sits’

The stem above is not compatible with *rada*, which as has been stated is ruled out in the presence of an agent. It also gets the ‘about to happen’ reading with *meh*. These facts align with considering the above stem an agentive unergative. I argue that agentive unergatives like these are derived from *event* roots. Arguably the AV is ruled out here because it can only do a finite set of things. When affixing to an already verbal intransitive it introduces an external argument. However, here the external argument is possible when the stem is bare, so the AV is not required. Arguably there is a null allomorph in Voice which alternates with the overt form of the prefix. Agentive bare stems are significantly rarer across the lexicon, and so we must stipulate that this set simply is derived necessarily with the null allomorph and moreover when the null allomorph is possible, the overt version is ruled out. We will revisit this proposal several times as we examine other stems which are necessarily always agentive.

Cases of ungrammaticality when bare

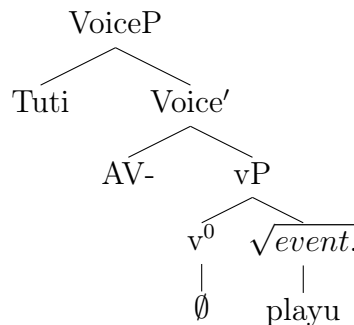
We stated in the previous section that the range of functions available to the AV prefix in its derivational capacity, is finite. This prefix can verbalize adjectival stems, or it can introduce an external argument to a verbal stem which does not already contain one. This

allowed us to rule out the introduction of the AV prefix, in cases where an eventive verb exists grammatically on its own when bare. However, now given this, we must explain the following:

- (138) Tuti m-[p]layu
 Tuti AV-run
 ‘Tuti runs’

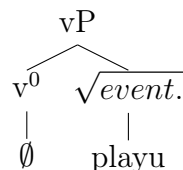
As with *lungguh* ‘sit’ above, we are dealing with an event root. However, in (138) above this root is not compatible with zero-derivation. So the AV is necessary to produce the grammatical form in (138). Arguably the introduction of the AV prefix is still correlating with the introduction of the external argument, but in this case it is not *creating* an agentive form where previously there existed a grammatical non-agentive form. Given this, it is easier to account for the fact that the result of the derivation is still intransitive. The event root $\sqrt{\text{playu}}$ does not come with an experiencer argument to promote to the role of agent. Subsequently, the resulting structure post-derivation with the AV allows for the base generation of a subject.

- (139) Tuti mlayu

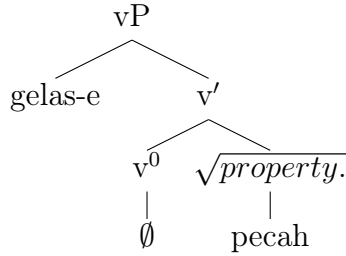


This behavior aligns with what we have seen of the AV prefix so far, when it is affixing to a stem that is already categorized as verbal, it introduces an external argument. There is just the added caveat that this stem is incapable of surfacing on its own as a zero derived verb. We could attribute this to the fact that there is no non-agentive version of *playu*, and unlike the other zero derived verbs we saw which were unaccusative, in this case there is no internal argument and subsequently no argument low enough to be introduced in the portion of the verb that is zero derived. For this to surface as zero derived it would need to be compatible with an unaccusative subject, because this is the only potential specifier position available to a verb this small. We can compare these two structures below:

- (140) an unergative stem without the AV:



- (141) an unaccusative stem without the AV:



The vast majority of agentive unergatives require this overt derivation with the AV prefix. The unergative stems in the previous section should be considered outliers. Their presence remains puzzling however, in that given our analysis they must contain covert structure that is larger than what is available to stems like *playu* through zero derivation. We will return to the topic of zero-derived agentive verbs in the following section when we discuss zero derived transitives.

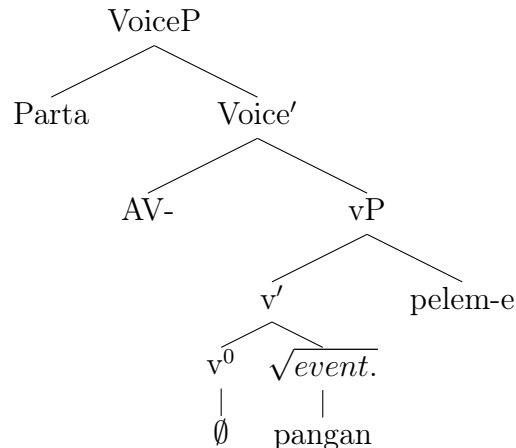
4.1.2 Transitive Bases

In this section we will look at the behavior of the AV prefix in conjunction with stems which are always at least transitive. Two overarching patterns exist within this set of bases, summarized below:

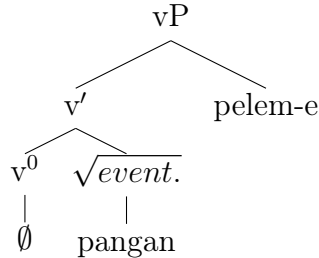
Bare stem	AV prefix
*	transitive
(di-)transitive	(di-)transitive

The stems discussed in this section are derived from event roots. The vast majority of transitive stems in Javanese are simply ungrammatical without the AV prefix. As with the agentive unergatives we saw above, in these cases the AV is correlating with the introduction of all necessary material for a grammatical transitive. This includes both the internal and external argument:

- (142) Parta m-[p]angan pelem-e
 Parta AV-eat mango-DEF
 'Parta eats the mango'



This fact again, aligns with the previous observations that when affixing to a stem which is categorized as verbal, the AV introduces an external argument. It also follows that these stems are ungrammatical when zero derived because only the internal argument is low enough to be introduced by the structure available under zero derivation:



The picture is complicated by the fact that some stems appear to surface as transitive without the AV prefix present:

- (143) *tuku* ‘buy’ Esti *tuku buku*
 Esti buy book
 ‘Esti buys a book’

It is further complicated by the fact that these stems are indeed compatible with the AV prefix, but its presence is conditioned by context:

- (144) Esti n-[t]*uku buku-ne kanca-ku*
 Esti AV-buy book-DEF friend-1.poss
 ‘Esti buys my friend’s book’

In (144) above the *nuku* derivative is used when object is more specific. This, according at least to the intuitions of my consultant and my language instructor, makes the sentence more ‘active’. There is more intentionality involved when there is a specific book in mind being bought. Despite the difficulty inherent in pinning down this meaning difference, the presence of a contrast means that we cannot simply say that there is a zero derived allomorph of the AV prefix in these cases. If that were true, we would expect either the result we saw with unergatives above where the AV was ungrammatical, OR we would expect true optionality, where both derivatives are equivalent in meaning. Moreover, it is not the case that the *tuku* version is used in cases where the object is dropped/covert/absent. The example in (143) is necessarily transitive.

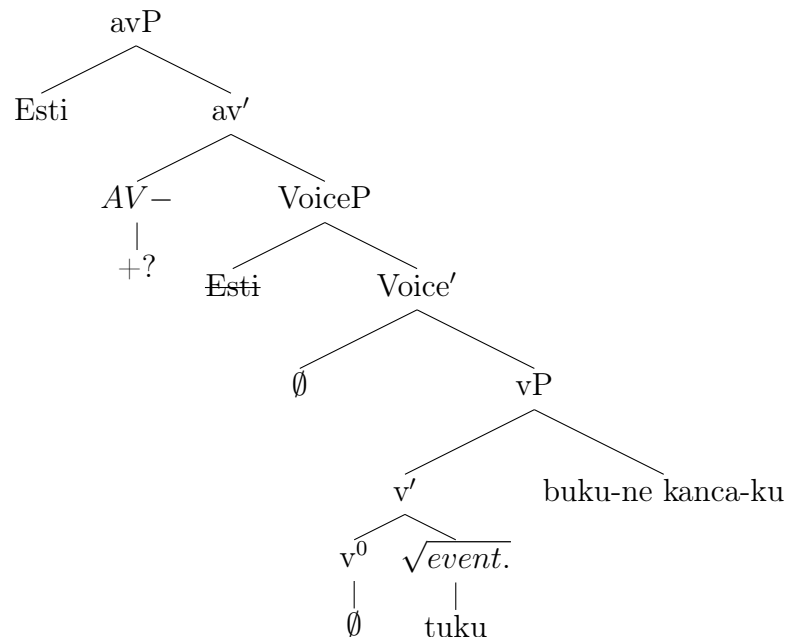
Another example of a change in meaning resulting from the addition of the AV prefix is given in (145) below. In this case the presence of the AV prefix indicates a greater degree of material ownership over the object in question, so it is used in descriptions of wealth and with things like real estate property:

- (145) a. *Aku duwe buku*
 1.sing have book
 ‘I have a book’
 b. *Dheweke n-duwe akeh omah*
 3.sing AV-have many house

‘He owns many houses’

In both the case with *tuku* ‘buy’ and the one with *duwe* ‘have’, the prefix is correlating with a greater degree of agency or intentionality attributed to the subject and their relationship to the direct object. In the case of buying this is indicated by the specificity of that object, in the case of ownership it is sort of situationally determined. It is difficult to determine how this contrast is manifested syntactically however.

If we want to unite these cases with the unergative cases and salvage an analysis wherein these cases contain a null allomorph of the AV prefix, we could say something like the following. We could argue that indeed the AV that introduces an external argument is ruled out here on account of the presence of this null variant. However, there is a structurally higher position at which the AV is attaching, and in this position it is effectively altering the theta role attributed to the external argument in some way that gives rise to the alternations we see.



We then have to account for the fact that this version of the AV is ruled out in the unergative cases seen previously. However, we might be able to do this by arguing that the theta role attributed to the transitive subject discussed here, is only available in transitive contexts, as it relates to the relationship between the subject and direct object in some meaningful way. This would rule out the availability of an overt AV prefix in the unergative cases like *sendhen* ‘lean’, all together.

Finally, we will turn to cases where ditransitives occur grammatically as bare stems. We never observe the AV deriving a ditransitive from a grammatical bare transitive, nor do we see an AV transitive related to a grammatical bare transitive. We only see instances of apparent optionality where a stem can surface grammatically as ditransitive, with or without the AV:

(146) Budi kirim Esti surat

Budi send Esti letter
 ‘Budi sent Esti a letter’

These stems differ from the transitives already discussed in that the AV prefix does appear to give rise to true optionality:

(147) Budi ng-[k]irim Esti surat
 Budi AV-send Esti letter
 ‘Budi sent Est a letter’

In these cases, it would appear that if these are possible with a null allomorph of the AV, the availability of that allomorph does not rule out the appearance of the AV prefix, they are simply in free variation. So now again we must rule out the sort of theta-role altering function we saw with the transitives above. I am going to argue again, that this is ruled out by some constraint inherent to this class of stems.

4.1.3 Nominal Bases

So far we have seen that the AV prefix can correlate with the formation of verbs of all valences, but it only appears to correlate with certain types of observable argument structure and valence change. When appearing on otherwise verbal bases, this prefix derived agentive unergatives from unaccusatives, as well as transitives from unaccusatives. It also correlated with the derivation of agentive unergatives and transitives from their ungrammatical bare stems. In this section we will look at the behavior of the AV prefix when occurring on nominal bases. These bases are defined as nominal based on their behavior as bare stems under reduplication as described in Chapter 3. This categorization is further supported by their behavior with the suffix *-an*, but that is the subject of Chapter 5. Within the system of roots I have defined these are derived from $\sqrt{\text{thing}}$ roots. We can observe three overarching patterns, which are summarized in the following table:

Bare stem	with AV prefix
nominal	intransitive
nominal	transitive
nominal	*

It is apparent that the presence of the prefix is always correlating with observable change, however the options are constrained in ways that are highly analogous to those observed in the previous section. The AV prefix, as in the previous section, can only overtly derive intransitive verbs or transitive verbs. We do not see the change from noun to ditransitive verb correlating with solely the introduction of this prefix. Unlike what we saw with verbal bases, the intransitives formed from nominal bases are necessarily non-agentive:

(148) *guyu* ‘laugh’
 Budi (rada) ng-guyu
 Budi (half) AV-laugh
 ‘Budi kind of laughs/laughed’

Again we know that these derived stems are intransitive due to their inability to surface with the OV prefix. We further know that *nguyu* laugh is non-agentive due to its compatibility with *rada*. The AV prefix can form transitives which fall into two main types: 1) *use* transitives 2) *make/prepare* transitives. Some examples of *use* transitive are given in (149-150):

- (149) *lim* ‘glue’
 Esti ng-lim buku-e
 Esti AV-glue book-DEF
 ‘Esti glues/glued the book (together)’
- (150) *sendhok* ‘spoon’
 Tuti ny-[s]endhok sup-e
 Tuti AV-spoon soup-DEF
 ‘Tuti spoons/spooned the soup’ (scoop with spoon)

These transitives are characterized by the fact that they all involve using the noun in question as some sort of instrument or tool. Though this appears to be very productive, there are many examples of nouns which cannot participate in this sort of derivation:

- (151) *garpu* ‘fork’
 a. *Esti ng-garpu pelem-e
 Esti AV-fork mango-DEF
 Int: Esti forks/forked the mango (pick up with fork?)
 b. Esti ng-anggo garpu-ne
 Esti AV-use fork-DEF
 ‘Esti uses/used the fork’

In cases where this mode of verbalizing is unavailable a paraphrase with the verb *anggo* ‘to use’ can be used in its place. It is unclear to me why a verb cannot be formed from *garpu* ‘fork’, however another analogous gap occurs with the word *drei* ‘screwdriver’, which is most likely a loanword from Dutch. This could have something to do with the unavailability of a denominal verb here, or it could simply be an accidental gap. We will see below and throughout that loanwords readily engage with these sorts of derivational processes so I am reluctant to give this as an explanation here.

The second type of transitive commonly formed through prefixation with the AV prefix is what I am calling a *make* or *prepare* transitive. These transitives denote an event of creating or in some way canonically treating the object described by the noun. Examples of this are in (152-154):

- (152) *jus* ‘juice’
 Budi n-jus alpokat-e
 Budi AV-juice avocado-DEF
 ‘Budi juices/juiced the avocado’ (make juice out of)
- (153) *tembang* ‘song’

- Esti n-[t]embang [insert specific song]
 Esti AV-song
 ‘Esti sings/sang [song]’
- (154) *sayur* ‘vegetable’
 Tuti ny-[s]ayur bayem
 Tuti AV-vegetable spinach
 ‘Tuti cooks spinach’ (make cooked/prepared, for a vegetable)

Unlike the ‘use’ transitives we discussed previously, the object of the ‘make’ transitive is in some way *created* by the action denoted by the verb. In the case of *sayur* ‘vegetable’ this is slightly more opaque, but the denominal still involves making a specific type of vegetable (namely a cooked one). The word *sayur* itself is like a generic term for vegetable.

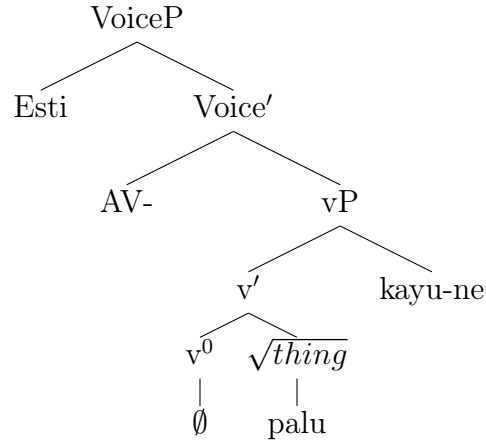
Given the transitive examples, it might be tempting to attempt to coerce a sort of ‘make’ usage from some intransitive de-nominals such as *nguyu* ‘to laugh’, which we have already seen. This could be construed as ‘produce/make a laugh’ and considered to be a non-agentive unergative⁷. I will argue based on the broader characterization of intransitive denominals that this is not the best way to think of these forms. This is due to the fact that some nouns appear to form unaccusative verbs which do not fall into any of the aforementioned categories, and which pose as outliers to the overarching generalization that the AV correlates at the very least with the introduction of an external argument (even if it is not agentive). Examples of such verbs are in (155) and (156).

- (155) *pati* ‘death’
 Tandur-an-e m-[p]ati
 plant-NOM-DEF AV-death
 ‘The plant dies/died’
- (156) *lenga* ‘oil’
 Tengen-ku ng-lenga
 hand-1.POS AV-oil
 ‘My hand is oily’

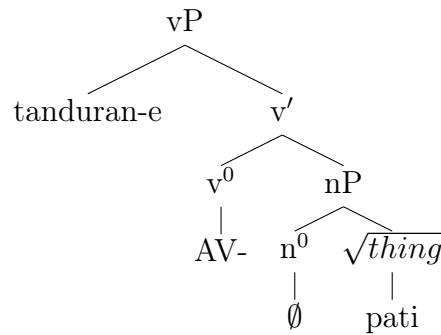
I will argue that as was seen in the previous section, we see an intransitive/transitive split in exactly the place where a verb is being derived from a non-verbal predicate versus a verb derived from a verbal predicate. Moreover, when the AV prefix affixes to a stem which is already categorized with a n^0 , the addition of the AV prefix derives an unaccusative. In the cases where we observe a ‘make’ or ‘use’ transitive, the AV is adding an external argument to a verbalized stem built from a $\sqrt{\text{thing}}$ root. This is essentially the equivalent to what was proposed for adjectives. :

- (157) transitive derived from a verbalized noun :
 Esti m-[p]alu kayu-ne
 Esti AV-hammer wood-DEF
 ‘Esti hammers the wood’

⁷Based off of Levin and Rappaport Hovav (1995) verbs of emission



- (158) unaccusative derived from an nP :
 Tanduran-e m-[p]ati
 plant-DEF AV-death
 'The plant died'



4.1.4 Interim Summary

To summarize what we have seen so far, I have proposed two homophonous versions of the AV prefix which carry out distinct derivational jobs and which exist in complimentary distribution. One variant of the prefix introduces an external argument to a stem which is already verbal, while the other verbalizes a non-verbal stem and derives an unaccusative. I have also argued for a 4 way distinction among roots:

- $\sqrt{property}$
- $\sqrt{experience}$
- \sqrt{event}
- \sqrt{thing}

The distinction between property and experience roots allows for us to account for the distribution of the AV when occurring on adjectival and unaccusative bases. Adjectival bases

built from experience roots, for my consultant, can only form zero derived adjectives and subsequently are only compatible with the verbalizing version of the AV. Verbal stems derived from experience roots become agentive unergatives when prefixed with the AV due to the nature of their subjects, and the fact that these subjects can be promoted to the external argument position introduced by the AV. All property roots which are compatible with external causation can verbalize directly and are subsequently compatible with the external argument introducing AV prefix.

Event roots are associated with stems which never grammatically surface as unaccusative, but instead are always agentive if verbal. This includes agentive unergatives like *mlayu* ‘run’, transitives like *mangan* ‘eat’, and ditransitives like *kirim* ‘send’. These stems are always directly categorized as verbs, and though they generally cannot surface grammatically without overt morphology, they always occur with the external argument introducing version of the AV. Thing roots can always be zero derived as nouns, however some thing roots can also be directly categorized as verbs and transitivize under prefixation with AV. Thing roots which are not compatible with direct verbalization become unaccusative verbs when prefixed with the AV.

4.2 The suffixes *-i* and *-ake*

The goal of this section is to repeat the methodology we applied to our examination of verbs derived using the AV prefix, but now with the consideration that voice prefixed forms can also serve as bases to the derivation. Given the analysis presented in the previous section for these forms, in conjunction with the lexical categories and root classes that have been established, we will consider a stem’s full derivational history when considering the role of suffixation. First the *-ake* suffix will be examined in detail, and then observations about the *-i* suffix which differ in meaningful ways to the distribution of *-ake* will be discussed in section (4.2.2).

4.2.1 Explaining the distribution of *-ake*

So far we have seen that the AV prefix performs the following two functions:

1. $Voice^0$: selects for a vP, introduces an external argument resulting in an agentive unergative or transitive
2. v^0 : selects for either nP or aP, verbalizes a non-verbal predicate resulting in an unaccusative

The distribution of the prefix and its resulting behavior is subsequently determined by the availability of a verbal or a non-verbal stem to which it can attach. Some roots necessarily categorize as adjectival or nominal where as some roots can categorize as verbal. This is summarized in the table below:

√ class	example √	cat. head	AV-prefix	prefixed example
property	<i>(biru</i> ‘blue’)	a	(IV: unaccusative)	(intransitive <i>mbiru</i> ‘get blue’)
	<i>pecah</i> ‘broken’	v	TV	<i>mecah</i> ‘to break X’
experience	<i>bundhet</i> ‘tangled’	a	IV: unaccusative	<i>mbundhet</i> ‘be tangled’
	<i>lara</i> ‘be sick’	v	IV: unergative	<i>nglara</i> ‘pretend to be sick’
event	<i>adeg</i> ‘to stand’	v	IV: unergative	<i>ngadeg</i> ‘to stand’
	<i>tulis</i> ‘to write’	v	TV	<i>nulis</i> ‘to write’
	<i>kirim</i> ‘to send’	v	DTV	<i>ngirim</i> ‘to send’
thing	<i>pati</i> ‘death’	n	unaccusative	<i>mati</i> ‘to die’
	<i>palu</i> ‘hammer’	v	TV	<i>malu</i> ‘to hammer’

The table above indicates what categorizing heads are necessarily compatible with each root type, and then what the result of prefixation with the AV subsequently yields. Just because a given root takes a categorizing head does not mean that the resulting form is grammatical when bare, that is not what is relevant here. What should be focused on is the relation between what categorizing head is available and the resulting AV prefixed derivative. Unaccusatives derived from adjectival property roots are in parenthesis since this is a point of interspeaker variation. For my consultant these roots *can* directly verbalize in addition to forming zero derived adjectives, so things like *biru* form transitives under prefixation. The range of AV functions listed above remains the same. Unergatives are derived from verbal experience roots, while transitives are derived from verbal property roots. I proposed that this is due to the role of +experiencer being attributed to the unaccusative subject in the case of experiencer roots, which allows this subject to serve as the external argument of the unergative produced with the AV, preventing an external argument from being base generated in the specifier of VoiceP and keeping the stem intransitive.

If we now expand our view of this system to include the AV prefix as well as the suffix *-ake*, we can observe the following additional derivatives⁸:

√ class	cat. head	AV-prefix	AV-ake	example
property	a	(IV: unaccusative)	(TV)	<i>mbirokake</i> ‘to make blue’
	v	TV	TV	<i>mecahake</i> ‘to cause to break’
experience	a	IV: unaccusative	TV	<i>mbundhetake</i> ‘to make tangled’
	v	IV: unergative	TV	<i>nglarakake</i> ‘to cause to feel hurt’
event	v	IV: unergative	*	*
	v	TV	DTV	<i>nulisake</i> ‘to write something for someone’
	v	DTV	DTV	<i>ngirimake</i> ‘to send something for someone’
thing	n	unaccusative	TV	<i>matekake</i> ‘to kill/cause to die’
	v	TV	DTV	<i>malokake</i> ‘to hammer something for someone’

This suffix always derives either a transitive or ditransitive stem where it is possible.

⁸There is no [n] initial allomorph of *-ake* to remedy vowel hiatus, instead there is an epenthetic glottal stop in careful speech. In causal speech the initial [a] is deleted. In writing which is modeled off of careful speech a ‘k’ is inserted between the root and the suffix to indicate the glottal stop.

The key facts we need to account for are cases where the AV and the addition of the suffix appear to both result in transitivization. This is because this suffix is clearly correlating with valence increasing from transitive to ditransitive in some cases, so we must rule out this valence increasing only some of the time. We will also have to explain cases where the suffix is appearing on stems which can be ditransitive on their own without the addition of suffixation, but I will turn to the transitive puzzle first.

Just from the table above we can see that a transitive is the result of *-ake* suffixation in almost all cases where the derivative is built from a non-*event* root. The one exception to this is the ditransitives derived from *thing* roots at the bottom of the table. Another way we could characterize the data in the table above is by looking at what stems are compatible with an unaccusative structure, somewhere in their derivational paradigm:

√ class	cat. head	bare unaccusative?	AV-prefix unaccusative?	AV-ake
property	a	*	(IV: unaccusative ✓)	(TV)
	v	✓	TV	TV
experience	a	*	IV: unaccusative ✓	TV
	v	✓	IV: unergative	TV
event	v	*	IV: unergative	*
	v	*	TV	DTV
	v	*	DTV	DTV
thing	n	*	unaccusative ✓	TV
	v	*	TV	DTV

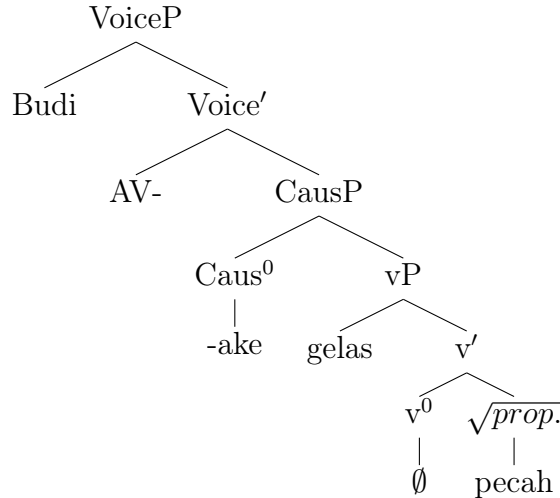
Every row with a check mark in the table above aligns with a row in which the *-ake* suffix derives a transitive. In each case there is an unaccusative present either as a bare stem or as the result of verbalization with the AV. I will argue that the key unifying factor among transitive *-ake* derivatives is this presence of an unaccusative structure.

I will argue that one of the functions of *-ake* is the following:

- *Caus*⁰ : select for an unaccusative vP, assign the role of causee to the unaccusative subject in the specifier of that vP

This suffix creates a causative by assigning the role of causee to an unaccusative subject in the specifier position of its complement. This is further illustrated in the tree below:

(159) Budi mecahake gelas ‘Esti breaks a glass’



The AV- prefix is required in these structures, it is in Voice and introduces the external argument as it did previously, but in this case that external argument is necessarily base generated. Given that the transitives produced through suffixation with *-ake* have this added element of functional structure and the denotation of causee we expect observable differences between AV- only transitives and *-ake* causatives. We do observe these:

- (160) a. Esti m-[p]ecah gelas-e
 Esti AV-break glass-DEF
 ‘Esti broke the glass’
 b. Esti m-[p]ecah-ake gelas-e
 Esti AV-break-ake glass-DEF
 ‘Esti broke the glass by accident’ / ‘Esti let the glass break’

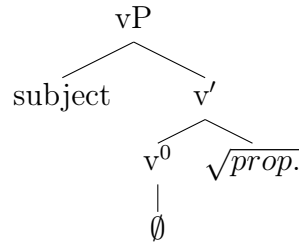
In the examples above the *-ake* causative is differentiated by an implicature that the causation was indirect. This is indeed cancellable but it is salient when no context is given to contradict the accidental nature of the breaking. Given the less direct nature of *-ake* causation, this also allows for agents which are less teleologically agentive to serve as subjects of *-ake* derivatives:

- (161) Angin-e gedhe m-[p]ecah-ake gelas-e
 wind-DEF big AV-break-ake glass-e
 ‘The strong wind broke the glass’

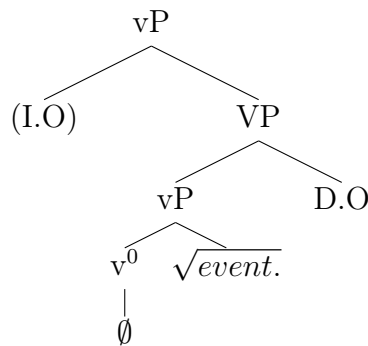
The reason why we see a ditransitive in the final row as opposed to a transitive is because though *thing* roots can be categorized directly with v^0 , the resulting structure is that of a transitive vP with a verb and its internal argument as the complement, just like we see in the cases of transitive event roots. As was argued in the previous section, the categorized vP structures we find with event roots are either argumentless unergative structures, or vPs with internal arguments. This is what disallows them from surfacing as zero derived stems, they are structurally deficient. So importantly, in exactly these cases that lack an unaccusative vP structure with a verb and an unaccusative subject, this is where we do not see *-ake* deriving a transitive. In order to account for the distribution of *-ake* that we observe I am going to argue that the structure instantiated by a transitive verb’s direct object is meaningfully

distinct from the structure instantiated by an unaccusative vP with an unaccusative subject:

(162) unaccusative vP:



(163) (di)transitive vP:

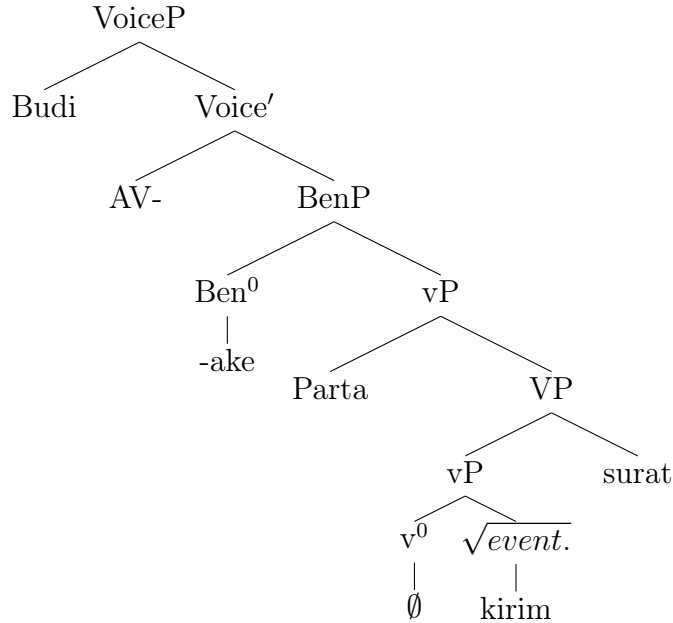


The structure above is meant to indicate that the same base structure is available in transitive and ditransitive cases, which is structurally distinct from unaccusative vP. Then there is variation as to the availability of a overt indirect object, as well as variation with respect to the need for that indirect object to be introduced by a prepositional phrase. Given these structures I am going to propose the following second function of the *-ake* suffix:

- Benefactee⁰ : selects for a transitive vP structure and assigns the role of benefactee to the indirect object which is base generated in the highest specifier position of vP

In cases where the base to suffixation is transitive this suffix also assigns a role to the object in the specifier position of its complement, however in this case that role is of *benefactee* and the argument is the verb's indirect object. This is illustrated in the tree below:

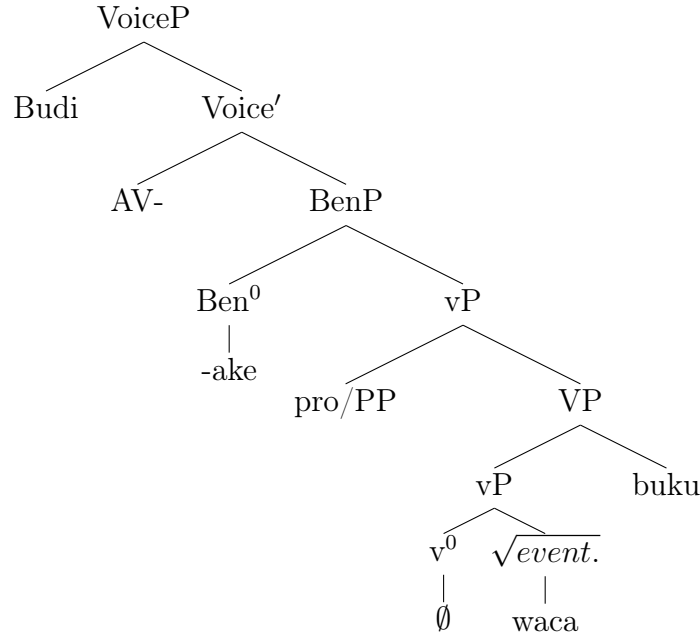
(164) *Budi ngirimake Parta surat* 'Budi sends Parta a letter (benefactive)'



As has been previously noted, one striking thing about benefactives derived using *-ake* is that the notion of a benefactee remains under object drop. Also striking is the fact that the benefactive relationship holds despite the fact that a true high applicative (with a double object word order) is not necessarily syntactically formed. This I am going to argue is because the suffix is not in and of itself creating any new structure or licensing the presence of any new arguments, it simply assigns the role of benefactee to an argument that is in the correct structural relation to it, whether that argument is overtly realized in that position or not. It must be a property of the verb itself that dictates which ditransitives are possible with the double object construction, and which require the indirect object to be introduced by a preposition, something that follows from the fact that ditransitives exist which require no overt morphology at all.

(165) *Budi macakake buku* ‘Budi reads a book (for someone)’

(166) *Budi macakake buku kanggo Esti* ‘Budi reads a book for Esti’



One remaining point to be addressed involves the fact that the existence of a transitive within the paradigm of a given stem, does not guarantee the presence of an *-ake* ditransitive. I have shown that *-ake* derives a transitive in every case where an unaccusative is possible somewhere in the broader paradigm of a given stem. This I have argued is because the *-ake* suffix attaches to an unaccusative vP whenever possible. This notion contains two key assumptions which I propose are general principles working in the Javanese lexicon:

1. Homophony is avoided in stems⁹
2. Morphemes which can select for more than one type of structure, select for the smaller one wherever possible

This is also predicated again on the notion that unaccusative vPs and transitive vPs are structurally distinct, with the latter containing inherently more structure. We can subsequently appeal to the above general principles, the first of which is observationally true of the lexicon, and the second of which builds in a bias towards deriving stems with fewer arguments, in order to account for the observed distribution of *-ake*. As an alternative we could have argued that transitives derived with the AV from unaccusative bases are structurally distinct, containing an unaccusative vP structure themselves which provides the base to the derivation for *-ake* transitives. However, the general principle above will regardless prove useful when accounting for the distribution of the *-i* suffix in the following section.

4.2.2 Explaining the distribution of *-i*

The *-i* suffix can be observed in many of the same contexts as the *-ake* suffix and subsequently I will argue that its core functions are the same. On the one hand this suffix derives transitives from unaccusative bases, attributing a specific role to the unaccusative subject, and on the

⁹We will see in the case of *-i* that there is necessarily homophony among *-i* suffixes

other it attributes a specific role to the indirect object of a ditransitive. The role in both of these cases is that of a locative goal:

1. Selects for an unaccusative vP and denotes that the subject is a locative goal
2. Selects for a (di)transitive vP and denotes that the indirect object of a ditransitive vP is a locative goal

These functions are illustrated in the examples below:

- (167) Esti n-[t]iba-ni jobin-e
Esti AV-fall-i floor-DEF
'Esti falls on the floor'
- (168) Budi ng-uncal-i Esti bal-e
Budi AV-throw Esti ball-DEF
'Budi throws Esti the ball'

It is true that in some cases the locative nature of the role attributed to an unaccusative subject is obscured, appearing to generate something that is simply causative like. The general function of the morpheme remains the same, though some idiosyncrasy must be built into the determining of what role is attributable to the subject for a given verb¹⁰.

- (169) Esti m-bundhet-i tali-ne
Esti AV-tangled-i rope-DEF
'Esti tangles the rope'

There are two things that set the distribution of the *-i* suffix apart from that of *-ake*. First there is the fact that the suffix can derive a locative version of a transitive from a clearly pre-existing transitive:

- (170) a. Esti n-[t]ulis surat
Esti AV-write letter
'Esti writes a letter'
- b. Esti n-[t]ulis-i kertas-e
Esti AV-write-i paper-DEF
'Esti writes on the paper'

Second this suffix is in competition with a homophonous suffix that denotes pluractionality and which does not affect argument structure:

- (171) Budi m-[p]ecah-i gelas
Budi AV-break-i glass
'Budi breaks glasses (repeatedly)'

¹⁰For a more in depth description of the range of usages observed with *-i* see Vander Klok and Evans (2022), Hemmings (2013), and Nurhayani (2014), though it can be noted that not all of the data contained within these resources was replicated with my consultant for her particular dialect of Javanese. More work is needed to understand the source of this variation.

- (172) Esti ng-adol-i buku
 Esti AV-sell-i book
 ‘Esti sells many books/ sells books repeatedly’

I take both of these patterns as evidence that, when there is a potential derivation with *-i* that does not necessitate a ditransitive, these are preferred. In these cases there is not a structural alternative to consider as there was with *-ake* transitives, which could have been derived from transitives that contained an unaccusative vP. In at least (171) and (172) above, the base to the derivation is necessarily a transitive with a vP structure which has the internal argument as the complement to the verb as opposed to an unaccusative subject. In each case the ditransitive variant is ruled out. It is also not the case again particularly in () and () that a locative ditransitive is inconceivable on semantic grounds. I am going to argue that the version of *-i* which denotes that an internal argument of a ditransitive is a locative goal, is ruled out if one of these other *-i* uses is available.

So given the facts above we must expand the original characterization of *-i*:

- Locative *-i* :
 1. Denotes that the subject of an unaccusative vP is a locative goal
 2. Denotes that the direct object of a transitive vP is a locative goal
 3. Denotes that the indirect object of a ditransitive vP is a locative goal
- Pluraction *-i* : creates a pluraction verb without impacting argument structure

The list above indicates that the result of an *-i* derivative is contingent on the size of the vP that this suffix selects for, and again there is a general principle at work which biases the derivation to favor the smallest possible input for a given stem. If a stem contains within its paradigm an unaccusative vP, this is the form that is selected for. The need for this suffix to occur with overt voice morphology and the introduction of an external argument means that in such cases a transitive is the result. In the case of an already transitive or ditransitive input no argument is actually added, only a relationship which denotes that the relevant argument is a locative goal. Finally, if a locative transitive is not possible, a pluractional verb is derived in favor of a ditransitive verb. What rules out a pluractional and rules in a ditransitive remains somewhat mysterious, though I think there is support for the claim that ditransitives which surface with *-i* are always to a certain extent syntactically ditransitive and subsequently if the semantics is also compatible with the notion of a locative indirect object, the suffix surfaces as such. These are also the contexts where a true applicative word order is most readily achieved as in (173):

- (173) Parta m-[w]eneh-i Tuti oleh-oleh
 Parta AV-give-i Tuti RED-gift
 ‘Parta gives Tuti a gift’

4.3 What about valence decreasing?

We have looked at several strategies employed by Javanese to increase the valence of a stem. This section briefly looks at a process of valence decreasing. Much additional work is needed to fully understand this process, but I think it could ultimately provide further insight into the nature of derived stems, serving as an additional tool for tracking how argument structure can be altered in this language. This section can serve as a jumping off point for that future work, situated within the analysis proposed more broadly here.

The *ke-* prefix has been described as an adversative passive in the literature (Davies 1995) but Nurhayani (2015) points out that this label is somewhat misleading. The *ke-* prefix more accurately forms an ‘accidental’ or non-agentive passive construction. This is what differentiates it from OV marking which only instantiates a passive word order but does not otherwise impact the nature of the syntactic subject:

(174)

- a. Budi ny-[c]olong buku-ne
Budi AV-steal book-DEF
‘Budi stole the book’
- b. buku-ne di-colong dening Budi
book-DEF OV-steal by Budi
‘The book was stolen by Budi’

(175) buku-ne ke-colong
book-DEF ke-steal
‘The book got stolen’

In the example above we can see the object of *colong* ‘to steal’ being fronted and the thief in question is basically left unspecified. It is in this way that this construction is seen as valence decreasing, an otherwise transitive verb is surfacing with only the direct object as a sort of undergoer of a stealing event carried out by an unnamed perpetrator. This is why the result is most readily translated with ‘got’. This construction is available for a subset of verbs which can function grammatically as transitives with only voice (AV or OV) morphology:

(176) *pecah* ‘broken’ → *kepecah* ‘get broken’

(177) *palu* ‘hammer’ → *kepalu* ‘get hammered (by mistake)’

(178) *waca* ‘to read’ → *kewaca* ‘get read (by mistake)’

(179) **bunder* ‘round’

(180) **biru* ‘blue’

In general the process seems unavailable for transitives derived from property roots which form zero derived adjectives as opposed to zero derived verbs, but it is widely possible with denominal transitives, transitives formed from verbal property roots, and stems which are always at least transitive such as *waca* ‘to read’.

This prefix is not possible with either the suffixes *-ake* or *-i*, however there is another construction used in cases of valence decreasing from ditransitive to transitive: *ke-√-an*.

- (181) Budi *ke-adol-an mobil-e*
 Budi *ke-sell-an car-3.poss*
 ‘Budi got is car sold (by someone)’

In the above example, the AV equivalent ditransitive stem would contain the *-ake* or *-i* suffix to denote the presence of an indirect object. Here it is simply understood that there is some unnamed seller who did the selling of the car on Budi’s behalf. There is also a usage of *ke-√-an* that demotes the subject of specifically locative transitive stems. In this case the fronted object is essentially a locative goal. In these cases the equivalent locative AV stems contain the suffix *-i*:

- (182) a. Esti *n-[t]eka-ni omah-e Budi*
 Esti AV-arrive-*i* house-3.poss Budi
 ‘Esti arrived at Budi’s house’
 b. *ketekan* ‘to have someone show up unannounced/ get arrived at’
- (183) *ketiban* ‘get dropped/fallen on’

So in both of the above cases it is puzzling that the equivalent AV stem would necessarily contain an additional derivational suffix to introduce the relevant argument and relation to the verb. It is unclear to me why something like ‘get dropped’ is unavailable in (183). Apparently if *-an* is present on a demoted transitive then only the locative interpretation is possible. Given the fact that the demoted ditransitive is not necessarily benefactive in (181), we could argue that in both transitive and ditransitive cases the relevant suffix is *-i*. We have discussed previously that the *-i* suffix requires the AV or OV to surface, so maybe in cases where *ke-* is present there is a reflex wherein *-i* must surface as *-an* for morphological well-formedness reasons.

Chapter 5

Deriving Nouns

Nominalization is a widely pervasive and productive process in Javanese. The most common overt nominalizing process is suffixation with the morpheme *-an*; this will also be the main focus of the current chapter. The *-an* suffix never co-occurs with voice prefixation and subsequently never appears with any of the other morphemes associated with valence increasing and verbalizing. The suffix can however co-occur with reduplication, with a variety of meanings contingent roughly on the lexical status of the reduplicated base.

We will apply the same methodology used for verbalizing morphology in Chapter 4 in that we will assume certain features about the input to the derivation, including lexical category and root class, then we will see how these features relate to the output of the derivation. A full examination of all nominalizing strategies is beyond the scope of the present investigation, however, we will observe how the underlying structures presented for verbal stems of various kinds presented in Chapter 4 help us to predict the meaning of nominalizations built with the suffix *-an*, something that was previously thought to be too idiosyncratic to predict in any meaningful way.

5.1 The suffix *-an*

As was stated above the *-an* suffix never co-occurs with overt verbalizing morphology, this includes the voice prefixes as well as the suffixes *-i* and *-ake*. The suffix can co-occur with reduplication and with the definite or possessive suffix(es)¹. Nothing ever occurs between the root and the suffix and two different strategies exist to prevent vowel hiatus when the suffix occurs on something that is vowel final. The suffix can trigger an [n] initial allomorph as is also seen with *-i* and *-e* suffixation. This allomorph triggers vowel lowering for high vowels, as is observed with the other suffixes as well:

(184) *waca* ‘(to) read’ → *waca-nan* ‘material for reading’

(185) *tuku* ‘(to) buy’ → *tuko-nan* ‘something that has been bought’

¹*-an* is also observed with the valence decreasing morpheme *ke-*, but it is unclear if this usage should be conflated with the nominalizing usage of *-an* in any meaningful way.

Alternatively, the *-an* suffix can trigger vowel deletion in order to prevent hiatus. The vowel lowering process occurs in these cases as well, and the initial [a] of the suffix is deleted:

(186) *tuku* →*tukon*

In many cases the deletion strategy appears to be preferred or at the very least more common. Typically, when both options are available there is no observed contrast in meaning, but exceptions to this exist and are mentioned if relevant.

Throughout the following section I will take as a starting point the lexical status of the bare stem to suffixation as defined in Chapter 3 as well as the root classes presented in Chapter 4. As has been observed, derivational processes can generally apply to stems of all kinds and implicational relationships arise only when many trans-paradigmatic factors are taken into account. Nominalization is no different, and the lexical starting point of the base will be used to anchor the broader discussion.

5.1.1 Verbal (and adjectival) Bases

Uhlenbeck (1978) in his description of nouns derived from verbal basis proposed the following three possible meanings attributable to verb derived *-an* forms:

1. **object** produced by the action indicated by the verb
2. **result** of the action indicated by the verb
3. **instrument** used in the action indicated by the verb

He also states that "It cannot be predicted which of the three semantic aspects will be present within the meaning of an *-an*-form. This seems to depend on the semantic nature of the verb basis." (Uhlenbeck 1978 p.156-157). If we follow his intuition that the resulting derivative is related to the nature of the base to which it attaches, and utilize our deepened understanding of the content of those basis I am going to argue that we can indeed predict the meaning of the resulting base, at least to a fairly high degree.

Within our system we have three root types which lexicalize as verbs, and which crucially cannot form grammatical bare nouns:

- √property
- √experience
- √event

Property and experience roots also can form zero derived adjectives, and though it is true that some √thing roots can directly verbalize and surface with the AV prefix, I am putting those aside for the present moment because they necessarily also exist as nouns without

overt derivation.

Looking first at $\sqrt{\text{event}}$ roots, we can see that the *-an* suffix is readily used to create nouns from stems that are otherwise transitive or ditransitive verbs. Importantly, these are not verbs which readily surface as intransitives when bare. I will turn to event roots which can surface as agentive unergatives shortly.

For stems of this kind when occurring with *-an*, the meaning follows straightforwardly from the relationship instantiated between the verb and the direct object. In the case of transitives, verbs either indicate some type of *production* or some type of *consumption*:

- (187) Consumption
- a. Esti m-[p]angan tela
Esti AV-eat sweet.potato ‘Esti eats sweet potato’
 - b. *panganan* ‘food’

In (187) above, the direct object of the verb *mangan* ‘to eat’ is the item which is consumed by the process of eating. The nominalization *panganan* is subsequently equivalent to ‘food’, a generic noun indicating the type of thing consumed when eating. We can contrast this with a verb of production like *mbangun* ‘to build’ :

- (188) Production
- a. Tuti m-bangun omah
Tuti AV-build house
‘Tuti builds a house’
 - b. *bangunan* ‘a building’

The verb meaning ‘to build’ indicates the creation of a built object, and the nominalization *bangunan* again is equivalent to a generic noun which indicates the result produced by the act of building, namely a ‘building’. Similarly we can see that in the case of ditransitives, the nominalization is still related to the would-be direct object of the verb in question:

- (189) a. Budi ng-[k]irim surat neng Esti
Budi AV-send letter PREP Esti
‘Budi sent a letter to Esti’
- b. *kiriman* ‘something that is sent, eg. a package’

These uses most closely align with the "object" and "instrument" meanings described by Uhlenbeck, wherein the *an*-form constitutes an actual object that is brought about by or integral to the performance of the action indicated by the verb. These derivatives also exhibit truly nominal behavior in that they are compatible with pluralizing reduplication, as well as possession with *duwe* ‘to have’:

- (190) *bangunan-bangunan* →plural buildings
- (191) Esti duwe *panganan-panganan*
Esti have RED-eat-an

‘Esti has some foods’

We can also note that the *production* and *consumption* difference is basically identical to the difference observed of denominal transitives in Chapter 4. It appears as though transitives more broadly can be divided in this way, and that whatever mechanism is responsible for instantiating this relationship between the verb and direct object, is also present in the resulting nominalization.

Given the fact that the *-an* forms above are necessarily related to the nature of a verbs internal argument, it raises the question: what happens in unergative cases where there is arguably no internal argument? In these cases it is tempting to say that the *-an* derivative can get an "object" or "production" interpretation equivalent to an instantiation of the action performed by the verb itself as in:

(192) *mlompat* ‘to jump’ → *lumpatan* ‘a jump’

This is reminiscent of zero-derived deverbal nouns in English like "a run" or "a dance". Similarly, we could propose that the *-an* derivative can receive a meaning similar to the "instrumental" or "consumption" interpretation as in (193):

(193) *lungguh* ‘to sit’ → *lungguhan* ‘a seat/place to sit’

However, it is unclear how these meanings would be consistently derived for these forms. Unergatives in Javanese, I have argued, are truly intransitive and so they lack the material contained in the transitives above which is capable of denoting the relation between the verb and a would be object. Moreover, we could have just as easily construed the meaning in (193) above to be in some way locative as opposed to instrumental.

We should expect the nominalizations derived from unergative stems to be more heterogeneous in their range of potential meanings. The lack of predictability that Uhlenbeck discussed should be more concentrated here, because there is less information contained in the verbal base to delimit the output to suffixation. I think this expectation is borne out:

(194) *sendhen* ‘to lean’ → *sendhenan* ‘something used for leaning’

(195) *mlayu* ‘to run’ → *playonan* ‘a running based game’

(196) *ngadeg* ‘to stand’ → *adegan* ‘a scene in a play or opera’

The above list is far from a comprehensive survey of the lexicon, but it still demonstrates the point that a greater degree of heterogeneity is exhibited (expectedly) by these stems. Another expectation of the current system, is that unaccusative intransitives which contain the structure of a low unaccusative subject internal to the vP, should exhibit less heterogeneity. Again this is because there is more information contained within the nominalization which can serve to constrain the potential meaning. We turn to *-an* forms which are arguably derived from such stems next. There are two possible meanings observed with derivatives built from stems of this kind. First there is something akin to Uhlenbeck’s "result" meaning, in that some derivatives form a nominalization which denotes the property an object would

have after having undergone the action denoted by the verb. These are in (197-198-184) below:

(197) *pecah* ‘broken’ → *pecahan* ‘(of something) broken or fragmented’

(198) *ajer* ‘melted/liquid’ → *jer-jeran*² ‘(of something) molten or dissolved’

It should be noted that these stems need a head noun in order to surface grammatically, they do not form plurals under reduplication and they cannot be the object of a ‘have’ construction with *duwe*:

- (199) a. **Esti duwe pecah-an*
Esti have broken-an
Intended: Esti has a broken thing
b. *Esti duwe pecah-an gelas*
Esti have broken-an glass
‘Esti has a broken/fragmented glass’

The other type of *-an* derivative we see equates to something like a noun which has an individual level property of the type denoted by the verb:

(200) *tiba* ‘to fall’ → *tiban* ‘fallen’

(201) *luwe* ‘hungry’ → *luwenan* ‘a (constantly) hungry person’

(202) *adhem* ‘cold’ → *adheman* ‘a cool/shady place’

In all of the examples above the property described in the *-an* derivative is a defining characteristic of the noun in question. For something like ‘hungry’ which is a state that animates experience it describes a person who is constantly hungry. For something like ‘cold’ this is used to denote a place that is cooler in temperature. The property of being ‘fallen’ in (200), is used to describe things like rainwater or on a more abstract level it can refer to structures which are temporary (equivalent to the English notion of a “pop-up shop”).

It should be noted that the distinction between the above meanings falls along the lines of the property vs. experience root distinction presented in Chapter 4. Property roots which form zero derived unaccusative verbs, when affixed with *-an* derive a “result” type participial construction. Experience roots which form zero derived unaccusatives, when affixed with *-an* form these derivatives which exhibit similar participial behavior but which carry this individual level property interpretation as opposed to a result interpretation. These also require an overt head noun and cannot be objects in a ‘have’ construction with *duwe*.

When it comes to property and experience roots which form zero derived adjectives, these form *-an* derivatives which are much more noun like:

(203) *bunder* ‘round’ → *bunderan* ‘a circle/roundabout’

²Vowel initial roots when suffixed with *-an* lose their initial vowel and get reduplicated due to being mono-syllabic, this is a morphotactic process and not a derivational one.

(204) *bundhet* ‘tangled’ → *bundhetan* ‘a tangle/knot’

These forms can be pluralized with reduplication and generally used like bonafide nouns. In terms of meaning they seem to denote an object which is fully defined by the property denoted by the stem, as opposed to something that simply possesses the property denoted by the stem.

To summarize what we have seen so far, it is apparent that though the *-an* suffix correlates with a wide range of derivative meanings, these seem strongly related to the nature of the base to which *-an* is attaching, as Uhlenbeck stated. Counter to Uhlenbeck however, I do think we can largely predict what the meaning of a given derivative will be now that we have a more nuanced understanding of the bases in question.

5.1.2 Nominal Bases

Little has been said previously about the relationship between *-an* suffixation and stems which themselves can function as nouns. Some descriptions exist in the literature but little else (see: Robson 2014, Uhlenbeck 1978). The first thing we notice if we look at the relationship between items that function as bare nouns and their behavior under *-an* suffixation, is the fact that the data appears again to be heterogenous. One set of derivatives appear to denote a smaller or imitation version of the affixed noun, something like a diminutive:

(205) *langit* ‘sky’ → *langitan* ‘ceiling’

(206) *gunung* ‘mountain’ → *gunungan* ‘small imitation of a mountain’ (puppet)

(207) *maca* ‘tiger’ → *macanan* ‘small imitation or toy tiger’

One thing that unites the base nouns in the set above is that these never co-occur with verbalizing morphology. That is not true for the next set which includes nouns which form intransitives under prefixation with the AV. Still, despite the existence of a verb somewhere in the paradigm of these stems, the resulting *-an* forms have widely varying meanings. One case where the *-an* derivative seems like it must be related to the verbal form in some way is with *layangan* ‘kite’. This seems distantly related to the base *layang* ‘letter’, but it is very fitting if we consider the existence of *nglayang* ‘to glide’. The

(208) *sepur* ‘train’ → *sepuran* ‘train ride/ticket’

(209) *layang* ‘letter, piece of mail’ → *layangan* ‘a kite’

(210) *guyu* ‘laughter’ → *guyonan* ‘a joke’

We see a similar level of idiosyncrasy with nominal bases which only verbalize upon suffixation with either *-ake* or *-i*, so again nouns which do engage with verbalizing morphology but which cannot transitivize directly with only voice morphology. Given the analysis presented in Chapter 4 for the derivation of *-ake* transitives, it makes sense that these pattern with the nouns above because they necessarily verbalize via the same mechanism. This follows if we assume that the nominalizations here are derived from verbalized forms of the base nouns in

question. I argue that this is the best option given that it unites the behavior of nouns like those above and *benik* ‘button’ below, and it helps to rule out a diminutive reading of these nouns which would otherwise be very difficult to explain.

(211) *benik* ‘button’ → *benikan* ‘something with buttons on it (type of clothing)’

It is worth noting that a diminutive version of nouns of the above type do exist, but they require reduplication as in (212). Arguably the nominalizing suffix is attaching to a nominal base in this case hence the diminutive reading being available. We will return to this point in the section on reduplication and the *-an* suffix.

(212) *sepur* ‘train’ → *sepur-sepuran* ‘a toy train’

Finally, the third type of derivative we see are akin to the object nominalizations we saw derived from transitive verbs. This is because each of these nouns has a corresponding denominal transitive stem, which is arguably providing the base to nominalization here. The resulting *-an* form denotes the physical object which is a result of the action denoted by the verb. So for example from *sekop* ‘shovel’ we get the transitive *nyekop* ‘to shovel’, and subsequently the *-an* nominalization *sekopan* equates to the physical output of shoveling such as a pile of dirt that has been removed in the process of digging a hole. Further examples of this are given below:

(213) *gunting* ‘scissors’ → *guntingan* ‘cuttings/cut material’

(214) *sekop* ‘shovel’ → *sekopan* ‘shovelings/shoveled material’ (pile of dirt)

(215) *sapu* ‘broom’ → *saponan* ‘sweepings/swept material’

(216) *tali* ‘rope, string’ → *talenan* ‘something tied, the result of tying’

So just as in the cases of deverbal nominalizations above derived from otherwise transitive bases, the relation between the verb and direct object dictates the meaning of the nominalization that we find upon suffixation with *-an*.

One thing to note regarding the data we have seen throughout this section, is that it runs counter to the broader principle proposed in Chapter 4 regarding a preference for the smallest derived structure possible. In the cases with *-an*, if that principle held we would arguably expect to see only diminutive forms, as opposed to nominalizations containing varying amounts of verbal structure. I am going to argue that the drive for small structures holds when lexical category is not in question. If you are deriving a verb from another verb, yes the principle stands. However, if the option exists to derive a stem of a different lexical category then that is preferred in order to increase the range of potential derivatives for a given stem’s inventory. This is why we see such a prevalence of deverbal nouns, despite a noun based nominalization being theoretically possible. Still, we expect these deverbal nominalizations to still build from the smallest amount of verbal structure possible, which drives some of the variation in meaning and behavior that we also see amid these forms.

5.1.3 A note about other nominalizing strategies

Though a full survey of the other nominalization strategies used in this language is beyond the scope of the present investigation, the following highlights some key features of these other processes that warrant further study. First, there is a highly productive process of zero-derived nominalization that occurs specifically with verbs. This process is observable through the presence of possessive morphology on the verb and requires overt voice prefixation when such prefixation is otherwise necessary for the grammatical use of that verb:

- (217) *tulis* ‘to write’
- a. Esti *(n)-[t]ulis surat
Esti AV-write letter
‘Esti writes a letter’
 - b. *(n)-[t]ulis-e Esti
AV-write-POSS Esti
‘Esti’s (way of) writing’

Zero nominalization is distinguishable from *-an* nominalization in this respect, as well as in the fact that zero nominalizations remain more verbal in general. Such nominalizations consistently denote the manner in which the subject, possessor, carried out the action described by the verb. Moreover, it is possible to refer to this manner in a temporal way, which is not possible with possessed *-an* nominalizations:

- (218) *Tangis-an-e Esti ora mandheg
cry-an-POSS.3 Esti NEG stop
Intended: Esti’s crying won’t stop
- (219) N-[t]angis-e Esti ora mandheg
AV-cry-POSS.3 Esti NEG stop
‘Esti’s crying won’t stop’

I stated in the previous section that the deverbal nominalizations seem to vary with respect to certain properties of nounhood. The deverbal nouns formed from transitives behaved as other canonical nouns in the language, but those derived from intransitives did not. These forms which occur with possession and overt AV morphology appear arguably even less noun-like, something like a gerund. More work is needed to understand how these observations correlate to the underlying structures present on these various derivatives.

Another strategy in need of further investigation involves reduplication. Reduplication as has been noted previously can also derive diminutive or “toy” forms of certain nouns derived from *thing* roots which have also been suffixed with *-an*. An example of this is given below:

- (220) *alas* ‘forest’ → *alas-alasan* ‘a small/imitation forest’

However, this combination of morphemes does not always derive this,

- (221) *kucing* ‘cat’ → *kucing-kucing-an* ‘to play a cat and mouse type game’ (reciprocal)

(222) *sapu* ‘broom’ → *sapon-saponan* ‘to sweep at each other’ (reciprocal)

In the second case above derived from *sapu* ‘broom’, we know that there does exist a transitive verb *nyapu* ‘to sweep’, so we can imagine that the potential for this verb is somehow making the reduplicated form above available. The same can not clearly be said for *kucing* though. It is true that there is an intransitive *ngucing* ‘to flex’, but it is difficult to see how this lends itself to a game of the nature described by the reduplicated form. Furthermore we see cases where a denominal verb exists for a given noun, and yet the reduplicated form with *-an* is diminutive as it is for nouns which do not interact with verbalizing morphology:

(223) *sepur* ‘train’ → *sepur-sepuran*

We saw above that the alternative diminutive form of *sepur*, which would be *sepuran* without reduplication is ruled out in favor of the deverbal noun meaning ‘train ride/ticket’. This could indicate that the reduplicated form with *-an* is some sort of homophony avoidance strategy. Given that I have argued in favor of a broader principle of homophony avoidance for this language, such strategies would follow from this. It could be that the nature of *nyepur* as an intransitive rules out a reciprocal derivative on semantic grounds as well. Otherwise we would still have to explain why the diminutive takes precedent over a reciprocal derivative for *sepur*, while it does not do so for *sapu*, since in both cases the *-an* only form is not a diminutive.

In other words, we could say that for *-an* derivatives that form deverbal nouns as opposed to diminutives, a diminutive *could* be made with the addition of reduplication in order to avoid homophony. However, this diminutive is only available if a reciprocal form is not available. More work is needed to better understand the interplay between these two processes, but this is a possible starting point that fits within the broader system I have proposed.

Chapter 6

Concluding Remarks

The present investigation set out to understand two core features of the Javanese lexicon:

1. The same derivational morphemes are compatible with stems of all types
2. Derivational morphemes appear to be multipurpose, performing multiple functions

The two above features serve to obscure the relationship between bases and their subsequent derivatives, and so ultimately this investigation sought to better understand that relationship. This was achieved through an in depth investigation of every stage of the derivation, from the properties of zero derived bases through the affixation of a variety of derivational morphemes. From this process I asserted that Javanese does benefit from the notion of lexical category, including the category of adjectives which previously were described as virtually indistinguishable from intransitive verbs. I further asserted that Javanese roots, like English roots, should be considered to have content and I presented an inventory of root types or classes:

- $\sqrt{\text{property}}$ eg. *pecah* ‘broken’
- $\sqrt{\text{experience}}$ eg. *adhem* ‘cold’
- $\sqrt{\text{event}}$ eg. *pangan* ‘eat’
- $\sqrt{\text{thing}}$ eg. *palu* ‘hammer’

This inventory allowed us to predict the places where the AV prefix derived unaccusatives from non-verbal predicates vs. where it derived transitive verbs and subsequently informed our understanding of the distribution of the suffixes *-ake*, *-i*, and *-an*. I also argued that there is a meaningful structural distinction between vP structures which contain an unaccusative subject vs. those which contain an internal argument in object position, as this again allowed us to account for the distribution of suffixal morphology. Beyond this structural distinction I proposed that several general principles inform the shape of the Javanese Lexicon:

- An avoidance of homophony
- A drive to derive the smallest structure possible

- A drive to derive a stem of a different lexical category whenever possible, which overrides the drive for small structures

These principles create a push and pull between the desire to increase the number and variety of derivatives for the paradigm of a given stem, and interests of economy which I propose can be likened in spirit to Dispersion Theory (Flemming 2004). Dispersion Theory asserts that the size and shape of a phoneme inventory is governed by a push and pull between the desire to increase the number of potential contrasts and a desire to keep contrasts sufficiently distant in perceptual space.

Some aspects of the Javanese derivational system that remain puzzling include areas where a given morpheme is unavailable for no obvious structural or semantic reason. It is true that accidental gaps occur amid derivational systems, but it is as yet unclear if we can attribute all the places where the AV prefix fails to occur as such accidents. Moreover, the factors that determine the distribution of the pluractional variant of the *-i* suffix remain mysterious. In general this is a language with a rich and productive derivational system, and issues remain to be addressed in every corner of it which I leave to future work.

References

- Alexiadou, A. & T. Lohndal. On the division of labor between roots and functional structure. In R. D'Alessandro & al. (eds) *The verbal domain*. Oxford University Press, 85-104.
- Arad, Maya. 2003. "Locality constraints on the interpretation of roots: the case of Hebrew denominal verbs," *Natural Language & Linguistic Theory* 21: 737–78.
- Carlson, Gregory N. 1977. *Reference to kinds in English*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Cole, Peter & Son, Min-Jeong. 2004. The argument structure of verbs with the suffix -kan in Indonesian. *Oceanic Linguistics* 43(2). 339–364.
- Cole, Peter, Yurie Hara, and Ngee Thai Yap. 2008. Auxiliary Fronting in Peranakan Javanese. In *Journal of Linguistics* 44: 1043.
- Connors, Tom J. 2008. *Tengger Javanese*, Yale University: PhD.
- Davies, W.D. 1995. Javanese adversatives, passives and mapping theory. *Journal of Linguistics*, 31, 15-31.
- Dowty, David. 1979. *Word Meaning and Montague Grammar: The Semantics of Verbs and Times in Generative Semantics and in Montague's PTQ*. Dordrecht: Reidel.
- Flemming, Edward. 2004. Contrast and perceptual distinctiveness. In Bruce Hayes, Robert Kirchner and Donca Steriade eds., *Phonetics in Phonology*, Cambridge: Cambridge University Press.
- Halle, Morris, and Marantz, Alec. 1993. Distributed Morphology and the Pieces of Inflection. In *The view from Building 20*, eds. Ken Hale and Samuel Jay Keyser, 111-176. Cambridge, Mass: MIT Press.
- Harley, Heidi 2005. "How do verbs get their names? Denominal verbs, manner incorporation and the ontology of verb roots in English," in Nomi Erteschik-Shir and Tova Rapoport (eds), *The Syntax of Aspect: Deriving Thematic and Aspectual Interpretation*. Oxford: Oxford University Press, 42–64.

- Hemmings, Charlotte. 2013. Causative and Applicatives: The case for polysemy in Javanese. *SOAS Working Papers in Linguistics* 16. 166-194.
- Horne, Elinor C. 1961. *Beginning Javanese*. New Haven: Yale University Press.
- Keeler, W. 1984. *Javanese: A cultural approach*. Athens, OH: Ohio University Center for International Studies.
- Levin, Beth, and Malka Rappaport Hovav. 1995. *Unaccusativity: at the syntax-lexical semantics interface*. Cambridge: MIT Press.
- Nurhayani, Ika. 2014. *A unified account of the syntac of valence in Javanese*. New York: Cornell University dissertation.
- Nurhayani, Ika. 2015. *javanese and Problems in the Analysis of the Adversative Passive*. *Jurnal Ilmial Masyarakat Linguistik Indonesia*. Volume 33.
- Ogloblin, Alexander K. 2005. *Javanese*. In *The Austronesian Languages of Asia and Madagascar*, eds. Alexander K Adelaar and Nikolaus P. Himmelmann. London: Routledge
- Poedjosoedarmo, G. 1986. *Role structure in Javanese*. Jakarta: Atma Jaya Press.
- Pylkkänen, L. 2002. *Introducing Arguments*. MIT Dissertation. Cambridge, MA.
- Rackowski, A. & Richards, N. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36: 565–599.
- Robson, Stuart. 1992. *Javanese Grammar for students*. Glen Waverley: Monash Papers on Southeast Asia.
- Robson, Stuart. 2014. *Javanese Grammar for students: 3rd Edition*. Glen Waverley: Monash Papers on Southeast Asia.
- Robson, Stuart, and Wibisono, Singgih. 2002. *Javanese-English Dictionary*. North Clarendon: Tuttle Publishing
- Sato, Yosuke. 2015. Argument ellipsis in Javanese and voice agreement. *Studia Linguistica* 69. 58-85.
- Suwadji. 1981. *Struktur dialek bahasa Jawa di Pesisir Utara Jawa Tengah*. Jakarta: Departemen Pendidikan dan Kebudayaan, Pusat Pembinaan dan Pengembangan Bahasa.
- Uhlenbeck, E.M. 1978. *Studies in Javanese morphology*. The Hague: Nijhoff.

Vander Klok, Jozina. 2012. Tense, aspect, and modality in Paciran Javanese, McGill University: Ph.D.

Vander Klok, Jozina & Bethwyn Evans. 2022. The evolution of non-syntactic functions of applicatives: *-i* suffixation in Javanese and neighbouring languages. In Sara Pacchiarotti and Fernando Zúñiga (eds.) *Applicative morphology: Neglected syntactic and non-syntactic functions*. [Trends in Linguistics] Berlin: Mouton de Gruyter. pp. 437-474.

Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca, NY: Cornell University Press.

Wijana, I Dewa Putu. 2021. Reduplication in Javanese. *Jurnal Ilmiah Masyarakat Linguistik Indonesia*. Volume 39.