DERIVING CATEGORICAL AND CONTINUOUS PROPERTIES OF JAVANESE SPEECH LEVELS*

Christopher Davis
University of the Ryukyus
cmdavis@grs.u-ryukyu.ac.jp

I propose an analysis of the Javanese speech level system in which (i) five classes of lexical items marked with distinct combinations of two binary speech level features combine syntagmatically to generate the three traditionally recognized speech levels (*Ngoko*, *Madya*, and *Krama*), and in which (ii) competition within particular paradigms of lexical alternatives generates continuous gradation within the Madya level. I connect these synchronic proposals to a diachronic account of the modern Javanese speech level system.

1. Overview

Javanese sentences are traditionally sorted into one of three speech levels, called *Ngoko*, *Madya*, and *Krama*. The choice between these three levels indexes the relationship between the speaker and addressee, with status, age, and intimacy the primary factors determining the choice between the levels (Wolff and Poedjosoedarmo 1982; Clynes 1989). The levels have the following canonical contexts of use:

- *Ngoko*: low status addressee, not older than the speaker, intimate relationship
- *Krama*: high status addressee, older than the speaker, non-intimate relationship
- *Madya*: a "halfway house" (Wolff and Poedjosoedarmo 1982, p.20) between Ngoko and Krama, canonically used in situations where the factors determining the choice of speech level are in conflict; for example, situations in which the speaker is talking to an addressee of higher social status but with whom the speaker is intimate.

Speech level contrasts are marked formally by alternant lexical items that are restricted as to which speech level they may occur in. The details of this system are spelled out in section 2, where I follow Clynes (1989) in arguing that the categorical distinction between Ngoko, Krama, and Madya speech levels can be derived by dividing such alternants into five distinct classes, characterized formally using two binary features. These classes show syntagmatic constraints on co-occurrence within a sentence, with conflicting feature values resulting in ungrammaticality. In section 3 I show how competition within particular paradigms of alternants generates gradient sublevels within the Madya level. On the basis of data from Wolff and Poedjosoedarmo 1982, I argue that paradigms of alternant items fall into ten distinct types, and that patterns of usage within sublevels of Madya are driven by different patterns of competition within these paradigm types. In section 4 I propose a diachronic account of the evolution of the contemporary speech level system, showing how the paradigm-based competition outlined in section 3 could seed reanalysis of individual lexical items, generating both new lexical classes and new paradigm types. Section 5 concludes.

*My thanks to the participants of AFLA 28, and in particular to Jozina Vander Klok, Mathieu Paillé, and Tamisha L. Tan for their helpful questions and discussion. I am indebted to Ibu Aisyah Alwi, Bapak Subiantoro, and Helvi Lystiani for their help with Javanese judgments. All errors are my own.

2. Categorical properties

The speech level of an utterance is signaled by the choice of lexical items belonging to classes which are themselves traditionally labeled in terms of the speech level with which they are canonically used.[1] These lexical items belong to sets of suppletive alternants with identical semantic content but differing in terms of the speech levels with which they are compatible. The most basic classification sorts such lexical items into either *ngoko* or *krama* classes.[2] For example, the verb 'eat' can be realized by either the ngoko alternant *mangan* or the krama alternant *nedha*. Similarly, the relativizer (glossed REL) can be realized using the ngoko alternant *sing* or the krama alternant *ingkang*. The Ngoko speech level is then signaled by the exclusive use of ngoko alternants, and the Krama speech level by the exclusive use of krama alternants, as illustrated by the following examples from Clynes 1989, pp.26–27, based on examples from Errington 1985:

(1) Mung Mas Poino iki sing teka dhèk wingi.
ngoko *ngoko!* *ngoko* *ngoko* *ngoko* *ngoko*
only older.male Poino here REL come time yesterday
'Only Mas Poino here came yesterday.' (*Ngoko* speech level)

(2) Kula saged mendhet pinten samenika?
krama *krama* *krama* *krama* *krama!*
1SG can take how.many now
'How many can I take now?' (*Krama* speech level)

The Madya level, meanwhile, is signaled by a mixture of ngoko and krama alternants. This is illustrated by the following example from Errington 1985, p.149, as cited and glossed by Clynes (1989, p.24):

(3) Kula isa tuku pinten
krama *ngoko* *ngoko* *krama*
1SG can buy how.many
'How many can I take?' (*Madya* speech level)

This simple taxonomy is complicated by the existence of what Clynes (1989) calls "style markers". First, among the ngoko and krama lexical alternants are a subset of items that are *not* compatible with the Madya level, a fact noted by Uhlenbeck (1970, pp.451–452). In other words, among the class of items traditionally labeled as ngoko there is a subset that is compatible *only* with the Ngoko speech level, while among those items traditionally labeled krama there is a subset that is compatible *only* with the Krama level. Such items are exemplified by *iki* 'here' in (1), and *samenika* 'now' in (2), neither of which are compatible with Madya speech level (unlike the other

[1] The following discussion is based primarily on chapter 1 of Clynes 1989. The system described is the one used by what Clynes terms "conservative 20th century" and "traditional" (19th century) aristocratic *priyayi* speakers associated with the court centers of Central Java. According to Clynes, the combinatoric rules that these two groups of speakers employ are essentially the same, with their language use differing only in "the social significance they attach to the madya style" (Clynes 1989, p.12).

[2] The terms *ngoko* and *krama* are thus ambiguous, referring in some instances to the *speech level* of an utterance, and in other instances, to classes of lexical items. I avoid this ambiguity by using an initial lower case when referring to lexical classes, and an initial upper case when referring to speech levels.

ngoko and krama items in these examples). They thus serve to unambiguously mark the sentences containing them as Ngoko or Krama, a fact I have marked with an exclamation point in the glosses. Second, there is a small set of lexical alternants that are *only* compatible with the Madya speech level.[3] For the examples above, there are madya alternants for the style markers in (1) and (2), which when used make these sentences unambiguously Madya speech level, despite no mixture of ngoko and krama lexical items elsewhere in the sentence, as seen in the following variants (also from Clynes 1989):

(4)

Mung	Mas	Poino	niki	sing	teka	dhèk	wingi.
ngoko			*madya!*	*ngoko*	*ngoko*	*ngoko*	*ngoko*
only	older.male	Poino	here	REL	come	time	yesterday

'Only Mas Poino here came yesterday.' (*Madya* speech level)

(5)

Kula	saged	mendhet	pinten	saniki?
krama	*krama*	*krama*	*krama*	*madya!*
1SG	can	buy	how.many	now

'How many can I take now?' (*Madya* speech level)

There are thus five distinct classes of lexical items that signal the speech level of an utterance, which following Clynes (1989) I treat in terms of two binary features, $\pm N$ and $\pm K$, as follows:[4]

(6) Lexical classes and their features

a.	strict krama (*Krama* level only)	$[-N, +K]$	eg. *samenika*
b.	non-strict krama (*Krama* or *Madya* level)	$[\quad +K]$	eg. *kula, saged,* etc.
c.	madya (*Madya* level only)	$[+N, +K]$	eg. *niki, saniki*
d.	non-strict ngoko (*Madya* or *Ngoko* level)	$[+N \quad]$	eg. *mung, sing,* etc.
e.	strict ngoko (*Ngoko* level only)	$[+N, -K]$	eg. *iki*
f.	neutral (compatible with all levels)	$[\quad]$	eg. *Mas*

There are six distinct lexical classes exemplified by the examples seen thus far; five of these impose restrictions on the speech level of the sentence, while the neutral class imposes no such restrictions. The traditional ngoko and krama classes are divided here into strict and non-strict subclasses, based on whether they can be used in the Madya speech level. The speech levels themselves are signaled by the combined features of a sentence, with Krama signaled by $-N, +K$, Madya signaled by $+N, +K$, and Ngoko signaled by $+N, -K$.[5]

The speech level of a sentence is calculated simply by looking at the speech level features

[3] Errington (1985, Table 4.1, pp.138–139) lists 36 such madya alternants along with their ngoko and krama counterparts.

[4] Clynes attempts to directly capture the meaning of the associated speech levels by using the features [±STATUS] and [±INTIMATE]. This reduction leads to difficulties that I leave aside for reasons of space.

[5] A sentence will thus be underdetermined as to its speech level in cases where the combined lexical items of the sentence fail to specify values for either or both of these features. This can be seen in the following variants of (1) and (2) in which the style markers have been removed. The former is $+N$ but does not specify a value for $\pm K$, while the latter is $+K$ but does not specify a value for $\pm N$. (Examples from Errington 1985, as cited and glossed in Clynes 1989, pp.26–27):

of all its constituents, which must obey the following combinatoric constraint:[6]

(7) Combinatoric constraint: The values of $\pm N$ and $\pm K$ must be consistent within a sentence.

This constraint is categorical, and leads to clear judgments of unacceptability when violated. This is illustrated on the basis of the following example from Clynes 1989, p.5, to which I have added feature values below individual lexical items in accordance with the discussion above:

(8)	Bu	Siti	sampun	nedha	ingkang	menika.	*Krama*
			$[-N,+K]$	$[+K]$	$[-N,+K]$	$[-N,+K]$	
	Bu	Siti	mpun	nedha	sing	niku.	*Madya*
			$[+N,+K]$	$[+K]$	$[+N]$	$[+N,+K]$	
	Bu	Siti	wis	mangan	sing	kuwi.	*Ngoko*
			$[+N,-K]$	$[+N,-K]$	$[+N]$	$[+N,-K]$	
	Bu	Siti	already	eat	REL	that	

"Bu Siti has already eaten that one."

The three variant sentences in (8) are distinguished by the speech level they are compatible with, and are otherwise semantically identical. There are three alternant lexical items meaning 'already', each of which is compatible with only one of the three speech levels. The alternant *sampun*, for example, is only compatible with the Krama level, while *mpun* and *wis* are only compatible with the Madya and Ngoko levels, respectively. The same situation holds for the lexical alternants meaning 'that'. Turning to the alternants meaning 'eat', *nedha* is compatible with both Krama and Madya levels, while *mangan* is compatible only with the Ngoko level. In the alternants for the relativizer REL, *sing* is compatible with Ngoko and Madya levels, while *ingkang* is compatible only with the Krama level.

All other combinations of the lexical alternants in (8) are (according to Clynes) ungrammatical,[7] which under the above account results from feature clashes in all other logical combinations (of which there are $3 \times 2 \times 2 \times 3 - 3 = 33$). The following ungrammatical variants of (8) (also from Clynes) illustrate, with clashing lexical features indicated (irrelevant features not shown):

(i) Mung Mas Poino ~~iki~~ sing teka dhèk wingi.
ngoko *ngoko* *ngoko* *ngoko* *ngoko*
only Mas Poino REL come time yesterday
'Only Mas Poino came yesterday.' (*Ngoko* or *Madya* speech level)

(ii) Kula saged mendhet pinten ~~samenika~~?
krama *krama* *krama* *krama*
1SG can take how.many
'How many can I take?' (*Krama* or *Madya* speech level)

[6] This constraint is spelled out by Clynes (1989, p.40) as "styles rule 1".

[7] Or, more theory-neutrally, unacceptable; whether the unacceptability should be treated as ungrammaticality or as something else (a semantic or pragmatic infelicity) depends on how we ultimately understand the categorical constraint against feature clashes.

(9) a. *Bu Siti wis nedha sing kuwi.
$-K$ $+K$ $-K$

b. *Bu Siti mpun nedha inkang niku.
$+N$ $-N$ $+N$

c. *Bu Siti sampun nedha inkang kuwi.
$-N, +K$ $+K$ $-N, +K$ $+N, -K$

(9a) has a clash in $\pm K$ features, (9b) a clash in $\pm N$ features, and (9c) a clash in both.

Table 1, from Clynes 1989, p.18, gives approximate counts of the number of lexical bases[8] belonging to different lexical classes, based on the number of forms in Poedjosoedarmo 1969 as a proportion of the approximately 20,000 bases in Horne 1974. The ngoko and krama counts in Table 1 include both strict and non-strict items. According to Clynes, the strict ngoko and strict krama class items (i.e. "style markers") include about 30 items each, leaving about 550 non-strict ngoko and krama items each. The vast majority of lexical items with non-trivial speech level features thus fall into the non-strict ngoko (i.e. $[+N]$) and non-strict krama (i.e. $[+K]$) classes.

	number (approx.)	% of lexicon
ngoko	580	3
krama	580	3
madya	30	$< 0.2\%$
deferential	210	1
neutral	c.20,000	93

Table 1: Lexical bases by class (from Clynes 1989, p.18)

The table includes two additional lexical classes, neither of which participates in the determination of a sentence's speech level. Items in the neutral class account for the majority of lexical bases, and are analyzed as lacking any speech level features. In addition to these unmarked bases, there are also a number of what Clynes calls *deferential* lexemes. These are traditionally divided into two sub-classes, called *krama inggil* (lit. 'high krama') and *krama andhap* (lit. 'low krama'). Despite these labels, the use of such items is orthogonal to the speech level system. Instead, they signal deference toward (*krama inggil*) or humilification of (*krama anddhap*) some grammatically determined referent,[9] thus corresponding to what have been variously termed "referent honorifics" (Comrie 1976), "propositional honorifics" (Harada 1976), "argument honorifics" (McCready 2019), and "content-oriented markers of politeness" (Portner et al. 2019). The speech level markers (i.e. items from the ngoko, madya, and krama lexical classes) are instead anchored directly to the context of utterance, signaling a relationship between the speaker and the addressee.

Items from the deferential class often occur as alternants to items that are themselves used to signal speech level.[10] For example, there is in addition to the two alternants for 'eat' (*mangan* and *nedha*) seen in (8) a third alternant form, *dahar*, which is traditionally labeled *krama inggil*. By

[8] The "base" is the root from which different lexemes are derived by various morphological processes; the actual number of words in each class is thus larger than these numbers suggest.

[9] As Uhlenbeck (1970, p.449) puts it, *krama inggil* items "enable a speaker to refer in a respectful way to third persons considered by him to occupy a higher social position than himself."

[10] Uhlenbeck (1970) states that there are also about 150 *krama inggil* items that occur as alternants with a neutral counterpart.

using this alternant, the speaker signals their deference toward the agent argument of the verb (the eater), but does not thereby signal anything about the sentence's speech level. This is illustrated by the examples in (10), from Clynes 1989, p.6,[11] which contain the deferential *dahar* 'eat' (marked DFR) in lieu of $[+N, -K]$ *mangan* or $[+K]$ *nedha* used in (8). They also contain the neutral item *lan* 'and' in the subject noun phrase.

(10)

Bu	Siti	*lan*	Bu	Marta	sampun	*dahar*	ingkang	menika.		*Krama*
					$[-N,+K]$	DFR	$[-N,+K]$	$[-N,+K]$		
Bu	Siti	*lan*	Bu	Marta	mpun	*dahar*	sing	niku.		*Madya*
					$[+N,+K]$	DFR	$[+N]$	$[+N,+K]$		
Bu	Siti	*lan*	Bu	Marta	wis	*dahar*	sing	kuwi.		*Ngoko*
					$[+N,-K]$	DFR	$[+N]$	$[+N,-K]$		
Bu	Siti	and	Bu	Marta	already	eat	REL	that		

'Bu Siti and Bu Marta already ate that one.' (+ the speaker honors Bu Siti and Bu Marta)

As can be seen, both the neutral and deferential items are compatible with all speech levels, and are hence unspecified for speech level features. Since the neutral and deferential classes are not involved in the determination of speech level, they will be ignored in the remainder of this paper. One important fact, however, is that according to Clynes (1989, p.17), "if a meaning has a neutral form, it will not have ngoko, madya or krama alternants, and vice versa." This is an important factor driving the paradigmatic competition discussed in the next section.

I summarize the discussion thus far by reviewing the core terminology introduced above:

- Alternant set: A set of synonymous lexical items differing in their speech level features.
- Speech level features: $\pm N, \pm K$. A given lexical item can be specified for either value of either or both features, or can be unspecified for either or both.
- Lexical class: Lexical items sharing the same speech level features.
- Speech level: A characteristic of entire sentences, determined formally by the speech level features of the lexical items in the sentence. Three categorically distinct speech levels exist, marked by the speech level features of the utterance: $+N, -K$ (Ngoko level), $+N, +K$ (Madya level), and $-N, +K$ (Krama level).

3. Continuous properties

The above section summarized the categorical properties of the Javanese speech level system, and provided a formal account of the system. This account generates a categorical distinction between the three traditionally recognized speech levels, and syntagmatic constraints that block the mixing of lexical items with incompatible feature values (i.e. mixing $+K$ and $-K$ or $+N$ and $-N$ lexical items). There is also, however, a widespread intuition that there is a continuous gradation in the Madya level, with a greater proximity to either Krama or Ngoko levels indicated by the proportion of krama (our [+K]) and ngoko (our [+N]) lexical items used:[12]

> [O]ne has to recognize the existence not of *one* particular MADYA-speech style,

[11] Clynes does not provide the Madya example; its acceptability was confirmed with two speakers of Javanese.

[12] As noted by Uhlenbeck (1970, p.453), the degree of respect expressed to the addressee can also be modified within a given speech level by the use or non-use of *krama inggil* deferentials referring to the addressee, although the use or non-use of these items does not affect the speech level as such.

> but a great variety of speech styles which have as common characteristic the use of the MADYA-forms, but which show *different proportions as to the occurrence of* KRAMA-forms. (Uhlenbeck 1970, p.453; emphasis added)

> Madyô is not a set of fixed forms, but is rather a cline rising from a level very close to Ngoko up to a level very close to Krômô. The height of the M level depends on the mixture of Ngoko and Krômô... *The greater the percentage of N forms ... the lower the M level.* (Wolff and Poedjosoedarmo 1982, p.17; emphasis added)

> In the intermediate madya style, the degree of relative "formality" or "distance" is directly dependent on *the relative proportions of ngoko and krama (general lexis) items* used. (Clynes 1989, p.45; emphasis added)

That is, *within* the categorically determined three-level system, there is a gradient between "more Krama-like" or "high" Madya, and "more Ngoko-like" or "low" Madya. The question I address in this section is how exactly this gradient status is to be calculated.

A simple answer, which implements the ideas quoted above, is this: Assume that speech levels are continuously valued on the interval [0,1], with Ngoko level valued at 0, Krama valued at 1, and Madya ranging over all the values between. Lexical items, in turn, would be numerically valued so that $[+N]$ items are valued 0 and $[+K]$ items are valued 1.[13] We could then calculate the numerical value of the speech level associated with a particular sentence by averaging the values of the lexical alternants used, which would in turn mean that Madya level sentences would have different intermediate values depending on the proportion of krama to ngoko items that they employ (c.f. McCready 2019 for a similar approach to calculating continuous values of speech register signaled by different combinations of honorifics in Thai and Japanese).

One problem with this solution is that, according to Wolff and Poedjosoedarmo (1982, p.36), different krama items do not "raise" the level of a given Madya sentence to the same degree: "These forms are not all equal, and the choice of some of them implies a much higher type of Madyo than the choice of others". In order to account for this variation in degree, we would need to assign different numerical values to different $[+K]$ lexical items. The same would hold true for $[+N]$ items, given that they "lower" the level of a Madya sentence by different degrees. This would leave us with no explanation for *why* certain $[+K]$ items raise (and $[+N]$ items lower) the level of a Madya utterance to a greater degree than others. Since there are over 1000 $[+K]$ and $[+N]$ items in the language, this would constitute a high amount of arbitrary continuous variation in the lexicon. We would also lose the numerical mapping suggested above, where 1 signals Krama and 0 signals Ngoko.

I will argue instead that this variation derives at least in part from differences in the particular paradigms that individual items are part of. Moreover, I will show that ngoko or krama items that are part of what I call *non-competitive* paradigms do not contribute to the calculation of Madya sublevels at all. The lexical alternants in example (8) exemplify three such paradigm types, which I label paradigm types 1–3 (P1, P2, P3). P1 is exemplified by the alternants for 'eat', P2 by the alternants for the relativizer REL, and P3 by the alternants for 'already' and for the demonstrative 'that'. P1 and P2 are two-member alternant sets, distinguished by which of the two alternants is used in the Madya level; in P1, only the krama alternant is compatible with the Madya level, while in P2, only the ngoko alternant is. P3 characterizes three-member alternant sets each of whose alternants is compatible with only one level.

[13] [+K,+N] items (the *madya* style markers) might be given the intermediate value 0.5, or might simply be ignored.

Each of these three paradigm types partitions the three speech levels among the alternants, so that there is only one alternant compatible with a given level. This means that, in the Madya level in particular, there is no competition between alternants. If the sub-level of Madya were determined simply by the proportion of $[+K]$ and $[+N]$ alternants, as suggested above, we would expect any Madya-level sentence containing 'eat' to bump the level up (since only the $[+K]$ *nedha* form is compatible with Madya), while an occurrence of the relativizer would necessarily bump the level down (since only the $[+N]$ form *sing* is compatible with Madya). This would be an odd state of affairs, and one that I show in this section does not seem to be empirically verified.

The data in the rest of this section comes from Wolff and Poedjosoedarmo 1982 (henceforth W&P), who had native speakers of Javanese assign utterances from a large corpus of Javanese dialogs into one of the three speech levels discussed above: Ngoko (N), Krama (K), and Madya. Madya utterances were further subdivided into one of three sub-levels: Low Madya (LM), Mid Madya (MM), or High Madya (HM). W&P note that inter-rater agreement for the categorically distinct levels (Ngoko, Madya, and Krama) was nearly 100%, while for the sub-levels of Madya there was significant variation. This is not surprising, given that the sub-levels of Madya are not categorically encoded in the grammar, and represent arbitrary and vague divisions of a continuum. The occurrence and non-occurrence of lexical items from a variety of alternant sets across these levels in their corpus is presented in a series of lists (W&P, pp.30–35) which form the empirical basis for the paradigms I describe below.

On the basis of the data in W&P, I argue that paradigms of lexical alternants can be sorted into ten types on the basis of what assortment of feature values the competing items in the paradigm have. I determined the feature values for individual lexical items on the basis of their occurrence or non-occurrence in the three categorical levels (Ngoko, Madya, and Krama) in the data in W&P, with non-occurrence taken to reflect a categorical restriction against use in that level.[14] Non-occurrence in the Ngoko level is taken as evidence for a $-N$ feature, while non-occurrence in the Krama level is taken as evidence for a $-K$ feature. Non-occurrence in the Madya level is taken as evidence for a $-K$ feature for items that occur in the Ngoko level, and as evidence for a $-N$ feature for items that occur in the Krama level, while *exclusive* occurrence in the Madya level is taken as evidence for a $[+N, +K]$ feature specification.

Table 2 illustrates the occurrence pattern of lexical alternants belonging to the three paradigm types exemplified in (8). The examples here and in the rest of this section follow the orthographic conventions used by W&P, which differ from the standard orthography used in the previous section (following that used by Clynes). The levels are labeled as described above: N for Ngoko, K for Krama, and LM, MM, and HM for the three sublevels of Madya (Low, Mid, and High). Each paradigm type is illustrated by a particular alternant set, along with the total number of sets fitting that paradigm. A checkmark ✓ on white background indicates occurrence of an item in a particular level, while an × on dark gray background indicates non-occurrence. There is no evidence from the distributions summarized in Table 2 that P1 $[+K]$ items are associated with higher levels of Madya (which would be evidenced by non-occurrence in the LM column), or that P2 $[+N]$ items are associated with lower levels of Madya (which would be evidenced by non-occurrence in the HM column). There is also no difference detectable in this data between the distribution of the Madya-compatible ngoko and krama items in P1 and P2 and the madya items in P3, all of which are seen to occur in all three Madya sublevels.

These three paradigm types (P1–P3) contain alternants that partition the speech levels, so

[14] This procedure makes it possible that accidental gaps in the corpus for certain items would lead to an incorrect negative featural specification; given the size of the corpus, however, such errors are likely few.

		Speech Level				
P1 (6 sets)		N	LM	MM	HM	K
'who'						
sinten	$[\quad +K]$	×	✓	✓	✓	✓
sôpô	$[+N, -K]$	✓	×	×	×	×
P2 (3 sets)						
REL						
éngkang	$[-N, +K]$	×	×	×	×	✓
séng	$[+N \quad]$	✓	✓	✓	✓	×
P3 (8 sets)						
'already'						
sampon	$[-N, +K]$	×	×	×	×	✓
ampon	$[+N, +K]$	×	✓	✓	✓	×
wés	$[+N, -K]$	✓	×	×	×	×

Table 2: Paradigm types 1–3, observed occurrence across speech levels

that there is only one item compatible with a given level. They are thus *non-competitive*, in that there is only one alternant possible for any given level. The other paradigm types that can be inferred from the data in W&P are by contrast *competitive*; in particular, these other paradigms provide more than one alternant that is compatible with the Madya level. The vast majority of alternants belong to what I will refer to as paradigm type 0 (P0), in which there are two alternants, one $[+K]$ and one $[+N]$. These sets exhibit one of three patterns, which are summarized in Table 3. In this table, an × on light gray background indicates an item's non-occurrence at a Madya sublevel despite being compatible with the Madya level as a whole, as indicated by that same item's occurrence in other Madya sublevels.

Pattern 1a accounts for the majority of the alternant sets reported in W&P. 47 such alternant sets are given in their Figure Four A, and it is noted that this is a partial list; as this is the only list that is said to be partial, we can infer that all the other sets occurring in their corpus were of this type. I thus take this to be the default pattern for this paradigm, and in turn the default paradigm type for the language as a whole. In this pattern, the Madya sublevels are split between the $[+K]$ and $[+N]$ forms, with the $[+K]$ form being used in both High and Mid Madya, and the $[+N]$ form in Low Madya (again, recall that these sublevels have no categorical status, and are simply approximations for regions along a continuum). Pattern 1b is a variation on this same pattern, with the % signaling that the $[+K]$ form is used "optionally" (W&P, p.34) in Low Madya, although it is unclear how this "optionality" was determined empirically. Pattern 2 illustrates the same trend ($[+K]$ for higher sublevels of Madya, $[+N]$ for lower sublevels), the difference being that now the $[+N]$ form is used in the Mid Madya level instead of the $[+K]$ form.

Given that (i) the split of Madya into three discrete sublevels is arbitrary and without theoretical significance, and (ii) that there was apparently significant inter-rater disagreement in the assignment of particular Madya utterances to these sublevels, it is unclear how much significance should be accorded to the differences between these patterns. In all three patterns, $[+K]$ items tend not to occur in Low Madya, while $[+N]$ items tend not to occur in High Madya. This distribution

		Speech Level				
P0, pattern 1a (47+ sets)		N	LM	MM	HM	K
'house'						
griyô	$[\quad +K]$	×	×	✓	✓	✓
omah	$[+N \quad]$	✓	✓	×	×	×
P0, pattern 1b (12 sets)						
'child'						
laré	$[\quad +K]$	×	%	✓	✓	✓
bocah	$[+N \quad]$	✓	✓	×	×	×
P0, pattern 2 (11 sets)						
'as'						
kadôs	$[\quad +K]$	×	×	×	✓	✓
kôyô	$[+N \quad]$	✓	✓	✓	×	×

Table 3: Paradigm type 0, observed occurrence across speech levels

contrasts with that of the $[+K]$ and $[+N]$ items in P1 and P2 in Table 2, where no such gaps are observed. The reason, I argue, is the lack of any competing forms in P1 and P2; in P1, the $[+K]$ form is the only alternant compatible with the Madya level, and is thus used across all its sublevels. Similarly, P2 forces the use of the $[+N]$ alternant in Madya level, and it is thus used across all sublevels thereof. P0, by contrast, has two distinct competitors that are compatible with the Madya level. This competition allows for the marking of sublevels of Madya, with higher levels signaled by the $[+K]$ alternant and lower levels signaled by the $[+N]$ alternant.

Table 4 summarizes four additional competitive paradigm types inferred from the data in W&P, which I label paradigm types 4–7 (P4, P5, P6, and P7). These are all three-member paradigm types which with the exception of P7 contain a dedicated madya (i.e. $[+N, +K]$) alternant along with a ngoko and krama alternant. In P4, the $[+N, +K]$ madya form competes with a $[+K]$ krama form, in P5 it competes with a $[+N]$ ngoko form, and in P6 there is a three-way competition.

P4 exhibits two distributional patterns. In both patterns, the dedicated madya form does not occur in High Madya level. This contrasts with the distribution of the dedicated madya form in P3, which occurs in all Madya sublevels. The reason for this divergence is straightforward: In P3, the madya form has no competitors, and is the only form compatible with the Madya level. In P4, the madya form competes with a krama form. This competition results in the choice of the krama form for "higher" levels of Madya, and thus the non-occurrence of the madya form at these levels. To put it another way: Given that the speaker is using the Madya speech level, and given that they are choosing a lexical item belonging to a P4 paradigm, the choice of the $[+N, +K]$ alternant will tend to signal a "less Krama-like" (hence "lower") level of Madya. Conversely, the choice of the $[+K]$ alternant, which can also be used in the Krama level, will tend to signal a "more Krama-like" (hence "higher") level of Madya; this tendency is seen in the first pattern for P4.[15] This tendency is missing in the featurally identical $[+K]$ in P1. Once again, given the lack

[15] Its absence in the second pattern (that is, the occurrence there of krama items in lower sublevels of Madya) may reflect a to-be-explained difference in two subtypes of P4, or may instead reflect the inherent noisiness and variability of the data, as noted above.

		Speech Level				
P4, pattern 1 (1 set)		N	LM	MM	HM	K
'how'						
kadôs pundi	$[\quad +K]$	×	×	×	✓	✓
(ke)pripon	$[+N, +K]$	×	✓	✓	×	×
(ke)priyé	$[+N, -K]$	✓	×	×	×	×
P4, pattern 2 (3 sets)						
'from'						
sakéng	$[\quad +K]$	×	✓	✓	✓	✓
(se)kéng	$[+N, +K]$	×	✓	✓	×	×
sekô / sôkô	$[+N, -K]$	✓	×	×	×	×
P5, pattern 1 (3 sets)						
'don't'						
sampon	$[-N, +K]$	×	×	×	×	✓
ampon	$[+N, +K]$	×	×	??	✓	×
ôjô	$[+N \quad]$	✓	✓	??	×	×
P5, pattern 2 (1 set)						
3SG pronoun						
priyambaqipon	$[-N, +K]$	×	×	×	×	✓
piyambaqé / kiyambaqé	$[+N, +K]$	×	×	×	✓	×
dèwèqé / dhèqé	$[+N \quad]$	✓	✓	✓	×	×
P6, pattern 1 (1 set)						
'place'						
panggènan	$[\quad +K]$	×	×	×	✓	✓
nggèn	$[+N, +K]$	×	✓	✓	×	×
nggôn	$[+N \quad]$	✓	✓	✓	×	×
P6, pattern 2 (2 sets)						
'most', 'alone'						
piyambaq	$[\quad +K]$	×	×	×	✓	✓
kiyambaq	$[+N, +K]$	×	×	×	✓	×
dhéwé	$[+N \quad]$	✓	✓	✓	×	×
P7 (1 set)						
'yes'						
inggéh	$[-N, +K]$	×	×	×	×	✓
nggéh / njéh	$[\quad +K]$	×	✓	✓	✓	✓
dhéwé	$[+N, -K]$	✓	×	×	×	×

Table 4: Paradigm types 4–7, observed occurrence across speech levels

of a competing Madya-compatible form, the use of a P1 $[+K]$ alternant signals nothing about the sublevel of Madya.

P5 is the mirror image of P4, with a dedicated $[+N, +K]$ madya form competing with a $[+N]$ ngoko form. The distributional pattern also mirrors that of P4, with the madya form not occurring in lower levels of Madya (LM), and the ngoko form not occurring in higher levels of Madya (HM).[16] Again, these gaps fall out straightforwardly from a competition-based view of Madya sublevels. The use of a P5 $[+N]$ form will tend to signal a more Ngoko-like / less Krama-like (lower) level of Madya, while the use of the $[+N, +K]$ alternant will tend to signal the opposite. The featurally identical $[+N]$ P2 alternant in Table 2 lacks this tendency, which again follows from the lack in that paradigm type of a competing alternant.

P6 has a three-way competition in the Madya level, with a dedicated $[+N, +K]$ madya form competing with both a $[+K]$ krama and $[+N]$ ngoko form. The observed distributions indicate that the $[+K]$ form shows a tendency to be used in higher sublevels of Madya, while the $[+N]$ form shows a tendency to be used in the lower sublevels. The $[+N, +K]$ form show two different patterns, apparently fluctuating between higher and lower Madya sublevels. While the pattern is not completely clear from this data, the competition-based account predicts that speakers would judge the use of the $[+N, +K]$ alternant to *simultaneously* signal a lower level of Madya than the $[+K]$ alternant and a higher level of Madya than the $[+N]$ alternant.

P7 lacks a dedicated madya alternant, and instead has both a strict and non-strict krama alternant. The Madya level is compatible with only one alternant, and this paradigm type is thus non-competitive in the Madya level. It *is* competitive in the Krama level, but since sublevels of Krama were not assessed, the impact of this competition cannot be assessed. The competition-based account for sublevels of Madya, however, leads us to expect a similar grading of sublevels of Krama to arise from the competition between the two krama alternants in this paradigm type.

Comparing Tables 2, 3 and 4, it is clear that it is only in cases of competition (i.e. when there is more than one alternant from the same paradigm compatible with the Madya speech level) that the use of a $[+N]$ form lowers the Madya level or a $[+K]$ form raises it. Moreover, we can see from Table 4 that the use of $[+N, +K]$ madya forms can either (i) lower the level of Madya when competing with a $[+K]$ form, as seen in P4, or (ii) raise the level of Madya when competing with a $[+N]$ form, as seen in P5.[17] This contrasts with the distribution of the P3 $[+N, +K]$ form, which (since it lacks any competitors) occurs in all sublevels of Madya.

For the sake of completeness, two additional paradigm types inferred from the data in W&P are given in Table 5. These paradigm types, of which there is only one example each, contain four alternants. I leave discussion of these two paradigms aside for reasons of space, but will return to them briefly in section 4.

To conclude, the data in this section show that an alternant's tendency to raise or lower the sublevel of a Madya sentence (as reflected in its occurrence or non-occurrence in the three sublevels distinguished by W&P) depends on the presence in the alternant set of a competing alternant that is also compatible with the Madya level. Given such a competitor, the choice of one or the other will signal a higher or lower level of Madya, as a function of those items' feature values: a $[+K]$ item will tend to bump the level up, and its competitor to bump it down, while a $[+N]$ item will tend to bump the level down, and its competitor to bump it up. As a special case, $[+N, +K]$ items (which are inherently unbiased as to their "Krama-ness" or "Ngoko-ness") will tend to bump the

[16] P5, pattern 1 items come from List One A (p.30), which does not indicate whether these items occurred in Mid-Madya (MM). I have put question marks into the corresponding cells to mark this lack of data.

[17] The situation in P6 is less clear.

		Speech Level				
P8 (1 set)		N	LM	MM	HM	K
'thus'						
mekaten	$[-N, +K]$	×	×	×	×	✓
ngaten	$[\quad +K]$	×	×	×	✓	✓
ngèten / ngoten / ngôten	$[+N, +K]$	×	✓	✓	×	×
ngéné / ngono / ngônô	$[+N, -K]$	✓	×	×	×	×
P9 (1 set)						
'not'						
sanès	$[\quad +K]$	×	✓	✓	✓	✓
senès / dédé	$[+N, +K]$	×	✓	✓	×	×
dudu	$[+N \quad]$	✓	✓	✓	×	×
udu	$[+N, -K]$	✓	×	×	×	×

Table 5: Paradigm types 8–9, observed occurrence across speech levels

level down when in competition with a $[+K]$ item, and tend to bump it up when in competition with a $[+N]$ item.

4. Diachronic speculations

In this section, I make suggestions regarding the evolution of the speech level system in Javanese. The starting point for these speculations is the fact that P0 sets (i.e. two-member sets whose members are marked as $[+K]$ and $[+N]$) account for the vast majority of alternant sets (over 500 sets = 1000 lexical bases, based on the lexical class counts in Clynes). The other paradigm types all involve at least one of what Clynes calls "style markers"; that is, the other paradigms include at least one $[+N, -K]$ (strict ngoko), $[-N, +K]$ (strict krama), or $[+N, +K]$ (madya) alternant. Clynes argues that the contemporary speech style system evolved from one that lacked these style markers.[18] In this section, I propose several pathways whereby competition between forms in the Madya level could have led to the development of the contemporary system from an earlier one containing only $[+N]$ and $[+K]$ items, in line with the reconstruction presented by Clynes.

The discussion in this section is premised on the assumption that Paradigm type 0 is both synchronically and diachronically the core of the speech level system. The original speech level system would have had only P0 alternant sets (two-way contrast of $[+K]$ and $[+N]$ alternants). This system would not have had any categorical restrictions on syntagmatic combinations of forms, given the lack of any negative speech level feature values in the lexicon. There would also be no corresponding categorical distinctions among speech levels; the speech level system would be a single continuum with no categorical boundaries. At extreme ends of this continuum, Krama and Ngoko speech levels would have been signaled by sentences that exclusively used either $[+K]$ or $[+N]$ forms, while the Madya level would have been signaled by sentences that mixed $[+K]$ and $[+N]$ forms. From this original system, the other paradigm types could then have developed diachronically from P0 by the following pathways:

[18] He also argues that contemporary Balinese preserves this earlier type of speech level system.

1. P0 ⇒ P1, P2
   Competition between $[+N]$ and $[+K]$ in the Madya level led, in some cases, to the creation of strict ngoko and strict krama forms (i.e. ngoko and krama style markers).
2. P0 ⇒ P6
   Alternative krama forms with differing levels of formality led to the creation of dedicated madya forms.
3. P6 ⇒ {P5, P4} ⇒ P3
   The new madya forms led to the development of additional paradigms, due to competition between the new madya forms with the older ngoko and krama forms.

The first pathway takes us from a P0 alternant set to either a P1 or P2 alternant set. To begin with, consider pattern 1 in Table 3, in which there is an apparent bias for the $[+K]$ alternant being used in Mid Madya sublevel. Imagine now that there is a drift in usage, such that the $[+K]$ form gets used in successively lower sublevels of Madya, until the point at which it is used at *all* such sublevels, with the $[+N]$ form being used only in the Ngoko level. This usage pattern could then seed a reanalysis of these alternants, such that the $[+N]$ form is taken to be incompatible with the Madya level; formally, a reanalysis of $[+N]$ to $[+N, -K]$. Note that there is no featural reanalysis of the $[+K]$ form that has "taken over" the Madya level. This process is illustrated in Table 6.

		Speech Level				
		N	LM	MM	HM	K
P0, pattern 1						
	$[+K]$	×	×	✓	✓	✓
	$[+N]$	✓	✓	×	×	×
⇐ Drift						
	$[+K]$	×	✓	✓	✓	✓
	$[+N]$	✓	×	×	×	×
	$[+K]$	×	✓	✓	✓	✓
Reanalysis	$[+N, -K]$	✓	×	×	×	×

Table 6: P0 seeding a shift to P1

The evolution of P2 from P0 is the mirror image of this process. I take the starting point to be a distribution like that exemplified by pattern 2 in Table 3, where there is an apparent bias toward using the $[+N]$ alternant in Mid Madya, with the $[+K]$ alternant restricted to higher sublevels of Madya. A drift in this usage pattern, with the $[+K]$ alternant being restricted to successively higher sublevels of Madya, could lead to the eventual takeover of all Madya sublevels by the $[+N]$ alternant, which in turn could seed a reanalysis of the paradigm, with the $[+K]$ alternant now only being compatible with the Krama level; formally, a reanalysis of $[+K]$ to $[-N, +K]$. Note that no featural reanalysis is necessary for the $[+N]$ form that has "taken over" the Madya level. This process is illustrated in Table 7.

The reanalysis triggered by this shift in usage represents a fundamental change in the nature of the speech level system, from a system with no syntagmatic constraints (because of the lack of negative feature values) and no categorical divisions between speech levels, to one with

syntagmatic constraints governing the co-occurrence of lexical items, and in which speech levels are broken into three distinct levels.

		Speech Level				
		N	LM	MM	HM	K
P0, pattern 2						
	$[\quad +K]$	×	×	×	✓	✓
	$[+N \quad]$	✓	✓	✓	×	×
⇒ Drift						
	$[\quad +K]$	×	×	×	×	✓
	$[+N \quad]$	✓	✓	✓	✓	×
Reanalysis						
	$[-N, +K]$	×	×	×	×	✓
	$[+N \quad]$	✓	✓	✓	✓	×

Table 7: P0 seeding a shift to P2

This first pathway gives us a system with strict ngoko and krama items, and the introduction of negative feature values, but the result is still limited to two-member alternant sets with no dedicated madya forms. As noted by Clynes (1989, p.25), most madya items "are clearly diachronically derived from their krama alternants by processes of elision and lenition". This generalization holds for all of the paradigm types in which madya forms occur, as seen in the examples in Tables 2 and 4: P3 krama *sampon* → madya *ampon* 'already'; P4 krama *sakéng* → madya *(se)kéng* 'from'; P5 krama *sampon* → madya *ampon* 'don't'; P6 krama *panggènan* → madya *nggèn* 'place'. Clynes (1989, p.25), citing Errington 1985, pp.140–145, also adduces evidence that these madya forms are synchronically distinct lexical items from their corresponding krama forms, with the elision/lenition processes reflecting diachronic rather synchronic processes.

Following suggestions made by Clynes (1989, p.207), the development of madya forms could have proceeded as follows. First, there would be a tendency to use elided P0 $[+K]$ forms to signal relative lack of distance, and thereby relative height along the Ngoko to Krama continuum. The elided form would thus tend to be used in the Madya level but not in the Krama level, while the full form would tend to be used in the Krama level but not in the Madya level. As suggested by Clynes, this tendency could then develop into an absolute restriction. Formally, such a reanalysis would first involve reanalysis of the elided and full forms as distinct lexical items with distinct feature values. The reanalysis could proceed in two distinct directions. The full form might be reanalyzed as a strict $[-N, +K]$ lexical item, leaving the elided form with its original $[+K]$ feature values. Such a reanalysis would take us from an original P0 alternant set to something like a P7 alternant set. I say "something like", because the observed distribution of P7 in Table 4 has a $[+N, -K]$ strict ngoko form rather than a $[+N]$ form. This divergence would result from additional reanalysis of the ngoko form, paralleling that posited in Table 6 taking us from $[+N]$ to $[+N, -K]$. Alternatively, the elided form might be reanalyzed as a $[+N, +K]$ madya lexical item, with the full form remaining $[+K]$. This alternative reanalysis would lead to the creation of a new lexical class (the $[+N, +K]$ madya class), and take us from an original P0 alternant set to a P6 alternant set.

P6 involves three-way competition in the Madya level between ngoko, madya, and krama alternants. This competition in P6 could seed additional reanalysis of the krama and ngoko forms, as the new madya form begins to take over the Madya level, pushing the krama and ngoko al-

ternants out and causing them to be reanalyzed as strict krama and ngoko forms. This pathway is schematized in Table 8. At the top is an idealized picture of the starting point, with the three

		Speech Level				
		N	LM	MM	HM	K
P6	$[\quad +K]$	×	×	×	✓	✓
	$[+N, +K]$	×	×	✓	×	×
	$[+N\quad]$	✓	✓	×	×	×
⇓						
	$[\quad +K]$	×	×	×	✓	✓
P4 (⇐ drift + reanalysis)	$[+N, +K]$	×	✓	✓	×	×
	$[+N, -K]$	✓	×	×	×	×
or						
	$[-N, +K]$	×	×	×	×	✓
P5 (⇒ drift + reanalysis)	$[+N, +K]$	×	×	✓	✓	×
	$[+N\quad]$	✓	✓	×	×	×
⇓						
	$[-N, +K]$	×	×	×	×	✓
P3 (additional drift + reanalysis)	$[+N, +K]$	×	✓	✓	✓	×
	$[+N, -K]$	✓	×	×	×	×

Table 8: P6 seeding a shift to P5, P4, and P3

competing forms in P6 dividing the three Madya sublevels between them: the ngoko form is used in Low Madya, the krama form in High Madya, and the madya form in Mid Madya. A drift of madya form usage to successively higher levels of Madya, along with a corresponding shift in the krama form, could seed a reanalysis to P5, while a mirror-image shift of the madya and ngoko forms could seed a reanalysis to P4. Shifting and reanalysis in both directions would give rise to P3. The idealized usage patterns for each paradigm in this pathway can be compared to the empirically observed patterns for the corresponding paradigms presented in the previous section; the fit is not perfect, but as already discussed, the assignment of sentences to Madya sublevels is variable and noisy.

A P7 alternant set, which we saw above could be derived from an original P0 set, could evolve directly into a P3 set by a reanalysis of the $[+K]$ krama item to a $[+N, +K]$ madya item. This shift would be seeded by competition between this item and the $[-N, +K]$ strict krama alternant within the Krama level; while this kind of competition has not figured into the discussion thus far, we can imagine that it parallels the competition seen in the Madya level. Alternatively, we can imagine a pathway taking us from P7 to P8. This would involve the reanalysis of elided forms of the $[+K]$ alternant (itself derived from an elided form of the $[-N, +K]$ alternant) as distinct lexical items, and the assignment of these new lexical items to the $[+N, +K]$ class. This process parallels directly the pathway taking us from P0 to P6, whereby elided $[+K]$ forms are lexicalized

and assigned to the madya class.

P9 (of which there is only one example) might result from competition within the Ngoko level. Assume an original P5 alternant set in which elision or non-elision of the $[+N]$ form can signal relative degree of formality. The elided form would tend to be used in lower levels and the non-elided form in higher levels, paralleling the suggestions above for krama forms. Such a usage pattern could seed reanalysis of the elided form as a distinct lexical item compatible only with the Ngoko level, and thus with $[+N, -K]$ feature values. The single example of this paradigm type is, however, complicated by the fact that the apparently strict ngoko form *udu* is listed in Robson and Wibisono 2002 as "dialectal"; this raises the possibility that the one purported P9 paradigm might be collapsing items from distinct dialects of the language.

5. Conclusions and further issues

In the preceding sections, I have shown that the Javanese speech level system relies on both syntagmatic and paradigmatic relationships. Section 2 showed that the three core speech levels are signaled syntagmatically. The syntagmatic system is categorical in two senses: it generates a categorical distinction between three distinct speech levels, and it generates categorical judgments of acceptability or non-acceptability regarding the combination of particular forms within a sentence. Section 3 showed how continuous variation within speech levels derives at least in part from paradigmatic competition between alternant forms. Section 4 linked these two systems diachronically, showing how competition within paradigms could have seeded reanalysis of paradigmatic alternants. This reanalysis takes us from a fundamentally continuous system with only one paradigm type and two lexical classes signaling speech level to the contemporary system of five lexical classes, multiple paradigm types, and categorical restrictions on the combination of forms.

There are two methodological shortcomings in this paper that should be noted. First, the analysis in section 2 relied primarily on the data and analysis of Clynes 1989, which relies on native speaker judgments of aristocratic (*priyayi*) speakers from the court centers of central Java. Section 3, by contrast, is based on corpus data that, while taken from the same region, represents (by design) a wider variety of speakers. This means that the alternants treated here as members of a single paradigm may in some cases represent mixing of different dialects, as was discussed in relation to P9 and the "dialectal" form *udu*. I also noticed one instance where the usage facts reported by W&P contradicted the judgments reported by Clynes, namely the alternant set for 'eat'. According to the data of Clynes (as reflected in (8)), this should be classified as a P1 alternant set, while according to the data of W&P it should be classified as a P0 alternant set. Judgments from Javanese consultants confirmed those of Clynes. The divergence with W&P might reflect dialectal variability in the corpus, rater error, or even a performance error on the part of some speaker in the corpus.

The second problem concerns the inference of speech level features from corpus data. The procedure I used to infer feature values for lexical items treated non-occurrence in one of the three categorical speech levels as resulting from a prohibition of use in that level. In other words, gaps in the corpus were used as negative evidence. This procedure is problematic in that accidental gaps in the corpus could lead to misanalysis. The solution to these problems would be to supplement the corpus data with native-speaker judgments of sentences that are predicted to be unacceptable, and to determine for individual paradigms which forms are used in which dialects (or by which speakers). This is made difficult by the widespread dialectal and social class variation in the speech level system, as well as the fact that the Krama speech level has been reported to be at risk of

endangerment (Vander Klok 2019).

The analysis presented in this paper, with lexical classes defined in terms of features, speech levels defined in terms of compatible feature values, and gradations within levels driven by paradigmatic competition between alternants, may provide a framework for understanding the similarities and differences between Javanese and other Western Indonesian languages reported to have speech levels (i.e. Balinese, Madurese, Sasak, and Sundanese). More tentatively, this framework might be extended to the honorific systems of typologically unrelated languages. Future work is needed in characterizing the deferential lexical class and its interaction with the speech level system. While I noted above that deferential items do not themselves participate in the determination of speech levels, they can modulate the degree of respect or formality shown to the addressee, when the item in question *refers to* the addressee or things immediately related to the addressee (c.f. Uhlenbeck 1970). Since deferential items are often in paradigmatic alternation with items that mark speech level (e.g. the alternants for 'eat' discussed in this paper), there is prima facie competition between these forms. Intuitively, it seems that the selection of the deferential in such cases will, when used to refer to the addressee, raise the sublevel of whatever speech level is otherwise indicated by the sentence or utterance as a whole. The details of this interaction, and the details of the deferential items themselves, I leave to future research.

References

Clynes, Adrian. 1989. Speech styles in Javanese and Balinese : A comparative study. Masters, Australia National University.

Comrie, Bernard. 1976. Linguistic politeness axes: Speaker-addressee, speaker-referent, speaker-bystander. *Pragmatics Microfiche* 1:1–12.

Errington, J. Joseph. 1985. *Language and social change in Java: Linguistic reflexes of modernization in a traditional royal polity*. Number 65 in Monographs in International Studies Southeast Asia Series. Athens, Ohio: Ohio University, Center for International Studies.

Harada, S. I. 1976. Honorifics. In *Japanese Generative Grammar*, ed. Masayoshi Shibatani, volume 5 of *Syntax and Semantics*, chapter Japanese Generative Grammar, 499–561. Academic Press.

Horne, Elinor Clark. 1974. *Javanese-English dictionary*. New Haven: Yale University Press.

McCready, Elin. 2019. *The semantics and pragmatics of honorification: Register and social meaning*. Oxford University Press, first edition.

Poedjosoedarmo, Soepomo. 1969. Wordlist of Javanese non-ngoko vocabularies. *Indonesia* 7:165–190.

Portner, Paul, Miok Pak, and Raffaella Zanuttini. 2019. The speaker-addressee relation at the syntax-semantics interface. *Language* 95:1–36.

Robson, Stuart, and Singgih Wibisono. 2002. *Javanese English dictionary*. Hong Kong: Periplus Editions.

Uhlenbeck, E.M. 1970. The use of respect forms in Javanese. In *Pacific linguistic studies in honour of Arthur Capell*, ed. S.A. Wurm and D.C. Laycock, number 13 in Pacific Linguistics Series C, 441–466. Canberra: Australian National University.

Vander Klok, Jozina. 2019. The Javanese language at risk? Perspectives from an East Java village. *Language Documentation* 13:300–345.

Wolff, John U., and Soepomo Poedjosoedarmo. 1982. *Communicative codes in central Java*. Ithaca, New York: Southeast Asia Program Publications.