WHAT'S HAPPENED TO US?
SOME DEVELOPMENTS
IN THE MALAY
PRONOUN SYSTEM

MARK DONOHUE
UNIVERSITY OF MANCHESTER

JOHN CHARLES SMITH
UNIVERSITY OF OXFORD

We examine the development of the Austronesian first-person pronouns in nonstandard varieties of Malay/Indonesian. The fact that of the inclusive and exclusive pronouns only the form of the inclusive, *kita?, is ever retained in the nonsingular forms when the inclusive-exclusive distinction is lost, or extended to a singular meaning, is taken as a sign of an oppositional relationship being retained even after the functional basis for that opposition has ceased to exist. Evidence for the markedness relationship that exists between the two pronominal forms is presented on the basis of qualitative and quantitative data, with both Classical Malay and Modern (written) Indonesian being examined. A discussion of the shifts in meaning that have been observed to occur, and the shift in the pronominal paradigm in the light of Blust's treatment of the politeness shifts in earlier varieties of Austronesian, offers a functional explanation for the kinds of changes observed. Additional data from local languages of Southeast Sulawesi, in which both forms are preserved, even though the function encoded is now one of plural/paucal, strengthen this argument.

1. SKEUOMORPHY. Work in diachronic morphology has tended to concentrate on the fate of forms rather than systems or subsystems. In addition, the study of functional or semantic changes and the study of formal changes have often proceeded quite separately from one another. Lass (1990) linked the two types of change, but inferred an essentially random relationship between form and function. Smith (1995, 1997), on the other hand, argued that the refunctionalization of a morphological opposition involves a degree of motivation. In this paper, we discuss the development of the Austronesian first-person nonsingular pronoun system in the light of these claims.

Oceanic Linguistics, Volume 37, no. 1 (June 1998)
© by University of Hawai'i Press. All rights reserved
Discussing the development of morphological oppositions that have ceased to encode functional oppositions, Lass (1990) suggested that “[t]his morphology is now, functionally speaking, junk.” He identified “three things that can in principle be done with it: [. . . ] (i) it can be dumped entirely; (ii) it can be kept as marginal garbage or nonfunctional/nonexpressive residue (suppletion, ‘irregularity’); (iii) it can be kept, but instead of being relegated as in (ii), it can be used for something else, perhaps just as systematic.” Borrowing a term from evolutionary biology, Lass referred to option (iii) as “exaptation.”

Following Lass’s work, Smith (1995, 1997) examined the evolution of the accusative and dative forms of the first- and second-person singular pronouns from Latin to Romance. The original functional opposition between the two cases may survive (as in Rumanian and Old Sardinian), or it may be lost. However, the data led Smith to question the idea that “junk” is a suitable notion to describe the situation that results from the disappearance of this distinction. Whenever the morphological contrast has been “dumped entirely” (as in standard French and most Italian dialects), it is always the dative form that has been discarded. When exaptation has taken place, the original contrast of case has been replaced by a distinction between a conjunctive pronoun (characteristically the complement of a verb) and a disjunctive pronoun (characteristically the complement of a preposition). But this process is likewise far from random, for it is always the original accusative case-form that has been exapted into the function of conjunctive pronoun, and always the original dative case-form that has taken on the role of disjunctive pronoun (as in Spanish, Portuguese, and some northern French dialects).

From these data, Smith concluded that the original opposition has been evacuated of all or almost all its concrete functional content (i.e., its exponence), but that a residual, more abstract, dichotomy between a “core” term and a “noncore” term remains—that is, an identity that, however diminished, is not yet junk. To describe this state, he borrowed the term “skeuomorphy,” used by art historians to refer to the stage at which a feature that was once functional has, as a result of progress in technology or design, become merely decorative (see Humphrey 1992:185–186). In the case of the Romance pronominal forms, the “core” item in the residual dichotomy can be objectively defined in terms of frequency and qualitative unmarkedness—the accusative is more frequent than the dative in Latin, it is more versatile (in terms of the grammatical contexts that require it), and it functions as a default case. Smith suggested that it is possible for a morphological opposition to lose its exponence, while retaining an abstract content that will influence any subsequent refunctionalization, and that the more frequent or unmarked term of the original opposition will assume the more frequent or unmarked function in the new one (in the Romance example, for instance, the conjunctive pronouns are more frequent than their disjunctive counterparts, and, it can be argued, are also qualitatively less marked, as complements of verbs rather than prepositions).

In this article, we shall use the idea of an opposition existing between morphological forms even when the original functional contrast has ceased to be relevant to examine the development of the Austronesian first-person nonsingular pronouns (originally, and still widely, first-person plural inclusive and first-person plural exclusive), specifically in the development of nonstandard varieties of
SOME DEVELOPMENTS IN THE MALAY PRONOUN SYSTEM

Malay/Indonesian. Further arguments for this position are presented from a consideration of the developments of the nonsingular pronouns in some of the local languages of Southeast Sulawesi.

2. AUSTRONESIAN PRONOUN SYSTEMS. Phenomena in Austronesian languages that are in many ways similar to the Romance data discussed in Smith (1995, 1997) constitute further evidence for the role of skeuomorphy in language change. Proto-Austronesian is assumed to have had a distinction between an inclusive first-person plural pronoun *i-[k]ita and an exclusive first-person plural pronoun *i-[k]ami (Dahl 1976, Blust 1987, Ross 1997), and this distinction is frequently maintained in the daughter languages, often with transparent historical connections. To name just one example, Amis, spoken on the east coast of Taiwan, uses [k]ïta and [k]ami, with inclusive and exclusive meanings, respectively (Chen 1987:135–136), perfectly reflecting both the form and the function of the Proto-Austronesian reconstructions.

In Proto-Malayic (Adelaar 1992)—the parent language of Standard Malay, Minangkabau (central west Sumatera), Banjar Malay (southwest Borneo), Seraway (Bengkulu/Palembang, south Sumatera), and Iban (northwest Borneo)—we find the inclusive/exclusive distinction maintained, and the form of the Proto-Austronesian pronouns similarly preserved; Adelaar (1992:124, 126, 137) reconstructs *kitaʔ and *kami for the inclusive and exclusive, respectively. (Adelaar (1992:3) defines Standard Malay as “the isolect on which Bahasa Indonesia and Bahasa Malaysia are based, and by which is meant ‘...the literary Malay which represents the direct descendant of the language used in the court of the Malacca sultanate... and which continued to be used in the court of the Sultans of Riau and Johore’ (Prentice 1987:23). Standard Malay is itself based on Classical Malay, which is the Malay of literary works from the sixteenth till the nineteenth century.”)

3. A SOCIOLINGUISTIC AND HISTORICAL SKETCH OF MALAY. The history of the Malay/Indonesian language is a long and interesting one. Written records go back to the sixteenth century, and its use in various communities at great distance from the Malay peninsula, extending back several centuries as a result of widespread trade, has led to multiple alternative speech traditions. The separate development of these varieties, which are most divergent from standard forms of the language in the eastern regions of Indonesia, has led to what may be regarded as nearly separate languages, as any speaker of standard Malay or Indonesian will testify on first visiting Ambon. The last century has, however, seen an opposite trend in the standardization of (at least) the written Malaysian form, and, since independence in 1945, the development of a standard form of the language in Indonesia, which has been introduced to most parts of the country through government schooling (see Heryanto 1995 for details of this process and an informed discussion of the political consequences of language development in Indonesia). Nevertheless, the local, nonstandard, varieties continue to be used in less formal circumstances.
It should not be thought that these nonstandard varieties are a recent development. Heryanto (1995:47), for instance, notes, referring to the "High" or Dutch-approved varieties, and the "Low" or locally used forms of the language, that, "from the latter part of the seventeenth century, there were already serious controversies concerning 'High' and 'Low' Malay, as well as the standardisation of lingua franca Malay, as the preferred language of trade, religious conversion and translated scripts, diplomacy, and, later, governmental administration." Similarly, it should not be thought that Malay/Indonesian varieties are uniform in their sociolinguistic status. Cumming (1991:10) writes that "The range of sociolinguistic statuses occupied by Malay varieties is probably as great as the geographic range of communities of speakers. In some communities it is the only language of monolingual speakers, while in others it is a marginal contact language used only by a small segment of the population. It is a national language in four countries—Indonesia, Malaysia, Brunei, and Singapore—but it evokes different attitudes, and fills different sociolinguistic niches, in each of them. In areas where Malay dialect is an old trade language, that variety of Malay may compete for social functions with both the standard national variety and a completely different local language." Any attempt to account for the different varieties of Malay in terms of one parameter only (standard vs. nonstandard, conservative vs. innovative, "pure" vs. mixed or creolized, etc.) is doomed to failure. In our use of the term "nonstandard," we are assuming any variety of the language that is not recognized by a national government.

The differences between these forms of Malay/Indonesian and the standard are manifest largely in terms of phonology and morphosyntax, with lexical differences mainly limited to certain high-frequency functors, which are usually quite different from the standard (for instance, the negative, tidak in standard Indonesian, is ndak in Sulawesi, nggak in Jakarta, seng in Ambon, sonde in Kupang, and tara in North Maluku and North Irian). Mechanisms for indicating aspect and passives are also quite different in the local varieties of Malay, as is the pronominal system. All varieties of Malay/Indonesian possess an abundance of first- and second-person pronouns, varying in degrees of familiarity or politeness conveyed, but in the nonstandard varieties we find reduction or alteration of the system of oppositions used. This is discussed in more detail in section 5.

Further details of the sociolinguistic and historical background to the Malay/Indonesian language lie beyond the scope of our present work, but much discussion can be found in the references cited at the end of this article.

4. FORMAL VARIETIES OF MALAY/INDONESIAN. The inclusive/exclusive opposition that existed in Proto-Austronesian survives intact in many, if not most, modern Austronesian languages, including modern standard Malay/Indonesian. The form preserved intact with the pronouns kita 'we (inclusive)' and kami 'we (exclusive)' relates transparently to the Proto-Austronesian forms presented in section 2. The distinction between these two pronominal forms is described as preserving the inclusive/exclusive contrast that was present in Proto-Austronesian. For example, discussing standard Indonesian, Macdonald and Soenjono (1967:120) state that the difference between the kami and kita forms of the pro-
nouns is that "[t]he pronoun *kita* is inclusive in that the speaker includes the hearer when he chooses this pronoun," whereas "[t]he pronoun *kami* is exclusive, on the other hand, in that the speaker excludes the hearer when he uses it."

More recently, Sneddon (1996:1), in describing "standard formal Indonesian," which "can loosely be identified as the language of government, administration and the mass media in the Republic of Indonesia," notes that "the inclusive first plural *kita* means 'we' where the person being spoken to is included, that is, 'I and you', while the exclusive first plural *kami* means 'we' where the person being spoken to is excluded, that is, 'I and others but not you'" (Sneddon 1996:160).

This is a succinct definition of the inclusive/exclusive distinction. Similar descriptions are found in pedagogical works on Malay or Indonesian, including those that focus on a more colloquial variety (e.g., Othman and Atmosumarto 1995:13). Other writers of Indonesian grammars are also in agreement about the difference between the *kami* and *kita* forms, and their view is borne out by an examination of the use of these forms in Sneddon's "standard formal Indonesian," represented by written documents.

In order to determine which of these forms may be considered the "core" term of the opposition, we examined both qualitative and quantitative data. Qualitatively, 'exclusive' is marked with respect to 'inclusive'. Specifically, an exclusive first-person plural may be defined as 'first person + third person', while an inclusive first-person plural may be defined as 'first person + second person (+ third person)'. In terms of the person hierarchy 'first person > second person > third person' (for motivation and discussion of this hierarchy, see Silverstein 1976), the inclusive combines the two least marked persons, while the exclusive combines the least marked with the most marked, excluding the middle term. In terms of combinations of person, the exclusive form will therefore be more marked. In discourse terms, too, a form that includes both (or all) discourse participants may be regarded as less marked than one that excludes one (or some) of them.

For quantitative data, we examined the relative frequency with which these pronominal forms occur in (written) discourse. We conducted frequency counts on material obtained from the Indonesia daily news on-line service for the months of March and April 1997 (available via the Internet at http://www.uni-stuttgart.de/indonesia/news/), representing a variety of discourse types, and found the distribution across an approximately 250,000-word corpus to be that shown in table 1.

![Table 1. Frequency of First-Person Plural Forms in Written Indonesian](image)

<table>
<thead>
<tr>
<th>TOTAL WORDS:</th>
<th>244,500</th>
<th>% OF TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL OCCURRENCES OF <em>kita</em></td>
<td>1071</td>
<td>0.44%</td>
</tr>
<tr>
<td>TOTAL OCCURRENCES OF <em>kami</em></td>
<td>487</td>
<td>0.20%</td>
</tr>
</tbody>
</table>

There are a total of 1558 tokens of first-person nonsingular pronominal forms in our corpus, of which 31% are the exclusive *kami*, and 69% the inclusive *kita*, or
a ratio of 2.2:1 in favor of the inclusive. Clearly, the inclusive form is unmarked in modern written discourse with respect to the exclusive form. The same pattern is found in older varieties of Malay. Proudfoot (1990), in his concordance to the Hikayat Inderaputera, a Classical Malay prose romance preserved in a manuscript dating from 1700, finds 170 occurrences of *kita* (1990:425–427) and 54 of *kami* (1990:370), a ratio of 3.1:1 in favor of *kita*. We can clearly infer that the use of *kita* in textual discourse is less marked in both the earlier texts and in modern writing. In the next section, we present evidence that in modern spoken Indonesian the gap in frequency of use of the two forms is wider still.

5. THE SPREAD OF *kita* AND THE LOSS OF *kami*. Despite this neat division of the first-person nonsingular pronominal category in Indonesian into an inclusive *kita* and an exclusive *kami*, we rapidly discover that this is not a complete account of the facts concerning first-person nonsingular pronouns in nonstandard varieties of Malay. In many colloquial varieties of Malay, the form derived from *kit$ has developed singular meanings; this is attested in North Maluku and (historically) North Irian Malay. This is analogous to the use of ‘we’ in formal English, and represents what we shall refer to (following the terminol-

---

1. The fact that the material we examined represents a variety of genres is not trivial; Hyslop (1997 pers. comm.) points out that in a language that maintains the distinction between the inclusive and exclusive forms the relationship between the two is to an extent dependent on the genre of the text. She observes that "one could expect a prevalence of inclusive forms in, for example, ceremonial speeches (we [incl] have come together today to . . .), and a prevalence of exclusive forms in narratives recounting experiences (we [excl] did this, and this . . .)." She surveyed texts from Ambae (Austronesian, Northern Vanuatu) – procedural and ceremonial texts on the one hand and narratives of personal experiences on the other, selected for their contrast—and arrived at the following figures for the ratio of inclusive to exclusive speech genres in Ambae:

<table>
<thead>
<tr>
<th>PROCEDURAL AND CEREMONIAL TEXTS</th>
<th>PERSONAL-EXPERIENCE NARRATIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>INCL 182</td>
<td>EXCL 3</td>
</tr>
<tr>
<td>EXCL 3</td>
<td>RATIO 60:7:1</td>
</tr>
<tr>
<td>PERSONAL-EXPERIENCE NARRATIVE</td>
<td></td>
</tr>
<tr>
<td>INCL 13</td>
<td>RATIO 0:1:1</td>
</tr>
</tbody>
</table>

These figures clearly show that there is a massive skewing in favor of the inclusive forms in the procedural and ceremonial texts, much higher than might be expected from a more mixed set of texts (Harlow 1990, for instance, reports a 1.8:1 ratio of inclusive to exclusive pronouns in traditional prose narratives for Māori, another language with a functional inclusive/exclusive distinction). Significantly, the data also show that in narratives of personal experiences the unmarked member of the opposition is the exclusive one. This illustrates the point that markedness is not simply a universal property of an opposition, but is dependent on genre for its instantiation. (Addressing this point, Foley and Van Valin 1984:2 note that the “standard” text used for analysis is a third person narrative, quite different from those examined by Hyslop, because “tracking of events and participants [. . .] shows up most clearly in this type of discourse.”) Furthermore, it is widely reported that grammaticality judgments monitor the type of speech used in narrative discourse. It is interesting to compare Hyslop’s figures for the frequency of occurrence of exclusive forms in personal-experience narratives in Ambae with Englebretson’s observations (reported in section 5) on the nonuse of exclusive forms in this environment in the colloquial Indonesian texts that he collected; here, even where we could expect to see a strong preference for the exclusive forms, the putatively inclusive form is preferred. It is clear that the inclusive-exclusive opposition is dysfunctional in that variety of Indonesian.
SOME DEVELOPMENTS IN THE MALAY PRONOUN SYSTEM

ogy in Blust 1977) as the FIRST MALAY POLITENESS SHIFT: the shift of the first-person plural inclusive pronoun to assume first-person singular (polite/formal) meaning, and in some cases simply to become the first-person singular pronoun. Importantly, while the inclusive pronoun kita has been attested with this shift, the exclusive kami is not reported in a first-person singular function.²

The basis of this shift can be seen in the reference of the inclusive pronoun: both first person and second person are included, both the speech act participants. When referring to an event, an addressee can trivially distinguish between first-person singular reference, when the speaker is referring to events that occurred without the addressee, and what occurred in the presence of the addressee; note, however, that this functional explanation does not offer any a priori reasons for the inclusive form being used, not the exclusive.

Once this shift occurred, however, most varieties of Malay introduced a means of differentiating the singular and the plural, commonly (though not exclusively) with -ong (< orang 'person')³ attached to the singular pronoun. Thus we find the widespread occurrence of dia 'third person singular', and diaorang, dorang, dong 'third person plural', < dia + o(ra)ng; North Irian Malay kita 'first-person singular' and kitong 'first-person plural'. At this point, we find the language no longer exhibits a reflex of *kami. There are two possible explanations for this. The first might be that, with the use of a general plural marker for all persons, the inclusive/exclusive distinction that was made earlier is no longer expressed, because there is no distinction between inclusive and exclusive in the first-person singular, and thus no means of having two distinct bases for the plural.⁴ Alternatively, the phonetic similarity of kamong (kami '1PL.EXCL' + the now-obligatory plural marker -ong) with the second-person plural (typically *kamu + -ong, yielding kamong)⁵ was too great, and so the first-person plural inclusive extended its range of meaning, and the form derived from *kami was dropped. Of course, in those varieties in which there is no overt plural marker in operation, such as Betawi or Sulawesi Malay, this explanation cannot be valid, and we must simply appeal to some inherent property of the inclusive form that the exclusive form lacks, even when refunctionalized.

Finally, we find that in most cases the FIRST MALAY POLITENESS SHIFT is hidden by further developments in the first-person singular pronoun. In most varieties of North Irian Malay, only older speakers use kita as a singular pronoun, younger speakers having replaced it with saya, borrowed from more standard varieties of Indonesian. In other varieties, in which the kita pronoun did not become the sole exponent of the 1SG pronoun, the original form (aku, beta, gua, etc.) has once again become the norm, and the polite option with kita is not in use.

In yet other cases, we find a second innovative use of the *kita? forms: as second-person singular polite pronouns (as reported in a number of languages of

². Outside its highly stylized use as a "royal" first-person singular pronoun, in certain speech styles; it is not, however, in general use as a first-person singular pronoun in any variety of Malay/Indonesian.

³. Possibly the result of influence from Hokkien (Min) Chinese, which also uses the word for person to indicate plurality in the pronouns: wa '1SG', wa-nay '1PL' ('1SG' + 'person') (Paul Kroeger, pers. comm.).
Indonesia, in Blust [1977:11]). This is found in the Malay of South and Southeast Sulawesi, where the same form *kita serves as both a first-person plural pro-

4. This situation has parallels in the pronominal systems of some languages of the Nimboran family (Voorhoeve 1975) in northern Irian Jaya. There are two main languages in this family, Nimboran and Kemtuik (earlier Kemtuk), each with several thousand speakers, and the Mlap language (referred to as Kwansu in earlier works; this is the name of the largest village) located between them. The Mlap pronominal system shows influence from both Kemtuik and Nimboran, but with changes in the meanings of both the bases and the suffix, resulting in a system that is quite different from either of the sources of its elements.

In Kemtuik we find a simple distinction between minimal and augmented pairs, with the augmented set being derived from the minimal set with the addition of -nag. There are basic pronouns for four categories, the 1, 12, 2 and 3 persons. Without -nag they refer to the minimal number of persons required to fill the persons; that is, one person in the case of gnam, mot, and nmot, and two persons (the speaker and the addressee) in the case of imot. With the addition of -nag the reference of the pronouns is any number greater than this; the reference of gnam-nag is exclusive, necessarily excluding the addressee, since the inclusive use has a more highly specified base in imot. The difference between imot and imot-nag is thus one of dual versus plural (more than two) reference, in contrast to the other pronouns, which are singular versus nonsingular. The Kemtuik pronominal system is set out below:

<table>
<thead>
<tr>
<th>Minimal</th>
<th>Augment</th>
</tr>
</thead>
<tbody>
<tr>
<td>I (EXCL)</td>
<td>gnam</td>
</tr>
<tr>
<td>12 (INCL)</td>
<td>imot</td>
</tr>
<tr>
<td>2</td>
<td>mot</td>
</tr>
<tr>
<td>3</td>
<td>nmot</td>
</tr>
</tbody>
</table>

Nimboran (Anceaux 1965) lies west of Kemtuik, and has a very different pronominal system both in form and function. There is, crucially, no distinction in number, the same forms being used (for instance) for 3SG and 3PL (go), or for 1SG and 1PL.EXCL (go). This system is laid out below:

<table>
<thead>
<tr>
<th>SG/PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (EXCL)</td>
</tr>
<tr>
<td>12 (INCL)</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
</tbody>
</table>

The Mlap language is spoken in three small villages, located between the larger Kemtuik and Nimboran language areas. The base forms of the pronouns show strong similarities with the Nimboran pronouns, yet the suffix -nag, found in Kemtuik, is also in use. The opposition that is marked by -nag in Mlap, however, is not the minimal/augmented one found in Kemtuik, but rather a simple singular/plural one. Since, however, one of the first-person pronominal bases (which show strong formal similarities with the Nimboran set, as mentioned earlier) is necessarily nonsingular in reference, being 1SG + 2SG, it must appear with -nag; there is no pronominal form *io in Mlap, only iomnag. The full set of pronouns is given below; the -m is an oblique marker, added to the pronominal bases before any other suffixation (either plural marking or other case marking, as in ga-m-se '1SG.ACCUSATIVE').

<table>
<thead>
<tr>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (EXCL)</td>
<td>n̄a</td>
</tr>
<tr>
<td>12 (INCL)</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>ko</td>
</tr>
<tr>
<td>3</td>
<td>no</td>
</tr>
</tbody>
</table>

We can see how the Mlap system represents a reanalysis of both the Nimboran pronominal bases and the Kemtuik derivational -nag, and that the resulting system and bases have taken on functional loads that are quite different from either of their sources.

5. It is extremely unlikely that kamorng ‘second-person plural’ arose from kami ‘first-person plural exclusive’ + -ong ‘plural’, because kami does not include the second person in its scope, and thus does not provide a starting point for the development.
noun and a second-person singular polite pronoun (Phil Quick [pers. comm.] also reports this phenomenon in Palu, Central Sulawesi). Again, the extension of meaning from 'you and I' to just one of these referents, 'you', is not surprising. This is found in a range of varieties of Malay, sometimes, as is reported for Sarawak Malay, with a phonetic difference between the 2SG.HON kita? and the 1PL.INCL kita. Needless to say, this shift in the meaning of *kita? is not compatible with the (full) FIRST MALAY POLITENESS SHIFT. In the following discussion, we shall examine the developments of the nonsingular pronouns, bearing in mind that the FIRST MALAY POLITENESS SHIFT offers a plausible, though not always attested, prior step to the spread of *kita?-derivatives in the plural, though it does not offer an explanation for why the originally exclusive form was chosen as the basis for regrammaticalization.

A common development from the FIRST MALAY POLITENESS SHIFT is that the exclusive pronominal form remains in use, but the inclusive form derived from *kita? has encroached on its territory, and can be used with both exclusive and inclusive reference (this is the case with Kelantan Malay kita [Brown 1927:5–6] and Kupang Malay katong/ko-tong [Steinhauer 1983:50]). In other regional varieties, this development has gone a stage further: the opposition between an inclusive form and an exclusive form has disappeared and there is only one first-person nonsingular pronoun, functioning as a general plural. In each case, this sole surviving form derives from the Proto-Malayic inclusive *kita? . This is the case in at least the following nonstandard varieties of the language, as shown in table 2.

TABLE 2. LOSS OF THE KAMI/KITA DISTINCTION IN NONSTANDARD VARIETIES OF MALAY

<table>
<thead>
<tr>
<th>VARIETY</th>
<th>FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Betawi (Jakarta Malay)</td>
<td>kite</td>
</tr>
<tr>
<td>(Southeast) Sulawesi Malay</td>
<td>kita</td>
</tr>
<tr>
<td>Ambonese Malay</td>
<td>katong</td>
</tr>
<tr>
<td>North Moluccan Malay</td>
<td>kitong</td>
</tr>
<tr>
<td>Merauke (South Irian) Malay</td>
<td>kita</td>
</tr>
<tr>
<td>Serui (North Irian) Malay</td>
<td>kitong</td>
</tr>
</tbody>
</table>

What is important to note here is not the form of the pronoun itself and its relationship to the forms used in standard Indonesian, but the fact that in all cases there is only one first-person nonsingular pronoun, and that in all cases it is derived from the *kita? inclusive pronoun (often with the addition of -ong, from orang 'person', grammaticalized as a plural marker in several eastern varieties of Malay) (Serui Malay does not entirely conform to this generalization, and its differences will be dealt with separately). The corollary of this observation is that we never find the *kami exclusive pronominal form used when the inclusive/exclu-
sive distinction has been lost, that is, functioning as a generic first-person nonsingular pronominal form.

Even in varieties of the language in which the *kamilkita* opposition is prescribed usage, we find cases in which the *kita* form is dominant. In a particularly detailed survey of spoken colloquial Indonesian (conducted in Yogyakarta, but based on speakers from a variety of backgrounds) Englebretson (1997 pers. comm.) notes that in his 36,825-word corpus of spontaneous conversation in natural context, there are no occurrences of *kami* as opposed to 327 of *kita*. He continues: "Many of these tokens occur in contexts where [...] one would ‘expect’ to get *kami*. For example, one of my segments consists of people sitting around a lunch table, and the conversation has turned to the subject of pickpockets. People started telling each other personal narratives about times when they had been robbed, or of robberies they had seen. A lot of the narratives took the form of: I was with a group of friends, we (*1PL.INCL*) went to such-and-such, we (*1PL.INCL*) saw such-and-such, etc. Semantically, one would expect the exclusive form *kami* here, since the speaker is recounting something that the rest of the interlocutors did not take part in. Yet she still uses the inclusive pronoun. When I was glossing this segment, I asked my assistant why they were using *kita* and not *kami*. My assistant said that, basically, people don’t ever say *kurni*, especially in informal contexts. My database certainly seems to bear this out.” These languages are ones in which the collapse of the distinction is absolute, in all cases favoring retention of the inclusive form. This is summarized in table 3, with pronominal forms taken from Betawi, but typical for all the nonstandard varieties of Malay/Indonesian discussed so far.

<table>
<thead>
<tr>
<th>TABLE 3. COLLAPSE OF THE INCLUSIVE/EXCLUSIVE DISTINCTION IN NONFORMAL VARIETIES OF MALAY/INDONESIAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proto-Malayic</td>
</tr>
<tr>
<td>1PL.INCL</td>
</tr>
<tr>
<td>1PL.EXCL</td>
</tr>
<tr>
<td>1SG</td>
</tr>
</tbody>
</table>

We can conclude that, in contrast to formal Malay/Indonesian, and many (if not most) of the Austronesian languages, the local varieties of Malay have, for one reason or another, lost the inclusive/exclusive distinction, and always in the same direction: the *kami* pronoun is lost, and the *kita* pronoun, often with some phonological or structural changes, is extended to cover all first-person nonsingular uses.

5.1 REINVENTING THE EXCLUSIVE PRONOUN. Interestingly, there are at least two cases recorded in which, having lost the inclusive/exclusive distinction encoded by *kita* and *kami*, the local variety of Malay has reinvented the distinction. In Kupang Malay, an eastern variety in which the use of *-ung* is regular in the formation of plural pronouns, we find that there is an (optional) opposition between an
inclusive form *katong/kotong, and an exclusive form based on the (necessarily exclusive) first-person singular form, beta, giving the first-person pronominal paradigm found in table 4.

**TABLE 4. KUPANG MALAY PRONOMINAL FORMS**

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>I INCL</td>
<td>—</td>
<td>katong/kotong</td>
</tr>
<tr>
<td>I EXCL</td>
<td>beta</td>
<td>betong</td>
</tr>
</tbody>
</table>

* from Steinhauer 1983:50

We may suppose that the development of this pronominal system followed the path given in table 5, with the initial spread of the I PL. INCL pronominal form to cover all nonsingular uses, and the subsequent reinvention of the inclusive/exclusive distinction by the formation of an exclusive form based on the singular (and, necessarily, exclusive) pronoun.

**TABLE 5. DEVELOPMENT OF THE FIRST-PERSON PRONOUNS IN KUPANG MALAY**

<table>
<thead>
<tr>
<th>SOURCE VARIETY OF MALAY</th>
<th>INTERMEDIARY STAGE</th>
<th>MODERN KUPANG MALAY</th>
</tr>
</thead>
<tbody>
<tr>
<td>I PL. INCL</td>
<td>*kita</td>
<td>katong/kotong</td>
</tr>
<tr>
<td>I PL. EXCL</td>
<td>*kami</td>
<td>katong/kotong</td>
</tr>
<tr>
<td>1 SG</td>
<td>*beta</td>
<td>beta</td>
</tr>
</tbody>
</table>

Another interesting case of the loss and (partial) reintroduction of an inclusive/exclusive contrast is found in the Malay of Serui, in which a form derived from *kitu? (kitong /tong, similar to, and, given the patterns of trade in the region, probably derived from, the North Moluccan Malay kitong) is used for both functions, yet an apparent reintroduction of the kita form (from Standard Indonesian?) for the inclusive only, without the complete loss of kitong in this function, has led to the partial reintroduction of the distinction (other forms of North Irian Malay have kitong for first-person plural, but lack the innovative use of kita found in Serui). In the reduced set pronouns, possibly verbal clitics, there is no inclusive/exclusive distinction. The first-person plural pronominal forms in Serui Malay are given in table 6.

---

6. "The reductive forms are usually reserved for the grammatical subject; in the object position, we often find the full forms" (van Velzen, 1995:327).
Despite the contrast in the inclusive forms, van Velzen (1995:326) notes that “[a] considerable number of speakers do not distinguish between the inclusive and exclusive forms. Some of them may do so, but not very consistently.” This points to a language basically without an inclusive/exclusive distinction that has only recently acquired a distinct form for the inclusive. The development of the pronominal forms in this variety is as seen in table 7, showing the initial loss of the inclusive/exclusive distinction (quite possibly before the existence of Serui Malay as an entity separate from North Moluccan Malay), and the later reintroduction of an inclusive form from an outside source.

Note that the developments from a stage that does not make the inclusive/exclusive distinction in Kupang Malay and Serui Malay have important differences (Kupang Malay has innovated a new exclusive form, whereas Semi Malay has borrowed a new inclusive form), but in both cases, despite the reintroduction of the exclusive form, the original *kami pronoun has been totally lost from the language. Not all developments of the exclusive pronoun away from its original function are so dramatic, and we also find examples of the loss of the inclusive/exclusive distinction in a language without the loss of the forms used to make that distinction, but with new meanings attached to them, discussed in the next section.
between the inclusive and exclusive forms. In addition to the loss of one of the
(originally nonsingular) forms (always *kami) described above, possibly as a
result of the changes arising from the reanalysis of the other (always *kita) as a
singular pronoun, there is another attested development of the Austronesian first-
person nonsingular pronouns that supports this claim of "core" status for the
inclusive and "noncore" for the exclusive.

When eliciting wordlists in local languages of Buton, the first author has dis-
covered that informants are regularly confused as to why, immediately after elicit-
ing a local equivalent for Indonesian *kita, he would then attempt to elicit a form
for Indonesian *kami as well, since the translation equivalent of this item had just
been recorded. An idealized, but typical, (translated) conversation would be simi-
ter to the one that follows:

MD: "How about the word *kita?"
Respondent(s): "*Ikita."

MD: "Good. What about *kami, then?"
Respondent(s): (Pause, glances, worried looks in MD's direction)
      "We've already told you. *Ikita. See, you've written it down."

MD: (Sigh) "Yes, that's for Indonesian *kita. But what about *kami?"
Respondent(s): "What do you mean exactly? It's the same, isn't it?"

(discussion ensues on how to differentiate these forms in Indonesian)

In order to elicit the local language's reflex of Proto-Austronesian *i-[k]ami, as
well as that of *i-[k]ita, the paraphrases *kita banyak orang 'we, a lot of people' or
kita semua 'all of us' for the reflex of *kita?, and *kita sedikit orang (saja) 'we, a
few people (only)' for the reflex of *kami proved to be the most reliable means
available. In this manner, forms such as *kita and *sami were elicited (in the Lasal-
imu and Kumbewaha languages, for instance), where asking for the equivalent of
Indonesian *kami tended to produce simply *kita as a response, not the form *sami (<
*kami, by a [semi-]regular *k > s sound change). René van den Berg (pers.
comm.) reports resorting to the same procedure in his survey of the western parts
of the Muna-Buton area. In languages that have not undergone this shift in mean-
ning of the nonsingular forms, such as Wolio, elicitation of the Indonesian inclu-
sive/exclusive distinction does not present any problems, resulting in *ingkami for
the exclusive, and *ingkita for the inclusive.

In addition to informing us about the loss of a distinction in the use of the pro-
nouns in Southeast Sulawesi Malay (reported in section 5), these findings also
offer an insight into the oppositions encoded in the local languages. Note the data
in table 8 showing elicitation cues (with translations of the meanings of the forms
as used in standard Indonesian), and the response elicited from informants in
Lasalimu (spoken in the village of the same name in eastern Buton, and typical of
many locations on Buton). We can conclude that the difference between (Indone-
sian) *kami and *kita is not recognized outside the educated elite, and that in at least
some languages of Southeast Sulawesi, the inclusive/exclusive opposition has
been exapted into an opposition of number. The distinction between inclusive and
exclusive forms has ceased to exist, and the form derived from *kita now encodes
TABLE 8. ELICITING PRONOUNS IN LANGUAGES OF SOUTHEAST SULAWESI

<table>
<thead>
<tr>
<th>ELICITATION CUE (INDONESIAN)</th>
<th>RESPONSE (LASALIMU)</th>
</tr>
</thead>
<tbody>
<tr>
<td>kita</td>
<td>kita, some hesitation</td>
</tr>
<tr>
<td>(1PL.INCL)</td>
<td></td>
</tr>
<tr>
<td>kami</td>
<td>confusion, kita</td>
</tr>
<tr>
<td>(1PL.EXCL)</td>
<td></td>
</tr>
<tr>
<td>kita banyak orang</td>
<td>kita</td>
</tr>
<tr>
<td>(1PL.INCL many people)</td>
<td></td>
</tr>
<tr>
<td>kita sedikit orang (saja)</td>
<td>sami</td>
</tr>
<tr>
<td>(1PL.INCL few people only)</td>
<td></td>
</tr>
<tr>
<td>kami banyak orang</td>
<td>kita</td>
</tr>
<tr>
<td>(1PL.EXCL many people)</td>
<td></td>
</tr>
<tr>
<td>kami sedikit orang (saja)</td>
<td>sami</td>
</tr>
<tr>
<td>(1PL.EXCL few people only)</td>
<td></td>
</tr>
</tbody>
</table>

a first-person plural, while the form that continues *kami is the exponent of first-person paucal (regardless, in each case, of whether or not the addressee is included). Donohue (1995:110), for instance, notes of Tukang Besi that “there are two pronominal sets corresponding to the first person nonsingular, ikami and ikita. ... These pronouns are used with a lot of overlap by most speakers, but the ikami forms usually refer to a small group of people, typically two to four, and the ikita forms typically refer to groups of four or more. Note that the fundamental difference between the two is not one of exclusive/inclusive, as is usually the case in Austronesian languages, but of ‘paucal’ versus ‘plural’ reference.” These languages show a different path for the development of the inclusive/exclusive distinction from that described earlier for the local varieties of Malay, in that the formal opposition is preserved, while the meaning difference associated with this opposition has changed. Compare table 9 with table 3 in section 5.

It is significant that the development of the paucal/plural distinction is found in a variety of languages, not all of which can be subgrouped together with a strong degree of confidence, and that in all cases the old exclusive pronoun is the one that has acquired the paucal meaning. A possible explanation for this can be found when we consider the relative markedness of paucal and plural forms.

Cross-linguistically we find that the basic nonsingular pronoun is the plural one, not a dual or trial (or paucal, which also covers the ranges expressed by the categories ‘dual’ and ‘trial’). Greenberg (1966:75–76) discusses the implicational hierarchy operating between singular, plural, and dual forms, concluding that the presence of a dual form is “dependent” on the existence of a plural form in the language, and that a dual form is marked with respect to the plural form, as well as being much less frequent in texts (though figures in Harlow and Thornton 1986 show that in certain genres of Māori discourse, specifically love songs and songs
of mourning in which the singer addresses the loved one or the deceased, the use of the dual outnumbers that of the plural by four to one).

**TABLE 9. REANALYSIS OF THE INCLUSIVE/EXCLUSIVE DISTINCTION IN TUKANG BESI (AND OTHER LANGUAGES OF SOUTHEAST SULAWESI)**

<table>
<thead>
<tr>
<th>PROTO-AUSTRONESIAN</th>
<th>TUKANG BESI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 PL. INCL.</td>
<td>*i-[k]ita  &gt; i kita</td>
</tr>
<tr>
<td>1 PL. EXCL.</td>
<td>*i-[k]ami &gt; ikami</td>
</tr>
<tr>
<td>1 SG</td>
<td>*i-aku &gt; i aku</td>
</tr>
</tbody>
</table>

* Since Tukang Besi has verbal agreement and possessive marking, there are other bound forms corresponding to the forms in this table as well, not illustrated here. They are described in Donohue (1995:109), but all reflect the same original forms.

Foley (1986:72), discussing pronominal systems in the Papuan languages of New Guinea, notes that, in addition to a singular/plural distinction, “... a few Papuan languages add a paucal, or sometimes a trial, in addition to a plural and a dual ... in which the dual and trial forms are derived from the plural ...”. He cites data from Kiwai (see table 10) showing a clear morphological derivation of the dual and trial forms from the relevant plural form, giving further support to the claim that the “core” nonsingular form is a basic plural, not a more specialized ‘dual’ or ‘paucal’.

Examining textual material in Tukang Besi to determine the frequency of the pronouns, we find that there is a clear basis for deciding the markedness relationship between the two forms in question. Examining a small corpus of 3,000 words—the texts found in Donohue (1995), representing a variety of discourse genres (see footnote 1 on page 70)—we find the frequencies of nonsingular pronominal forms (both bound and free) as shown in table 11.

**TABLE 10. KIWAI PRONOUN FORMS**

<table>
<thead>
<tr>
<th>SG</th>
<th>DUAL</th>
<th>TRIAL</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>mo</td>
<td>nimo-to</td>
<td>nimo-ibi</td>
</tr>
<tr>
<td>2</td>
<td>ro</td>
<td>nigo-to</td>
<td>nigo-ibi</td>
</tr>
<tr>
<td>3</td>
<td>nou</td>
<td>nei-to</td>
<td>nei-bi</td>
</tr>
</tbody>
</table>

* from Foley 1986

We find a total of 31 nonsingular pronominal forms, of which only 19% are paucal in reference, and 81% plural, giving a ratio of 4.2:1 in favor of the plural. Just as we could see (from the much larger database examined) that the inclusive forms were unmarked in discourse with respect to the exclusive forms in written
TABLE 11. FREQUENCY OF FIRST-PERSON NONSINGULAR FORMS IN TUKANG BESI

<table>
<thead>
<tr>
<th>Total Words</th>
<th>3,000</th>
<th>Percentage of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Occurrences of IPlur</td>
<td>25</td>
<td>0.83%</td>
</tr>
<tr>
<td>Total Occurrences of IPauc</td>
<td>6</td>
<td>0.20%</td>
</tr>
</tbody>
</table>

Indonesian, both modern and classical, so too can we state that the plural forms are unmarked relative to the paucal in the Tukang Besi materials examined.

7. LINGUISTIC OPPOSITION AND CHANGE. From this survey of the evolution of first-person nonsingular pronouns in Malay/Indonesian, and the (non-Malayic) languages of Southeast Sulawesi, we have been able to identify five different paths of development. In the first of these, exemplified by Kelantan Malay, both forms are present, but the usage of kita is extending to cover what was previously the domain of kami, but is now shared between the two forms. Both pronouns are still synchronically attested, but the range of one is clearly expanding. This is shown schematically in table 12.

TABLE 12. INCLUSIVE AND EXCLUSIVE IN KELANTAN MALAY

<table>
<thead>
<tr>
<th>EXCL</th>
<th>INCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>*kami</td>
<td>*kita?</td>
</tr>
<tr>
<td>kami</td>
<td>kita</td>
</tr>
</tbody>
</table>

The second case is an extension of this one, attested in numerous local varieties of Malay/Indonesian, and sees the kami form disappearing completely, with kita (or its reflex, see table 2) now the sole first-person nonsingular form. These varieties now exhibit a general plural form, and have no lexical means of encoding the inclusive/exclusive distinction. This is shown in table 13, illustrated with data from Betawi.

TABLE 13. INCLUSIVE AND EXCLUSIVE IN BETAWI

<table>
<thead>
<tr>
<th>EXCL</th>
<th>INCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>*kami</td>
<td>*kita?</td>
</tr>
<tr>
<td>kite</td>
<td>kite</td>
</tr>
</tbody>
</table>
In a logical development from this stage, we have two attestations of the (partial) reintroduction of the inclusive/exclusive distinction. Interestingly, neither of these cases involves the use of a form cognate with *kami (even when borrowing is employed as a strategy to reestablish the distinction). In Kupang Malay, we find that the first-person singular pronoun beta, necessarily exclusive, can take the plural marker -ong and be used as a first-person nonsingular exclusive pronoun. This is shown in table 14.

**TABLE 14. INCLUSIVE AND EXCLUSIVE IN KUPANG MALAY**

<table>
<thead>
<tr>
<th>EXCL</th>
<th>INCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>*kami</td>
<td>*kita?</td>
</tr>
<tr>
<td>beta</td>
<td>katong</td>
</tr>
</tbody>
</table>

The other attested development that reinstates the inclusive/exclusive distinction is found in Serui Malay (see table 15), in which the form kita is borrowed from Standard Indonesian as an alternative for the inclusive form. The “native” Serui Malay first-person nonsingular kitong, itself derived from *kita?, is still used for both inclusive and exclusive reference.

**TABLE 15. INCLUSIVE AND EXCLUSIVE IN SERUI MALAY**

<table>
<thead>
<tr>
<th>EXCL</th>
<th>INCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>*kami</td>
<td>*kita?</td>
</tr>
<tr>
<td>kitong</td>
<td>kitong/kita</td>
</tr>
</tbody>
</table>

Finally, we have seen examples from Southeast Sulawesi in which the formal opposition between forms is maintained, but in which the distinction between them is regrammaticalized, no longer encoding inclusive/exclusive, but now plural/paucal. Although the features encoded by the opposition have changed, the relationship between them remains constant, with the unmarked form preserving its unmarked status. This can be seen in table 16.

As was pointed out before, we find that, in all of these different patterns of development, it is the original inclusive form that is preserved and generalized to other parts of the nonsingular paradigm, or to become the unmarked member of
the new paradigm. Furthermore, when the distinction is reintroduced by borrow-
ing from a language that has both an inclusive and an exclusive form, the borrow-
ing is of the inclusive form; we have no examples of the exclusive form being
borrowed to reintroduce this distinction.

In relevant respects, the evolution of the first-person nonsingular pronouns in
Austronesian parallels the development into Romance of the oblique forms of the
Latin first- and second-person singular pronouns discussed by Smith (1995, 1997). When one of the forms disappears, it is the “noncore” form (the form that
is qualitatively marked and less frequent); and, when the opposition is refunc-
tionalized, it is the “core” form of the original opposition that takes on the “core”
function in the new opposition. These findings provide further support for the
claim made by Smith (1997) that “[e]ven after a morphological opposition has
closed to encode a particular functional opposition, it can still retain a more
abstract value that can guide its refunctionalization; and it might be the case that it
cannot be refunctionalized unless this residual opposition is present.” They also
demonstrate that skeuomorphy is not simply an idiosyncratic property of the
Romance data, but is also a characteristic of refunctionalization processes in Aus-
tronesian languages. The fact that the same principle appears to guide develop-
ments in separate categories in totally unrelated language families implies that the
possibility of meaningful morphological opposition existing without exponence
should be taken into account more generally in work on historical linguistics.

Having shown that we have a loss of the otherwise near-universal (among Aus-
tronesian languages) opposition between inclusive and exclusive first-person non-
singular pronouns, we can speculate on the reasons why it happens to be Malay/
Indonesian that is the language so affected. It has long been established that varieties
of Malay are prone to borrow pronouns from other sources; witness 1SG saya (< Sanskrit) in Standard Malay/Indonesian (as well as the use of kita in this function as
reported in section 5), 2SG ose (< Portuguese) in Ambonese Malay, 1SG gua (< Hok-
kien [Min] Chinese) in Betawi, 2SG lu (< Hokkien [Min] Chinese) in both Betawi
and Kupang Malay, and yu (< English) in the formal speech of educated people in
many regions. It is worth noting that none of these source languages, nor any of the
other languages that have had a strong influence on the linguistic history of Malay/
Indonesian, maintains an inclusive/exclusive distinction in the nonsingular pro-
nouns. It is speculative, but tempting, to suggest that we have a case here of a linguist-
ic paradigm (lack of opposition between inclusive and exclusive forms of the first-
person nonsingular pronoun) being borrowed, without the form itself being trans-
mitted. It is, however, beyond the scope of this article to pursue this hypothesis fur-
ther, and it remains a topic for future research.

REFERENCES


———. 1997. Merauke Malay: some observations on contact and change. MS, Department of Linguistics, University of Manchester.


Ross, Malcolm. 1997. Reconstructions of Austronesian case markers and pronouns. MS, Department of Linguistics, Research School of Pacific and Asian Studies, Australian National University, Canberra, Australia.


van Velzen, Paul. 1995. Some notes on the variety of Malay used in the vicinity of Serui. In Tales from a concave world: Liber amicorum Bert Voorhoeve, ed. by Connie Baak, Mary Bakker, and Dick van der Meij, 311–343. Leiden: Projects Division, Department of Languages and Cultures of Southeast Asia and Oceania, Leiden University.


Mark Donohue
Department of Linguistics
University of Manchester
Manchester M13 9PL
England
donohue@cheops.anu.edu.au

John Charles Smith
St. Catherine’s College
Oxford, OX1 3UJ
England
johncharles.smith@stcatz.ox.ac.uk