ALEXANDER ADELAAR

Much ado about *di-*

1. Introduction

Discussions about the origin of the Malay passive marker *di-* have been ongoing for the last sixty years.¹ They have somewhat intensified recently because of their wider implications for typological linguistics. Throughout the years, several explanations have been proposed for the origin of *di-*, some of which are not necessarily mutually exclusive:

1. *Di-* is originally a cliticized form of the third person pronoun *dia*.
2. *Di-* developed from a preposition *di*, which became cliticized to the root of the following passive verb.
3. *Di-* developed from an earlier passive prefix **ni-* (still attested in seventh-century Old Malay inscriptions) through denasalization of initial *n-*.
4. Old Malay had the verbal prefixes *bər-* and *di-*, but in the seventh-century court language of the South Sumatran inscriptions, these had been replaced by respectively *mar-* and *ni-* which were Batak loan affixes; sub-dialectally and in later inscriptions, however, *bər-* and *di-* remained in use and eventually outlived *mar-* and *ni-* (Aichele 1942-43; see also Mahdi 2005).

One other explanation has recently been put forward by Wolff (2002):

5. The prefix is a conflation of a third person plural enclitic pronoun *da-* (< Proto Malayic **sida* 'third person plural') and the passive prefix *ni-* (Wolff 2002).

¹ I am grateful to Anthony Jukes and John Hajek for their useful feedback on an earlier version of this article. They are in no way responsible for the final version.

Alexander Adelaar is Associate Professor and Reader at the Melbourne Institute of Asian Languages and Societies (MIALS) and holds a PhD from Leiden University. His main field of academic interest is Austronesian historical and descriptive linguistics, especially regarding the languages of Madagascar, Taiwan and Borneo. He is the author of *Proto-Malayic*, Canberra: Pacific Linguistics, 1992, and *Salako or Badameà: Sketch grammar, texts and lexicon of a Kanayatn dialect (West Kalimantan)*, Wiesbaden: Harrassowitz, 2005. Professor Adelaar may be reached at MIALS, University of Melbourne, VIC 3010, Australia. <s.adelaar@unimelb.edu.au>

In a recent article in *BKI*, René van den Berg (2004) discusses most of these theories. He believes that *di-* is an inherited prefix and continues an older **ni-*, which is still attested in Old Malay (explanation 3). This theory was proposed for the first time by Teeuw in 1959. Van den Berg challenges on several counts my position (based on Aichele 1942-43) that *di-* developed from the preposition *di* (explanation 2).

I have addressed this matter in various studies involving the history of the Malayic linguistic subgroup² (Adelaar 1992a:162-3, 1992b, 2002:22, forthcoming). I generally agree with Van den Berg's assessment and rejection of explanations 1 and 5 and will not discuss them any further. Explanation 4 has recently been given some support by Mahdi. I basically maintain the view that I expressed in Adelaar (1992a, 1992b): I do not a priori exclude the possibility of such a course of events, but nor do I find it totally convincing. To assume Batak influence in Srivijaya seems rather speculative, given our limited knowledge about this empire and the early history of Malay. More important is the argument proposed by Teeuw (1959) that the nature of the linguistic data does not require us to look for solutions outside the history of the Malayic subgroup itself.

In Adelaar (1992:162-3) I opt for explanations 1 and 2, but in later publications (Adelaar 1992b:404, 2002:22, forthcoming) I have come to favour explanation 2 only, and this is also the position I take here. As Van den Berg in his recent article challenges this position and argues for explanation 3, my discussion here concentrates on the validity of explanations 2 and 3.

In Section 2 I give a short explanation of my arguments for claiming that the passive marker *di-* originated from a preposition. In Section 3 I give a summary of Van den Berg's arguments in favour of *di-* as a continuation of **ni-*, and I test several aspects of his argumentation. Finally, in Section 4, I summarize and evaluate my findings.

2. *Why I think that the Malay passive marker di- was originally a preposition *di*

There are various Old Malay inscriptions. The very oldest among them, the seventh-century inscriptions from South Sumatra and Bangka Island, exhibit a passive verbal prefix *ni-* and an intransitive verbal prefix *mar-* which in general correspond to modern Malay *di-* and *bər-* respectively. Later Old Malay inscriptions (some of which were found in West Java and in the Philippines) already have *di-* and *var-* (where the 'v' probably reflects *b*, see Vikør 1988).

² The Malayic subgroup consists of Malay in all its regional and sociolinguistic variety and furthermore of various languages that are sufficiently close to it so as to form a genetic subgroup, including Minangkabau, Iban, Kendayan, Urak Lawoi', and Kerinci.

As far as Malay *di-* is concerned, I have argued that this developed from the preposition *di*. In modern Malay *di* functions as a locative preposition, but historically it must have had a much wider application. Minangkabau and Kelantan Malay, for instance, use it also as an agent preposition; Kendayan uses it as a preposition to indicate source (*di mae?* [from where] 'where from?') as well as an agent preposition. (In Malay, agent and source are expressed by the prepositions *oleh* and *dari* respectively.)

Kendayan provides strong evidence for the shift from a preposition to a passive prefix. In this language, the agent immediately precedes the verbal base and may be marked by the preposition *di-*. If there is no agent (and only if there is no agent), the following verbal base is marked by a prefix *di-*. Compare the following sentences from the Salako dialect of Kendayan (note that in this language nasalization also applies in passive constructions denoting a completed action):

(1)	<i>Bini-e</i>	<i>dah</i>	<i>mati</i>	<i>di</i>	<i>kayo</i>	<i>munuh</i>	[N-bunuh]
	wife-3.POS	PERF	dead	PREP	enemy		N-kill
	'His wife was killed by the enemy'						
(2)	<i>Bini-e</i>	<i>dah</i>	<i>mati</i>	<i>dimunuh</i>			[di-N-bunuh]
	wife-3.POS	PERF	dead				PASS-N-kill
	'His wife was killed'						
(3)	* <i>Bini-e</i>	<i>dah</i>	<i>mati</i>	<i>di</i>	<i>kayo</i>	<i>dimunuh</i>	[di-N-bunuh]
	wife-3.POS	PERF	dead	PREP	enemy		PASS-N-kill
	'His wife was killed by the enemy'						

I do not believe that *di-* is a continuation of Old Malay *ni-* (which in turn reflects Proto Malayic **ni-*, and Proto Malayo-Polynesian **ni-/ *<in>*), because a change from initial **n* to *d* (called 'denasalization') would be unprecedented as it is not a regular sound change in the phonological history of Malay.

Of course, there is also another Malay prefix that seems to have undergone denasalization, namely the intransitive verbal prefix *bər-*, which corresponds to Old Malay *mar-*. Various scholars (including Van den Berg) have therefore argued that the development from **ni-* and **mar-* to modern *di-* and *bər-* respectively can be brought under the same denominator and are both the result of denasalization. However, I contend that the change from **mar-* to *bər-* is due to other developments. I argue (Adelaar 1992:163) that *bər-* must have developed from Proto Malayic **mar-* via an intermediate stage, in which the reduction of the antepenultimate vowel to *ə* triggered the emergence of (an 'epenthetic') **b* after **m*; the resulting initial **mb* cluster was subsequently reduced to *b-* (that is, **mar-* > **mər-* > **m(b)ər-* > **(m)bər-* > *bər-*; the specific phonological condi-

tions for this development will be demonstrated in more detail in Section 3.2). On this account, there are no analogical developments in Malay morphological history that support a phonological change from **ni-* to *di-*.

I also do not believe that *di-* is the continuation from a Proto Malayic prefix **di-*. If it were, it would have undergone antepenultimate neutralization, a very strong tendency in the phonological history of Malay to weaken all vowels before the penultimate syllable to ə, for example Sanskrit *kāpala* 'skull' > *kəpala* 'head'; (Proto Malayic) **sumangət* 'spirit, inspiration' > *səmangat*; **biruang* 'bear' > *bəruang*.² This tendency also affected prefixes and infixes, which usually take an antepenultimate position as the following root almost always consists of (minimally) two syllables, for example the verbal prefix **maN-* > *məN-*, and the verbal infix <*um*> became <*əm*>, as can still be seen in fossilized forms like *turun-t<əm>urun* 'offspring, descendants'. In summary, then, if *di-* was inherited, it should not have kept its *i*, and it would have become **də-*. This argument also strengthens my first point, that *di-* cannot be a continuation of **ni-*: if it were, it would be inherited and its vowel would have become ə. The fact that it still has *i* indicates that it has been added to the Malay affix system recently enough to escape the neutralization tendency.

3. An assessment of Van den Berg's arguments for the claim that *di-* continues **ni-*

Van den Berg's belief that *di-* originated from **ni-* through denasalization is based on various arguments. These can be summarized as follows:

1. Sound changes that have generally affected the lexicon have not necessarily also affected prefixes.
2. There is evidence for maintenance of **i* in prefixes.
3. It is unconvincing that **mar-* would have developed into *bər-* via an intermediate stage where initial **m* was followed by an epenthetic *b*.
4. *Bər-* and *di-* have both undergone the same sound change applying only to prefixes and consisting of denasalization of the initial voiced consonant.
5. *Di-* and *ni-* were in free variation in Old Malay.
6. Denasalization of **ni-* to *di-* is more common than it seems because a parallel development is seen in South Sulawesi languages.
7. The development from a preposition to a voice marker is typologically unlikely.

² Unless indicated otherwise, lexical data from standard and dialectal Malay as well as information on the (Sanskrit, Arabic, Portuguese and Dutch) sources of loanwords in Malay are all based on Wilkinson 1959. Proto Malayic information is taken from Adelaar 1992a, and data on the Kendayan (Salako dialect) from Adelaar 2005.

3.1 As a rule, affixes are not exempt from sound changes affecting the lexicon

As I demonstrated above, antepenultimate neutralization of vowels to schwa has had a very strong effect on both the lexicon and affixes of Malay, and *di-* is the only prefix that has not undergone it. While acknowledging these facts, Van den Berg (2004:541) maintains that *di-* is inherited and reflects **ni-* because ‘there is no a priori reason why *i* in prefixes could not have been exempt from this vowel shift’.

I disagree that there is no a priori reason to assume that affixes have undergone the same phonological changes as the lexicon. While it is true that the sound changes they exhibit occasionally deviate from those reflected in the lexicon, in most cases they do not, and there is certainly no license to ignore rules of sound change wherever they do not fit into our overall theory. The default situation remains that we have to expect the same sound changes until there is clear evidence that the conditions for these changes do not apply. If sound changes do not apply it is often due to ‘analogical formation’ in a specific grammatical environment (Bynion 1977:44). Although *di-* is a passive marker, I cannot find any analogy or other conditioning factor that could explain why it has escaped the expected vowel neutralization.

In Section 5.2.3 Van den Berg uses the *Tukang Besi* suffix *-su* ‘my’ as an example of ‘a sporadic sound change limited to one particular affix’. Although Proto Malayo-Polynesian **k* as a rule remained *k* in *Tukang Besi*, this suffix allegedly reflects Proto Malayo-Polynesian **-ku* ‘first person singular genitive’. In this case there is very little reason indeed to assume, with Van den Berg, that *-su* is a reflex of Proto Malayo-Polynesian **-ku*. It seems much more likely that there is no sound correspondence at all involved and that *-su* is borrowed or has acquired its present meaning recently through a semantic shift.

3.2 Antepenultimate **i* was maintained as *i* in proclitics but not in prefixes

In order to demonstrate that *di-* has been exempt from the antepenultimate neutralization rule, Van den Berg (2004:541) tries to find other prefixes that have maintained **i* as *i*, as they would strengthen the theory that antepenultimate **i* was maintained in prefixes. He argues that the personal article *si* is comparable to *di-* in that both are proclitics because they can never stand alone. (The fact that orthographically *si* is written as a separate word, whereas *di-* is added (without hyphen) to the following verbal stem, is of course of little weight in a linguistic discussion.)

Van den Berg probably has good reason to characterize both of them as proclitics. However, by doing so, he actually makes the crucial observation that neither of them has become fully grammaticalized, which supports rather than contradicts my point that *di-* has only recently become a passive verb marker. Proclitics are not a fully integrated part of the structure of the word

they belong to, and this is why they sometimes exhibit phonological features that are not shown in prefixes. In the case of *di-* and *si*, this explains why they have maintained **i* as *i*. If *di-* and *si* have remained proclitics instead of having become fully grammaticalized prefixes, this actually emphasizes how ill-adapted *di-* has remained to the word structure of Malay. It strongly suggests that it has been added relatively recently to the Malay affix inventory, and this in turn implies that it cannot have developed from Old Malay *ni-*, which is a reflex of Proto Malayic **ni-* and Proto Malayo-Polynesian **ni-/*<in>*.

The fact that Mualang (West Kalimantan) has *də-* or *da-*, Orang Akit (Riau Archipelago) has *də-*, and Orang Darat (Riau Archipelago) has *j-* (before an initial vowel) as a prefix corresponding to Malay *di-*, only shows that in these Malayic varieties the passive marker has become more grammaticalized than in Malay (compare Section 5.1.3 in Van den Berg).

There is in fact a fully grammaticalized prefix that reflects historical **i*. As expected, it has undergone the antepenultimate vowel neutralization. The reciprocal prefix *sə-* is usually not recognized as an independent prefix in Malay/Indonesian grammars because it has become homophonous with the numeral prefix *sə-* 'one; (of) the same' (as in *kawan sə-rumah* 'housemate'), and also because it is almost always preceded by *bər-*. However, it must reflect an earlier Proto Malayic and Proto Austronesian reciprocal prefix **si-*. Compare the following examples:

<i>tumpu</i> 'take-off, abutment'	<i>bər-sə-tumpu</i> 'to take off against each other (for instance in tug-of-war game) (Kamus Dewan 1994)
<i>tubuh</i> 'body'	<i>bər-sə-tubuh</i> 'to have sex' (Echols and Shadily 1989)
<i>kongkol</i> 'to discuss, gossip'	<i>bər-sə-kongkol</i> 'to conspire' (Echols and Shadily 1989)

Some other Malayic varieties still exhibit *i* in their cognate prefixes, such as Minangkabau, which has a reciprocal prefix *(ba)si-*, and Kendayan, which has *(ba)si-* and *(ba)si(N)-* forming medial verbs (Adelaar 1992b:395-6). These varieties have undergone antepenultimate neutralization to a much smaller extent than standard Malay.

It is noteworthy that *(bər-)sə-* occurs immediately before the lexical root, in contrast to *si*, which may occur before a prefixed word, for example *si tərdaḳwa* 'the accused', or *si pəmalas* 'lazy-bones'; and also in contrast to *di-*, for example *di-pər-tanya-kan* 'to question (passive form)'. As a fully grammaticalized prefix, *(bər-)sə-* demonstrates clearly that inherited prefixes containing **i* have reacted as positively as any other prefix to the tendency to neutralize antepenultimate vowels to *ə* in Malay phonological history.

3.3 More evidence for the claim that Proto Malayic *mar- became Malay bər-

In Section 5.2.1, Van den Berg attributes historical importance to the parallel between Old Malay *ni-/*Malay *di-* and Old Malay *mar-/*Malay *bər*. According to him, both *bər-* and *di-* have undergone a similar sound change which only applies to prefixes and which consists of denasalization of the initial voiced consonant. He does not accept my theory that the change from **m/-* to *b/-* is due to an intermediate stage in which an epenthetic *b* was inserted, as such insertion is otherwise only attested in clusters that are intervocalic (for instance Arabic *jumlaḥ* > Jakarta Malay *jumbəlè*), and not in clusters at the beginning of a word.

As shown in the previous sections, there are no sound rules in Malay that can explain an alleged change from **ni-* to *di-*. On the other hand, in the case of **mar-*, there is in fact strong evidence that as an antepenultimate initial syllable, **mar/-* was prone to phonological readaptation to *bər/-* in the history of Malay. The vowel in **mar/-* was first reduced to ə through antepenultimate neutralization, which brought about a short syllable with a nasal onset and a coda consisting of a liquid. Malay has various trisyllabic roots beginning with *mər/-* that have variant forms with initial *bər/-*, and trisyllabic roots beginning with *məl/-* that have variant forms with initial *bəl/-*. These roots seldom belong to the core vocabulary and often refer to names of trees, plants, or fish. Compare the following sets:

mərjagong, bərjagong, mənjangong 'a tree (*Ixonanthes* spp.)'
mərtulang, bərtulang, məntulang, mətulang 'a tree, *Canthium didymum*'
bərtimun, məntimun, kətimun 'cucumber gourd'
məragan, bəragan 'life-like in death'
bəlibis, məliwis (the latter used in Java) 'whistling teal'
məlalah, (Perak Malay) bəlalah 'ravenous for food'
məleter, (Penang Malay) bəleter 'to talk long-windedly'
burong bəragi, (Kedah Malay) burong məragi 'painted snipe'

Another pattern of variation involves trisyllabic roots beginning with *mər/-* or *məl/-* and corresponding roots beginning with *məN/-*, compare

mərbulan, məmbulan 'a tree, *Endospermum malaccense*'
mərbuloh, məmbuloh 'a tree, *Gynotroches axillaris*'
mərbatu, məmbatu 'a tree-name (...)'
mərcali, məncali 'a tree, used for making masts for junks; *Lophopetalum* sp.?'
mərpitis, məmpitis, a plant, *Cratoxylon* sp.'
mərtua, məntua 'parent-in-law'.

The tendency for antepenultimate initial **mar/-* sequences to change to *bər/-* can be observed in the adaptation of some loanwords. Sanskrit *mara-bahaya* 'dangers of all kinds' must have been reinterpreted as *bər-bahaya* 'dangerous'. Portuguese *marinho* 'maritime' generally became *mərinju* 'land bailiff, municipal officer' but comes out as *bərinju* (same meaning) in Baba Malay. However, there are also examples showing a tendency towards the formation of trisyllables with *mər/-*, for example Sanskrit *bharyapati* 'dove, pigeon', which became *mərpəti*, or Portuguese *manteiga* 'butter', which became *mərtega*/*məntega*.

The several possibilities of variation do not always fall into clear-cut formal categories, as is also demonstrated in a case like *rəmuđu* and *bərudu* 'tadpole'. By and large, however, they indicate that initial antepenultimate *mər/-* sequences are phonologically unstable.

Another phenomenon in Malayic word structure is that original intervocalic consonant clusters consisting of a nasal and a following liquid have often become syllabic through insertion of an epenthetic voiced stop followed by ə between the constituent consonants. The voiced stop is homorganic to the nasal it follows, as is shown in the following examples, which are almost all loanwords:

Proto Malayic <i>*timrah</i> 'tin foil'	Minangkabau (<i>*timbərah</i> >) <i>timbarah</i> 'id.' (Adelaar 1992a:110-1)
Arabic <i>zamrud</i> 'emerald'	<i>jamrud</i> , <i>jəmərut</i> , (Echols and Shadily 1989) 'id.' <i>jəmbrud</i> , 'id.'
English <i>general</i>	<i>jəndəral</i> , <i>jənəral</i> 'id.'
Non-standard Portuguese <i>ingrês</i> 'English'	<i>Inggəris</i> ³ 'id.'
Arabic <i>jumlah</i> 'sum, whole, body; crowd'	Malay <i>jumlah</i> , Jakarta Malay <i>jumbəlè</i> (Abdul Chaer 1976) Banjar Malay (Kuala subdialect) <i>jumbələh</i> 'number, amount' (Abdul Jebar 1977)

This syllabification is more frequent in spoken language and dialect forms than in written Malay. The following cases show the emergence of an epenthetic *b*, although they do not give clear evidence of an expansion of the number of syllables (which may be due to their spelling, or to the syllabic structure of the lending form):

³ In this instance the epenthetic [g] may already have been inserted into the Portuguese lending form.

<i>cəmərut</i> 'sullen, glum, morose'	<i>cəmbərut</i> 'id.' (Echols and Shadily 1989)
<i>kimlo</i> 'k.o. soup' (< Chinese)	<i>kimblo</i> 'id.' (Echols and Shadily 1989)
Dutch <i>kameraad</i> 'comrade'	<i>kambrat</i> 'id.' (Echols and Shadily 1989)

Occasionally, this avoidance of a nasal in too close proximity to a liquid can even be observed in final syllables. Observe the emergence of an epenthetic *b* after the nasal in the following loanwords:

Dutch <i>emmər</i> 'bucket'	Indonesian Malay <i>ember</i> 'id.'
Dutch <i>nummər</i> 'number'	Indonesian Malay <i>nomor</i> ,
	(substandard) <i>nombor</i> , <i>nombər</i> 'id.'

The tendency for antepenultimate **mər/-* syllables to become *bər/-* (whether as a prefix or a syllable of the root) is not unrelated to the tendency for original intervocalic nasal + liquid clusters to become syllabic and acquire a homorganic voiced stop. Both cases involve an environment lacking the presence of a full-length vowel (a vowel other than ə), which is required to prevent a nasal from being too close to a following liquid. In spoken Malay, *bər-* even tends to lose its vowel altogether if it is prefixed to an initial vowel, for example *bər-isi* → [brisi] 'to contain'. Wherever a nasal and a following liquid are in too close proximity, this 'unfortunate' phonotactic constellation tends to be solved by the insertion of an epenthetic voiced stop homorganic to the nasal. This is what happened in original intervocalic nasal + liquid clusters. It is also what happened in unstressed syllables such as the antepenultimate syllable **mar/-* and the prefix **mar-*, in which the vowel was reduced to ə.⁴ In the latter case, however, the resulting initial nasal + voiced stop cluster **mb-* was subsequently reduced to *b* only because such clusters are not permitted word-initially in Malay phonotactics.

3.4 No denasalization has taken place

Van den Berg proposes 'denasalization' as a sound change that occurred in only two cases, both prefixes. As I showed above, in my analysis there is no ground for such a sound change in Malay.

As I demonstrated in the previous section, the phonological history of Malay provides a straightforward explanation for the development from **mar-* to *bər-*. On the face of it, the development seems to involve denasalization, because the prefix originally had **m-* and has ended up with *b-*. But it would be misleading to qualify it that way because it ignores the fact that it was effectuated by a chain of intermediate changes, none of which includes denasalization.

⁴ See Blust 1995 and Hajek and Goedemans 2003 for the correlation between the shortness of schwa and lack of stress.

With the elimination of **mar-* to *bər-* as the only other instance of denasalization, and in the light of the arguments and alternative explanations that I propose in Section 5.1.2 and 5.7, it becomes increasingly clear that *di-* has never undergone denasalization either.

3.5 In the Old Malay inscriptions, *di-* and *ni-* are not in free variation

According to Van den Berg (2004:543), Old Malay *di-* and *ni-* were in free variation:

Even the Gandasuli inscription from AD 832 contains a form *di-rakṣa* ‘is protected’ (De Casparis 1950:61-5). This may point to free variation in Old Malay, which makes it even more likely that there is continuity between Old Malay and Standard Malay for this particular form and meaning.

However, there is no free variation of *di-* and *ni-* in the Gandasuli inscription. While this text does have one instance of *di-*, it lacks any instances of *ni-*. In this respect it agrees with the Manila inscription, which also has *di-*, but no *ni-*. The South Sumatran inscriptions, on the other hand, only have *ni-*. This allows for the establishment of two historical dialects: seventh-century Old Malay as found in the inscriptions of South Sumatra and Bangka, which are characterized by the use of the passive prefix *ni-*, and Old Malay from the ninth century and later, which includes the Gandasuli and Manila inscriptions and is characterized by the use of *di-*. The idea that Old Malay may have had two dialects is by no means new (see Mahdi 2005).

3.6 The evidence from South Sulawesi languages

In Section 5.2.4, bringing in evidence from South Sulawesi languages, Van den Berg argues that denasalization of **ni-* to *di-* is not as unusual as it seems.⁵ Some of these languages also have a passive marker *di-*. Van den Berg does not agree that this must be due to Malay influence, especially not as it is found in various South Sulawesi languages which are spoken in the mountainous interior of South Sulawesi (such as Tae’ and Duri) and have remained relatively isolated from Malay influence. On the other hand, so he argues, coastal languages have *ri-* (compare Bugis and allegedly Konjo) or *ni-* (compare Makassar), and it is in fact these languages (especially the cosmopolitan Bugis and Makassar languages) that tend to have been exposed to influence from Malay.

Van den Berg’s account of the distribution of these passive markers also, I contend, allows for a different reading. To put the record straight, Konjo does not have *ri-*,⁶ and the Bugis passive marker *ri-* is little else than a historical

⁵ Note that Proto South Sulawesi also had a voice marker **ni-* reflecting the Proto Austronesian undergoer voice marker **ni-*/*<*in*>.

⁶ Coastal Konjo (which is the dialect that is usually referred to as ‘Konjo’ in the literature) has a passive marker *ni-* instead (Friberg 1996:157).

di- that has acquired its present form as the result of a general sound change ($*d > r$) in the history of Bugis. This alters the South Sulawesi 'landscape' to an important extent, as Bugis has had a very strong influence on most other minor South Sulawesi languages, whether coastal or inland. It is quite possible that Bugis borrowed **di-* from Malay at an early stage, and that this voice marker was subsequently borrowed from Bugis into various other South Sulawesi languages, before it finally changed into *ri-* in Bugis itself. This would still be a case of borrowing from Malay into South Sulawesi interior languages, be it one in stages.

It is also possible that South Sulawesi *di-* (Bugis *ri-*) originated in the same way as Malay *di-*, that is, that it developed from the preposition *di* (Bugis *ri*), which (at least in Bugis and Makassar) is a general preposition used for locative, source, goal and (in passive clauses) agent. Both scenarios are more plausible than one that involves denasalization, for the simple reason that, as is the case with the Malayic subgroup, such a sound change has nowhere else taken place in the phonological history of the South Sulawesi subgroup, neither in the lexicon nor in affixes (Mills 1975).

3.7 Is it typologically unlikely that *di-* has developed from an agent preposition?

I have claimed that the evidence provided by several varieties of Malayic (and particularly by Kendayan varieties) strongly suggests that Malayic *di-* is the cliticized form of an earlier proposition *di*, the function of which at some point in time also included that of agent marker. Van den Berg rejects such a development on the grounds of typological improbability, pointing out that recent typological surveys by Haspelmath (1990) and Heine and Kuteva (2003) do not include it as a possible grammaticalization pathway.

However, recent support for my claim has emerged from several other languages that have undergone similar developments.

The Belangin language is a variety of Malayic spoken in the Landak regency of West Kalimantan. It seems to be closely related to Kendayan, although it differs from it in some important ways. It also uses a passive marker and an agent preposition that are historically related. Interestingly they are not related to Malay *di/di-* but are derived from Proto Malayo-Polynesian **uliq* 'to go back', and are related to Malay *oleh*, a verbal root which means 'to get, obtain' and which also functions as an agent-marking preposition. The passive marker is *lih*, a proclitic which occurs preverbally, and the agent preposition is *ulih* or *lih*. The following example (from my personal field data) exhibits both of them:

⁷ Akamine (2005) uses the qualification 'minimal' instead of 'singular', but in the case of Sinama *ku* the distinction is irrelevant.

(4) <i>Ia</i>	<i>lih</i>	<i>munuh</i>	[N-bunuh]	(u)lih	<i>munsuh</i>
3s	passive.marker		N-kill	Agent	enemy
'She was killed by the enemy'					

Ross (forthcoming) draws attention to the fact that the development from a preposition to a voice marker has also happened in Bajau. The following information on Sinama is taken from Akamine (2005:391-2). Sinama is a variant of Sama Bajau spoken in the Sulu Archipelago, Tawitawi Province, southern Philippines. It has a complex voice and mode system that expresses the major verbal alternations. It also has a preposition *leq*, which takes genitive pronouns and marks reason or (in undergoer constructions) agent. (Incidentally, this preposition also seems to reflect Proto Malayo-Polynesian **uliq*.) Alongside its regular voice and mode system, Sinama also has a 'derived-transitive' construction, in which *leq* is procliticized to an actor-voice verb that is marked by the (actor) prefix *N-*. The construction is characterized by its perfective aspect, its realis mode, and by the fact that its grammatical subject is totally affected. In such constructions *leq* may occur twice, as a verbal prefix and as an agent preposition. Akamine (2005:391) gives the following example:

(5) <i>leq-ngadjal</i>	<i>leq</i>	<i>ku</i>	<i>manuk</i>
Leq-N:cook	Agent	1s.GEN ⁷	chicken
'I have cooked the chicken'			

Akamine explains that, although the verb *ngadjal* (< *N-qadjal*) appears to be in actor voice, the grammatical subject (*manuk*) is not the actor: the actor is *ku*, marked by *leq*. He also shows that the actor can be topicalized, but in this case *leq-* is not prefixed to the verb:

(6) <i>leq</i>	<i>ku</i>	<i>ngadjal</i>	<i>manuk</i>
Agent	1s.GEN	N:cook	chicken
'I have cooked the chicken'			

It is noteworthy that in Sinama as well as in Kendayan, the agent marker and the voice marker never co-occur preverbally.

In Kendayan, the verb cannot be marked by *di-* if it is preceded by an agent (whether the latter is marked by *di* or not). As already pointed out in example sentence (3) (repeated below), the following construction is ungrammatical:

- (3) **Binie* *dah* *mati* *di* *kayo* *dimunuh*
 bini-e dah mati di kayo di-N-bunuh
 wife-3.POS PERF dead PREP enemy PASS-N-kill
 ‘His wife was killed by the enemy’

In Sinama, the agent marked by *leq* may precede or follow the verb. However, if it precedes the verb, the latter does not obtain the voice marker *leq-*.

- (7) **leq* *ku* *leq-ngadjal* *manuk*
 Agent 1s.genitive Leq-N:cook chicken
 ‘I have cooked the chicken’

So, it appears that neither Sinama nor Kendayan allow their homonymous agent preposition and voice marker to co-occur in preverbal position. The obvious explanation for this is that these homonymous grammatical elements are historically related and have not yet become independent enough to co-occur.

Incidentally, it appears that Van den Berg (2004:548) has misinterpreted my views on the further history of Malay *ni-* and *di-* when he writes that my argument is not only typologically implausible, but ‘also leads to a rather bizarre diachronic scenario’:

If I understand Adelaar correctly, Proto-Malayic probably had a passive marker **ni-* (tentative reconstruction) which in Old Malay also developed a variant *di-*. As this prefix shows no trace in present-day isolects, it must have disappeared from the language. At the same time or some time later, a morpheme appears which in present-day isolects has the same form as one of the earlier variants (*di-*) and essentially the same meaning, but allegedly derives from a completely different source (a preposition). This scenario is hard to believe.

Such a scenario would indeed be hard to believe, but then again, I do not recognize it as my own. It was with a sense of relief that, upon rereading my own work, I discovered that I had never written anything of the sort. Here Van den Berg may have let some of his own conclusions slip into his reading of my perspective. At any rate, in my work, *ni-* and *di-* have always been considered as historically unrelated, and I have never suggested that Proto Malayic developed a variant *di-* which was subsequently lost in present-day forms of Malayic.

4. Conclusion

In the preceding pages I have addressed the various arguments that Van den Berg adduces in order to show that *di-* developed from an earlier *ni-* through denasalization. On the basis of the following conclusions I remain convinced that *di-* was originally a preposition *di*.

In contrast to Van den Berg, I do believe that as a matter of principle we have to assume that prefixes undergo the same sound changes as the lexicon. There are exceptions, but these have to be accounted for by complementary regular sound changes or by rules of analogy.

It appears that antepenultimate vowel neutralization also applies to prefixes with historical **i*, such as reciprocal *sə-*. The fact that it does not apply to *di-* is due to the fact that this is not a fully grammaticalized prefix but has remained a proclitic.

No denasalization has taken place in the histories of *bər-* and *di-*. *bər-* can be explained as a development from **mar-* through a series of regular sound changes, but these do not include denasalization. *di-* developed from a preposition *di*.

In the Old Malay inscriptions, *ni-* and *di-* are not in free variation but on the contrary are markers of distinct historical dialects of Old Malay.

The fact that some South Sulawesi languages also have a passive marker *di-* is not very significant. It may have been borrowed from Malay and via Bugis into other South Sulawesi languages, or it may have developed from a preposition *di* (> Bugis *ri*). A development from South Sulawesi **ni-* through denasalization is unlikely because such a sound change is unprecedented in these languages.

The historical development of a voice marker from a preposition may be typologically unusual, but Kendayan is not alone in showing it. It is also observed in several other Indonesian languages. These languages do not use the same signifier as Kendayan for this historical double function, and at least in one of them (Sinama) the development cannot be attributed to contact but must be an independent development.

My conclusion that *di-* originated from the preposition *di* may seem unlikely from a typological perspective. But that is just as well: imagine what linguistics would be like if all its developments could be predicted from a few authoritative inventories of linguistic typology. With several more languages showing the development of an agent preposition into a voice marker, it may be a good time to readjust our views on the pathways along which passive markers come into being.

Typology clearly has a role to play in the study of the linguistic history of a particular language, but the success of this study is equally dependent on the exhaustive use of historical data drawn from various other disciplines

including epigraphy, philology, dialect variety, sociolinguistics, language contact, and history and prehistory. The spread of *di-*, for instance, can only be understood if one considers that throughout the centuries Malay has played the roles of court language, literary standard, and national language vis-à-vis other varieties of Malayic, and that it has also played these roles in various originally non-Malayic-speaking areas of Indonesia.

References

- Abdul Chaer Mad'ie
1976 *Kamus dialek Melayu Jakarta - bahasa Indonesia*. Jakarta: Nusa Indah.
- Abdul Jebar Hapip
1977 *Kamus Banjar - Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Adelaar, K. Alexander
1992a *Proto-Malayic: a reconstruction of its phonology and part of its morphology and lexicon*, Canberra: Pacific Linguistics.
1992b 'The relevance of Salako for Proto-Malayic and for Old Malay epigraphy', *Bijdragen tot de Taal-, Land- en Volkenkunde* 148:381-408.
2002 'Salako morphology and the interrelation between voice, mood and aspect', in: K. Alexander Adelaar and Robert Blust (eds), *Between worlds; Linguistic papers in memory of David John Prentice*, pp. 1-27. Canberra: Pacific Linguistics.
2005 *Salako or Badameà; Sketch grammar, texts and lexicon of a Kanayatn dialect (West Kalimantan)*. Wiesbaden: Harrassowitz.
forthcoming 'On the classifiability of Malayic', in: David Gil (ed.), *Malay and Indonesian linguistics*. London: Routledge.
- Aichele, Walther
1942-43 'Die altmalaiische Literatursprache und ihr Einfluß auf das Altjavanische', *Zeitschrift für Eingeborenen-Sprachen* 33:37-66.
- Akamine, Jun
2005 'Sama (Bajau)', in: Alexander Adelaar and Nicolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, pp. 377-96. London: Routledge. [Routledge Language Family Descriptions.]
- Berg, René van den
2004 'Some notes on the origin of Malay *di-*', *Bijdragen tot de Taal-, Land- en Volkenkunde* 160:532-54.
- Blust, Robert A.
1995 'Notes on Berawan consonant gemination', *Oceanic Linguistics* 34:123-38.
- Bynon, Theodora
1977 *Historical linguistics*. Cambridge: Cambridge University Press. [Cambridge Textbooks in Linguistics.]

- Casparis, J.G. de
1950 *Inscripties uit de Çailendra-tijd*. Bandung: Nix. [Prasasti Indonesia 1.]
- Echols, John M. and Hassan Shadily
1989 *An Indonesian–English dictionary*. (3rd edition, revised and edited by John U. Wolff and James T. Collins, in cooperation with Hassan Shadily). Ithaca NY: Cornell University Press.
- Friberg, Barbara
1996 'Konjo's peripatetic person markers', in: H. Steinhauer (ed.), *Papers in Austronesian Linguistics No.3*, pp. 137-71. Canberra: Pacific Linguistics.
- Hajek, John and Rob Goedemans
2003 'Word-initial geminates and stress in Pattani Malay', *The Linguistic Review* 20:79-94.
- Haspelmath, Martin
1990 'The grammaticalization of passive morphology', *Studies in Language* 14:25-72.
- Heine, Bernd and Tania Kuteva
2003 *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
- Kamus Dewan*
1994 *Kamus Dewan*. Kuala Lumpur: Dewan Bahasa dan Pustaka, Kementerian Pendidikan.
- Mahdi, Waruno
2005 'Old Malay', in: Alexander Adelaar and Nicolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, pp. 182-201. London: Routledge. [Routledge Language Family Descriptions.]
- Mills, Roger Frederick
1975 *Proto South Sulawesi and Proto Austronesian phonology*. [PhD thesis, University of Michigan, Ann Arbor.]
- Ross, M.D.
forthcoming 'Notes on the prehistory and subgrouping of Malayic', in: J. Bowden and N.P. Himmelmann (eds), *Papers in Austronesian linguistics*. Canberra: Pacific Linguistics.
- Teeuw, A.
1959 'The history of the Malay language; A preliminary survey', *Bijdragen tot de Taal-, Land- en Volkenkunde* 115:138-56
- Vikør, Lars S.
1988 *Perfecting spelling: Spelling discussions and reforms in Indonesia and Malaysia, 1900-1972*, with an appendix on Old Malay spelling and phonology. Dordrecht/Providence: Foris. [KITLV, Verhandelingen 133.]
- Wilkinson, R.J.
1959 *A Malay–English dictionary (Romanised)*. London: Macmillan.
- Wolff, John U.
2001 'The origin of the Malay affixation *di-nya*'. Paper, Fifth International Symposium on Malay/Indonesian Linguistics, Leipzig, 16-17 June.