This paper will explore the idiosyncratic voice system found in both standard Malay and Indonesian and explain it in a novel way using pseudo-incorporation (Massam 2001) and cliticization; it will also test analyses against ditransitive data, which has not been considered in previous accounts. The language exhibits an active clause (1), as well as two passive-like constructions, which I will call “passive” (2) and “object voice” (3), following Cole et al. (2008):

(1) Active: kami tidak akan mem-baca buku ini  
1PL not will meN-read book this  
“We will not read this book.”
(2) Passive: buku ini tidak akan di-baca (oleh Siti)  
book this not will Pass-read by Siti  
“This book will not be read by Siti.
(3) Object voice: buku ini tidak akan kami baca  
book this not will 1PL read.  
“This book will not be read by us.” (adapted from Cole et al. 2008)

The active is problematic as the marker on the verb, ‘meN-’, is optional. While the passive can be derived in the usual way, the object voice is unique in that the preverbal agent is restricted to pronominal forms, a property that has been largely ignored in previous accounts, which all solely use properties of v to explain the voice alternations.

Both Cole et al. (2008) and Aldridge’s (2008) primary research goals are to account for A’-extraction in the language, which is limited to the surface subject, i.e. the agent in the active and the patient in other two constructions (see Chung 1976 for extensive evidence); while other arguments cannot extract, extraction of adjuncts is possible. Cole et al. (2008) adopt an agreement approach, arguing that v agrees with either the agent (‘meN-’) or the patient (a null affix) in their lexically marked Case features. Other arguments, whose Case clashes with v, cannot raise, while Caseless adjuncts (adverbs and PPs), can freely do so. Functional heads usually agree not in Case but other phi-features such as person, number, and gender; an account without a language-specific form of agreement would be preferable.

For Aldridge (2008), Indonesian exhibits an unstable mixed voice system, in transition between the ergative system found in other Austronesian languages (object voice) a nominative-accusative alignment (the active and passive); extraction asymmetries are accounted for by claiming that the active is derived from an old antipassive and that it retains some of its idiosyncratic properties, though the analysis offers no diachronic evidence for this. Additionally, neither analysis has a satisfying account for the optionality of ‘meN-’.

All analyses limited to a functional voice projection also have no explanation as to why the agent must be pronominal in the object voice. Here it will be argued that the object voice is generated by pseudo-incorporation of the agent (Massam 2001, Levin 2014); that it cannot be extracted in the object voice follows trivially from this new analysis. Following Fortin (2006), I will treat the active “prefix” ‘meN-’ as an object clitic, with the lexical patient dislocated: it is for this reason that it cannot extract. While this analysis makes some interesting predictions that interact nicely with Hopper and Thompson’s (1980) transitivity diagnostics, I will present novel ditransitive data that still challenge for all analyses of meN-’, and discuss possible adaptations to Fortin (2006). The present analysis is superior in that these additional facts are explained without attributing ad hoc properties to the functional heads responsible for voice.
Sources