

Why You Can Drink a Book in Malay/Indonesian (but Not in English)

David Gil

Compare sentence (1) in colloquial (Jakarta or Riau) Indonesian with its English counterpart in (2):

(1) Badut minum buku
clown drink book

(2) The clown is drinking the book

Although similar, these sentences differ in two important ways:

- (3) *Formally*: While the Indonesian sentence consists just of three morphemes, denoting an activity and its participants, the English sentence contains several additional morphemes, denoting semantic categories such as number, definiteness, tense and aspect.
- (4) *Semantically*: Whereas the Indonesian sentence may be used to describe a situation in which a clown is drinking from a glass while reading a book, the English sentence is semantically anomalous, suggesting that *the book* is the patient of *drink*.

This paper argues that the formal; and semantic differences between the sentences summarized in (3) and (4) are systematically related to each other, and proposes an explanation for why this relationship should obtain.

In last year's ISMIL paper ("Can You Drink a Book in Malay/Indonesian?"), I presented the results of an ongoing cross-linguistic truth-condition experiment showing that speakers of Malay/Indonesian do indeed accept interpretations such as that described in (4), in which the semantic relationship between two sister constituents is unspecified with respect to thematic roles. Such interpretations were characterized as *associational*, resulting from the application of the association operator, in the absence of any more specific semantic rules making reference to grammatical features such as case marking and linear order. In the conclusion to that paper, I presented preliminary evidence suggesting that the availability of associational interpretations in a given language stands in inverse correlation to its *specificity index*, a numerical measure showing the extent to which a language partitions semantic space into finer units, by the (optional or obligatory) expression of various semantic categories such as number, definiteness, tense and aspect: the higher the specificity index, the lower the availability of associational interpretations.

This year's paper picks up from where last year's left off. In the first part of the paper, additional cross-linguistic evidence from a total of 16 different languages (creoles, West African and Southeast Asian) is presented showing that the inverse correlation between the specificity index and the availability of associational interpretations is indeed empirically robust.

However, this correlation is in fact rather surprising, given that most of the elements that contribute to a high specificity index have nothing to do with thematic roles. For

example, in English sentence (2), it is not obvious how the presence of definite articles, singular nominal number, and present progressive marking could be responsible for the unavailability of the associational interpretation available in Indonesian sentence (1). The second part of this paper proposes a solution to this puzzle, in the form of a principled explanation for the inverse correlation between the specificity index and the availability of associational interpretations.

The explanation lies in the following two unidirectional implications both relating semantic properties to a single syntactic one, namely, the inventory of syntactic categories:

- (3) If a language has a highly differentiated inventory of syntactic categories, then it tends to have a high specificity index
- (4) If a language has a highly differentiated inventory of syntactic categories, then it tends to have a low availability of associational interpretations.

The basic idea behind both implicational relationships is that languages tend to make use of the syntactic categories that they have, otherwise there wouldn't be any point in having them. In particular, if a language has a noun-verb distinction, then this distinction is likely to be reflected by grammatical markings making reference to this distinction, such as nominal number and definiteness marking and verbal tense and aspect marking, thereby raising the specificity index of the language. Similarly, this distinction forms the basis for the marking of thematic roles, specifying the semantic relationships between a verb and its nominal arguments, thereby decreasing the availability of associational interpretations. Thus, different syntactic category inventories are what underlie the inverse correlation between the specificity index and the availability of associational interpretations. In particular, whereas the rich syntactic category inventory of English results in the low availability of associational interpretations, rendering sentence (2) semantically anomalous, the impoverished syntactic category inventory of colloquial Indonesian makes associational interpretations possible, which is why, as in sentence (1), you can drink a book in Malay/Indonesian.