

Can You Drink a Book in Malay/Indonesian?

Variation Across Isolect, Socio-Economic Status, and Age

David Gil

Riau Indonesian has been said to exhibit a number of syntactic and semantic features that characterize it as typologically exceptional, seemingly unlike most other natural languages, at least as commonly described. Semantically, it has been claimed that when two expressions X and Y with meanings P and Q respectively are combined, the meaning of the collocation X Y is derived from that of its constituent parts by means of the association operator, A (P, Q), which says that the meaning of X Y is associated in an unspecified way with the meanings of X and Y respectively. For example, if *ayam* means 'chicken' and *makan* means 'eat', *ayam makan* means A (CHICKEN, EAT), or anything that has to do in some way with 'chicken' and with 'eat'. In particular, the semantic representation A (CHICKEN, EAT) lacks any specification of thematic roles: the chicken could assume the role of agent, patient, or whatever might make sense in the context of the utterance.

This paper presents the results of an experiment designed to measure, objectively across a variety of languages, the availability of *apparently associational interpretations*: interpretations that appear to be obtainable from the association operator without reference to thematic roles or other semantic categories. Two kinds of apparently associational interpretations are sought: (a) those in which what looks like a bare noun preceding a bare verb is interpreted as the patient (rather than the agent); and (b) those in which what looks like a bare noun in construction with a bare verb is interpreted as an oblique argument or even a non-argument (in the absence of prepositions or other such markings) — as in the title of this paper. The experiment presents subjects with a sentence in the target language and two pictures; subjects are asked which of the two pictures is best described by the sentence. The experiment is ongoing; as of March 2005, over 800 subjects in 12 languages had been tested.

Whereas my ALT 6 paper in Padang presents some of the cross-linguistic results of the above experiment; the present paper focusses on Malay/Indonesian, and specifically on patterns of variation with respect to isoelect, socio-economic status and age. The major patterns of variation to emerge from the experimental data are summarized in the following three statements, representing the availability of apparently associational interpretations across different populations:

- (1) isoelect: Minangkabau > Jakarta Indonesian > Siak Malay > Riau Indonesian
> Kuala Lumpur Malay
- (2) SES: low socio-economic status > high socio-economic status
- (3) age: young children > older children and adults

This paper argues that the above three patterns may be accounted for in a unified fashion, in terms of the *Standard Language Interference Principle*, in accordance with which subjects, in their responses, exhibit interference from Standard Malay/Indonesian.

Given that Standard Malay/Indonesian is characterized by low availability of apparently associational interpretations, the Standard Language Interference Principle predicts that the greater the interference, the lower the availability of apparently associational interpretations. All three patterns in (1) – (3) bear out this prediction. The isoelect hierarchy in (1) provides an accurate characterization of how sociolinguistically distinct from Standard Malay/Indonesian each isoelect is and hence how likely it is to resist interference from the standard language: at one end, Minangkabau is for all

practical purposes an independent language, hence it is subject to less interference from Standard Indonesian, and therefore it has the highest availability of apparently associational interpretations; at the other end, Kuala Lumpur Malay is nothing more than the basilectal counterpart of Standard Malay, hence it experiences the greatest amount of interference from Standard Malay, and accordingly it has the lowest availability of apparently associational interpretations. The SES hierarchy in (2) reflects the state of affairs in which subjects of lower socio-economic status are less exposed to the standard language than are subjects of higher socio-economic status, and therefore more willing to accept apparently associational interpretations. Finally, the age hierarchy in (3) reflects the fact that young children first acquire a colloquial variety of Malay/Indonesian, while only later, in school and elsewhere, are they exposed to the standard language, which then affects their performance on the experiment by lowering their acceptance of apparently associational interpretations.

Data from naturalistic corpora suggests that all Malay/Indonesian speakers, of whatever islect, socio-economic status and age, make use of utterances involving apparently associational interpretations; in fact, there are cases where the same person produces such a construction in the naturalistic corpus and then later, as an experimental subject, rejects a similar construction. Thus, it would appear that the patterns of variation in (1) – (3) are largely and perhaps even entirely an artefact of the experimental method and concomitant interference from the standard language. To the extent that this is the case, it means that all of the isolects of Malay/Indonesian under consideration here are characterized by a uniformly high degree of availability of apparently associational interpretations. However, different populations still vary greatly with respect to how closely their performance on the experiment reflects their actual competence: of the groups tested so far, high-SES Kuala Lumpur Malay adults make the worst subjects, young low-SES Minangkabau children the best.