Describing Malayic:
When Incommensurability Meets the Language-Dialect Continuum
David Gil
Max Planck Institute for the Science of Human History

Linguistic theories differ with respect to whether they consider the grammars of different languages and dialects to be commensurate. Simplifying somewhat, the following three approaches to linguistic description and analysis are discernable:

(1) (a) Different languages and dialects are described as a single system with variation.
     (b) Different languages are described as different but commensurate systems; different dialects of the same language are described as a single system with variation.
     (c) Different languages are described as different but incommensurate systems; different dialects of the same language are described as a single system with variation.

Broadly speaking, approach (1a) is adopted by some of the more radical versions of generative grammar; approach (1b) is characteristic of traditional linguistics; while approach (1c) is associated with the practice of many contemporary functionalists and typologists.

Approaches (1b) and (1c) make crucial reference to the distinction between languages and dialects, positing different descriptive modes for variation across and within languages. In particular, approach (1c) is predicated on the assumption that while the grammars of different languages are incommensurate, those of different dialects of the same language are commensurate. In a series of recent publications, Haspelmath (2010, 2015, 2016) concludes that because of their incommensurability, the descriptive categories of individual languages cannot be used for cross-linguistic comparisons, and that instead, typologists require recourse to an ontologically completely different kind of entity, which he refers to as "comparative categories". However, we all know that in reality there is no clear-cut and objective distinction between languages and dialects; rather, there are lots of cases where two speech varieties straddle the boundary between different languages and different dialects of the same language. What, then, are the implications of the language-dialect continuum with regard to the approaches outlined in (1) above, and, more specifically, with respect to the notion of comparative category?

The Malayic language family is a hotbed of language-dialect indeterminacy, providing a valuable test case for the examination of these methodological issues. The indeterminacy is reflected in the mess presented by the commonly used language names. Two big language names, "Malay" and "Indonesian", are commonly used to refer to a partly overlapping set of speech varieties many of which exhibit substantial diversity, in some cases resembling that associated with distinct languages. On the other hand, in some regions, such as southern Sumatra, there is a proliferation of different language names mostly derived from toponyms which seem, in many cases, to refer to dialects of the same language. This situation poses myriad practical methodological problems for any

linguistic practitioner in the region, such as what to call the speech variety one is working on, whether to describe the speech variety on its own terms or with reference to other related varieties, and how to analyze instances of variation and contact involving multiple speech varieties.

This paper takes a close look at the methodological problems posed by the language-dialect continuum in Malayic, and their implications with regard to the three approaches outlined in (1) above. This paper argues instead for a fourth hybrid approach combining aspects of the three approaches in (1). The proposed hybrid approach dispenses with the distinction between languages and dialects, replacing it with the single and more general notion of languoid:

(2)   Some features of different languoids are described as a single system with variation; other features of different languoids are described as different but commensurate systems; yet other feature of different languoids are described as different but incommensurate systems.

Although the approach outlined in (2) above makes no reference to languoid distance, it is likely that upon implementation, a correlation will emerge in which commensurability will be more commonly associated with closer languoids such as dialects of a single language, while incommensurability will be more often characteristic of more distant languoids, such as in typological studies of world-wide scope. Nevertheless, the results of this paper and the hybrid approach in (2) point towards a refinement of the distinction between language-specific descriptive categories and comparative concepts as proposed by Haspelmath, allowing for the possibility that descriptive categories observable in particular languages may also form the basis for cross-linguistic comparisons.

Haspelmath, Martin (2010) "Comparative Concepts and Descriptive Categories in Cross-Linguistic Studies", *Language* 86:663–687.
Haspelmath, Martin (2015) "Defining vs. Diagnosing Linguistic Categories: A Case study of Clitic Phenomena", in J. Blaszczak, D. Klimek-Jankowska and K. Migdalski eds., *How Categorical Are Categories? New Approaches to the Old Questions of Noun, Verb and Adjective*, De Gruyter Mouton, Berlin and New York.
Haspelmath, Martin (2016) "The Serial Verb Construction: Comparative Concept and Cross-Linguistic Generalizations", *Language and Linguistics* 17.