

# LANGUAGE, DIALECT AND RIOTOUS SOUND CHANGE: THE CASE OF SA'BAN

Robert Blust

University of Hawai'i

## ABSTRACT

Sa'ban (also known as Saban), an obscure and poorly described member of the North Sarawak group of languages, can be shown on lexicostatistical grounds to be a dialect of Kelabit-Lun Dayeh. However, it has undergone extraordinarily rapid and extensive sound changes which have destroyed intelligibility with any of its better-known sister dialects/languages. These changes are surprising for several reasons. First, contrary to universal tendencies, they exhibit a strong pattern of 'erosion from the left' which appears to be at least partly independent of prosody. This canonical feature, as well as the unusual phonetic content of some changes which must have taken place over a very short time, forces us to confront the question whether all sound changes are phonetically motivated. Second, despite many recurrent correspondences with other dialects of Kelabit-Lun Dayeh, these changes exhibit massive irregularity. Finally, sound change has triggered a major restructuring of verb morphology in relation to voice marking. This restructuring has produced not only new patterns of affixation, but also many new problems in relating the members of morphological paradigms to an underlying base of constant shape, and so almost certainly have complicated the task of first language acquisition.

**1. Language and dialect.** The language/dialect distinction has vexed linguists for generations. Under what conditions can we say that two speech communities are dialects of a single language, and when can we say they are

two different languages? Answers to this question have appealed to essentially two types of evidence: 1) intelligibility, and 2) cognate percentage in basic vocabulary. Both types of evidence are problematic, and as will be seen, there is no guarantee that they will agree.

Intelligibility offers a commonsense approach to the language/dialect question: if the speech of two communities is mutually comprehensible the two should be considered dialects of a single language; if not, they should be considered different languages. In many ways this approach is appealing, but in practice it is fraught with complications. First, intelligibility may be non-mutual. The fact that speakers in community A profess to understand speakers in community B is no guarantee that the experience will be reciprocated. Second, intelligibility may be delayed. Americans arriving in Australia to live often have difficulty understanding the local dialect for a period of weeks or months, particularly when dealing with certain semantic domains (hardware, automobiles, etc.). But the differences seem to 'wear off' with exposure, and full intelligibility generally is achieved with no special effort on the part of the new arrival. Finally, in chaining or network situations there may be no meaningful or non-arbitrary way to distinguish dialects from languages. If a chain of dialects ABCDE exists such that mutual intelligibility decreases with distance, A and E may be mutually unintelligible, but so long as intermediate dialect gradations exist it is impossible to draw a language boundary anywhere within the chain. However, if dialects B, C and D should become extinct the result would be two languages. Somewhat ironically, in such a situation dialect death can be said to result in language birth.

Another criterion that has been used to define the language/dialect distinction is percentage of cognate basic vocabulary. Dyen (1965) used the expression 'language limit' to mark the boundary between two communities which speak

dialects of a single language vs. two communities which speak different languages. He suggested that the language limit be set at 70% cognation, a figure which reportedly shows a high degree of correspondence with the limits of mutual intelligibility. Wurm (1971:552), on the other hand, as well as other linguists working in Australia, have suggested 81% as corresponding closely with the limits of mutual intelligibility.

The belief that a lexicostatistically-defined language limit will tend to correlate closely with the limits of mutual intelligibility is based on the tacit assumption that rates of linguistic change are roughly homogeneous throughout a language: if lexical change is rapid it is unlikely that phonological or grammatical change will be slow, and vice versa. Historical linguists have not made much progress in the study of rates of linguistic change, and the assumption of homogeneity in all components of a language is an article of faith more than a well-established product of carefully conducted research.

With upwards of 1,000 languages that represent the linguistic residue of a generally rapid expansion out of Taiwan over island Southeast Asia and the Pacific during the past four or five millenia, the Austronesian language family offers a vast laboratory for the study of language change. Impressionistically it is not difficult to think of languages in which some components appear to be conservative, but others innovative. Atayal of northern Taiwan is quite conservative in preserving the distinctive system of multiple voice marking known in the Austronesian literature as 'focus', but it is lexically and phonologically highly innovative. Trukese of the eastern Caroline islands of Micronesia, on the other hand — like other members of the large Oceanic subgroup — is grammatically quite innovative, and phonologically extremely innovative, although it is lexically rather conservative. Sa'ban,

spoken in the border region of northern Sarawak and Kalimantan, exemplifies a case of extremely rapid sound change which has run far ahead of lexical replacement and given rise to extensive morphosyntactic complications that surely cannot have facilitated the task of first language acquisition. As a general consequence of these changes Sa'ban must be classified on lexical grounds as a dialect of Kelabit, but on grounds of morpheme structure, phonology, morphology or intelligibility it is clearly a distinct language. The principal aim of this paper is to raise these statements from the level of general impressions to that of well-supported claims.

**2. The linguistic position of Sa'ban.** The Austronesian (AN) family as a whole divides into perhaps ten primary branches. Nine of these branches are represented among the 15 surviving languages in Taiwan (Blust 1999). All of the remaining 1,000-1,200 AN languages outside Taiwan belong to the Malayo-Polynesian (MP) subgroup, which in turn divides into Central-Eastern Malayo-Polynesian (the AN languages of eastern Indonesia and the Pacific exclusive of Palauan and Chamorro), and Western Malayo-Polynesian (WMP), a large collection of MP languages in the Philippines and western Indonesia. Although WMP may turn out to be an innovation-defined subgroup, it is best regarded for the present as a collection of heterogeneous MP languages linked by certain exclusively shared features of verb morphology. The most notable of these shared features is the process of homorganic nasal substitution used to form active verbs from unaffixed word bases, as in Malay /pukul/ 'hit' : /me-mukul/ 'to hit' : /di-pukul/ 'be hit'. Nasal substitution and its complementary process nasal accretion, takes various forms in individual languages. Presumably under the influence of a family-wide disyllabic canonical target, a number of languages in Indonesia have come to form active verbs by nasal substitution alone, without an accompanying