

International Workshop
Current Issues In Research On Languages In Borneo And Malay Language,
7 September, UMS

**REVIEW OF *ETHNOLOGUE* DESCRIPTIONS OF
LANGUAGES IN SABAH—PRELIMINARY FINDINGS**

Dr. Mark Miller, SIL Malaysia

Dr. Jeannet Stephen, PPIB, UMS

Prof. Dr. Jacqueline Pugh-Kitingan,
Kadazandusun Chair, UMS

What is the *Ethnologue*?

- *Ethnologue: Languages of the World* is a catalogue of the all the world's known living languages (currently the count is 7,102)
- Widely regarded as the most comprehensive source of information of its kind, including:
 - region(s) of use within countries;
 - Alternate names and dialect names;
 - population of speakers and/or ethnic population;
 - genetic classification;
 - language use and viability;
 - writing script(s) used;
 - available literature and other products of language development.

History of the *Ethnologue*

- Began in 1951 by SIL linguist Richard S. Pittman (10 mimeographed pages, containing info on 46 languages or groups of languages).
- Dr. Pittman continued to expand his research through the seventh edition (1959), which listed 4,493 languages
- Under Barbara F. Grimes as research editor from 1967 to 2000, the language count expanded from 4,493 to 6,809 in size, and info about each language expanded too.
- With the 17th edition, the *Ethnologue* shifted to a web-centric paradigm, enabling annual updates. Over 12,000 updates have been made to the *Ethnologue* database since the 17th edition of the *Ethnologue* was released one yr ago.
- Much of the focus has been (and continues to be) on the less commonly known languages.

Some features of the *Ethnologue*

- Over 200 color language maps (for Malaysia, there are maps for Peninsular Malaysia, Sabah, and Sarawak - Brunei)
- Extensive bibliography of references cited
- Each language is identified by a unique 3-letter code from ISO 693-3, to ease the identification of the same language called by different names across country borders, e.g., [tau] for Tausug – Suluk, AND to distinguish languages called by the same name (e.g. ‘East Coast Bajau’)

ISO 639-3 language identifiers

- In 1998, the International Organisation for Standardization (ISO) adopted a standard for three-letter language identifiers: ISO 639-2. But ISO 639-2 had identifiers for less than 400 languages.
- The standard named ISO 639-3 was approved in 2006 and published in 2007. It provides unique 3-letter codes for over 7,500 languages (which includes ancient and constructed languages)
- SIL International was named as the registration authority for ISO 639-3 and administers the annual cycle for changes and updates as a function separate from the *Ethnologue*.

Criteria for deciding ‘language’ vs. ‘dialect’

Not all scholars agree on this. The ISO 693-3 standard applies the following basic criteria to deciding between ‘language’ and ‘dialect’:

- Is there “inherent understanding” of the other variety?
- Where spoken intelligibility between varieties is marginal, is there nevertheless a common literature or common ethnolinguistic identity with a central variety that both can understand?
- Where there is enough spoken intelligibility between varieties to enable communication, the identification of well-established distinct ethnolinguistic identities can still be a strong indicator of ‘separate’.

Previous Ethnologue Reviews

- There is a continued need for updating the information found in the *Ethnologue*. How do we best accomplish that? An *Ethnologue* review or “audit” entails a focused time frame for bringing scholars together to review existing entries and make needed changes.
- In recent years, SIL has conducted *Ethnologue* audits in Brazil (mostly an internal audit, involving SIL scholars) and the Philippines. The most comprehensive effort to date has been in Nepal (audit conducted in 2010-11, resulting in a publication in 2012 published by SIL International Nepal in cooperation with the Central Department of Linguistics, Tribhuvan University, Nepal).

History of Survey of Sabah Languages

- An initial survey in 1978-81 by SIL members of the languages and dialects spoken in Sabah resulted in a “tentative classification” of the languages and dialects of Sabah, published in *Languages of Sabah: A Survey Report* (King & King, eds., 1984) [based upon a lexicostatistical analysis of hundreds of word lists taken throughout the state, as well as Recorded Text Testing at selective locations]
- A survey of the coastal languages of Sabah was conducted by SIL Malaysia branch members in 1998-99, which focused only on Suluk, Iranun, Tidung, the Malayic language grouping, and the Sama-Bajau language grouping.
- Much of the information contained in the *Ethnologue* for languages of Sabah draws from the findings contained in these surveys, although some changes in classification (as well as plain errors) have been introduced over the years. Much of the data is simply outdated, and there has not been much reported on language use and viability.
- Most of the current info in the *Ethnologue* has been reported on the findings of SIL field linguists. There is a need for more input from local scholars and members of the language communities themselves. In summary, there is the great need for more collaboration between SIL, local academic and gov’t bodies, and the language communities to update and improve the existing *Ethnologue* entries.

Ethnologue audit for Sabah languages

- Originally SIL Malaysia intended to conduct the first audit with UNIMAS for languages of Sarawak. But SIL lost its director for language programs in Sarawak during that time. The decision was made to engage UMS as primary partner to conduct the first Malaysia review (SIL and UMS signed an MoU in 2012).
- John Epele, who is the SIL language assessment coordinator for Asia Area, made visits to UMS in 2014 (in Sept 2014, Feb & April 2015) to meet with key people at PPIB, including the Dean of PPIB (Datuk Dr. Vincent Pang) also Prof. Dr. Jacqueline Pugh-Kitingan as Holder of the Kadazandusun Chair. UMS showed strong interest. Plans began to come together in spring 2015.