

Paper presented at 13th Annual Internationalisation and Localisation Conference, 2-3 October 2008, Marino Institute of Education, Dublin, Ireland

Preservation of Sarawak Ethnic Languages: The Sarawak Language Technology (SaLT) Initiative

Alvin W. Yeo, Suhaila Saeed, Sarah Flora, Nurfauzia Jali, Suriati Khartini,
Edwin Mit, Ting Su Hie, Jennifer Wilfred, and Ng Giap Weng

Universiti Malaysia Sarawak
Kota Samarahan, 94300, Sarawak Malaysia
6082-583580
alvin [AT] fit.unimas.my

27

ABSTRACT

The population of speakers of indigenous languages all over the world is decreasing. This drop in numbers is due to the pressures of dominant global languages (such as English which is the lingua franca of international commerce, research, and the Internet), rural-urban migration, and exogamy (inter-ethnic group marriages). Similarly, the number of speakers of Sarawak's 63 languages is also declining. Thus, the Sarawak Language Technologies (SaLT) Research Group at Universiti Malaysia Sarawak has initiated a number of research and development projects with the end goal of revitalising and maintaining the ethnic languages of Sarawak. The ongoing projects include building corpora of languages (Iban, Melanau, and Kelabit), as well as, research and development of technologies which contribute to the implementation of software for the ethnic languages. Specifically, these projects include development of morphological analysers and Part of Speech (POS) taggers which adds to knowledge on Iban-English translation, and in human computer interaction using Melanau speech and text. Other projects in the pipeline include a wiki approach in building a Bidayuh lexicon, and a web-based Sarawak Malay language dictionary. Through the initiative and projects, a multidiscipline approach to preserve indigenous languages is being developed and refined.

Keywords

Ethnic languages, revitalisation and maintenance, and Sarawak Language Technologies, Information and Communication Technologies

1.0 Introduction



Figure 1.0: Map of Sarawak in Malaysia