Comparing orthography design in Barayin (Chad) and Kodi (Indonesia)

Joseph Lovestrand

University of Oxford England

Abstract— This research stems from the emergence of a problem which is the lack of knowledge of millennial children about folklore from their respective regions, especially children in Southeast Sulawesi. Preservation of folklore which is one part of traditional literature is very important because one of the functions of folklore is as children's character education. One of the folklores from Southeast Sulawesi that needs to be preserved is "Tula-Tula Metulungino Mie Kodosa folklore" originating from Muna. The purpose of this study was to determine the character education contained in Tula-Tula Metulungino Mie Kodosa folklore. This research is a literature study. The method used is descriptive qualitative method. The source of the research data comes from the book, Buton and Muna Folklore in Southeast Sulawesi, which was published by the Ministry of Education and Culture in the Language Development and Language Development Center in 1998 in Jakarta, 144 pages thick. The result showed that character education contained in the Tula-Tula Metulungino Mie Kodosa folklore, that is helpful, kind, patient, and never give up. This folklore is important for millenial generation in digital era because it educates them to have good characters and helpfull in preventing the folklore extinction.

Keywords— Character education, Muna Folklore, Tula-Tula Metulungino Mie Kodosa

1 Introduction

It is estimated that half of the world's 7000 languages, including half of Indonesia's 700 languages, are on the path to extinction [1]. With the loss of these languages, the unique cultural heritage expressed in each language is also likely to disappear, as well as any traditional knowledge its speakers hold.

One key tool for strengthening endangered languages and promoting their continued use is literacy. Literacy increases the status of a language, so that it is less likely to be considered inferior to regional languages that have a standard writing system. Literacy can provide more domains for use of a language. Literacy, or at least the development of a writing system, is a prerequisite to developing mother tongue-based multilingual educational opportunities for children who do not speak the standard language of instruction at home. Mother tongue-based multilingual education has been shown to have the potential for significantly improving the overall levels of educations in minority groups, including significant gains in using the official

language [2] [3] It is possible for a language to go extinct even after having developed a robust literacy program. However, even in this case, the existence of literacy facilitates the documentation of the language before its death, so that at least a partial record is available for research and as an archive for heritage speakers of the language.

Linguists play a crucial role in the development of a new writing system. In some cases, the process might be relatively straightforward, but the potential complications should not be underestimated. Developing orthographies for each of the thousands of languages that are exclusively used orally is a significant challenge because each orthography needs to be adapted to each individual language and to the society that will use it.

Part of the challenge of orthography design lies in the attempt to reduce a phonological system to written symbols. However, there is just as much, if not more challenge in navigating the social aspects of orthography design [4]. The non-linguistic aspects of orthography design receive less attention from linguists, despite their importance in supporting the preservation of endangered languages. This paper compares orthography design projects in two very different parts of the world in order to illustrate the types of complications that can arise, as well as to compare what partial solutions are found in each area. By learning about orthography design projects in different parts of the world, we might progress to more complete solutions for communities transitioning from being exclusively oral to being literate.

The first language considered, Barayin, is introduced in Section 2. Barayin is a Chadic (Afroasiatic) language spoken by about 5,000 people in Chad. The second language, Kodi, is introduced in Section 3. Kodi is a Malayo-Polynesian (Austronesian) language spoken by about 20,000 people on the island of Sumba in eastern Indonesia. In Section 4, the main linguistic issues related to each orthography system are presented. Section 5 explains how the national government policies had a role in the development of each orthography. Section 6 explains the role of local community-based organizations in the same process. Section 7 is a conclusion summarizing what can be learned from comparing these two programs.

2 Method

The Barayin language is spoken in Chad (Africa) by about 5,000 people [5]. However, there are four dialects, only two of which share mutual intelligibility [6]. These two dialects, with about 3,000 total speakers, are the only ones that have been the subject of linguistic analysis [7]. Barayin has five contrastive vowel sounds: /a, e, i, o, u/. Table 1 shows the sixteen phonemic consonants of Barayin by their place and manner of articulation. Barayin has contrastive vowel length and contrastive consonant length.

Table 1 Barayin consonants

	bilabial	alveolar	palatal	velar
voiceless stop	p	t		k
voiced stop	b	d	ɟ	g
fricative		s		
nasal	m	n	ɲ	ŋ
approximant	w		j	
lateral		l		
tap		r		

Barayin is a tonal language. The meaning of words can change based on their pitch. Several of the minimal pairs for contrastive tone are found in the pronominal system. Three examples of pairs of words that only differ in their tone are given in examples (1) through (3).

(1) [nìl:à]   LL   'them (3PL)'
     [níl:à]   HL   'you all (2PL)'

(2) [ŋ̀]   L   'I (1SG.SBJ)'
     [ŋ́]   H   'the two of us (1DU.INCL.SBJ)'

(3) [kà]   L   'he (3SG.M.SBJ)'
     [ká]   H   'also'

3 Result and Discussion

The Kodi language is spoken in Sumba, Indonesia by about 20,000 people [5]. Kodi has five contrastive vowel sounds: /a, e, i, o, u/ and 20 consonant sounds, as shown in Table 2 [8]. Kodi has contrastive consonant length and phonetic vowel lengthening in stressed syllables. Words in Kodi often vary in how many syllables they contain. There are three different word-shortening processes illustrated in the examples (4) through (6).

(4) Initial-vowel deletion: [mbaɾojo] ~ [mbrojo]   'far'
                               [pʰaneːɣeŋ] ~ [pʰneːɣeŋ]   'to speak'

(5) Final-vowel deletion: [ʔuɾa] ~ [ʔuɾ]   'rain'
                               [mboːnoho] ~ [mboːnoh]   'hot'

(6) Metathesis: [waːtʰu] ~ [waut]   'stone'
                               [maːtʰi] ~ [mait]   'to die'

Table 2 Kodi consonants

	bilabial	dental	palatal	velar	glottal
voiceless stop	pʰ	tʰ	c	kʰ	
implosive	ɓ	ɗ			
prenasalized stop	ᵐb	ⁿd	ⁿɟ	ⁿg	
fricative				ɣ	h
nasal	m	n	ɲ	ŋ	
approximant	w		j		
lateral		l			
tap		r			

The brief descriptions of aspects of Barayin and Kodi grammar in Sections 2 and 3 illustrate linguistic features that potentially complicate the writing systems. Since Barayin is a tonal language, simply using a segmental approach to representing the sounds of the language will lead to some ambiguities that confuse the reader. Something needs to be done to mark the tonal difference in these words. However, there are relatively few such ambiguous words in the language, so it is not necessary to mark tone on every word [9].

Two different strategies are used for marking contrasting tone in Barayin. As shown in example (7), if the contrasting tone is on a vowel, the word with high tone is marked with an acute accent. However, in the pair of words in example (8), the contrast is on a syllabic nasal. To avoid writing diacritics over a consonant, it was decided to mark the word with high tone by adding an unpronounced letter <i>.

(7) <nilla> [nìl:à] 'them (3PL)'
     <nílla> [níl:à] 'you all (2PL)'

(8) <ŋ>   [ŋ̀]   'I (1SG.SBJ)'
     <iŋ> [ŋ́]   'the two of us (1DU.INCL.SBJ)'

Kodi is not a tonal language, but it has a different kind of phonological challenge. In many contexts, words are pronounced with one less syllable. Kodi people disagree whether the language should be written the way it is pronounced (with fewer syllables), or if all syllables should always be written, even if they are not pronounced in some contexts.

In terms of "orthographic depth" an orthography that reflects the pronunciation is called a "shallow orthography", and an orthography that preserves the same spelling for the same word in all contexts is a called a "deep orthography" [10]. Currently there is no standard solution to whether Kodi should be written in a shallow or deep orthography. Each person spells the words as they wish. In order to make any recommendations on what system might work best for Kodi, there need to be studies of natural discourse in Kodi to understand how frequently each form is used and what contexts correlate with each form.

The government policies and practices of each country have an impact on how orthographies are developed. Barayin is spoken in the Republic of Chad. In 2009, the government of Chad set guidelines for which characters should be used to write Chadian languages based on the *Alphabet national du Tchad* shown in Table 3.

Table 3 Alphabet national du Tchad (lowercase Latin letters)

a	b	ɓ	c	ch	d	ɗ	e	ɛ	ə	f	g	gb	h	ɦ	i
ɨ	j	k	kh	kp	l	m	mb	mv	n	ṇ	nd	ng	nj	ŋ	o
ɔ	p	r	r̠	s	sl	t	u	v	vb	w	y	yʼ	z	zl	ʼ

The letters from the national alphabet are meant to be used when creating new orthographies for Chadian languages. This may seem restrictive, but it can also simplify the decision making process. For example, when designing the Barayin alphabet, the national alphabet proposes the letter <ṇ> for the palatal nasal sound /ɲ/. Other options, such as digraphs or other diacritics, do not need to be considered.

In contrast, in Kodi, there is still more than one opinion on how to write implosive sounds and no national or regional guidelines to follow. The two main ways that implosives are being written are either in a digraph such as <bh> for the bilabial implosive /ɓ/, or with a simple <b> [8]. The use of <b> is possible because Kodi does not have the regular voiced plosive sound that the plain letter <b> normally stands for. In other words, the same letter can be used with a different pronunciation from other languages. In other languages in Sumba, you can find implosives writte marked by an apostrophe or underline, or written with the phonetic symbol <ɓ>. There is no general standard for what letters should be used in alphabets of Indonesian languages. While the Indonesian government is not involved in specific guidelines for alphabets, they supported the Kodi orthography financially through a grant from the INOVASI project. This grant was facilitated by a national non-project organization (*yayasan*) called *Suluh*

Insan Lestari in 2018 and 2019. The project developed reading material for use in eight primary schools in the Kodi-speaking area, and trained government teachers to use the materials to help children learn to read and write in their mother tongue as a subject in the school curriculum. This kind of program is not available in Chad, and government schools do not generally provide support for the use of local languages.

The Barayin literacy project was initiated by a community-based organization, ADPLB (*Association pour le développement et la promotion de la langue baraïn*). This organization began when Barayin speakers saw neighboring languages reading and writing their languages and decided they would not be left behind. Their organization joined a regional literacy organization, FAPLG (*Fédération des associations pour la promotion des langues du Guéra*). These local organizations are able to raise grassroots support for mother tongue literacy, recruit volunteers, train teachers, and make decisions about how to write the language and what books to publish. In Sumba, there are no local organizations dedicated to mother tongue literacy, and no language committees that can make decisions about the Kodi language on behalf of the community. The mother tongue literacy project in Kodi is facilitated by a national organization based in Jakarta in partnership with a Sumbanese organization that is generally involved in development work, but not specifically focused on Kodi or on literacy. In other words, there is no organization that represents the Kodi language. This creates a challenge when there is a need for a decision, for example, how to write implosives (Section 5) or whether to use a shallow or deep orthography (Section 4). It also means that the literacy activities in Kodi will likely stop once the funding for the involvement of outside organizations comes to an end.

4 Conclusion

There are always challenges when developing a new orthography. The linguistic challenges (like deciding how to write tone or what kind of orthographic depth to use) require a partnership between researchers who can analyze the structure of the language, and well-informed community members who can represent the community's preferences for writing the language. Without a local organization that represents the language, this process can leave the researcher unsure of how to resolve open questions about writing the language.

References

[1]   Campbell *et al.*, *New knowledge: Findings from the Catalogue of Endangered Languages (ELCat)*. Mānoa: University of Hawaiʻi, 2013.

[2]   S. L. Walter and K. G. Chuo, *The Kom Experimental Mother Tongue Education Project Report for 2011*. 2011.

[3]   S. L. Walter and D. E. Dekker, "Mother tongue instruction in Lubuagan: A case study from the Philippines," *Int. Rev. Educ.*, vol. 57, pp. 667–683, 2011.

[4]   M. Cahill, "Non-linguistic factors in orthographies. In Michael Cahill & Keren Rice (eds.)," *Dev. Orthogr. Unwritten Lang.*, pp. 9–26, 2014.

[5]   D. M. Eberhard, G. F. Simons, and C. D. Fennig, "Ethnologue: Languages of the World. Twenty-second edition," in *SIL International*, 2019.

[6]   J. Lovestrand, "The dialects of Baraïn (East Chadic)," *SIL Electron. Work. Pap.*, 2011.

[7]   J. Lovestrand, "The linguistic structure of Baraïn (Chadic)," 2012.

[8]   M. Balle and J. Lovestrand, "How to write implosives in Kodi. In Yanti (ed.)," in *Konferensi Linguistik Tahunan Atma Jaya (KOLITA) 17*, 2019.

[9]   S. Bird, "Strategies for representing tone in African writing systems: A critical review," *Writ. Lang. Lit.*, vol. 2, pp. 1–44, 1999.

[10]   F. Seifart, *Orthography development. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.)*. Berlin: New York: Mouton de Gruyter, 2006.