

Developing an orthography for Onya Darat (western Borneo)

Practical and theoretical considerations

URI TADMOR

ABSTRACT

Onya Darat is a language spoken, with great dialectal variation, in the interior of western Borneo. It is the southernmost member of Land Dayak, a branch of the Austronesian language family. This article reports on the development of a writing system for Onya Darat. In addition to five vowels and 19 simple consonants, Onya Darat also exhibits three series of complex oral-nasal segments: prenasalized oral stops, preoralized nasals, and postoralized nasals. An analysis of the Onya Darat sound system reveals that of these three series only postoralized nasals are distinctive and therefore need to be represented in the writing system. The proposed orthography, developed with the aid of native speakers, represents all and only the phonemes of Onya Darat without resorting to diacritics or special characters.

KEYWORDS

Onya Darat, Land Dayak, orthographies, writing systems, phonology, Borneo, West Kalimantan

1 ORTHOGRAPHY DEVELOPMENT¹

Developing an orthography or writing system for a language is rather different from developing a systematic transcription. A transcription - whether phonemic or phonetic, narrow or wide - is aimed first and foremost at serving

¹ An earlier version of this paper was presented at the International Conference on Minority Languages, Central University for Nationalities, Beijing, 23-26 November 2007. I am grateful to Ardy Suhardi and to the people of Balai Berkuak without whose help and generosity this paper would not be possible.

URI TADMOR completed his PhD in linguistics at the University of Hawai'i, where he later headed the Indonesian language program. Since 1999 he has been Director of the Indonesia Office of the Max Planck Institute for Evolutionary Anthropology. He is also Adjunct Associate Professor of Linguistics at the University of Delaware, where he supervises graduate students specializing in languages of Indonesia. His research focuses on topics related to historical linguistics and sociolinguistics, and especially to language contact. E-mail: uri.tadmor@degruyter.com.

the academic community. On the other hand, an orthography should also be of service to the speech community of the relevant language, and must satisfy various practical and other demands.

Various issues, both linguistic and extralinguistic, must be considered when developing a writing system. From a linguistic point of view, an orthography must be phonemic. It should represent all and only phonemic contrasts in the language, whether segmental (consonants, vowels) or suprasegmental (tone, stress). It should not represent subphonemic contrasts, such as phonetic labialization of consonants before round vowels or phonetic palatalization before high vowels. Ideally, each phoneme of the languages should be represented by one character. This is known as the “one phoneme, one symbol” principle.

While ideal, the “one phoneme, one symbol” principle is not always practical. Sometimes a phonemic contrast has such a low functional load that it may not be worthwhile to make the effort to represent it in the orthography. This is especially true if expressing this contrast would require a special character or a diacritic. This touches upon the second great principle of orthography development: simplicity. The simpler the orthography, the easier it would be for native speakers to learn it and to use it correctly and efficiently.

While phonemicity and simplicity are the two golden rules of orthography development, they are not the only considerations. Seifart (2006) lists a variety of other issues, such as marking word boundaries, emblematic value, relation to existing orthographies (of neighboring or dominant languages), standardization dialect variation, and a range of technical production issues.² To these considerations I would like to add another one - intuitiveness. It is very important for the orthography to represent the language in a way that makes sense to native speakers. Sometimes an issue may be resolved by abstract linguistic analysis, but if it raises the objections of native speakers, it would not be as suitable as an alternative solution that may be more complicated or less elegant from a purely scientific point of view. As will be seen below, the intuitiveness principle has played an important role in the development of the Onya Darat orthography.

2 THE LANGUAGES AND ITS SPEAKERS

Onya Darat³ is a language spoken, with much dialectal variation, along several adjacent rivers in southwestern Borneo. Its main dialects are Kualan, Samandang, Baram, and Simpang. Together they constitute the southernmost member of Land Dayak, a language group spoken over a vast area in the

² For issues specific to developing orthographies for use with computers, see Anderson et al. 2005.

³ Speakers of Onya Darat actually have no ethnonym (neither autonym nor exonym) and no specific name for their language. However, they refer to themselves as Onya Darat (literally ‘Inland People’) to distinguish themselves from the mainly coastal Malays and from other groups, and this is the name I use for their language.

interior of western Borneo. Land Dayak languages belong to the western Malayo-Polynesian branch of the Austronesian language family, but their exact subgrouping within this branch is yet to be determined. Indeed, it is one of the least studied groups in Austronesian, and the entire body of literature on it consists of a handful of articles and dictionaries and one book (Rensch et al. 2006).⁴

3 DIALECT CHOICE

Onya Darat dialects and subdialects exhibit great variation, especially when it comes to phonology. They differ not only with regard to allophonic alternations and other phonetic details, but also with regard to morphophonemic alternations (and arguably even in their inventories of phonemes). It would therefore be impossible to create an orthography that would represent all dialects, and a choice had to be made regarding which dialect to use as the basis of the orthography. There is no standardized variety of Onya Darat, and no dialect is considered by speakers to be 'better' than the others. Moreover, none of the dialects is used as a lingua franca. Since the degree of mutual intelligibility among dialects is high, when speakers of different dialects converse each simply uses his or her native dialect. At times, some misunderstandings can arise due to unusual phonological forms or unique dialectal lexical items, but overall the communication is not hampered.

The dialect chosen to serve as the basis for the orthography was Kualan. This dialect is spoken (in many subdialects) along an eponymous river in the Ketapang regency of the Indonesian province of West Kalimantan. Two main reasons lie behind the choice of Kualan. First, although no precise demographical data exist, it probably has the largest number of speakers among all Onya Darat dialects. Second, it is the indigenous language of the district capital of Balai Berkuak, and as such is well known among speakers of all other dialects.

4 PREVIOUS STUDIES AND PUBLICATIONS ON AND IN ONYA DARAT

As already mentioned Land Dayak languages in general have been very poorly studied, and no linguistic studies have yet been published on Kualan or on any other Onya Darat dialect. The first attempts to write down the language date back to the 1970s, when American missionaries wanted to produce texts for religious instruction and for use in churches. Since about 1980, some texts in a language that appears to be a mixture of the Kualan and Samandang dialects have been produced and distributed locally; examples can be seen in Appendix A and Appendix B. In the mid 1990s, the Institute of Dayakology Research and Development in Pontianak (the provincial capital of West Kalimantan) initiated the publication of a series of booklets of folk stories of the region, in

⁴ It is remarkable yet telling that Blust 2009, a monumental work of 824 pages, barely mentions Land Dayak languages, despite the fact that the author has himself been working on languages of Borneo for decades.

local languages as well as in translations. Some of the booklets were in the Simpang dialect of Onya Darat (for a sample, see Appendix C).

5 THE SEGMENT INVENTORY OF KUALAN

Kualan has five vowels (Table 1) and nineteen simple consonants (Table 2).

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

Table 1. Vowels of the Kualan dialects

Place of articulation	Labial	Dental/Alveolar	Palatal	Velar	Glottal
Voiced stops	b	d	j	g	ʔ
Voiceless stops	p	t	c	k	
Nasals	m	n	ɲ	ŋ	
Liquids		l, r			
Fricatives		s			h
Glides	w		y		

Table 2. Simple consonants of the Kualan dialect

Three simple consonants are represented by IPA symbols not normally used in Romanized alphabets: /ɲ/, /ŋ/, and /ʔ/. In Indonesian, the national language of Indonesia, /ɲ/ is represented by the digraph <ny>, and /ŋ/ by the digraph <ng>. This practice is generally followed in the orthographies of local languages of Indonesia, and it was also applied in the orthography proposed here. The glottal stop is written less regularly in languages of Indonesia. Sometimes it is written as an apostrophe <'>, sometimes as <q>, sometimes as <k> (especially but not exclusively when it is in final position), and sometimes it is not represented in writing at all. For representing the glottal stop in Kualan, <q> was deemed the best choice, since this letter is not used for any other sound in the language, but is known to speakers and is present on the keyboard of typewriters (and computers) sold in Indonesia.

If these were the only issues, creating an orthography for Kualan would be a straightforward affair. However, in addition to the simple consonants in Table 2, Kualan also exhibits three series of complex oral-nasal segments: prenasalized oral stops, preoralized nasals, and postoralized nasals. The issues of phonemic status of these segments, whether they should be represented in the orthography, and if so – how, constituted the main challenges in creating an intuitive orthography for Onya Darat. They are discussed in detail in the following sections.

6 COMPLEX ORAL-NASAL SEGMENTS: PHONOLOGY

6.1 PRENASALIZED ORAL STOPS

If the onset of syllable consists of one of the nasals /m/, /n/, /ŋ/, or /ɲ/, the vowel nucleus is nasalized. If a final oral stop is preceded by such a nasalized vowel, the oral stop is prenasalized. This is not an insertion rule, since the nasal element is not an independent segment. It does not independent and cannot be separated from the oral element by any process. Rather, the prenasalization is best viewed as an assimilation process: a final oral stop becomes similar to a preceding nasalized vowel by being prenasalized. Some examples are provided in Table 3.

Final consonant	Plain	Prenasalized
p	/cirup/ [cirup] ‘grass’	/ɲanap/ [ɲanãmp̄] ‘hunt’
t	/kuret/ [kuret] ‘skin’	/rojet/ [rojẽnt̄] ‘sky’
k	/ibuk/ [ibuk] ‘short’	/comek/ [comẽk̄] ‘disgusted’

Table 3. Plain and prenasalized final oral stops in Kualan

Prenasalized (oral) stops are therefore not phonemic. Simple and prenasalized stops are allophones in complementary distribution: prenasalized stops occur in final position after nasalized vowels, and plain stops occur in all other environments. The phonetic prenasalization of final stops after a nasalized vowel is still productive, as demonstrated by recent loanwords:

/menit/ [menĩnt̄] ‘minute’ (< Indonesian *menit*)
 /tomat/ [tomãnt̄] ‘tomato’ (< Indonesian *tomat*)

6.2 PREORALIZED NASALS

If a final nasal consonant follows an oral vowel, it undergoes preocclusion. This is the mirror image of the process described above that produces prenasalized final stops. Similarly, this is not an insertion rule, because the occlusion is not a separate segment. This is also an assimilatory process: a final nasal becomes similar to a preceding oral vowel by being preoralized. Some examples are provided in Table 4.

Final nasal	Plain	Preoralized
m	/tonam/ [tonãm] ‘plant’	/ɲorum/ [ɲõrub̄m] ‘night’
n	/pongan/ [poŋãn] ‘thing’	/dien/ [died̄n] ‘3pl’
ŋ	/kadonãŋ/ [kadonãŋ] ‘swim’	/baãŋ/ [baãŋ̄] ‘all’

Table 4. Plain and preoralized final nasals in Kualan

Like prenasalized stops, preoralized nasals are not phonemic. Simple and preoralized nasals are allophones in complementary distribution: preoralized nasals occur in final position after oral vowels, and plain nasals occur in all other environments. The phonetic preocclusion of final nasals after an oral vowel is still productive, and occurs in recent loanwords:

- /pelam/ [pelab̥m] 'film' (Indonesian *filem*, Jakarta dialect [péləm])
- /ben/ [bed̥n̄] 'music band' (Indonesian *band* [ben])
- /mesen/ [mesed̥n̄] 'machine' (Indonesian *mesin* [məsin])

6.3 POSTORALIZED NASALS

Unlike prenasalized stops and preoralized nasals which only occur in final position, postoralized nasals occur only at the beginning of syllables. There are four preoralized nasals: *m̄b*, *n̄d*, *ŋ̄j*, and *ŋ̄g*, as in the following examples: *m̄bo* 'elder sibling', *tund̄u* 'above', *maŋ̄joa* 'daylight', *piŋ̄gan* 'dish'. Preoralized nasals are similar to prenasalized stops and preoralized nasals in that the occlusion does not constitute an independent segment, and cannot be separated from the nasal element by any process. Phonetically, the occlusion is barely audible, and an acoustic cue for its presence is the lack of the expected nasalization of the following vowel. There is no way to predict whether a plain or postoralized nasal will occur at the beginning of a final syllable; this is lexically determined. Postoralized nasals are therefore phonemes separate from plain nasals (and plain oral stops). While not as common as plain nasals, they are common enough for there to occur numerous minimal pairs. The distinction between plain and postoralized nasals therefore has a fairly high functional load and should be represented in the orthography. Examples for minimal pairs are *ona?* 'child': *on̄da?* 'want'; *mo* '2nd person singular pronoun': *m̄bo* 'elder sibling'; *tama?* 'enter': *tamb̄a?* 'grave'; *mana?* '1st person plural pronoun': *mand̄a?* 'despise'; *mano?* 'chicken': *mand̄o?* 'burn'.

7 COMPLEX ORAL-SEGMENTS: ORTHOGRAPHY

As mentioned above, two different orthographies have been developed for Onya Darat. In the earlier orthography, created in the 1970s by American missionaries (henceforth 'the missionary orthography'), prenasalized stops and preoralized nasals are not represented. This makes sense, since, as explained above, they are allophones of plain oral stops and nasals respectively, and occur in predictable environments. However, phonemic postoralized nasals are not represented either in the missionary orthography, and are written with the same symbols as those of plain nasals: /m/ and /m̄b/ are both represented by <m>; /n/ and /n̄d/ by <n>; /ŋ/ and /ŋ̄j/ by <ny>; /ŋ/ and /ŋ̄g/ by <ng>. A vowel following a plain nasal is marked with an acute accent mark, to indicate that it is nasalized, for example in the words *ronyét* 'sky, heaven' and *oná'* 'child'. Therefore, at least in principle, the distinction between plain and postoralized nasals is maintained, since nasalized vowels only follow plain

nasals. However, this solution is far from satisfactory, because it does not conform to native speakers' intuitions. The nasalization of vowels following plain nasals is completely predictable, automatic, and unconscious. Native speakers are not aware that these vowels are nasalized. When asked, they say that the words *manoq* 'chicken' and *mand̄oq* 'burn' have the same vowels but two different kinds of *n*. Indeed, the missionary orthography is in this respect so counter-intuitive that even its creators cannot adhere to it, and in their publications they fail to use an acute accent over most nasalized vowels.

A short text sample written in an early version of the missionary orthography can be seen in appendix A. In the sample, the acute accent appears in the words *ronyét* 'sky, heaven', and *oná* 'child'. But the accent mark is missing on most nasalized vowels, for example in the words *samuyá* 'all', *-neh* '3rd person singular pronominal clitic (appears in the words *ise' neh* 'its contents', *mulaneh* 'its beginning', and *ngaodanneh* 'name it'), *tonah* 'land', *nudong* 'cover', *nyen* 'distal demonstrative', *ngaodanneh* 'name it', *manyoa* 'day', *ngorum* 'night', *nya* 'for', *ngamagi* 'to separate' (unmarked nasalized vowels are in bold type). It may be that the creators of this orthography only intended to mark nasalized vowels in final syllables, because postoralized nasals only occur at the onset of final syllables. However, if the nasalized vowels are indeed phonemes independent from plain vowels, they should be written wherever they occur, not only in certain syllables or in certain words. Moreover, even in final syllables the creators of the missionary orthography fail to mark most nasalized vowels, as in *samuyá*, *mulaneh* 'its beginning' (and the other forms that exhibit the clitic *-neh*), *tonah* 'land', and *nyen* 'distal demonstrative'.

Appendix B contains a more recent version of the same text (the first verses of Genesis), produced in 2000. The translation itself is much better, and the spelling is better. However, nasalized vowels - even in final syllables - are still inconsistently marked. Already in the first line, we can see that the nasalized vowel in the word *ronyét* is marked with an acute accent, but the accent is missing in the words *mulaneh* and *tonah*, even though the vowels of their final syllables are nasalized.

A second orthography was developed in the 1990s by the Institute of Dayakology Research and Development (henceforth the IDRDR orthography). The creators of this orthography may have been unaware of the earlier missionary orthography, or perhaps they chose to ignore it. An example of the IDRDR orthography can be seen in appendix C. With regard to complex oral-nasal segments, the IDRDR orthography takes the following approach.

Prenasalization of stops is not marked. This can be seen in the word *ronyét* 'sky, heaven', which appears in the title, and in other words. Not assigning prenasalized stops unique symbols separate from plain stops conforms to the phonemic principle, because - as already discussed - prenasalized stops are allophones of plain stops which occur in final position following a nasalized vowel. On the other hand, preocclusion of nasals is marked, as in the words *ngorupm* 'night', *diētn* '3rd person plural pronoun', *lapm* 'in', *kawatn* 'friend',

and many others. This contradicts the phonemic principle, since preoralized nasals are allophones of plain nasals which occur in predictable positions. They cannot be confused with plain nasals, because they occur in different environments. Therefore they should not be part of a phonemically-based orthography. There is also an internal inconsistency: the prenasalization of final oral stops and preoralization of final nasal stops are parts of one and the same assimilatory process. Logically, if one is represented in the orthography, the other should be too, and if one is not represented, the other should not be represented either.

It may be argued that marking preoralized stops does serve an orthographic-phonemic purpose: it indicates that the preceding vowel is not nasalized, which means that a nasal occurring before the vowel is postoralized. For example, in the word *mepm* 'don't', which appears on the last line, preoralization indicates that the initial consonant is the postoralized nasal /m̄b/ rather than the plain nasal /m/. If the initial nasal were a plain nasal, the vowel would be nasalized, and the final nasal wouldn't be preoralized. This, however, would be a rather convoluted and counter-intuitive way of marking post-occlusion of nasals, and was probably not the intention of the orthography's creators. Marking the presence of postoralized nasals this way requires representing the preocclusion of final nasals orthographically even when there is no preceding postoralized nasal, as in the words *dietn* '3rd person plural pronoun' and *kawatn* 'friend'. Moreover, such a system would not mark phonemic postocclusion if the word happens to end with a vowel, as in the words *somu* /sombu/ 'above' and *onu* /ondu/ 'day' (all these examples appear in Appendix C).

Postoralized nasals in the IDR orthography are written with the same symbols as plain nasals. For example, in the words *monik* /mond̄i?/ 'come' and *nyamot* /n̄ambot/ 'receive guests' the characters *m* and *n* stand for plain nasals, while in the words *dango* /daŋo/ 'hut' and *simak* /sima?/ 'go up' the characters *ng* and *m* stand for plain nasals. There is no attempt to mark vowel nasalization even as sporadically as the missionary orthography. This situation is not satisfactory from a theoretical point of view, because it ignores phonemic distinctions. It is also unsatisfactory from a practical point of view, since it can lead to ambiguities and misunderstandings. As already mentioned above, there are quite a few minimal pairs that are only distinguished by the presence of plain or postoralized nasals, and a good example can be seen in Appendix B. The word *onakko* /ond̄ak-ko/ (which appears close the bottom of the page) means 'I want', but in the IDR orthography the exact same spelling can also stand for /onak-ko/, which has a completely different meaning: 'my child'.

When I first started collecting Kualan texts with the assistance of Ardy Suhardi, a native speaker of the dialect, it soon became apparent that neither the missionary orthography nor the IDR orthography was adequate for our purposes. The orthographies were not suitable for the native-speaking consultant, because they did not match his intuitions. Nor were they adequate

for me as a linguist, since they did not follow the phonemic principle. (In fact, both the missionary orthography and the IDR orthography have other shortcomings, but they are beyond the scope of the present paper.) We therefore decided to develop a third orthography (henceforth the phonemic orthography) which we found more suitable for our needs. As it happens, none of the native speakers we have worked with has ever used either the missionary orthography or the IDR orthography for writing *Kualan*; they simply never wrote the language down. So the new orthography we developed did not have to replace or compete with another system.

The approach in developing our orthography (henceforth 'the phonemic orthography') is very simple. All phonemes are assigned unique characters (letters or letter combinations), and no allophones are assigned unique characters. Prenasalized stops, which are predictable allophones of plain stops, are spelled with the same characters as plain stops: in the word *nganap* / *ɲanap* / [ɲanap̃] 'hunt' the final consonant is spelled the same way as the final consonant the word *ontap* / *ontap* / [ontap] 'hard'; in the word *ronyet* / *ronyet* / [ronyent̃] 'sky, heaven' the final consonant is spelled the same as in *orut* / *orut* / [orut] 'boat'; and in *comek* / *comek* / [comeñk] 'disgusted' the final consonant is spelled the same as in *colok* / *colok* / [colok] 'torch'. Similarly, preoralized nasals, which are predictable allophones of plain nasals, are spelled with the same characters as plain nasals: in the word *ngorum* / *ɲorum* / [ɲorub̃m] 'night' the final consonant is spelled the same way as the final consonant of the word *tonam* / *tonam* / [tonam] 'to plant'; in the word *dien* / *dien* / [died̃n] '3rd person plural pronoun' the final consonant is spelled the same as in *meen* / *meen* / [meen] 'like this'; and in *baang* / *baan* / [baang̃] 'all' the final consonant is spelled the same as in *ngomong* / *ɲomon* / [ɲomon] 'speak'.

Not representing predictable allophones in the orthography was a simple procedure to implement. A somewhat greater challenge was how to distinguish in writing between plain nasal phonemes and postoralized nasal phonemes. Whatever the solution was, it had to accomplish several things: to accommodate the native speakers' intuition that they were indeed separate phonemes; to be iconic and easy to remember; and be easy to write (or type on a typewriter). The solution agreed upon was to write postoralized nasals as digraphs, consisting of symbols for plain nasals followed by symbols for their homorganic voiced stops: <mb>, <nd>, <nyj>, <ngg>. Using digraphs to represent preoralized nasals as several advantages:

The digraphs are unique. They contrast with plain nasals, which are spelled *m*, *n*, *ny*, *ng* respectively. They do not coincide with sequences of nasals and voiced stops, because such sequences do not occur in the language. (Indeed, it is possible that postoralized nasals developed historically from such sequences.)

The digraphs are iconic. Postoralized nasals start out as nasals and end as occlusives, as do the digraphs.

The digraphs are easy to write and to type, and need no special symbols or diacritics.

This solution has one major disadvantage: nonnative speakers who learn Onya Darat might produce spelling pronunciations of postoralized nasals as sequences of plain nasals followed by homorganic voiced stops. For example, they may pronounce the word *tundu* 'above' as [tundu] instead of as [tundu]. (This problem does not arise for native speakers, since, as mentioned above, sequences of plain nasals followed by homorganic voiced stops do not occur in their language.) But causing a possible problem for nonnative learners is a small price to pay for the advantages of phonemicity, simplicity, intuitiveness, iconicity, and ease of writing and printing. After all, the primary purpose of an orthography is to serve its native speakers, not to serve linguists, missionaries, or other nonnative speakers.

The digraph solution proved suitable for our purposes: it is easily learned by native speakers, and because it is phonemic, it can also be used in scientific writing by linguists. Since earlier orthographies for Onya Darat have not caught on (perhaps partially because they were counter-intuitive), the phonemic orthography can now be used by speakers of all Onya Darat dialects.

8 CONCLUSION

This article presented the various factors that were taken into consideration while developing a simple, phonemic orthography for the Onya Darat language. The resulting orthography avoids using diacritics or special characters and conforms to native speakers' intuitions. Prenasalized stops and preoralized nasals, which are allophones in complementary distribution with plain stops and nasals, are not represented in the orthography. On the other hand postoralized nasals, which are phonemically distinct from plain nasals, are distinguished by using digraphs.

REFERENCES

- Anderson, Deborah, Rick McGowan, and Ken Whistler. Incorporating comments by Lorna Priest), 2005. *Recommendations for creating new orthographies*. Unicode Technical Note #19, accessed online at <http://www.unicode.org/notes/tn19/>.
- Blust, Robert. 2009. *The Austronesian languages*. Canberra: Research School of Pacific and Asian Studies, The Australian National University. [Pacific Linguistics 602.]
- Buko' Kajadi. 2000. "Onya Darat translation of Genesis". Published by Gereja Firman Allah di daerah Semandang / Kualan and Y.M.M.P.
- Kitab Asal Mulaneh. 1981. "Onya Darat translation of Genesis 1-11:9". Publisher not indicated.
- Nokng. 1996. "Tabuok ka Ronyet (Tabuok goes to heaven)", in: F.X. Beleng, Livinus Prianidi, Sosimus, Laurensius Salem, Nico Andasputra, and Stephanus Djuweng (eds), *Tabuok ka Ronyet; Seri Cerita Rakyat Kalbar*. Pontianak: Institute of Dayakology Research and Development.

- Rensch, Calvin R., Carolyn M. Rensch, Jonas Noeb, and Robert Sulis Ridu. 2006. *The Bidayuh language; Yesterday, today and tomorrow*. Kuching: Dayak Bidayuh National Association.
- Seifart, Frank. 2006. "Orthography development", in: Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel (eds), *Essentials of language documentation*, pp. 275-299. Berlin and New York: Mouton de Gruyter. [Trends in Linguistics Studies and Monographs 178.]
- Tadmor, Uri and Ardy Suhardi. 2002-2009. "Onya Darat database". [A project of the Jakarta Field Station of the Max Planck Institute for Evolutionary Anthropology.]

APPENDICES


APPENDIX A

An early example of the missionary orthography (Kitab Asal Mulaneh 1981)

First verses of Genesis

Kejadian 1 1

Allah Ngaodeh Ronyét
ba Tonah ba Samuya' Ise'neh



1) Asal mulaneh Allah ngaodeh ronyét ba tonah.

2) Tonah ogon baempah, ge' puang: Cicit-dalit nudong laot aih, aba Roh Allah manginang di tunu piing.

3) Allah ngomong, jaIDoh: "Odehlah tanpa'!"
Lalu tanpa' pun odeh. 4) Allah mantau tanpa' nyen bagas. Lalu banagiDoh tanpa' dai sagam.

5) Mangka Allah ngaodanneh tanpa' nyen manyoa, aba sagam nyen ngorum. Dah odeh samonu ba engkolam, nyenlah onu yang kamaro'.

6) Allah ngomong, jaIDoh: "Odehlah sapat di tuda piing nya' ngamagi piing dai piing."

7) Mangka Allah nangkoi sapat na' ngamagi piing yang odeh di sigin sapat nyen dai piing yang odeh

APPENDIX B

A more recent example of the missionary orthography (Buko' Kajadi 2000)
First verses of Genesis

BUKO' KAJADI

Pasal 1*Gesah Allah ngaodeh dania ba samuyá' ise'neh*

Sae'seh, jong asal mulaneh, Allah ngaodeh ronyét ba tonah, aba samuyá'neh yang odeh. ²Jong nyén sae'seh dania ento ogon baempah. Ge'pun jong nyén dania ge' puang, cicit-dalit, sagam, ge' libo' lai laot dalam. Ge'pun Roh Allah manginang di tunu piing nyén.

³Lalu Allah ngomóng. JaiDoh: "Odehlah tampa'!" Lalu taka "lar", tampa' pun omát odeh. ⁴Lalu Allah mantau tampa' nyén seja ka bagasam. Lopas nyén, lalu banagi Doh tampa' dai sagam. ⁵Tampa' nyén naodan Doh manyoa-onu, aba sagam nyén ngorum. Seta dah santo' samonu, terus ka engkolamneh, nyéném deh onu yang kamaro'neh jong nyén sae'seh.

⁶Lopas nyén Idoh ngomóng ge'. JaiDoh: "Odehlah sapat di tuda piing, nyá' ngamagi piing dai piing." ⁷Lalu pas ka omóngDoh nyén, odehem sapat antara piing yang odeh di sigin ronyét, aba piing yang odeh di somú ronyét. ⁸Lalu Allah ngaodan sapat antara nyénseh ronyét. Lalu onu nyén pun ngorum, lalu ka engkolam, nyéném deh onu yang panganukahneh.

⁹Lopas nyén Allah ngomóng ge'. JaiDoh: "Bakolom piing yang di sigin ronyét ka site' arong, ge' sampaya timol tonah yang rangkai!" Lalu piing pun bakolo. ¹⁰Lalu Allah ngaodan tonah yang rangkai nyénseh darat, aba piing yang bakolo nyénseh naodan Doh laot. Ge'pun Allah mantau samuyá'neh nyén bae'-bagas bangam.

¹¹Lopas nyén Allah ngomóng ge'. JaiDoh: "Timolom di tonah sagala bansa patanam-patumoh, sagala koyuh-koyuh, sagala ka pongan yang baige' empah podi deen, sagala ka koyuh yang babuah, yang buahneh baige'. Ge' sampaya jong ige'neh dah balantang paji, satiap bansaneh nyén ajah-ajah ngasel buah dop ném." Lalu pada' patanam-patumohseh, aba sagala koyuh-koyuh nyén pun odeh. ¹²Seja omát timolom di tonah pada' patanam-patumohseh, aba koyuh-koyuh, samuyá' bansa koyuh yang buahneh baige'. Samuyá' ige'neh nyén, jong io dah balantang paji, ajah-ajah ngurut bansa dop ném. Ge'pun Allah mantau samuyá'neh nyén seja

APPENDIX C

Example of the IDRD orthography

Beginning of the folktale 'Tabuok ka Ronyet' (*Tabuok goes to heaven*)
(Nokng 1996)

TABUOK KA RONYET

Onu udah ngorupm, sagapm kararopatneh. Me taok-taok takng ka monik apm kawatn-kawatn Tabuok seh, ka dango idoh. Uдах monik ka koetn seh, nyapa Tabuok.

"O... Tabuok, odeh mo," je dietn nyen seh.

"Osih nyen?" je oso Tabuok nyamot lapm dango cuh.

"Ome," je kawatn-kawatn Tabuok.

"Ome osih," je oso Tabuok agek.

"Ome bah, kawatn-kawatn Tabuok," je bala kawatn Tabuok seh.

"Oh, dih. Simak apm datn neh," je oso doh.

Uдах nyen oso doh pun nganyongok ka panampbakng. Kawatn Tabuok seh pun simak apm ka somu. Lalu oso doh tamak ka dalapm, nyapa Tabuok. Mpbah neh idoh seh agek ka elakng-elakng neh ngarowot bila jare, di tonyu cuh.

"O, Tabuok," je oso doh.

"Oei," je Tabuok.

"Odeh kawatn-kawatn omo, montde di soah cuh," je oso doh agek.

Tabuok pun lalu ka luar, nomus kawatn-kawatn idoh seh. Uдах nyen idoh mincjakng kawatn-kawatn idoh simak ka padokng.

"Onih agah kadietn seh," je idoh.

"Ah, bagas bah dih?" je kawatn-kawatn idoh.

"Ntdo, ome toh kojakng gesah kojakng carita," je kawatn-kawatn idoh gek.

"Yo, marapm ngopikng neh," je Tabuok agek.

"Onak me omo aba ome blayar ka Jawa," je kawatn-kawatn idoh ngajak Tabuok.

"Oei...., onakko dih," je idoh.

"Ah, omo mepm kadodap bah Tabuok. Odop udah muntduh,

APPENDIX D

Sample transcription in the phonemic orthography
Opening of the Kualan folktale 'The Turtle and the Monkey'
(Tadmor and Suhardi 2002-2009)

(Abbreviations: 12PL 'first plus second person plural pronoun'; 3DU= 'third person dual pronoun used when both referents are members of the same generation'; 3PL≠ 'third person plural pronoun used when referents are not members of the same generation'; 3sg≤ 'third person pronoun used when referent is member of the same generation or younger generation with regard to another person'; ACT 'active prefix'; AGT 'agent-deriving prefix'; COHORT 'cohortative particle, "let's"': EXPL 'particle used to introduce explanatory statements'; INTR 'intransitive verb marker'; PF 'perfective/imperative particle'; REL 'relative particle'; TOP 'topic marker'; TRU 'truncated form'.)

jadi-eh jaman haeq-eh odeh ngkura ba boroq.
become-TOP period formerly-TOP exist turtle with k.o.monkey
Once upon a time there were a turtle and a monkey.

doduh badukah-eh bakawan dah lambat-am.
3DU= INTR-two-TOP INTR-friend after slow-PF
They were both old friends.

jadi, nyang odan-neh boroq ka raja pangalampur, raja rangkaq.
become REL name-3SG≤ turtle EXPL king AGT-lie king greedy
The monkey was a big liar and very greedy too.

jadi, jay boroq nyangkaq ngkura.
become say k.o.monkey tell turtle
One day the monkey said to the turtle:

"yoh odup betanding nonam boraq, san!"
COHORT 12PL INTR-match ACT-plant banana TRU-in.law
"Let's have a banana planting competition!"

"nyom!", jay ngkura nyambot.
COHORT say turtle ACT-answer
"Okay!", answered the turtle.

koih denaq ngomang tompang boraq.
go.there 3PL≠ ACT-look.for sapling banana
So they went looking for banana saplings.

jadi ngkura-eh ngundah onaq boraq nyang geq ocek naq-neh tompang.
become turtle-TOP ACT-gather child banana REL still small for-3SG≤ sapling
The turtle gathered young banana saplings,

jaq boroq-eh nyobot puun boraq nyang dah maa kaih-neh, dah bebuah-am.
but k.o.monkey-TOP ACT-yank.out tree banana REL after very big-3SG≤ after INTR-fruit-PF
but the monkey uprooted large banana trees that were already bearing fruit.