ISO-639-3 and PARADISEC

Nick Thieberger
University of Melbourne
<table>
<thead>
<tr>
<th>Item details</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item ID</td>
<td>AA1-001</td>
</tr>
<tr>
<td>Title</td>
<td>Pak Nongeng tells about the first four people, food taboos, and various other topics</td>
</tr>
<tr>
<td>Description</td>
<td>Pak Nongeng tells about the first four people, food taboos, and various other topics</td>
</tr>
<tr>
<td>Origination date</td>
<td>1987-11-17</td>
</tr>
<tr>
<td>Archive link</td>
<td><a href="http://catalog.paradisec.org.au/repository/AA1/001">http://catalog.paradisec.org.au/repository/AA1/001</a></td>
</tr>
<tr>
<td>Collector</td>
<td>Alexander Adelaar</td>
</tr>
<tr>
<td>Countries</td>
<td>Indonesia - ID</td>
</tr>
<tr>
<td>Language as given</td>
<td>Salako, Badamea</td>
</tr>
<tr>
<td>Subject</td>
<td>Kendayan - knx</td>
</tr>
<tr>
<td>Region / village</td>
<td>Sasak village, Samalantao, Sambas Regency, West Kalimantan</td>
</tr>
</tbody>
</table>

http://catalog.paradisec.org.au/collections/AA1/items/001
You gotta be in it to win it

May 29, 2011, 8.43 am by Nick Thieberger

Peter Austin's blog post deals with online endangered language archive searchability. As one of the targets of his latest post, PARADISEC apparently does not provide him with the results he wants in searching a catalog. Searching for 'Educational material' in a catalog makes lots of assumptions about the way that catalog has been constructed, one of which must be that the term is provided by the catalog or that the typical depositor would use the term in their freeform description of the item. Strangely, the answer he offers is not to provide the infrastructure on which such searches may succeed in future, but to advocate a folksonomy in which such searches will always be sure to fail.

The post is an advertisement for what is undoubtedly a very nice interface to a set of material held by ELAR, but we should also bear in mind the large amount of funding that ELAR/ELDP have had, so we would hope for at least a nice looking webpage after eight years now. It is also interesting that ELAR holds only 70 collections after ELDP has funded 216 projects, what has happened to the rest of the material, or am I being too commodifying to think of such a thing?

The comments on the post raise OLAC – a great service that provides information for the broader community (including linguists), but especially speakers who can access it via...
Standard terms or folksonomy?

• Looking for ‘songs’ will not find ‘song’

• Looking for ‘kastom’ will not find ‘Custom description’ or ‘Custom narrative’ or ‘Custom story’, let alone ‘Folk Tale’, ‘Narrative’, ‘Myth narrative’, ‘Narration’, ‘Narrative from visual prompt’, and many more.
‘Language: Arandic’
‘Arrernte’
‘Arandic’ does feature in lists like Glottolog and Multitree (which deal with language groupings).
OLAC resources in and about the Western Arrarnta language

The combined catalog of all OLAC participants contains the following resources that are relevant to this language:

- Primary texts
- Lexical resources
- Language descriptions
- Other resources about the language

Other known names and dialect names: Akara, Akirre, Aranda, Western Arrarnta, Anunta, Southern Aranda, Western Aranda

Use faceted search to explore resources for Western Aranda language.

Primary texts


Lexical resources


Language descriptions


Other resources about the language


http://www.language-archives.org/language/are
MARC

• 515 items

• Groupings of language at various levels

• Klingon

http://www.loc.gov/marc/languages/language_code.html
Marc list as published
http://www.loc.gov/marc/languages/language_name.html viewed 20/12/12

Agta

Assigned collective code [phi]
(Philippine (Other))
UF Cagayan Agta, Central
Central Cagayan Agta
Agta (Casiguran)
USE Dumagat (Casiguran)
Agta (Umirey)
USE Dumagat (Umirey)
10/12/11
Dear Linguists,


Yes, there is a variation in the language name. My proposal to create a subject heading cross-reference from "Dupaningan" to "Dupaninan" is pending.

For those interested, the call number for Laura's thesis is Hawn AC1 .H3 no.5072. It is available on the web: http://micro189.lib3.hawaii.edu/ezproxy/details.php?dbId=2335

-- Aloha,

XXXX, Librarian
Cataloging Department
Thomas Hale Hamilton Library
University of Hawai‘i at Mānoa
Linguasphere

70,900 linguonyms
4,500 outer languages
100 referential zones
94,049 languoids

http://www.glottolog.org
Choose by property

Name wala
- whole name
- part of a name
- include non-English names
- include trees other than main tree

ISO 639-3

Glottolog ID (alphanumeric)

Submit

search results

The following languoids match the query with their primary name:

Austronesian: Wala [lg] (Lengua Wala, Wala language)

The following languoids have at least one alternate name which matches the query:

Atlantic-Congo: Dagaari Dioula [dgd] (Wala)
Atlantic-Congo: Wali (Ghana) [wtx] (Wala)

http://www.glottolog.org
The World Loanword Database contains information on loanwords, source words and other words in 395 languages.

Language: Ceq Wong

Affiliation: Austro-Asiatic, Aslian

ISO 639-3 codes: cwg

http://wold.livingsources.org/language/26
Rosetta

Lingual nodes

• Unique identifier
• Metadata (human readable name for the node)
• Actual documents in the language

Rosetta

- Subject Verb Object [SVO] = RDF triple
- Reification of triples
- [[SVO] V O]
Book C
Title: Why French and Italian are related to Latin
Mapping to a standard

1:1 – [given term] to ISO-639-3 = ‘sameAs’

1. Related (rel:related)
   1.1. Same as (rel:sameAs)
   1.2. Almost same as (rel:almostSameAs)
   1.3. Broader than (rel:broaderThan)
       1.3.1. Super class of (rel:superClassOf)
       1.3.2. Has part (rel:hasPart)
           1.3.2.1. Has direct part (rel:hasDirectPart)
   1.4. Narrower than (rel:narrowerThan)
       1.4.1. Sub class of (rel:subClassOf)
       1.4.2. Part of (rel:partOf)
           1.4.2.1. Direct part of (rel:directPartOf)

http://lux13.mpi.nl/relcat/site/index.html
Kendayan rel: superClassOf Salako

HasReliabilityIndex 5
IsReliable
IsPartiallyReliable
IsUnreliable

Linguasphere uses an asterisk to indicate ‘items of data which are unreliable or which require corroboration’

Item details

Item ID: AA1-001
Title: Pak Nongeng tells about the first four people, food taboos, and various other topics
Description: Pak Nongeng tells about the first four people, food taboos, and various other topics
Origin date: 1987-11-17
Origin date free form: 1987
Archive link: http://catalog.paradiseec.org.au/repository/AA1/001
URL: 
Collector: Alexander Adelaar
Countries: Indonesia - ID
Language as given: Salako, Badamea
Subject: Kendayan - knx

Reliability:
- Reliable
- Partially Reliable
- Unreliable

Language(s): Kendayan - knx
Content: Kendayan - knx
Dialect: 
Region / village: Soso village, Sambalant, Samba Regency, West Kalimantan

Map of the region.