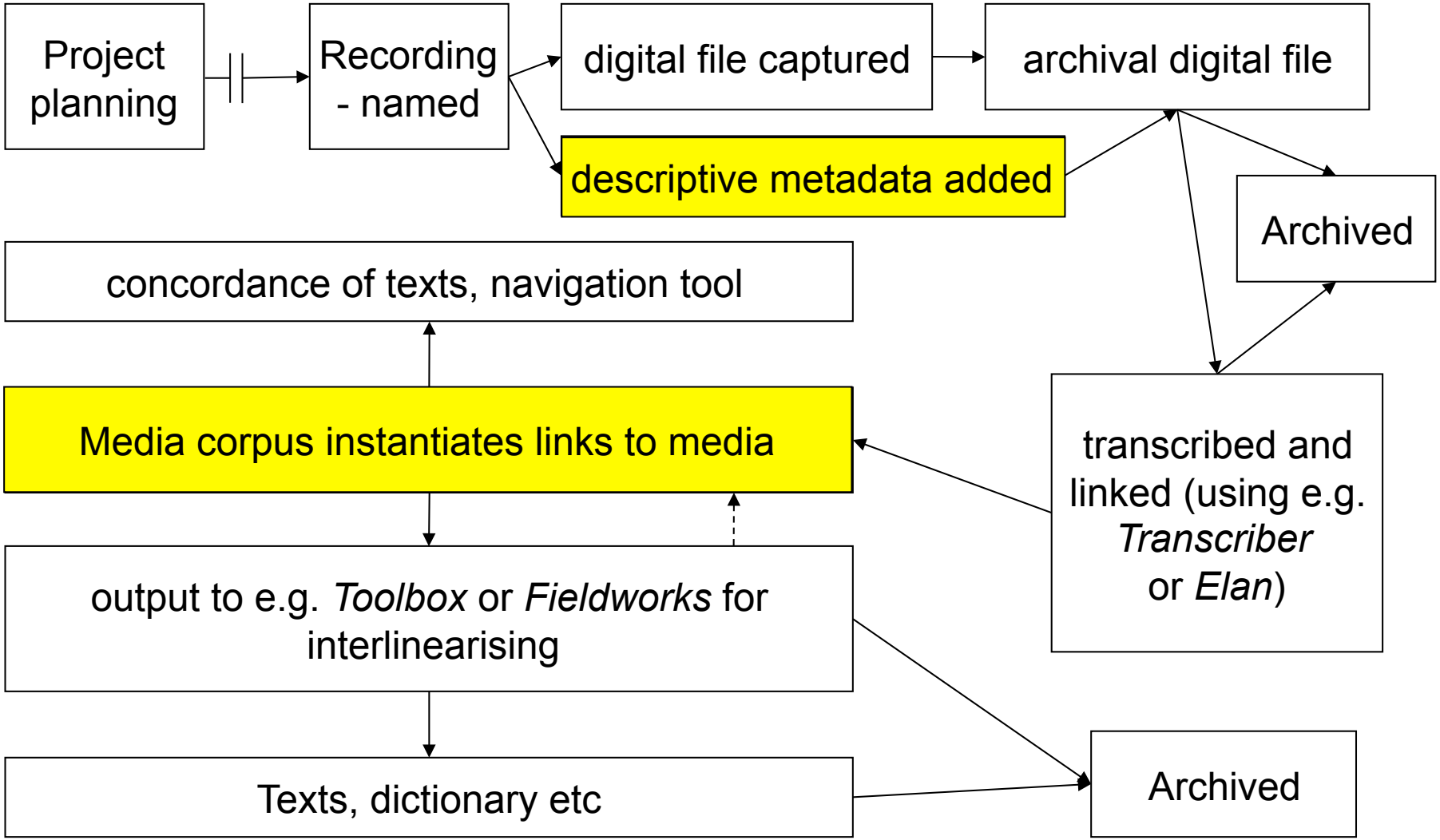
Pathways to reusability for fieldwork records

Exploring data from language documentation,

Berlin, May 10-11, 2013

Nick Thieberger
University of Melbourne

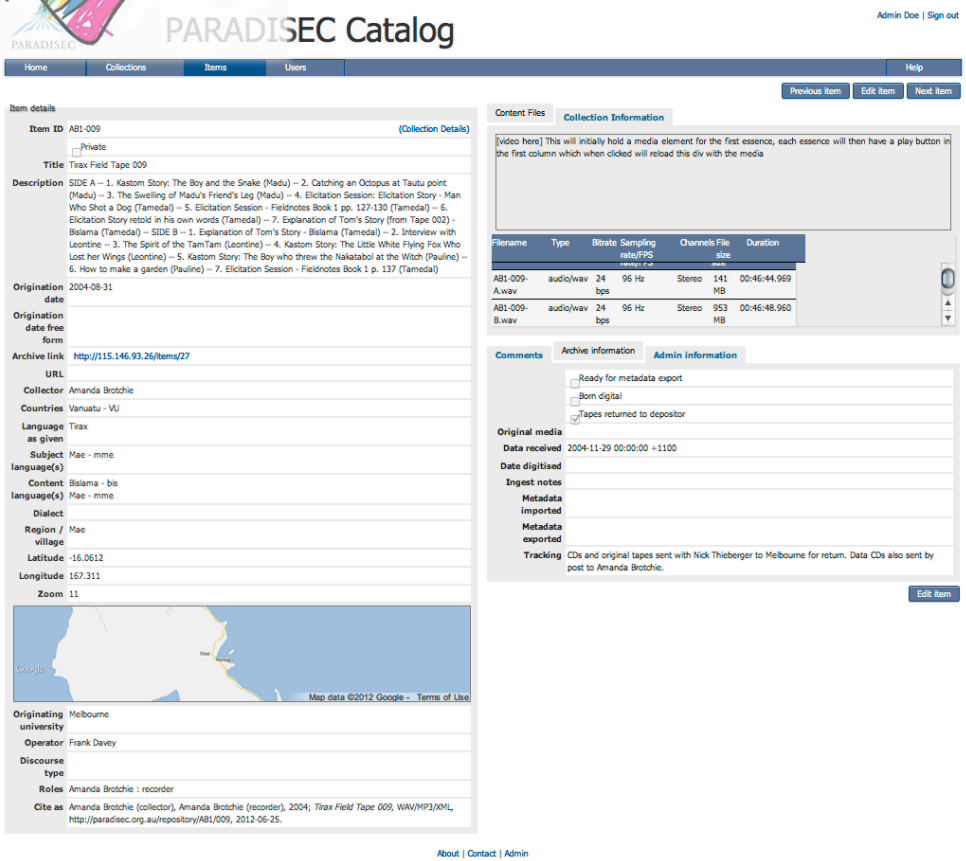
Typical workflow resulting in well-formed data



PARADISEC

Pacific and Regional Archive for Digital Sources in Endangered Cultures

Safeguarding research in Australia's region



PARADISEC Catalog

Admin Doe | Sign out

Home | Collections | Items | Users | Help

Previous item | Edit item | Next item

Item details

Item ID AB1-009 (Collection Details)

Private

Title Tirax Field Tape 009

Description SIDE A -- 1. Kastom Story: The Boy and the Snake (Madu) -- 2. Catching an Octopus at Tautu point (Madu) -- 3. The Swelling of Madu's Friend's Leg (Madu) -- 4. Elicitation Session: Elicitation Story - Man Who Shot a Dog (Tamedal) -- 5. Elicitation Session - Fieldnotes Book 1 pp. 127-130 (Tamedal) -- 6. Elicitation Story retold in his own words (Tamedal) -- 7. Explanation of Tom's Story (from Tape 002) - Bislama (Tamedal) -- SIDE B -- 1. Explanation of Tom's Story - Bislama (Tamedal) -- 2. Interview with Leontine -- 3. The Spirit of the TamTam (Leontine) -- 4. Kastom Story: The Little White Flying Fox Who Lost her Wings (Leontine) -- 5. Kastom Story: The Boy who threw the Nakatabol at the Witch (Pauline) -- 6. How to make a garden (Pauline) -- 7. Elicitation Session - Fieldnotes Book 1 p. 137 (Tamedal)

Origination date 2004-08-31

Origination date free form

Archive link http://115.146.93.26/items/27

URL

Collector Amanda Brotchie

Countries Vanuatu - VU

Language as given Tirax

Subject language(s) Mae - mme

Content language(s) Bislama - bis
Mae - mme

Dialect

Region / village Mae

Latitude -16.0612

Longitude 167.311

Zoom 11

Map data ©2012 Google - Terms of Use

Originating university Melbourne

Operator Frank Davey

Discourse type

Roles Amanda Brotchie : recorder

Cite as Amanda Brotchie (collector), Amanda Brotchie (recorder), 2004; *Tirax Field Tape 009*, WAV/MP3/XML, http://paradisec.org.au/repository/AB1/009, 2012-06-25.

Content Files | Collection Information

[video here] This will initially hold a media element for the first essence, each essence will then have a play button in the first column which when clicked will reload this div with the media

Filename	Type	Bitrate	Sampling rate/FPS	Channels	File size	Duration
AB1-009-A.wav	audio/wav	24 bps	96 Hz	Stereo	141 MB	00:46:44.969
AB1-009-B.wav	audio/wav	24 bps	96 Hz	Stereo	953 MB	00:46:48.960

Comments | Archive information | Admin information

Ready for metadata export

Born digital

Tapes returned to depositor

Original media

Data received 2004-11-29 00:00:00 +1100

Date digitised

Ingest notes

Metadata imported

Metadata exported

Tracking CDs and original tapes sent with Nick Thieberger to Melbourne for return. Data CDs also sent by post to Amanda Brotchie.

Edit item

About | Contact | Admin

Documentation index

Jaru

K.012

Map E 5210

ethcode DDJ

State WA

list view | go to intro | show all | find by loc | speaker # | resources | language info | find aiatsis code

Documentation (0- 20) as of 2/12/2002 = **10**

- ○ No documentation known
- ○ No more language documentation possible

- ○ Detailed dictionary (eg Arrernte, Kayardild)
- ◉ Medium dictionary (eg Bardi)
- ○ Small dictionary/ wordlist (eg Gamilaraay)
- ○ Simple wordlist (eg Bates, Curr) 3

linguist Tsunoda, Tasaku

- ◉ Extensive text collection
- ○ Several texts (<10)
- ○ Elicited/example sentences 3

- ○ Detailed grammar (eg PhD Thesis, publishec
- ○ Middle-sized grammar (eg. Handbook)
- ○ Grammar sketch or many technical articles
- ○ Few technical articles only 0

language programmes

- ○ Substantial ethnolinguistic work (eg thesis)
- ○ Ethnolinguistic description
- ○ Some ethnolinguistic information 0

- ◉ More than several hours of audio
- ○ Less than several hours of audio
- ○ Less than an hour of audio 3

notes

- ○ Literature, AV > 1000 words
- ○ Literature, AV <1000 words
- ◉ Literature, AV < 300 words 1

AUSTLANG
Australian Indigenous Languages Database

Language information

hide map enlarge map refresh map

AIATSIS Code: K12

AIATSIS reference name: Jaru

Documentation information is based on materials held at AIATSIS and also listed in OZBIB (a linguistic bibliography of Aboriginal Australia and Torres Strait Islands). Click 'Search MURA' and 'Search OZBIB' under the 'Resource' tab to view these materials. Click here to read the details of documentation scoring system.
Documentation score (0-16) = 13

Type	Documentation Status	Documentation Score
Word list	Large (more than 200 pages)	4
Text Collection	Small (20-100 pages)	2
Grammar	Large grammar (more than 200 pages)	4
Audio-visual	More than 10	3

Manuscripts/field notes in the AIATSIS library:
tape transcription/field note available

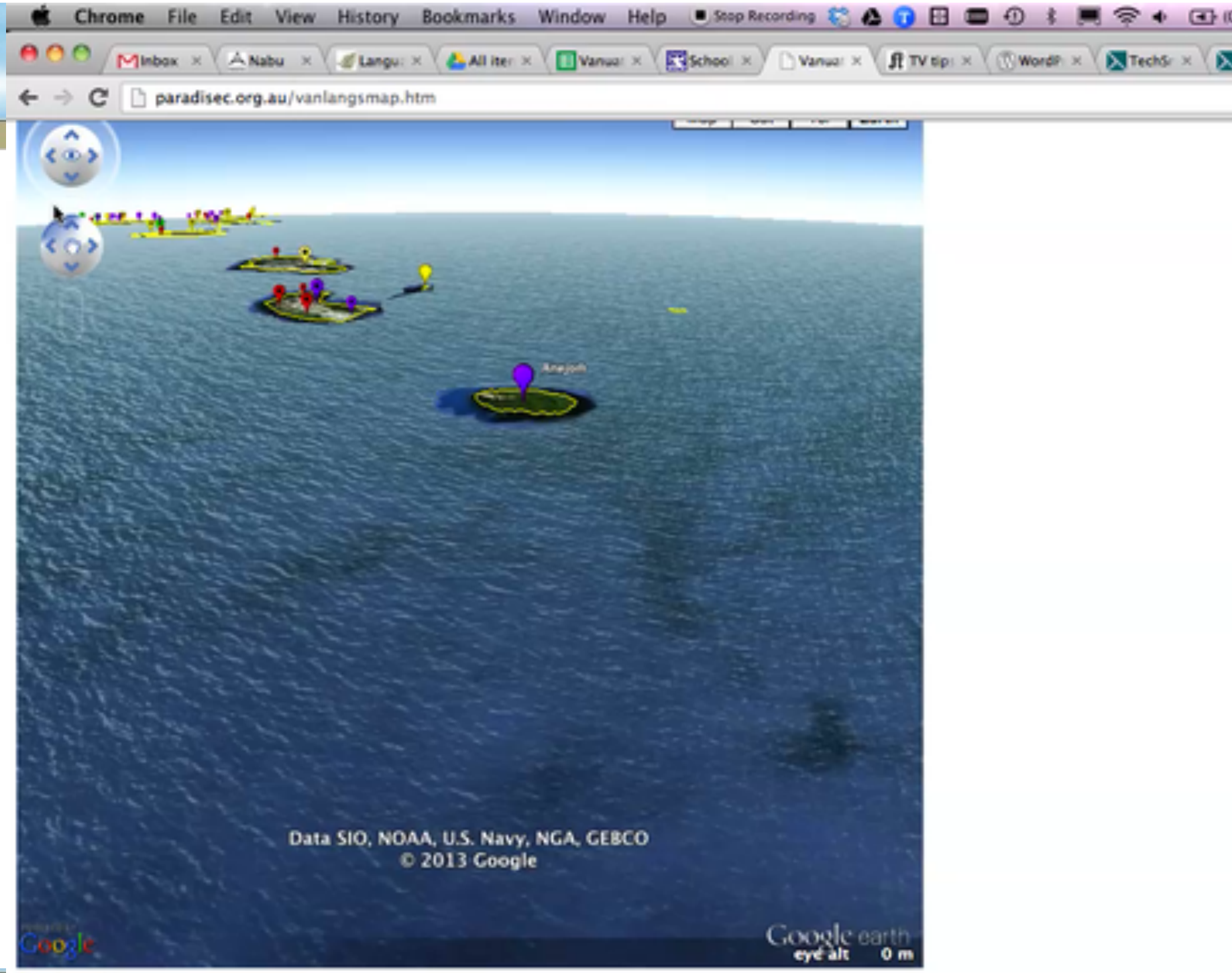
Grammar:
Tsunoda, Tasaku. 1981. The Djaru language of Kimberley, Western Australia: Pacific Linguistics B78. Canberra: Pacific Linguistics.

Dictionary:
Kimberley Language Resource Centre. Jaru Dictionary, ASEDA 0248. Kimberley Language Resource CentreTsunoda's Jaru, ASEDA 0510

Information as of: 2013-03-20 10:26:25

http://austlang.aiatsis.gov.au/

PARADISEC



View Larger Map

A proper collection

- representative range of participants recorded
- the use of both audio and video recording
- a mix of discourse styles
- and so on

- AND that the collection is described to a level that allows the items within it to be located by interested outsiders.

Services for language collections

- Open Language Archives Community (OLAC)
- CLARIN (Eu)
- DARIAH (Eu)
- Bamboo (USA)
- HuNI (Au)
- HCS vLab (Au)

Problems with collections we are given

Metadata only at the highest level

- 'collection of recordings made on June 5th 1966 in village X'

No dates, no personal names of speakers/performers etc.

Digital files or analog objects (tapes, papers) with no metadata at all

Files buried in several levels of directories (which each have meaningful names) but all files are named identically within each directory so that they have to be renamed to inherit the semantics associated with their location in a particular directory

- 20110305
  - IMG_001 >
  - IMG_002 >
  - IMG_003 >
  - IMG_004 >
- 20110315
  - IMG_001 >
  - IMG_002 >
  - IMG_003 >
  - IMG_004 >
- 20110321
  - IMG_001 >
  - IMG_002 >
  - IMG_003 >
  - IMG_004 >

Files buried in several levels of directories (which each have meaningful names) but all files are named identically within each directory so that they have to be renamed to inherit the semantics associated with their location in a particular directory

- 20110305
  - IMG_001 > 20110305-001
  - IMG_002 > 20110305-002
  - IMG_003 > 20110305-003
  - IMG_004 > 20110305-004
- 20110315
  - IMG_001 > 20110315-001
  - IMG_002 > 20110315-002
  - IMG_003 > 20110315-003
  - IMG_004 > 20110315-004
- 20110321
  - IMG_001 > 20110321-001
  - IMG_002 > 20110321-002
  - IMG_003 > 20110321-003
  - IMG_004 > 20110321-004

Multiple copies of the same files with different names and no indication which is the final version (and unknown file formats)

AIETAPES.CPY
AITAPES.DOC
AITAPES.txt
Backup of mancass.wbk
Backup of Manus Cassettes Work.wbk
Backup of Manus Cassettes.wbk
Backup of Manus Recordings Copy.wbk
Backup of Manus Recordings.wbk
Backup of Manus Tapes All.wbk
Copy of Manus Cassettes.doc
DIRN-MAN.DAT
FilmBoxes.doc
KM1.DOC
MANCAS S 94.DOC
MANCASS feb 17 1994.DOC
MANCASS FEB 94.DOC
mancass.doc
MANCASS.txt
MANJOUR.DOC
MANTAPES jan 14 1998.TXT
MANTAPES.DOC
MANTAPES.TXT
Manus Cassettes.doc
Manus Tapes All - tara convert.docx
Manus Tapes All.doc
Manus Tapes Copy.doc
Manus Tapes Remnants.doc
Manus Tapes.doc
NGAITAPE.CPY
NGTAPES.DOC
NGTAPES.txt
NGTPS.DOC

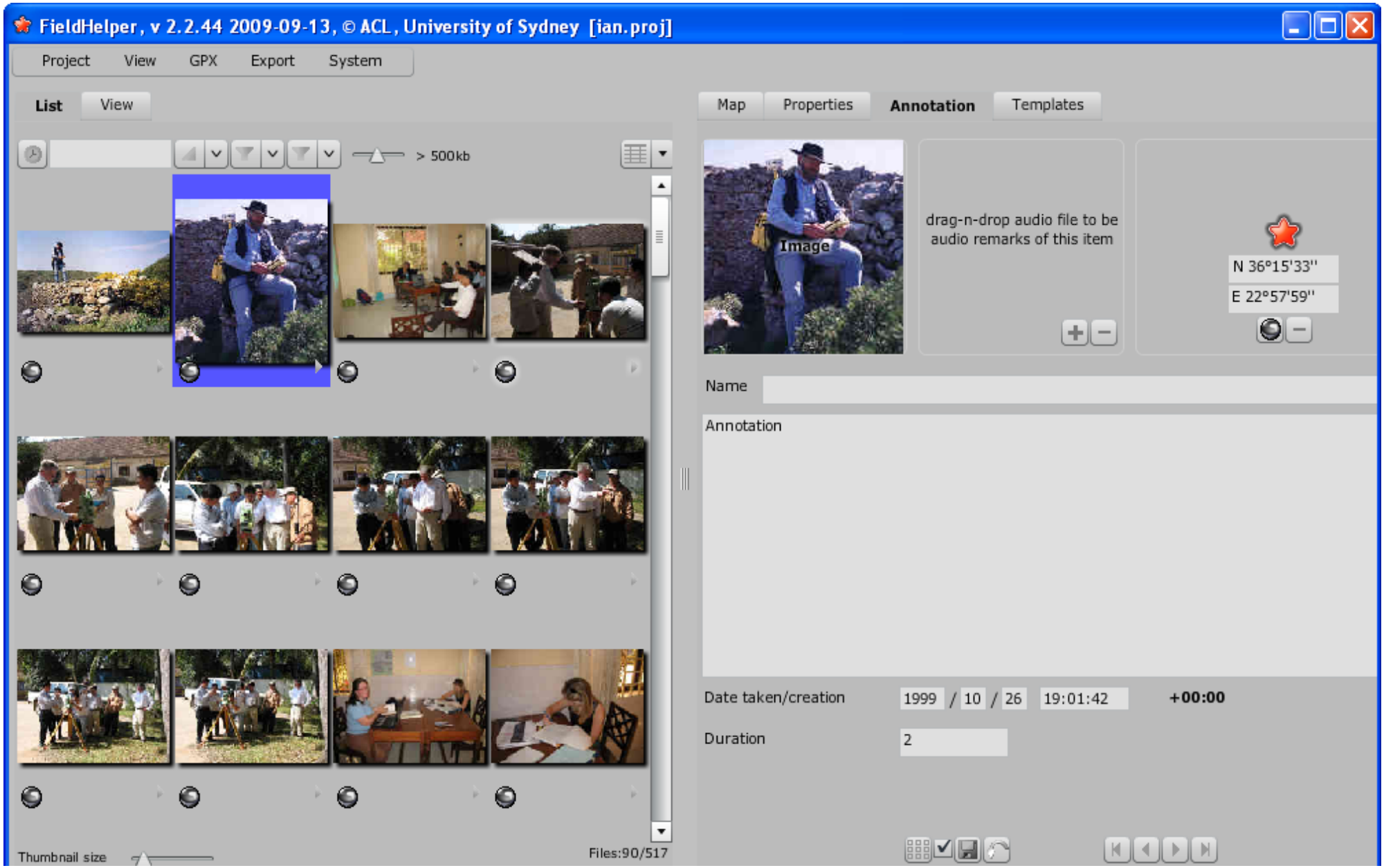
Metadata + recording

-> discoverable

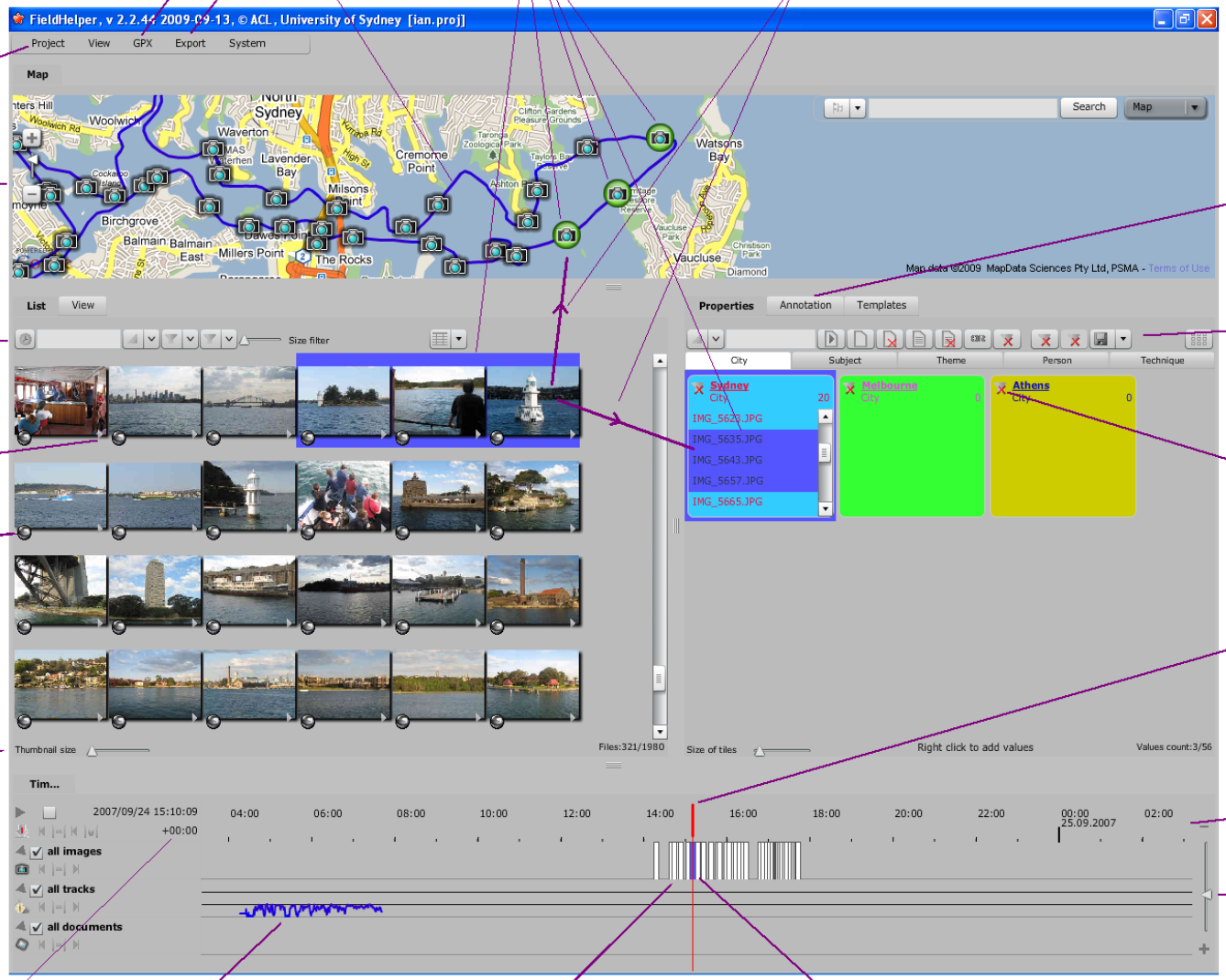
-> annotatable (by speakers and by linguists)

Enriched collection

How to get metadata out of linguists?


FieldHelper, v 2.2.44 2009-09-13, © ACL, University of Sydney [ian.proj]
Project
View
GPX
Export
System
List
View
> 500kb
Thumbnail size
Files:90/517
Map
Properties
Annotation
Templates
Image
drag-n-drop audio file to be audio remarks of this item
N 36°15'33"
E 22°57'59"
Name
Annotation
Date taken/creation
1999 / 10 / 26
19:01:42
+00:00
Duration
2

FieldHelper Vsn 2 Basics



Match files with GPS tracks to set locations

Export XML to METS, KML and EXIF file headers

GPS track

Selected files highlighted

"Drag and tag"
(drag and drop files to set values)
(drag to map to set location)
(drag on map to change location)

Open projects, scan files, import metadata from headers, import/export definitions

All panels can be detached and repositioned

Annotate (text and audio)

Filter and sorting controls

Controls to manage metadata panels/values

View and edit metadata in XML or form view

Filter by this value

Flags files with coordinates

'Play' timeline

Control over thumbnails and labelling

Draggable timeline

Separate or combined time bars for each device

Zoom timeline scale

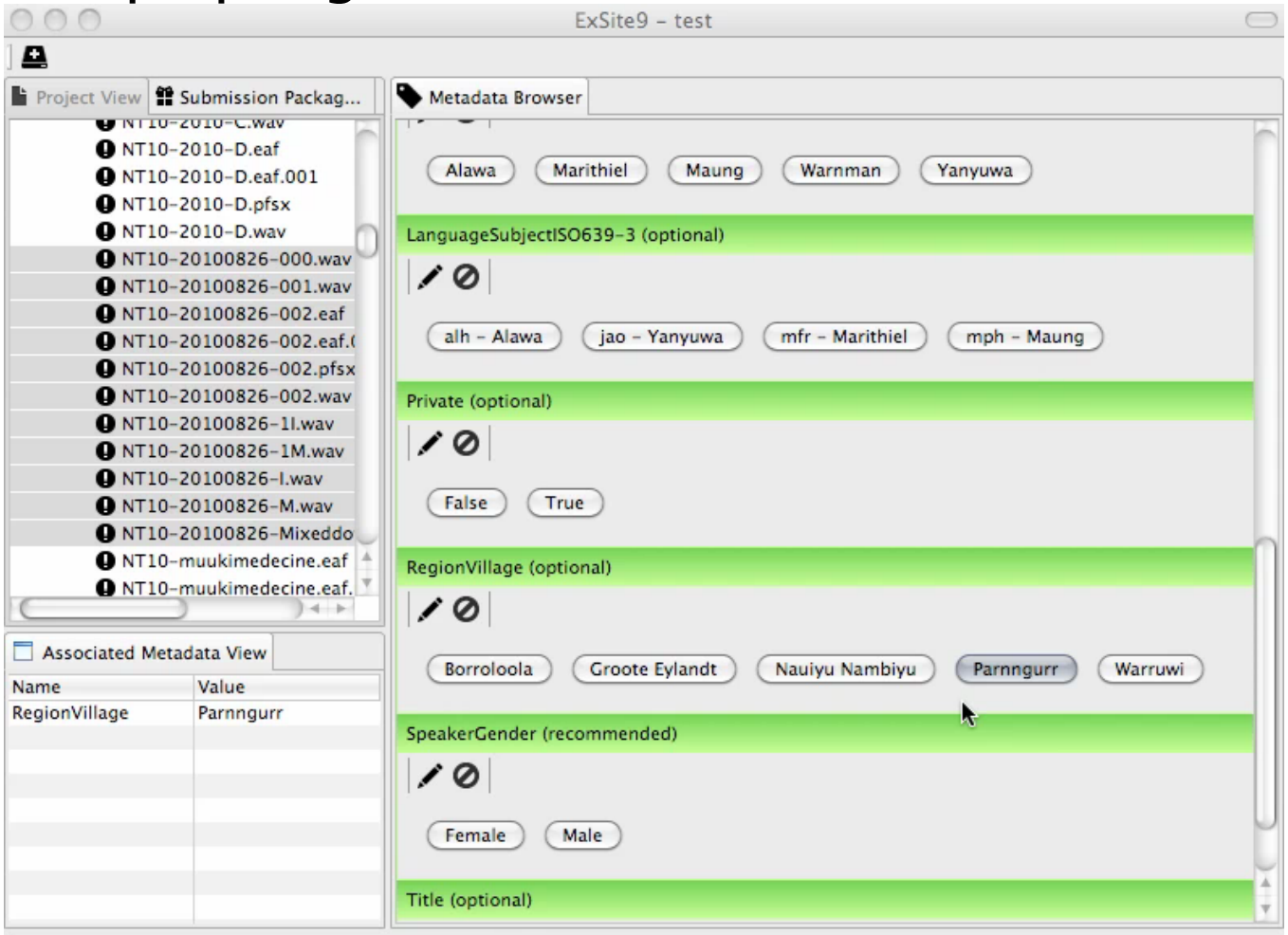
GPS track

Images corresponding with GPS track

Files selectable on timeline

Click on time to adjust time zone and/or offset between GPS and camera

Laptop organisation of collections – ExSite9



http://youtu.be/4_NmUXrWZIE

http://www.intersect.org.au/exsite9

	SayMore	Ramp	Arbil	ExSite9
Freely available	√	[seems to be only for SIL members]	√	√
Output to a standard XML file	?	?	√ IMDI / CMDI	√ XML
Bundle files	?	?	√	√
Project metadata (contextual metadata over a whole set)	?	?		√
Open Source	.net ?	.net ?	√	√
Cross platform	-	?	√	√
Connect to external devices to copy files	√	?	√	-
Creation of Submission Information Package (SIP)	√	√	Via Lamus	√
Automatic harvest of file details, e.g., EXIF from images	√	?	√	√ [few]
Workflow management	√	?	-	-
File format conversion	√	?	-	-
Other	Transcribe, BOLD, subtitling			

XML schema — ExSite9

<metadata_category inextensible="false" name="Title" type="Free Text" use="required">
<description>This title should be a sensible title, unique to each item, briefly summarising the contents of the item, for example "Ilocano songs recorded in Burgos, Ilocos Sur, Philippines, 17 April 1993"</description>
</metadata_category>
<metadata_category inextensible="false" name="Description" type="Free Text" use="required">
<description>Description may include but is not limited to: an abstract, table of contents, reference to a graphical representation of content, or a free-text summary account of the content. [DCMT] Description may also offer an annotation, or a qualitative or evaluative comment about the resource, such as a statement about suitability for a particular application or context.</description>
</metadata_category>
<metadata_category inextensible="true" name="Private" type="Controlled Vocabulary" use="required">
<description>Choose either “false”, meaning that the metadata for the item should be publicly available, or “true”, meaning that the metadata for the item should be hidden (perhaps because you plan to check it and edit it later).</description>
<value>True</value>
<value>False</value>
</metadata_category>
Format

INTERNET ARCHIVE WayBackMachine
http://lacito.vjf.cnrs.fr/archivage/main.htm Go
14 captures
29 Apr 03 - 5 Jan 05
MAR APR JUL
2002 29 2003 2004
Close
Help ?

The LACITO Archive

An archive of natural speech in "rare" languages

The LACITO Archive provides free access to documents of connected, spontaneous speech, mostly in "rare" or endangered languages, recorded in their cultural context and transcribed in consultation with native speakers. Its goal is to contribute to the documentation and study of a precious human heritage: the world's languages. At present, the archive contains some 79 documents in 16 languages (see index).

A sound archive with synchronized transcriptions

For linguistic science, language is first and foremost *spoken* language. The medium of spoken language is sound. The LACITO archive gives access to original recordings simultaneously with transcriptions and translations, as a guarantee of authenticity and as a resource for further research.

A structured, open architecture

The archived data is structured in accordance with the latest data-processing standards, in an open format, and may be downloaded for research purposes. The software used to prepare and disseminate it is open-source.

Copyright

The Archive is an ongoing project of the research group "Oral Tradition: languages and civilizations" of the French National Center for Scientific Research.

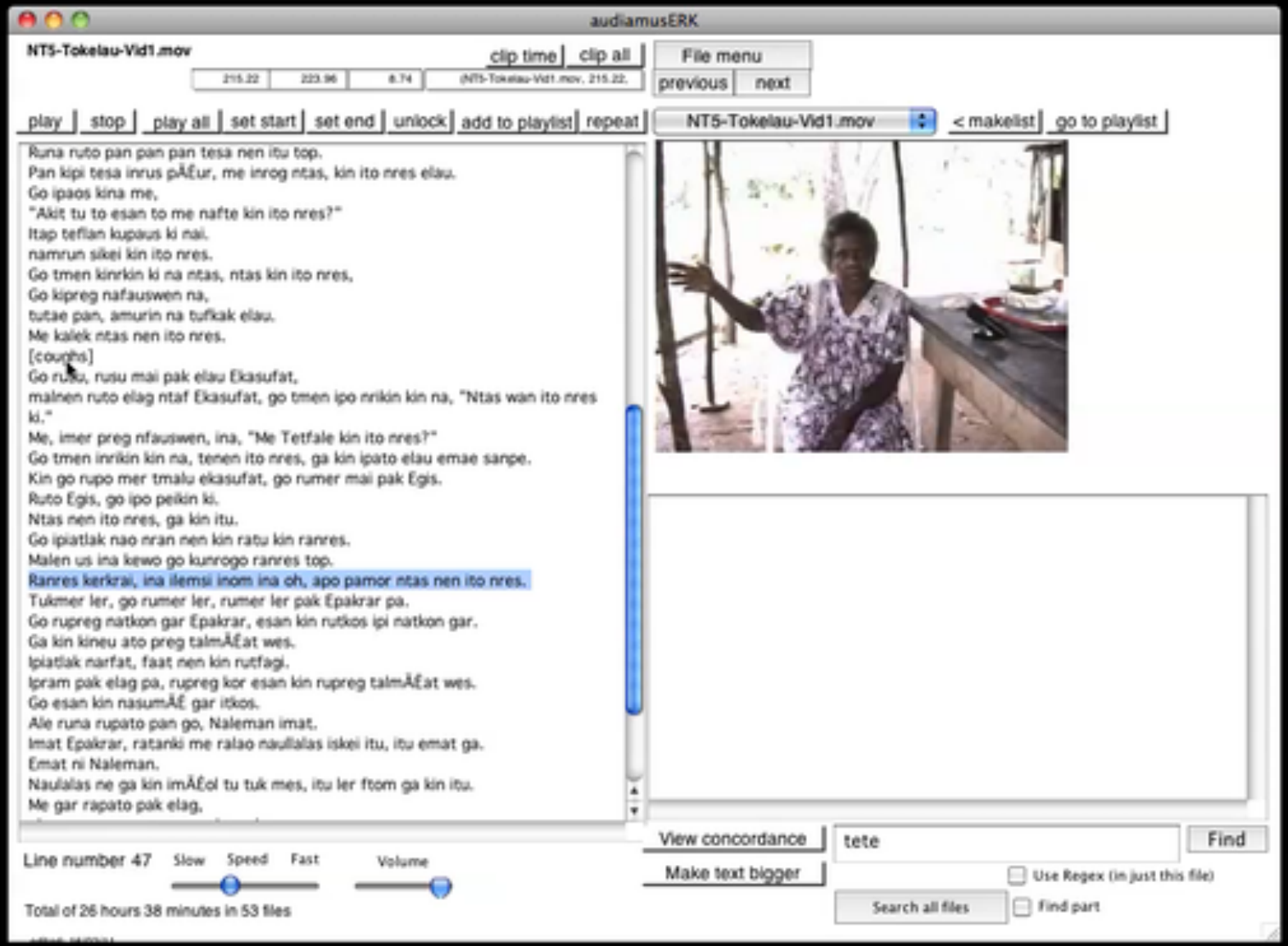
Copyright of Archive documents is held by the authors. Non-commercial use for research purposes is authorized; please contact the project for any other use.

http://web.archive.org/web/20030323205742/http://lacito.vjf.cnrs.fr/archivage/
April 2003

PARADISEC

Pacific and Regional Archive for Digital Sources in Endangered Cultures

Safeguarding research in Australia's region



EOPAS – text and media corpus presentation

- Upload text in interlinear format
- Upload media
- Present text and media online
- Citable text and media
- www.eopas.org

Planned

- Allow media in PARADISEC to be called
- Provide online annotation

p1

Ruto los ena, ruto pan los Ermag me gar,

Ruto		los	ena,		ruto		pan	los	Ermag	me	gar,
ru=	to	los	e-	na	ru=	to	pan	los	Ermag	me	gar
3P.RS=	PROG	wash	LOC-	hesit	3P.RS=	PROG	go	wash	Erromango	and	3P.

The washed there, they would go and wash at Erromango, but they,


```
<idGroup>
  <id>098:001</id>
  <aud>NT1-20003-B 1173.403
    1180.39</aud>
      <txGroup>
        <tx>Ruto</tx>
        <mr>ru=</mr>
        <mg>3P.RS=</mg>
        <mr>to</mr>
        <mg>PROG</mg>
      </txGroup>
      <txGroup>
        <tx>los</tx>
        <mr>los</mr>
        <mg>wash</mg>
      </txGroup>
      <txGroup>
        <tx>ena,</tx>
        <mr>e-</mr>
        <mg>LOC-</mg>
        <mr>na</mr>
        <mg>hesit</mg>
      </txGroup>
      <txGroup>
        <tx>ruto</tx>
        <mr>ru=</mr>
        <mg>3P.RS=</mg>
        <mr>to</mr>
        <mg>PROG</mg>
      </txGroup>
      <txGroup>
        <tx>pan</tx>
        <mr>pan</mr>
        <mg>go</mg>
      </txGroup>
      <txGroup>
        <tx>los</tx>
        <mr>los</mr>
        <mg>wash</mg>
      </txGroup>
      <txGroup>
        <tx>Ermag</tx>
        <mr>Ermag</mr>
        <mg>Erromango</mg>
      </txGroup>
      <txGroup>
        <tx>me</tx>
        <mr>me</mr>
        <mg>and</mg>
      </txGroup>
      <txGroup>
        <tx>gar,</tx>
        <mr>gar</mr>
        <mg>3P.</mg>
      </txGroup>
      <fg>The washed there, they
    would go and wash at Erromango,
    but they,</fg>
  </idGroup>
```

Pacific and Regional Archive for Digital Sources in Endangered Cultures

Safeguarding research in Australia's region

eopas

EthnoER online presentation and annotation system

Nick Thieberger
Logout

Home | Browse Transcripts | Browse Media | Upload Transcript | Upload Media | Contact Us | Help

Making thatch ; Efate, South (erk) ; Vanuatu

Set Displayed Tra

Current Segment URL: http://eopas.rnld.unimelb.edu.au/transcripts/55

Media Player

Making thatch , 2011-02-15 00:00:00 UTC

Making thatch , 2000-03-31 00:00:00 UTC

Participants	**speaker** : Toukelau Takau
Language	erk
Country	VU
Date	2000-03-31 00:00:00 UTC
Depositor	Nick Thieberger
Description	A short description of making thatch out of sago plam (rowat).
Original Format	Toolbox
Download	EOPAS Original (Toolbox)

Edit transcript metadata

Unlink Media Item
Delete transcript

p1

Malen amur rowat,

Malen	a			na	katur		rowat,
malnen	a		-i	-n	na	ka=	tur	rowat
as	1		t -TS	-3S.O	COMP	1S.IRS=	sew	sago_palm

When I wan atch,

go apo pan kafan slat rowat.

go	apo	pan	slat	rowat,	kafan		slat	rowat.
				t	ka=	fan	slat	rowat
				_palm	1S.IRS=	go:IR	take	sago_palm

I'll get rowat.

rowat inom.

			pei	ŋhasel	rowat,	aŋhasel		rowat	inom.
			pei	ŋhasel	rowat	a=	ŋhasel	rowat	i=	nom
		IR	first	derib	sago_palm	1S.RS=	derib	sago_palm	3S.RS=	finish

pare it until it's done.

		lkin				itu.	
			pel	-ki	n	i=	tu
		RS=	bend	-TR	3S.O	3S.RS=	stand

ai. Apregptaki, taiptakiø

								lop.	Apan		tai	lop,	mai.	Apregptaki,				
								lop	a=	pan	tai	lop	mai	a=	preg	pta	-ki	
p5	but	1S.RS=	PSP	again	go	cut	bamboo	1S.RS=	go	cut	bamboo	come	1S.RS=	make	make.good	-TR		

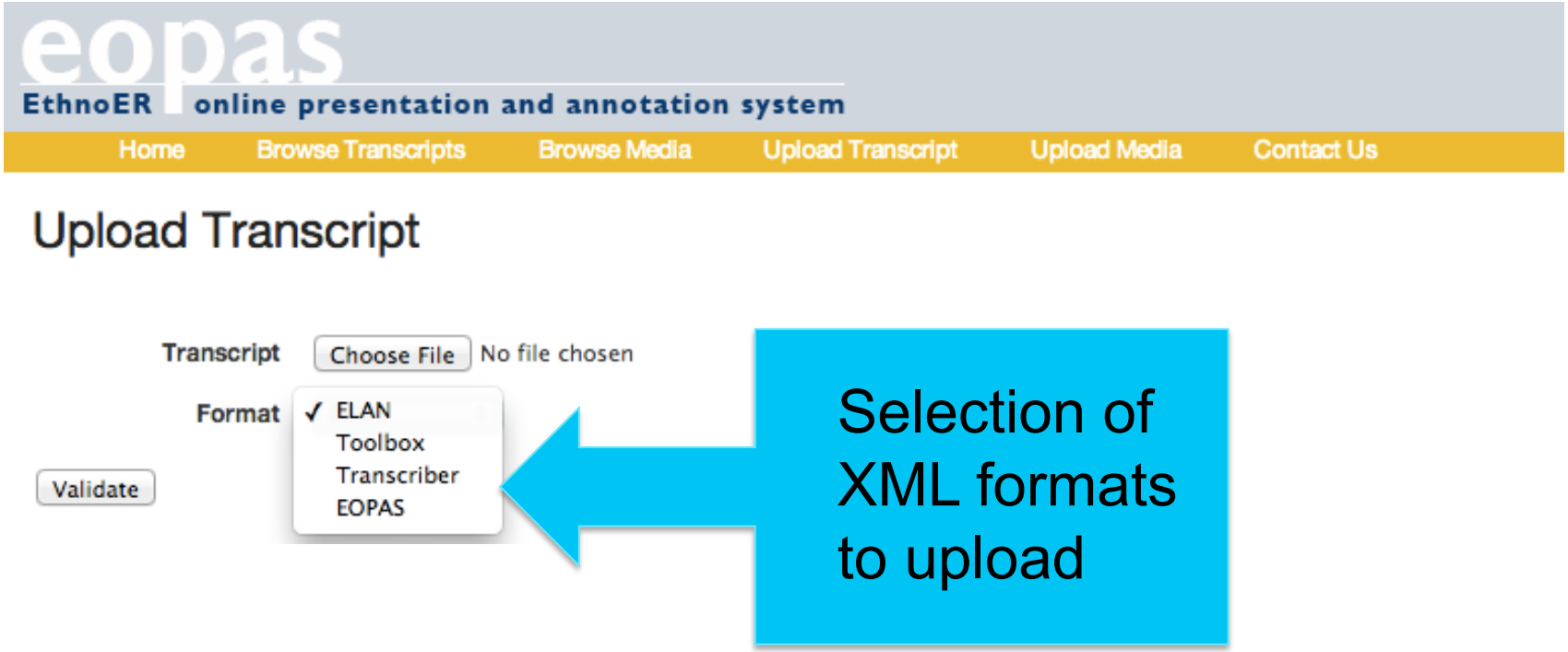
taiptakiø

tai	pta	-ki	-ø
cut	make.good	-TR	-3S.O

I get it ready, cut it to make it ready. Then I'll cut bamboo. I cut bamboo and come.

lleg nen kin kafo tur rowat, apreg, atonkin rupitkaskei. Lop rupitkaskei.

Import Transcripts and media


eopas
EthnoER online presentation and annotation system
Home
Browse Transcripts
Browse Media
Upload Transcript
Upload Media
Contact Us
Upload Transcript
Transcript
Choose File
No file chosen
Format
✓ ELAN
Toolbox
Transcriber
EOPAS
Validate
Selection of XML formats to upload

eopas
EthnoER online presentation and annotation system

Home | Browse Transcripts | Browse Media | Upload Transcript | Upload Media | Contact Us

Upload Transcript

Transcript Choose File No fil

Format Toolbox

Validate

Validation of
XML on upload

eopas
EthnoER online presentation and annotation system

Home | Browse Transcripts | Browse Media | Upload Transcript | Upload Media | Contact Us

Uplaod Media Item

Format * ✓ audio video

Media File * e No file chosen

Title *

Description *

Recorded on * 7 May 2013

Copyright Holder *

License * CC-AU-BY-SA CC Attribution Share-alike

Make item private?

Create

* required attributes

Media upload and transcode

Date	Depositor	Title	Description	Language	Country	Media	Action
1966-01-31	Nick Thieberger	Abiamio Nakikai 57 phrases	A story recorded by John Harris in Kiwai, from PNG in the mid 1960s.	kiw	PG	Kiwai recording	Show Edit Delete
2000-04-07	Nick Thieberger	Angels and Erromango 90 phrases	A story about a woman who flies to Efate and whose wings are stolen by a man who wants her for his wife. Her children eventually find her wings and return them to her and she flies home.	erk	VU	20003b	Show Edit Delete
1998-10-03	Nick Thieberger	Asaraf 14 phrases	Asaraf was a giant who could walk from Efate to Erromango and who pushed Erromango away from Efate.	erk	VU	98007A	Show Edit Delete
2011-04-11	Nick Thieberger	CP Drowning 58 phrases		mwf	AU	9-08-CP-drowningMurrinhPatha	Show Edit Delete
1998-10-20	Nick Thieberger	Daisy Wayane, A life story 54 phrases	Daisy Wayane talking about the various jobs she has had in her life.	erk	VU	98010-B	Show Edit Delete
1998-09-29	Nick Thieberger	Darkness, light and Christianity 4 phrases	A brief description of the coming of Christianity	erk	VU	98003A	Show Edit Delete
1996-03-07	Nick Thieberger	English Police 30 phrases	Kalsarap talking about his time with the English Police in the 1930s. He is reading from his notes, and made some slight edits to the transcript.	erk	VU	NT1-004-B	Show Edit Delete
2011-04-12	Nick Thieberger	Harry Laoto's story 21 phrases		erk	VU	98010-B	Show Edit Delete
1998-09-29	Nick Thieberger	How girls can behave 22 phrases	A description of how girls used to be expected to behave.	erk	VU	98003b	Show Edit Delete

http://www.eopas.org/transcripts?language_code=erk

Date	Depositor	Title	Description	Language	Country	Media	Action
2000-04-07	Nick Thieberger	Angels and Erromango 90 phrases	A story about a woman who flies to Efate and whose wings are stolen by a man who wants her for his wife. Her children eventually find her wings and return them to her and she flies home.	erk	VU	20003b	Show Edit Delete
1998-10-03	Nick Thieberger	Asaraf 14 phrases	Asaraf was a giant who could walk from Efate to Erromango and who pushed Erromango away from Efate.	erk	VU	98007A	Show Edit Delete
1998-10-20	Nick Thieberger	Daisy Wayane, A life story 54 phrases	Daisy Wayane talking about the various jobs she has had in her life.	erk	VU	98010-B	Show Edit Delete
1998-09-29	Nick Thieberger	Darkness, light and Christianity 4 phrases	A brief description of the coming of Christianity	erk	VU	98003A	Show Edit Delete
1996-03-07	Nick Thieberger	English Police 30 phrases	Kalsarap talking about his time with the English Police in the 1930s. He is reading from his notes, and made some slight edits to the transcript.	erk	VU	NT1-004-B	Show Edit Delete
2011-04-12	Nick Thieberger	Harry Laoto's story 21 phrases		erk	VU	98010-B	Show Edit Delete
1998-09-29	Nick Thieberger	How girls can behave 22 phrases	A description of how girls used to be expected to behave.	erk	VU	98003b	Show Edit Delete
1998-09-30	Nick Thieberger	Independence 26 phrases	Discussion of the time of Independence in Vanuatu.	erk	VU	98004a	Show Edit Delete
1998-10-23	Nick Thieberger	Katapel and Liportani 25 phrases		erk	VU	98009A	Show Edit Delete

Pacific and Regional Archive for Digital Sources in Endangered Cultures

Safeguarding research in Australia's region

eopas

EthnoER online presentation and annotation system

Nick Thieberger
Logout

Home | Browse Transcripts | Browse Media | Upload Transcript | Upload Media | Contact Us | Help

Making thatch ; Efate, South (erk) ; Vanuatu

Set Displayed Tracks

Current Segment URL: http://eopas.rnld.unimelb.edu.au/transcripts/55

Media Player

Making thatch , 2011-02-15 00:00:00 UTC

Making thatch , 2000-03-31 00:00:00 UTC

Participants	**speaker** : Toukelau Takau
Language	erk
Country	VU
Date	2000-03-31 00:00:00 UTC
Depositor	Nick Thieberger
Description	A short description of making thatch out of sago plam (rowat).
Original Format	Toolbox
Download	EOPAS Original (Toolbox)

Edit transcript metadata

Unlink Media Item
Delete transcript

p1 Malen amurin na katur rowat,

Malen	amurin		na	katur	rowat.
malnen	a=	n			
as	1S.RS=	v			

When I want to sew

and 1S.RS= PSP ... palm

and I will get sago p

p3 Aler mai, kafo pei ñ

Aler		mai,	kafo		pei	ñasel	rowat,	amasel		rowat	inom.	
a=	ler	mai	ka=	fo	pei	ñasel	rowat	a=	ñasel	rowat	i=	nom
1S.RS=	return	come	1S.IRS=	PSP:IR	first	derib	sago_palm	1S.RS=	derib	sago_palm	3S.RS=	finish

I come back, I'll prepare the rowat, prepare it until it's done.

p4 Amer þelkiø, aþelkin itu.

Amer		þel
a=	mer	þel
1S.RS=	again	be

p5

me	a=	po					mai.	Apregptaki,			
but	1S.RS=	PSP				o	come	mai	a=	preg	pta	-ki
								1S.RS=	make	make.good	-TR

taiptakiø

tai	pta	-ki	-ø
cut	make.good	-TR	-3S.O

I get it ready, cut it to make it ready. Then I'll cut bamboo. I cut bamboo and come.

Ileg nen kin kafo tur rowat, apreg, atonkin rupitkaskei. Lop rupitkaskei.

Playable media

Metadata

eopas
EthnoER online presentation and annotation system

Nick Thieberger
Logout

Home | Browse Transcripts | Browse Media | Upload Transcript | Upload Media | Contact Us | Help

Set Displayed Tracks

Current Segment URL: http://eopas.rnld.unimelb.edu.au/transcripts/55#!/p2/w3/m1

Concordance

go a= po	pan	slat rowat ka= fan s...
me a= po mer	pan	tai lop a=
me a= po mer	pan	tai lop a=
me namet -e -n i=	pan	pato elag Elakmamiel
pan tkal naur sees	pan	
ra=	pan	panpan ra= pato ŋhel...
...ken kusu i= siwer i=	pan	pato siwer ur nskau ...
... ka= mer pak talŋhat	pan	ka= fan lel wak ɸur...
...at menal i= na go i=	pan	na i= tkal nagis pei...
i= mai pak euut me i=	pan	soksok -i i= na me a...
ale wak i= na i=	pan	me i= ta ler mau
ra=	pan	panpan ra= mer kop n...
malfane kusu i=	pan	totan
...ki= tok takinrog kin	pan	ki= nrog -o -ø na i=...
...os i= na i= pa i= ta	pan	nrik tem -e -n go ra...
...os i= na i= pa i= ta	pan	nrik tem -e -n go ra...
...al ki= pe to ŋhit i=	pan	na i= pak nagis i= s...
...e ŋhit i= na Watetjo	pan	na i= pak nagis i= s...
ra= en	pan	i= pak nagis i= skei...
...u= slat -i -ø kin po	pan	tan -ki -ø
me malnen kin i=	pan	i=
me malnen kin i=	pan	i=
...au teetwei ga i= to	pan	i= to pak Ermag
...ok pak na nmal nawen	pan	na i= le mar i= lek ...
i= to esan to	pan	pan i= mur ke= fak E...

p2
go apo pan slat rowat, kafan slat rowat.

go	apo		pan	slat	rowat,	kafan
go	a=	po	pan			
and	1S.RS=	PSP	go			

and I will get sago palm leaves (rowat), I'll get rowat.

p3
Aler mai, kafo pei ŋhasel rowat, aŋhasel rowat inom.

Aler	mai,	kafo		pei	ŋhasel	rowat,	aŋhasel		rowat	inom.	
a=	ler	mai	ka=	fo	pei	ŋhasel	rowat	a=	ŋhasel	rowat	i=	nom
1S.RS=	return	come	1S.IRS=	PSP:IR	first	derib	sago_palm	1S.RS=	derib	sago_palm	3S.RS=	finish

I come back, I'll prepare the rowat, prepare it until it's done.

p4
Amer ɸelkiø, aɸelkin itu.

Amer		ɸelkiø,	aɸelkin	itu
a=	mer	ɸel		
1S.RS=	again	bend		

Then I'll fold it, it is folde

M po mer pan tai lop.

| | | |

eopas
EthnoER online presentation and annotation system

Nick Thieberger
Logout

Home | Browse Transcripts | Browse Media | Upload Transcript | Upload Media | Contact Us | Help

Concordance

Making thatch ; Efate, South (erk) ; Vanuatu

Media Player

Making thatch , 2011-02-15 00:00:00 UTC

Making thatch , 2000-03-31 00:00:00 UTC

Participants	**speaker** : Toukelau Takau
Language	erk
Country	VU
Date	2000-03-31 00:00:00 UTC
Depositor	Nick Thieberger
Description	A short description of making thatch out of sago plam (rowat).
Original Format	Toolbox
Download	EOPAS Original (Toolbox)

Edit transcript metadata

Unlink Media Item
Delete transcript

Set Displayed Tracks

- [x] Show transcription
- [ ] Show morphemes
- [x] Show translation

Current Segment URL: http://eopas.rnld.unimelb.edu.au/transcripts/55#!/p2/w3/m1

sel rowat, aṁasel rowat inom.
are the rowat, prepare it until it's done.

p4
Amer ..., ap̃elkin itu.
Th... it is folded.

p5
Me... an tai lop. Apan tai lop, mai. Apregptaki, taiptakiø
I g... ut it to make it ready. Then I'll cut bamboo. I cut bamboo and come.

..., atonkin rupitkaskei. Lop rupitkaskei.
...h, I measure it so that they are all equal.

...ampe neu rufitkaskei. Inom.
...y thatch all the same. That's all.

...ieces.

...en rusosoø ki 'pin'.
...ns'.

p10
Amas pin ina inom.
I clean the pins until it is done.

p11
Go apo tur rowat.
And then I'll sew the thatch.

p12
Atrus tefla inom go malfane atae na, atae tur
I'll sew it like this, and now I can sew

p13
fifti, atae turtur ralimilim.
fifty, I can sew fifty.

p14
ko atae tur tifli iskei.
or I can sew one hundred.

Ability to turn off morphemic view

eopas
EthnoER online presentation and annotation system

Nick Thieberger
Logout

Home | Browse Transcripts | Browse Media | Upload Transcript | Upload Media | Contact Us | Help

Set Displayed Tracks

Current Segment URL: http://eopas.rnld.unimelb.edu.au/transcripts/55#!/p2/w5

Concordance

Making thatch ; Efate, South (erk) ; Vanuatu

Media Player

Making thatch , 2011-02-15 00:00:00 UTC

Making thatch , 2000-03-31 00:00:00 UTC

Participants	**speaker** : Toukelau Takau
Language	erk
Country	VU
Date	2000-03-31 00:00:00 UTC
Depositor	Nick Thieberger
Description	A short description of making thatch out of sago plam (rowat).
Original Format	Toolbox
Download	EOPAS Original (Toolbox)

Edit transcript metadata

Unlink Media Item
Delete transcript

p2

go apo pan slat rowat, kafan slat rowat.

go	apo		pan	slat	rowat,	kafan		slat	rowat.
go	a=	po	pan	slat	rowat	ka=	fan	slat	rowat
and	1S.RS=	PSP	go	take	sago_palm	1S.IRS=	go:IR	take	sago_palm

and I will get sago palm leaves (rowat), I'll get rowat.

p3

Aler mai, kafo pei ṁasel rowat, aṁasel rowat inom.

Aler		mai,	kafo		pei	ṁasel	rowat,	aṁasel		rowat
a=	ler	mai	ka=	fo	pei	ṁasel	rowat	a=	ṁasel	rowat
1S.RS=	return	come	1S.IRS=	PSP:IR	first	derib	sago_palm	1S.RS=	derib	sago_palm

I come back, I'll prepare the rowat, prepare it until it's done.

p4

Amer p̃elkiø, ap̃elkin itu.

Amer		p̃elkiø,			ap̃elkin				itu.	
a=	mer	p̃el	-ki	-ø	a=	p̃el	-ki	n	i=	tu
1S.RS=	again	bend	-TR	-3S.O	1S.RS=	bend	-TR	3S.O	3S.RS=	stand

Then I'll fold it, it is folded.

p5

Me apo mer pan tai lop. Apan tai lop, mai. Apregptaki, taiptakiø

Me	apo		mer	pan	tai	lop.	Apan		tai	lop,	mai.	Apregptaki,			
me	a=	po	mer	pan	tai	lop	a=	pan	tai	lop	mai	a=	preg	pta	-ki
but	1S.RS=	PSP	again	go	cut	bamboo	1S.RS=	go	cut	bamboo	come	1S.RS=	make	make.good	-TR

taiptakiø			
tai	pta	-ki	-ø
cut	make.good	-TR	-3S.O

I get it ready, cut it to make it ready. Then I'll cut bamboo. I cut bamboo and come.

p6

Ileg nen kin kafo tur rowat, apreg, atonkin rupitkaskei. Lop rupitkaskei.

Ileg		nen	kin	kafo		tur	rowat,	apreg,		atonkin				rupitkaskei.		Lop
i=	leg	nen	kin	ka=	fo	tur	rowat	a=	preg	a=	ton	-ki	-n	ru=	pitkaskei	lop
3S.RS=	straight	that	REL	1S.IRS=	PSP:IR	sew	sago_palm	1S.RS=	make	1sgRS=	measure	-TR	-3sgO	3P.RS=	equal	bamboo

rupitkaskei.	
ru=	pitkaskei
3P.RS=	equal

It is right so I can make the thatch, I measure it so that they are all equal.

Reference to morpheme-level

eopas
EthnoER online presentation and annotation system

Nick Thieberger
Logout

Home | Browse Transcripts | Browse Media | Upload Transcript | Upload Media | Contact Us | Help

Set Displayed Tracks

Current Segment URL: http://eopas.rnld.unimelb.edu.au/transcripts/55#t=76.253,87.067

Concordance

Making thatch ; Efate, South (erk) ; Vanuatu

Media Player

Making thatch , 2011-02-15 00:00:00 UTC

Making thatch , 2000-03-31 00:00:00 UTC

Participants	**speaker** : Toukelau Takau
Language	erk
Country	VU
Date	2000-03-31 00:00:00 UTC
Depositor	Nick Thieberger
Description	A short description of making thatch out of sago plam (rowat).
Original Format	Toolbox
Download	EOPAS Original (Toolbox)

Edit transcript metadata

Unlink Media Item
Delete transcript

p10
Amas pin ina inom.

Amas		pin	ina		inom.	
a=	mas	pin	i=	na	i=	nom
1S.RS=	clean	pin	3S.RS=	INCH	3S.RS=	finish

I clean the pins until it is done.

p11
Go apo tur rowat.

Go	apo		tur	rowat.
go	a=	po	tur	rowat
and	1S.RS=	PSP	sew	sago_palm

And then I'll sew the thatch.

p12
Atrus tefla inom go malfane atae na, atae tur

Atrus			tefla	inom		go	malfane	atae		na,	atae
a=	tur	-us	tefla	i=	nom	go	malfane	a=	tae	na	a=
1S.RS=	sew	-3S.O	similar	3S.RS=	finish	and	now	1S.RS=	know	COMP	1S.RS

I'll sew it like this, and now I can sew

p13
fifti, atae turtur ralimilim.

fifti,	atae		turtur		ralimilim.
fifti	a=	tae	tur	tur	ralimilim
fifty:BI	1S.RS=	know	sew	sew	fifty

fifty, I can sew fifty.

p14
ko atae tur tifli iskei.

ko	atae		tur	tifli	iskei.	
ko	a=	tae	tur	tifli	i=	skei
or	1S.RS=	know	sew	hundred	3S.RS=	one

or I can sew one hundred.

p15
Go kafo tae na ralimilim kefo tae ƀakor nanre nasuṁ

Go	kafo		tae	na	ralimilim	kefo		tae	ƀakor	nanre	nasuṁ
go	ka=	fo	tae	na	ralimilim	ke=	fo	tae	ƀakor	nanre	nasuṁ
and	1S.IRS=	PSP:IR	know	COMP	fifty	3S.IRS=	PSP:IR	know	cover	side	house

And I know that fifty can cover a side of a house.

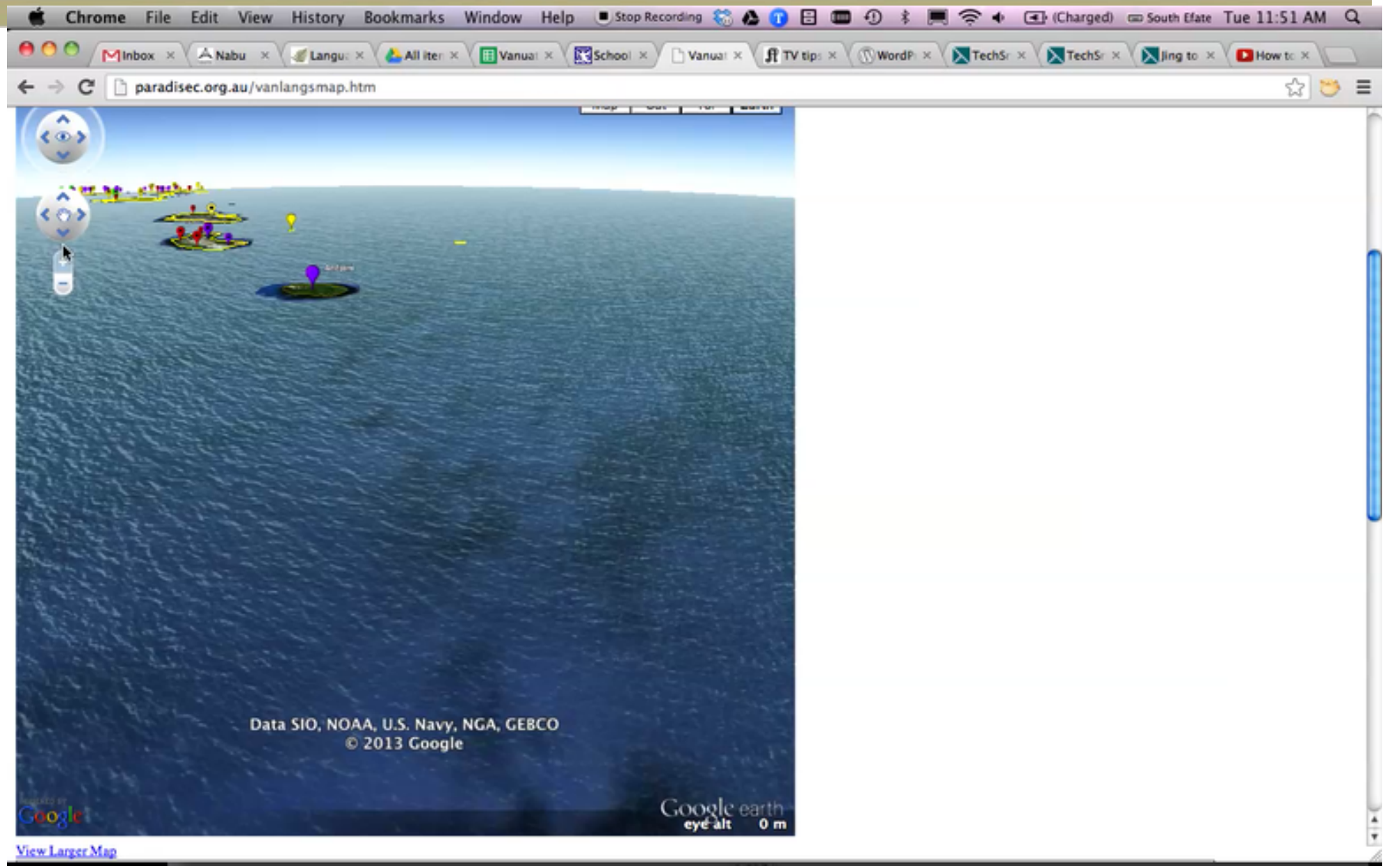
Go ralimilim kip atae ƀakor nanre nasuṁ

Reference to timed chunk

PARADISEC

Pacific and Regional Archive for Digital Sources
in Endangered Cultures

Safeguarding research
in Australia's region



Corpus presentation

- Valuing corpora as scholarly output
- Ranking corpora as equivalent to: monograph; journal article; book chapter; working paper.

blog discussion here:

http://www.paradisec.org.au/blog/2012/11/counting-collections/

- As has been observed for a number of Humanities projects, a lack of tools frustrates the use of existing datasets, as "it is maybe not the critical mass of digitized data that is missing – though it is debatable enough whether the quality of its web availability is appropriate – but rather the tools that would allow humanities scholars to use those digitized materials in any useful way for research." (van Zundert 2012:167)

Conclusions

- Small gaps in fieldwork methods can cause a research project to risk losing its data
- Two gaps identified here
  - Metadata entry
  - Corpus presentation

Conclusions

- Confluence of
  - large datasets of primary linguistic material
  - newly developed services for scholarly outputs

PARADISEC

Pacific and Regional Archive for Digital Sources in Endangered Cultures

Safeguarding research in Australia's region

LANGUAGE DOCUMENTATION & CONSERVATION

LD&C Home | Current Issue | Past Issues | Subscribe | Submissions | About LD&C | Editors | Contact | Search

- New! *LD&C Special Publication No. 2: Fieldwork and Linguistic Analysis in Indigenous Languages of the Americas* is now available here.
- New! Beginning in January 2010, *LD&C* submissions will be published upon acceptance.
- Volume 3, Issue 2 (December 2009) is available here
- Presentations from the 1st International Conference on Language Documentation & Conservation (March 12–14, 2009 in Honolulu, Hawai'i) are available here.

About the journal

Welcome to the homepage of *Language Documentation & Conservation* (*LD&C*), a fully refereed, open-access journal sponsored by the National Foreign Language Resource Center and published exclusively in electronic form by the University of Hawai'i Press.

LD&C publishes papers on all topics related to language documentation and conservation, including, but not limited to, the goals of language documentation, data management, fieldwork methods, ethical issues, orthography design, reference grammar design, lexicography, methods of assessing ethnolinguistic vitality, archiving matters, language planning, areal survey reports, short field reports on endangered or underdocumented languages, reports on language maintenance, preservation, and revitalization efforts, plus software, hardware, and book reviews.

Please browse our site to learn more about this timely new journal.

Subscribe to *LD&C* (free)...
Learn more about *LD&C*...

ISSN 1934-5275

NFLRC

sponsored by National Foreign Language Resource Center

HAWAI'I

published by University of Hawai'i Press

licensed under Creative Commons (licensing information)

http://www.nflrc.hawaii.edu/ldc/

rnld

Resource Network for Linguistic Diversity

Home | About us | Advocacy | Background | Events | FAQs | In the news | Mailing list | Resources | Training | Useful links | Contact us | Log in

Home

What we provide

- **Networking**— we facilitate international networking through our active **email discussion list**, **Facebook group**, and links to **blogs and networks**. Archived RNLD discussions are publicly accessible and searchable through the **Linguist List**.
- **Resources**— we disseminate information about **funding opportunities**, **capacity building programs**, **publications**, **films**, **news**, and the methods and technologies needed to **document**, **archive**, **revitalise and maintain** indigenous languages.
- **Training**— we partner with indigenous communities and organisations to build capacity in language documentation and revitalisation. Our flexible training materials and practices ensure that the skills which people gain are used, retained and shared more widely within a community.
- **Train-the-trainer**— we promote autonomy and sustainability through workshop methods which mentor skilled language workers in becoming trainers in linguistic documentation and revitalisation methods within their own communities.
- **Advocacy**— we advocate on issues such as bilingual education, indigenous literacy, language policies and recognition for Indigenous peoples, raise awareness through public **events** on occasions such as International Mother Language Day, and organize informal gatherings for linguists and community language activists.

To read more about RNLD, please choose from the menu items above

http://www.rnld.org

Workshop on digital tools and methods for language documentation

Dates: 3rd-4th December 2013

Venue: University of Melbourne, Australia

Highlights

Keynote Speakers:

Alexandre Arkhipov (Moscow State University) will speak on methods used by his research to build an integrated documentation and analysis system.

Andrea Witt (Head of the TEI-SIG, Institut für Deutsche Sprache, Mannheim) will speak on the Text Encoding Initiative-Special Interest Group and TEI for linguists.

Research, records and responsibility (RRR): Ten years of the Pacific and Regional Archive for Digital Sources in Endangered Cultures (PARADISEC)

Dates: 2nd-3rd December 2013

Venue: University of Melbourne, Australia

A PARADISEC conference

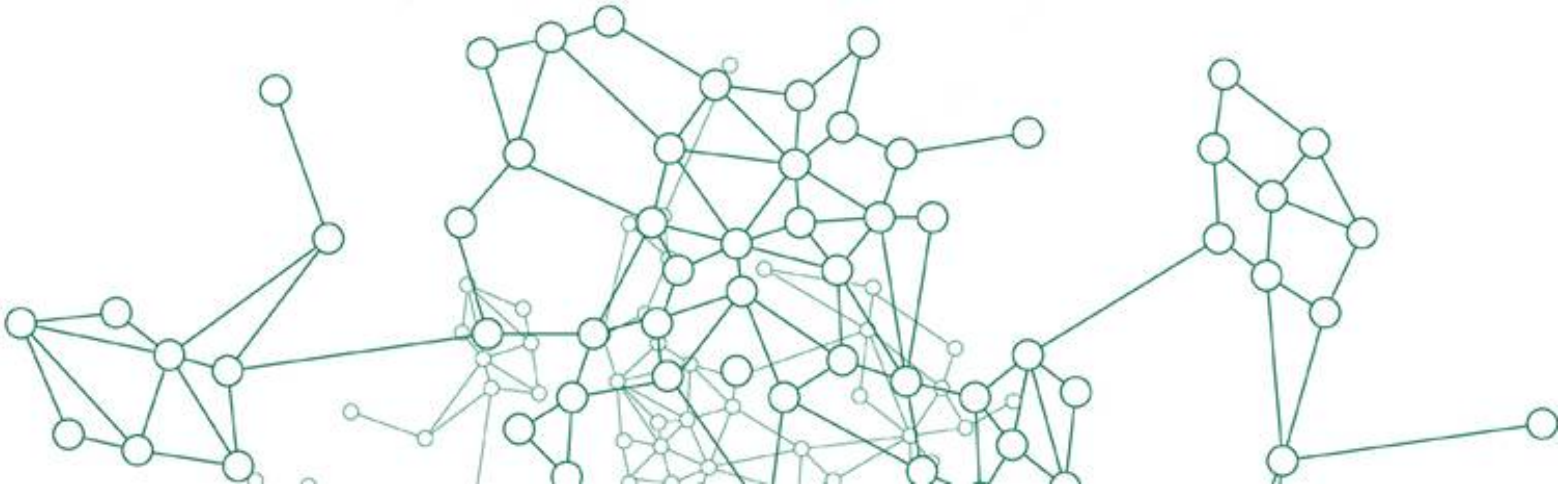
Highlights

Keynote Speaker

Shubha Chaudhuri
Associate Director General (Academic)
Archives and Research Centre for Ethnomusicology
American Institute of Indian Studies
Gurgaon, India

Thanks to the Vanuatu linguists who supplied information for the documentation index of Vanuatu

Australian Research Council – LIEF grants LE110100142, LE0560711, LE0453247
ARC DP0450342, DP0984419

The speaker's attendance at this conference was sponsored by the Alexander von Humboldt Foundation.

http://www.humboldt-foundation.de