

Citing the spoken word: Adventures in language documentation

Nicholas Thieberger

School of Languages and Linguistics
The University of Melbourne

Department of Linguistics
University of Hawai'i at Mānoa

Me, selwan ag kupi eñae, tiawi itraus traus
traus traus, natrauswen ga itaos nlag. Itrausi
pan kaipa. Me komam uta laap kin uto mau, a?
Malen umat, inom.

*But when you are far away the old man can talk
and talk and talk, his story is like the wind.
He tells it and it is gone. But there aren't
many of us left. When we die, it will be
finished.*

†Kalfañun Mailei, 1998, Erakor Village, Efate, Vanuatu
(NT1-98009-A, 1932.36 1942.16)





“Il n’est pas d’avantage possible à la linguistique de représenter fidèlement le parler des langues vivantes même à l’aide de ces transcriptions arbitraires, tout à fait insuffisantes, imaginées faute de mieux, et qu’aucun accord n’est parvenu à uniformiser. ... tout cela dont la linguistique est incapable, le phonographe dès maintenant en offre la possibilité.”

“Nor is it possible for linguistics to faithfully represent the speech of living languages even with the help of its arbitrary and totally insufficient transcriptions, made up for want of better and which no agreement has ever made uniform. ... all that of which linguistics is incapable, the phonograph from now on makes possible.”
(Léon Azoulay 1900:175)

‘ there is the special methodological interest in the frank and full presentation of all available linguistic evidence ’ (Malinowski 1935: 4)

‘If I could, by a good phonographic record, counterfeit the living voice of Tokulubakiki: [...] I should certainly be better able to *translate* the text in the sense of imparting to it its full cultural flavour and significance. Again, if by cinematographic picture I could reproduce the facial expression, the bodily attitude, the significant gestures, this would add another contextual dimension.’ (Malinowski 1935: 26)

“I'm not particularly interested in 'smoothed-over' versions of native culture. I like the stuff in the raw, as felt and dictated by the natives. [...] The genuine, difficult, confusing, primary sources. These must be presented, whatever else is done. [...] There are too many glib monographs, most of which time will show to be highly subjective performances. We need to develop in cultural anthropology that *anxious respect for documentary evidence* that is so familiar to the historian, the classical scholar, the Orientalist. [...] where is the raw evidence? I can't tell whether a given statement is common native knowledge or is merely your interpretation of one man's say-so.”

Letter written by Edward Sapir to Fay Cooper-Cole, 25 April 1938, quoted in Regna Darnell. 1990. *Edward Sapir: Linguist, Anthropologist, Humanist*. Berkeley: University of California Press.

‘There is no reason whatever why, in the future, an exact and physiological study of speech should not use the apparatus of sound films for reproducing fully contextualised utterances.’ (Malinowski 1935: 26)

Malinowski, Bronislaw. 1935. *Coral gardens and their magic*. Vol II
London: G. Allen & Unwin.

Descriptive grammars written since Sapir and Malinowski may have been built on a corpus of recordings,

but that is not evident from the finished product

Perspectives on grammar writing

Payne: 'good linguistic descriptions provide the raw data for much research into the organization of the human mind'

In recent years the 'emphasis has shifted back to the centrality of data.'

Studies in Language, 2006, 30:2 p.236

second person, i.e., the linguist), 171.1.3 (asking addressee to convey prohibition to other person); negative suggestion 20.19.3 (if the sense is approximately 'let's not go there'); negation of morally/legally improper behaviour ('mustn't') 54.1.6, 54.4.1-2, 57.4.4, 67.2.2-3, 68.4.6, 68.5.3, 75.1.3 (three), 75.2.2, 76.1.4, 76.2.2/8, 77.1.3, 77.2.1/2/5, 83.1.2/3 (three), 84.1.5, 85.1.1/4/7, 88.3.1/2/3 (note that most of these exx. are in a sequence of short texts about customary kinship-based behaviour); inability (usually translated as English cannot with present[.] tense) 20.7.3, 20.8.5, 42.6.4 (all involving sense 'cannot see'); negative intention (1st person) or prediction (any person) 17.13.4 ('I won't spear you'), 28.13.4 (1Sg), 62.5.2 (1Sg), 90.2.7 ('He will not [be forced to] go to jail', perhaps in moral/legal sense but if so not focusing on subject's responsibility, rather on unnamed coercing agents), 164.2.4 ('we[ExMDu] might not return' with particle /ari/ 'maybe', §12.11), 164.3.3 (1ExMDu), 165.1.7 (3Pl, prediction).

Past Potential Negative is uncommon in the texts. It occurs in 120.6.1-3, and there is a good ex. from Hughes in TNT 27. Past Potential Negative is used in negating either clause in a counterfactual condition, see discussion §16.7. The ex. in TNT 27 is as follows in my transcription:

(15.xlii) yagi-maga: bamba:--'bi-ni bamba:--'ga:ru:--'
 neg- you-hit-[it-]for-them you left them
 'You shouldn't have hit them, you should have
 left them [alone]' (TNT 27)

Both verbs are in the same form (Past Potential), but /yagi-/ negates only the first, so that the second verb is positive (but still potential) and represents the antithesis of the prohibition. Postposition /-maga:/ is common with such contexts; see §12.21.

There are also some special uses of the "Future Negative" form with /yagi/ which are best dealt with in other sections; for negative purposive clause see 89.2.3 and discussion §16.5, for negated clauses in hypothetical (as opposed to counterfactual) conditionals see ex. in 62.1.2/3 and discussion §16.7, and for gerundial 'while not VERB-ing' in sense 'before VERB' see ex. 47.19.3 and discussion §16.6.

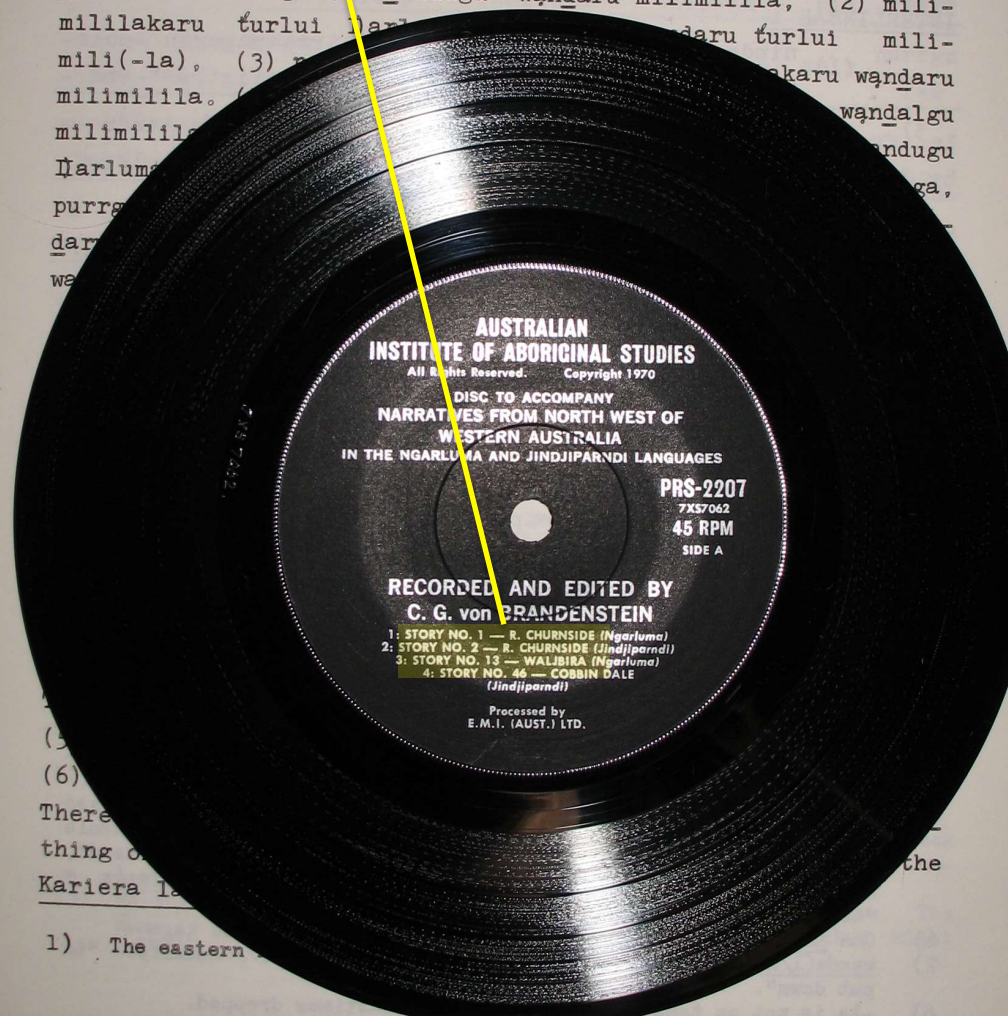
Most textual exx. of /wa:='ri/ have a co-occurring inflected verb, the overall construction being Past (Actual) Negative or Present Negative (Chapter 8). The Past Negatives in NMET are 3.2.5, 5.15.4/5, 7.6.2/3, 14.18.3, 15.11.2, 17.13.5, 19.9.2, 21.10.2, 28.10.4, 41.7.3/4, 47.1.5, 47.4.5/6, 47.5.4 (two), 47.6.2/3, 47.7.3/4 (three), 47.9.1, 47.11.2/3, 47.13.2, 47.17.6, 47.22.3/7, 51.1.3, 69.5.7, 69.14.1/3, 70.10.2, 71.6.3, 73.4.1, 73.6.1 (if transcription correct), 100.4.1 (two), 113.3.3, 113.7.2, 117.1.2, 117.6.5, 119.4.4, 120.3.2, 120.5.1/3/4, 157.15.1, 161.5.2, 161.26.1, 162.17.1, 162.18.4, 162.24.6, 163.2.2/3, 166.10.2/4, 166.22.2 (two), 167.1.2, 170.1.1. There is, in addition, an intriguing example of Past (Actual) Negative where one might have expected a Past Potential Negative for semantic reasons, suggesting that in the past tense the Potential form is more marked than its counterpart (Future) is in the nonpast system: 10.2.3 ('he father could not see'); contrast 'inability' exx. of /yagi/, top of this page.

Comment by Robert Churnside-Parraruru

Ngarluma - Tape 2,81 - Diary III 288

TEXT 1

(1) ninda mirlbana nurla Jiirramagardumulala warngai
takaru Darlumagu nadutandugu wāndaru milimilila, (2) mili-
mililakaru turlui Darlumagu wāndaru turlui mili-
mili(-la), (3) mililakaru wāndaru milimilila.
milimilila. wāndalgu
Darlumagu nadutandugu
purruga,
dar
wa



Brandenstein, Carl Georg von. 1970. *Narratives from the North-West of Western Australia in the Ngarluma and Jindjiparndi languages*. (Volumes 1–3 + audio-disc.) Canberra: Australian Institute of Aboriginal Studies.

Interlinear Text Archive

by J. Randolph Valentine
Department of Anthropology
University of Western Ontario
London, Ontario N6A 5C2



This stack is a very simple interlinear text archive, designed especially for use with the Rook® Grammar Authoring System.

Version Tue, Mar 25, 1997

Click mouse to continue...



Warnman story (6) *

Tilpayarna wantimanyiyarna. Nyarnilangka. Jijiyarna
yanarni. Wantinyayarna jumajinya nyarningka ngulyu.

Tilpayarna wantimanyiyarna. Nyarnilangka. Jijiyarna
Still-pa-yarna wanti-manyi-yarna nyarni -la-ngka jiji -yarna
Still-pa-1plexS sit -pres -1plexS this,here-emph -LOC child-1plexS

yanarni. Wantinyayarna,
ya -na -rni wanti-nya -yarna
go-past-? sit -past-1plexS

jumajinya nyarningka ngulyu.
juma-jinya nyarni -ngka ngulyu
big -become this,here-LOC really

(Click on bold text to get more detailed information)

English Only

Hear it

Next

whole



First version: 1990

Version Tue, Mar 25, 1997



ROOK

AUSTRALIAN INSTITUTE OF ABORIGINAL
AND TORRES STRAIT ISLANDER STUDIES

Australia's languages

Australian indigenous languages information
stacks

• Version 1.0.1 © AIATSIS 1995 •



Fieldwork in Eratap and
Erakor villages, Vanuatu
1996–

Grammar based in a corpus

Examples cited to original
sources

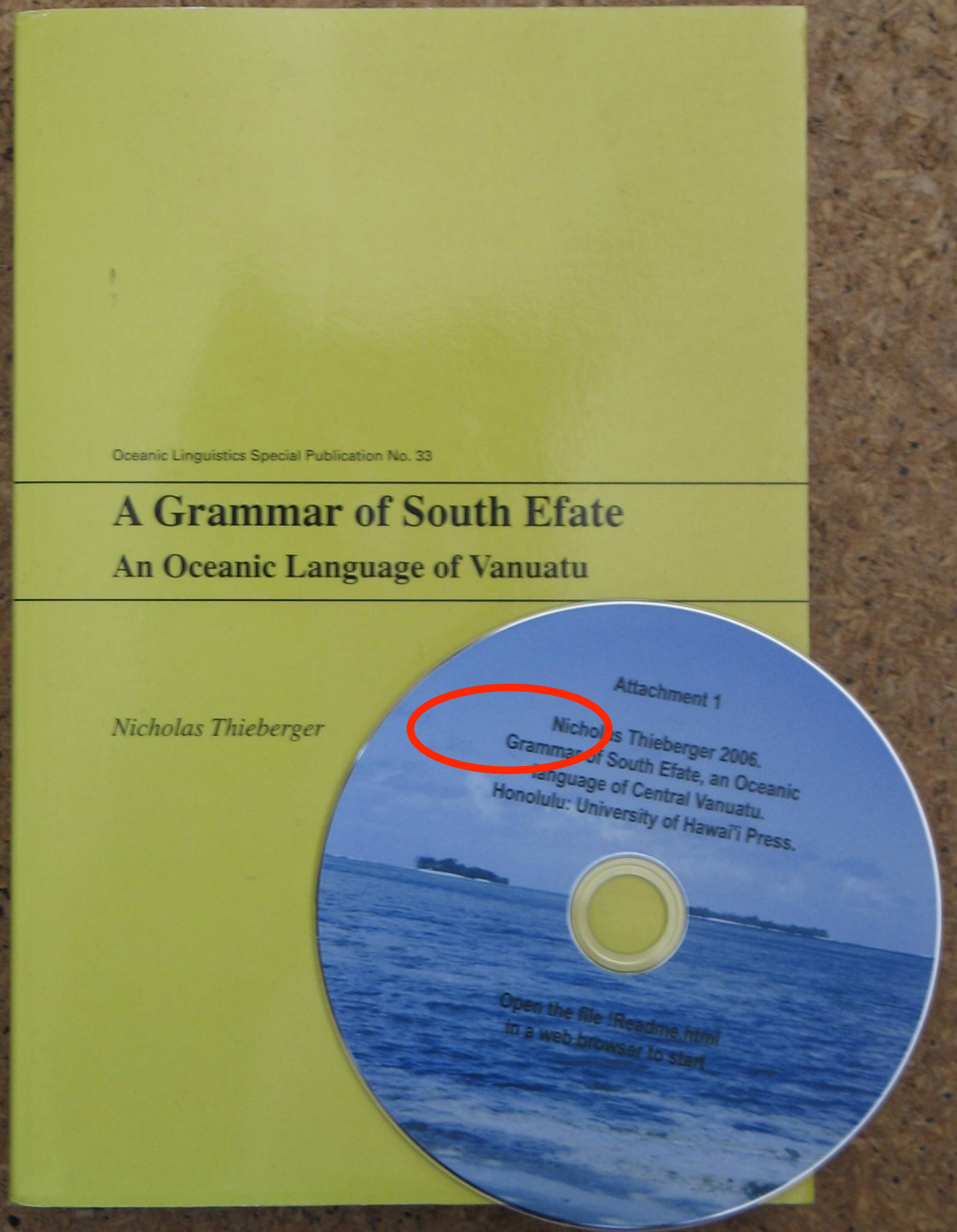
All material archived

Fieldwork in Eratap and
Erakor villages, Vanuatu
1996–

Grammar based in a corpus

Examples cited to original
sources

All material archived



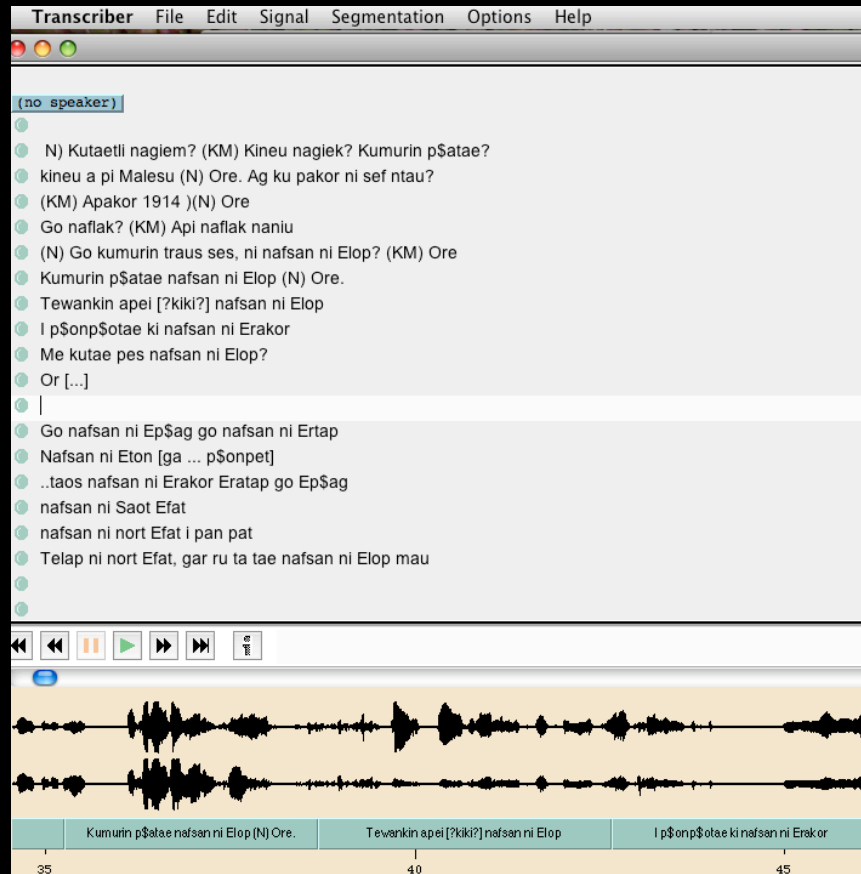
**How to create a citable set of
textual material linked to
primary media?**

Time-aligned transcripts

Transcriber: <http://sourceforge.net/projects/trans/>

Elan: <http://www.mpi.nl/tools/elan.html>

Transana, Exmaralda, Anvil ...



Audiamus

Created for interacting with field recordings via their transcripts

Version 1 – HyperCard – 1999,

Version 2 – RunTime Revolution 2002

<http://bit.ly/audiamus>

Kalsarap.mov

clip time

clip all

0

0

0

(Kalsarap.mov, 0, 0)

start

end

duration

reference

File menu

previous

next

play

stop

play all

set start

set end

unlock

add to playlist

repeat

Kalsarap.mov

< makelist

go to playlist

Text 3 Litrapong a natopu, by Kalsarap Namaf (Play all)

Story from tape 20001bx told by Kalsarap Namaf. Transcribed and translated into Bislama by Manuel Wayane. The story concerns a natopu or spirit called Litrapong, also known in Bislama as a Lisepsep. The telling of this story is also recorded on video.

Akit tumau wi tae esan ipi, go

akit tu= mau wi tae esan i= pi go

1plin 1p.inRS= all know place 3sgRS= be and

Litrapog ne, amurin na kanrikmus ki,

Litrapog ne a= mur -i -n na ka= nrik -mus ki -ø

name this 1sgRS= want -TS -3sg0 COMP 1sgIRS= tell -2p10 PREP -3sg0

We all know that place, and this Litrapong, I want to tell you about.

Litrapog ne, ipi naflak ni apu.

Litrapog ne i= pi naflak ni apu

name this 3sgRS= be clan of grandfather

Gar apu rato pusrek ser tete nrak.

gar apu ra= to puserek ser tete nrak

3pl grandfather 2/3d.RS= HAB talk every some time

Litrapong. She is of grandfather's clan. Those two, grandfather and Litrapong, would talk every now and then.

Ser nrak rato pusrek rato ur naor iskei.

ser nrak ra= to puserek ra= to ur naor i= skei

every time 2/3d.RS= HAB talk 2/3d.RS= HAB follow place 3sgRS= one

Atae naliati iskei

a= tae aliat i= skei

1sgRS= know day 3sgRS= one

Every time they would talk, they would be at one place. I know one day

We all know that place, and this Litrapong, I want to tell you about.

126 Litrapog ne, ipi naflak ni apu. Litrapog ne i= pi naflak ni apu name

133 ne i= pi naflak ni apu name this 3sgRS= be clan of grandfather Gar

142 be clan of grandfather Gar apu rato pusrek ser tete nrak. gar apu ra=

149 pusrek ser tete nrak. gar apu ra= to puserek ser tete nrak 3pl grandfather

357 3sgRS= come ilakor mai saof apu i= lakor mai saof apu 3sgRS= approximate come

362 apu i= lakor mai saof apu 3sgRS= approximate come visit grandfather when we v

575 him herbal medicine. Ser nawes napu, na, a ga preg nalkis ikano nrogtiawi. ser

585 nalkis ikano nrogtiawi. ser nawes napu a= ga preg nalkis i= kano nrogtiawi ev

View concordance

Make text bigger

apu

Find

Search just this file

Find part

Line number 1

Slow Speed Fast

Volume

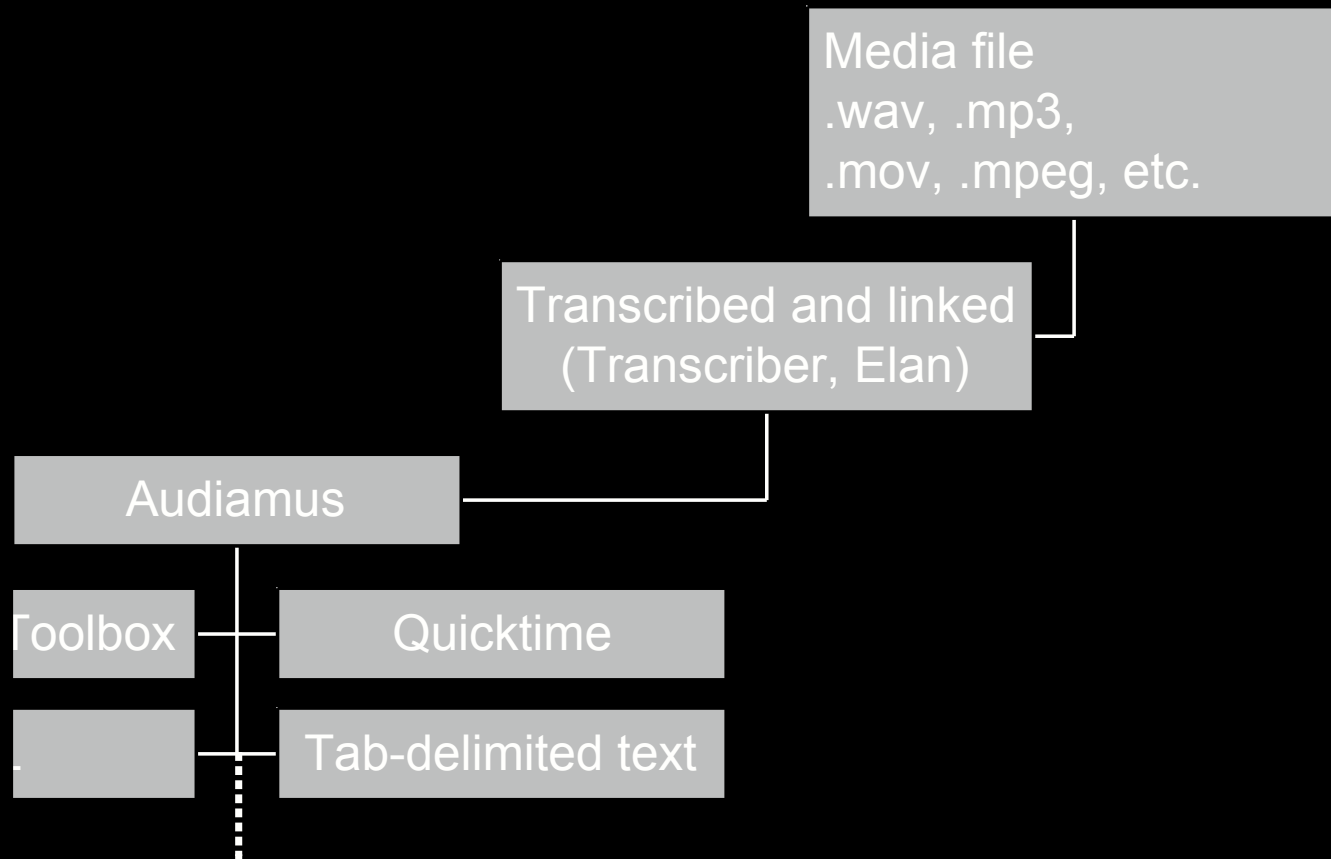
Total of 25 hours 42 minutes in 51 files

Audiamus

- ❖ Analysis of a linguistic media corpus via text
- ❖ Building a citable corpus of media via linked transcripts
- ❖ A working tool (now) that gives access to a media corpus via a time-aligned textual index
- ❖ Takes the output of Transcriber or Elan as input
- ❖ Built in proprietary software, Runtime Revolution (rapid prototyping)
- ❖ Win/Mac (&Linux) standalone versions

❖ <http://bit.ly/audiamus>

Audiamus workflow



Audiamus

Ability to cite primary data

			clip time	clip all
539.887	548.86	8.973	(NT1-98003-B, 539.887, 548.86)	

Komam u=weswes u=maos u=mur-i-n na ko=to
1p.ex 1p:RS=work 1p:RS=tired 1p:RS=want-TS-3sg say 1p.exlRR=stay

wi ko=**fo** **pei** maos.
good 1p.exlRR=PSP:IR first tired

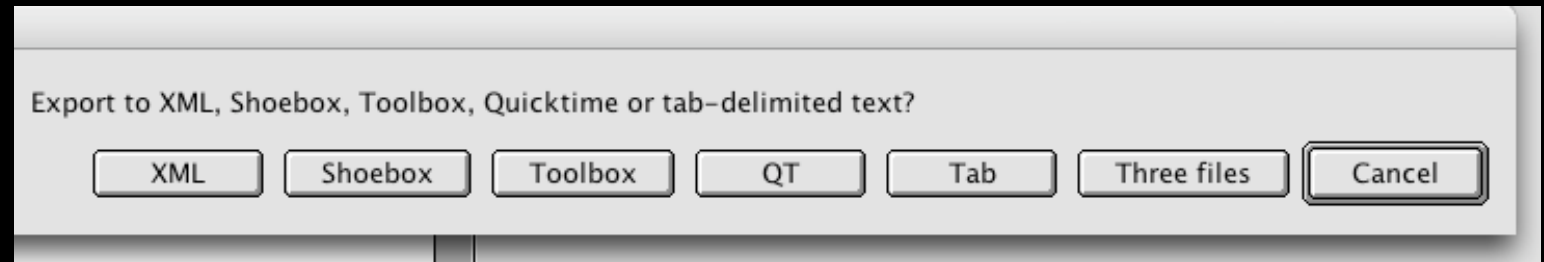
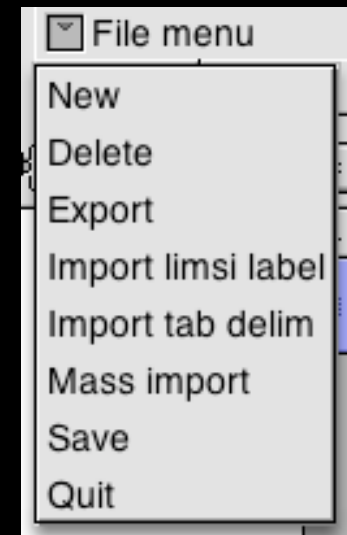
We worked and we got tired, if we wanted to stay healthy, we would (have to) get tired first. (NT1-98003-B, 539.887, 548.86)

Audiamus

Import

Tab-delimited text

Limsi label (Transcriber)



Export

Tab-delimited text

XML

Toolbox

Quicktime

Outputs from *Audiamus*

Toolbox

_sh v3.0 498 SE Text

\aud 98009b 22.94 28.64

\tx (T) A, kagaag traus nmatu iskei,

\aud 98009b 29.16 37.82

\tx go nanwei iskei, rato. Me ratik ki tesa.

\aud 98009b 37.96 47.06

\tx Me ipiatlak teesa iskei ipaakor faat elau.

Outputs from *Audiamus*

Quicktime

```
{QTtext}{font:Geneva}{plain}{size:12}{textColor: 65535, 65535, 65535}{backColor:  
0, 0, 0}{justify:center}{timeScale:1000}{width:160}{height:48}{timeStamps:absolute}  
{language:0}{textEncoding:0}  
[00:00:00.]  
{SPEAKER = HT}Ipiatlak nmatu iskei,  
[00:00:02.920]  
nmatu tiawi iskei ipiatlak tesa inru rana rato panpan. [....]  
[00:00:08.120]
```

I will then go and cut bamboo. I
cut bamboo and bring it, I get it
ready

A ler mai, ka fo pei m\$asel roat,
a m\$asel roat inom



Benefits of Audiamus

Local control of data

Immediate access (no web connection)

Ability to interact with and improve the transcript

Ability to 'clip' citation forms

Imports and exports several formats

Find and concordance function

Creation of 'playlists' organised by theme rather than by media

Freely distributed under GNU General Public License

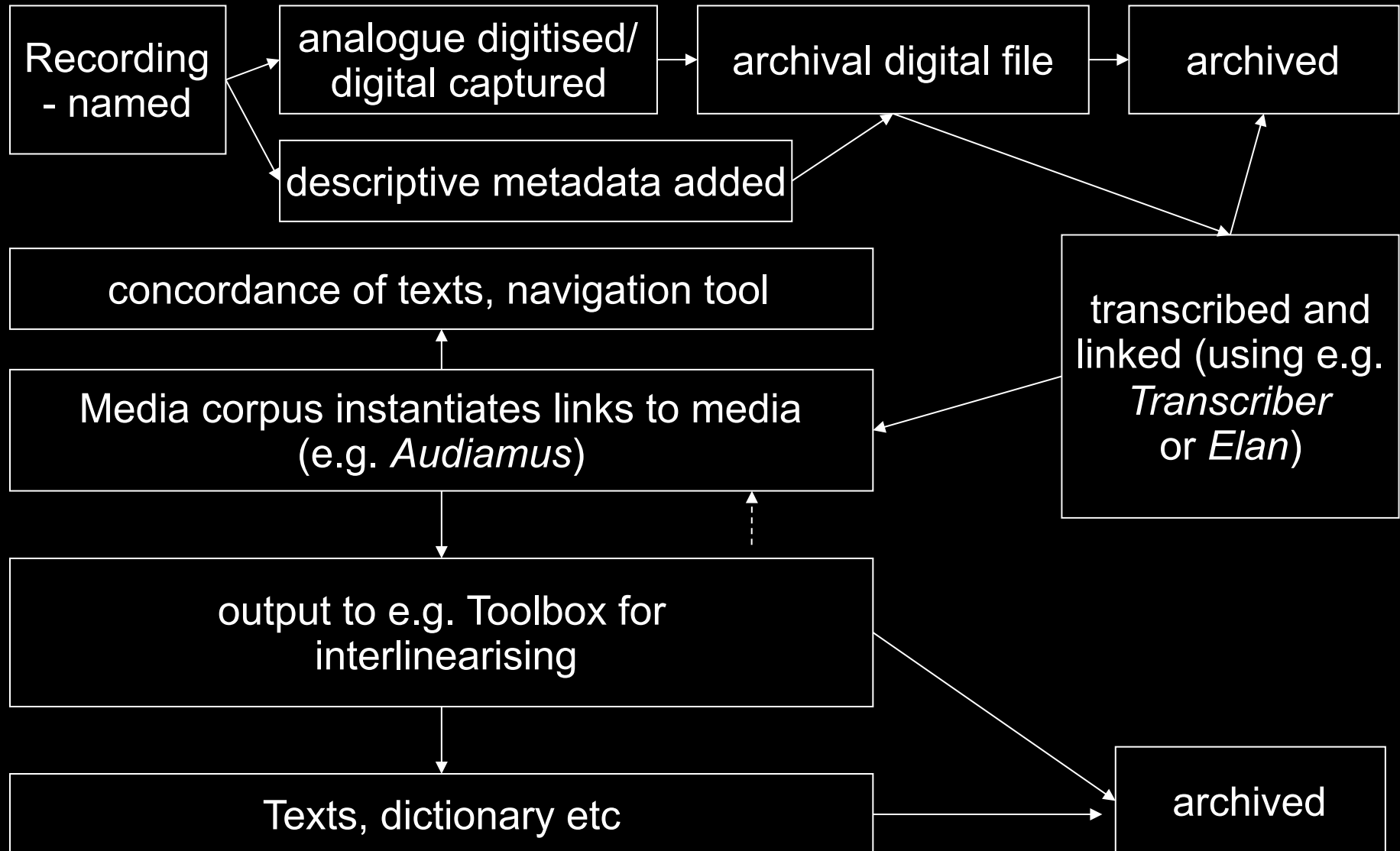
Problems of Audiamus

Data has to be imported, ideally it would be in XML and read by an Audiamus-like browser

Audiamus implies a workflow and so version control is a problem

Proprietary software means development can only be done by buying Runtime Revolution

Typical workflow resulting in well-formed data



Citability

Citation of the spoken word requires a persistent form of the cited form

Persistence of name and location is provided by a digital archive

How will WOLP maintain persistent identification of texts?

How will WOLP license the use of texts?

Cataloging texts

DOI? Handle.net?

What repository to use?

Example of PARADISEC

149 collections

6,517 items

43,301 files

4.67 TB

2653 hours of audio data

Data representing 648 languages from 60
countries



Online catalogue: paradisec.org.au/catalog

Collection	Item	Item list	Search items	Update items	Contact	Import/export	Login	Search	Discard changes	Save changes
------------	------	-----------	--------------	--------------	---------	---------------	-------	--------	-----------------	--------------

	coll ID	item ID		
* Persistent identifier	NT1-98002	NT1 98002		
Catalogue link	http://store.apac.edu.au/paradisec/repository/NT1/98002			
* Item name	CD09-10			
URL				
Hide metadata	<input type="checkbox"/>			
Language as given	South Efate, Bislama			
Date created	1998-09-27	Copy to Date ISO		
* Date created in ISO format	1998-09-27	Copy to Date created		
* Countries	<div>Select ... Inherit from collection <input checked="" type="checkbox"/> VANUATU VU</div>			
** Languages	<div>Inherit all from collection <input checked="" type="checkbox"/><table><tr><td>Efate, South</td><td>erk</td></tr></table></div>		Efate, South	erk
Efate, South	erk			
Dialect				
Region / village	Erakor village			
Discourse Type	Select ...			
* Collector	<input checked="" type="checkbox"/> Thieberger, Nick [details]			
Operator	<input checked="" type="checkbox"/> Bone, Corin [details]			

Collection level information [\[details\]](#)
Collection digitised in 2001. NICK THIEBERGER COLLECTION DATA 51 CDs, total of 38 hours of language material original medium: audio cassette date of recordings: 1996-2000 All burned as .AIFF data CDs. Archive copy of each CD in .AIFF data

Date of assessment		Form received <input checked="" type="checkbox"/>				
Location	returned to depositor					
Countries	VANUATU VU					
Languages	<table><tr><td>Efate, South</td><td>erk</td></tr><tr><td>Lelepa</td><td>lpa</td></tr></table>		Efate, South	erk	Lelepa	lpa
Efate, South	erk					
Lelepa	lpa					

☒ Inherit collection values
Copies the "collector", "access conditions", "media" and "originating uni" values, lat/long, and all countries and languages into this item record

Subject language	Copy from content
Content language	Copy from subject

<input checked="" type="checkbox"/> Bislama	bis
<input checked="" type="checkbox"/> Efate, South	erk

Fields marked with * are mandatory for imp.xml generation.
Fields marked with ** are optional.

Online catalogue: paradisec.org.au/catalog

**** Languages**
☐ Inherit all from collection

☒ Efate, South

☒ Bislama

☒ Efate, South

Dialect

Region / village Erakor village

*** Collector** ☒ Thieberger

Operator ☒ Bone, Corin

**** Note (publishable)** (item notes)

Comments (internal) (internal Paradise)

Role

☒ recorder

☒ speaker

Filename

Files

☒ NT1-001-A

☒ NT1-001-B

☒ NT1-001-C

00003 /u1/archive/SealedBWF/NT1-001-A.wav

00003 /u1/archive/SealedBWF/NT1-001-B.wav

2003-10-17

2003-10-17

7-08

7-08

Select bounding-box

100km

Layers

Honshu

Java & Bali

Kai Islands

Language Groups of The Pacific *

Micronesia

New Britain

New Caledonia *

New Hebrides *

Northeastern Irian Jaya

Northern Celebes

Northern Australia

Northern Borneo

Northern Mainland Southeast Asia

☒ Show item density

Longitude 166.52 - 169.89

Latitude -20.25 - -13.71

Tracking 1 CD given to the Vanuatu Cultural Centre

Access conditions standard, as per PDSC Access form

Originating uni Melbourne

tape received ☐

tape labelled ☒

metadata entered ☐

digitised ☒

ready for imp.xml ☐

CD burnt ☐

data received

date digitised 2001-01-01

imp.xml generated ☒

CD id

	cassettes	r-to-tapes (metr
number	0	0
length	0	0
total	0	0
search set total	0	0
	reel-to-reel speed	
	reel-to-reel time	0



OLAC Record

oai:paradisec.org.au:NT1-98002

Open Archives Initiative

Metadata

<i>Title:</i>	CD09-10
<i>Access Rights:</i>	standard, as per PDSC Access form
<i>Bibliographic Citation:</i>	Thieberger, Nick (recorder), Kalman, Petro (speaker), Kalopog, Kali (speaker), Tenene, Waia (speaker), Thieberger, Nick (depositor) 1998; CD09-10, WAV/MP3/XML/JPG http://paradisec.org.au/repository/NT1/98002 2010-01-01
<i>Contributor (depositor):</i>	Thieberger, Nick
<i>Contributor (recorder):</i>	Thieberger, Nick
<i>Contributor (speaker):</i>	Kalman, Petro Kalopog, Kali Tenene, Waia
<i>Coverage (Box):</i>	northlimit=-17.7235; southlimit=-17.8235; westlimit=168.217; eastlimit=168.317
<i>Coverage (ISO3166):</i>	VU
<i>Date (W3CDTF):</i>	1998-09-27
<i>Date Created (W3CDTF):</i>	1998-09-27
<i>Description:</i>	There are time-aligned transcripts of this item. . Language as given: South Efate, Bislama
<i>Format:</i>	Digitised: yes; Media: audiocassette;
<i>Identifier:</i>	NT1-98002
<i>Identifier (URI):</i>	http://paradisec.org.au/repository/NT1/98002
<i>Language:</i>	Bislama South Efate
<i>Language (ISO639):</i>	bis erk
<i>Rights:</i>	standard, as per PDSC Access form
<i>Subject:</i>	South Efate language
<i>Subject (ISO639):</i>	erk
<i>Subject (OLAC):</i>	language_documentation

Open Language Archives Community (OLAC)

<http://www.language-archives.org/language/erk>



Resources in and about the South Efate language

ISO 639-3: [erk](#)

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 resources in the language](#)

Other known names and dialect names: Efate, South, Erakor, Fate, Southern Efate

Primary texts

1. *Fila Story: A Oa Erua; B&C: Lamenu (cont.); Baki; Pango-Erakor-Eratap*. Capell, Arthur (recorder). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au/AC1-203](http://oai.paradisec.org.au/AC1-203)
2. *Oriya, Lowolwol*. Capell, Arthur (researcher). n.d. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au/AC1-411](http://oai.paradisec.org.au/AC1-411)
3. *Prodigal Son in Efate Dialects*. Capell, Arthur (recorder). 1900. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au/AC1-416](http://oai.paradisec.org.au/AC1-416)
4. *Texts in languages from Efate*. Capell, Arthur. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au/AC2-VEFAT403](http://oai.paradisec.org.au/AC2-VEFAT403)
5. *The Prodigal Son Story in Samoa Point/Havannah Harbour*. Capell, Arthur (compiler); Newton, Peter (depositor); Capell, Arthur (researcher). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au/AC2-VEFAT404](http://oai.paradisec.org.au/AC2-VEFAT404)
6. *The Gospel According to Luke in the language of Efate*. Newton, Peter (depositor); Capell, Arthur (researcher). 1877. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au/AC2-VEFAT501](http://oai.paradisec.org.au/AC2-VEFAT501)
7. *Catechism in the language of Efate*. Newton, Peter (depositor); Capell, Arthur (researcher). 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au/AC2-VEFAT502](http://oai.paradisec.org.au/AC2-VEFAT502)
8. *Notes on Nguna-Efate dialects*. Graham Miller, J. 1999. Pacific And Regional Archive for Digital Sources in Endangered Cultures (PARADISEC). [oai:paradisec.org.au/AC2-VNEFAT104](http://oai.paradisec.org.au/AC2-VNEFAT104)
9. *South Efate Language Primer*. Newton, Peter (depositor); Capell, Arthur (researcher). 1999. Pacific And Regional Archive for Digital

Data access (paradise.org.au/repository)



Welcome to the PARADISEC repository
Please log in to identify yourself

Username:

Password:

[Log in](#)



PARADISEC repository

[logout](#) | [collections](#) | [news](#) | [FAQ](#)

Repository collections

Here's a listing of the collections viewable by your username *nxt563*. Refer to the [repository website glossary](#) for an explanation of terms and abbreviations repository listings.

Collection	Items	Files	Size	Duration	Catalog Metadata
AB1	11 items	44 files	29.38GB	14:47:59.60	metadata
AC1	58 items	307 files	64.81GB	33:15:06.00	metadata
AC2	1,025 items	12,567 files	18.06GB	00:00:00.00	metadata
AM2	14 items	56 files	36.65GB	19:08:27.40	metadata
AM3	13 items	46 files	32.41GB	16:19:28.90	metadata
AM4	1 items	2 files	1.50GB	00:45:17.99	metadata
AP2	4 items	14 files	6.00GB	03:01:18.90	metadata
AR1	13 items	42 files	23.00GB	14:04:24.40	metadata
AR2	3 items	8 files	6.09GB	05:58:15.00	metadata
AS1	6 items	20 files	10.81GB	05:26:47.90	metadata
BE1	1 items	14 files	2.15MB	00:00:00.00	metadata
BH1	21 items	42 files	3.75GB	03:40:37.50	metadata
BM1	2 items	28 files	3.02GB	04:40:35.40	metadata
BP1	2 items	12 files	6.01GB	03:01:29.80	metadata
BP2	9 items	32 files	15.73GB	07:55:18.80	metadata
BP3	3 items	6 files	1.65GB	00:49:47.11	metadata
CB3	5 items	16 files	5.68GB	02:51:47.30	metadata
CH1	42 items	202 files	23.28GB	14:54:39.10	metadata
CM1	10 items	270 files	175.46GB	00:00:00.00	metadata

Data access (paradisec.org.au/repository)

PARADISEC repository

[logout](#) | [collections](#) | [news](#) | [FAQ](#) | [glossary](#) | [admin](#)

[collections](#) » [NT1](#) » [98002](#)

Item **NT1/98002**

Here's a listing of the files of item NT1/98002 viewable by your username *nxt563*. Refer to the [repository website glossary](#) for an explanation of terms and abbreviations used in the repository listings.

File	Size	Duration	Class	Status	Metadata	Catalog Metadata
NT1-98002-98002A.mp3	37.39MB	00:40:50.19	audio	online	metadata	metadata
NT1-98002-98002A.wav	413.37MB	00:40:55.74	audio	online	metadata	metadata
NT1-98002-98002B.mp3	37.35MB	00:40:47.97	audio	online	metadata	metadata
NT1-98002-98002B.wav	413.00MB	00:40:53.52	audio	online	metadata	metadata
NT1-98002-98002az.xml	39.74KB	--	xml	online	metadata	metadata
NT1-98002-98002bz.xml	46.42KB	--	xml	online	metadata	metadata
NT1-98002-A.tab	23.65KB	--	unknown	online	metadata	metadata
NT1-98002-A.xml	39.73KB	--	xml	online	metadata	metadata
NT1-98002-A001.jpg	1.60MB	--	image	online	metadata	metadata
NT1-98002-A002.jpg	1.65MB	--	image	online	metadata	metadata
NT1-98002-A003.jpg	1.64MB	--	image	online	metadata	metadata
NT1-98002-A004.jpg	1.68MB	--	image	online	metadata	metadata
NT1-98002-A005.jpg	1.68MB	--	image	online	metadata	metadata
NT1-98002-A006.jpg	1.66MB	--	image	online	metadata	metadata

Reuse

Recording > Transcript

Preg nafkal skot nañer nig Emlakul.

NT1-98002-A 316.43 320.317

Reuse

Recording > Transcript > Interlinear

\id 061:005

\aud NT1-98002-A 316.43 320.317

\tx Preg nafkal skot nañer nig Emlakul.

\mr preg nafkal skot nañer nig Emlakul

\mg make war with people of p.name

\fg *Fight with the people from Malakula.*

South Efate

[Lexicon](#) [English - South Efate](#) [Categories](#)

a e f g i k l m n o p r s t u w

This dictionary is a work in progress. This edition can be cited as: Nicholas Thieberger (2007-01-16) Dictionary of South Efate (<http://paradisec.org.au/SELexicon/index-english/main.htm>). This work is the result of a collaborative project between speakers of South Efate, mainly in Erakor village, Efate, Vanuatu, and [Nicholas Thieberger](#) who has written a grammar of the language. It was produced using [Shoebbox](#) and [LexiquePro](#).

English - South Efate

a b c d e f g h i j k l m n o p q r s t u v w y

A - a

abalone	kaiaraskei
abort	kispun
above	elag
absent	lalu
.	puel
accompany	ptaan
.	plak
act	gien
active	polkirkir
adam's apple	npatnkafik
add	skar
admire	lewi
adopt	pas
adult	pafof
adze	limur
.	kram pel
aeroplane	man ni nmalfa
after	me
.	ntakun
afterbirth?	nal ni teesa
afternoon	kotfan
again	mer
against	ktekit
.	putkau
age-mate	tak
agree	trok
aim, point with finger	tum

nakper *prep.* ahead, forward, front. *Kalek nakper selwan kana pueg namran.* You look ahead when you want to do something.

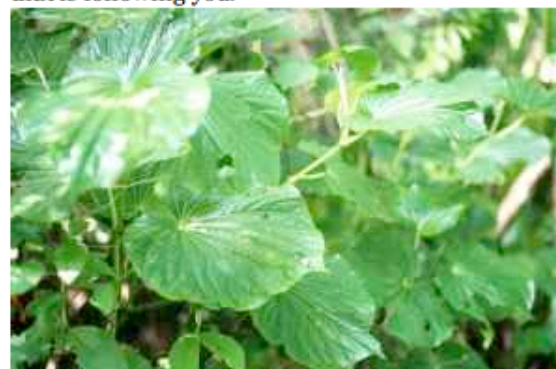
nakpes *n.* tree sp., young leaves eaten together with **ntali** fruit. Also use the wood to make fish traps by planting sticks vertically in the water to create a fence.

nakrik *n.* Fig tree. *Ficus virgata*.

nakrok *n.* vine sp. Strong vine.



nakser *n.* wild kava, creeping thick vine, leaf eaten, but not the root. Its leaves can stop a devil that is following you.



nakte *pron ?.* mine (1st person singular possessive).

nakur *n.* Indian Mulberry, not used as a dye today, but it was used as a yellow dye for mats and hair. Lavard (1915) says it was used for sickness, and that it was mixed with lime to make a red paint. Also

South Efat

Lexicon English - South Efat Categories

a e f g i k l m n o p r s t u w

This dictionary is a work in progress. This edition can be cited as: Nicholas Thieberger (2007-01-16) Dictionary of South Efat (<http://paradisec.org.au/SELexicon/index-english/main.htm>). This work is the result of a collaborative project between speakers of South Efat, mainly in Erakor village, Efat, Vanuatu, and [Nicholas Thieberger](#) who has written a grammar of the language. It was produced using [Shoobox](#) and [LexiquePro](#).

English - South Efat

a b c d e f g h i j k l m n o p q r s t u v w y

A - a

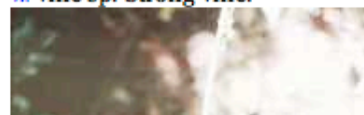
abalone

kaiaraskei

nakpes *n.* tree sp., young leaves eaten together with **ntali** fruit. Also use the wood to make fish traps by planting sticks vertically in the water to create a fence.

nakrik *n.* Fig tree. *Ficus virgata*.

nakrok *n.* vine sp. Strong vine.



EFATE.LEX:2

\lx **nakser**
 \ps **n**
 \ge wild_kava
 \de wild kava, creeping thick vine, leaf eaten, but not the root. Its leaves can stop a devil that is following you.
 \nt info from William Wayne and Sailas Kalfapun 3/10/98
 \syl nak.ser
 \CV CVC.CVC
 \sl 2
 \pic 19, flm 3 11/4/00 JM
 \sd plants
 \pc 3_21.jpg
 \aud NT1-98015-B 1896.74 1900.431
 \sf nakser.wav

its leaves can stop a devil

aim. point with finger **tum**

hair. Lavard (1915) says it was used for sickness, and that it was mixed with lime to make a red dye for mats and

South Efat

a e f g i

This dictionary is a work in progress. This edition can be cited as the result of a collaborative project between speakers of South Efat was produced using [Shoebox](#) and [LexiquePro](#).

English - South Efat

abcdefghijklmnopqrstuvwxyz

A - a

abalone

kaiaraskei

EFATE.LEX:2

\lx	nakser
\ps	n
\ge	wild_kava
\de	wild kava, creeping thick vine, leaf eaten, but
\nt	info from William Wayne and Sailas Kalfap
\syl	nak.ser
\CV	CVC.CVC
\sl	2
\pic	19, flm 3 11/4/00 JM
\sd	plants
\pc	3_21.jpg
\aud	NT1-98015-B 1896.74 1900.431
\sf	nakser.wav

Transcriber 1.5.1

File Edit Signal Segmentation Options Help

(no speaker)

- [lead in silence]
- id 31 October, Endis reading words
- sfiu
- kat porpres npatin
- p\$roro
- nrig
- ntan
- [silence]
- nafet
- puk
- kuom
- nakser
- p\$aup\$al
- ptom
- fusuer
- kat
- urur
- telekor
- kuyaf
- nagal
- sisi



aim. point with finger tum

hair. Lavard (1915) says it was used for sickness, and that it was mixed with lime to make a red paint. Also dye for mats and

South Efa

Lexicon English - South Efa

a e f g i k l m n o p r s t u w

English - South Efa

a b c d e f g h i j k l m n o p q r s t u v w y

A - a

abalone	kaiaraskei
abort	kispun
above	elag
absent	puel
.	lalu
accompany	plak
.	ptaan
act	gien
active	polkirkir
adam's apple	npatnkafik
add	skar



nafinaotan *n.* chiefliness, also the chiefly line. *See:* **naot**.

nafinrin *Variant:* **nafinr**. *n_inposs.* rib.

nafis *n.* Fijian asparagus, navisa (Bislama), small sugar cane.

nafit *n.* slave. This could be someone who has had their land taken away as punishment and is sent away from their own village. *See:* **manmun**. *Usage:* archaic.

nafitiawian *n.* old age. *See:* **tiawi**.

nafkal *n.* battle, war. **Selwan inlaun rupakruk rutae preg nafkal**. They can come together for big battles. **Preg nafkal skot narner nig Emlakul**

nafkou *n.* messenger. **Pafi nafkou pak Eton**. You be the messenger to Eton.

naflak *n.* clan. Note: you are in the same clan as your mother. You have to marry someone from another clan. Some clans no longer exist because of no children in the line. Some of the clans and their characteristics are: **kram** a shell which sits on the beach in good time but buries itself in the sand in bad times; **nmial** a wild yam. The shoots don't follow the rope you put to train it. **namkanr** a plant which grows quickly, takes firm hold, but dies soon after; **ntal** taro, skin is a

Online dictionary with sound and image links, constructed from well-formed underlying data

Waya Tenene, Story of Erakor and Mare



People from Mare came here in the time of the Colonial Government.

The French government sent them as police.

They came to be police on Efate.

On the French side, the French government had a fight in Malakula and it sent them to Malakula.

To fight the people of Malakula.

Those from Mare who came, those from New Caledonia who came because the French government called them.

The police came here.

Teni Emar nen kin rumai pak san tetuei mal ni kolonial kafman.

Franis kafman isent kir rumai reki na polis.

Rumai pi polis ni Efate.

Nanre ni Franis, Franis kafman go ipi atlak nafkal nen ito Emlakul go itpekir pak Emlakul.

Preg nafkal skot namier nig Emlakul.

Teni Emar nen kin rumai, teni Caledoni nen rumai pa esa nlaken Franis kafman isosor nafet.

Polis rumai pak esa.

Me, selwan ag kupi eñae, tiawi itraus traus
traus traus, natrauswen ga itaos nlag. Itrausi
pan kaipa. Me komam uta laap kin uto mau, a?
Malen umat, inom.

*But when you are far away the old man can talk
and talk and talk, his story is like the wind.
He tells it and it is gone. But there aren't
many of us left. When we die, it will be
finished.*

†Kalfañun Mailei, 1998, Erakor Village, Efate, Vanuatu
(NT1-98009-A, 1920.92, 1930.76)

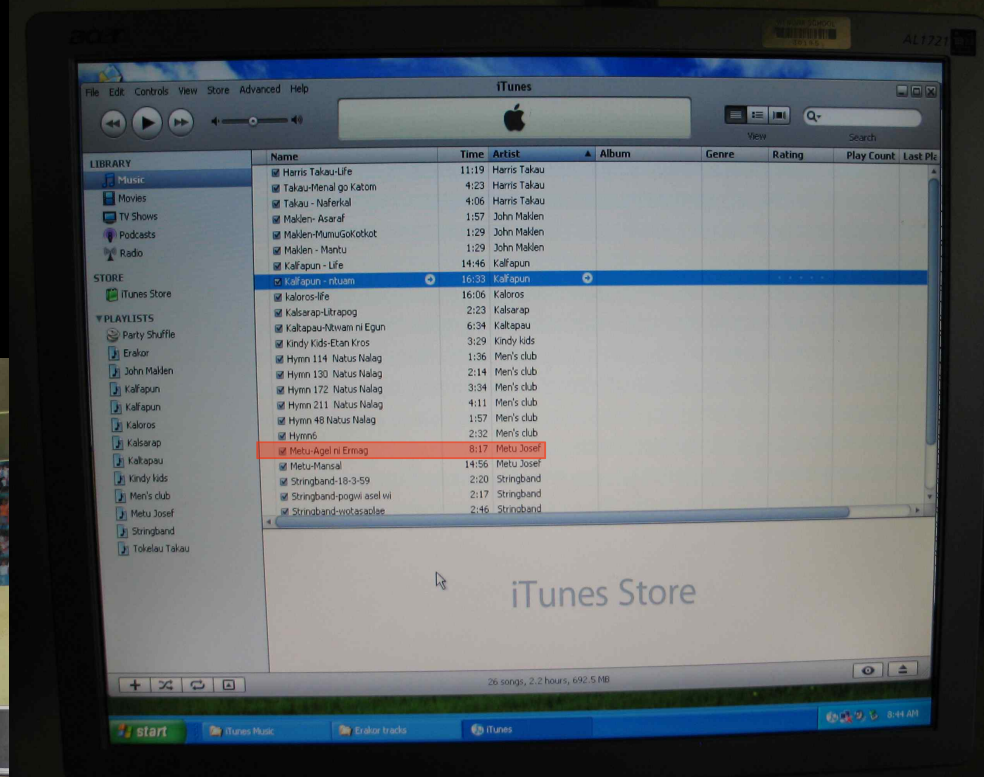


NT1-98009-A.wav 1932.36 1942.16

iTunes installation



Erakor school, Efate, Vanuatu



Example of reuse of interlinear glossed text (IGT)

Ipitlak nmatu inru, rato elag Eṗuf.

i=		pitlak		nmatu		i=		nru		ra=		to		elag		Eṗuf
3sRS=		have		woman		3sRS=		two		3d.exRS=		stay		above		p.name

There are two women, they lived up at Buḡa.

NT1-98004-A 210.34, 212.2

Interlinearising tools

TRANSC(ript) for CPM

IT for DOS and Macintosh in the 1980s

Shoebox and its successor Toolbox

ITE (Jacobson n.d.)

Fieldworks

TypeCraft

For a detailed listing of annotation tools see:

http://annotation.exmaralda.org/index.php/Linguistic_Annotation



Le Lacito

La recherche

Langues et pays

Publications

Archives orales

Projets associés

Les membres

Rechercher
Sur le WEB du Lacito

11/11/2016



[Accueil](#) > [Accueil Archive](#)

To English


Le contenu de l'archive ([mode d'emploi](#) / [rechercher](#))

L'archive du LACITO ("Archivage") donne accès à des exemples de parole spontanée, prioritairement dans des langues « rares », enregistrés dans leur contexte social et transcrits en ... [Lire la suite](#)

L'archive publique contient plus de 200 documents en 45 langues, annotés par une vingtaine de chercheurs (liste non exhaustive car ne comprenant pas les données en cours de numérisation).

- Afrique
- Balkans
- Caucase
- Océanie
- Népal et Asie
- Moyen-Orient et Amérique du Sud

Corpus de langues d'Afrique (présentation) ▲

Langue	Famille	Pays	Chercheur(s)	Archive
grand-comorien	Bantu	Comores	Marie-Françoise Rombi	
lhanzu	Bantu	Tanzanie	Margaret Dunham	
langi	Bantu	Tanzanie	Margaret Dunham	
mahorais	Bantu	Comores	Marie-Françoise Rombi	
mbugwe	Bantu	Tanzanie	Margaret Dunham	
nyilamba	Bantu	Tanzanie	Margaret Dunham	
ouldémé	Tchadique	Cameroun	Véronique de Colombel	



[Accueil](#) > [Accueil Archive](#)

Titre: Iron (Langue: Langi)

Mot à mot : ☒ Traductions: ☒



Lecture en continu: ☒

S1 **_susu valangi tadoma n iseki toayyawa**

Nous les Valangi, nous allions dans la forêt pour ramasser
We Valangi went to the forest to collect

susu va-langi tu-a-dom-a na i-saka-i t-o-a-yuw-a

pers1pp PI2-langi PV1pp-acc-aller-sfx ass PI5-forêt-loc PV1pp-rel-acc-ramasser-sfx

S2   **_kur iseki dapakesari yuwa nkowa nkwala kuko**

la-bàs, dans la forêt DoWakEsari, pour tourner la-bàs
there, in the DoWakEsari forest, to turn there

ku-ra i-saka-i danakesari yuw-a Ø-ŋkɔwa nɔ-walal-a ku-ko

PD17-dém é PI5-forêt-loc DaWakEsari ramasser-sfx PI9-maïs rel-tourner-sfx PI17-déict

Les membres

Rechercher

Sur le WEB du Lacito

11/11/2016



S3   ƙotayuwire ma takodza na kayi

Le ramassage fini, nous rentrions à la maison.
When the collecting was finished, we came back home.

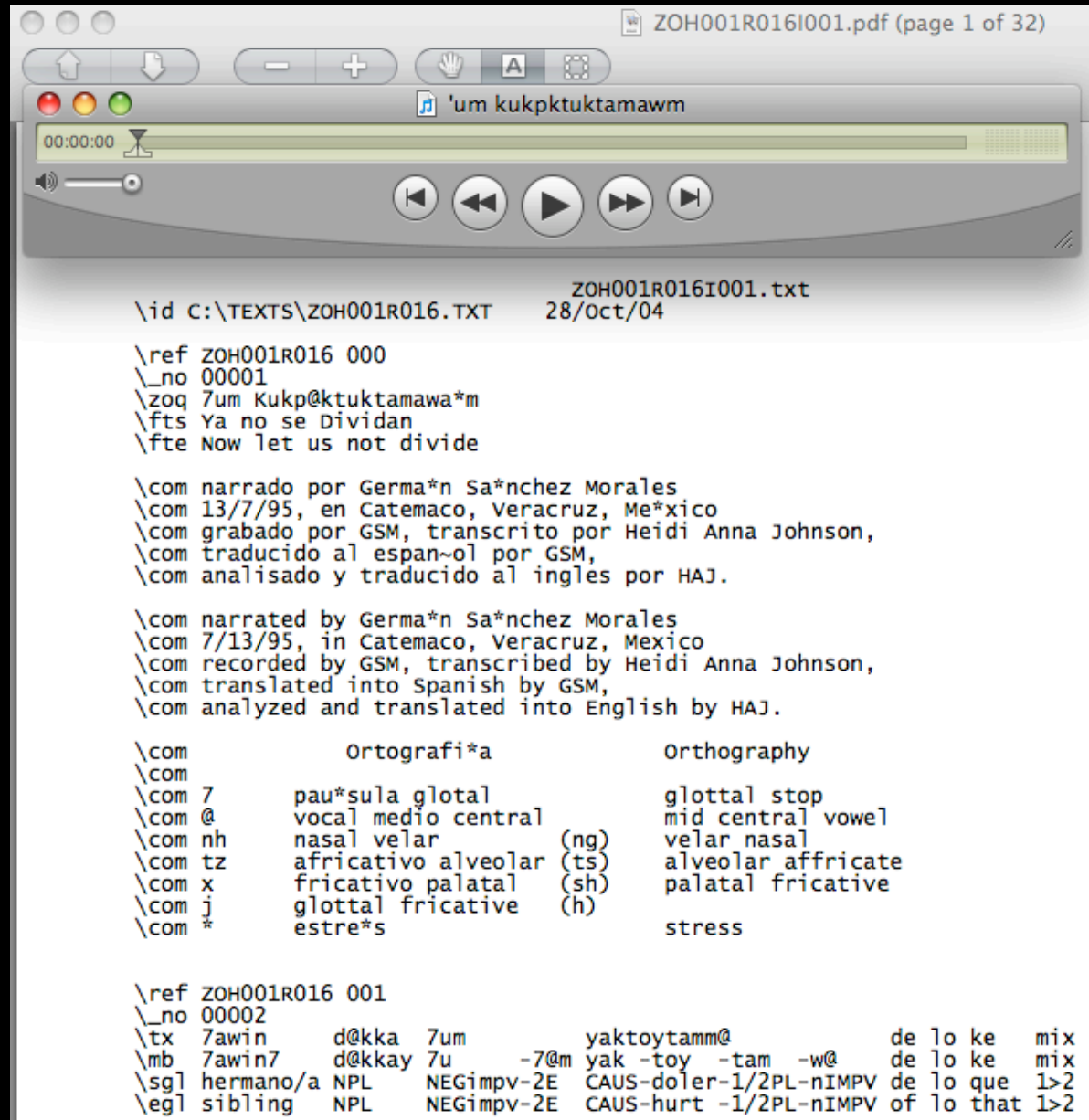
k-ɔ-t-a-yuw-ire ma t-a-k-ɔdʒ-a na Ø-kaya-i

PV17-rel-PV1pp-acc-ramasser-perf puis PV1pp-acc-PI15-venir-sfx ass PI9-maison-loc

S4  _ma takolba iruso

Puis nous préparons de l'alcool.

AILLA text / media formats



ZOH001R016I001.pdf (page 1 of 32)

'um kukpktuktamawm

00:00:00

\id C:\TEXTS\ZOH001R016.TXT ZOH001R016I001.txt
28/oct/04

\ref ZOH001R016 000
_no 00001
\zoq 7um Kukp@ktuktamawa*m
\fts Ya no se Dividan
\fte Now let us not divide

\com narrado por Germa*n Sa*nchez Morales
\com 13/7/95, en Catemaco, Veracruz, Me*xico
\com grabado por GSM, transcrito por Heidi Anna Johnson,
\com traducido al espan~ol por GSM,
\com analizado y traducido al ingles por HAJ.

\com narrated by Germa*n Sa*nchez Morales
\com 7/13/95, in Catemaco, Veracruz, Mexico
\com recorded by GSM, transcribed by Heidi Anna Johnson,
\com translated into Spanish by GSM,
\com analyzed and translated into English by HAJ.

\com	Ortografi*a	Orthography
\com		
\com 7	pau*sula glotal	glottal stop
\com @	vocal medio central	mid central vowel
\com nh	nasal velar (ng)	velar nasal
\com tz	africativo alveolar (ts)	alveolar affricate
\com x	fricativo palatal (sh)	palatal fricative
\com j	glottal fricative (h)	
\com *	estre*s	stress

\ref ZOH001R016 001
_no 00002
\tx 7awin d@kka 7um yaktoytamm@ de lo ke mix
\mb 7awin7 d@kkay 7u -7@m yak -toy -tam -w@ de lo ke mix
\sgl hermano/a NPL NEGimpv-2E CAUS-doler-1/2PL-nIMPV de lo que 1>2
\egl sibling NPL NEGimpv-2E CAUS-hurt -1/2PL-nIMPV of lo that 1>2

INDIGI TUBE

[Home Page](#) | [Radio: Listen Live](#) | [Radio Archive](#) | [Watch Videos](#) | [Contact Us](#)

Search, 2 Results!



▶ [Ngaanyatjarraaku Turku Purting](#)

Live recording of Warburton Music Festival, featur



▶ [Patjarr Kapi Piti Tjukurrpa](#)

A combination of tjurkurpa and oral history, Pulpur
Patjarr Rockhole.

Search!



Patjarr Kapi Piti Tjukurrpa

From: NG MEDIA

Oral Histories & Bush Trips

Location: **Patjarr Rock Hole, WA** A combination of tjurkurpa and oral history, Pulpuru
Davies tells the story of the Patjarr Rockhole.

Duration: 14.26

Made In 2009

Language: Noaanvatiarra

10 Indigitube, Indigitube Radio and all associated trademarks, including Broadcast Material is © IRCA

<http://www.indigitube.com.au/>

Why not use youtube?

Text / media link not provided for

No commitment to longevity

No path from linguistic fieldwork tools
(especially interlinear glossed text)

Lack of control

Ideally linguistic institutions would be curators

IP issues not in control of the depositor

Text and media

<http://eopas.erc.unimelb.edu.au:8080/exist/eopas3/browse.xq>

Media Player



Display Options

☐ Interlinear ☒ Translation

Concordance

Downloads: [Original](#) | [Eopas](#) | [CMML](#) | [Toolbox](#) | [Transcriber](#) | [Elan](#) | [PDF](#) | [RTF](#)

Wit Go Kusu ▢

2006-11-27, Tokelau, erk, VU, ...

date	2006-11-27
creator	Tokelau
contributor	recorder: Nick Thieberger
coverage	VU ?
subject	(erk)
language	erk ?
type	linguistic-type: Primary Text discourse-type: Narrative
mediafile	http://maenad.itee.uq.edu.au/exist/mediafiles/annodex/21116960.anx

- s1 + Amurin gag puserek, Nick, kafo gag pusereki kusu go wit.
I want to tell you, Nick, I'll tell you about the rat and the octopus.
- s2 + Naliati iskei, elau imat.
One day, it was low tide.
- s3 + Elau imat, itrau mat pe mat pe mat me.
It was low tide, a really low tide.
- s4 + Ipiatlak nskau ienkot naur iskei
There was a reef around an island.
- s5 + Naur ses iskei ito elau me elau imat
This island was in the sea, but the tide was out.
- s6 + Malen elau imat kerkrai,
When the tide was very low,
- s7 + mat top, go nskau igar userek
very low, the reef was dry right around the island
- s8 + pan tkal naur ses.
right up to the small island.
- s9 + Me kusu, kusu iskei, ito Efate
But the rat, this rat, it was from Efate.

Text and media

<http://eopas.erc.unimelb.edu.au:8080/exist/eopas3/browse.xq>

- Media Player



- Display Options

☒ Interlinear ☒ Translation

- Concordance

Downloads: [Original](#) | [Eopas](#) | [CMML](#) | [Toolbox](#) | [Transcriber](#) | [Elan](#) | [PDF](#) | [RTF](#)

Wit Go Kusu ▶

2006-11-27, Tokelau, erk, VU, ...

date	2006-11-27
creator	Tokelau
contributor	recorder: Nick Thieberger
coverage	VU ?
subject	(erk)
language	erk ?
type	linguistic-type: Primary Text discourse-type: Narrative
mediafile	http://maenad.itee.uq.edu.au/exist/mediafiles/annodex/21116960.anx

▶ s1 + Amurin gag puserek, Nick, kafo gag pusereki kusu go
a= mur -i -n gag puserek Nick ka= fo gag puserek -ki kusu go
1sgRS= want -TS -3sgO 2sgBEN talk Nick 1sgIRR= PSP.IR 2sgBEN talk -TR rat and
wit.
wit
octopus
I want to tell you, Nick, I'll tell you about the rat and the octopus.

▶ s2 + Naliati iskei, elau imat.
naliati i= skei elau i= mat
day 3sg.R.SBJ= one sea 3sgRS= low_tide
One day, it was low tide.

▶ s3 + Elau imat, itrau mat pe mat pe mat me.
elau i= mat i= trau mat pe mat pe mat me
sea 3sgRS= low_tide 3sgRS= very low_tide INT low_tide INT low_tide and
It was low tide, a really low tide.

▶ s4 + Ipiatlak nskau ienkot naur iskei
i= piatlak nskau i= enkot naur i= skei

Text and media

<http://eopas.erc.unimelb.edu.au:8080/exist/eopas3/browse.xq>

Media Player



Display Options

☒ Interlinear ☒ Translation

Concordance

lel tete naik elau
na ke= fak elau
gakit sao ntas elau
lel tete naik elau
na ke= fak elau
kin ra= to elau Ekaꝑum
-n i= puel elau go i=
of malnen elau i= mat
elau i= mat
panpan elau i= mat
me elau i= mat
naliati i= skei elau i= mat
elau i= mat

Downloads: [Original](#) | [Eopas](#) | [CMML](#) | [Toolbox](#) | [Transcriber](#) | [Elan](#) | [PDF](#) | [RTF](#)

Wit Go Kusu ▶

2006-11-27, Tokelau, erk, VU, ...

s1 + Amurin gag puserek, Nick, kafo gag pusereki kusu go
a= mur -i -n gag puserek Nick ka= fo gag puserek -ki kusu go
1sgRS= want -TS -3sgO 2sgBEN talk Nick 1sgIRR= PSP.IR 2sgBEN talk -TR rat and
wit.
wit
octopus

I want to tell you, Nick, I'll tell you about the rat and the octopus.

s2 + Naliati iskei, elau imat.
naliati i= skei elau i= mat
day 3sg.R.SBJ= one sea 3sgRS= low_tide

One day, it was low tide.

s3 + Elau imat, itrau mat pe mat pe mat me.
elau i= mat i= trau mat pe mat pe mat me
sea 3sgRS= low_tide 3sgRS= very low_tide INT low_tide INT low_tide and

It was low tide, a really low tide.

s4 + Ipiatlak nskau ienkot naur iskei
i= piatlak nskau i= enkot naur i= skei
3sgRS= have reef 3sgRS= lie_around island 3sgRS= one

There was a reef around an island.

s5 + Naur ses iskei ito elau me elau imat
naur ses i= skei i= to elau me elau i= mat
island small 3sgRS= one 3sgRS= stay sea but sea 3sgRS= low_tide

This island was in the sea, but the tide was out.

s6 + Malen elau imat kerkrai,
malnen elau i= mat kerkera
as sea 3sgRS= low_tide strong

When the tide was very low,

Fragment of IGT encoded in XML

```
<text xsi:type="orthographic">amurin</text>
```

```
<morphemes>
```

```
<morpheme>
```

```
<text xsi:type="morpheme">a</text>
```

```
<text xsi:type="gloss">1sgRS</text>
```

```
</morpheme>
```

```
<morpheme>
```

```
<text xsi:type="morpheme">mur</text>
```

```
<text xsi:type="gloss">want</text>
```

```
</morpheme>
```

```
<morpheme>
```

```
<text xsi:type="morpheme">-i</text>
```

```
<text xsi:type="gloss">-TS</text>
```

```
</morpheme>
```

```
<morpheme>
```

```
<text xsi:type="morpheme">-n</text>
```

```
<text xsi:type="gloss">-3sgO</text>
```

```
</morpheme>
```

```
</morphemes>
```


We need a standard schema for IGT

No schema (standard representation) for IGT

So ODIN relies on heuristics to locate IGT on the web
(as the basis for making typological claims)

<http://www.csufresno.edu/odin/>

Inferring the nature of IGT from contextual clues

Eopas sample converted from Toolbox XML

```
<database>
<itmgroup>
<itm>105</itm>
<idgroup>
  <id>001</id>
  <aud>200518.aud 20.507 27.202</aud>
  <txgroup>
    <tx>Amurin</tx>
    <mr>a=</mr>
    <mg>1sgRS=</mg>
    <mr>mur</mr>
    <mg>want</mg>
    <mr>-i</mr>
    <mg>-TS</mg>
    <mr>-n</mr>
    <mg>-3sgO</mg>
  </txgroup>
  <fg>I want to tell you a story.</fg>
</idgroup>
```

```
<Interlinear-text>
<phrases>
  <phrase id="s1" startTime="20.507" endTime="27.202">
    <text xsi:type="translation">I want to tell you a story</text>
    <text xsi:type="orthographic">Amurin na kagag</text>
    <words>
      <word>
        <text xsi:type="orthographic">amurin</text>
        <morphemes>
          <morpheme>
            <text xsi:type="morpheme">a=</text>
            <text xsi:type="gloss">1sgRS=</text>
          </morpheme>
          <morpheme>
            <text xsi:type="morpheme">mur</text>
            <text xsi:type="gloss">want</text>
          </morpheme>
          <morpheme>
            <text xsi:type="morpheme">-i</text>
            <text xsi:type="gloss">-TS</text>
          </morpheme>
          <morpheme>
            <text xsi:type="morpheme">-n</text>
            <text xsi:type="gloss">-3sgO</text>
          </morpheme>
        </morphemes>
      </word>
```

....

Uploading...

The screenshot shows the eopas EthnoER online presentation and annotation system interface. The browser window title is "http://localhost:8080/exist/eopas/uploader.xq - Microsoft Internet Explorer". The address bar shows "http://localhost:8080/exist/eopas/uploader.xq". The page header includes the eopas logo, the text "EthnoER online presentation and annotation system", and a user name "Nick Thieberger" with a "logout" link. A navigation bar contains links for "Home", "Advanced Search", "Upload", "Help", and "About EthnoER".

Step 1: Locate and validate your Files

Select the type of transcript you want to upload, locate your transcript file, as well as the corresponding media file.

Type	Transcript File	Media File
Toolbox	<input type="text"/> Browse...	<input type="text"/> Browse...

Validate

Your transcript file *ofgomakou.xml* is valid (against Schema */db/eopas/schemas/toolbox.xsd*) and can be uploaded with the media file after completing Step 2

Step 2: Metadata

We've filled in a few of these fields automatically. You might like to revise or add more information:

Title:	<input <="" td="" type="text" value="is it a shrimp or what?"/>	
Creator:	<input type="text" value="Nick Thieberger"/>	
Contributer:	speaker	<input type="text" value="Speaker Name"/>
Coverage:	Vanuatu (VU)	<input <="" td="" type="text" value="?"/>
Subject:	Efate, South (ek)	<input type="text" value=""/>
Language:	Efate, South (ek)	<input <="" td="" type="text" value="?"/>
Date:	<input type="text" value="2006-09-24+10:00"/>	
Linguistic Type:	Primary Text	
Discourse Type:	Narrative	

Preview

<http://eopas.esrc.unimelb.edu.au:8080/exist/eopas3/transcript/15582013>

Browsing...

The screenshot shows a Mozilla Firefox browser window. The address bar displays the URL `http://maenad.itee.uq.edu.au/exist/e>`. The page header features the 'eopas' logo and the text 'EthnoER online presentation and annotation system'. To the right of the logo are input fields for 'username' and 'password', and a 'login' button. Below the header is a navigation bar with links: 'Home | Browse | Advanced Search | Upload | Help | About EthnoER'. A search bar is located on the right side of this bar. The main content area is titled 'Browse Collection' and contains a paragraph: 'Below is a list of the transcripts currently stored in the database. Click on the table header to sort or on the various values to filter the list. Clicking the title will open the transcript.' Below this text is a table with five columns: 'date', 'creator', 'title', 'lang', and 'country'. The table lists 14 transcripts. The status bar at the bottom of the browser window shows 'Done'.

Browse Collection

Below is a list of the transcripts currently stored in the database. Click on the table header to sort or on the various values to filter the list. Clicking the title will open the transcript.

date	creator	title	lang	country
1998	Nick Thieberger	Koaiseno	erk	VU
2003	Nick Thieberger	Agel ni Ermag	erk	VU
2005	Tokelau Takau	Wit Go Kusu, the octopus and the rat	erk	VU
1998	Nick Thieberger	Litog story	erk	VU
1998	Tokelau Takau	Natopu karu, 'Another spirit'	erk	VU
1996	Nick Thieberger	English Police	erk	VU
2005	Tokelau Takau	Wit Go Kusu, the octopus and the rat	erk	VU
1998	Nick Thieberger	Litapurog	erk	VU
1998			erk	VU
2001-11-10	John Bowden	Tujuh burung, Seven birds	wmh	TL
2006-09-29+10:00				
2006-10-02+10:00	Mailei, Kalfapūn	Flesaur and Flepōg, two natopu	erk	VU

Done

Viewing and concordancing...

Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://localhost:8080/exist/eopas/transcript.xq?id=18405154

eopas
EthnoER online presentation and annotation system

username
password
login

Home | Browse | Advanced Search | Upload | Help | About EthnoER

Search

Media Player



Display Options

☒ Interlinear ☐ Translation

Metadata

creator Nick Thieberger
contributor annotator: Nick Thieberger
coverage VU
subject (erk)
language erk
type linguistic-type: Primary Text
discourse-type: Narrative
date 2006-09-22+10:00

Concordance

...r Isiwer raki raki elau ,
...er siwer raki raki elau ,
...[mar] ilek wit pak elau , me inrik wit ki...
...apan panpan le pak elau , pan.
...upan rupan los los elau .
...upan rupan los los elau .
...fan gakit sao ntas elau .
...ti mai mai pak pak elau .
...l. Imurin na kefak elau .
...kor ipakor fat fat elau .
...afan lel tete naik elau ."
Elau imat, itrau mat ...
Naliati iskei, elau imat.
Me elau imat.

Downloads: Original | Eopas | CMML | Toolbox | Transcriber | Elan

Wit Go Kusu

s1 Amurin gag puserek, Nick, kafo gag
a= mur -i -n gag puserek Nick ka= fo ag
1sgRS= want -TS -3sgO 2sgBEN talk Nick 1sgIRR= PSP,IR sgBEN
pusereki kusu go wit.
puserek -ki kusu go wit
talk -TR rat and octopus

s2 Naliati iskei, elau imat.
naliati i= skei elau i= mat
day 3sgR.SBJ= one sea 3sgRS= low_tide

s3 Elau imat, itrau mat pe mat pe mat me.
elau i= mat i= trau mat pe mat pe mat me
sea 3sgRS= low_tide 3sgRS= very low_tide INT low_tide INT low_tide and

s4 Ipiatlak
nskau ienkot naur iskei
i= piatlak nskau i= enkot naur i= skei
3sgRS= have reef 3sgRS= lie_around island 3sgRS= one

s5 Naur ses iskei ito elau me
elau imat
naur ses i= skei i= to elau me elau i= mat
island small 3sgRS= one 3sgRS= stay sea but sea 3sgRS= low_tide

s6 Malen elau imat kerkrai,
malen elau i= mat kerkrai
as sea 3sgRS= low_tide strong

s7 mat go
top, nskau igar userek
mat top go nskau i= gar userek
low_tide big and reef 3sgRS= dry go_around

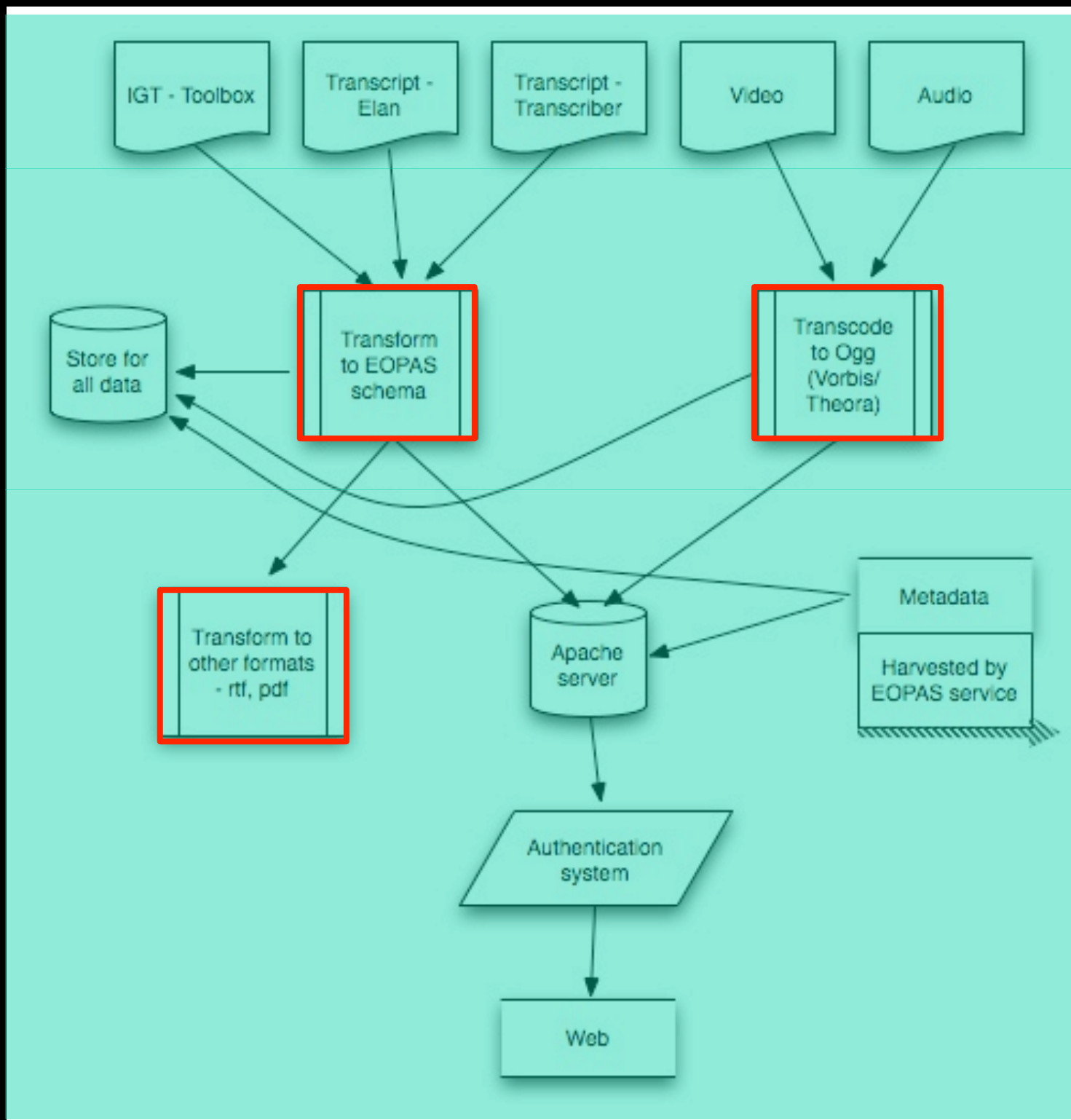
s8 pan tkal naur ses.
pan tkal naur ses
until touch island small

s9 Me kusu, kusu iskei, ito Efat
me kusu kusu i= skei i= to Efat
and rat rat 3sgRS= one 3sgRS= stay Efat

Done

EOPAS - <http://bit.ly/eopas2010>

- login functionality: authenticated access to upload/edit/download
- upload functionality of video/audio
- transcoding video/audio to Ogg Vorbis and Ogg Theora formats for delivery
- storage of video/audio
- download video/audio
- upload functionality of text (ELAN, Transcriber, IGT from Toolbox)
- handling of unicode and other i18n characters
- validation of text
- transcoding text
- storage of text
- download text in different formats (ELAN, pdf, rtf)
- aligned display of text with video/audio
- search
- citability on the level of syllables linked to video/audio
- editing (but not authoring!) of text
- open source publication of system
- creation of project in Sourceforge




```
<!DOCTYPE HTML> <html lang="en">
  <head> <meta http-equiv="Content-Type" content="text/html;
charset=UTF-8"> </head>
  <body> <div id="main">
    <video id="WitGoKusu.mp4" src="WitGoKusu.ogv" autobuffer
controls> </video>
    <div> <p class="segment"> <span class="ident"> </span> <span
class="text erk">
```

```
<a href="javascript:jumpToTimeoffset(4.877,10.963);"> Amurin gag
puserek, Nick, kafo gag pusereki kusu go wit.<i> I want to tell you,
Nick, I'll tell you about the rat and the octopus. </i></a>
</span> </p> <p class="segment"> <span class="ident"> </span>
<span class="text erk">
```

[...]

Confluence

Several projects are now aiming to produce textual material in small languages, e.g.,

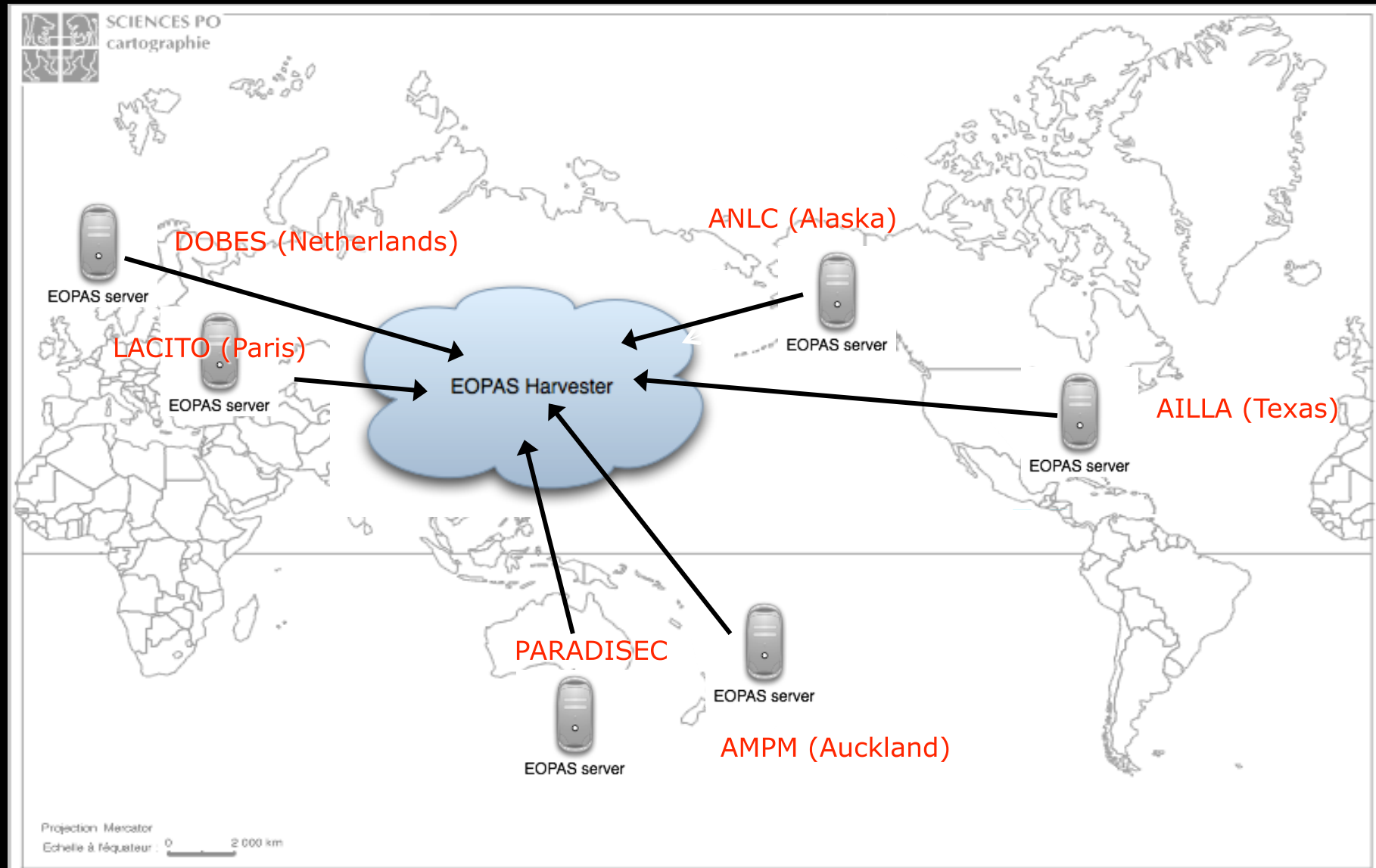
WOLP & Firebird

PARADISEC

AILLA

Potentially each archive and funding project will need to present texts and media

Central service harvesting text collections for search and display



Conclusion

Need for standards for encoding textual material as interlinear text linked to media

Ideally building these standards into a normal linguistic workflow

Creating multiple outputs from a single research effort

Exciting possibilities for return of material and for representation of small languages

How do we catch the wind?

Me, selwan ag kupi eñae, tiawi itraus traus traus traus, natrauswen ga itaos nlag. Itrausi pan kaipa. Me komam uta laap kin uto mau, a? Malen umat, inom.

But when you are far away the old man can talk and talk and talk, his story is like the wind. He tells it and it is gone. But there aren't many of us left. When we die, it will be finished.

†Kalfañun Mailei, 1998, Erakor Village, Efate, Vanuatu

