



Onderzoeksdata en toegangsrechten

Jacqueline Ringersma

Max Planck Instituut voor Psycholinguïstiek

Nijmegen



Max Planck Instituut voor psycholinguïstiek

Max Planck Gesellschaft

78 onderzoeksinstituten (Duitsland)

3 buiten Duitsland:

2 Italië (kunst)

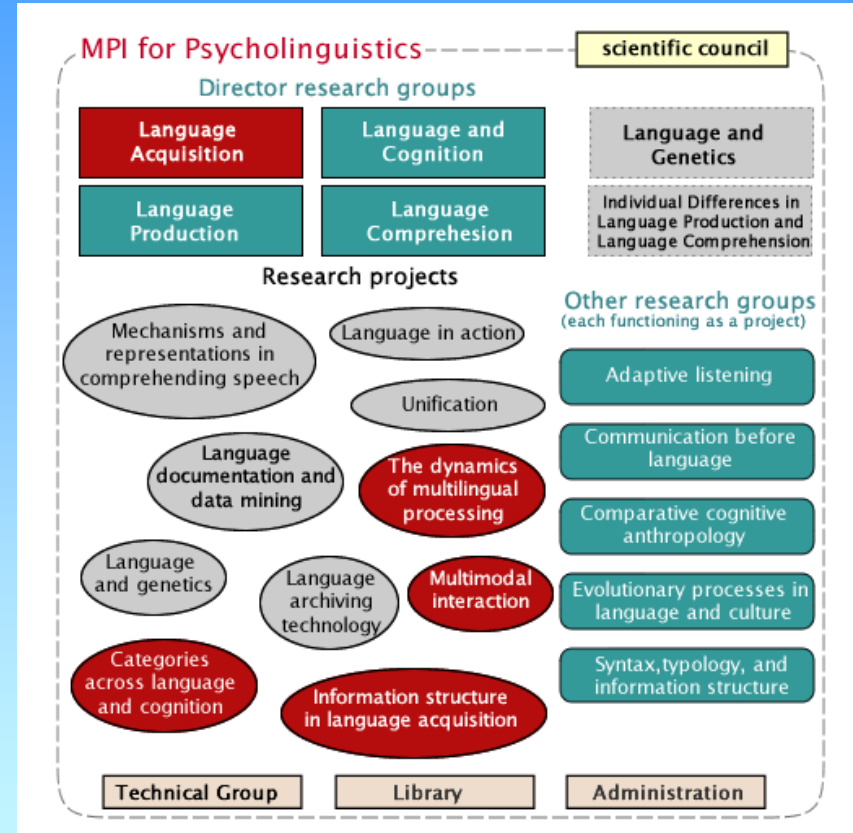
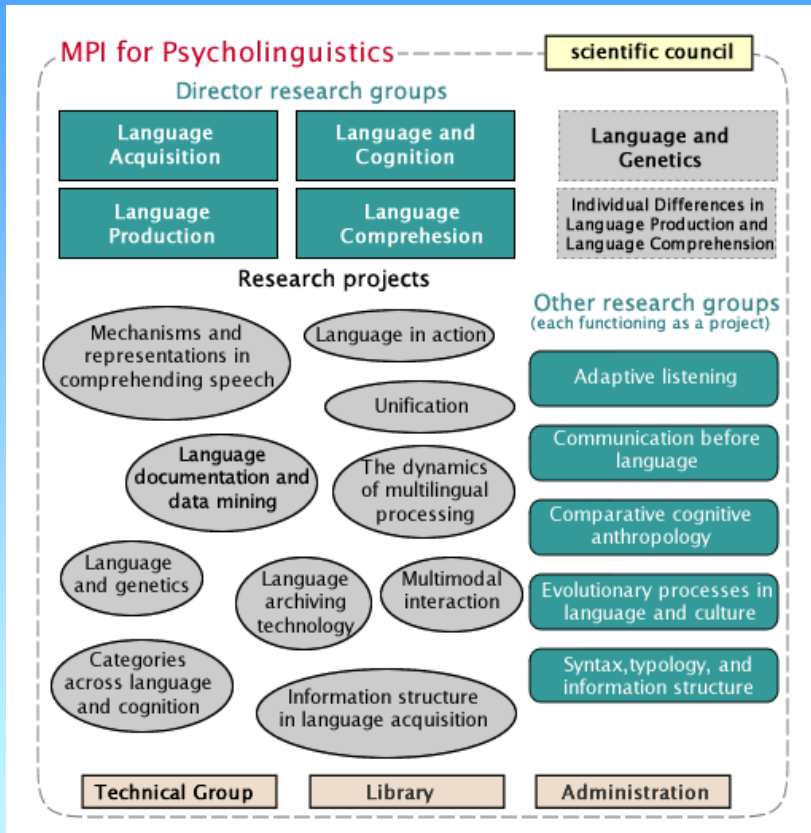
1 in Nederland (psycholinguïstiek)

Onderzoek naar cognitieve, mentale processen die gebruikt worden bij het produceren van taal en spraak, het begrijpen van taal en taalverwerving. Tevens wordt er onderzoek gedaan naar de relatie tussen taal, denken, gedachtegoed en cultuur.





Max Planck Instituut voor psycholinguïstiek

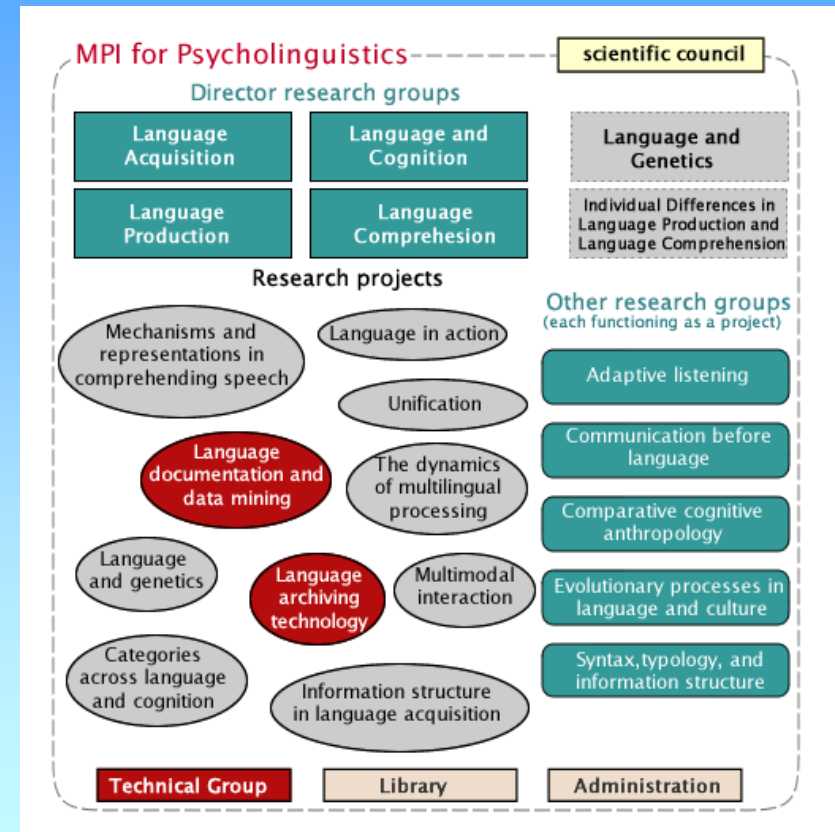




Technische Groep

1. Ondersteuning voor onderzoekers
2. Digitaal onderzoeksdata archief
3. Eigen onderzoek naar data-mining en onderzoeksinfrastructuur
4. Archief software ontwikkeling

Ongeveer 25 mensen, waarvan:
2 archiefbeheerders
2 archief software ontwikkelaars
10 LAT software ontwikkelaars





Onderzoeksdata in een archief?





Onderzoeksdata: Documentatie van (bedreigde) talen

Audio resources: weergave van gesproken taal

Video resources: informatie over de sociale en linguïstische context

Verrijkingen: Transcripties, vertalingen, woordenboeken etc.

DOKUMENTATION BEDROHTER SPRACHEN
DOCUMENTATION OF ENDANGERED LANGUAGES

DOBES

Search | Sitemap | Contact | Disclaimer | Copyright

Home > Documentation Projects > Beaver > Beaver - Data Examples

Beaver

[Language](#) | [Geography](#) | [People & Culture](#) | [Project](#) | [Team](#) | [Data examples](#) | [Links](#)

New: For an example of our clickable stories, click here.

audio file:
beginning of a personal narrative told by Albina Fournier

MyStoryBegin.mp3
transcript: MyStory.pdf

video file:
beginning of a story about a bad day told by Mary Francis

fisherhome.mov
transcript: fisher.pdf

video file:
beginning of a story about three girls who wanted to marry the stars told by Dominique Habitant

starshome.mov
transcript: stars.pdf

Home
DoBeS Programme
DoBeS Map
Documentation Projects
Archive Information
Archive Access
Archive Upload
Training Courses
Meetings
Ethical and Legal Aspects
Advisory Boards
Volkswagen Foundation
Regional Archives

Documents
Help and Advice
Links
Related Events
Language Archives
Tools
LAN (Archives
Newsletter)
DELAMAN Network
DAM-LR Project
MPI for Psycholinguistics
MPI Archive



MPI onderzoeksdata

Browsable Corpus
housed by the MPI for Psycholinguistics
(HTML version)

The Browsable Corpus gives you access to the IMDI-corpora housed at the MPI for Psycholinguistics. It allows you to browse through their hierarchical tree structure, whereby each node in this structure gives you access to further nodes. The nodes are characterized through metadata information, and they often have additional files associated with them: information files and, at the lowest level in the hierarchy, audio/video and annotation files. Metadata and information files are openly accessible, while audio/video and annotation files are usually protected for ethical and legal reasons. If you want to access protected files, please read the accompanying metadata files; they display the necessary contact information.

The Browsable Corpus is a preliminary test version for accessing XML-based IMDI-files with normal web-browsers. If you notice any errors, please contact the [Linguistics Group](#). Only browsers newer than Netscape 4.7 are supported.

If you don't see a tree on the left side of the page you may try our [advanced](#) version of the IMDI browser.

Beschreven Corpora



Multimedia lexica

(WORKSPACE) LEXUS-Lexicon Scheme Viewer

Sample lexicon for Rosset Demo 10/Feb/2006

Lexeme: *tiplu sing*

Class: 1 'sing sing' - see style of traditional song 2 assert, berde, 3 perform_singing

Example: 1 there's no singing; 2 tiplu we daa dookoo

Free Translation:

Note: 1. No clear relationship to berde! This song style contrasts with you, stand and other styles both musically and lexically. Tiplu was only sung by men and boys, in performances that last all night. Performers wear special grass skirts and carry spears, 2. includes cockroach, stag beetles, wasps, ants, mosquitoes, white ants (and anything with ... weel) but excludes bees, bees, cane-cane, 3.



Fotos

Workspace

File: r03_V20_v5.asf (MP12340424)

text | 14 | NONE controls | SW

Class: 1 | Note times: NONE

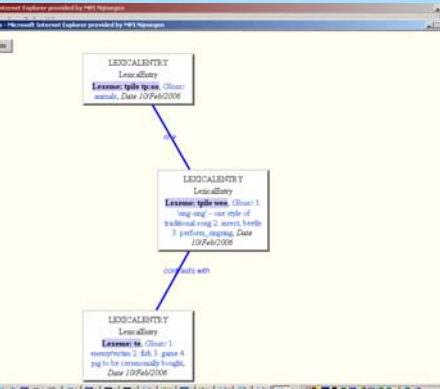
00:00:00.000 - 00:00:00.000

00:00:00.000 - 00:00:00.000

Just sing one! I also have never heard it (tricking?) **I was singing it** think it will go down here (in recorder) that's where you started a **beginning of the mbwaa** OK, you started to sing V-A; flower's opening at it. I also made a mistake because of such a crowd. It down please. I know this man I made a mistake, I went across. Because there were a lot of people overtake me I never finished it we wanted to see you, you sing many have you in memory? This one is second? You sing Myla's mbwaa that one. sing the beginning part, start at the beginning. I want sing it. people are coming closer. No, just sing it differently (not loud) you go away, you Yaliméni person OK, I haven't heard it, sing it in my hearing might sing it in church. You are "drinking/singing the Tiplu" pyu tiplu That the beginning of the V-A? That's the beginning of the mbwaa section (mbwaa) I started doing the Myla tiplu we again, because why? I might sing it now you say it properly, he might record it. He's going down (to forepart parts) He you drink V-A's mbwaa first. If I started taro piece and salt-water as complementary food; no other food have on this tiplu we. When Kiyé...

Geannoteerde media files

Typed relatives in de lexica





Onderzoeksdata archief

Waar wordt de data opgeslagen?

Digitale en online beschikbare databases/archieven:

DoBeS & MPI-archief, AILLA (Austin), Paradise© (Sydney)

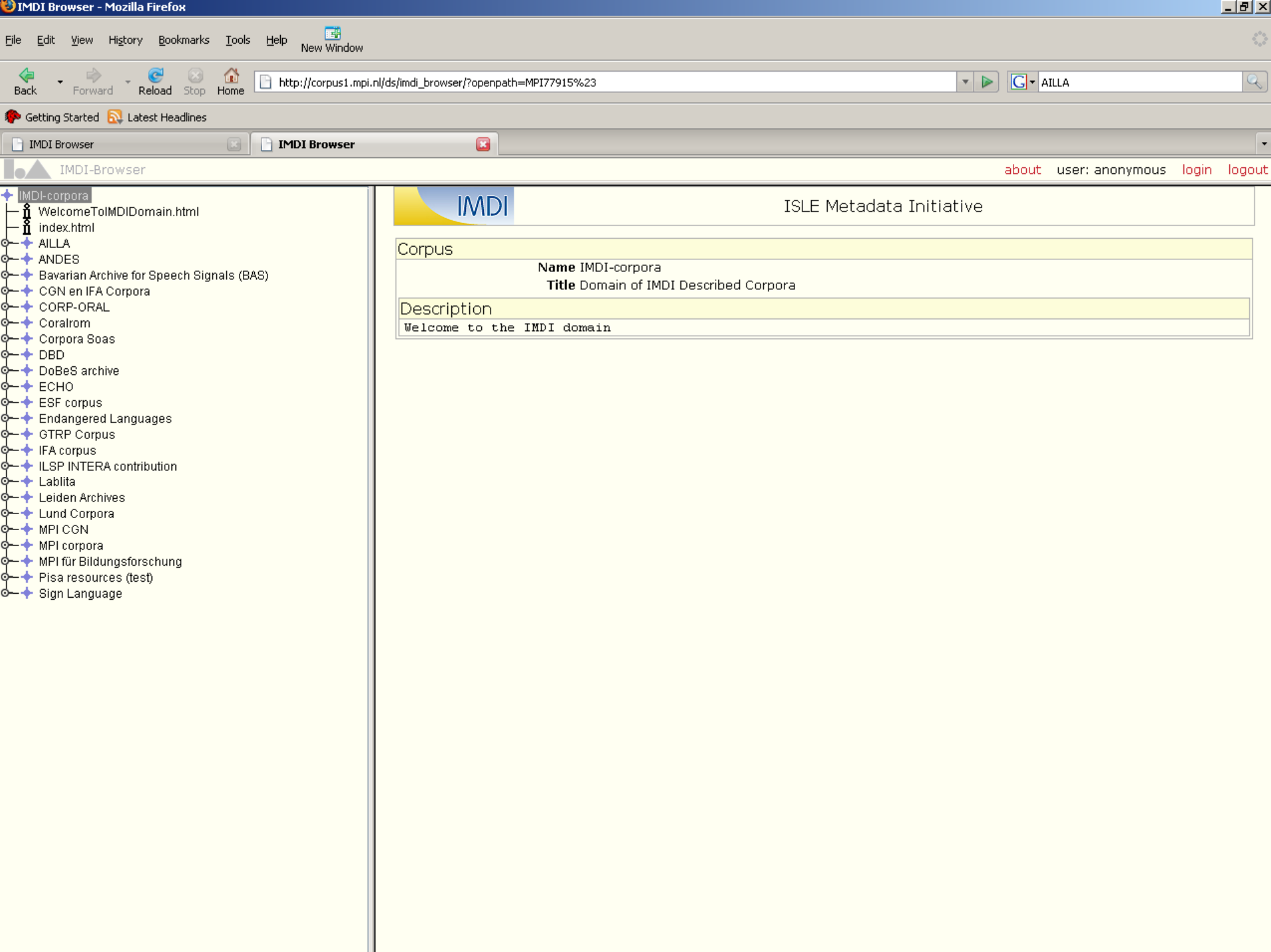
Archief voor linguïstische data (MPI en DoBeS)

Data in het archief

Meer dan 400.000 objecten, ongeveer 20 terabytes

Organisatie:

Metadata beschrijvingen, data base



- IMDI-corpora
 - WelcomeToIMDIDomain.html
 - index.html
 - AILLA
 - ANDES
 - Bavarian Archive for Speech Signals (BAS)
 - CGN en IFA Corpora
 - CORP-ORAL
 - Coralrom
 - Corpora Soas
 - DBD
 - DoBeS archive
 - ECHO
 - ESF corpus
 - Endangered Languages
 - GTRP Corpus
 - IFA corpus
 - ILSP INTERA contribution
 - Lablita
 - Leiden Archives
 - Lund Corpora
 - MPI CGN
 - MPI corpora
 - MPI für Bildungsforschung
 - Pisa resources (test)
 - Sign Language

IMDI ISLE Metadata Initiative

Corpus	
Name	IMDI-corpora
Title	Domain of IMDI Described Corpora
Description	
Welcome to the IMDI domain	



Onderzoeksdata archief

Voor wie is de data?

Onderzoeksteam

Collega onderzoekers

Algemene publiek – onderwijs, informatie

Taalgemeenschap – delen van kennis, herinvoering van de taal



Onderzoeksdata archief

Toegang tot de data

1. Browsing

Database weergave in de vorm van een online tree browser



Onderzoeksdata archief

Toegang tot de data

1. Browsing

Database weergave in de vorm van een online tree browser

2. Door middel van een search op de metadata

3. Door middel van een search in de content van de transcripties

- Corporal
- Corallrom
- Corpora Soas
- DBD
- DoBeS archive
- ECHO
- ESF corpus
- Endangered Languages
- DoBeS archive**
 - WelcomeToDOBES.html
 - introduction.html
 - Aweti
 - Bakairi, Katuyana and Mawe project
 - Beaver Archive
 - Cashinahua project
 - Chaco languages
 - Chintang and Puma Documentation Project
 - Chipaya project
 - Chontal
 - ECLING project
 - Enets and Forest Nenets
 - Gente del Centro - People of the Center
 - Gorani
 - Hocank team
 - Iwaidja team
 - Jaminjungan and Eastern Ngumpin
 - Kola Saami Documentation Project
 - Kuikuro Team
 - Lacandon Cultural Heritage
 - Marquesan team
 - Movima
 - Salar Monguor Project
 - Saiba/Logea
 - Savosavo
 - Semang
 - Semoq Beri and Batek
 - Sri Lanka Malay
 - Taa
 - Teop
 - Tima
 - Tofa team
 - Totoli
 - Trumai
 - Tsafiki project
 - Uru-Chinaya languages

Metadata search

Show occurrences

within 1 selected corpus: DoBeS archive [12372 sessions]

Key word search **Standard search** Advanced search

Search in all fields: fish

Search **Save** **Print** **Content search**

148 matches within 12372 selected sessions found.

< 1-50 >

Search complete

- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Cultural Data/Everyday Culture/in-house-05
- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Cultural Data/Food/buriti-trip
- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Cultural Data/Food/corn
- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Cultural Data/Food/fishing_086
- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Cultural Data/Food/fishing_trip
- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Cultural Data/Food/makaw04
- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Cultural Data/Food/manioc
- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Linguistic Data/Elicitation/113_birds
- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Linguistic Data/Elicitation/fish_names
- /IMDI-corpora/Endangered Languages/DoBeS archive/Aweti/From Language and Culture/Linguistic Data/Monological/Explanations/118_fishing
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/grammar/TAM/perfectives004
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/topical dictionary/bodyparts_lex001
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/verbs/paradigms/paradigm_clean_fish001
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/verbs/paradigms/paradigm_kill_fish001
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/word lists/animals/fish001
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/word lists/animals/fish002
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/word lists/animals/fish003
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/word lists/animals/fish004
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/word lists/animals/fish005
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/word lists/animals/fish006
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/word lists/animals/fish007
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/word lists/hunting, weapons & tools/hunting005
- /IMDI-corpora/Endangered Languages/DoBeS archive/Beaver Archive/Data/linguistic/elicited/word lists/hunting, weapons & tools/hunting007

Domain: YeleAnnexDemo

EAF (3)

Find

#hits	:	5	<div style="width: 100%; height: 10px; background-color: blue;"></div>
#annotations with a hit	:	5	
#annotations investigated	:	5260	

Ready

Options

Hit 1 - 5 of 5 hits

*(touches temple) (sings verses) (sings **tpile wee**) mo yi mo mm*
ee! mwada yii lili nyi yââmênê pi (tpile wee) (tpile wee sung) nyââ, kn:ââ, mbwaa kn:ââ u ngwo a nyi chapî mbwaa kn:ââ
mwada yii lili nyi yââmênê pi (tpile wee) (tpile wee sung) nyââ, kn:ââ, mbwaa kn:ââ u ngwo a nyi chapî mbwaa kn:ââ (sings tpile wee)
(tpile wee sung) nyââ, kn:ââ, mbwaa kn:ââ u ngwo a nyi chapî mbwaa kn:ââ (sings tpile wee) wu n:i u pwo a kwo, ala d:oo mbwaa u kada kada ngee wowo, awêde ala ngwo anê dy
 or of Tada, but passed it on to Andrew. Note that although associated with a clan, K.Kmiyé, the **tpile wee** expertise can pass through the male line (Willy's F was Andrew's a michó) Although



Text

Grid

Subtitle

Timeline

Video display

min



Media information

min

Resource: demo_r03_v20_s5.eaf
Media file: demo_r03_v20_s5.m...

Elapsed time: 00:01:24:000

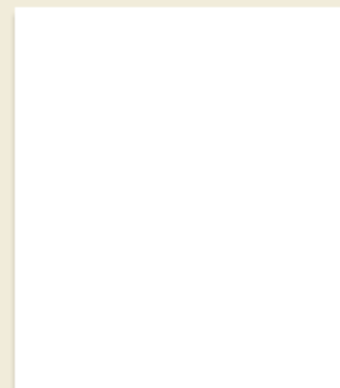
Selected chunk:

Begin time: -

End time: -

Text: -

Mini Data Frame



Tier: none

Font size: 14

Play selection

Clear selection



Play screen by screen

Play continually

Tier text font:

Arial Unicode MS

Timeline

Play video

Gloss1

Gloss2

Kp

Pius

K

Ricky

man outside view 2

(sings tpile wee)

|ee?|



Onderzoeksdata archief

Toegang tot de data

1. Browsing

Database weergave in de vorm van een online tree browser

2. Door middel van een search op de metadata

3. Door middel van een search in de content van de transcripties

4. Via een Google overlay

5. Via speciale webportals

Search

Fly To Find Businesses Directions

From e.g., Utica, NY

To e.g., Tucson, AZ

"Tsafiki" lat=-2.86703106042, lon=

Places Add Content

- [Yéli Dnye](#)
- [Yucatec](#)
Yucatec (Yucatec Maya) is a Mayan
- [Language Acquisition](#)
- [Chinese](#)
The Chinese acquisition is part of
- [Czech](#)
L2 (second language) acquisition
- [Dutch](#)
L1 (first language) acquisition of Dutch:
- [English](#)
L2 (second

Layers

- Primary Database
- Geographic Web
- Roads
- 3D Buildings
- Street View
- Borders and Labels
- Traffic
- Weather
- Gallery
- Global Awareness
- Places of Interest
- More
- Terrain

Google Earth

File Edit View Tools Add Help

Search

Fly To Find Businesses Directions

From e.g., Utica, NY

To e.g., Tucson, AZ

"Tsafiki" lat=-2.86703106042, lon=

Yéli Dnye

Pioneers of Island Melanesia -

Yéli Dnye (also known as Rossel, Yela, Yele, Yelejong, Yeletnye) is a Papuan language spoken on Rossel Island, Louisiade Archipelago, Papua New Guinea. Although surrounded by Austronesian languages, Yéli Dnye shows little evidence of influence by them, making this language an isolate. Yéli Dnye is known as the language with the world's most complex phonemic inventory.

Google Earth

File Edit View Tools Add Help

Search

Fly To Find Businesses Directions

From e.g., Utica, NY

To e.g., Tucson, AZ

"Tsafiki" lat=-2.86703106042, lon=

Yéli Dnye

Pioneers of Island Melanesia -

Yéli Dnye (also known as Rossel, Yela, Yele, Yelejong, Yeletnye) is a Papuan language spoken on Rossel Island, Louisiade Archipelago, Papua New Guinea. Although surrounded by Austronesian languages, Yéli Dnye shows little evidence of influence by them, making this language an isolate. Yéli Dnye is known as the language with the world's most complex phonemic inventory.

Project leader is [Stephen Levinson](#).

[browseable corpus link](#)
[LEXUS example](#)
For this demonstration of LEXUS use "demo" as username and "demo" as password.
[ANNEX example 1](#) [ANNEX example 2](#) [ANNEX example 3](#)
[Pioneers of Island Melanesia webpage](#)

12°47'10.20" S 154°13'44.59" E

ANNEX Manual

user: anonymous logout

File: demo_r03_v19_s2.eaf

text 14 hide controls hide media hide QT CSV

tsGest NONE NONE

hide times

00:00:00.000 - 00:00:00.000 00:00:00.000

No media? Click [here](#) for possible reasons.

http://corpus1.mpi.nl/ids/annex/runLoader?modeid=MP1361092%23

Taa Portal

A Pan-Dialectal Documentation of Taa

[Home](#) [About](#) [Culture](#) [Histories](#) [Home](#) [Mapping](#) [Stories](#)

About

This portal was developed by the **MPI for Psycholinguistics** in collaboration with the **Taa DOBES documentation project**. The portal makes use of the **Wordpress** CMS and a REST interface to the **IMDI metadata catalog**. The graphical design is based on the patagonia theme designed by **Wp Themes Planet**.

By **admin** June 23, 2009

Comments are closed.

SEARCH

PAGES

- [About](#)
- [C](#) [About](#)
- [Histories](#)
- [Home](#)
- [Mapping](#)
- [Stories](#)



Toegangsrechten

Issues in het debat over toegangsrechten

(Cultureel) gevoelige data

Van wie is de data (Intellectual property rights – IPR - debate)

Onderzoeksdoeleinden

Nationale en internationale regelgeving

Gedragcodes



AMS – Access Management System

Op alle resources in de database kunnen onderzoekers zelf de toegangsrechten zetten:

Per gebruiker of groep gebruikers

Per resource type (video, audio, annotatie, image, info file)

Met beperkte of oneindige geldigheid

Met of zonder gedragscode ter ondertekening

AMS is online, bereikbaar vanuit de browser



AMS – Access Management System

The screenshot shows the AMS web interface. On the left is a tree view of nodes, with 'Marquesan team' selected. The main content area displays the 'Node Authorization Overview' for 'Marquesan team'. It includes a navigation menu, a dropdown for selecting a user/group, and a table of rules for the selected node. The table lists users and their assigned roles and permissions.

IMDI Browser
IMDI-Browser about user: jacrin@mpi.nl

Welcome to AMS
- lat access management system -

User Management | Node Authorization | My Account | Mailing | ?

Node Authorization Overview

- view access privileges to LAT Resources 'Marquesan team: Marquesan team' -

Focus overview on selected User/Group:

Rules of Node 'Marquesan team'

- /MPI0#/MPI239013#/MPI77915#/MPI80685# -

There are no licenses assigned to Node 'Marquesan team'

[Manage Node Licenses](#)

Iris Helming	<ul style="list-style-type: none">Domain ManagerDomain EditorRead Info Files AllowRead Annotations AllowRead Images AllowRead Audio Files AllowRead Video Files Allow	Edit
Gaby Cablitz	<ul style="list-style-type: none">Domain CuratorDomain ManagerDomain Editor	Edit
marquesan_group	<ul style="list-style-type: none">Read Info Files AllowRead Annotations AllowRead Images AllowRead Audio Files AllowRead Video Files Allow	Edit
Mona Wich	<ul style="list-style-type: none">Domain ManagerDomain Editor	Edit

[Add new Rules](#)
[Force Exnort](#)



AMS – Access Management System

Voor wie is de data?

Onderzoeksteam

Collega onderzoekers

Algemene publiek – onderwijs, informatie

Taalgemeenschap – delen van kennis, herinvoering van de taal

AMS werkt voor:

Onderzoeksteam

Collega onderzoekers

Taalgemeenschap – delen van kennis, herinvoering van de taal

Maar niet voor:

Algemene publiek – onderwijs, informatie



RRS – Resource request system

Toegangsregeling voor algemeen publiek

1. X bezoekt browser en via metadata vindt hij een interessante resource
 2. Maar hij heeft geen toegang 
 3. Hij vraagt toegang aan via een web formulier
-



RRS – Resource request system

Opdracht: Maak een flowchart voor de (geautomatiseerde) afhandeling van een verzoek tot toegang tot een 'single' resource

1. Wat willen we van de aanvrager weten?

Naam, organisatie, waarom wil hij de resource bekijken, hoe lang wil hij toegang, is aanvrager al een bekende van AMS?

2. Wat weten we van de resource?

Unieke en persistente identifier (handle)

Eigenaar (met e-mail adres)

Gedragscode op de resource



RRS – Resource request system

3. Afhandeling (waar je aan moet denken)

Controle of gebruiker al bekend is in AMS en of gebruiker al de juiste gedragscode getekend heeft (heeft gebruiker wellicht al toegang?)

of gebruiker moet worden toegevoegd aan AMS: een username en een pw toegekend krijgen/of zelf laten kiezen (?)

Gebruiker moet ergens in het proces de gedragscode tekenen (en dit moet worden opgenomen in AMS zodat bij een volgende aanvraag niet opnieuw een code moet worden getekend)

Eigenaar van de resource moet toestemming geven en mogelijk extra informatie vragen aan de gebruiker

Rechten in AMS moeten worden gezet

Database moet worden ge-update voor zover het de toegangsrechten betreft.



RRS – Resource request system

3. Afhandeling (waar je aan moet denken)

Let op:

Sommige resources zijn open, bij dergelijke resources moet RRS niet in actie komen

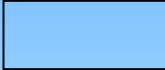
Andere resources zijn semi-open, de aanvrager moet wel de gedragscode tekenen, maar heeft daarna toegang zonder verdere toestemming


Tijdens de afhandeling moet de aanvrager tijdig en redelijke feedback ontvangen van het RRS

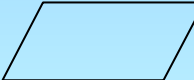


Flowcharts

Start en einde: cirkel, ovaal 

Proces: rechthoeken 

Beslissing: Ruitvormen 

Input/output: parallellogrammen 

Flow richting: pijlen 