Data exchange for web site content and resource sharing

Jacqueline Ringersma
Karin Kastens
Jos van Berkum

Max Planck Institute for Psycholinguistics
Renewing the MPI-PL web site

“bigger, better, MPG greener ..”
“structure the content”
“more user friendly”

“consistent information”
“interaction with data repositories”
Web site in Plone CMS

Content Management System

Collaborative environment

Differentiated user roles

Easy workflow

Content control system

Adjustable content types

‘Easy’ interaction with other applications

Some in-house Plone experience

Trusted company (Zest)

Open source
A ‘Plone’ website:

A set of ‘content types’ structured together:

- folders
- pages
- news items
- events
- collections, collages

with templates (styles) for every content type

and… filled with content
Special content types for organizational units:

person, project, group
Content types

Special content types for research output:

publication lists and presentation lists
Data exchange to fill CT from other applications:

MPI-PL people data base
MPDL Publication repository PubMan

Data exchange for primary research data:

MPI-PL primary data archive

Background:

Consistency of information
Single data entry in specialized application
Information can be retrieved on multiple web pages
Standard view and display
Data exchange

MPI primary research archive
- Primary data
- Secondary data

MPI people database

PubMan
- Book
- Chapter
- Journal article
- Journal
- Poster
- Proceedings
- Proceedings paper
- Report
- Series
- Special issue
- Talk
- Teaching
- Thesis
- Working paper

www.mpi.nl
- Person pages
- Project pages
- Group pages
From people database to person page

For every staff member in the MPI-PL

Standard home page: information retrieved from the MPI Oracle database

Standard folder structure (Publications, Presentations, Research, Teaching, Tools, Press, CV)

Free text
From people database to person page

Max Planck Institute for Psycholinguistics

Jos van Berkum

The overall goal of my research is to understand the nature of sentence- and discourse-level language comprehension, with a specific focus on how people relate the unfolding utterance to its wider context in real time. Topics include establishing coherence in context, the anticipation of upcoming meanings, the language-grammar interface, mood-based modulation of language processing, the interaction of static and linguistically configured knowledge, and pragmatic reasoning. I'm interested in the functional architecture of the systems involved as well as in their realization in the brain. Because of this, I use both classical but proven techniques from experimental psycholinguistics (e.g., self-paced reading, questionnaires) and methods from cognitive neuroscience (EEG, MEG, fMRI).

Email: Jos.vanBerkum@mpi.nl
Function: Research Project Coordinator
MPI Group: Language Production Group
Affiliations: Néerlandais Centre for Cognitive Neuroimaging

Last checked 2009-07-19 by Jos van Berkum
The overall goal of my research is to understand the nature of sentence- and discourse-level language comprehension, with a specific focus on how people relate the unfolding utterance to its wider context in real time. Topics include establishing reference in context, the anticipation of upcoming meaning, the language-valence interface, mood-based modulation of language processing, the interaction of static and linguistically configured knowledge, and pragmatic reasoning. I am interested in the functional architecture of the systems involved as well as in their realization in the brain. Because of this, I use both classical and modern techniques from experimental psycholinguistics (e.g., self-paced reading, questionnaires) and methods from cognitive neuroscience (EEG, fMRI, TMS).
Person pages are created on the basis of the Oracle data base field ‘name’.

Process is controlled by a python script.
Multi media from MPI-PL research data archive

Research (primary) data
Online available
Each file has:

PID (handle)
access right setting
Multi media from MPI-PL research data archive

Direct links from web page to data file

Alternative: supplementary material to publications,
→ PubMan
background

- Early adopter for PubMan since 2007
- First MPI to be migrated from eDoc → PubMan
- Live with PubMan since April 2009
Workflow at MPI Psycholinguistics

**ONE** entry point for publications,

**MULTIPLE** outlets

Researchers, secretaries, or librarians enter publication data AND upload full texts

LIBRARIANS perform quality check before ‘releasing’ item for world-wide visibility

Metadata always visible

Visibility of full texts can be set to public, private, or restricted
Objective: Standardized layout of publication lists

Standardized listings for

Person pages

Project pages

Informal verbal interaction is the core matrix for human social life. A mechanism for coordinating this basic mode of interaction is a system of turn-taking that regulates who is to speak and when. Yet relatively little is known about how this system varies across cultures. The anthropological literature reports significant cultural differences in the timing of turn-taking in ordinary conversation. We test these claims and show that in fact there are striking universals in the underlying pattern of response latency in conversation. Using a worldwide sample of 10 languages drawn from traditional indigenous communities to major world languages, we show that all of the languages tested provide clear evidence for a general avoidance of overlapping talk and a minimization of silence between conversational turns. In addition, all of the languages show the same factors explaining within-language variation in speed of response. We do, however, find differences across the languages in the average gap between turns, within a range of 250 ms from the cross-language mean. We believe that a natural sensitivity to these tempo differences leads to a subjective perception of dramatic or even fundamental differences as offered in ethnographic reports of conversational style. Our empirical evidence suggests robust human universals in this domain, where local variations are quantitative only, pointing to a single shared infrastructure for language use with likely ethological foundations.

Link to source (DOI) >
Downloadable fulltext >

Supplementary materials
Stivers_2009_universals_suppl.pdf
PubMan item view – adding content to PubMan

Short view of an item with
one file and one locator

<table>
<thead>
<tr>
<th>State</th>
<th>Medium View (all Items)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Differential roles for left inferior frontal and superior temporal cortex in mu...</td>
</tr>
<tr>
<td>Creator(s)</td>
<td>Willems, R. M.; Ozyurek, Asli; Hagoort, Peter</td>
</tr>
<tr>
<td>Date(s)</td>
<td>Accepted: 2009</td>
</tr>
<tr>
<td>Genre</td>
<td>Journal Article</td>
</tr>
<tr>
<td>File(s)</td>
<td>1 File</td>
</tr>
<tr>
<td>Locator(s)</td>
<td>1 Locator</td>
</tr>
</tbody>
</table>
**Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language**

**Contact**
Moderator

**Files**

<table>
<thead>
<tr>
<th>Name</th>
<th>Willems_2009_NeuroImage.pdf (Publisher version) [MD5]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>Visibility</td>
<td>Public</td>
</tr>
<tr>
<td>MIME-Type / Size</td>
<td>application/pdf / 2MB</td>
</tr>
</tbody>
</table>

**Locators**

<table>
<thead>
<tr>
<th>Locator</th>
<th>Supplementary table S1 (Supplementary material)</th>
</tr>
</thead>
</table>

**Content category**
Supplementary material

**Full text uploaded into PubMan**

**Link to supplementary material**
.. Authors and affiliations

MPI department

MPI project
.. Via REST interface

![eSciDoc SearchAndExport Service](image_url)

**Description**

**REST Example**

**Search Query Form**

CQL search query: `esicdoc.metadata="test" and esicdoc.content-model.objid=esicdoc:persistent4`

**Indexes**

You can find here all indexes are allowed.

**Sorting**

Sorting Key

Sorting Order: ascending

**Record Span**

Start Record (default is first)

Maximum Number of records (default is 1)

**Database Index**

Language: all languages

**Export Options**

Export Format: APA

Target URL: `http://pubman.mpdl.mpg.de/search/Se...`

**Output Format**

PDF

HTML

RTF

ODT

SNIPPET
Exporting of PubMan Data for website

.. Export url


... produces xml

.. Snippet format

dcterms:bibliographicCitation xmlns:dcterms="http://purl.org/dc/terms/"<!--[CDATA[<span class="Default">Cutler, A., Otake, T., &amp; McQueen, J. M.<span class="DisplayDateStatus"> (2009).</span> Vowel devoicing and the perception of spoken Japanese words.</span><span class="Italic">&nbsp;The Journal of the Acoustical Society of America</span><span class="Default">, </span><span class="Italic">125</span><span class="Default">(3), 1693-1703.</span>]]></dcterms:bibliographicCitation>
PubMan url

Default:
Nightly full import

Ad hoc import – urgent matters

Pubman import

~ Up to Site Setup
Publications and presentations are put in the 'publications' folder in the root of the website. This folder is emptied and refreshed nightly. When it does not exist (anymore) it is created.
The default url can be changed in the pubman settings control panel.
Go to the publications folder.

Full import
Warning: Using this import will remove all existing publications and presentations and import them fresh using the given url.

URL to the pubman data
The default url that is given imports all publications. This is done nightly when the cronjob has been turned on. The default url can be changed in the pubman settings control panel.
http://pubman.mpdl.mpg.de/search
/SearchAndExport?q=escidoc.content-model.oid=escidoc:permanent%20AND%20escidoc.context.oid%20AND%20escidoc.content.sortKeys&exportFormat=APA&outputFormat=snippet&language=all&sortKeys=
sortOrder=ascending&startRecord=0&maximumRecords=5

Import Data

Single publication import
Import a single publication or presentation. If the publication or presentation already exists on this website, it is removed and then added with the new data. This is non-destructive: all other publications and presentations are left untouched.

Escidoc object id [escidoc]: Import Single Item

Non-destructive import
Do a non-destructive import. This only adds or updates publications and presentations, and -- unlike the full import -- does not throw anything away. You can use this for importing say the changes of the last two weeks.

URL to the pubman data
Import Data
Publication metadata (reference, abstract, links) are stored in CMS.
Each publication is stored in a text string via python script.
Publication lists are selected by criteria

- Author name
- Organizational unit (departments, projects)
- Publication type
- Year or publication status
- Local tag (used to select special publications OR exclude publications)
Selection by author name
Selection by project name
in press


Published articles

by rion — last modified Apr 23, 2008 05:48 PM

Journal article

- Dimroth, C. (2009). L’acquisition de la finitude en allemand L2 à différents âges. AILE (Acquisition et Interaction en Langue étrangère)/LIA (Languages, Interaction, Acquisiton), 1(1), 113-135. more >
- Dimroth, C. (2003). Age effects on the process of L2 acquisition? Evidence from the acquisition of negation and finiteness in L2 German. Language Learning, 50(1), 117-150. more >

Key: By year

- sub-type: in press
- publication list: view by status and publication year
- view: sub-type, display, add publication list, state, published

Key: By type

- sub-type: Journal article
- publication list: view by status and publication year
- view: sub-type, display, add publication list, state, published

Institution: Max Planck Institute for Psycholinguistics

Karin | Site Setup

PO Box 310
6500 AH Nijmegen
The Netherlands

Phone: +31-24-3521443
Fax: +31-24-3521213
Room: 245
Manage profile
Conclusions

- Central data repositories
  - Institute people data base
  - Primary data archive
  - PubMan

- Complex information bundles

- Consistent standardized and formatted output

- Multiple output portals: person, project & group

- No heavy load in CMS of web site (speed)

- Website is an incentive for researchers to use PubMan or MPI-PL archive
Presentation lists on the web site

Do you want to use our python script? Check our

http://www.mpi.nl/people/kastens-karin/presentations

http://www.mpi.nl/people/ringersma-jacquelijn/presentations