

StressTyp2: A database for the accentual patterns in the world's languages

Rob Goedemans (Leiden)

Jeffrey Heinz (Delaware)

Harry van der Hulst (UConn)

University of Leiden, Netherlands

August 16, 2014

Support from NSF#1123692 and #

Overview

- Goals
 - Facilitate typological studies of stress and accent
 - Accessibility, extensibility and flexibility
 - A key to the literature (not an imposition of our own views)
- Data from 699 languages (859 lects)

A Brief History

1994 StressTyp launched by Goedemans and van der Hulst. They have updated and expanded it continuously since then.

2007 Heinz combined Bailey's 1995 database with Gordon's 2002 database to create the Stress Pattern Database.

2010 van der Hulst hosted a conference at UConn to initiate the StressTyp2 project, whereby researchers would combine and coordinate efforts.

2011 NSF support for the StressTyp2 project was obtained.

Today's 'interactive' demo

1. Browsing the database via the web
2. Searching the database via the web
3. The structure of the ST2 database
4. Archiving versions with the Linguistic Data Consortium
5. Future work to do
6. People deserving much thanks

Browsing the database via the web

<http://st2.ullet.net/stresstyp2>

Browsing the database via the web

- Lect and pattern centric
 - Language/lect distinction
 - Distinctions between primary, secondary, exceptional, and compound stress
 - Distinctions between dominant and subordinate stress patterns.
 - StressTyp Codes and Syllable Priority Codes

Some features for pattern description

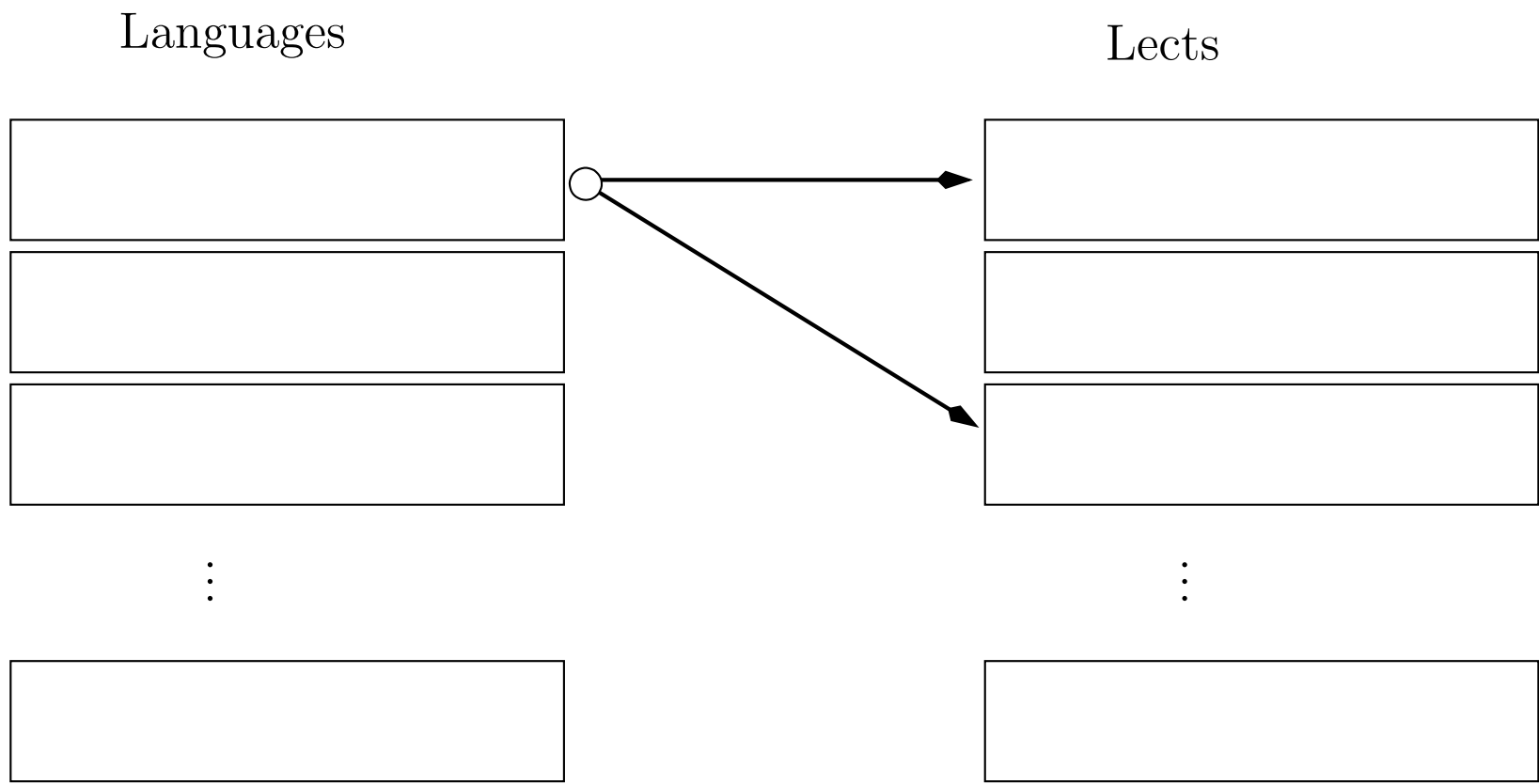
- English prose
- Theoretical descriptions
- Computational descriptions and implementations for 'in toto' patterns

Searching for patterns and descriptions

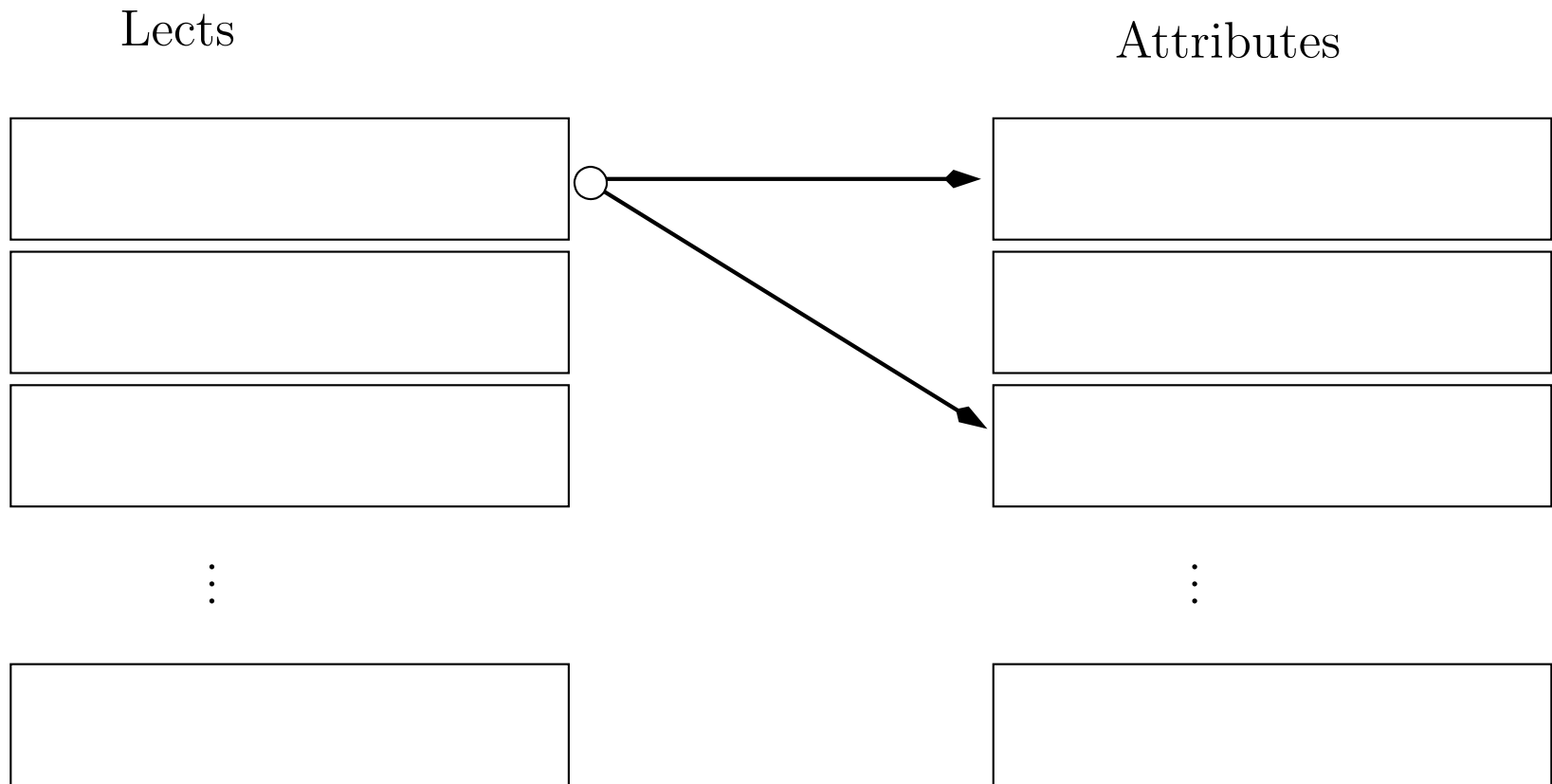
- Quick search allows simple queries for lects and/or patterns
- Advanced queries requires some knowledge of the structure of the db

Structure of the ST2 database

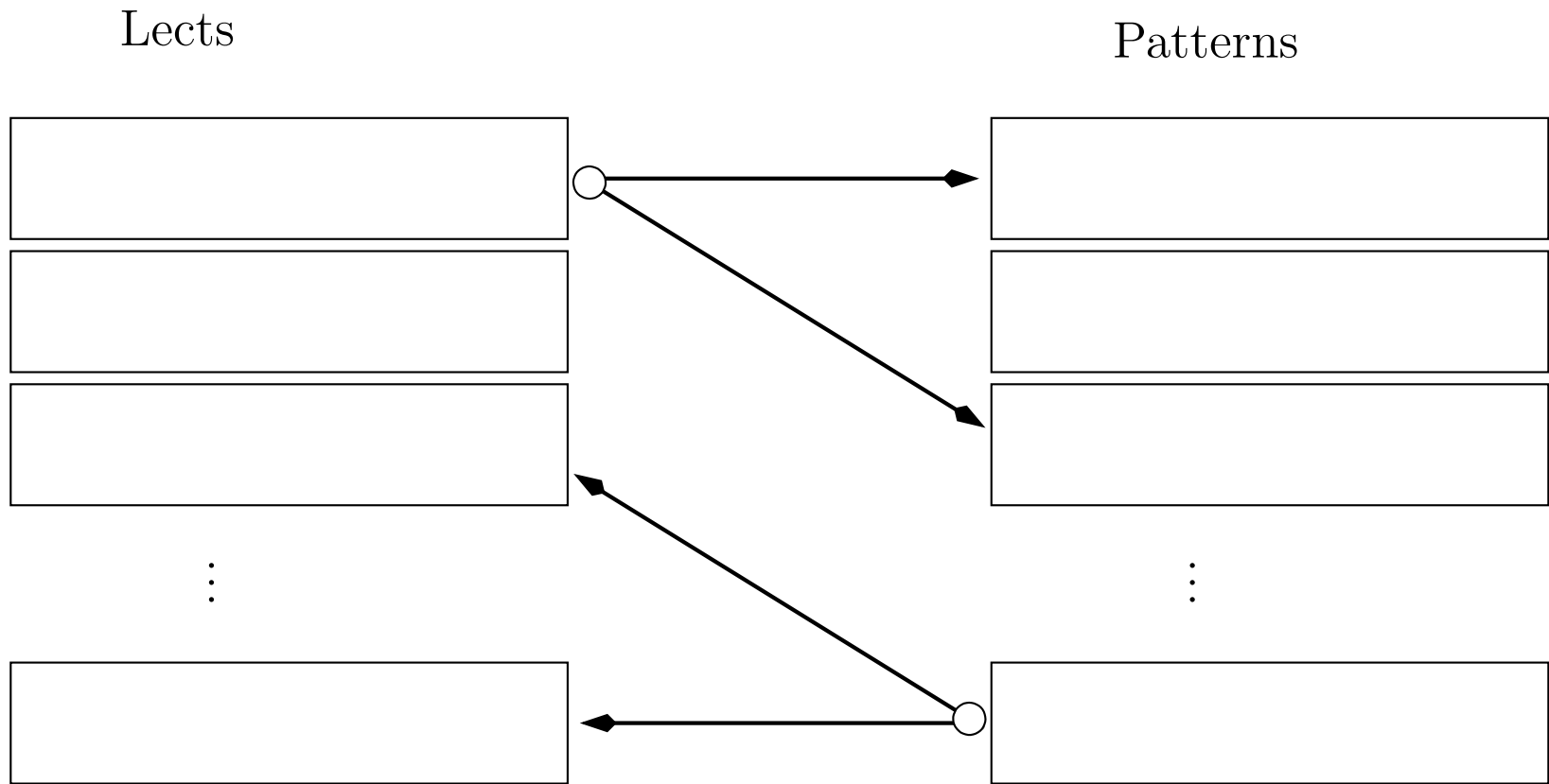
- MySQL relational database



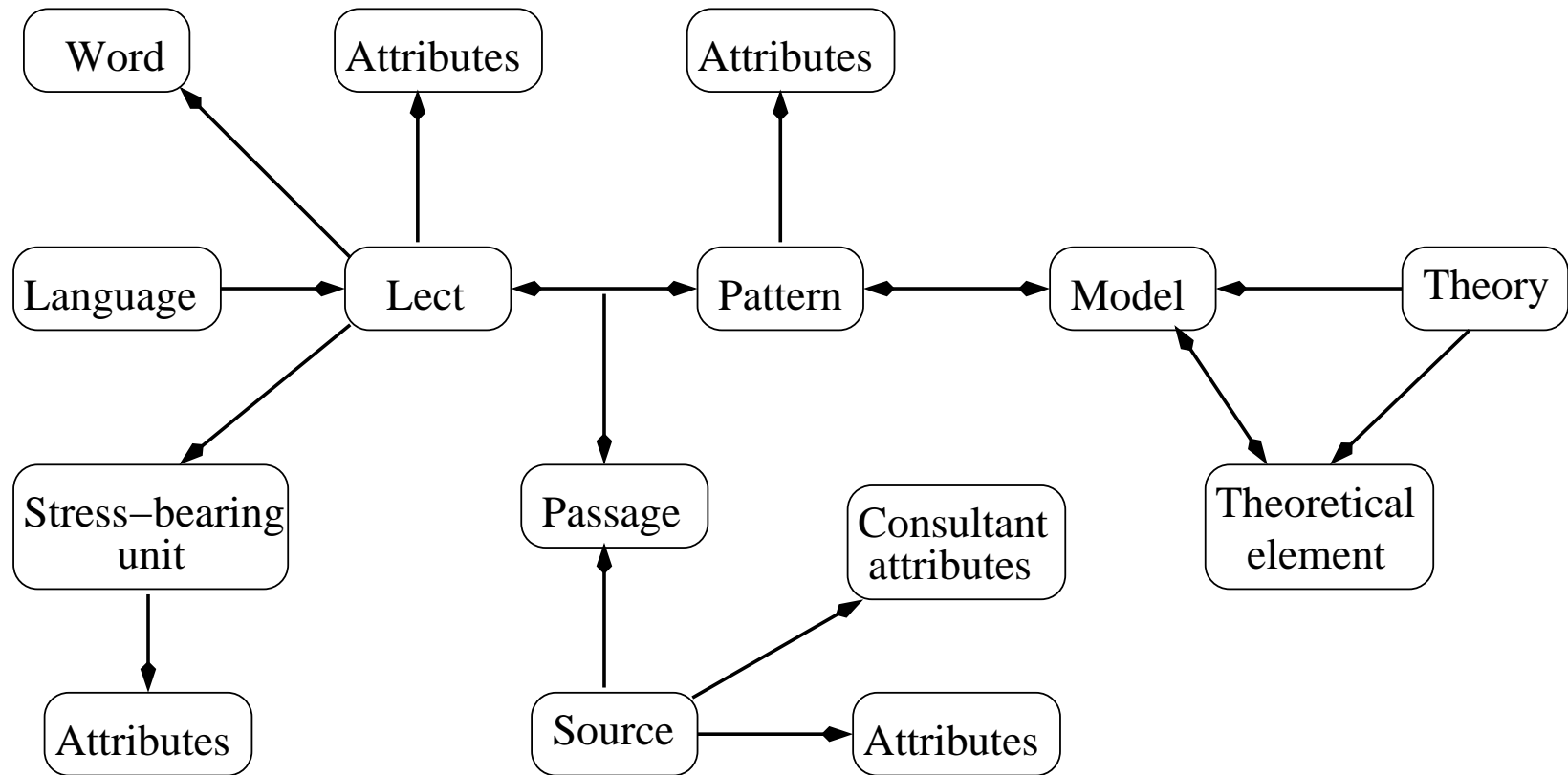
Structure of the ST2 database



Structure of the ST2 database



Structure of the ST2 database



What will be archived with LDC

- The SQL database itself
- A list of useful queries and their outputs
- Transducers for the in toto patterns.
- Documentation explaining the structure of the database.

Future efforts – Add more information

- More languages, lects, and patterns
- Source quality
- Phonetic attributes
- Morphological factors
- ...

Future efforts – Web and DB developments

- Some dialects are listed with a “(SPD)” or “(STC)” suffix
- Many languages are missing ISO codes
- Excerpts are not yet displaying
- Notes conflate developer issues with public notes and so are confusing
- Need to address the issue of multiple lect names
- Put information into the fields which qualify nature of source
- Currently, disagreeing authors cannot be displayed within the same lect
- ...

People deserving much thanks

Web developers

- Coen Zimmerman (Leiden)
- Gordon Hemsley (Delaware)
- Manex Agirrezabal (Basque Country, visiting UD)

Data analysis and management

- Adam Jardine (Delaware)
- Amanda Payne (Delaware)

Data collection and analysis

- Aida Talic (UConn)
- Abigail Thornton (UConn)

Thank you for your interest and feedback!

Please email us with comments, problems, and questions.

`stresstyp2@gmail.com`

`R.W.N.Goedemans@hum.leidenuniv.nl`

`heinz@udel.edu`

`harry.van.der.hulst@uconn.edu`