UPCOMING RESOURCE
AN ONLINE LANGUAGE CHANGE DATABASE

UNIVERSITY OF HELSINKI
WHO WE ARE

• Reassessing language change: the challenge of real time
  • Academy of Finland funded project, 2014–2018

• PI **Terttu Nevalainen**, VARIENG Research Unit
  • Postdocs Tanja Säily & Turo Vartiainen

• Collaborators **Peter Trudgill** (sociolinguistics),
  Jefrey Lijffijt & Jukka Suomela (computer science)

UNIVERSITY OF HELSINKI
PROBLEM: FRAGMENTED RESEARCH

• Access to data & previous research on language change

• Knowledge of historical data

• At present, research on real-time language change is largely not cumulative
• Comparative, real-time baseline data for modelling change in progress

• Language Change Database (LCD)

• Open access, part of Varieng e-resources
Socio-linguistic typologies (Trudgill 2011)

Baseline for other studies (Tagliamonte & D’Arcy 2009)

Replication with other data sets

Statistical modelling (Blythe & Croft 2012)

Systematic reviews, meta-analyses

UNIVERSITY OF HELSINKI
WHAT THE DATABASE WILL CONTAIN

• Abstract/summary of work
• Metadata
  • Bibliographical details
  • Linguistic items & topics studied
  • Corpora & time periods studied
• Data
  • Tables, figures – or link to original publication
  • Overview of results
ITEM: *outtaken* (preposition and subordinator)

CONTRIBUTOR: Matti Rissanen


ABSTRACT: Brief description of the occurrence of the Middle English subordinator and preposition *outtaken*. This connective appears in Middle English and disappears in Early Modern English.
Diachronic coverage of the discussion: ME

Origin: Native. Cf., however, Latin except.

First appearance as a connective: LME

Dialectal distribution: Northern dialects

Corpora used: HCM; CMEPV

Summary: The connective appears in LME, approximately at the same time as the verb *outtake*. This supports the hypothesis that the form is a loan translation from Latin rather than the result of grammaticalisation in ME. It is typically a Northern or East Midland form.
Table 1. Outtaken, except, save and unless in late Middle and Early Modern English: the Helsinki Corpus (HC) and the Corpus of Middle English Prose and Verse (Middle English Compendium = MEC) (2002)

<table>
<thead>
<tr>
<th></th>
<th>outtake(n)</th>
<th>save</th>
<th>except</th>
<th>unless</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>prep.</td>
<td>subord.</td>
<td>prep.</td>
<td>subord.</td>
</tr>
<tr>
<td>HC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1350-1420</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>1420-1500</td>
<td>5</td>
<td>0</td>
<td>17</td>
<td>5</td>
</tr>
<tr>
<td>1500-1570</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>5</td>
</tr>
<tr>
<td>1570-1640</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>1640-1710</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>MEC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1350-1400</td>
<td>7</td>
<td>0</td>
<td>60</td>
<td>20</td>
</tr>
<tr>
<td>1400-1500</td>
<td>19</td>
<td>4</td>
<td>188</td>
<td>73</td>
</tr>
</tbody>
</table>
NEXT STEPS

• Design the structure of LCD
  • Platform, format, descriptive model…

• Pilot data insertion with research assistant (autumn 2014)

• Roundtable on *Language Change and Community* in early 2015
BIG-PICTURE ISSUES

- Change towards **analyticity**? (Lass 2000)
- Community type (Trudgill 2011)
- Register change (Szmrecsanyi 2012)
- Rate of change
EXTENDING THE PARADIGMS OF CHANGE

• Vartiainen: lexical **semantic** and constructional change
  • Subjectification (Traugott 2010)

• Säily: **derivational** vs. inflectional change
  • Redefining syntheticity (Danchev 1992)

• New materials and **methods**
QUESTIONS FOR YOU

• Structure of database?
• Copyright issues?
• Interested in contributing?
  • lcd-project@helsinki.fi
REFERENCES


