The weakification of strong preterites in West-Germanic: an interdisciplinary approach

Freek Van de Velde\textsuperscript{1}
Katrien Beuls\textsuperscript{2}
Isabeau De Smet\textsuperscript{1}
Dirk Pijpops\textsuperscript{1, 3}

\textsuperscript{1} University of Leuven, QLVL
\textsuperscript{2} AI-Lab, Vrije Universiteit Brussel
\textsuperscript{3} FWO Flanders
Strong and weak preterites

- Germanic languages have two morphological strategies for building preterites (not counting analytic perfects, he has written a book):

  1. Strong inflection:
     - English *sing* – *sang*
     - Ablaut, based on Indo-European aspectual system (perfect > preterite)

  2. Weak inflection
     - English *work* – *worked*
     - Dental suffix, based on a analytic formation [VERB + *d^h_1eh*, *d^h_1oh* - ('did')]
Changes

• Various changes occur:
  – irregularisation (Eng. buy – bought)
  – one strong ablaut class to another (Du. heffen – hief < hoef (Germ. hob, hub))
  – weak to strong (Du. vragen – vroeg < vraagde (vs. Germ. fragte))
  – strong to weak (Eng. carve – carved < cearf (Du. kerfde < karf))

⇒ Long-term drift, over many centuries
Quantifying the weakification

• Lieberman et al. (2007):
  – tracked all originally strong Old English verbs (that still exist)
  – noted when they weakened (Middle or Modern English)
  – reference grammars
  – binary encoding (strong = 1, weak = 0)
  – 6 log-frequency bins

• Carroll et al. (2012):
  – German
  – same method
  – Old, Middle, Early New, New High German
Quantifying the weakification

- Dutch data (2017)
  - Old, Middle, Modern (1500-1800) and present-day Dutch (1800-now)
  - controlled for type-token frequency and vowel pattern (ABA, ABB or ABC)
Lieberman et al. 2007: Constant rate of regularisation through time, only dependent on frequency

Carroll et al. 2012: Constant rate does not work for German... neither for Dutch

If the constant regularisation rate were true, the lines would follow the same power law curve and overlap.
Lieberman et al. 2007: Constant rate of regularisation through time, only dependent on frequency

\[ \Rightarrow \text{lines follow the same power law curve (linear on log-log plot) and overlap} \]

Lieberman et al. 2007: three measurement points:

Replication with fourth measurement point:
Lieberman et al. 2007: Constant rate of regularisation through time, only dependent on frequency

But the constant rate breaks down when we add an extra measurement point for E. Mod. Eng.:
Socio-demographical factors

• Can we attribute these changes to demography?
Socio-demographical factors

• Can we attribute these changes to demography?
• Lupyan & Dale (2010):
  - Smaller languages: more morphological complexity
  - Bigger languages: less morphological complexity
Socio-demographical factors

• Can we attribute these changes to demography?

• Lupyan & Dale (2010):
  - Smaller languages: more morphological complexity
  - Bigger languages: less morphological complexity

• Bentz & Winter (2013):
  - Languages with more L2-speakers: smaller case systems
Socio-demographical factors

• Can we attribute these changes to demography?

• Lupyan & Dale (2010):
  - Smaller languages: more morphological complexity
  - Bigger languages: less morphological complexity

• Bentz & Winter (2013):
  - Languages with more L2-speakers: smaller case systems

⇒ Languages adapt to the cognitive constraints of their speakers (Christiansen & Chater 2008)
⇒ Morphosyntactic complexity is reduced by high degree of language contact (involving adult learners)
Historical demographic data

• Problem: no clear data on population size or migration

• We can work with urbanisation:
  – In pre-industrial times, population growth is too high to be explained solely by natural growth (De Vries 1984:199-266, Howell 2006:208)
  – Migration, leading to koineization (Kerswill 2002), due to an influx of L2 speakers
    • Language diversity was higher in Medieval and Early Modern cities
    • Dialects were often mutually unintelligible

• Data Bairoch et al. (1988)
Average of largest city in each century covering the linguistic periods in each area

\[
\log(\text{inh}) \Rightarrow \begin{array}{ccc}
\text{English} & 0.96^* & 0.97^* & 0.77 \text{ (n.s.)} \\
\text{Dutch} & 0.94 \text{ (n.s.)} & 0.99^{**} & 0.82 \text{ (n.s.)} \\
\text{German} & 0.90 \text{ (n.s.)} & 0.81 \text{ (n.s.)} & 0.99^* \\
\end{array}
\]
Observing & Simulating

Empirical research

Observed data

Simulated data

Theoretical explanation

Computer simulations
Observing & Simulating

Empirical research

Computer simulations

- Weakification from inferior type & token frequency
- Conserving Effect
- Class Resilience

General Applicability

Pijpops, Beuls & Van de Velde (2015)
Observing & Simulating

Different rates of weakification

Empirical research

Demography

Computer simulations
Observing & Simulating

Parameters:
- Number of series: 20
- Number of agents: 100
- Time: 5,000,000 time units (average interactions per agent)
- Replacement rate: 1/5,000, 1/10,000, 1/20,000, 1/100,000
- Replacement number: 1
- Verbal replacement: none
Conclusions

• No constant rate of weakification

• Different rates can be explained by language/dialect contact
Thanks!


References