Language-specific differences in regularization rates of the Germanic preterite

Isabeau De Smet[1]
Katrien Beuls[2]
Dirk Pijpops[1, 3]
Freek Van de Velde[1]

[1] University of Leuven, QLVL
[2] AI-Lab, Vrije Universiteit Brussel
[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 + *$d^h eh_1$-, *$d^h oh_1$- ('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

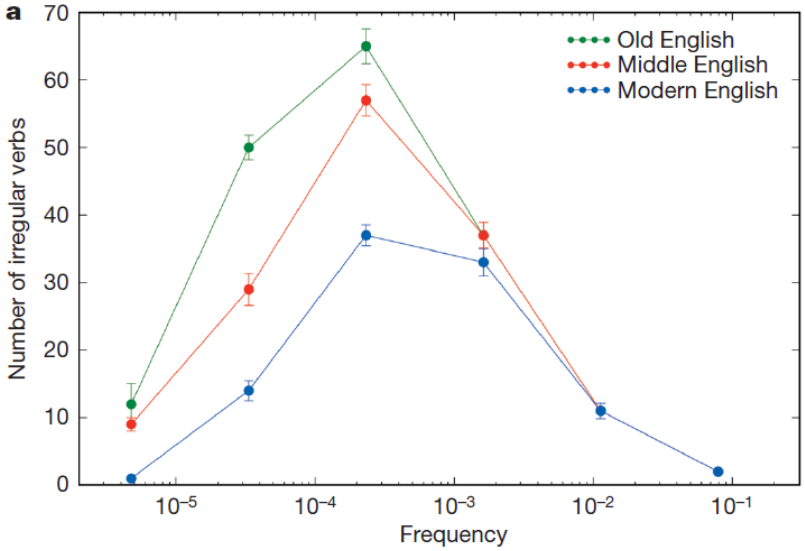
Quantifying the weakification

- 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

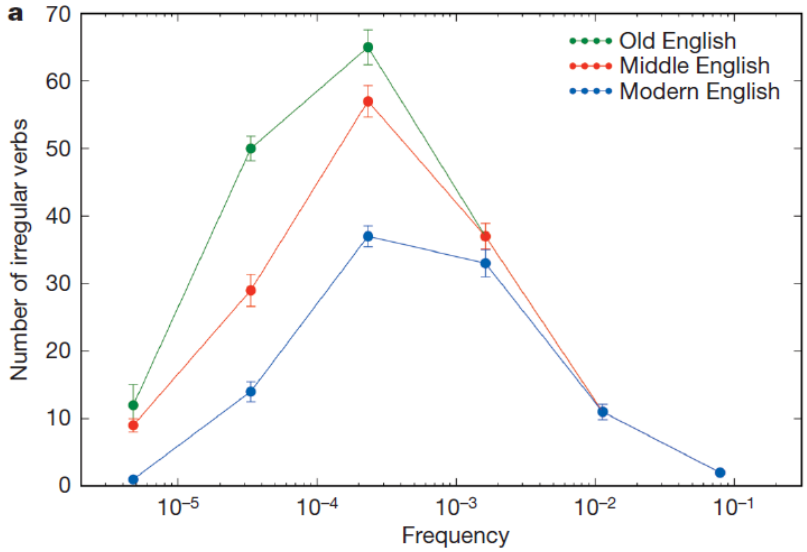
Quantifying the weakification

- 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)

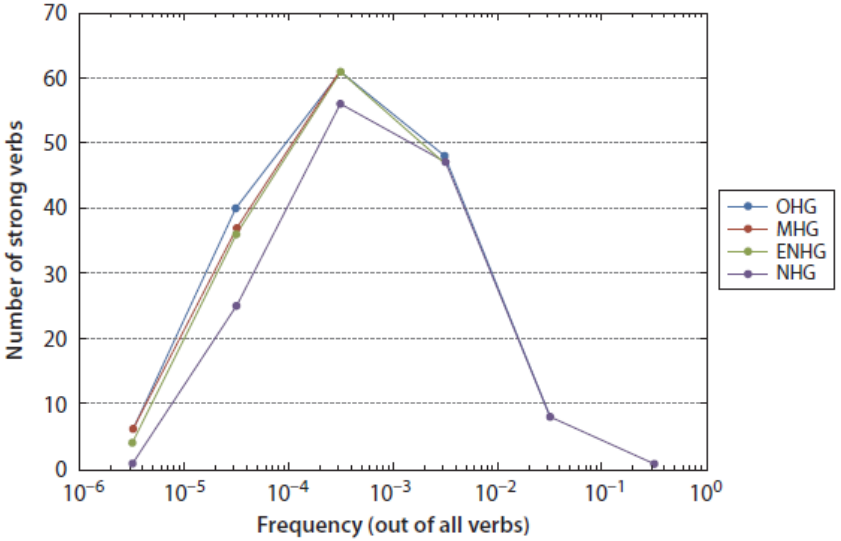
ENGLISH: Lieberman et al. 2007



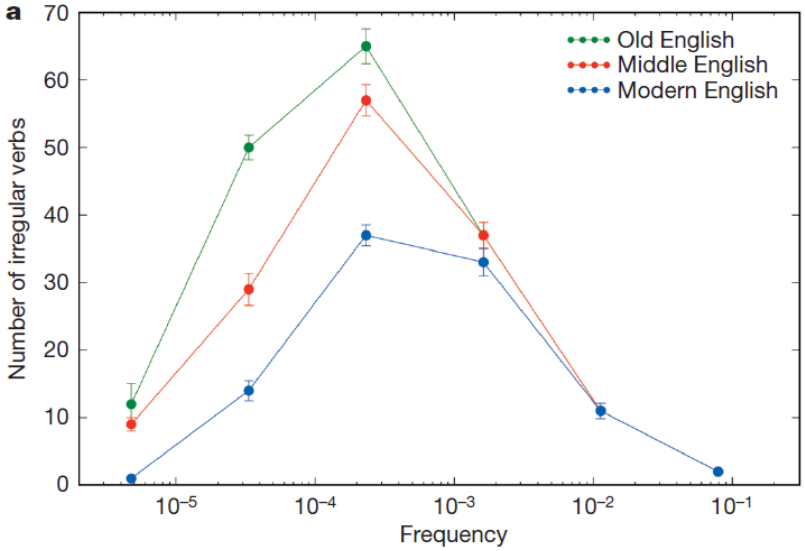
ENGLISH: Lieberman et al. 2007



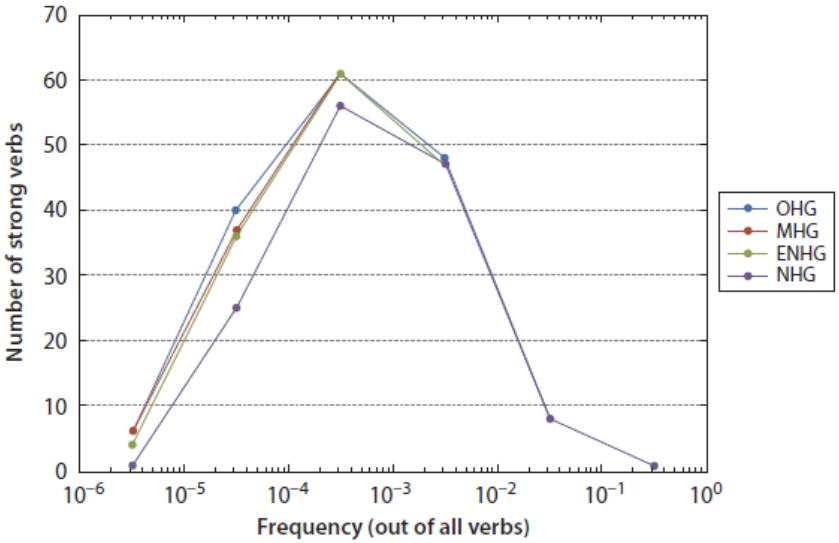
GERMAN: Carroll et al. 2012



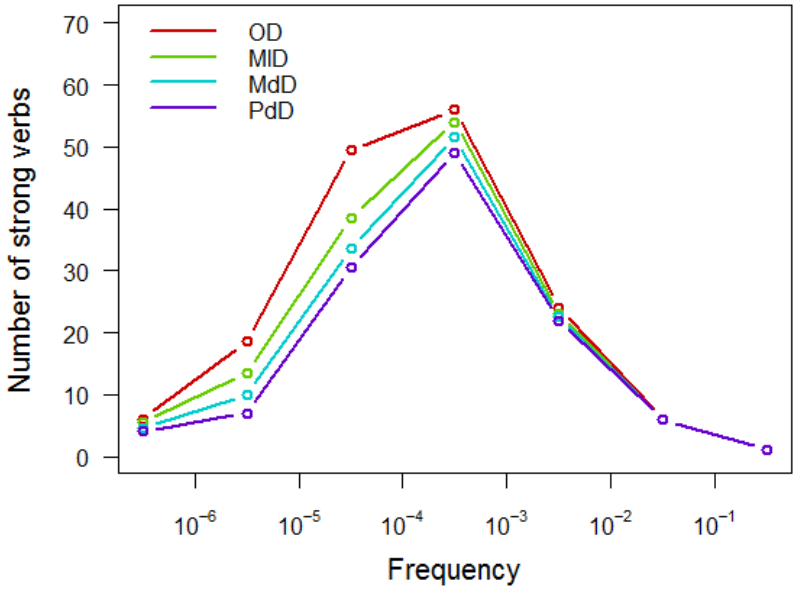
ENGLISH: Lieberman et al. 2007



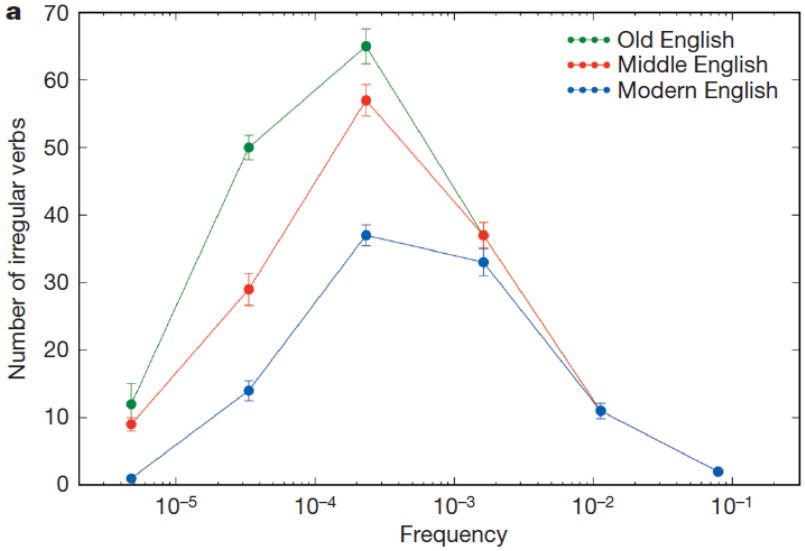
GERMAN: Carroll et al. 2012



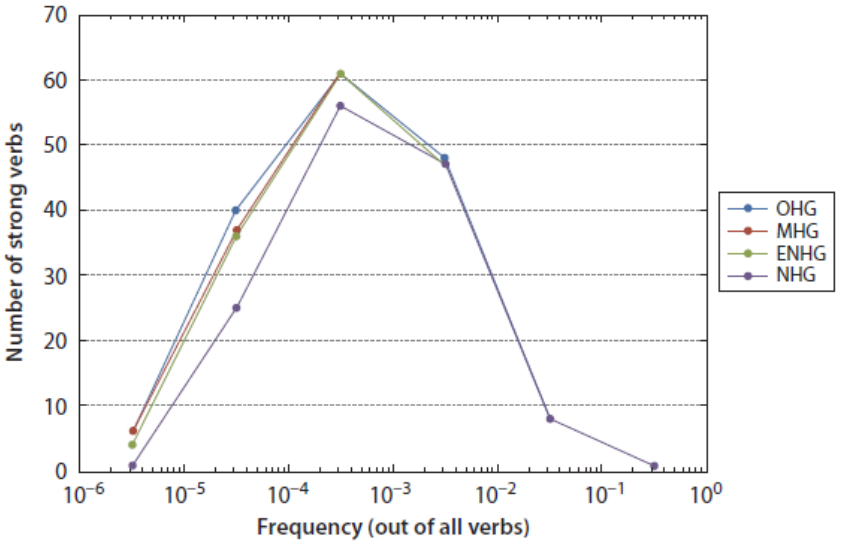
DUTCH



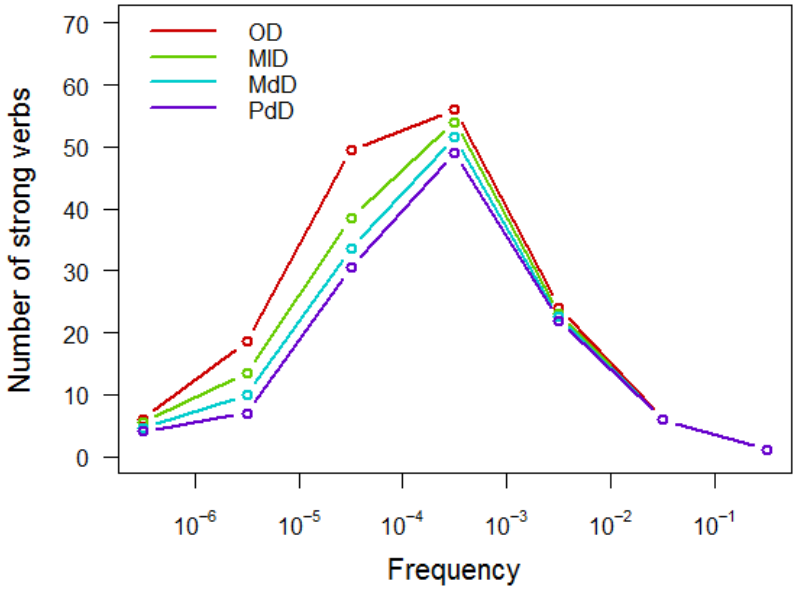
ENGLISH: Lieberman et al. 2007

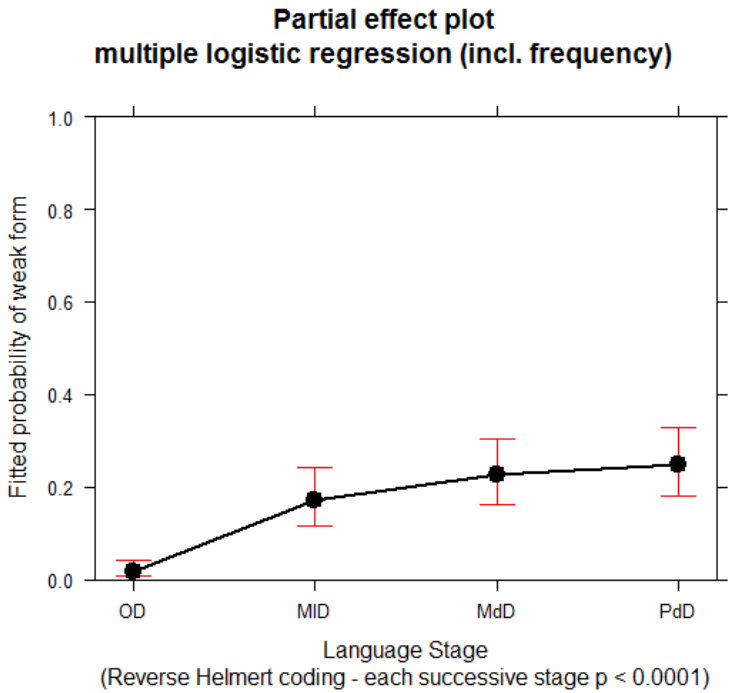


GERMAN: Carroll et al. 2012

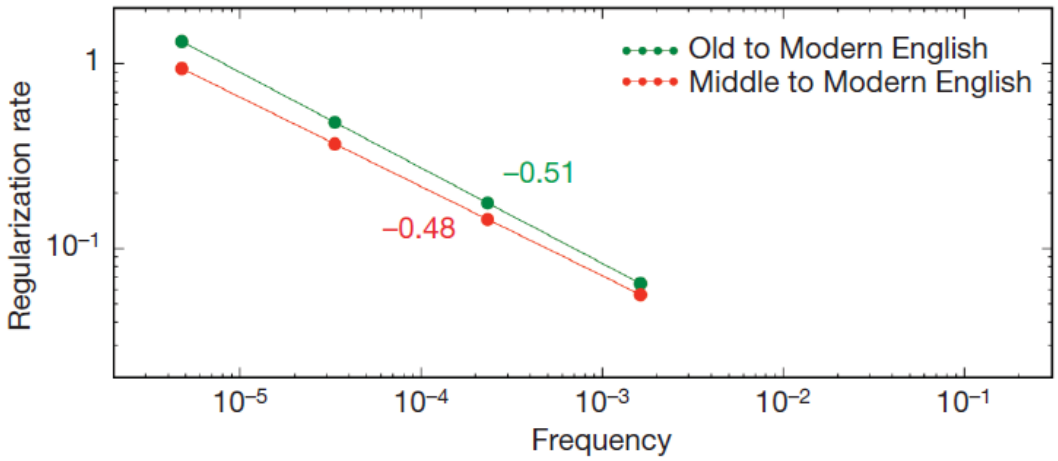


DUTCH



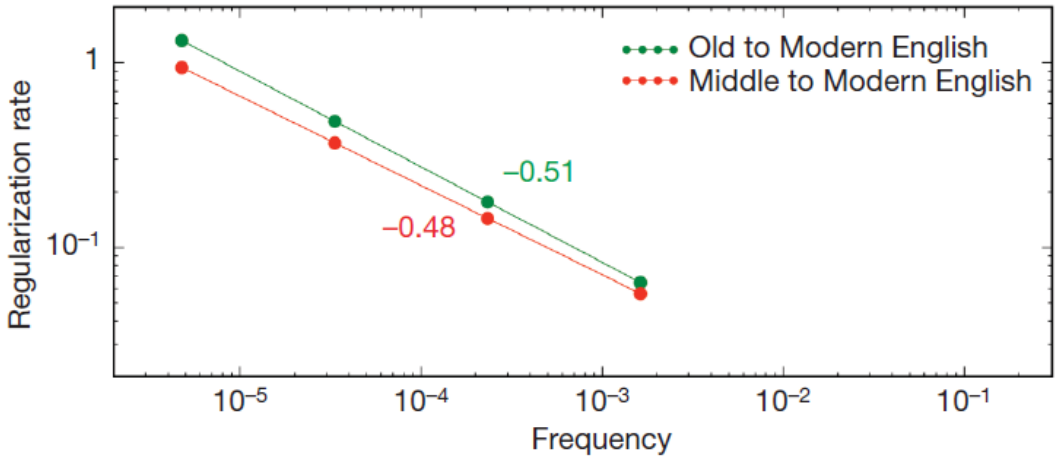


Lieberman et al. 2007: Constant rate of regularisation through time, only dependent on frequency

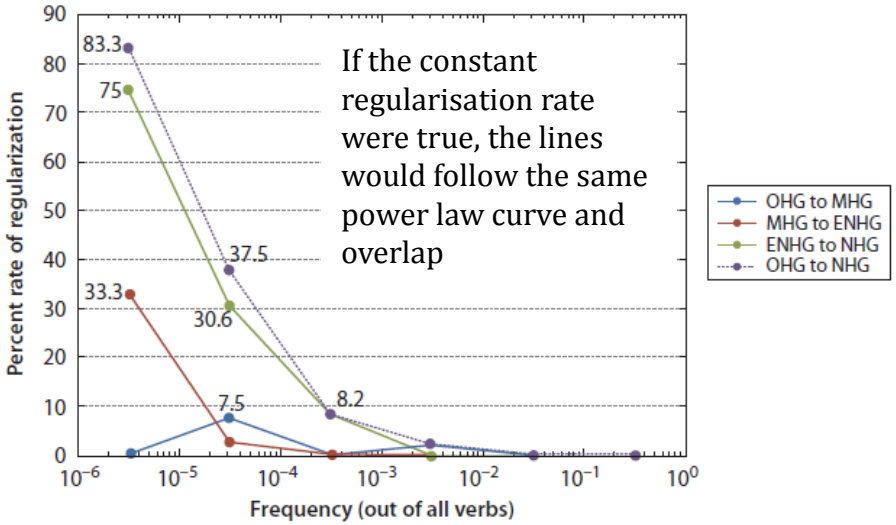


⇨ lines follow the same power law curve (linear on log-log plot) and overlap

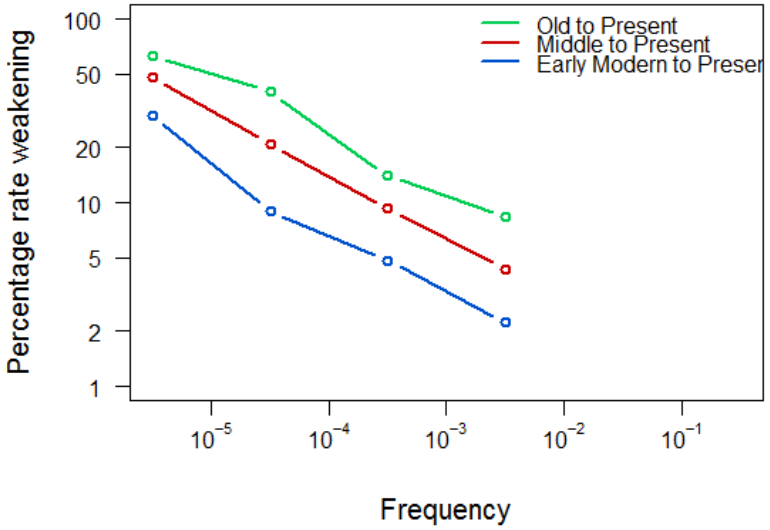
Lieberman et al. 2007: Constant rate of regularisation through time, only dependent on frequency



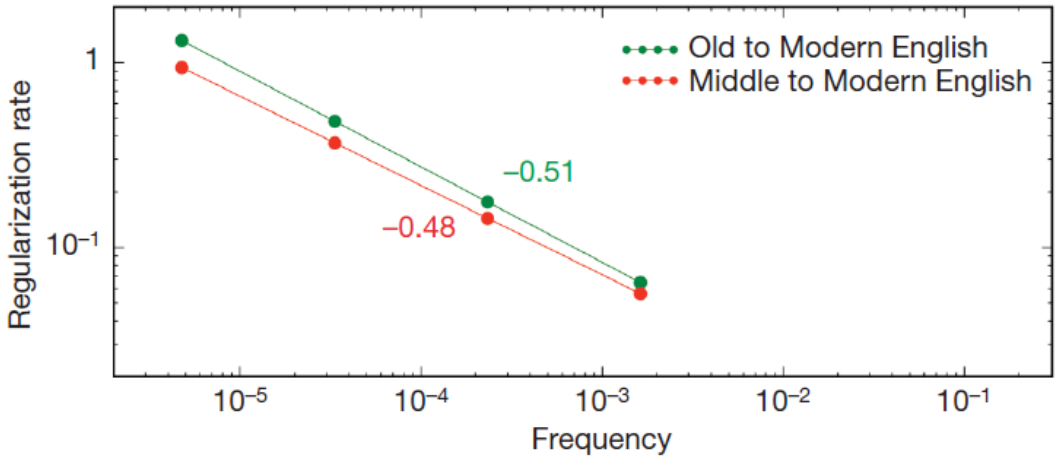
Carroll et al. 2012: Constant rate does not work for German



... neither for Dutch



Lieberman et al. 2007: Constant rate of regularisation through time, only dependent on frequency

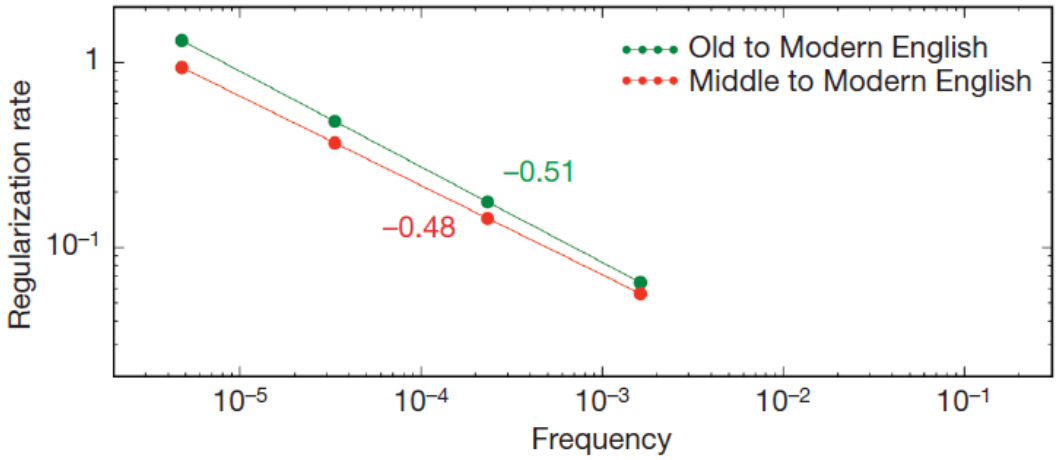


⇨ lines follow the same power law curve (linear on log-log plot) and overlap

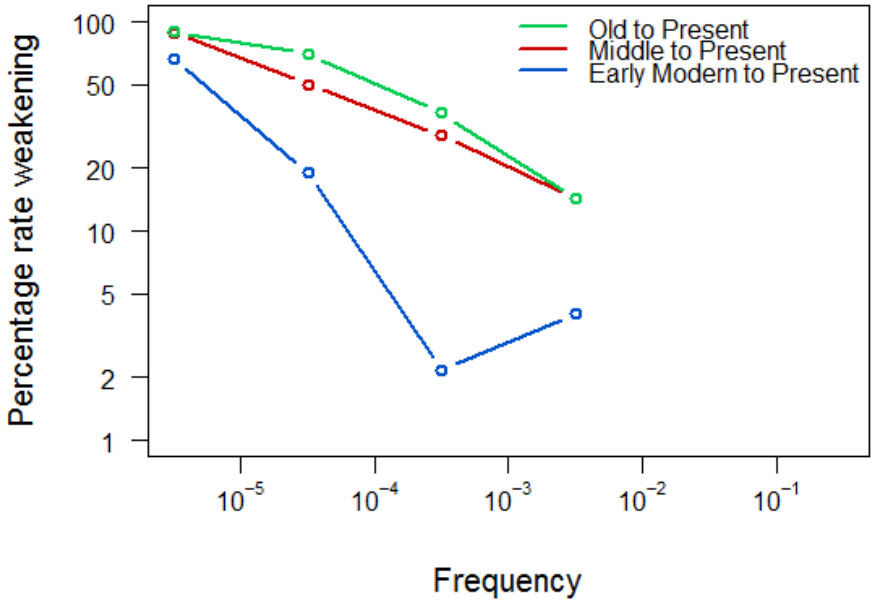
Lieberman et al. 2007: three measurement points:

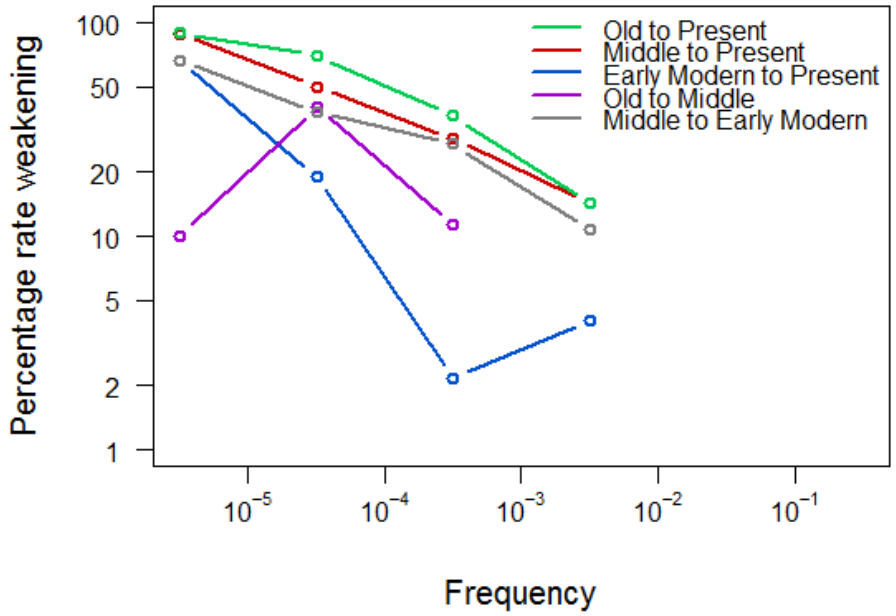
Replication with fourth measurement point:

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?

"[A] social characteristic with structural consequences is dialect or language contact. Increased exposure to different varieties often – though not always – corresponds to patterns of morphological and other leveling or simplification (...). The ENHG period, when verb regularization picks up dramatically in the history of German, is a period notable for increased geographical mobility, in particular urbanization."
(Carroll et al. 2012: 169)

Historical demographic data

- Problem: no clear data on population size or migration

Historical demographic data

- Problem: no clear data on population size or migration

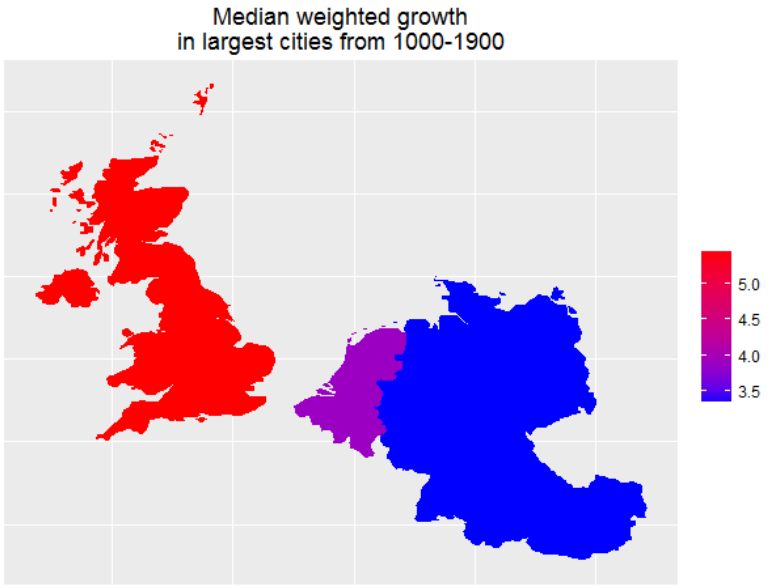
- We can work with urbanisation:
  - In pre-industrial times, population growth is too high to be explained solely by natural growth (De Vries 1984:199-266, Howell 2006:208)
  - Migration, leading to koineization (Kerswill 2002), due to an influx of L2 speakers
    - Language diversity was higher in Medieval and Early Modern cities
    - Dialects were often mutually unintelligible
- Data:
  - De Vries (1984); Chandler (1987); Bairoch et al. (1988); Mitchell (1998).

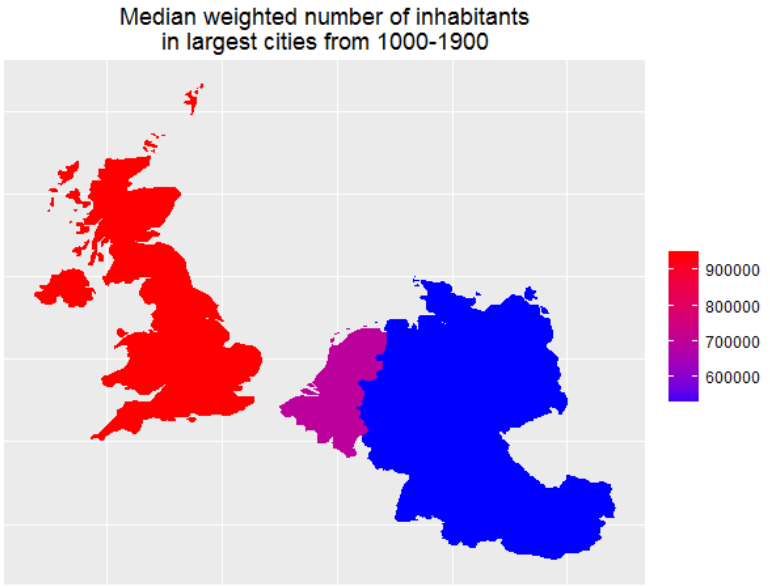
Historical demographic data

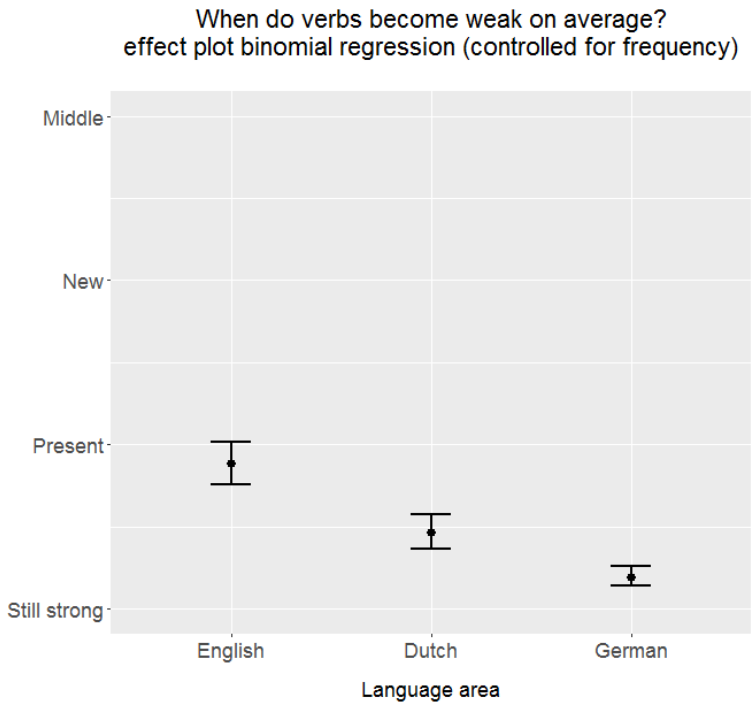
- 3 Areas:
  - ENGLISH: United Kingdom
  - DUTCH: Belgium & Netherlands
  - GERMAN: Germany & Austria
- 4 Periods:
  - Old: mean number of inh. of 3 largest cities in 1200, weighted for their rank
  - Middle: mean number of inh. of 5 largest cities in 1500, weighted for their rank
  - New: mean number of inh. of 7 largest cities in 1700, weighted for their rank
  - Present: mean number of inh. of 9 largest cities in 1900, weighted for their rank

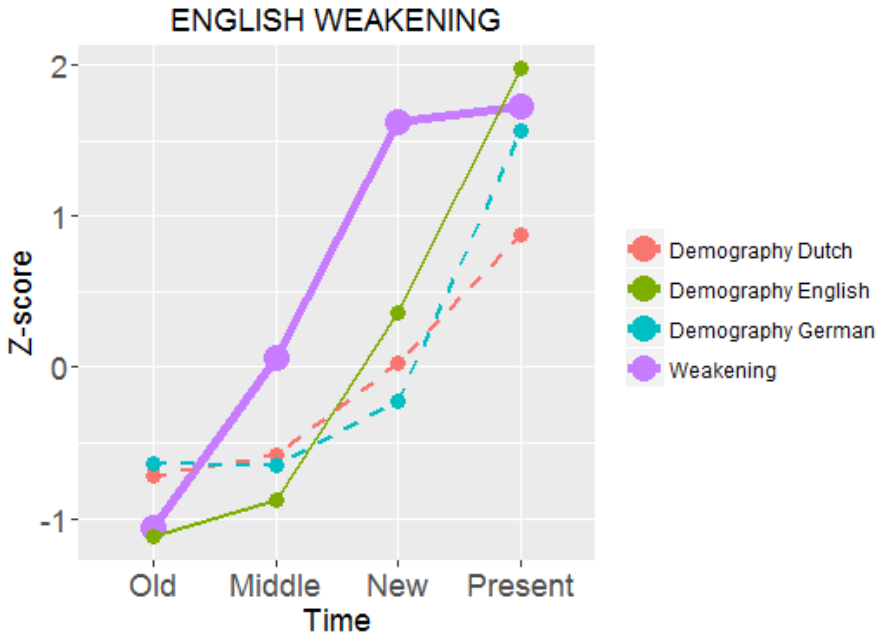
- Cumulative percentage of weak verbs over total number of originally strong verbs

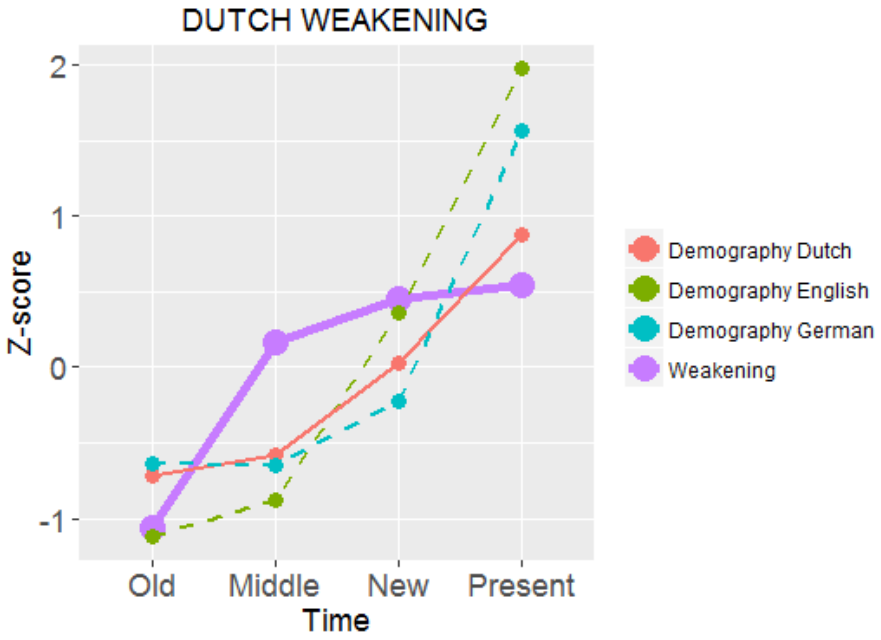
Dutch between English and German (Van Haeringen 1956)

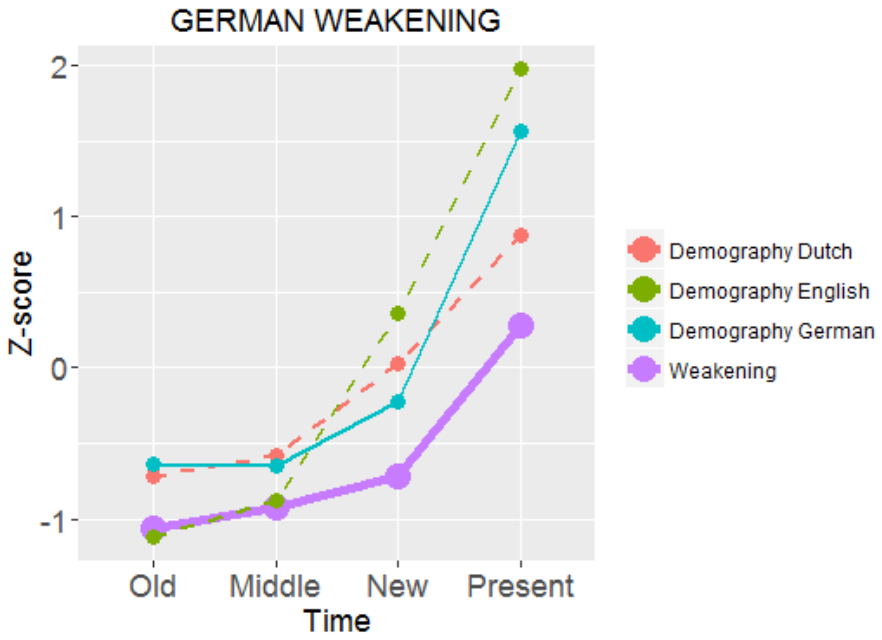






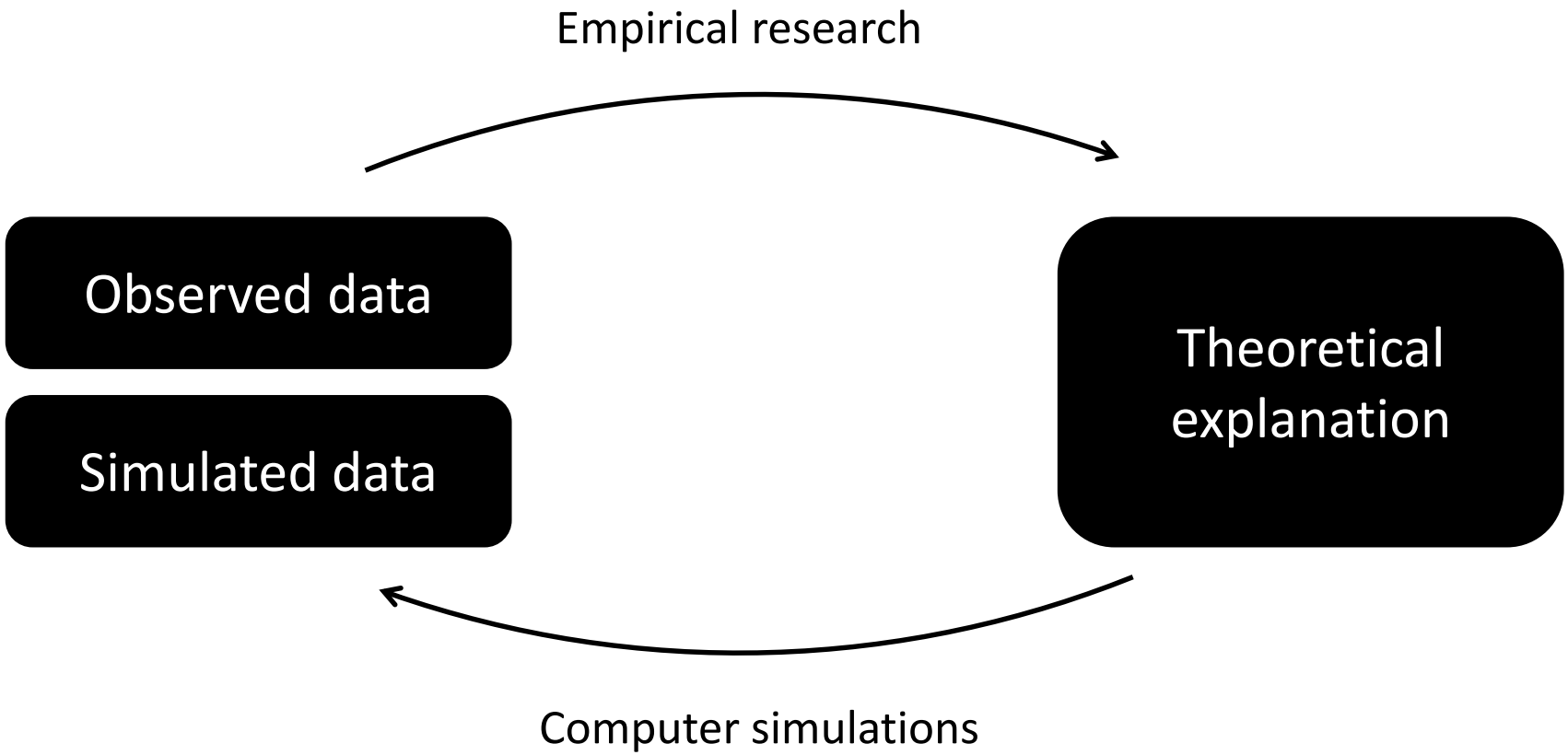




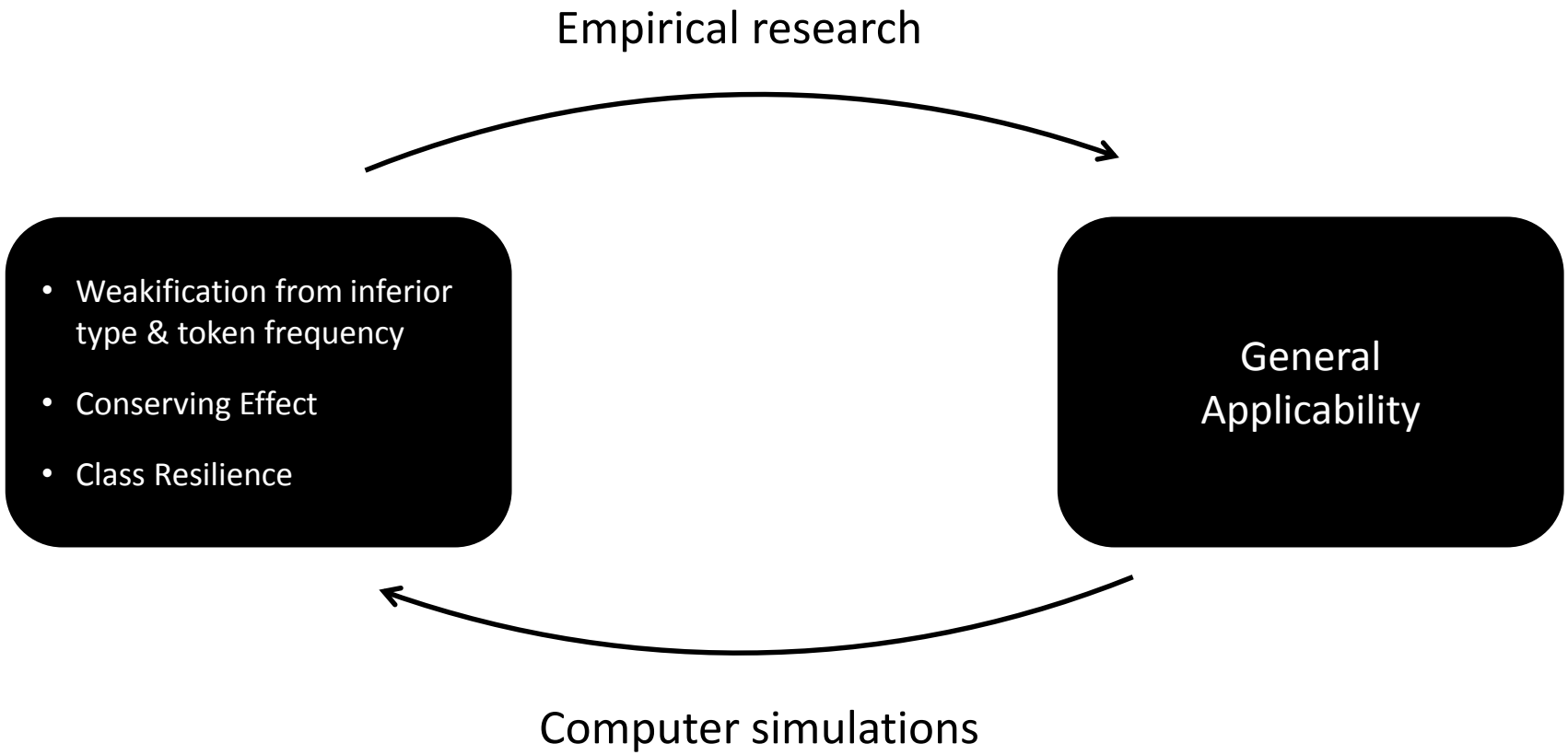


Pearson corr.		Weakening		
		English	Dutch	German
Demography	English	0.87 (p=0.13)	0.72 (p=0.28)	0.96 (p=0.03)
	Dutch	0.86 (p=0.31)	0.72 (p=0.28)	0.97 (p=0.03)
	German	0.69 (p=0.31)	0.56 (p=0.44)	0.99 (p=0.01)

Observing & Simulating

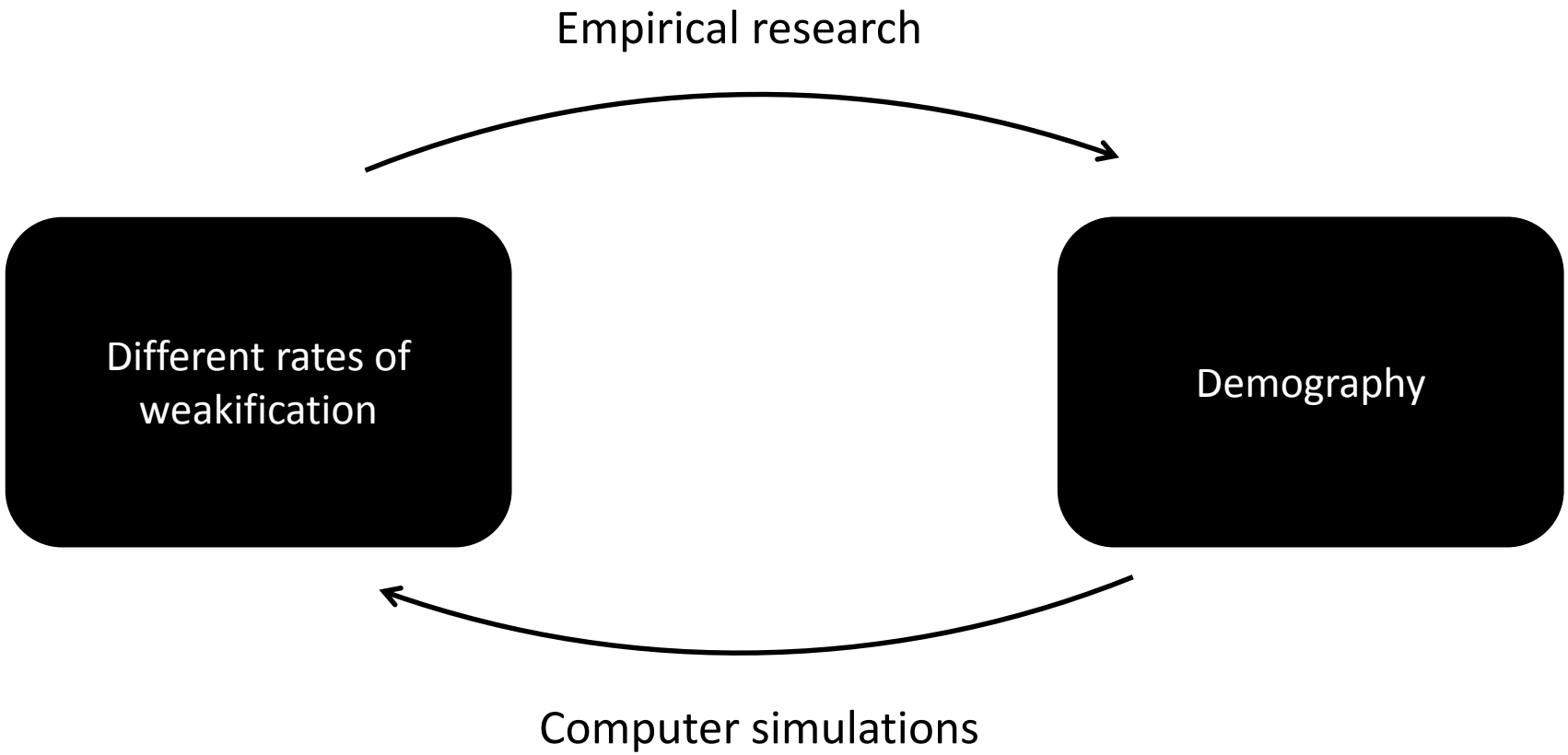


Observing & Simulating



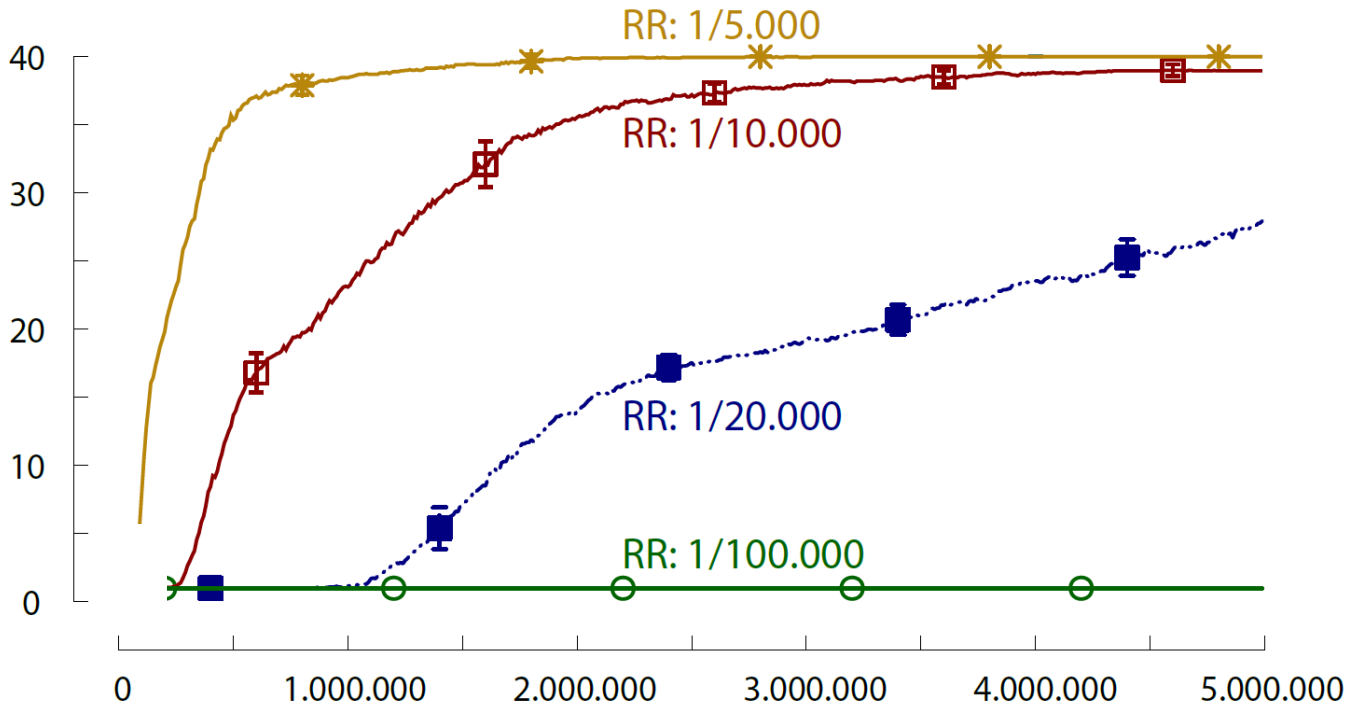
Pijpops, Beuls & Van de Velde (2015)

Observing & Simulating



Pijpops, Beuls & Van de Velde (2015)

Observing & Simulating



Parameters:

- Number of series: 20
- Number of agents: 100
- Time: 5.000.000 times 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

Pijpops, Beuls & Van de Velde (2015)

Conclusions

- No constant rate of weakification
- Different rates can be explained by language/dialect contact

Thanks!

Acknowledgement: thanks to Ryan Carroll, Ragnar Svare and Joseph Salmons for sharing their dataset

More:

Pijpops, Dirk, Katrien Beuls & Freek Van de Velde. 2015. The rise of the verbal weak inflection in Germanic. An agent-based model. *CLIN Journal* 5: 81-102.

De Smet, Isabeau. 2016. *De verzwakking van het preteritum in het Nederlands*. Master's thesis, University of Leuven.

References

- Bairoch, P., J. Batou & P. Chèvre. 1988. *La population des villes Européenes de 800 à 1850.* Geneva: Librarie Droz.
- Carroll, R., R. Svare & J. Salmons. 2012. 'Quantifying the evolutionary dynamics of German verbs'. *Journal of Historical Linguistics* 2: 153-172.
- Chandler, T. 1987. *Four thousand years of urban growth*. Lewiston: St. David's University Press.
- De Smet, I. 2016. *De verzwakking van het preteritum in het Nederlands*. Master's thesis, University of Leuven.
- De Vries, J. 1984. *European urbanization, 1500-1800*. London: Methuen & Co.
- Howell, R.B. 2006. 'Immigration and koineisation: the formation of early Modern Dutch urban vernaculars.' *Transactions of the Philological Society* 104(2): 207-227.
- Kerswill, P. 2002. 'Koineization and accomodation: Koineization as language change'. In: J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *The handbook of language variation and change*. Oxford: Blackwell. 669-702.
- Lieberman, E., J.-B. Michel, J. Jackson, T. Tang & M.A. Nowak. 2007. 'Quantifying the evolutionary dynamics of language'. *Nature* 449: 713-716.
- Mitchell, B.R. 1998. *International historical statistics: Europe 1750-1993*. London: Macmillan.
- Pijpops, D, K. Beuls & F. Van de Velde. 2015. 'The rise of the verbal weak inflection in Germanic. An agent-based model'. *Computational Linguistics in the Netherlands Journal* 5: 81-102.
- Trudgill, P. 2002. 'Linguistic and social typology'. In: J.K. Chambers, P. Trudgill & N. Schilling-Estes (eds.), *The handbook of language variation and language change*. Oxford: Blackwell. 707-728.