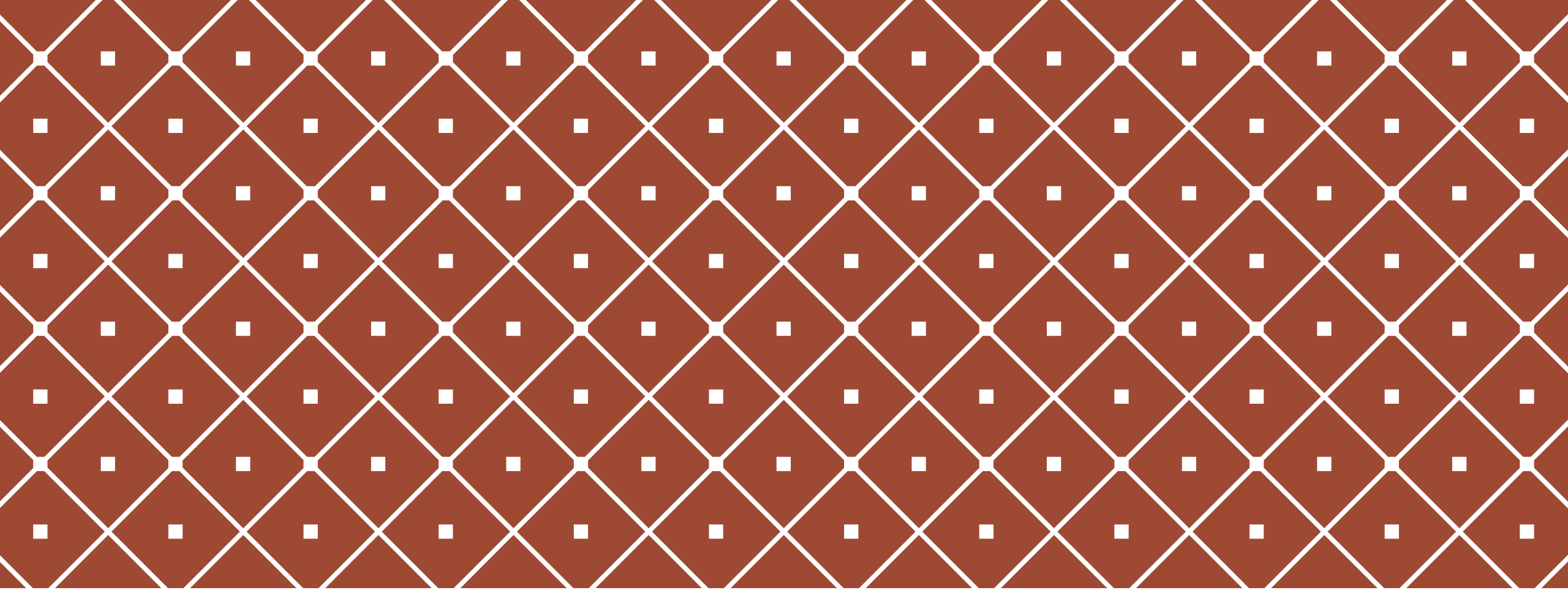
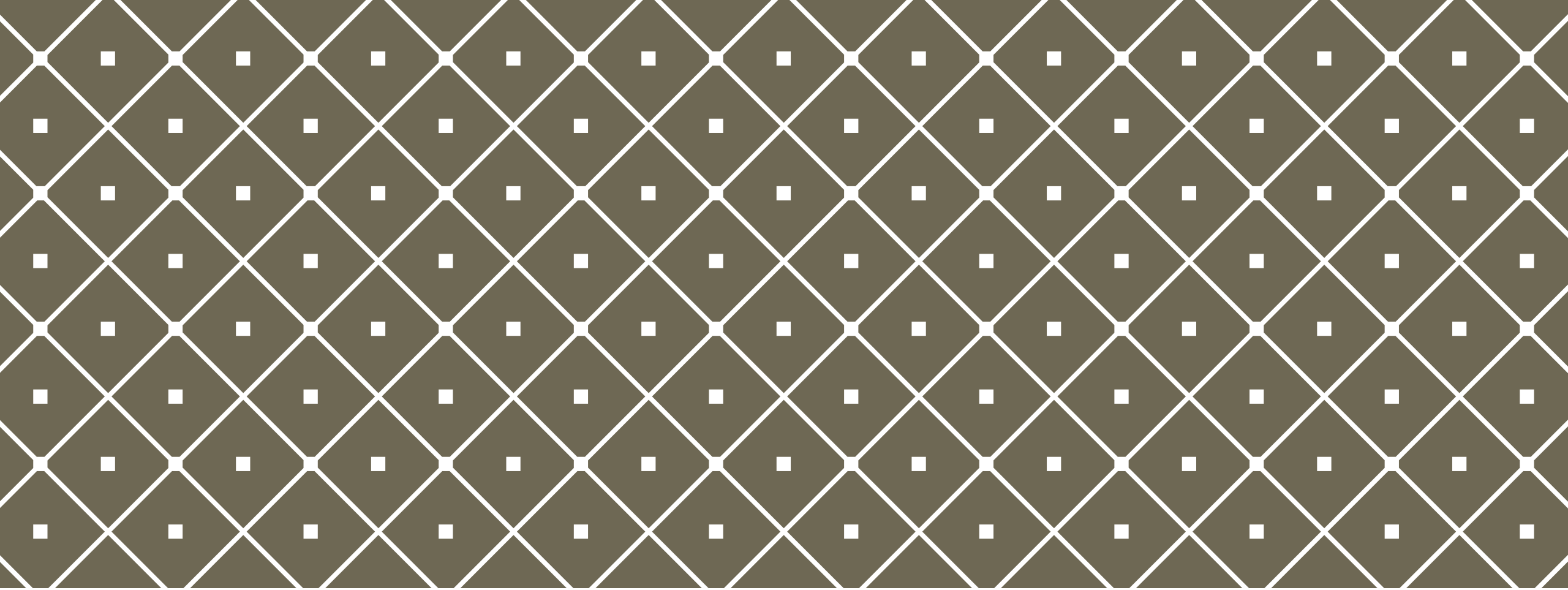


COMPUTER SIMULATIONS IN LANGUAGE VARIATION RESEARCH

Dirk Pijpops



IN THEORY

Rational paradigm

Theory
↓
Thesis
↓
Discussion

Mathematics, filosophy, theoretical physics

Chomsky (1995), Bybee (2010),…

Empirical paradigm

Theory
↓
Hypothesis
↓
Data

Experimental physics, chemistry, psychology

Grondelaers et al. (2009),…

Simulational paradigm

Theory
↓
Simulation
↓
Data

Astronomy, nuclear physics, evolutionary biology, economics

Steels (2015), Bloem et al. (2015), Pijpops et al. (2015), Lestrade (2015),…

Data-driven paradigm

Data
↓
Exploration
↓
Trends

Sociology?

Speelman et al. (forthc.), Levshina & Heylen (2014),…

SIMULATION = COMPUTER PROGRAM

! Be careful to interpret its results, e.g. economics

1. Find an effect to be explained
2. Draw up a possible cause
3. Implement cause in a computer program, making only minimal assumptions
4. See if effect emerges in computer program

⇨ If so, this shows that the cause CAN IN PRINCIPLE give rise to the effect. It does NOT show that it also does so in reality.

WHAT IF A DIFFERENT SIMULATION SHOWS THE SAME EFFECT?

1. Find an effect to be explained
2. Draw up a possible cause
3. Implement cause in a computer program, **making only minimal assumptions**
4. See if effect emerges in computer program

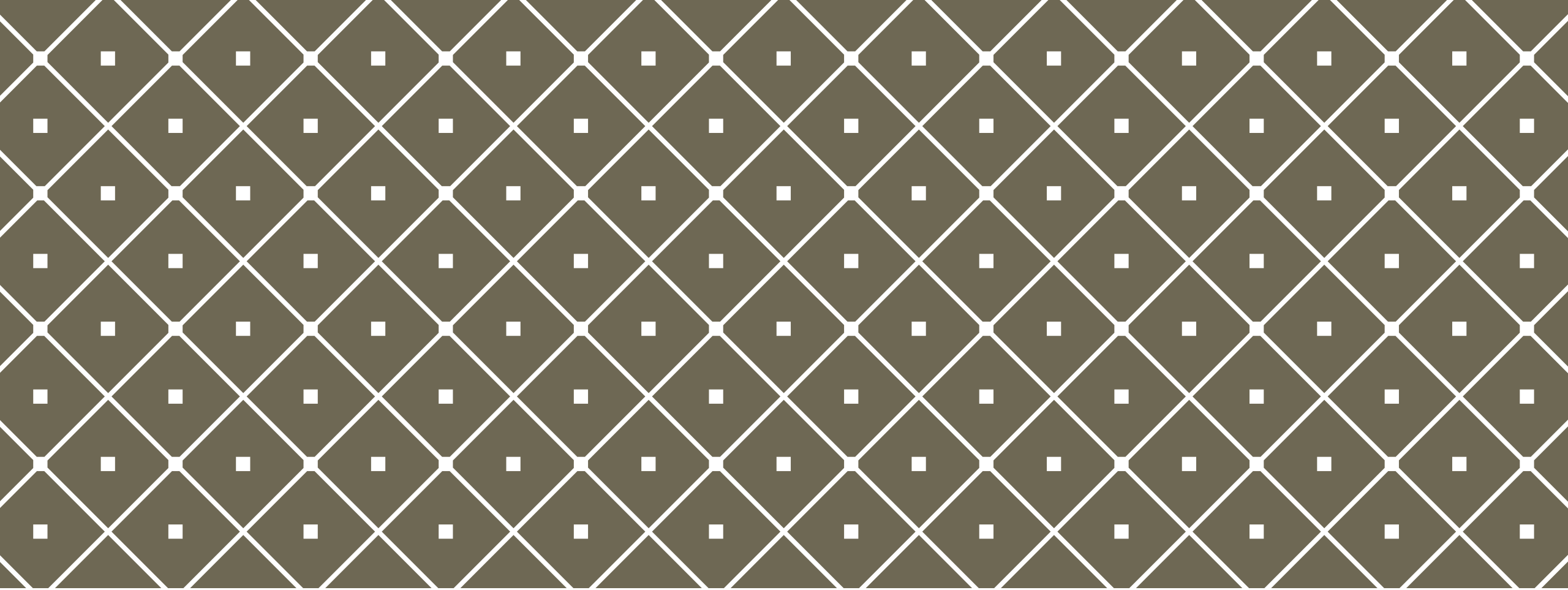
⇨ If so, this shows that the cause CAN in principle give rise to the effect. It does NOT show that it also does do in reality.

OCKHAM'S RAZOR

Heuristic: given two possible explanations for the same effect, choose the one that makes the fewest assumptions.

E.g. Palissy & Henry III

SIMULATION = COMPUTER PROGRAM
= IMPLEMENTED THEORY



IN PRACTICE

TO DO

1. Find an effect to be explained
2. Design the simulation at a conceptual level on paper
3. Implement the simulation in computer code
4. Evaluate the results

STEP 1: FIND AN EFFECT TO BE EXPLAINED

TO DO

1. Find an effect to be explained
   - Well-described by empirical studies
   - No single, universally accepted explanation

⇨ Rise of the verbal weak inflection in Germanic

Disclaimer: the following glosses over many of the complexities of the development of the Germanic past tense system.

GERMANIC PAST TENSE

- Strong inflection: classes of vowel alternations

ij → ee	i → a	i → o	ee → a	….
– kijk-keek	– zit-zat	– vind-vond	– steek-stak	
– lijd-leed	– lig-lag	– schrik-schrok	– eet-at	
– krijg-kreeg	– bid-bad	– zing-zong	– breek-brak	
– …		– ….	– …	

- Weak inflection: dentalsuffix

  + de/te

  – lach-lachte
  – waai-waaide
  – praat-praatte
  – …

⇨ Both are regular

GERMANIC PAST TENSE

- Strong inflection
  - Oldest
  - Indo-European aspectual systeem
- Weak inflection
  - Youngest
  - Germanic innovation
- Competition has been going on for thousands of years and still continues
  - *lach-loech, waai-woei, ? loop-loopte*
  - New verbs are conjugated weakly. Was not always the case: *schrijf-schreef*

GERMANIC PAST TENSE

- Clear why the weak inflection is winning now
  - Strong inflection is a lot less regular
  - Weak inflection has a much higher type frequency

- However, neither was the case in Proto-Germanic

RESEARCH QUESTION

How could the weak inflection have grown to overthrow the strong inflection, given that

i. The weak inflection had to start from a position vastly inferior in both type and token frequency

(↔ Hare and Elman 1995; Yang 2002)

ii. The strong inflection was still clearly regular?

(↔ Colaiori et al. 2015; Pijpops and Beuls 2015)

EXPLANATIONS FOR THE SUCCESS OF THE WEAK SUFFIX

1. General applicability of the dental suffix

2. Restrictions on the strong system

3. Disintegration of the strong system

(Ball 1968: 164; Bailey 1997: 17)

EXPLANATIONS FOR THE SUCCESS OF THE WEAK SUFFIX

1. **General applicability of the dental suffix**
2. Restrictions on the strong system
3. Disintegration of the strong system

(Ball 1968: 164; Bailey 1997: 17)

ARGUMENTS AGAINST GENERAL APPLICABILITY AS A SUFFICIENT EXPLANATION

1. The addition of the weak inflection only further complicates past tense inflection
   - A purely strong system is simpler than the present hybrid system, with all its exceptions
   - A purely weak system is admittedly even simpler, yet still isn't in sight
   - ⇨ Additional assumption needed: disintegration of the strong system, e.g. *houd-hield*
2. For a language user, general applicability is only useful if there are any verbs which he/she cannot (yet) conjugate strongly
   - If every strong class starts more frequent, it will be acquired faster
   - ⇨ Additional assumption needed: restrictions on the strong system

PROPOSALS

1. **General applicability of the dental suffix**
2. Restrictions on the strong system
3. Disintegration of the strong system

⇨ Disintegration of the strong system may be result, rather than cause

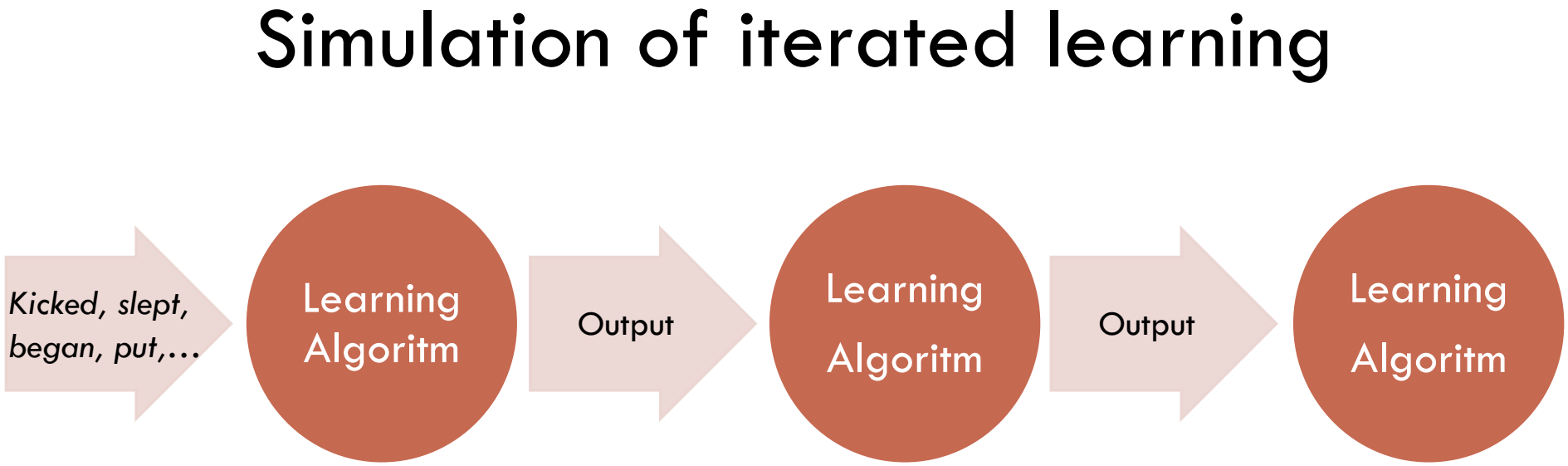
STEP 2: DESIGN THE SIMULATION AT A CONCEPTUAL LEVEL

TO DO

2. Design the simulation at a conceptual level
   - What kind of simulation do we want?
   - What do we put it?
   - What do we not put in?
   - How do we decide whether it is succesful or not?

WHAT KIND OF SIMULATION DO WE WANT?

- Acquisition as the main driving force behind language change:



WHAT KIND OF SIMULATION DO WE WANT?

Goal: what kind of learning algoritm do we need in order to obtain realistic effects in

i. The mistakes that children make, e.g. U-shaped learning
ii. Language changes in the recent past

Connectionists: Neural Netwerk vs. Formal linguists: explicit rules + memory of exceptions

Rumelhart & McClelland (1986), Macwhinney & Leinbach (1991), Plunkett & Marchman (1991, 1992), Hare & Elman (1995), Plunkett & Juola (1999)

Pinker & Prince (1988), Ling & Marinov (1993), Marcus et al. (1995), Taatgen & Anderson (2002), Yang (2002)

WHAT KIND OF SIMULATION DO WE WANT?

- Language use as the main driving force behind language change:

Agent-based simulations

WHAT KIND OF SIMULATION DO WE WANT?

- General applicability is only useful when **producing** language, not when learning it

⇨ Agent-based simulation

WHAT DO WE PUT IN?

- Single, generally applicable weak suffix vs. multiple strong classes
- Weak suffix starts inferior in type and token frequency to any individual ablaut class
- Verbs show a realistic (Zipfian) frequency distribution
- Agents are gradually replaced

WHAT DO WE NOT PUT IN?

- Any restrictions on the strong system: each verb can be conjugated strongly
- Any irregular verbs, or ways to become irregular
- Any other possible advantage to the weak inflection
  - ↳ Agents will never forget strong verb forms (↔ Taatgen and Anderson 2002: 124)
  - ↳ No advantage of linear segmentability: Hearers recognize equally easy

    sing-ed 'sing + PAST'

    s-ou-ng 'sing + PAST'
  - ↳ No social structure or social preference

⇨ Explicitly unrealistic

KEEP IT SIMPLE STUPID

- Only finite past tenses
- No influence of phonetic resemblance

(Landsbergen 2009: 18-19)

HOW DO WE DECIDE WHETHER IT IS SUCCESSFUL OR NOT?

Evaluation criteria: General applicability can explain

1. Rise of the Weak Inflection (Carroll et al. 2012; Cuskley et al. 2014)
2. Gradual Rise (Cuskley et al. 2014)
3. Conserving Effect (Bybee 2006: 715; Lieberman et al. 2007)
4. Class Resilience (Mailhammer 2007; Carroll et al. 2012: 163-164)

⇨ Emergence should not be dependent on specific parameter settings

⇨ Define AND delimit

STEP 3: IMPLEMENT THE SIMULATION

TO DO

3. Implement the simulation in computer code
   - Choose an environment/language, e.g. Babel2 (Lisp), R, Python,…
   - If you are forced to make a choice between several implementations
     i. Switch back to Step 2 to make a choice
     ii. If that fails, explore (and report on!) both options

GET A FUNCTIONING STRONG SYSTEM

Strong vowel alternations: extracted from Corpus of Spoken Dutch

I	ij → ee	krijg → kreeg
II-a	ie → oo	vlieg → vloog
II-b	ui → oo	kruip → kroop
III-a	i → o	vind → vond
III-b	e → o	trek → trok
III-c	e → ie	sterf → stierf
IV/V-a	ee → a	geef → gaf
V-b	i → a	zit → zat
VI	aa → oe	draag → droeg
VII-a	aa → ie	laat → liet
VII-b	a → i	hang → hing

GET A NUMBER OF VERBS TO BE EXPRESSED

Verbs: extracted from Corpus of Spoken Dutch

(all can be conjugated strongly, no irregulars, realistic frequency distribution)

- vind 1518
- zit 1157
- krijg 359
- lig 208
- …
- stink 11
- draag 11
- eet 10
- …
- bid 1

World	
vind	34%
zit	26%
…	…
draag	**0.25%**
sterf	0.20%
…	…
bid	0.02%

vind → vond 233

zit → zat 162

draag → droeg 9

draag → draagde 1

bid → bad 1

sterf → sterfde 3

droeg +1

droeg 90%

draagde 10%

droeg

World	
vind	34%
zit	26%
…	…
draag	**0.25%**
sterf	0.20%
…	…
bid	0.02%

vraag → vroeg 80

slaap → sliep 15

blaas → blies 3

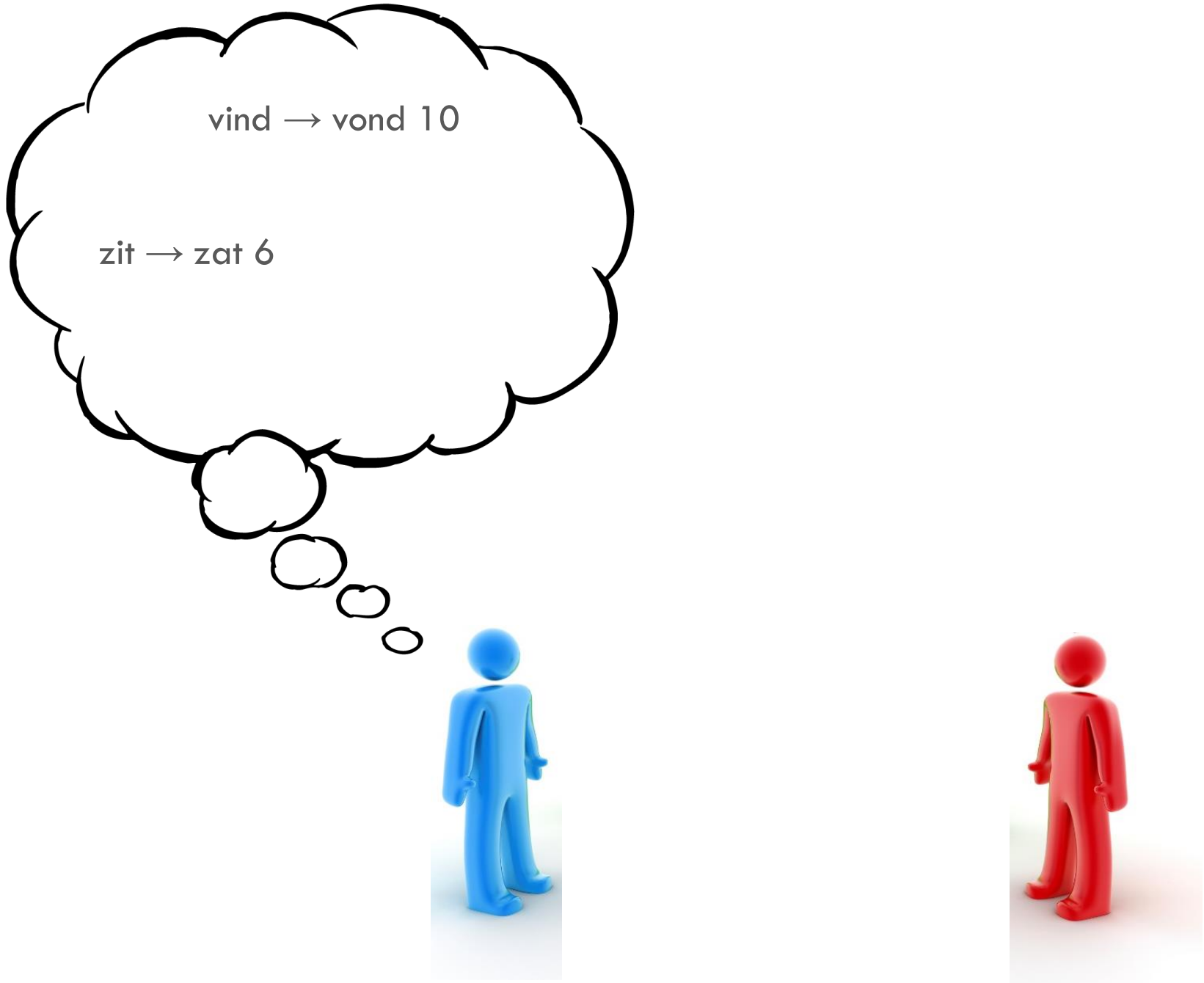
sterf → sterfde 2

droeg 80%
drieg 18%
draagde 2%

droeg +1

droeg

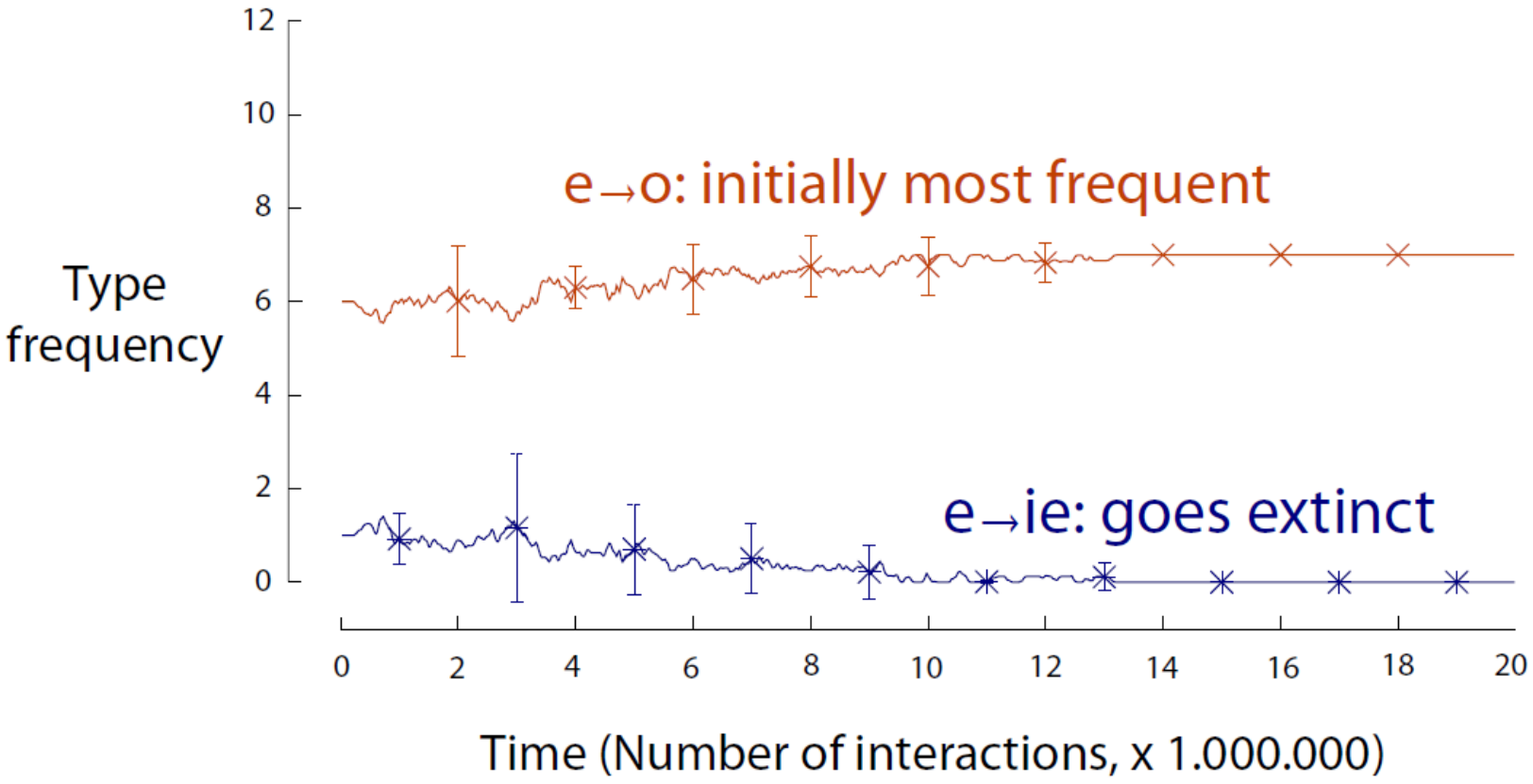
World	
vind	34%
zit	26%
…	…
draag	**0.25%**
sterf	0.20%
…	…
bid	0.02%



DEFINE THE STARTING SITUATION

Starting situation: only strong classes

- All starting agents know perfectly how to conjugate each verb
- Have access to all strong classes
- Weak inflection does not exist



- Either both competing classes hold each other in balance
- Or the initially most frequent one prevails

⇨ Initial frequency fully determines outcome

BRING IN THE WEAK INFLECTION

Starting position of the weak inflection

- ~~Preterito-presentia~~ (Bailey 1997: 578)
- Take the starting position of the feeblest strong class, i.e. e ⟶ *ie*
  - Inferior in initial type & token frequency to any other class
  - Direct competition with more frequent III-b class (e ⟶ o)
  - Went extinct in the previous simulation

The only difference between the weak inflection and the *e* ⟶ *ie* class is that the weak suffix can in principle be applied to all verbs

STEP 4: EVALUATE THE RESULTS

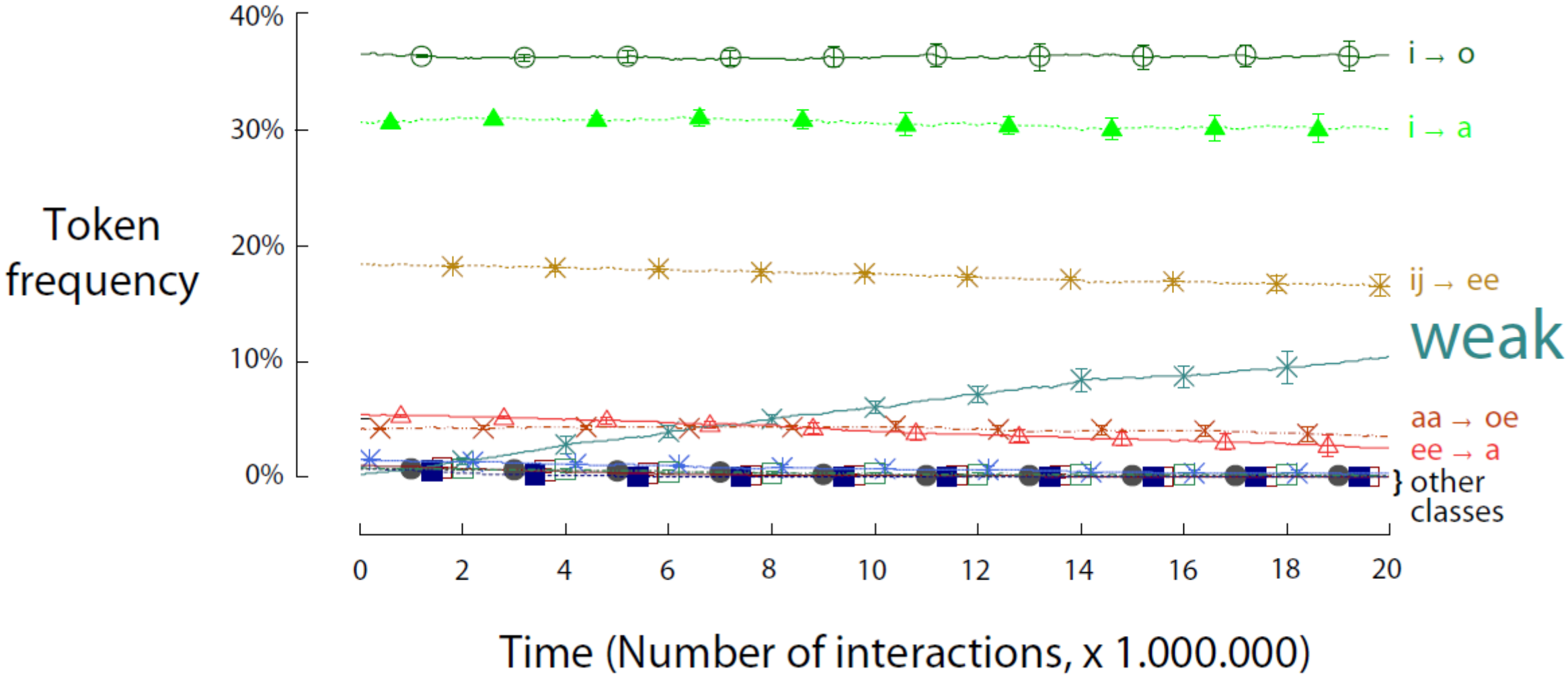
HOW DO WE DECIDE WHETHER IT IS SUCCESSFUL OR NOT?

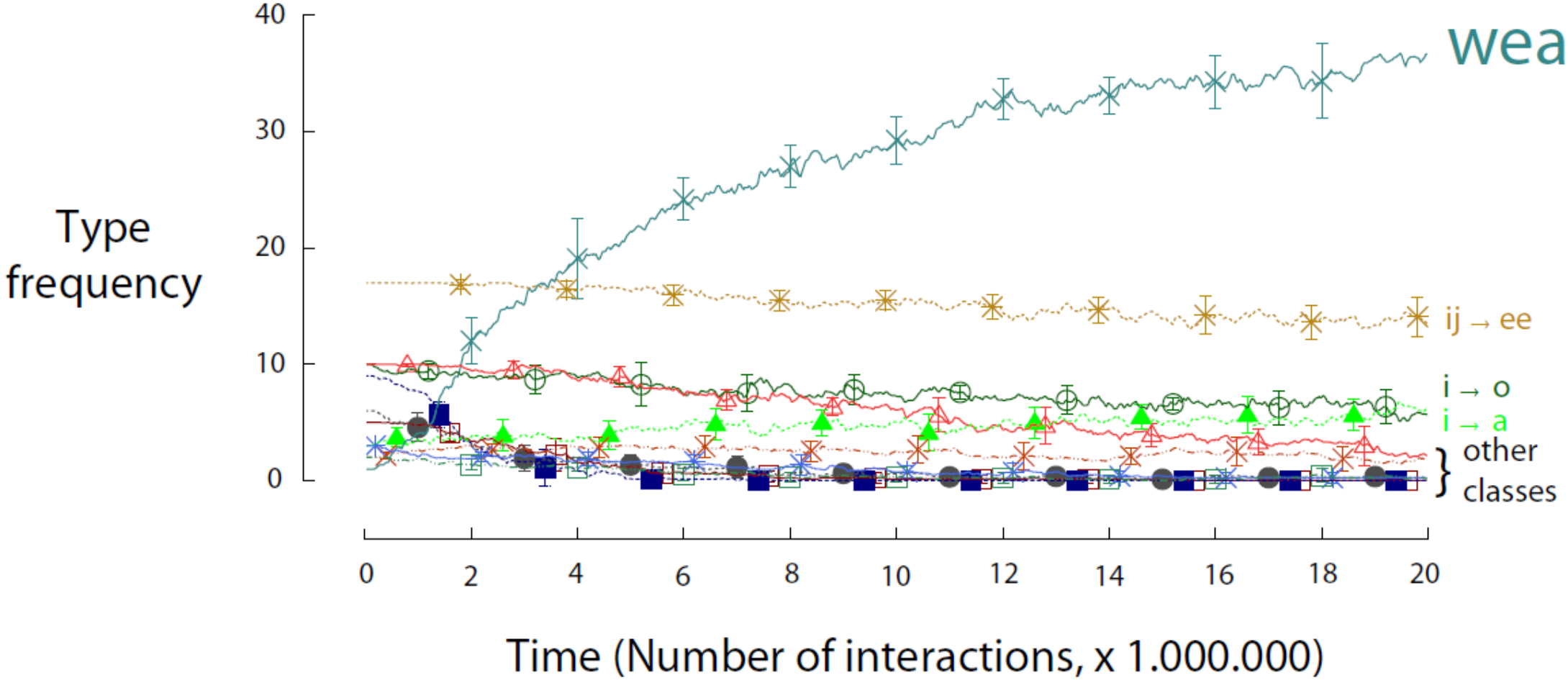
Evaluation criteria: General applicability can explain

1. Rise of the Weak Inflection (Carroll et al. 2012; Cuskley et al. 2014)
2. Gradual Rise (Cuskley et al. 2014)
3. Conserving Effect (Bybee 2006: 715; Lieberman et al. 2007)
4. Class Resilience (Mailhammer 2007; Carroll et al. 2012: 163-164)

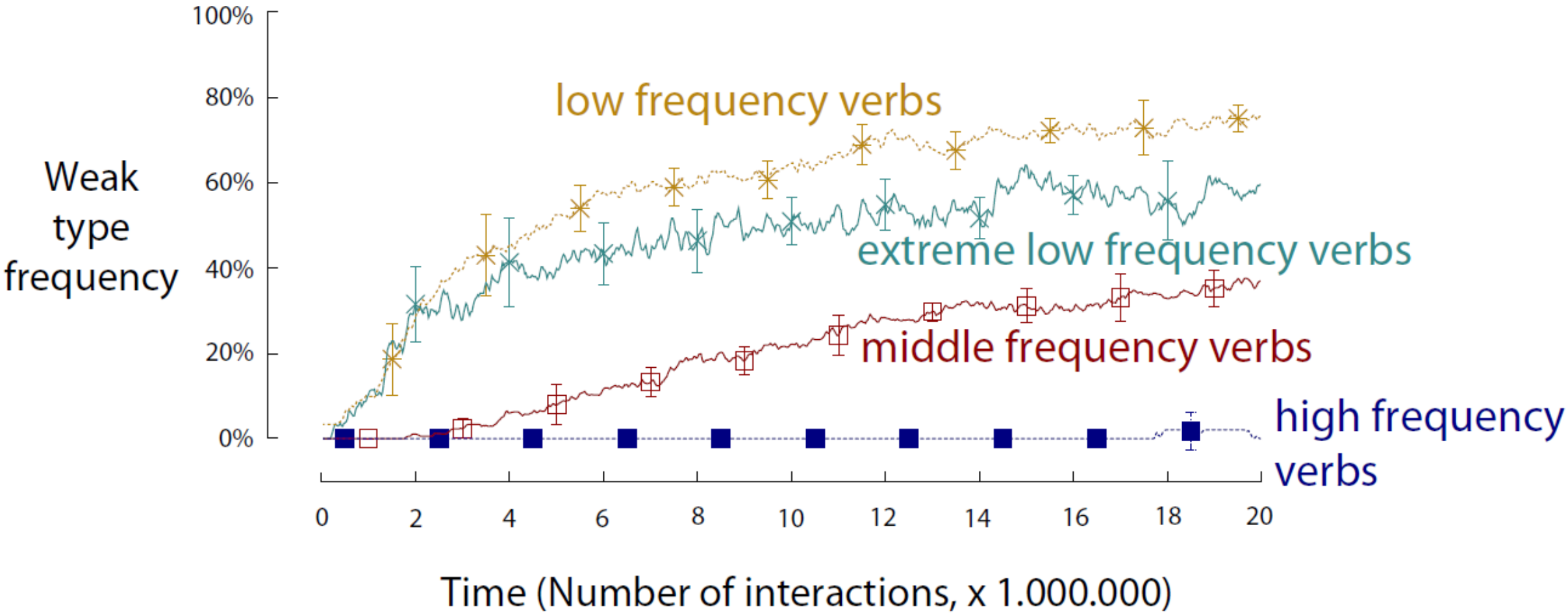
⇨ Emergence should not be dependent on specific parameter settings

⇨ Define AND delimit

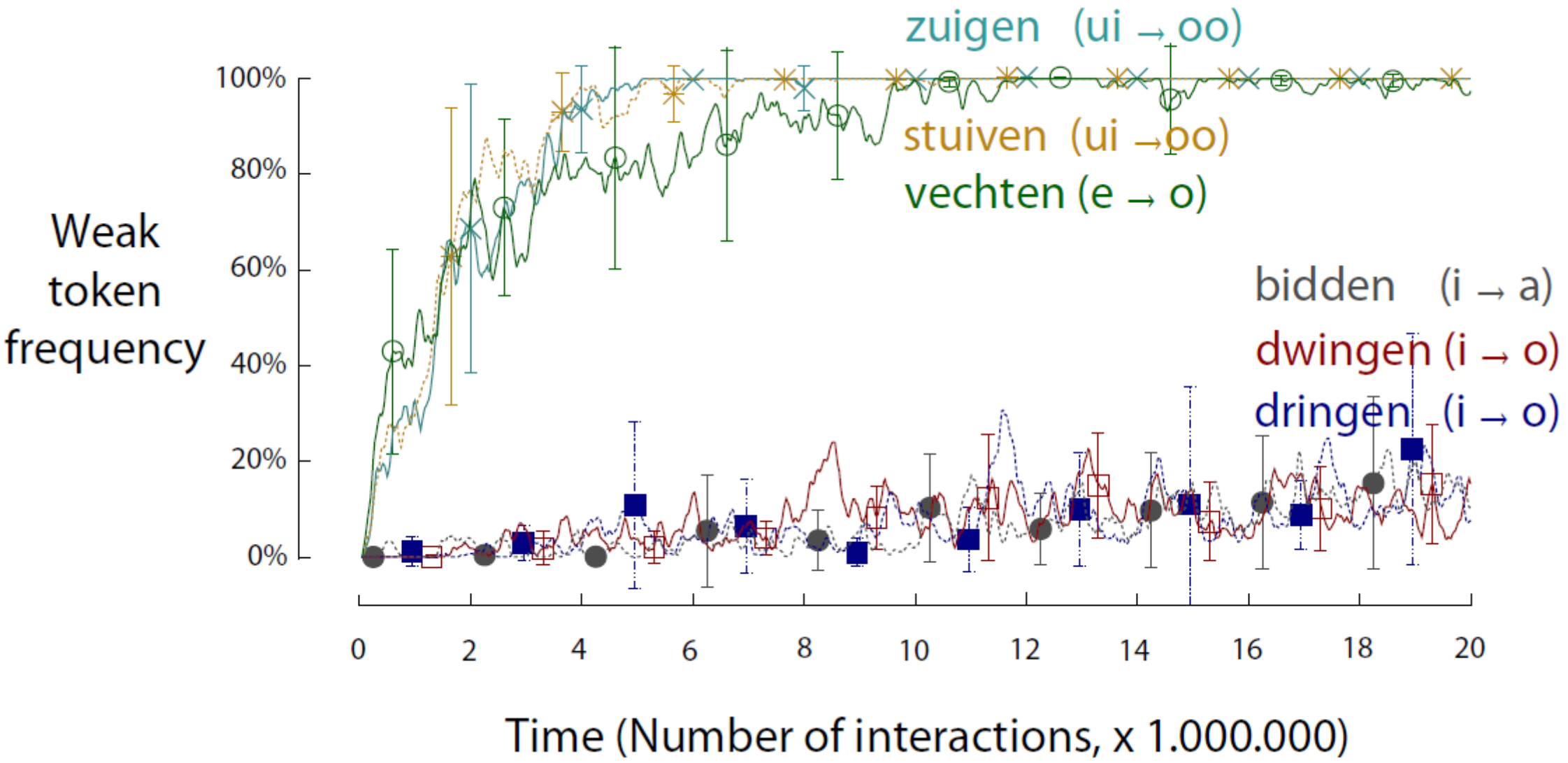

Token frequency
40%
30%
20%
10%
0%
i → o
i → a
ij → ee
weak
aa → oe
ee → a
} other classes
0
2
4
6
8
10
12
14
16
18
20
Time (Number of interactions, x 1.000.000)



1. Rise of the Weak Inflection in token and type frequency
2. Gradual Rise



3. Conserving Effect

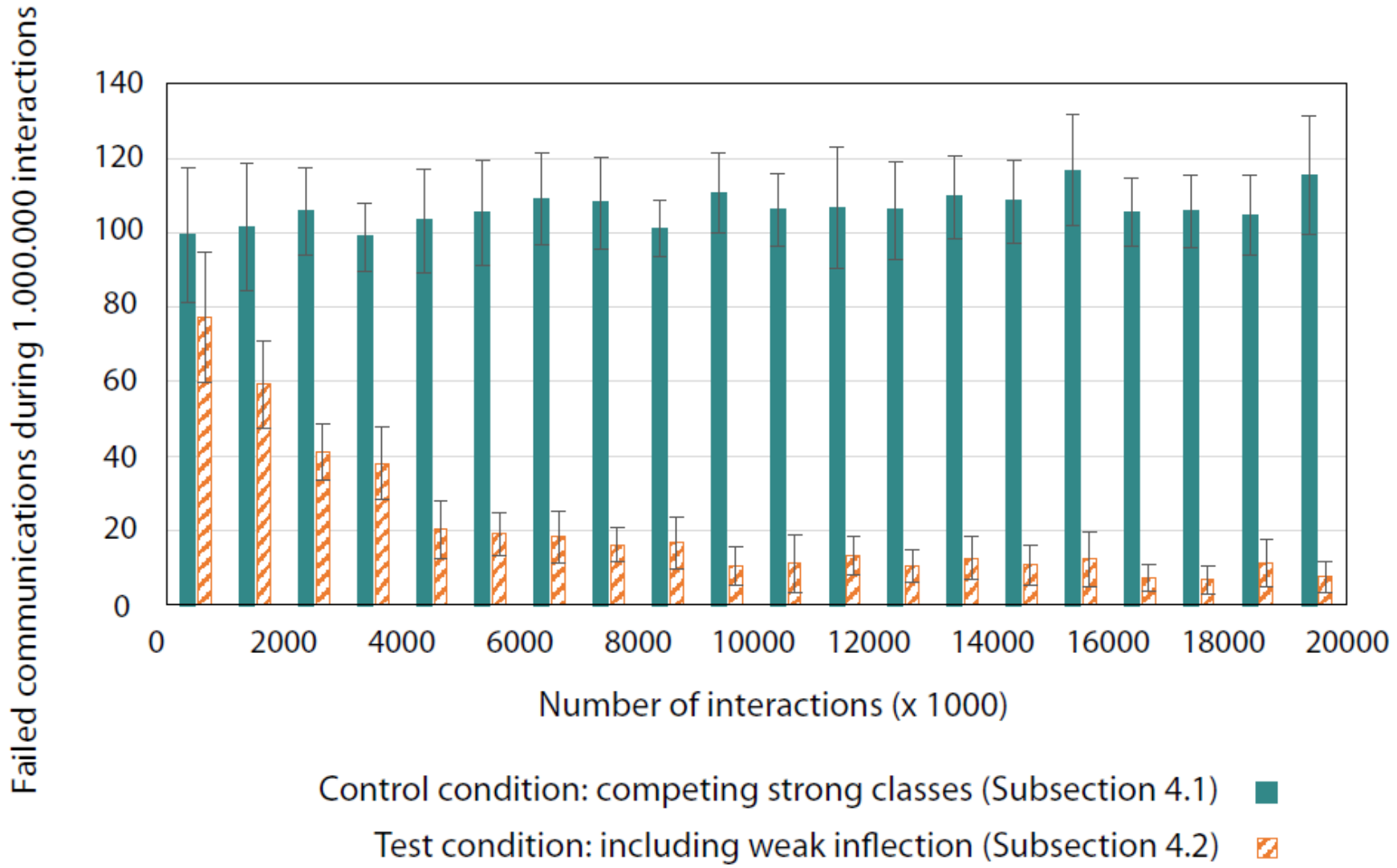


4. Class resilience

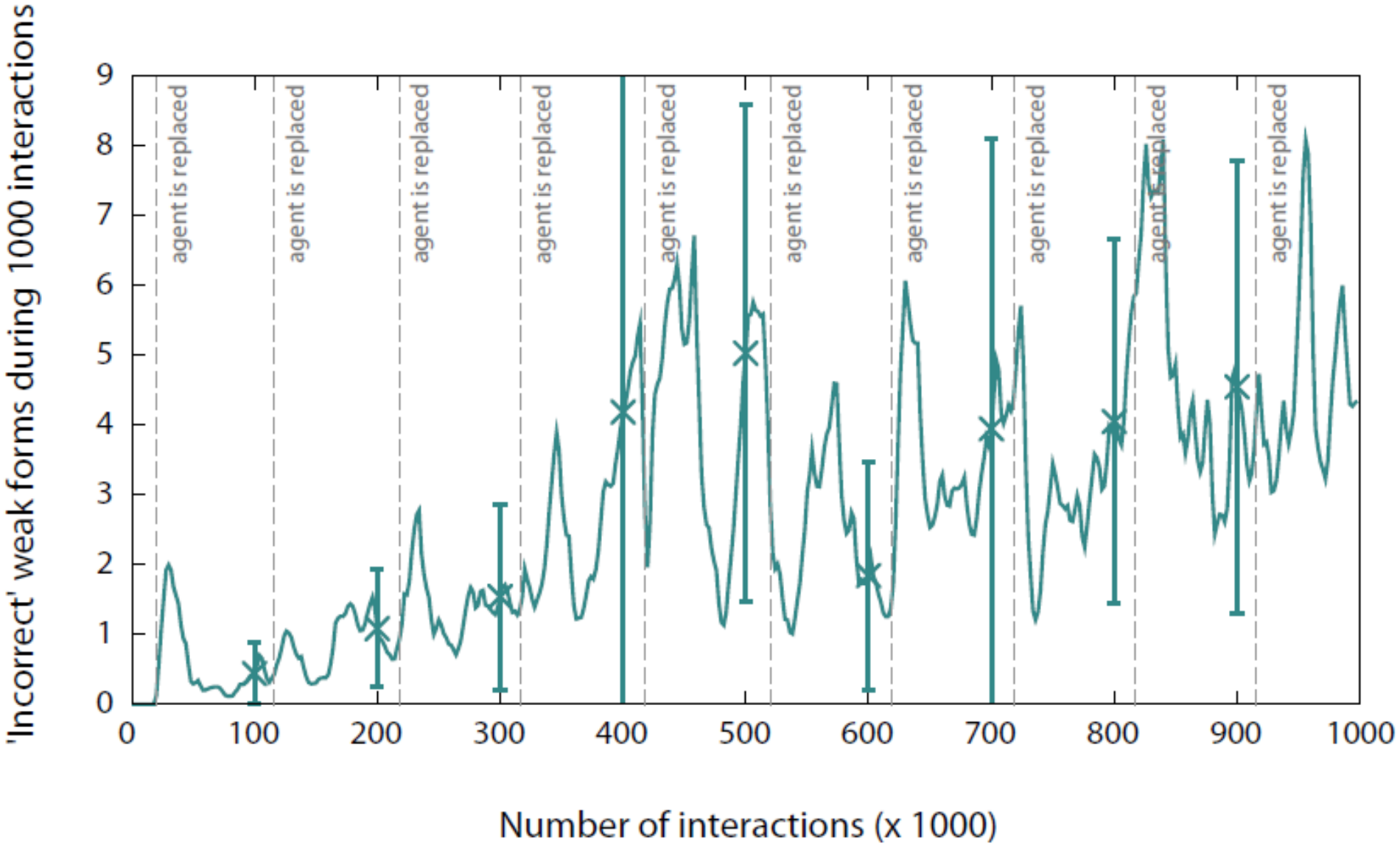
EFFECTS OF THE PARAMETERS

- Number of agents: more agents, slower rise
- Replacement rate: lower replacement rate, slower rise
  - ⇨ Emergence of the evaluation criteria is not dependent upon specific parameter settings
  - ⇨ To kill off the weak inflection, the replacement rate needs to be set extremely high

1. Addition of the weak inflection only complicates past tense inflection



2. For a language user, general applicability is only useful if there are any verbs which he/she cannot (yet) conjugate strongly



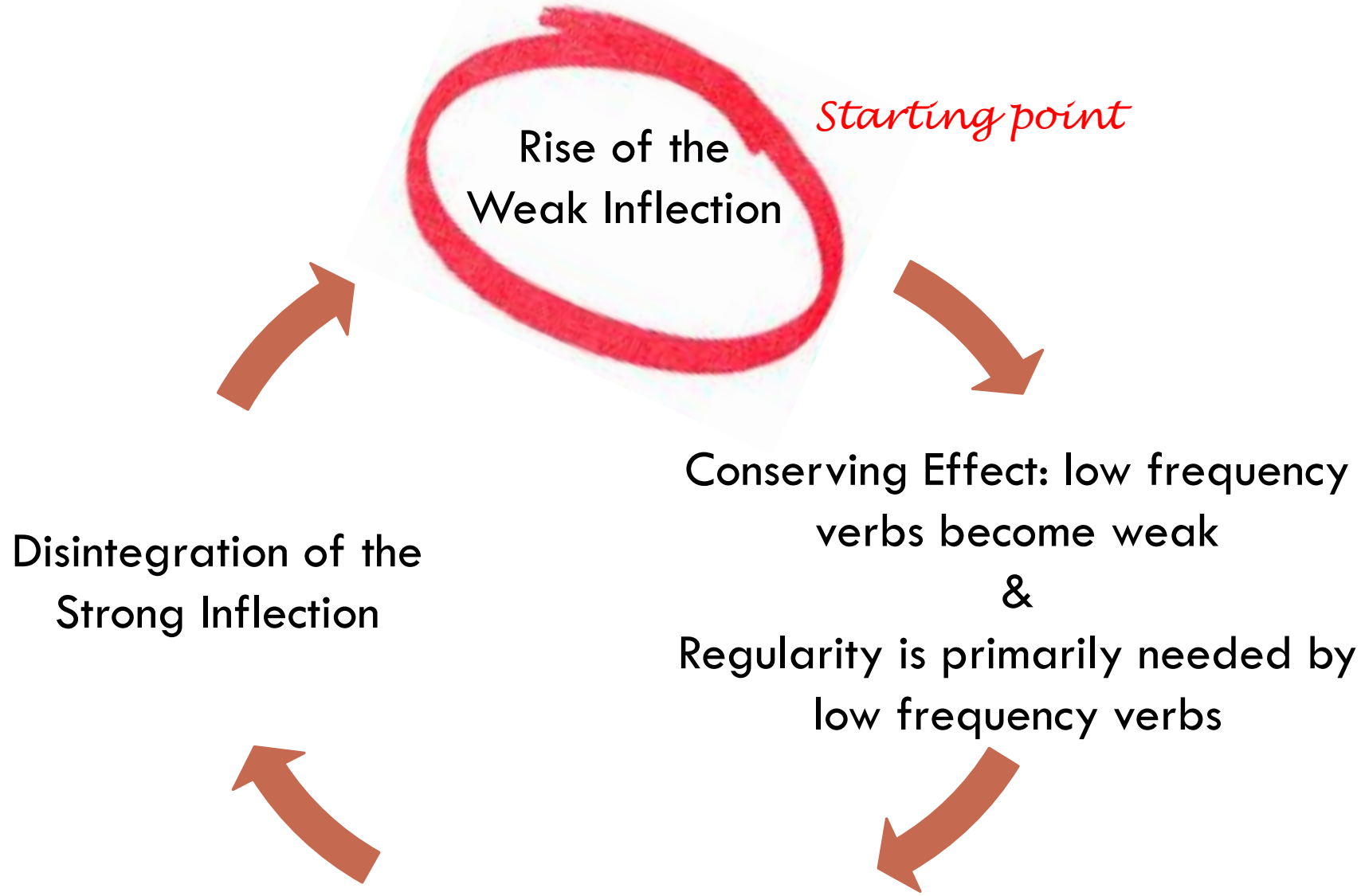
CONCLUSIONS

- The only thing that set the weak inflection apart from the strong classes in our simulation was its general applicability
- This suffices to explain
    1. Rise of the Weak Inflection
    2. Gradual Rise
    3. Conserving Effect
    4. Class Resilience

PROPOSALS

1. **General applicability of the dental suffix**
2. Restrictions on the strong system
3. Disintegration of the strong system

⇨ Disintegration of the strong system may be result, rather than cause


Rise of the
Weak Inflection
Starting point
Conserving Effect: low frequency
verbs become weak
&
Regularity is primarily needed by
low frequency verbs
Less pressure to
maintain the regularity
of the strong system
Disintegration of the
Strong Inflection

FOR FURTHER INFORMATION

Pijpops, Dirk, Katrien Beuls and Freek 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.

REFERENCES

Bailey, Christopher Gordon. 1997. The Etymology of the Old High German Weak Verb. University of Newcastle upon Tyne.

Ball, Christopher. 1968. The Germanic dental preterite. *Transactions of the Philological Society* 67. 162–188.

Beckner, Clay, Joan Bybee, William Croft, Richard Blythe, Morten H Christiansen, Nick C Ellis, Jinyun Ke, Diane Larsen-Freeman, John Holland and Tom Schoenemann. 2009. Language is a complex adaptive system: Position paper. *Language Learning* 59(1). 1–26. doi:10.1111/j.1467-9922.2009.00533.x.

Bentz, Christian and Bodo Winter. 2013. Languages with More Second Language Learners Tend to Lose Nominal Case. *Language Dynamics and Change* 3(1). 1–27.

Bybee, Joan. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82(4). 711–733.

Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.

Carroll, Ryan, Ragnar Svare and Joseph Salmons. 2012. Quantifying the evolutionary dynamics of German verbs. *Journal of Historical Linguistics* 2(2). 153–172.

Colaiori, Francesca, Claudio Castellano, Christine Cuskley, Vittorio Loreto, Martina Pugliese and Francesca Tria. 2015. General three-state model with biased population replacement: Analytical solution and application to language dynamics. *Physical review. E, Statistical, nonlinear, and soft matter physics* 91(1-1). 12808.

Collitz, Hermann. 1912. *Das schwache Praeteritum und seine Vorgeschichte*. Göttingen: Vandenhoeck and Ruprecht.

Croft, William. 2000. *Explaining language change: An evolutionary approach*. Essex: Pearson Education Limited.

Cuskley, Christine, Martina Pugliese, Claudio Castellano, Frascesca Colaiori, Vittorio Loreto and Franscesca Tria. 2014. Internal and External Dynamics in Language: Evidence from Verb Regularity in a Historical Corpus of English. *Plos One* 9(8). e102882.

Eerten, Laura van. 2007. Over het Corpus Gesproken Nederlands. *Nederlandse Taalkunde* 12(3). 194–215.

Gilbert, Nigel. 2008. *Agent-based models*. Los Angeles: Sage.

Hare, Mary and Jeffrey Elman. 1995. Learning and morphological change. *Cognition* 56(1). 61–98.

Hill, Eugen. 2010. A case study in grammaticalized inflectional morphology: Origin and development of the Germanic weak preterite. *Diachronica* 27(3). 411–458.

Landsbergen, Frank. 2009. Cultural evolutionary modeling of patterns in language change: exercises in evolutionary linguistics. Utrecht: LOT.

Lieberman, Erez, Jean-Baptste Michel, Joe Jackson, Tina Tang and Martin Nowak. 2007. Quantifying the evolutionary dynamics of language. *Nature* 449(7163). 713–716.

Ling, Charles and Marin Marinov. 1993. Answering the connectionist challenge: a symbolic model of learning the past tenses of English verbs. *Cognition* 49(3). 235–290.

Loewe, Richard. 1898. Das schwache Präteritum des Germanischen. *Indogermanische Forschungen* 8. 254–266.

Lupyan, Gary and Rick Dale. 2010. Language structure is partly determined by social structure. *PloS one* 5(1). e8559.

Macwhinney, Brian and Jared Leinbach. 1991. Implementations are not conceptualizations: revising the verb learning model. *Cognition* 40(1-2). 121.

Marcus, Gary, Ursula Brinkmann, Harald Clahsen, Richard Wiese and Steven Pinker. 1995. German inflection: the exception that proves the rule. *Cognitive Psychology* 29(3). 189.

Meid, Wolfgang. 1971. *Das germanische Praeteritum*. Innsbruck: Institut für vergleichende Sprachwissenschaft der Universität Innsbruck.

Noord, Rik van. 2015. Modeling the learning of the English past tense with memory-based learning. *Computational Linguistics in the Netherlands (CLIN)*. Antwerp, 6 February.

O'Neil, Wayne. 1978. The evolution of the Germanic Inflection Systems: A Study in the Causes of Language Change. *Orbis* 27. 248–286.

Pijpops, Dirk and Katrien Beuls. 2015 Agent-gebaseerde modellering in de historische taalkunde. Een model van regularisatiedruk op de Nederlandse werkwoorden. *Handelingen der Koninklijke Zuid-Nederlandse Maatschappij voor Taal- en Letterkunde en Geschiedenis*. 69. 5-23.

Pinker, Steven and Alan Prince. 1988. On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* 28(1). 73–193.

Plunkett, Kim and Patrick Juola. 1999. A Connectionist Model of English Past Tense and Plural Morphology. *Cognitive Science* 23(4). 463–490.

Plunkett, Kim and Virginia Marchman. 1991. U-shaped learning and frequency effects in a multi-layered perception: Implications for child language acquisition. *Cognition* 38(1). 43–102.

Plunkett, Kim and Virginia Marchman. 1993. From rote learning to system building: acquiring verb morphology in children and connectionist nets. *Cognition* 48(1). 21–69.

Ringe, Don. 2006. A sociolinguistically informed solution to an old historical problem: the Gothic genitive plural. *Transactions of the Philological Society* 104(2). Oxford, UK. 167–206.

Roberge, Paul. 2010. Contact and the History of Germanic Languages. *The Handbook of Language Contact*, 406–431.

Rumelhart, David and James McClelland. 1986. On learning the past tense of English verbs. In David Rumelhart & James McClelland (eds.), *Parallel distributed processing: explorations in the microstructure of cognition*, 216–271. Cambridge: MIT Press.

Shields, Kenneth. 1982. The origin of the Germanic dental preterite: A new proposal. *Leuvense Bijdragen* 71. 427–440.

Steels, Luc. 2011. *Design Patterns in Fluid Construction Grammar*. Amsterdam: John Benjamins.

Taatgen, Niels and John Anderson. 2002. Why do children learn to say "Broke"? A model of learning the past tense without feedback. *Cognition* 86. 123–155.

Tops, Guy. 1974. *The origin of the Germanic dental preterit*. Leiden: Brill.

Trijp, Remi van, Luc Steels, Katrien Beuls and Pieter Wellens. 2012. Fluid construction grammar: The new kid on the block. *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*. Avignon: ACL.

Yang, Charles. 2002. *Knowledge and learning in natural language*. Oxford: Oxford University Press.

EXTRA SLIDES

40 series of 20.000.000 interactions, 10 agents, replacement rate of 1/20.000

WHY A CORPUS OF MODERN DUTCH?

- No corpora of Proto-Germanic, corpora of Middle-Dutch or Gothic arguably as 'bad' as one from Modern Dutch
- CGN is annotated and more representative of frequency distributions in spoken language
- In principle, any model which complies to the building blocks (slide 14-16) and leads to the emergence of the 4 evaluation criteria will do
  - ⇨ Realistic frequency distributions important
- Intuitively interpretable, but explicitly not a realistic model of Proto-Germanic

WHY IS THE STRENGTH OF A CLASS DETERMINED BY TOKEN INSTEAD OF TYPE FREQUENCY?

- No Advantages for the weak inflection: Type frequency would be more beneficial for the weak inflection than token frequency (Conserving Effect)
- KISS: More design choices need to be made for type frequency, e.g. how do you exactly measure it? What to do with verbs that show variation? Does one occurrence of ‘vraagde’ count for as much as 1000 occurrences of ‘vroeg’?

WHY DO THE FREQUENCIES OF THE GRAMMATICAL CONSTRUCTIONS ONLY PLAY A ROLE IF THE AGENT HAS NEVER HEARD THE VERB BEFORE?

Alternative: formula that takes into account both the frequencies of the lexical and grammatical constructions.

- More realistic, but also more complex: necessitates the inclusion of two more parameters
- Current approach makes the agents highly conservative. If anything, this impedes the rise of the weak inflection

IS IT NOT REDUNDANT FOR THE AGENTS TO KEEP BOTH THE GRAMMATICAL AND LEXICAL CONSTRUCTIONS IN MEMORY?

Yes it is, but is is also very minimal in its assumptions. It only assumes that any pattern that is recognized by humans will become more entrenched in their memory if they encounter it more often.

The alternative is a rule-list approach, which assumes that regular and irregular forms are handled fundamentally different by agent memory. That is a quite expensive assumption.

WHAT IF: NO RULES, ONLY ANALOGY?

Model does not need to change: the frequency of the grammatical constructions is exactly equal to the sum of the frequencies of its verb forms.

WHY DO WE USE ONLY ONE DENTAL SUFFIX IF THERE ARE GERMANIC LANGUAGES WITH MULTIPLE WEAK CLASSES, E.G. ICELANDIC?

One ‘dental suffix’ means that, each time an agent hears any past form with a dental suffix, this dental suffix becomes **more entrenched** in its memory. It is this dental suffix that is available for all verbs to form their past tense.

Conversely, if an agent hears a past form conjugated according to the first strong class (Dutch *ij* → ee), then **only this class** becomes more entrenched in the agent’s memory, and not the second class (Dutch *ie* → oo). This is the simulation’s core assumption of **general applicability**.

If you disagree with this assumption, you are wellcome to build another simulation. If you can show the same effects, using less assumptions, you have **disproven the current simulation**.

WHY AN AGENT-BASED MODEL (AND NOT ONE OF ITERATED LEARNING?)

- General applicability is usage property
- Usage-based view on language change (Croft 2000, Bybee 2010)
- Language as a Complex Adaptive System (Gilbert 2008, Beckner et al. 2008)

- Models of iterated learning focus on the acquisition of the Germanic past tense, as a case study of language acquisition in general:

Rumelhart and McClelland (1986), Pinker and Prince (1988), Macwhinney and Leinbach (1991), Plunkett and Marchman (1991, 1993), Ling and Marinov (1993), Hare & Elman (1995), Marcus et al. (1995), Plunkett and Juola (1999), Taatgen and Anderson (2002), Yang (2002), van Noord (2015)

WHAT ARE SOME OF THE QUESTIONS THE PRESENT SIMULATION CANNOT ANSWER?

- Origin of the dental suffix (o.a. Loewe 1898; Collitz 1912; Ball 1968; Meid 1971; Tops 1974; Shields 1982; Ringe 2006: 179-785; Hill 2010)

- What originally made the strong system so succesful?
  - Shorter verb forms
  - Germanic first-syllable stress

  ⇨ Influx of L2-learners: advantages of the weak inflection – general applicability and greater linear segmentability – proved more decisive

(cf. O'Neil 1978; Roberge 2010; Lupyan and Dale 2010; Bentz and Winter 2013)