

Herfstvergadering Koninklijke Zuid-Nederlandse Maatschappij voor Taal- en
Letterkunde en Geschiedenis, 24 oktober 2014, Gent

Dirk Pijpops, KU Leuven, onderzoeksgroep QLVL



AI Lab



Agent-gebaseerde modellering

Een testbaan voor de historische taalkunde?



Overzicht

- ▶ Agent-gebaseerde modellen
- ▶ The verb-inflection game



Agent-gebaseerde modellen

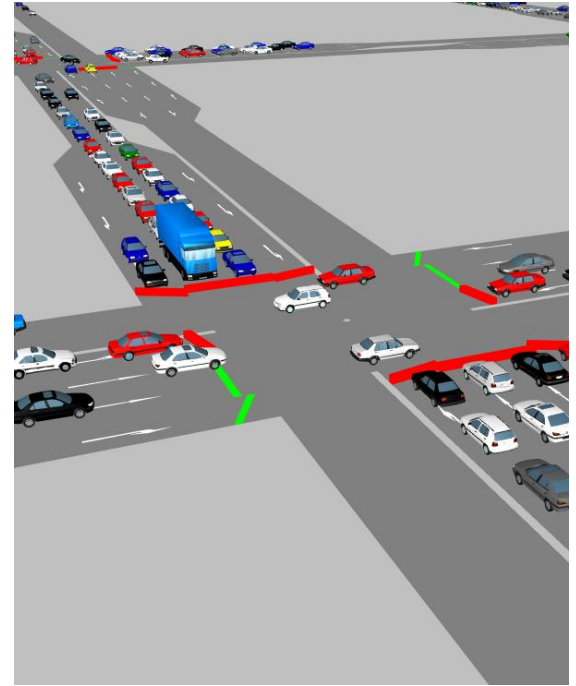
Wat?



Agent-gebaseerde modellen



Agent-gebaseerde modellen

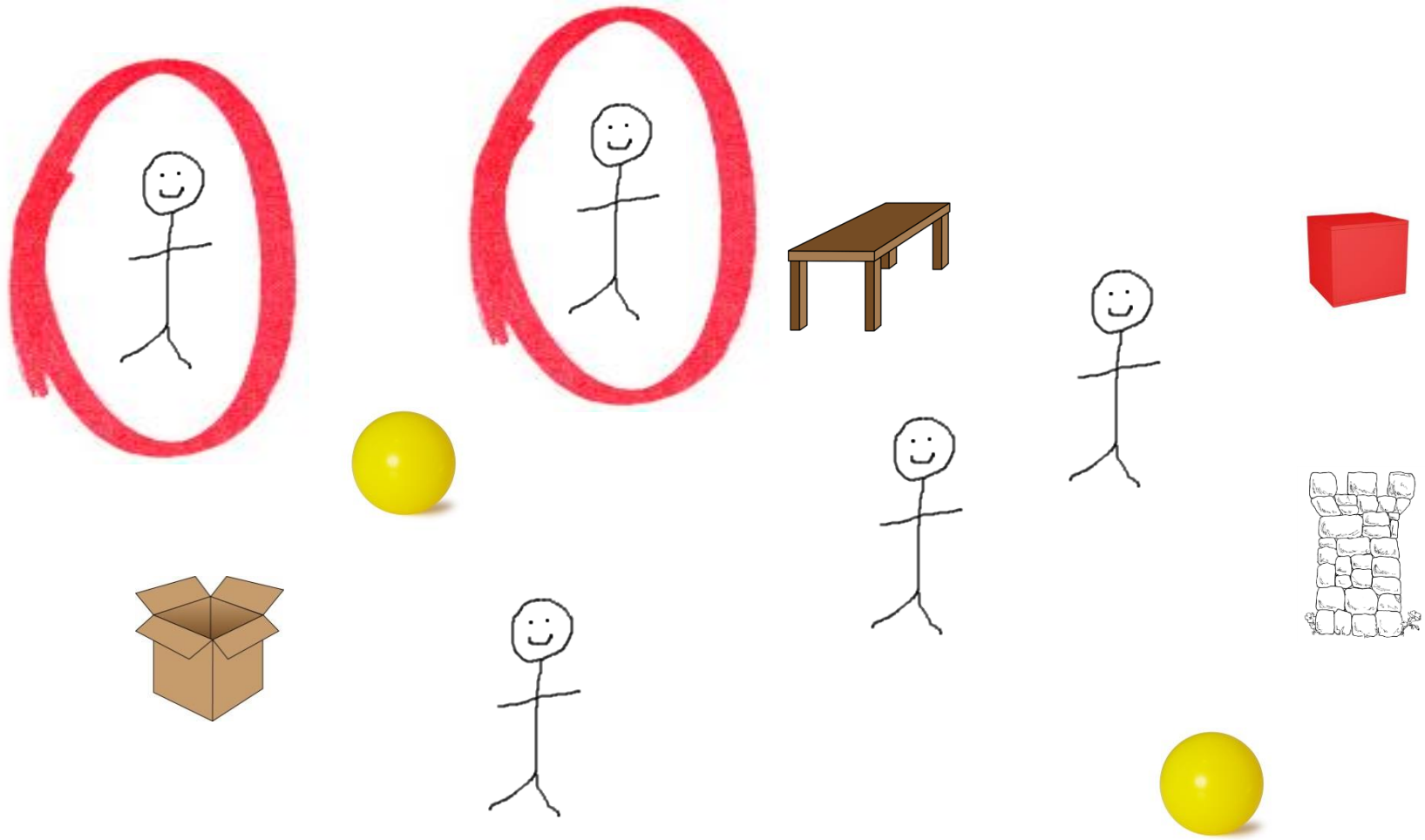


Agent-gebaseerde modellen

Evolutionaire linguïstiek



Taalspel

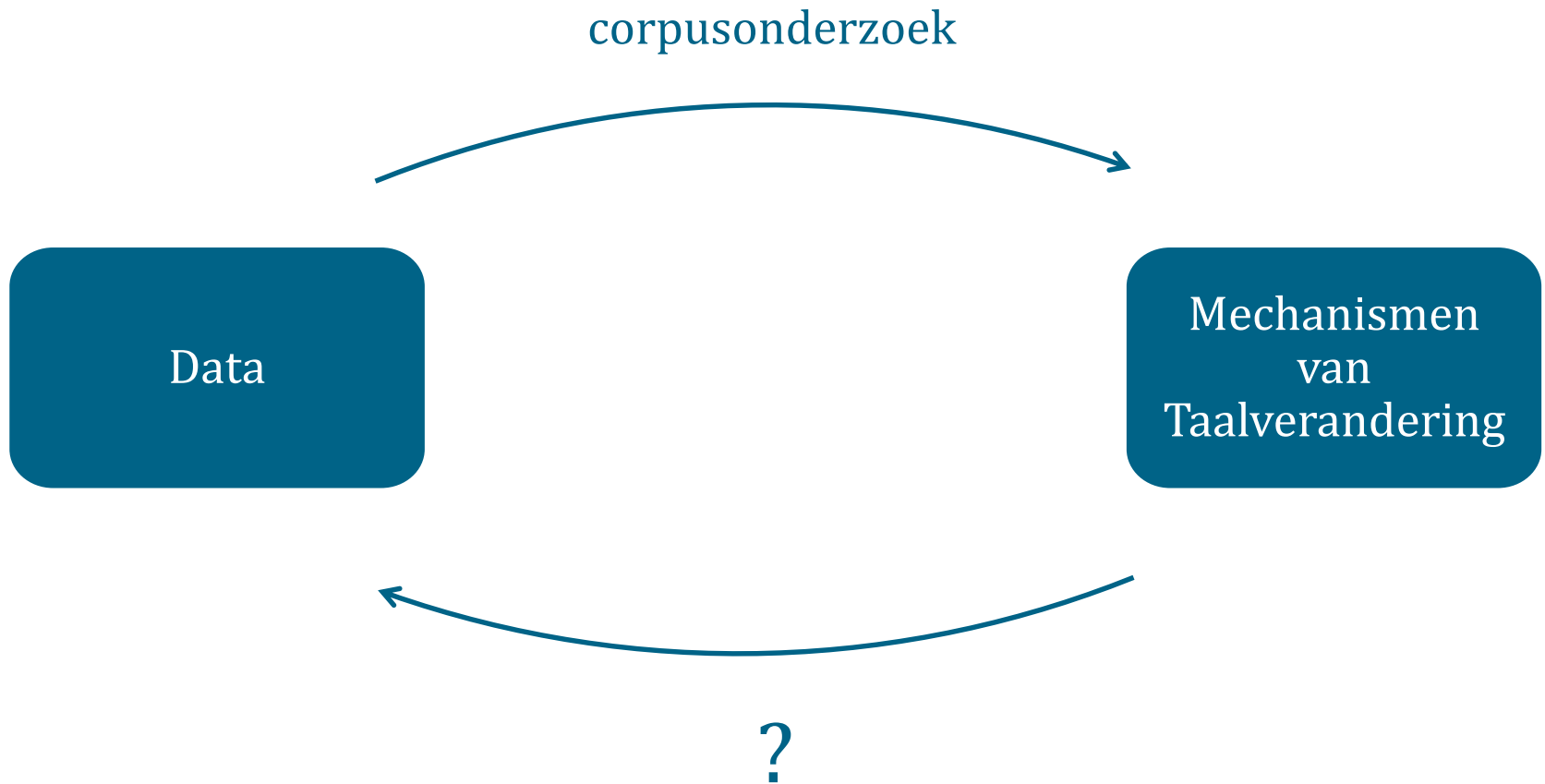


Agent-gebaseerde modellen

Waarom?



Agent-gebaseerde modellen



Agent-gebaseerde modellen

⇒ Kan het zo werken?

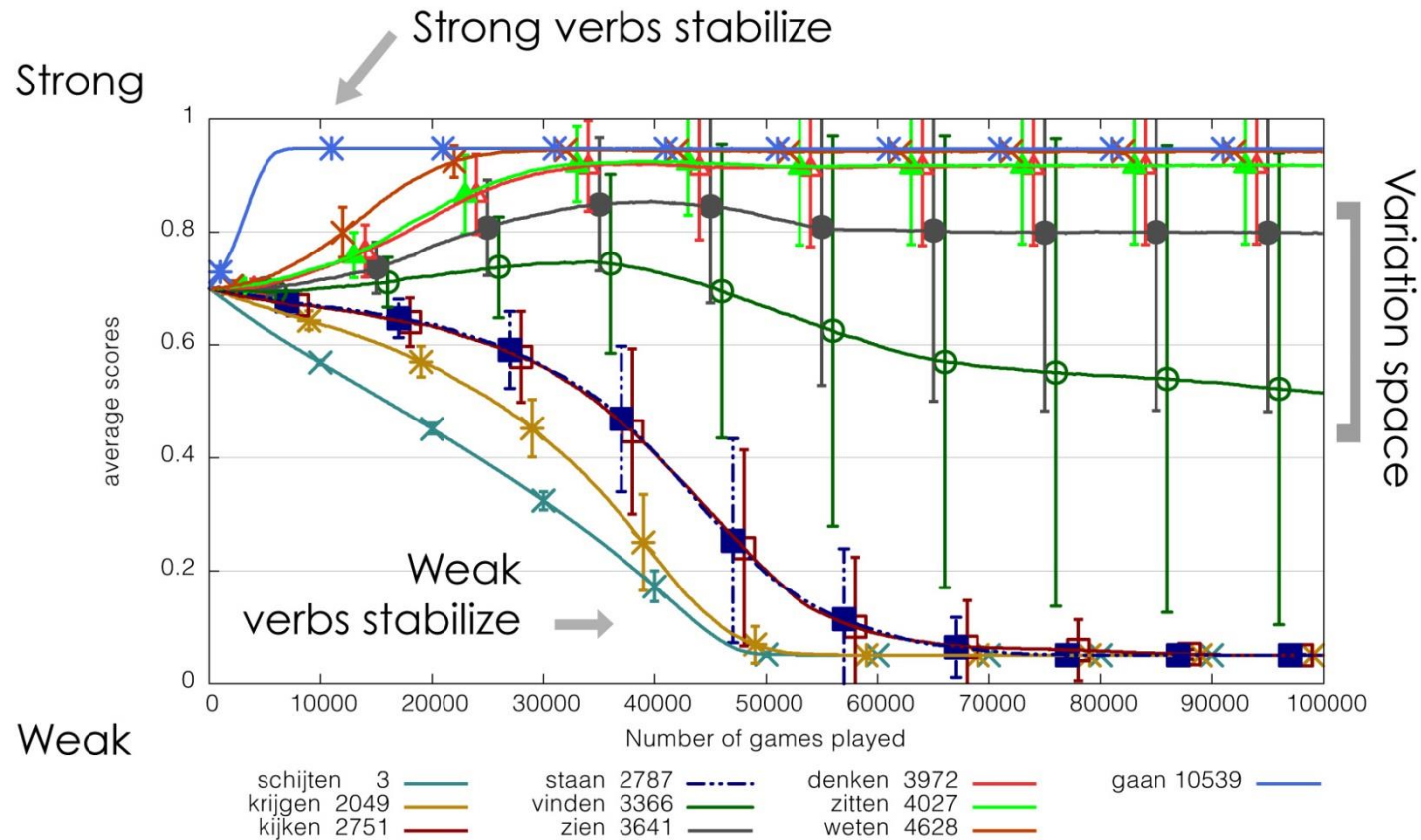
- ▶ van Trijp, Remi. 2012. Not as Awful as it Seems: Explaining German Case through Computational Experiments in Fluid Construction Grammar. *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*. Avignon: France.
- ▶ van Trijp, Remi. 2012. Self-Assessing Agents for Explaining Language Change: A Case Study in German. *Proceedings of the 20th European Conference on Artificial Intelligence*. Montpellier: France.
- ▶ van Trijp, Remi. 2013. Linguistic Selection Criteria for Explaining Language Change: A Case Study on Syncretism in German Definite Articles. *Language Dynamics and Change*, 3(1):105-132. doi: 10.1163/22105832-13030106
- ▶ van Trijp, Remi. 2014. Fitness Landscapes in Cultural Language Evolution: A Case Study on German Definite Articles. To appear in *The Evolution of Language* (Evolang-X).



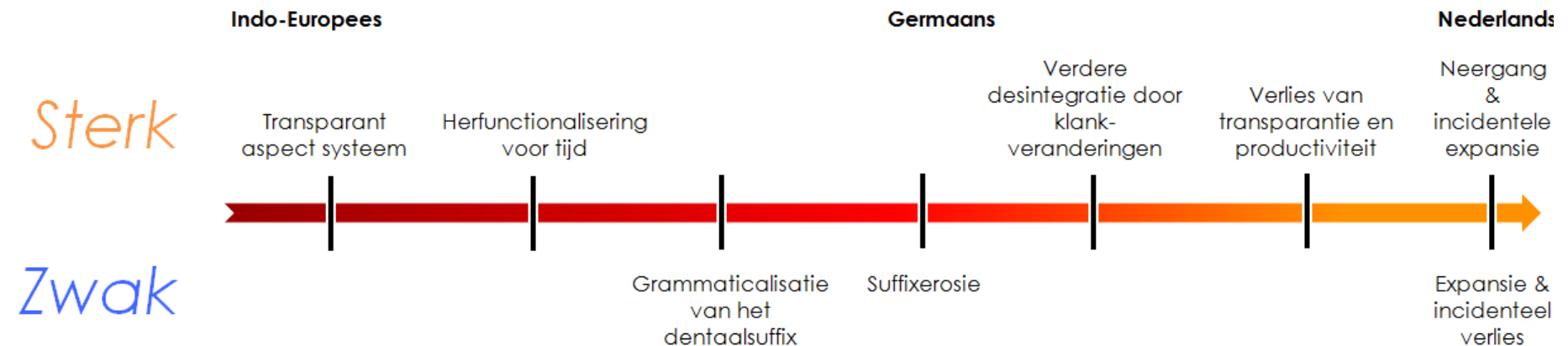
The verb-inflection game



The verb-inflection game



Sterke en zwakke werkwoordsvervoegingen



- ▶ Pijpops, Dirk, Katrien Beuls, Pieter Wellens & Freek Van de Velde. 2014. Entrenchment vs. Transparency. Modelling the Dutch strong-weak past tense competition in an agent-based simulation. *Poster presentation at the Evolang X Workshop 'How grammaticalization processes create grammar'*. Vienna, 14 April.

Sterke en zwakke werkwoordsvervoegingen

- ▶ Sterke ablaut alternatie *loop* → *l-ie-p*
- ▶ Zwakke suffix inflectie *loop* → *loop-te*
- ▶ Sterke Verankering vs. Zwakke Transparantie

⇒ Kunnen deze twee krachten elkaar in balans houden, en zo ja, hoe?

- ▶ Lieberman, E., Michel, J.-B., Jackson, B., Tang, T., & Martin, A.. 2007. Quantifying the evolutionary dynamics of language. *Nature* 449 (7163), 713-716.
- ▶ Pugliese, M., Loreto, V., Cuskey, C., Castellano, C., Colaiori, F. & F. Tria. 2014. 'The role of coordination in regularization'. E.A. Cartmill, S. Roberts, H. Lyn & H. Cornish (eds.) *The Evolution of Language*. London: World Scientific. 260-261.



Basisontwerp: wereld

- ▶ 200 events met CGN-frequenties
- ▶ 10 agents
- ▶ Event gebeurt → 2 agents willekeurig geselecteerd om taalspel te spelen

- ▶ Van Eerten, Laura. 2007. Over het Corpus Gesproken Nederlands' *Nederlandse Taalkunde* 12(3): 194-215.



Basisontwerp: lexicon

- ▶ Elke agent
- ▶ Continu geüpdatet
- ▶ Sterke verankering

Event	Strong form	Strong frequency	Weak form	Weak frequency	Score
kijken	“keek”	0	“kijkte”	0	0.7
vragen	“vroeg”	0	“vraagde”	0	0.7
...					
hangen	“hing”	0	“hangde”	0	0.7
...					
wreken	“wrook”	0	“wreekte”	0	0.7



Alignment strategieën

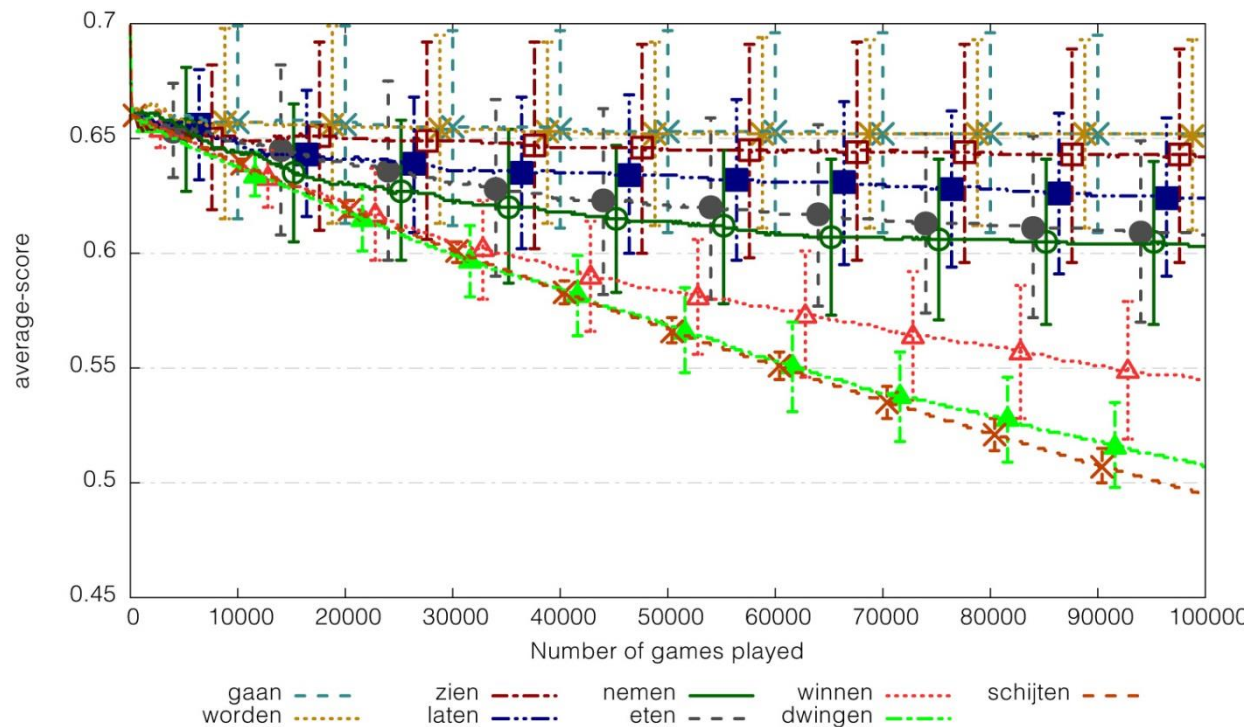
- ▶ Hoe passen agents hun scores aan na een interactie?
- ▶ Zwakke transparantie
- ▶ Verschillende probabilistische strategieën vergeleken
 - ▶ Frequentieverhouding
 - ▶ Lineaire alignment
 - ▶ ...



Alignment strategieën: frequentieverhouding

► Score van *lopen* =
$$\frac{\# \textit{liep}}{\# \textit{liep} + \# \textit{loopte}}$$





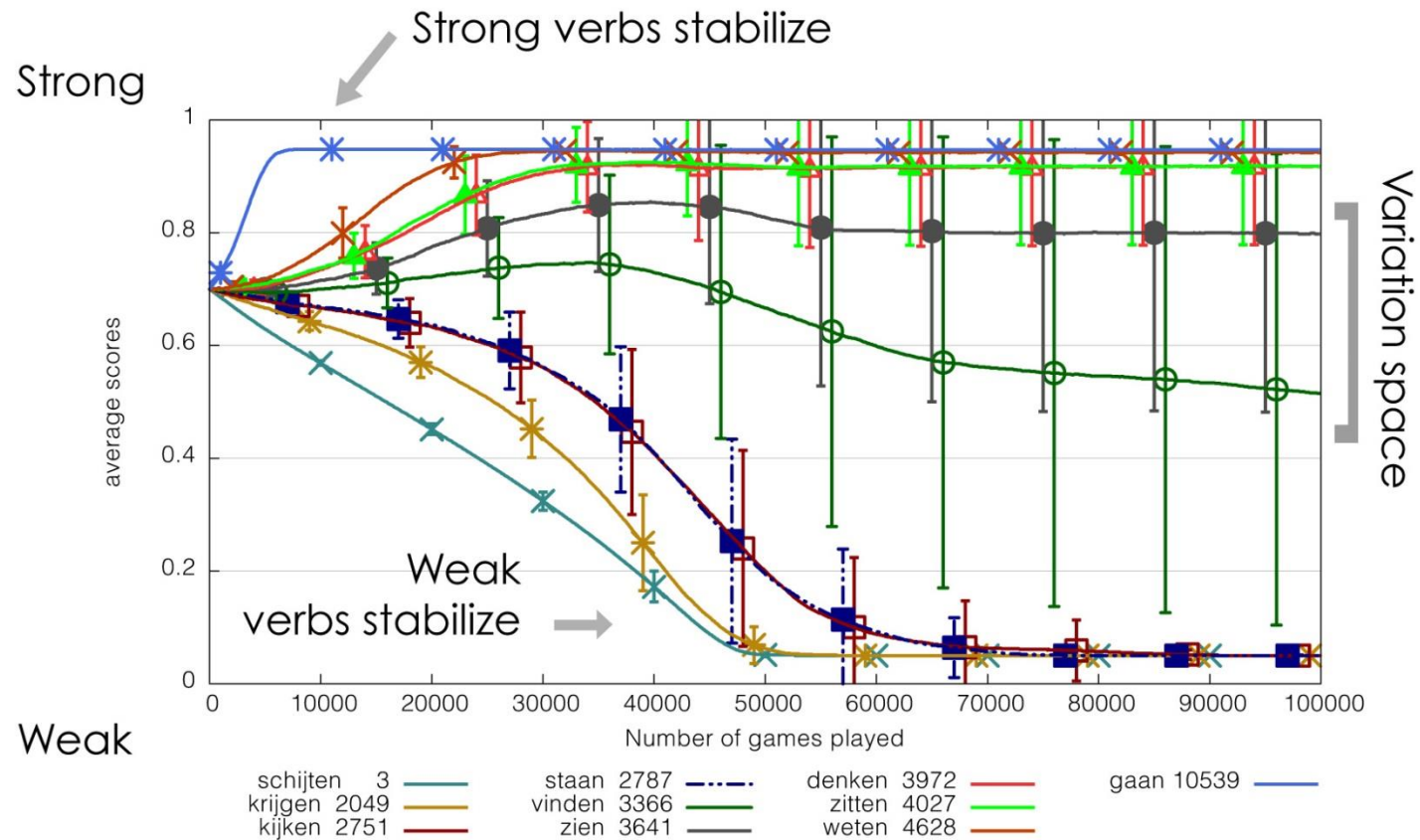
Alignment Strategieën: lineaire alignment

- ▶ Hoor “liep” score “lopen” + 0.01
- ▶ Hoor “loopte” score “lopen” - 0.01
 alle scores - 0.0005



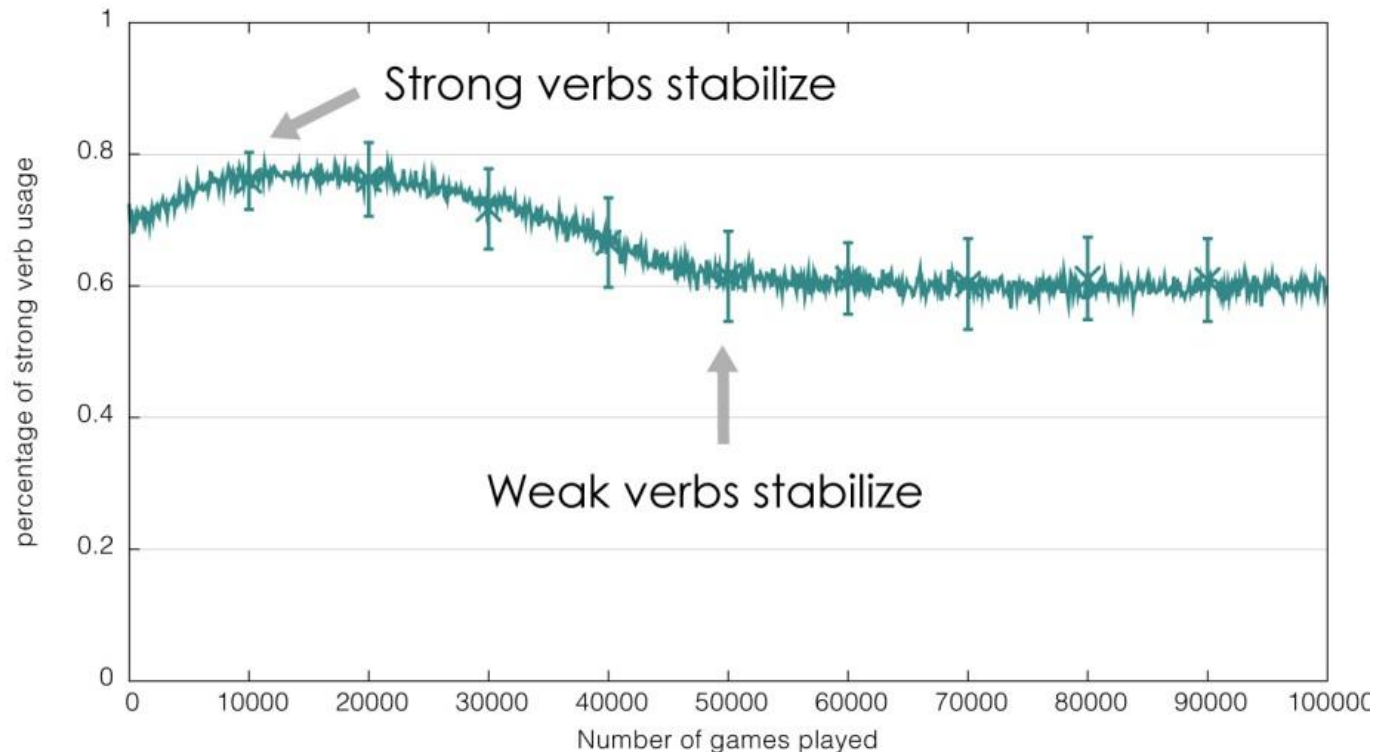
Alignment Strategieën: Delta accumulatie

► Ontwikkeling gemiddelde scores



Interpretatie basismodel

► Verhouding token frequentie



Interpretatie: Parameter aanpassingen

- ▶ Aantal agents: Des te meer agents, des te trager de convergentie
- ▶ s_{init} : Des te hoger de beginscores, des te meer werkwoorden sterk blijven
- ▶ δ_{gen} : Stel de sterkte van het algemeen effect in
- ▶ δ_{spec} : Stel de sterkte van het werkwoordspecifieke effect in
- ▶ Prestigieuze agents: veroorzaken sterke initiële onenigheid
versnellen convergentie op lange termijn
vergroten variatieruimte



Conclusies

- ▶ Verankering vs. Transparantie: lineaire alignment
- ▶ Hoogfrequent vs. Laagfrequent
- ▶ Tokenfrequentie vs. Typefrequentie
- ▶ Verdere uitbreiding mogelijk, bv. prestigieuze agents
- ▶ Volgende stappen: sterke ablautklassen + productiviteit
=> constructiegrammatica in FCG
- ▶ Strik, O. & R. Knooihuizen. In prep. 'Relative productivity potentials of Dutch verbal inflection patterns'.



Bedankt aan

- ▶ Katrien Beuls
- ▶ Pieter Wellens
- ▶ Freek Van de Velde
- ▶ Luc Steels
- ▶ Remi van Trijp
- ▶ Jullie



The verb-inflection game

- ▶ Lisp

- ▶ Babel2-framework

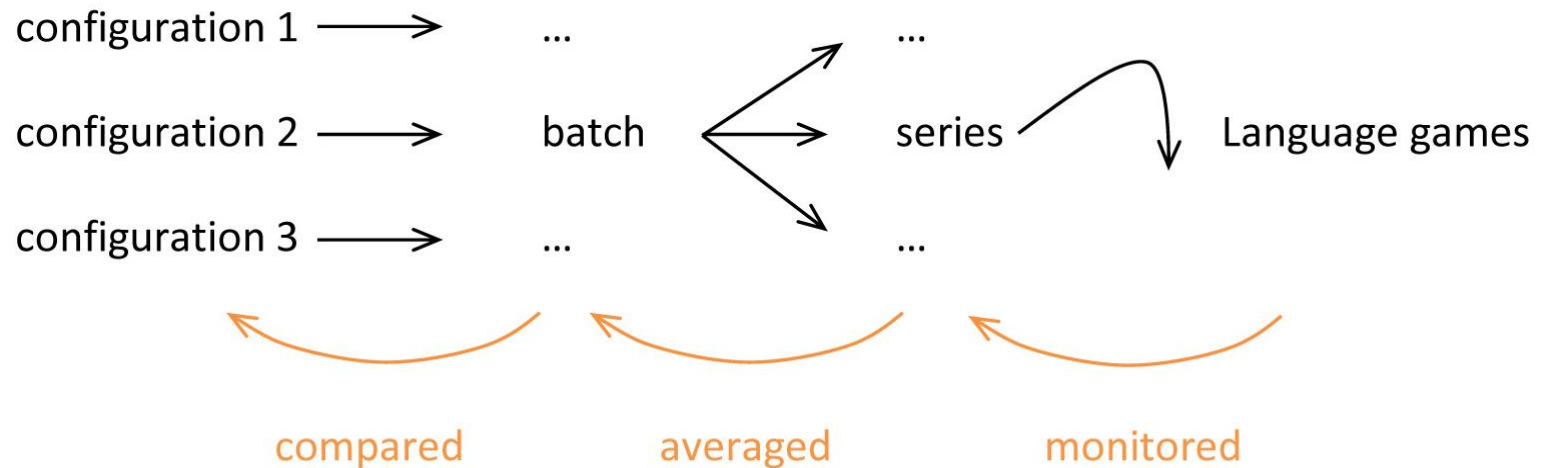
gratis te downloaden op <http://emergent-languages.org/Babel2/index.html>

- ▶ Competitie tussen sterke en zwakke
werkwoordsvervoeging

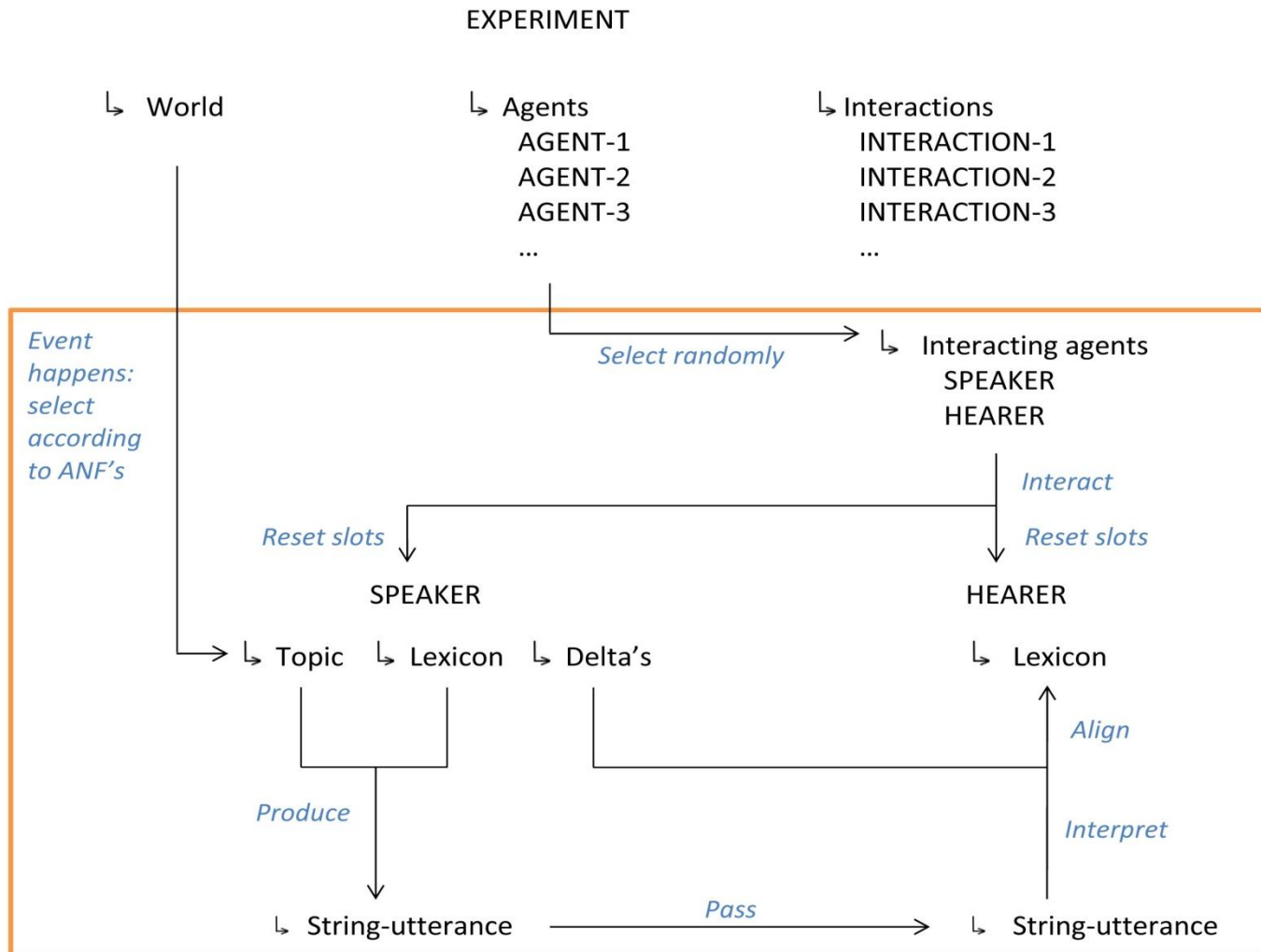


Basisontwerp: batch

- Batches met 40 series van elk 100.000 interacties

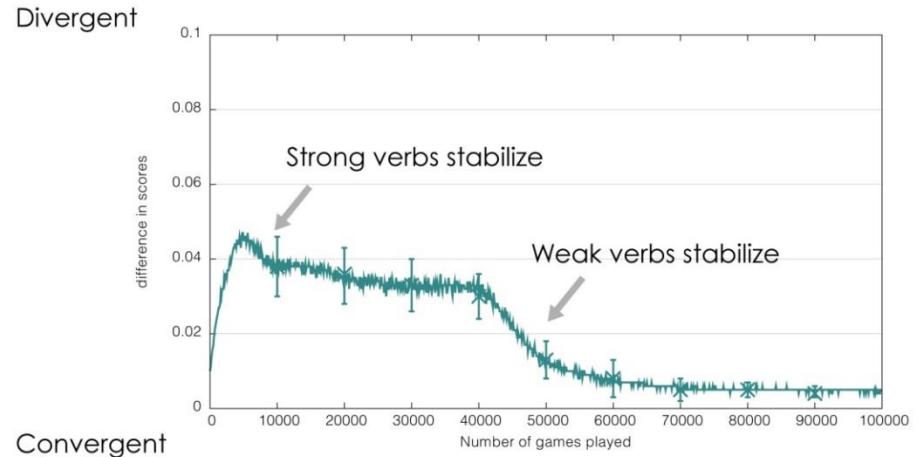


Basisontwerp: interactie

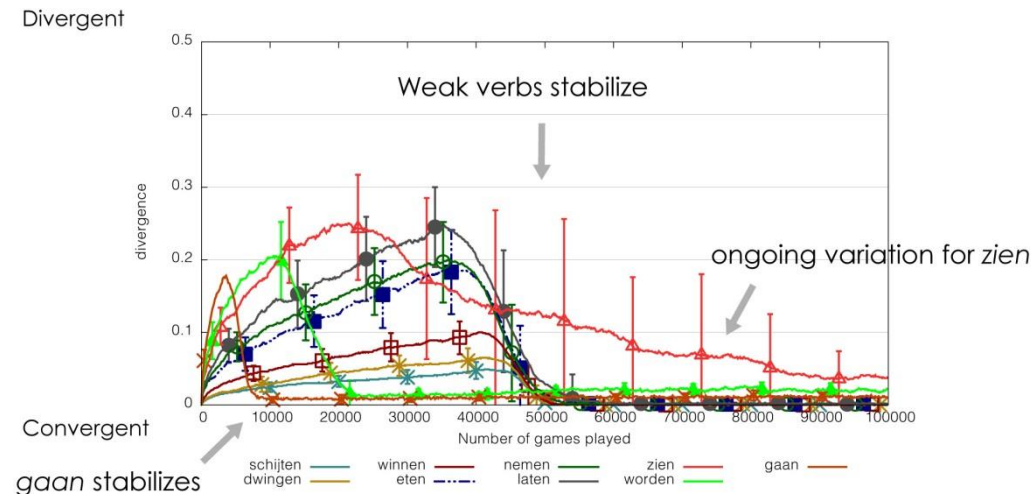


Interpretatie basismodel

- ▶ Hoorder-spreker convergentie



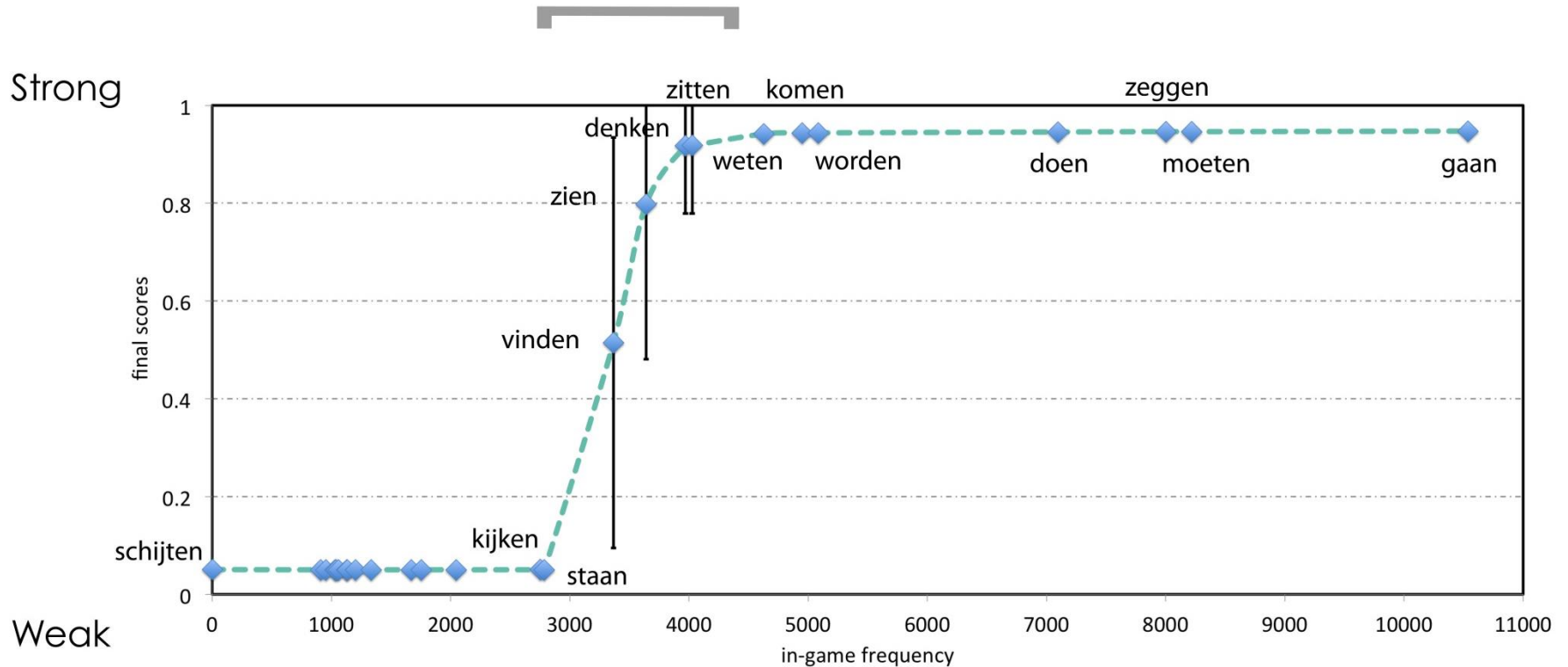
- ▶ Werkwoords-convergentie



Interpretatie basismodel

► Eindscores

Variation space



Alignment strategieën: Frequentieverhouding

- ▶ Score herberekend na elke interactie:

$$s_n = \frac{Count_{strong}(ev_n)}{Count_{strong}(ev_n) + Count_{weak}(ev_n) + \delta_{gen} \sum_{i=1}^{200} Count_{weak}(ev_i)}$$

