Agent-gebaseerde modellering in de historische taalkunde. Een model van regularisatiedruk op de Nederlandse werkwoorden

Pijpops, Dirk; Beuls, Katrien

Published in:
Handelingen van de Koninklijke Zuid-Nederlandse Maatschappij voor Taal- en Letterkunde en Geschiedenis

Publication date:
2015

Document Version
le PDF de l'éditeur

Agent-gebaseerde modellering in de historische taalkunde. Een model van regularisatiedruk op de Nederlandse werkwoorden[1]

Dirk PIJPOPS & Katrien BEULS

Abstract

This article presents an accessible introduction to agent-based modelling for non-computational linguists, more specifically historical linguists, who may find it to be a useful addition to their methodological toolbox. We will discuss both the possibilities and limitations of this technique and in doing so, attempt to show how computational modelling can be complementary with the empirical methods of historical linguists. Furthermore, this paper presents a walkthrough of the various steps in the design of such a model. To serve as an example, we will present a basic agent-based model of the competition between regular and irregular verb inflection in Dutch.

1. INLEIDING

Het belangrijkste doel van deze paper is een toegankelijke inleiding te bieden van agent-gebaseerde modellering, specifiek voor (historische) taalkundigen die niet bekend zijn met deze methodologie. Een bijkomend doel is de voorstelling van een eenvoudig agent-gebaseerd model dat daarbij als voorbeeld dient. De klemtoon ligt dan ook op de methodologische aspecten van deze techniek, veeleer dan op de resultaten van het voorbeeldmodel.

Agent-gebaseerde modellen zijn computationele simulaties waarbij de invloed van het gedrag van individuele delen – het microniveau – op het grotere geheel – het macroniveau – wordt bestudeerd. Concreet zijn de individuele delen agents die met elkaar interageren volgens typisch erg eenvoudige regels. Vervolgens wordt onderzocht hoe deze eenvoudige interacties op het macroniveau grotere structuren of veranderingen kunnen laten ontstaan. Binnen andere wetenschappen worden deze modellen dan ook voornamelijk gebruikt om emergente fenomenen te besturen, zoals het gedrag van insectenkolonies (Bonabeau, Dorigo & Theraulaz 1999) of van het internet (Dhamdher & Dovrolis 2009). Binnen de linguïstiek lijkt hun gebruik voorlopig grotendeels beperkt tot het veld van de evolutionaire taalkunde, waar men in de eerste plaats geïnteresseerd is hoe taal kan ontstaan, en niet zozeer hoe ze veran-

[1] We zouden graag beide anonieme beoordelaars bedanken voor hun erg nuttige suggesties bij de herwerking van dit artikel. Daarnaast bedanken we ook Freek Van de Velde om een eerdere versie na te lezen.

dert (zie o.a. Steels 1998; Cangelosi & Parisi 2002; Morten & Kirby 2003; Beuls & Steels 2013; Steels & Garcia-Casademont 2015. Voor agent-gebaseerde modellen die wel focussen op taalverandering, zie o.a. Landsbergen 2009; Landsbergen et al. 2010; Blythe & Croft 2012; van Trijp 2013; Pijpops, Beuls & Van de Velde 2015; Bloem, Versloot & Weerman 2015, en verwijzingen in Bhattacharjee 2004). De theoretische achtergrond daarbij is die van culturele evolutie.[2] Tot nu toe is er weinig sprake van uitwisseling tussen de evolutionaire en historische taalkunde. Dat is betreurenswaardig, aangezien enerzijds de historische taalkunde al een meerdere mechanismen heeft blootgelegd waardoor nieuwe taalstructuren kunnen ontstaan, zoals grammaticalisatie, heranalyse of analogiewerking, terwijl anderzijds de evolutionaire taalkunde dankzij haar interdisciplinaire karakter een breed scala aan methodes kent die tot op heden nauwelijks zijn toegepast in de historische taalkunde. Deze paper is een bescheiden poging tot zo'n uitwisseling.

Taalverandering is bij uitstek een emergent fenomeen. In historische corpora kunnen we vaststellen hoe gedurende duizenden jaren een taal als het Nederlands in eenzelfde richting evolueert (een verschijnsel dat *drift* genoemd wordt, zie Sapir 1921; van der Horst 2013). Het ligt het echter minder voor de hand de achterliggende mechanismen van zulke taalveranderingen te testen aan de hand van empirische technieken zoals corpusstudies of experimenteel onderzoek met menselijke participanten. Corpusstudies zijn minder geschikt aangezien de effecten van de veronderstelde mechanismen in corpora moeilijk in afzondering kunnen worden bestudeerd (Landsbergen et al. 2010: 367-368). Voor experimenten met menselijke participanten ligt de moeilijkheid erin dat deze veranderingen zich over een erg lange termijn uitstrekken, en dus moeilijk terug te brengen zijn tot het gedrag van de individuele taalgebruikers. Zo stellen we bijvoorbeeld op lange termijn toenemende syncretismen vast in de naamvalsystemen van Indo-Europese talen (Baerman 2009; Hawkins 2004), maar het is lastig de geopperde verklaringen tegen elkaar af te wegen (zie daarvoor van Trijp 2013, 2014).

Het is bij dat laatste dat de mogelijkheden van agent-gebaseerde modellen liggen. Hun doel is de veronderstelde mechanismen te implementeren om te onderzoeken welke de verwachte resultaten opleveren, of op welke manier ze dat doen. Ook stellen ze de onderzoeker in staat de invloed na te gaan van verschillende parameters, startcondities, opstellingen van de agents enzovoort. Gegeven deze mogelijkheden, hebben agent-gebaseerde modellen echter ook enkele duidelijke beperkingen. Zo kan dit soort modellen hoogstens aantonen hoe iets gebeurd kan zijn, niet of het ook werkelijk zo gebeurd is. Agent-geba-

[2] Culturele evolutie beoogt het ontstaan en de evolutie van culturele fenomenen zoals taal te verklaren aan de hand van dezelfde basismechanismen als de biologische, Darwinistische evolutieleer. Een bespreking van culturele evolutie valt buiten het bestek van deze paper, maar zie Croft (2000), Steels (2011) en Mesoudi (2011).

seerde modellering zal dan ook nooit corpusonderzoek kunnen vervangen (Landsbergen 2009: 17). Daarnaast kan het model dat hier voorgesteld wordt, geen voorspellingen doen over de toekomst van onze taal, noch over het gedrag van al te specifieke taalelementen. De reden hiervoor is dat het model met opzet sterk simplificeert. Een al te complex model verstoort immers een scherp beeld op de effecten van elk geïmplementeerd mechanisme (Landsbergen 2009: 18-19; Landsbergen et al. 2010: 367-368).[3]

Computationele modellen van taalverandering impliceren een theorie over de drijvende kracht achter die verandering. Zo sluiten enerzijds modellen van *imperfect learning* (Niyogi & Berwick 1997; Yang 2000) aan bij kind-gebaseerde theorieën van taalverandering (Halle 1962; Andersen 1973).[4] Hierbij leren opeenvolgende generaties een taal, zodat taalverandering ontstaat doordat leerders de grammatica van hun voorgangers niet perfect overnemen. Anderzijds sluiten modellen als de *language games* (Loreto et al. 2011; Steels 2011; Wellens 2012) en het *Utterance Selection Model* (Baxter et al. 2006, 2009) nauw aan bij taalgebruiksgebaseerde theorieën van taalverandering (Croft 2000; Bybee 2010). In deze modellen gaan de agents ook na hun 'volwassenwording' door met het aanpassen van hun taalsysteem. De kiem van taalverandering ligt dus niet zozeer in opeenvolgende fouten in de verwerving van een taalsysteem, maar eerder in communicatie tussen 'volwassen' agents, d.i. in taalgebruik. Beide vormen van modellen hoeven elkaar echter niet uit te sluiten; binnen eenzelfde model kunnen zowel verwerving als gebruik een belangrijke rol spelen (cf. Pijpops, Beuls & Van de Velde 2015). Het model dat hier als voorbeeld zal dienen, is een voorbeeld van de tweede groep van computationele modellen. Daarom zullen we de opbouw van het model beschrijven aan de hand van termen uit de taalgebruiksgebaseerde taalkunde.

Het onderwerp van dit model is de competitie tussen de regelmatige en onregelmatige werkwoordinflectie van het Nederlands (van Santen 1997; Salverda 2006; Knooihuizen & Strik 2014).[5] Op macroperspectief zien we een

[3] In principe is het wel mogelijk het gedrag van specifieke taalelementen te simuleren in een agent-gebaseerd model, maar op dit moment ligt de interesse van modelleerders vaak meer bij algemene mechanismen dan specifieke gevallen. Dat kan later echter veranderen.

[4] Een belangrijk deel van deze modellen van *imperfect learning* sluit dan ook aan bij de theoretische achtergrond van de generatieve grammatica, maar zeker niet allemaal. Ook in bv. connectionistische modellen als Hare & Elman (1995) ontstaat taalverandering als gevolg van taalverwerving. Aangezien deze computationele modellen geen gebruik maken van een populatie van agents, maar van de opeenvolgende applicatie van een leeralgoritme, worden ze meestal niet agent-gebaseerd genoemd.

[5] Deze competitie komt grotendeels overeen met de tegenstelling tussen de zwakke en sterke vervoeging (Grimm 1819), zij het niet geheel. Getuige hiervan zijn het voorkomen van zwakke onregelmatige werkwoorden als *zoeken-zocht* en *denken-dacht*, en de sterke ablautklassen, die in het Nederlands zelfs nog een beperkte mate van productiviteit vertonen (Knooihuizen & Strik 2014). Toch is de zwakke vervoeging, een innovatie van het Proto-Germaans, doorheen de ontwikkeling van onze taal er veel beter in geslaagd haar transparantie te bewaren, terwijl de sterke vervoeging, een erfenis uit het aspectsysteem van het Proto-Indo-Europees (Ringe 2006), uiteengevallen is in subklassen en uitzonderingen. Voor een model dat de invloed van deze klassen wel in rekening brengt, zie Pijpops, Beuls & Van de Velde (2015).

toenemende regularisering van de Nederlandse werkwoordsinventaris (van Santen 1997). Een vast aantal werkwoorden slaagt er echter in blijvend weerstand te bieden aan deze regularisatiedruk en blijft verankerd (d.i. *entrenched* als in Bybee 2006) in hun oorspronkelijke, onregelmatige vorm (Cuskley et al. 2014). Met dit agent-gebaseerd model willen we onderzoeken hoe deze tegengestelde krachten, oorspronkelijke verankering en regularisatie, elkaar in evenwicht kunnen houden. We vermoeden dat de sleutel daartoe ligt in een *conserving effect* van frequentie (Bybee 2006: 715). In corpora stellen we inderdaad vast dat hoogfrequente werkwoorden er beter in slagen hun onregelmatige vorm te behouden (Lieberman et al. 2007; Carroll, Svare & Salmons 2012; Cuskley et al. 2014), maar het is moeilijker aan te tonen dat deze correlatie noodzakelijk ook een causaal verband inhoudt, en indien dat zo is, hoe dat verband exact in zijn werk gaat.

Deze paper is als volgt opgebouwd. Sectie 2 stelt de verschillende stappen in de opbouw van een agent-gebaseerd model voor. Vervolgens bespreekt Sectie 3 de resultaten van enkele mogelijke keuzes in het ontwerp van het voorbeeldmodel. Ten slotte bevat Sectie 4 enkele afsluitende bemerkingen.

2. Opbouw van een agent-gebaseerd model

2.1. Stap 1: Een te verklaren effect

Agent-gebaseerde modellering wordt vaak bekritiseerd omdat een strikte evaluatie zou uitblijven. Daarom is het aan te raden in deze eerste stap al duidelijk te definiëren wat je verwacht vast te stellen in de resultaten. De keuze van onderwerp valt dan ook best op een effect dat al grondig beschreven is aan de hand van corpusonderzoek, zodat je als onderzoeker een duidelijk beeld hebt van de resultaten die je model moet vertonen.

In het huidige model is het te verklaren effect de regularisatie van de laagfrequente werkwoorden, met tegelijkertijd een *conserving effect* (Bybee 2006) bij de hoogfrequente werkwoorden (van Santen 1997; Lieberman et al. 2007; Carroll, Svare & Salmons 2012). We verwachten co-existentie van beide varianten (d.z. de regelmatige en onregelmatige vervoeging) op het typeniveau, gecombineerd met fixatie op het tokenniveau. De onderscheidende factor dient daarbij de frequentie van het werkwoord te zijn. Dit wil zeggen dat we verwachten dat beide varianten langdurig naast elkaar kunnen bestaan, waarbij sommige werkwoorden onregelmatig blijven als gevolg van hun hoge frequentie, terwijl minder frequente werkwoorden overgaan naar de regelmatige vervoeging. Op het niveau van het individuele werkwoord dient echter wel in het merendeel van de gevallen gekozen te worden voor één van beide varianten (Cuskley et al. 2014: 3). Hier mogen de onregelmatige en regelma-

tige werkwoordsvorm niet naast elkaar blijven bestaan.[6] De onderzoeksvraag van het model luidt dan ook:

i. Welke aannames zijn nodig op conceptueel niveau (zie beneden) om regularisatie van de laagfrequente werkwoorden en een *conserving effect* bij de hoogfrequente werkwoorden te verklaren, en hoe dienen deze mechanismen concreet geïmplementeerd te worden?

2.2. Stap 2: Conceptueel niveau

In deze stap bevinden we ons op het niveau van de theorie. Hier wordt een hypothese geformuleerd die de onderzoeksvraag kan beantwoorden, waarbij het doel is het minimum aan aannames te maken dat strikt noodzakelijk is. Om bijvoorbeeld syncretismen in het Duitse naamvalsysteem te verklaren, neemt van Trijp enkel aan dat taalgebruikers hun taal pogen te optimaliseren (van Trijp 2013: 107-108, 2014: 340). Ook Landsbergen et al. benadrukken de noodzaak zo zuinig mogelijk te zijn op het aantal aannames in de hypothese, om op die manier het model zo eenvoudig mogelijk te houden (Landsbergen et al. 2010: 368). In het huidige model is de enige aanname dat bij transparant taalgebruik, d.i. het gebruik van de regelmatige vervoeging, naast de specifieke werkwoordsvorm ook een constructie wordt opgeslagen, die net als de werkwoordsvorm onderhevig is aan de effecten van frequentie.[7] Dit wil zeggen dat de taalgebruiker meer geneigd zal zijn te kiezen voor de vorm/constructie die hij het vaakst hoort. Voor het overige wordt het model zo eenvoudig mogelijk gehouden. We nemen voorlopig geen sociale structuur aan, noch een leerderscomponent. Er zullen geen nieuwe agents aan de simulatie worden toegevoegd. Mocht deze hypothese te weinig aannames maken, en het model dus te eenvoudig blijken, kunnen zulke uitbreidingen nog steeds gerealiseerd worden in Stap 5.

2.3. Stap 3: Implementatieniveau

In deze stap vindt het eigenlijke codeerwerk plaats. Het doel is de aannames op conceptueel niveau in een concreet model te gieten. Een belangrijk voordeel van computationele modellering is net dat het je als onderzoeker dwingt om de mechanismen van het conceptuele niveau concreet te maken (Blythe & Croft 2012: 270), en je zo voor keuzes stelt waar je anders niet bij stilstaat,

[6] In de realiteit geeft naast frequentie ook de fonetische vorm van een werkwoord vaak de doorslag voor regularisatie (Hare & Elman 1995: 64-70; Cuskley et al. 2014: 4-6). Voorlopig zijn we echter enkel geïnteresseerd in het effect van frequentie, en blijft ons model dus blind voor fonetische structuur.

[7] Zoals reeds aangegeven, gebruiken we termen uit de taalgebruiksgebaseerde taalkunde, zoals *constructie*, om het model te beschrijven. In principe is het model zelf echter theorieneutraal, en kan ook over *regels* gesproken worden. In deze paper hanteren we de definitie van constructies als in Goldberg (2006: 6).

maar die toch in hoge mate de resultaten kunnen beïnvloeden. Dit wordt duidelijk in Sectie 3, waarin drie zogenaamde aanpassingsstrategieën besproken worden. Alle drie komen ze overeen met wat in de vorige stap geformuleerd is, maar de exacte implementatie verschilt, zodat ook de resultaten verschillen.

In wat volgt, wordt de opbouw van het voorbeeldmodel voorgesteld. Het model dient de volgende bouwstenen te bevatten:

1. Een wereld waarin gebeurtenissen voorkomen waarover de agents moeten communiceren. De ene gebeurtenis moet vaker voorkomen dan de andere, zodat de agents bepaalde werkwoorden vaker zullen gebruiken dan andere.
2. Een competitie tussen regelmatige en onregelmatige werkwoordsvormen binnen het taalsysteem van elke agent, waarbij het gebruik van de regelmatige vervoeging naast een werkwoordspecifiek ook een constructioneel effect heeft.

Om de wereld samen te stellen, zijn alle werkwoorden geselecteerd die in de ANS genoemd worden in het overzicht *onregelmatige werkwoorden-a* (Haeseryn et al. 1997: 84-94). Hiervan zijn de 200 geselecteerd die in het Corpus Gesproken Nederlands (CGN) het frequentste bleken (van Eerten 2007).[8] Dit zijn de gebeurtenissen die in de wereld voorkomen. Op elk punt in de tijd zal één gebeurtenis voorkomen, waarbij de kans op elke gebeurtenis direct gecorreleerd is met de frequentie van het werkwoord in het CGN. Zo is er steeds meer dan 10% kans dat de meest frequente gebeurtenis, *gaan*, voorkomt (zie Tabel 1). De frequentie van deze gebeurtenissen vertoont ruwweg een Zipfiaanse verdeling (Zipf 1932).

Tabel 1: gebeurtenissen en hun kans voor te komen op elk punt in de tijd

Rang	Gebeurtenis	Kans
1	gaan	10,5%
	...	
20	nemen	1,13%
	...	
200	schijten	0,003%

Na een gebeurtenis wordt onder de agents willekeurig één spreker geselecteerd die de gebeurtenis moet uitdrukken met een werkwoordsvorm in de verleden tijd. De spreker zoekt de gebeurtenis op in zijn lexicon, en stoot daarbij telkens op een keuze uit een regelmatige en een onregelmatige vorm. In het

[8] We maken geen gebruik gemaakt van een historisch corpus, omdat het enkel belangrijk is dat de werkwoorden een realistische frequentieverdeling vertonen. Aangezien ons model blind is voor fonetische structuur (zie boven) is de keuze van werkwoorden verder niet van belang.

lexicon van elke agent wordt voor elk werkwoord ook een score tussen 0 en 1 bijgehouden die aangeeft in hoeverre dit werkwoord verankerd is in zijn onregelmatige vorm. Deze score komt overeen met de kans dat de agent voor de onregelmatige vorm zal kiezen. Werkwoorden met score 1 zullen door de agent dus steeds onregelmatig vervoegd worden, werkwoorden met score 0 steeds regelmatig. Aangezien de geselecteerde werkwoorden onregelmatig zijn, hebben alle werkwoorden bij de start een even sterke voorkeur voor hun onregelmatige vorm. Bij het begin van elke simulatie staan alle scores op 0,7.[9]

De hoorder interpreteert vervolgens de uiting van de spreker, en past de scores in zijn eigen lexicon aan aan wat hij hoort. Telkens hij een bepaalde werkwoordsvorm hoort, raakt deze vorm steviger verankerd in zijn lexicon. Dit wil zeggen dat wanneer de hoorder een onregelmatige werkwoordsvorm hoort, de score van dat werkwoord zal stijgen. Hoort hij een regelmatige werkwoordsvorm, dan daalt de score van dat werkwoord. Bij het horen van een onregelmatige werkwoordsvorm is dit werkwoordspecifieke effect het enige. Aangezien de regelmatige vervoeging zijn transparantie echter behouden heeft, vindt bij het horen van een regelmatige vorm ook een tweede, constructioneel effect plaats. Dit constructioneel effect stelt voor dat, wanneer een agent een regelmatige werkwoordsvorm hoort, de regelmatige vervoeging als constructie steviger verankerd raakt. Dat houdt in dat de kans op het gebruik van de regelmatige vervoeging als geheel stijgt, oftewel dat de score van alle werkwoorden lichtjes daalt. De precieze formules waarmee scores aangepast worden, zijn van groot belang voor ons model. In Sectie 3 zullen verschillende aanpassingsstrategieën worden getest, en hun resultaten vergeleken.

In de typologie van Baxter et al. (2009: 267-272) en Blythe & Croft (2012: 272-276) is dit model een vorm van *replicator selection* met *neutral interactor selection*. *Replicator selection* houdt in dat er kwalitatief verschil is tussen een van beide varianten, in dit geval de regelmatige en onregelmatige vervoeging. Het huidige model onderscheidt zich echter op twee vlakken van wat gewoonlijk *replicator selection* genoemd wordt. Ten eerste is het onderscheid er geen van sociale waardering oftewel *(c)overt prestige* (Labov 1966: 108; Trudgill 1972; Blythe & Croft 2012: 272), maar is het van taalstructurele aard. De onregelmatige waardering start steviger verankerd, terwijl de regelmatige vervoeging over een constructioneel effect beschikt. Ten tweede is een van beide varianten niet superieur aan de andere. Beide hebben een voordeel, namelijk stevigere initiële verankering versus constructioneel effect, dat elk in verschillende omstandigheden zwaarder kan doorwegen. Op dat vlak is dit model meer vergelijkbaar met het model van van Trijp, waarbij de beide varianten,

[9] Dit getal is slechts een parameterinstelling (zie Subsectie 3.3), die je kunt laten variëren. Cruciaal is dat de parameter boven 0,5 staat, om aan te geven dat de onregelmatige vervoeging initieel steviger verankerd is.

het Oudhoogduitse en Nieuwhoogduitse naamvalsysteem, ook elk op verschillende vlakken beter scoorden (van Trijp 2013: 111-128).

2.4. Stap 4: Evaluatie

In deze stap worden de resultaten van het model vergeleken met wat vastgesteld is in empirische studies. Meer bepaald wordt gekeken of de resultaten voldoen aan de eisen gesteld in Stap 1. Is dat niet het geval, dient onderzocht te worden of dat falen te wijten is aan problemen op implementatie- of conceptueel niveau. Problemen op implementatieniveau zijn vaak reeds te voorzien in Stap 3. Wanneer je als onderzoeker tot een bepaalde praktische keuze wordt gedwongen, hou je best de overige wegen open, zodat je de resultaten van de verschillende implementaties later kunt vergelijken. Problemen op conceptueel niveau zijn fundamenteler van aard. Ze houden in dat de hypothese in Stap 2 herbekeken moet worden, en dus elementaire onderdelen aan het agent-gebaseerd model aangepast of toegevoegd moeten worden.

De vraag of het falen van een agent-gebaseerd model ligt aan problemen op het implementatie- of conceptueel niveau is echter vaak moeilijk te beantwoorden. Meer nog, de moeilijkheid hiervan lijkt te wijzen op een fundamenteel methodologisch probleem. Indien een model immers faalt de verwachte resultaten op te leveren, kan steeds beweerd worden dat dit slechts te wijten is aan een foute implementatie, en niet aan het tekortschieten van de hypothese. Dit lijkt te betekenen dat deze modellen enkel kunnen bevestigen, en niet falsifiëren, wat een grote rem zou zijn op het nut van modellering voor de historische taalkunde.

Agent-gebaseerde modellen kunnen echter wel degelijk gebruikt worden voor falsificatie. Wat hiervoor nodig is, is de bevestiging van een model dat minder aannames maakt dan de te falsifiëren hypothese. Volgens het scheermes van Ockham dient de complexere hypothese vervolgens verworpen te worden:

> [I]f the weaker mechanism [...] is sufficient to account for the data, then it should be adopted over the stronger mechanism. (Blythe & Croft 2012: 277, met verwijzing naar Labov 2001: 192)

Op die manier weet van Trijp (2014: 340) aan te tonen dat het onnodig is een universele naamvalshiërarchie aan te nemen om syncretismen in het Duitse naamvalsysteem te verklaren, en laten Landsbergen et al. (2010: 385-386) zien dat verschillen in frequentie of verschillen in mutatiebereidheid elk apart reeds volstaan om unidirectionaliteit in betekenisverandering te verklaren.

2.5. Stap 5: Eventuele uitbreidingen

Eenmaal het model gebouwd en geëvalueerd is, kunnen uitbreidingen gerealiseerd worden. Deze houden in dat bovenstaande stappen ten minste deels opnieuw doorlopen dienen te worden. Mogelijke redenen om het model uit te breiden zijn o.a. (i) het model heeft gefaald, en er wordt geoordeeld dat meer aannames nodig zijn op conceptueel niveau om de te verwachte resultaten te produceren; (ii) een bijkomende vereiste wordt toegevoegd in Stap 1, waaraan het model in zijn huidige versie nog niet voldoet; of (iii) je bent eenvoudigweg geïnteresseerd hoe deze uitbreiding de resultaten van je model beïnvloedt.

3. RESULTATEN

Deze sectie toont de resultaten van het voorbeeldmodel. Om te laten zien hoe verschillende keuzes op implementatieniveau de resultaten van het model kunnen beïnvloeden, worden de effecten van drie verschillende aanpassingsstrategieën vergeleken. Deze aanpassingsstrategieën preciseren hoe de agents de scores in hun lexicon aanpassen. Elke strategie komt overeen met wat in de vorige sectie beschreven is.

3.1. Frequentie-gebaseerde aanpassing

Elke aanpassingsstrategie is fundamenteel gebaseerd op de frequentie waarmee de agent een werkwoordsvorm hoort. De frequentie-gebaseerde aanpassingsstrategie leidt de kans op een vorm echter zo direct mogelijk af van het aantal keer dat de agent de vorm gehoord heeft. De score van het werkwoord *w* laat zich in deze strategie als volgt berekenen:

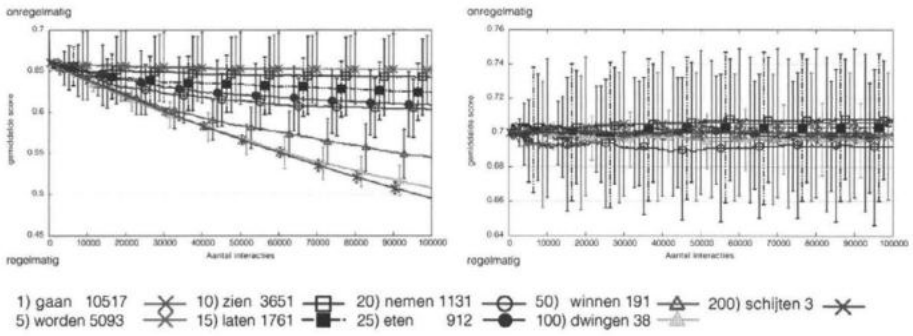
$$score_w = \frac{\#\,onreg\ vormen_w}{\#\,onreg\ vormen_w + \#\,reg\ vormen_w + \delta_{constr} * totaal\ \#\,reg\ vormen}$$

Figuur 1: aanpassingsformule voor de frequentie-gebaseerde strategie

Het eerste deel van deze verhouding laat zich makkelijk begrijpen: $\frac{\#\,onreg\ vormen_w}{\#\,onreg\ vormen_w + \#\,reg\ vormen_w}$. De score is gelijk aan het aantal keer dat de agent de onregelmatige vorm van het werkwoord gehoord heeft, gedeeld door het totaal aantal keer dat hij eender welke vorm van het werkwoord gehoord heeft. Naast dit werkwoordspecifieke effect, is er ook een constructioneel effect van de regelmatige vervoeging opgenomen, dat het totaal aantal regelmatige vormen dat de agent gehoord heeft, in de vergelijking binnenbrengt:

$\delta_{constr} * totaal\ \#\ reg\ vormen$.[10] Op die manier daalt de score van een werkwoord ook lichtjes wanneer de agent een regelmatige vorm van een ander werkwoord hoort, d.w.z. stijgt de kans op de regelmatige vorm van het werkwoord.

Grafiek 1 toont hoe werkwoorden van een verschillende frequentie zich gedragen onder deze aanpassingsstrategie.[11] Op de y-as zien we de gemiddelde score van het werkwoord. Op de x-as bevindt zich de tijdschaal, van 0 tot 100.000 interacties. We zien hoe de minst frequente werkwoorden langzaam toegeven aan de regularisatiedruk. De meest frequente werkwoorden lijken er vrij goed in te slagen hun beginstatus te behouden, maar raken niet steviger verankerd in hun onregelmatige vorm. We stellen dus wel degelijk regularisatie vast met een *conserving effect*, maar op tokenniveau doet zich geen fixatie voor. Er zijn veel werkwoorden die langdurig variatie blijven vertonen. Dit uitblijven van fixatie op tokenniveau is niet te wijten is aan een te sterk constructioneel effect, d.i. een te hoge instelling van de δ_{constr} parameter. Dat kan gezien worden in Grafiek 2, waar δ_{constr} op 0 ingesteld staat, en zich nog steeds geen fixatie voordoet.



Grafiek 1: gemiddelde scores van enkele werkwoorden onder de frequentie-gebaseerde strategie. $\delta_{constr} = 0{,}001$[12]

Grafiek 2: gemiddelde scores van enkele werkwoorden onder de frequentie-gebaseerde strategie. $\delta_{constr} = 0$

[10] Aangezien het totaal aantal keer dat de agent een regelmatige werkwoordsvorm gehoord heeft (*totaal # reg vormen*) typisch veel groter is dat het aantal keer dat hij een vorm van het werkwoord in kwestie gehoord heeft ($\#\ onreg\ vormen_w$ en $\#\ reg\ vormen_w$) dient δ_{constr} hier laag ingesteld te worden. Zo niet is het werkwoordspecifieke effect te verwaarlozen t.o.v. het constructionele effect.

[11] Om het overzicht te bewaren in deze grafiek, kan de evolutie van alle 200 werkwoorden niet getoond worden. Geen van hen vertoont echter een radicaal verschillend gedrag. Elk van de grafieken in deze paper tonen de resultaten van 40 simulaties van elk 100.000 interacties, met als parameterinstellingen: aantal agents = 10, initiële scores = 0,7. De overige parameterinstellingen worden bij de grafiek zelf vermeld.

[12] Bij elk werkwoord is zijn rang en gemiddelde frequentie tijdens 100.000 interacties vermeld.

3.2. Delta-multiplicatie aanpassing

Bij de delta-multiplicatie aanpassingsstrategie worden de scores geüpdatet volgens de vergelijkingen in Tabel 2.

Tabel 2: aanpassingsformules voor de delta-multiplicatie strategie

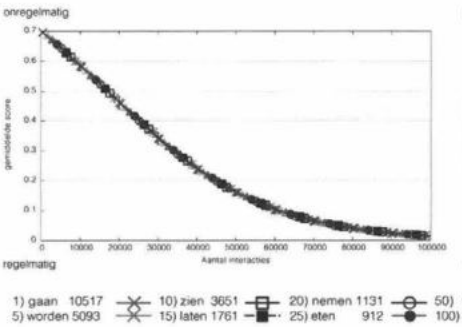
	Werkwoordspecifiek effect	**Constructioneel effect**
Onregelmatige vorm	$score_w = score_w * (1 - \delta_{spec}) + \delta_{spec}$	Afwezig
Regelmatige vorm	$score_w = score_w * (1 - \delta_{constr})$	$score_w = score_w * (1 - \delta_{constr})$

De redenering achter deze formules is de volgende. Ten eerste dienen dienen de werkwoordspecifieke effecten elkaars spiegelbeeld te zijn. In Sectie 2.2 hebben we immers als enige aanname van ons model gedefinieerd dat bij transparant taalgebruik ook een constructie wordt opgeslagen. Dit verschil tussen tussen de onregelmatige en regelmatige vervoeging is al in acht genomen via het constructionele effect, en verdere verschillen zijn dus niet toegelaten. Ten tweede willen we met deze aanpassingsstrategie testen wat het effect is, wanneer het horen van een nieuwe, onverwachte vorm een groter effect heeft dan het horen van het verwachte alternatief, dat men al duizendmaal gehoord heeft.

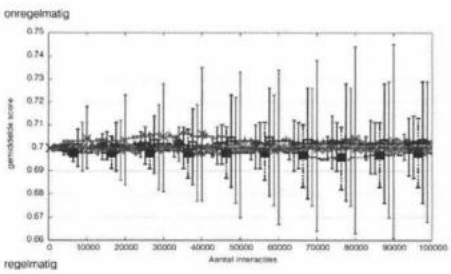
Beide punten kunnen duidelijker worden gemaakt met een voorbeeld, onder de parameterinstellingen van de Grafieken 3-4. Stel dat de score van het werkwoord *komen* bij een agent op 0,8 staat, d.w.z. dat er 80% kans is dat hij zelf de onregelmatige vorm *kwam* gebruikt. *Kwam* is dus de verwachte vorm. Als deze agent *kwam* hoort, gaat de score omhoog met 0,002. Hoort hij onverwachts *komde*, dan gaat de score omlaag met 0,008. Stel daarentegen dat de score op 0,2 staat, d.w.z. dat er 80% kans is dat de agent zelf de regelmatige vorm *komde* gebruikt, dan zal de score bij het horen van de verwachte, regelmatige vorm *komde* dalen met 0,002. Bij het horen van de onverwachte, onregelmatige vorm *kwam* zal ze echter omhooggaan met 0,008.

De resultaten onder deze aanpassingsstrategie zijn te zien in Grafiek 3. Alle werkwoorden geven even snel toe aan de regularisatiedruk, ongeacht hoe frequent ze zijn. Hier doet zich dus wel degelijk regularisatie voor, maar dan zonder *conserving effect*. Op het tokenniveau zien we wel fixatie, maar op het typeniveau geen co-existentie: alle werkwoorden worden regelmatig. Deze resultaten zijn opnieuw niet te wijten aan een verkeerde keuze van parameterinstellingen. In Grafiek 4 wordt het constructionele effect bijvoorbeeld volledig uitgeschakeld, en ook hier ontstaat geen verschil tussen werkwoorden van verschillende frequenties. Het falen van de frequentie-gebaseerde en delta-

multiplicatie strategieën ligt vermoedelijk eerder in de aard van de aanpassingsformules. Zoals reeds opgemerkt voor de delta-multiplicatie strategie, wordt de score bij beide strategieën sterker aangepast wanneer de agent een onverwachte vorm van het werkwoord hoort dan wanneer hij de verwachte vorm hoort. Dit leidt ertoe dat, wanneer we het constructionele effect uitschakelen, we op tokenniveau een co-existentie van beide varianten vinden (cf. Grafiek 2 en 4, Blythe & Croft 2012: 282).



Grafiek 3: gemiddelde scores van enkele werkwoorden onder de delta-multiplicatie strategie.
$\delta_{constr} = 0{,}0005$, $\delta_{spec} = 0{,}01$



Grafiek 4: gemiddelde scores van enkele werkwoorden onder de delta-multiplicatie strategie.
$\delta_{constr} = 0$, $\delta_{spec} = 0{,}01$

3.3. Delta-accumulatie aanpassing

In tegenstelling tot de delta-multiplicatie strategie, gaan de scores in de delta-accumulatie aanpassingsstrategie steeds met een vaste waarde omhoog of omlaag, volgens de formules in tabel 3.[13] De onverwachte vorm wordt dus niet langer bevoordeeld t.o.v. de verwachte vorm.

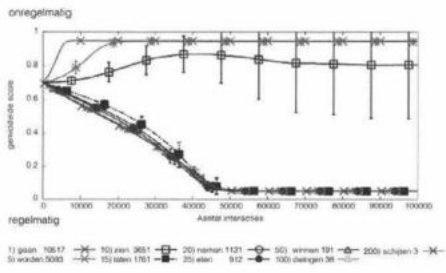
Tabel 3: aanpassingsformules voor de delta-accumulatie strategie

	Werkwoordspecifiek effect	**Constructioneel effect**
Onregelmatige vorm	$score_w = score_w + \delta_{spec}$	Afwezig
Regelmatige vorm	$score_w = score_w - \delta_{spec}$	$score_w = score_w - \delta_{constr}$

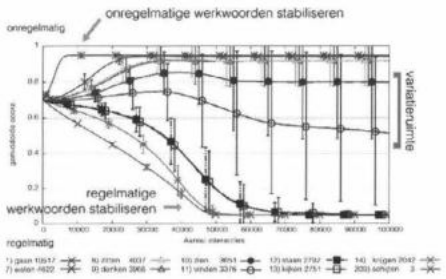
Bij deze strategie stellen we een regularisatie vast met *conserving effect*. De laagfrequente werkwoorden worden regelmatig, terwijl de hoogfrequente werkwoorden steviger verankerd raken in hun onregelmatige vorm (zie Grafiek 5). Op typeniveau vinden we dus co-existentie: de regelmatige en onregelmatige vervoeging blijven naast elkaar bestaan met elk een eigen habitat in het lexicon,

[13] Aangezien de scores probabiliteiten voorstellen en dus tussen 0 en 1 dienen te blijven, zijn het minimum en maximum om respectievelijk 0,05 en 0,95 gezet.

nl. de laag- en hoogfrequente werkwoorden. Op tokenniveau vinden we bij de meeste werkwoorden fixatie: ze worden of onregelmatig, of regelmatig. Slechts enkele werkwoorden slagen erin variatie te blijven vertonen tussen hun regelmatige en onregelmatige vorm. Grafiek 6 is identiek aan Grafiek 5, met als enige verschil dat andere werkwoorden worden getoond. Buiten het meest en minst frequente werkwoord, zien we in deze grafiek ook een aantal werkwoorden rond de variatieruimte tussen regelmatig en onregelmatig. Bij veruit de meeste werkwoorden ontstaat fixatie, maar variatie is niet geheel onmogelijk, wat we ook in werkelijk taalgebruik vaststellen (Cuskley et al. 2014).



Grafiek 5: gemiddelde scores van enkele werkwoorden onder de delta-accumulatie strategie.
δ_{constr} = 0,0005, δ_{spec} = 0,01



Grafiek 6: emiddelde scores van enkele werkwoorden onder de delta-accumulatie strategie.
δ_{constr} = 0,0005, δ_{spec} = 0,01

Opvallend is dat de individuele werkwoorden bij hun overgang naar een van beide vervoegingen S-curves lijken te vertonen (Grafiek 5 en 6). Dit soort curves is al erg vaak vastgesteld in taalveranderingen (Bailey 1973; Labov 1994: 64; Chambers & Trudgill 1998: 162-164). In Stap 2 hebben we dit niet als een vereiste voor ons model opgenomen, aangezien we niet op de hoogte zijn van corpusstudies die aantonen dat bij de regularisatie van individuele werkwoorden zich inderdaad ook zulke S-curves voordoen. Gezien hun veelvuldig opduiken bij andere taalveranderingen, lijkt dit ons echter zeker niet onmogelijk.

Merk op dat het oninteressant is te kijken hoeveel werkwoorden of welke werkwoorden precies regelmatig dan wel onregelmatig worden. Dat kun je makkelijk laten variëren door de parameters aan te passen. Wat wel interessant is, is te kijken waar de grens van de vrijheid ligt die de parameters je verschaffen. Bij de delta-multiplicatie strategie kun je bijvoorbeeld de parameters naar believen aanpassen, maar zal je nooit een scheiding krijgen tussen de laag- en hoogfrequente werkwoorden. Omgekeerd kun je bij delta-accumulatie weliswaar aan de hand van de δ_{spec} en δ_{constr} parameters of de initiële scores ervoor zorgen dat de overgang van onregelmatig naar regelmatig bij een lagere of hogere frequentie plaatsvindt, maar het is onmogelijk geen *conserving*

effect van frequentie te genereren. Frequentere werkwoorden zullen dus steeds onregelmatig blijven, terwijl infrequentere werkwoorden regelmatig worden. Agent-gebaseerde modellen dienen dan ook enkel om te testen wat wel en niet mogelijk is binnen bepaalde aannames. Als je daarentegen een voorspelling wil doen of een specifiek werkwoord al dan niet regelmatig zal worden in het Nederlands, kun je beter gebruik maken van statistische modellen als Lieberman et al. (2007), Carroll, Svare & Salmons (2012) of Cuskley et al. (2014).

Hieronder worden de effecten van alle parameterinstellingen kort opgesomd voor de delta-accumulatie strategie. Hun effecten zijn mutadis mutandis vergelijkbaar voor de andere strategieën.

- Aantal agents: Bij een hoger aantal agents vertraagt de fixatie op tokenniveau, zoals ook vastgesteld door Blythe & Croft (2012: 293).[14]
- Initiële scores: De hoge initiële scores stellen het voordeel van initiële verankering van de sterke vervoeging voor. Des te hoger deze scores, des te meer werkwoorden onregelmatig blijven.
- $\delta_{constr}/\delta_{spec}$: Met de δ_{constr} parameter kan de grootte van het constructioneel effect ingesteld worden, terwijl de δ_{spec} parameter de grootte van het werkwoordspecifieke effect bepaalt. Des te groter de $\delta_{constr}/\delta_{spec}$ ratio, des te groter het voordeel van de transparante, regelmatige vervoeging, en des te meer werkwoorden regelmatig worden.

De laatste implementatie van ons model is dus succesvol gebleken. De aanname dat de taalgebruiker bij transparant taalgebruik naast de specifieke werkwoordsvorm ook een constructie opslaat, waarbij beide representaties onderhevig zijn aan frequentie-effecten, kan volstaan om regularisatie met een *conserving effect* te veroorzaken, waarbij op typeniveau co-existentie optreedt, en op tokenniveau fixatie. Uiteraard kunnen in dit model meer dingen gemeten worden dan de individuele ontwikkeling van de werkwoorden. Er kan bv. ook nagegaan worden hoeverre de agents overeenkomen welke werkwoorden onregelmatig dan wel regelmatig horen te zijn, wat de tokenfrequentieverhouding tussen beide vervoeging is, etc. Om dit artikel bondig te houden, worden deze maatstaven niet besproken, maar zie Pijpops & Beuls (2015). Vervolgens kan verdergegaan worden met Stap 5, door verdere uitbreidingen aan het

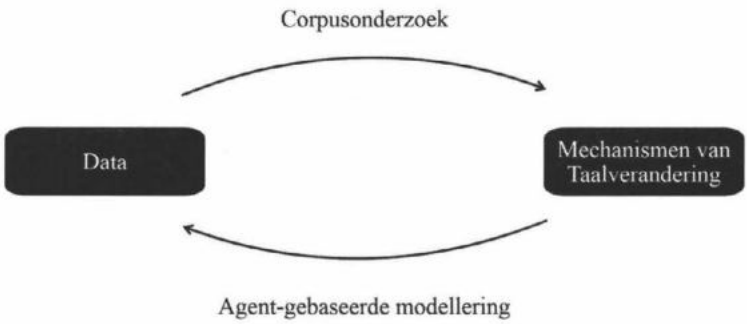
[14] Voor een uitgebreide bespreking van het effect van o.a. de grootte van de agent-populatie op de snelheid van taalverandering, zie Nettle (1999), maar ook Ke (2004: 162-164)structured population which was presented in an earlier paper (\u2018Using Social Impact Theory to simulate language change\u2019, Lingua 108, 95\u2013117, 1999 voor enkele kritische bemerkingen.

model toe te voegen. Zo'n uitbreiding is al gerealiseerd, namelijk door een sociale structuur aan de agent-populatie toe te kennen en de effecten daarvan na te gaan. Om dezelfde reden verwijzen we hiervoor echter ook naar Pijpops & Beuls (te versch.).

4. CONCLUSIES

Het model dat hier is voorgesteld, is met opzet sterk simplificerend (Landsbergen 2009: 18-19). Participia worden volledig genegeerd, net als het bestaan van de nog steeds min of meer regelmatige ablautklassen van de sterke vervoeging, en de invloed van de fonetische structuur van het werkwoord. Deze invloed is echter zonder twijfel van belang (Hare & Elman 1995: 64-70; Carroll, Svare & Salmons 2012; Knooihuizen & Strik 2014). Dit model kan dan ook beter gezien worden als een algemeen model van de competitie tussen regelmatig en onregelmatig taalgebruik, eerder dan een specifiek model van de Nederlandse zwakke en sterke werkwoorden. Voor meer geavanceerde agent-gebaseerde modellen van ontwikkelingen in het Nederlands, zie o.a. Landsbergen (2009), Bloem, Versloot & Weerman (2015) en Pijpops & Beuls (2015).

Ten slotte willen we twee bemerkingen opnieuw benadrukken (zie daarvoor ook de Boer 2012). Ten eerste zijn corpusonderzoek en agent-gebaseerde modellering complementair, zoals voorgesteld in Figuur 2. Enerzijds dient de bouw van een agent-gebaseerd model steeds voorafgegaan te worden door empirisch onderzoek. Zonder zo'n onderzoek is immers geen evaluatie van het model mogelijk. Anderzijds stelt agent-gebaseerde modellering de onderzoeker in staat de veronderstelde mechanismen elk apart te implementeren, en zo na te gaan wat hun effect in afzondering is. Vooral dit laatste is moeilijk in corpusonderzoek (Landsbergen et al. 2010: 367).



Figuur 2: corpusonderzoek en agent-gebaseerde modellering als complementaire technieken

Ten tweede is een veelgehoorde kritiek op agent-gebaseerde modellen dat een strikte evaluatie onmogelijk is, aangezien alle resultaten mogelijk zijn door de parameters te laten variëren. Om die reden is een evaluatie van het huidige model aan de hand van de vraag precies hoeveel en welke werkwoorden regelmatig dan wel onregelmatig geworden zijn, inderdaad nutteloos. In Sectie 3 hebben we echter aangetoond dat (i) niet zomaar alle resultaten mogelijk zijn binnen eenzelfde model, en (ii) een evaluatie op basis waarvan sommige modellen worden verworpen en andere bevestigd, wel degelijk uitgevoerd kan worden. Bij de interpretatie van een agent-gebaseerd model dien je steeds te kijken naar wat wel of niet mogelijk is binnen de vrijheid die je parameters je geven (Hare & Elman 1995: 91-92). Een model waarin alles mogelijk is, is net zo oninteressant als een model waarin niets mogelijk is.

BIBLIOGRAFIE

ANDERSEN, HENNING. 1973. Abductive and Deductive Change. *Language* 49(4). 765-793.

BAERMAN, MATTHEW. 2009. Case Syncretism. *The Oxford handbook of case*, 219-230. Oxford: Oxford University Press.

BAILEY, CHARLES-JAMES. 1973. *Variation and linguistic theory.* Washington D.C.: Center for Applied Linguistics.

BAXTER, GARETH, RICHARD BLYTHE, WILLIAM CROFT & ALAN MCKANE. 2006. Utterance selection model of language change. *Physical Review E* 73(4). 046118.

BAXTER, GARETH, RICHARD BLYTHE, WILLIAM CROFT & ALAN MCKANE. 2009. Modeling language change: An evaluation of Trudgill's theory of the emergence of New Zealand English. *Language Variation and Change* 21(2). 257-296.

BEULS, KATRIEN & LUC STEELS. 2013. Agent-Based Models of Strategies for the Emergence and Evolution of Grammatical Agreement. *PLoS ONE* 8(3). e58960.

BHATTACHARJEE, YUDHIJIT. 2004. From heofonum to heavens. *Science* 303(5662). 1326-1328.

BLOEM, JELKE, ARJEN VERSLOOT & FRED WEERMAN. 2015. An agent-based model of Germanic verbal cluster word order change. *Computational Linguistics in the Netherlands (CLIN)*. 8 februari, Antwerpen.

BLYTHE, RICHARD & WILLIAM CROFT. 2012. S-curves and the mechanisms of propagation in language change. *Language* 88(2). 269-304.

BONABEAU, ERIC, MARCO DORIGO & GUY THERAULAZ. 1999. *Swarm intelligence: From natural to artificial systems*. New York: Oxford University Press.

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.

CANGELOSI, ANGELO & DOMENICO PARISI (eds.). 2002. *Simulating the evolution of language*. Londen: Springer.

CARROLL, RYAN, RAGNAR SVARE & JOSEPH SALMONS. 2012. Quantifying the evolutionary dynamics of German verbs. *Journal of Historical Linguistics* 2(2). 153-172.

CHAMBERS, JACK & PETER TRUDGILL. 1998. *Dialectology*. 2e editie. Cambridge: Cambridge University Press.

CROFT, WILLIAM. 2000. *Explaining language change: An evolutionary approach*. Essex: Pearson Education Limited.

CUSKLEY, CHRISTINE, MARTINA PUGLIESE, CLAUDIO CASTELLANO, FRASCESCA COLAIORI, VITTORIO LORETO & FRANSCESCA TRIA. 2014. Internal and External Dynamics in Language: Evidence from Verb Regularity in a Historical Corpus of English. *Plos One* 9(8). e102882.

DE BOER, BART. 2012. Modelling and Language Evolution: beyond fact-free science. In Luke McCrohon, Tomomi Fujimura, Kazuo Okanoya, Koji Fujita, Reiji Suzuki, Roger Martin & Noriaki Yusa (eds.), *The Evolution of Language: Proceedings of the 9th International Conference*, 83-92. Evolang-9 Organizing Committee.

DHAMDHER, AMOGH & CONSTANTINE DOVROLIS. 2009. An agent-based model for the evolution of the internet ecosystem. *COMSNETS. 1st International Conference on Communication Systems and Networks*, 1-10.

GOLDBERG, ADELE EVA. 2006. *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.

GRIMM, JACOB. 1819. *Deutsche Grammatik*. Göttingen: Dieterichsche Buchhandlung.

HAESERYN, WALTER, KIRSTEN ROMIJN, GUIDO GEERTS, JAAP DE ROOIJ & MAARTEN VAN DEN TOORN. 1997. *Algemene Nederlandse Spraakkunst*. Groningen: Nijhoff.

HALLE, MORRIS. 1962. *Phonology in Generative Grammar*. Chicago: Word.

HARE, MARY & JEFFREY ELMAN. 1995. Learning and morphological change. *Cognition* 56(1). 61-98.

HAWKINS, JOHN. 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

KE, JINYUN. 2004. Self-organization and Language Evolution: System, Population and Individual. Dissertation City University of Hong Kong.

KNOOIHUIZEN, REMCO & OSCAR STRIK. 2014. Relative productivity potential of Dutch verbal inflection patterns. *Folia linguistica historica* 35(1). 173-200.

LABOV, WILLIAM. 1966. Hypercorrection in the lower middle class as a factor in linguistic change. In William Bright (ed.), *Sociolinguistics*, 84-113. Den Haag: Mouton.

LABOV, WILLIAM. 1994. *Principles of linguistic change, vol. 1: Internal factors*. Oxford: Blackwell.

LABOV, WILLIAM. 2001. *Principles of linguistic change, vol. 2: Social factors*. Oxford: Blackwell.

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

LANDSBERGEN, FRANK, ROBERT LACHLAN, CAREL TEN CATE & ARIE VERHAGEN. 2010. A cultural evolutionary model of patterns in semantic change. *Linguistics* 48(2). 363.

LIEBERMAN, EREZ, JEAN-BAPTSTE MICHEL, JOE JACKSON, TINA TANG & MARTIN NOWAK. 2007. Quantifying the evolutionary dynamics of language. *Nature* 449(7163). 713-716.

LORETO, VITTORIO, ANDREA BARONCHELLI, ANIMESH MUKHERJEE, ANDREA PUGLISI & FRANCESCA TRIA. 2011. Statistical physics of language dynamics. *Journal Of Statistical Mechanics-Theory And Experiment*. 04006.

MESOUDI, ALEX. 2011. *Cultural Evolution: How Darwinian Theory Can Explain Human Culture and Synthesize the Social Sciences*. Chicago: University Of Chicago Press.

MORTEN, CHRISTIANSEN & SIMON KIRBY (eds.). 2003. *Language Evolution*. Oxford: Oxford University Press.

NETTLE, DANIEL. 1999. Is the rate of linguistic change constant? *Lingua* 108(2). 119-136.

NIYOGI, PARTHA & ROBERT BERWICK. 1997. A dynamical systems model for language change. *Complex Systems* 11. 161-204.

PIJPOPS, DIRK, KATRIEN BEULS & FREEK VAN DE VELDE. 2015. Fighting in broken formation. The competition between the Dutch strong ablaut classes and weak suffix inflection in an agent-based model. *Societas Linguistica Europaea (SLE) conferentie*, 48. 3 september, Leiden.

PIJPOPS, DIRK & KATRIEN BEULS. 2015. Generating the Conserving Effect without Language Acquisition. An agent-based model. *Posterpresentatie op de Thematische Conferentie van de Belgische Kring voor Linguïstiek (BKL-CBL)*. 8 juni 2015, Brussel.

RINGE, DON. 2006. *From Proto-Indo-European to Proto-Germanic*. Oxford: Oxford University Press.

SALVERDA, REINIER. 2006. Over de sterke werkwoorden in het Nederlands, Engels en Duits. In Matthias Hüning, Ulrike Vogl, Ton van der Wouden & Arie Verhagen (eds.), *Nederlands tussen Duits en Engels*, 163-181. Leiden: Stichting Neerlandistiek Leiden.

SAPIR, EDWARD. 1921. *Language. An introduction to the study of speech*. New York: Harcourt.

STEELS, LUC. 1998. Synthesising the origins of language and meaning using coevolution, self-organisation and level formation. In James Hurford, Michael Studdert-Kennedy & Chris Knight (eds.), *Approaches to the Evolution of Language*, 384-404. Cambridge: Cambridge University Press.

STEELS, LUC. 2011. Modeling the cultural evolution of language. *Physics of Life Reviews* 8(4). 339-356.

STEELS, LUC & EMILIA GARCIA-CASADEMONT. 2015. Ambiguity and the origins of syntax. *Linguistic Review* 32(1). 37-60.

VAN EERTEN, LAURA. 2007. Over het Corpus Gesproken Nederlands. *Nederlandse Taalkunde* 12(3). 194-215.

VAN DER HORST, JOOP. 2013. *Taal op drift: lange-termijnontwikkelingen in taal en samenleving*. Amsterdam: Meulenhoff.

VAN SANTEN, ARIANE. 1997. Hoe sterk zijn de sterke werkwoorden. In Ariane van Santen & Marijke van der Wal (eds.), *Taal in tijd en ruimte [Feestbundel Cor van Bree]*, 45-56. Leiden.

VAN TRIJP, REMI. 2013. Linguistic Assessment Criteria for Explaining Language Change: A Case Study on Syncretism in German Definite Articles. *Language Dynamics and Change* 3. 105-132.

VAN TRIJP, REMI. 2014. Fitness landscapes in cultural language evolution: a case study on german definite articles. In Erica Cartmill, Seán Roberts, Headi Lyn & Hannah Cornish (eds.), *The Evolution of Language: Proceedings of the 10th International Conference*, 334-341. World Scientific.

TRUDGILL, PETER. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society* 1(2). 179-195.

WELLENS, PIETER. 2012. Adaptive Strategies in the Emergence of Lexical Systems. Dissertation Vrije Universiteit Brussel.

YANG, CHARLES. 2000. Internal and external factors in language change. *Language Variation and Change* 12(3). 231-250.

ZIPF, GEORGE KINGSLEY. 1932. *Selected Studies of the Principle of Relative Frequency in Language*. Harvard: Harvard University Press.