

Een classificatie van simulatiegebaseerd onderzoek naar taalverandering

door Anthe SEVENANTS en Dirk SPEELMAN

Abstract

In this article we propose a classification for agent-based computer simulations of language change. The classification hinges on four major categories which describe a trigger on the local level and a global, emergent effect. These categories, which coincide with linguistic subdisciplines, are assessed on the basis of four dimensions: (1) whether meaning plays a role; (2) what agent types (M. Abrams 2013) are used; (3) whether agents are connected in a social structure; (4) whether the focus lies on new forms or the disappearance of old ones. The differences between the categories in the proposed classification reflect their different goals, but because of the versatile nature of simulation research, classification remains a difficult task.

INLEIDING

In de taalkunde wordt het veld van computersimulaties steeds belangrijker. Taalkundigen beseffen namelijk steeds meer dat computersimulaties een interessant theoretisch laboratorium zijn waarin onderzoeksvragen niet geremd worden door de beperkingen van traditioneel empirisch onderzoek. Met computersimulaties kan een onderzoeker duizenden virtuele sprekers — ook wel *agents* genoemd — programmeren en ze met elkaar laten communiceren. Deze ‘ideale wereld’ biedt heel wat voordelen voor een onderzoeker. Allereerst kan de onderzoeker minutieus volgen hoe de agents spreken, met wie ze dat doen, en op welk moment. Daarbovenop zijn computersimulaties ook niet beperkt in tijd en ruimte (Gong en Shuai 2013, 14); een onderzoeker kan met simulaties perfect een hypothese testen die we duizenden jaren terug in de tijd moeten situeren, of die een periode van duizenden jaren bestrijkt (Fagyal e.a. 2010, 2063). Gong en Shuai (2013, 14) benadrukken wel dat computersimulaties geen teletijdmachines zijn: ze kunnen niet *letterlijk* terug in de tijd, maar ze bieden wel de mogelijkheid om een omgeving op te zetten die de (veronderstelde) omstandigheden van vroeger nabootst. Ook dwingen simulaties af dat de geteste hypotheses erg concreet zijn (A. D. M. Smith 2014, 282). Computers vereisen erg precieze instructies, waardoor onderzoekers hun hypotheses zo expliciet mogelijk moeten definiëren. Wanneer een hypothese niet expliciet genoeg geoperationaliseerd kan worden, kan dit wijzen op lacunes in een theorie (Valenzuela, Hilferty, en Vilarroya 2016, 267). Ten slotte is het voor computersimulaties ook noodzakelijk dat de gesimuleerde werkelijkheid

vereenvoudigd wordt (A. D. M. Smith 2014, 282), waardoor de zogenaamde “noise”, die we associëren met empirisch onderzoek uitgesloten wordt. Hierdoor wordt het gemakkelijker om te focussen op het probleem in kwestie.

De echte kracht van een simulationeel model zit hem echter in de mogelijkheid om taal te modelleren als een complex-adaptief systeem (Kirby 1999, viii; A. D. M. Smith 2014, 284). Dat houdt kortweg in dat, prescriptivistische academies ten spijt, taal geen centraal georganiseerd fenomeen is, maar het resultaat van lokaal taalgebruik tussen sprekers dat uiteindelijk voor verandering zorgt in het taalsysteem. Computersimulaties zijn ideaal voor het definiëren van dergelijk ‘lokaal gedrag’. Het is vervolgens aan de onderzoeker om het emergente effect van dit lokale gedrag in het model (de veranderingen in het taalsysteem) te observeren en te verklaren. In empirisch onderzoek is het veel moeilijker om een emergent systeem te beschrijven. In de sociolinguïstiek, bijvoorbeeld, beschikken we wel over grote dataverzamelingen, zoals het Corpus Gesproken Nederlands (Van Eerten 2007), maar die bieden slechts een fragmentarisch beeld van het complex-adaptief systeem waarin we geïnteresseerd zijn. Een dergelijk complex fenomeen onderzoeken zou een panelstudie van immense schaal vereisen, wat niet alleen organisatorisch onmogelijk is, maar ook onbetaalbaar. Het is onbegonnen werk voor een onderzoeker, of zelfs voor een groep van onderzoekers, om zo veel mensen intensief te volgen en hun taalgebruik te registreren (Beuls en van Trijp 2016, 3), laat staan om het lokale gedrag te verbinden aan emergente effecten. In die zin laten computersimulaties dus toe om de plausibiliteit van bepaalde theorieën te onderzoeken die onmogelijk te controleren zijn met empirisch onderzoek.

Tegelijkertijd is het belangrijk om te benadrukken dat computersimulaties daarom niet het ultieme antwoord zijn voor elk onderzoek. Zo is het bijvoorbeeld erg moeilijk en zelfs onwenselijk om verschillende fenomenen of invloeden in één simulatiemodel te onderzoeken. Als er te veel verschillende invloeden zijn in een model, kunnen we niet één proces als oorsprong aanduiden voor het gedrag van dat model, wat de voordelen van de gecontroleerde simulatieomgeving tenietdoet (A. D. M. Smith 2014, 282). Een traditioneel corpusonderzoek met regressieanalyse kan veel beter omgaan met verschillende invloeden. Daarnaast kunnen simulatiemodellen “hoogstens aantonen hoe iets gebeurd kan zijn, niet of het ook werkelijk zo gebeurd is” (Pijpops en Beuls 2015, 6). Simulatieonderzoek probeert namelijk niet om definitieve antwoorden te vinden op onderzoeksvragen; simulaties zijn geen ‘alwetende goden’, maar louter exploratieve middelen om te onderzoeken welke processen verantwoordelijk *kunnen* zijn voor een bepaald fenomeen in de echte wereld (Pijpops en Beuls 2015, 6). Empirisch onderzoek probeert bloot te leggen welke tendensen spelen in de echte wereld, maar het ontstaan van die tendensen is vaak moeilijk of niet te ontwarren. Simulatieonderzoek,

daarentegen, laat wél zien hoe tendensen kunnen ontstaan, maar is niet in staat om te bewijzen of die tendensen in de échte wereld daadwerkelijk zo ontstaan zijn. De gedachtegang is dat de uitkomst van een simulatie *aanwijzingen* kan geven over het door de simulatie gemodelleerde proces in de *echte* wereld (Stanford en Kenny 2013, 122) wanneer de simulatie bepaald gedrag wel of niet vertoont.

Wat ons vooral interesseert in een simulatiemodel zijn de ‘minimale voorwaarden’ (Baronchelli 2014). Dat wil zeggen dat we op zoek gaan naar welke veronderstellingen ingebouwd moeten worden in een model opdat dat model bepaald emergent gedrag vertoont. Om die minimale voorwaarden te vinden, moeten we een breed scala aan theoretisch mogelijke veronderstellingen uitproberen, waarvan het exacte bereik niet per se rechtstreeks uit een externe bron hoeft voort te komen. Het hele doel van een simulatiestudie is om uit te zoeken welke omstandigheden leiden tot een bepaald fenomeen, maar ook welke omstandigheden dat niet doen. Deze interpretatie van een model verschilt sterk van de traditionele manier van modellen maken. Het ontwerp van een Hidden Markov Model, bijvoorbeeld, begint altijd met het bepalen van kansen vanuit empirische data. Een lineair regressiemodel beschrijft dan weer een trend op basis van observatiedata. Een simulatiemodel draait dit proces om: in plaats van een model te maken vanuit data, proberen we een breed bereik van mogelijke modellen uit om te bepalen welke modellen leiden tot het fenomeen dat we onderzoeken. Het is dit bereik (en de voorwaarden aan dit bereik verbonden zijn) dat centraal staat in simulatieonderzoek. De ‘plausibiliteit’ van een hypothese duidt binnen simulatieonderzoek daarmee eerder op haar ‘mogelijkheid’ of ‘werkbaarheid’: zie het als het succes van een *proof of concept*.

Hoewel het veld van simulatieonderzoek nog relatief nieuw is, is er toch al heel wat onderzoek verricht met deze methodologie. Wat echter nog ontbreekt binnen dit onderzoeksparadigma is een structuur om het bestaande onderzoek duidelijk af te bakenen. Een dergelijke structuur kan nuttig zijn voor onderzoekers die geïnteresseerd zijn in simulatiemethodologie, maar niet weten waar hun eigen onderzoek in zou passen.

We stellen daarom een classificatie voor met vier categorieën die tracht dit hiaat op te vullen. Het is echter niet de bedoeling om een volledig exhaustief overzicht te geven, maar wel om een structuur aan te reiken die als houvast kan dienen. Verder kijken we ook naar de verschillende dimensies die gebruikt kunnen worden om het simulatieonderzoek over die categorieën heen te vergelijken. Ten slotte zullen we de voorgestelde onderverdeling gebruiken om een overzicht te geven van het huidige simulatieonderzoek, en bespreken we de mogelijke tekortkomingen van de classificatie. Eerst moeten we echter de strekking van de classificatie nauwkeurig bepalen.

1. AFBAKENING VAN DE CLASSIFICATIE

De hier voorgestelde classificatie behandelt simulaties van fenomenen uit een breed scala van taalkundige subdisciplines, beperkt tot de simulaties die betrekking hebben op taalverandering. Voor bepaalde subdisciplines zoals taalevolutie is die beperking evident. Andere velden, zoals bijvoorbeeld dat van de cognitieve taalkunde, hebben in het algemeen een bredere strekking. Concreet betekent dit dat modellen van bijvoorbeeld taalverwerving (b.v. die besproken in Alis-hahi 2010) niet behandeld zullen worden.¹ Daarnaast zullen alleen simulaties gebruikmakend van de computer aan bod komen. Taalkundige experimenten met menselijke participanten (bv. Sneller en Roberts 2018; Kirby, Cornish, en Smith 2008) worden hier beschouwd als ‘empirisch onderzoek’, hoewel ze ook soms simulaties genoemd worden. Ten slotte is het belangrijk om te vermelden dat de classificatie ook beperkt is tot agent-gebaseerde modellen. Dit betekent dat we alleen modellen bespreken die individuen simuleren in een groter systeem. Zogenaamde ‘geaggregeerde’ modellen (M. Abrams 2013, 2), modellen die het veronderstelde emergente effect rechtstreeks modelleren in plaats van met individuele agents (b.v. Nowak, Plotkin, en Jansen 2000; D. M. Abrams en Strogatz 2003), worden niet behandeld.

2. CATEGORIEËN EN TE VERGELIJKEN DIMENSIES

Onze onderverdeling op macroniveau van taalkundig simulatieonderzoek bestaat uit twee hoofdonderdelen: enerzijds de trigger van taalverandering, en anderzijds het effect dat de trigger heeft op het talige systeem. In de volgende categorieën wordt de trigger aangeduid in **vet**; het effect wordt onderlijnd:

1. **Betekenisonderhandeling** zorgt voor meer structuur in het taalstelsel
2. **Geheugen- en transmissiefouten** zorgen voor meer structuur in het taalsysteem
3. **Frequentiemechanismen** zorgen voor nieuwe taalvormen
4. **Transmissie in een gemeenschap** zorgt voor standaardisering

Deze categorieën (of combinaties van triggers en effecten) werden op voorhand bepaald (‘top-down’), waarna voor elke categorie passende voorbeelden werden gezocht. Men zou kunnen stellen dat de onderverdeling daarom wat ‘willekeurig’ is. Die kritiek is terecht, maar het blijft wel zo dat deze analyse het veld op een granulaire doch vrij algemene manier kan onderverdelen;

¹ Verderop in het artikel bespreken we een model dat ook met taalverwerving te maken heeft. De taalverwerving in dat model maakt echter deel uit van een onderzoeksvraag die specifiek met taalverandering te maken heeft. Modellen die alleen te maken hebben met taalverwerving worden natuurlijk uitgesloten van deze classificatie.

de gekozen categorieën tonen zowel het lokale gedrag in het model (in **vet**) als het emergente resultaat (onderlijnd). Sommige onderzoeken zijn echter erg veelzijdig en lijken daarom onder meerdere categorieën te kunnen vallen. Verderop geven we voorbeelden van dergelijke studies en bespreken we de gevolgen voor de robuustheid van de classificatie.

Een interessant gevolg van de gekozen categorieën is dat ze geassocieerd kunnen worden met verschillende subdisciplines van de taalkunde: de betekenisonderhandelingscategorie kan worden geassocieerd met taalevolutie, de geheugen- en transmissiefoutencategorie is gerelateerd aan cognitieve taalkunde, de frequentie categorie komt overeen met historische taalkunde, en de transmissie categorie valt onder de noemer van sociolinguïstiek. Hoewel de associaties niet zo rigide zijn als ze hier voorgesteld worden – onderzoeken kunnen (delen van) de theorie of methodologie van een andere subdiscipline gebruiken wanneer de onderzoeksvraag hierom vraagt – is het handig voor onderzoekers die nog niet vertrouwd zijn met het simulatieveld om een vorm van houvast te hebben. Dit is ook waarom de associaties opgenomen zijn in Tabel 1.

Onze onderverdeling van taalkundig simulatieonderzoek zal verder focussen op de verschillende eigenschappen of dimensies van bestaand onderzoek, geïllustreerd met twee voorbeeldonderzoeken.² Ten eerste zal voor elke categorie worden vastgesteld of betekenis een belangrijke rol speelt in het onderzoek van die categorie. We bestempelen een onderzoek als ‘betekenisgerelateerd’ wanneer er een parallellisme is tussen welke informatie door agents moet worden uitgewisseld en welke informatie *effectief* uitgewisseld wordt. Als dit parallellisme onbelangrijk of onverkend is in een studie, beschouwen we de studie als een studie met ‘mechanische processen’ in plaats van een studie die gebruik maakt van betekenis.

Ten tweede beoordelen we of de onderzoeken van een categorie *simple* dan wel *moderate* agents (M. Abrams 2013) gebruiken. Het verschil tussen *simple* en *moderate* agents komt neer op hoe uitgebreid de cognitieve implementatie van de agents is. We spreken daarom over ‘moderate agents’ als de formalisatie van een zintuig of cognitief vermogen essentieel is voor het onderzoeksdoel. Als de agents van het onderzoek niet over cognitieve vermogens beschikken, of als de implementatie van die vermogens dusdanig geabstraheerd is dat ze niet meer herkenbaar zijn als cognitie, spreken we van ‘simple agents’. Het is natuurlijk mogelijk dat beide types agents worden gebruikt binnen één categorie. (M. Abrams 2013) definieert ook de ‘complex agent’, een agent die uiterst gesofisticeerde cognitieve capaciteiten bezit. Dergelijke capaciteiten zijn in het onderzoek dat tot nu toe gevoerd is nog niet gebruikt.

² Voor lezers die meer onderzoek willen lezen uit een bepaalde categorie, zijn ook extra onderzoeken gegeven in Tabel 1. Deze vallen echter grotendeels buiten de bespreking.

Verder onderzoeken we op welke manier ‘gemeenschap’ geïmplementeerd is in de simulaties. Aangezien er altijd een notie van gemeenschap geïmpliceerd is in een agent-gebaseerd model — agents moeten ten minste met elkaar willen communiceren in dezelfde omgeving — volstaat het niet om simpelweg de vraag te stellen of er *een* vorm van gemeenschap geïmplementeerd is. Daarom focussen we op het *type* gemeenschap dat we terugvinden in de modellen van een welbepaalde categorie. Daarvoor kijken we naar de netwerkstructuur in die gemeenschap. We spreken over een ‘gestructureerde’ gemeenschap als de agents in een model op een betekenisvolle wijze met elkaar verbonden zijn. ‘Betekenisvol’ impliceert hier dat de agents op één of andere manier met elkaar verbonden zijn in een netwerk, of dat ze duidelijk verdeeld kunnen worden door middel van een lectale³ factor. Een voorbeeld van structuur in een gemeenschap is bijvoorbeeld de structuur gedefinieerd in Barabasi (1999), waarin een netwerk omschreven wordt dat een aantal ‘centrale’ leden kent met veel connecties, alsook perifere leden. We beschouwen een gemeenschap als ‘ongestructureerd’ wanneer de agents *niet* betekenisvol verbonden zijn, en behoren tot een oppervlakkige, vlakke gemeenschap zonder stammen, leiders of andere sociale structuren.

Een laatste eigenschap op basis waarvan de categorieën zullen worden vergeleken is of het onderzoek van een categorie zich richt op het ontstaan van nieuwe vormen dan wel het verdwijnen van oude vormen. In de praktijk maakt het niet uit welke vormen ‘oud’ zijn en welke ‘nieuw’. We bekijken simpelweg of het aantal vormen of constructies toegenomen is of niet, of of er een op- of neerwaartse trend kan worden waargenomen. Uiteraard correleert deze dimensie sterk met het vooropgestelde effect (onderlijnd) van elke categorie in de classificatie.⁴

2. CLASSIFICATIE

In dit onderdeel stellen we de classificatie voor die gebouwd is met de categorieën en dimensies die hiervoor werden geïntroduceerd. De volgende subsecties komen telkens overeen met één kolom uit de Tabel 1 (van links naar rechts). Per categorie worden telkens alle eigenschappen besproken aan de hand van twee voorbeeldstudies. Tabel 1 geeft een overzicht van de volledige classificatie.

³ “Alle types van taalvariëteiten of lecten: dialecten, regiolecten, nationale variëteiten, registers, stijlen, idiolecten, wat dan ook” (Geeraerts 2005, 168, eigen vertaling)

⁴ Oorspronkelijk werd voor elke categorie ook bekeken of de spreker dan wel de hoorder het belangrijkste was in de simulaties van die categorie. Het criterium bleek echter niet interessant genoeg om categorieën te vergelijken en werd daarom geschrapt.

Tabel 1: Een overzicht van de verschillende categorieën, hun geassocieerde onderzoeksvelden en hun waarden voor de vijf dimensies in de classificatie. Voorbeeldstudies zijn ook gegeven. Opgelet: deze tabel toont alleen de prototypische waarden voor elke categorie.

	Betekenis- onderhandeling	Geheugen en transmissie	Frequentie- mechanismen	Transmissie gemeenschap
Geassocieerd onderzoeksveld	taalevolutie	cognitieve taalkunde	historische taalkunde	sociolinguïstiek
Betekenis?	ja	nee	nee	nee
Agenttype (Abrams 2013)?	moderate	moderate	moderate	simple
Gestructureerde gemeenschap?	nee	nee	nee	ja
Meer of minder vormen?	minder	minder	meer	minder
Besproken voorbeeldstudies	Steels, Loetzsch, en Spranger (2016), Beuls en Steels (2013)	Kirby (1999), Stanford en Kenny (2013)	Lestrade (2016), Pijpops, Beuls, en Van de Velde (2015)	Fagyal e.a. (2010), Pierrehumbert, Stonedahl, en Daland (2014)
Onbesproken voorbeeldstudies	Dekker en De Boer (2020), Steels (1995)	K. Smith, Kirby, en Brighton (2003), Griffiths en Kalish (2007)	Pijpops (2022), Bloem (2021)	Civico (2019), Sevenants en Speelman (2021)

2.1. Betekenisonderhandeling zorgt voor meer structuur in het taalsysteem

In deze sectie bespreken we deze studies met de volgende onderzoeksvragen:

- Steels, Loetzsch, en Spranger (2016)

Kunnen robots die de fysieke werkelijkheid kunnen percipiëren met elkaar een communicatiesysteem opzetten waardoor ze de objecten in die fysieke werkelijkheid aan de andere robots kunnen signaleren?

- Beuls en Steels (2013)

Welke cognitieve processen zijn noodzakelijk bij de ondubbelzinnige communicatie over de objecten in een omgeving opdat een vorm van grammaticale overeenkomst optreedt?

In deze categorie van taalkundig simulatieonderzoek is het onderzoeksmodel op die manier gebouwd dat de virtuele agents betekenis onderhandelen op lokaal niveau. Hieruit ontstaat een vorm van structuur op het niveau van het taalsysteem in zijn geheel. Voorbeeldstudies van deze categorie zijn Steels, Loetzsch, en Spranger (2016), een studie waarin een groep robots probeert om een gemeenschappelijke woordenschat op te bouwen voor een verzameling van geometrische objecten in de echte wereld, en Beuls en Steels (2013),

een studie waarin agents proberen om gebeurtenissen in hun virtuele realiteit voldoende nauwkeurig te beschrijven, zonder ten prooi te vallen aan ambiguïteit. Om deze ambiguïteit te vermijden kunnen agents ervoor kiezen om ‘markers’ toe te voegen aan bepaalde eigenschappen van hun spraak, waardoor het model zich leent tot een onderzoek naar de oorsprong van grammaticale congruentie.

Zoals kan worden afgeleid uit de naam van de categorie, en ook uit de beschrijvingen van de voorbeeldonderzoeken, is betekenis de hoeksteen van dit soort onderzoek. In Steels, Loetzsch, en Spranger (2016) zien we bijvoorbeeld dat het hele associatieproces van betekenis met geconventionaliseerde vormen wordt onderzocht. Beuls en Steels (2013) houdt dan weer rekening met het minimaliseren van semantische ambiguïteit. Deze onderzoekscategorie kan dus zonder twijfel als betekenisgerelateerd worden bestempeld.

Het betekenisaspect brengt ons naadloos bij de noodzakelijke implementatie van cognitie. Betekenis is inherent geassocieerd met cognitieve capaciteiten, en het is essentieel voor de onderzoeksvragen in dit soort simulatieonderzoek dat agents een manier hebben om concepten in hun virtuele geheugens op te slaan, en om die concepten later ook weer te herkennen in hun gemodelleerde werelden. Moderate agents zijn daarom een absolute vereiste in het onderzoek van deze categorie. We vinden dit soort agents ook terug in de voorbeeldstudies. In Steels, Loetzsch, en Spranger (2016) kunnen de robotagents hun omgeving begrijpen en interpreteren. In Beuls en Steels (2013, 2) beschikken agents over perceptie en kunnen ze patronen, objecten en basale features herkennen.

Verder beschikt onderzoek uit deze categorie over een ongestructureerde gemeenschap. Zoals al eerder vermeld behandelen studies uit deze categorie onderzoeksvragen die te maken hebben met de fundamenteën van menselijke communicatie. Daarom zou het onwenselijk zijn om ook ingewikkelde sociale invloeden in te bouwen in de modellen van die studies, aangezien dit de interpretatie van de modellen zou bemoeilijken. Natuurlijk is de keerzijde van deze noodzakelijke vereenvoudiging dat de ecologische validiteit⁵ (Haugen 1972) van het model wordt beïnvloed. Uit andere vormen van taalkundig onderzoek weten we dat, in tegenstelling tot de veronderstellingen die worden gemaakt in dit soort onderzoek, verschillende gemeenschappen andere betekenissen geven aan dezelfde vorm. Denk bijvoorbeeld aan de zogenaamde ‘valse vrienden’ uit het model van Martin (2001, 715) van lexicale variatie tussen Belgisch en Nederlands Nederlands. Daarnaast willen verschillende sprekers niet per se dezelfde vormen gebruiken voor dezelfde betekenis. Dit verklaart bijvoorbeeld ook het bestaan van dialectatlassen, zoals die van van

⁵ In een taalkundige context wordt ‘ecologie’ gedefinieerd als “de sociale en culturele setting waarin een taal gebruikt wordt” (Garner 2017, 92, eigen vertaling)

der Sijs (2011). De ongestructureerde gemeenschap vinden we ook terug in Steels, Loetzsch, en Spranger (2016) en Beuls en Steels (2013).

Ten slotte zien we in studies van deze categorie een afname van het aantal talige vormen, en daardoor een toename van structuur. In Steels, Loetzsch, en Spranger (2016) zien we hoe een grote verscheidenheid van lexicale vormen evolueert naar een gemeenschappelijke woordenschat. In Beuls en Steels (2013) grammaticaliseren willekeurige morfologische markers en worden die zo deel van één grammatica. Natuurlijk is er vóór de uiteindelijke standaardisering eerst een grote hoeveelheid andere vormen die niet als standaard eindigen. Deze explosie van vormen zouden we kunnen karakteriseren als groei, wat indruist tegen onze vooropgestelde notie van reductie en standaardisering. Omdat deze studies echter als einddoel hebben om te onderzoeken hoe een supralokale consensus kan worden gevormd, en onze classificatie daarbovenop alleen focust op het einddoel van een experiment, herbergen we deze categorie onder een afname van vormen.

3.2. Geheugen- en transmissiefouten zorgen voor meer structuur in het taalsysteem

In deze sectie bespreken we deze studies met de volgende onderzoeksvragen:

- Kirby (1999)

Hoe beïnvloeden complexiteit en verwerkingsmoeite in een taal haar eigenschappen, en bijgevolg taalkundige universalia zoals zinsvolgorde?

- Stanford en Kenny (2013)

Kunnen in een model waarin jonge en volwassen taalleerders dezelfde capaciteiten hebben om taalinnovaties te leren, de volwassen taalleerders toch wisselender succes hebben dan de jonge taalleerders?

Deze categorie herbergt modellen die onderzoeken hoe de beperkingen van onze cognitie onze taal in de richting van structuur stuwen. Die cognitieve beperkingen uit zichzelf als beperkingen van ons geheugen, verwerkingsmoeilijkheden of problemen in de fase van L1-taalverwerving. Voorbeelden van dergelijke studies zijn Kirby (1999) en Stanford en Kenny (2013). Kirby (1999) onderzoekt hoe voordelen m.b.t. verwerking aanleiding geven tot taalkundige universalia. Stanford en Kenny (2013) onderzoeken of een model waarin jonge en volwassen taalleerders dezelfde capaciteiten hebben om taalinnovaties te leren, toch kan leiden tot de wisselende mate van succes die we associëren met volwassen taalleerders. In dit specifieke geval wordt de acquisitie van de *Northern Cities Shift* onderzocht, een reeks fonologische verschuivingen die voor een toename van structuur zorgt.

Betekenis is geen essentieel onderdeel van dit soort onderzoek. Het is mogelijk dat betekenis veralgemeent als gevolg van cognitieve bottlenecks, maar dit is een specifieke onderzoeksvraag die kan worden gesteld *binnen* dit onderzoekskader; het is dus geen vertrekpunt voor dit soort studies. Kirby (1999) focust op de complexiteit van een gegeven taaluiting, niet op haar betekenis. Stanford en Kenny (2013) hebben ook geen betekenis ingebouwd in hun studie. Hun agents communiceren slechts met enkele klanken; de betekenis van die klanken is compleet ongerelateerd aan de onderzoeksvraag. Dit soort onderzoek is dus eerder mechanisch dan betekenisgerelateerd.

In de bespreking van de vorige categorie zagen we dat het belang van betekenis vanzelf leidde tot het belang van cognitie. Aangezien betekenis in deze categorie niet zo belangrijk is, is het misschien te verwachten dat cognitie hier ook geen grote rol speelt. Dat is echter wel het geval. De cognitieve focus van deze categorie vereist de implementatie van cognitie, en daarom dus ook moderate agents. Hoewel andere aspecten van cognitie worden benadrukt in deze categorie in tegenstelling tot de vorige, blijft cognitie wel belangrijk. In Kirby (1999) zijn agents moderate omdat ze kunnen inschatten hoe moeilijk het is om een bepaalde woordvolgorde of constructie te verwerken. In Stanford en Kenny (2013) beschikken agents over een geheugen van klanken die ze recent gehoord hebben. Dat geheugen gaat langzaam achteruit. Agents kunnen hun klankgeheugen gebruiken om klanken van andere sprekers te identificeren, maar ook om innovatieve vormen te leren.

Verder is in deze categorie de implementatie van gemeenschap gewoonlijk beperkt tot de ongestructureerde gemeenschap. Kirby (1999) gebruikt bijvoorbeeld die ongestructureerde vorm. Stanford en Kenny (2013) maken echter uitgebreid gebruik van de gestructureerde gemeenschap. In hun model definiëren ze twee steden, Chicago en Saint Louis, elk met hun eigen klinkersysteem. Het ene systeem heeft een gelijke afstand tussen klinkers, het andere niet. Het suboptimale klinkersysteem ondergaat vervolgens *raising*, en contact tussen de virtuele bewoners van beide steden zorgt ervoor dat het klinkersysteem met raising van de ene naar de andere stad overspringt.⁶ Hoewel gemeenschap erg belangrijk is voor deze studie, behoort de studie toch niet tot de transmissie categorie (zie verderop). Dat komt doordat het hoofddoel van de studie, zoals eerder vermeld, is om te onderzoeken of het mogelijk is een simulatie te bouwen waarin kinderen en volwassen dezelfde cognitieve capaciteiten hebben. De standaardisering of gemeenschap is binnen dit kader slechts bijkomstig. Hoe dan ook is de gestructureerde gemeenschap geen noodzakelijk onderdeel van dit soort onderzoek, hoewel het, zoals duidelijk geworden is, wel deel kan uitmaken van een studie als de onderzoeksvraag erom vraagt.

⁶ Door een gebrek aan dicht contact hebben slechts enkele agents van de andere stad het klinkersysteem volledig kunnen overnemen, zoals was voorspeld in de onderzoeksvraag (Stanford en Kenny 2013, 143–45).

Ten slotte concentreert het onderzoek rond cognitieve beperkingen zich op een afname van het aantal taalkundige vormen dat gebruikt wordt door sprekers. Het is logisch dat cognitieve beperkingen er niet voor zullen zorgen dat taal complexer wordt (zoals gedemonstreerd in Kirby 2000). In Kirby (1999) moeten complexere constructies het onderspit delven tegen eenvoudige constructies. In Stanford en Kenny (2013) wint het eenvoudige systeem terrein ten koste van het complexe systeem. Je zou kunnen zeggen dat de introductie van het klinkersysteem met raising in de tweede stad het aantal vormen daar juist heeft doen verhogen. Dat klopt, maar tegelijkertijd moet dat proces worden bekeken in de context van een verandering *in progress* die, uiteindelijk, het systeem juist zal vereenvoudigen. Die vereenvoudiging valt wel gewoon binnen onze vooropgestelde definitie van de neerwaartse trend.

3.3. Frequentiemechanismen zorgen voor nieuwe taalvormen

In deze sectie bespreken we deze studies met de volgende onderzoeksvragen:

- Lestrade (2016)

Welke processen zijn noodzakelijk bij de ondubbelzinnige communicatie over de gebeurtenissen in een omgeving opdat een vorm van argument marking optreedt?

- Pijpops, Beuls, en Van de Velde (2015)

Welke minimale voorwaarden zijn noodzakelijk voor de verspreiding van het zwakke vervoegingssysteem in het Germaans, ten koste van het sterke vervoegingssysteem?

Een derde categorie bakent een terrein van onderzoek af waarin frequentiemechanismen zorgen voor nieuwe taalvormen of systemen waarin nieuwe taalvormen voorkomen (zoals vervoegingssystemen). Voorbeelden van dit soort studies zijn Lestrade (2016) en Pijpops, Beuls, en Van de Velde (2015). Lestrade (2016) onderzoekt de ontwikkeling van *argument markers*. Pijpops, Beuls, en Van de Velde (2015) proberen te verklaren hoe het zwakke Germaanse vervoegingssysteem dominant is kunnen worden in de Germaanse talen, maar het sterke systeem nooit volledig heeft kunnen verdringen.

Zoals in de vorige onderzoekscategorie is betekenis niet essentieel voor deze vorm van onderzoek. Het hangt af van de exacte onderzoeksvraag of de studie een formalisatie van betekenis vereist. Voor Lestrade (2016) is die formalisatie van betekenis belangrijk, aangezien de agents in zijn model moeten kunnen inschatten of een bepaalde taaluiting semantisch ambigu is of niet. De ingeschatte semantische ambiguïteit van een uiting bepaalt of ze al dan niet een argument marker toevoegen aan hun uiting. Pijpops, Beuls, en Van de Velde (2015) zijn niet expliciet geïnteresseerd in de exacte betekenis van

een woord. In hun modelarchitectuur is er geen sprake van semantische ambiguïteit; de focus ligt uitsluitend op de morfologie. Die morfologie is in dit onderzoek ongerelateerd aan betekenis.⁷ De voorbeeldstudies tonen dat, hoewel betekenis niet onlosmakelijk verbonden is met dit soort onderzoek, het ook niet compleet onmogelijk is dat ze een rol speelt als de onderzoeksvraag erom vraagt.

De situatie is vergelijkbaar voor de soort agents. De noodzakelijke hoeveelheid abstractie varieert sterk, van duidelijk moderate (Lestrade 2016) naar amper moderate (Pijpops, Beuls, en Van de Velde 2015). De agents van Lestrade (2016) moeten gebeurtenissen in hun virtuele wereld kunnen begrijpen en die linken aan hun eigen woordenschat. De agents van Pijpops, Beuls, en Van de Velde (2015) kunnen ook gebeurtenissen ‘begrijpen’, maar de gebeurtenissen in hun modelwereld zijn veel minder complex. Dat de agents in deze studie amper moderate zijn is geen gebrek; het is juist in het belang van de onderzoekers om de implementatie van het model zo eenvoudig mogelijk te houden (in functie van de onderzoeksvraag), aangezien dit de interpretatie van het model bevordert. Dit is dus wederom een dimensie waarin de onderzoeksvraag een belangrijke rol speelt. Aangezien de agents van onze voorbeeldonderzoeken nog steeds over een bepaalde graad van cognitie beschikken, beschouwen we deze categorie als ‘moderate’.

Ook dit soort simulatieonderzoek leunt doorgaans op de ongestructureerde gemeenschapsstructuur. Frequentiegebaseerde simulaties kunnen wel gebruik maken van een gestructureerde gemeenschap — zo zien we deze soort gemeenschap opduiken in de bijkomende studies van Pijpops (2022) en Bloem (2021). Zij maken gebruik van twee verschillende groepen met andere grammaticale voorkeuren. Toch is het, omwille van dezelfde redenen omtrent eenvoud die we al eerder aanhaalden, aangewezen om de gemeenschap zo eenvoudig mogelijk te houden (tenzij het uiteraard past in de onderzoeksopzet). De prototypische gemeenschapsvorm blijft dan ook de ongestructureerde gemeenschapsvorm, zoals die voorkomt in onze voorbeeldstudies.

Zoals de naam van deze categorie doet vermoeden, gaat het in deze categorie over een toename van het aantal beschikbare vormen. In Lestrade (2016) wordt de toename van *geen* argument-markingstrategie naar *een* argument-markingstrategie onderzocht. In Pijpops, Beuls, en Van de Velde (2015) wordt onderzocht hoe het Germaans van een naar twee werkwoordvervoegingssystemen ging. Hoewel die laatstgenoemde studie het bestaan van het zwakke systeem veronderstelt (de studie probeert de oorsprong van het systeem niet te achterhalen), wordt het zwakke systeem verondersteld in een prille fase. Dat betekent dat de simulatie, hoewel ze niet letterlijk draait rond het ‘ontstaan’ van het zwakke systeem, wel probeert te onderzoeken hoe de

⁷ Een beoordelaar wijst erop dat in bepaalde theoretische frameworks zoals constructiegrammatica vorm altijd gerelateerd is aan betekenis. In dit onderzoek is dat niet het geval.

taal evolueerde van één actief gebruikt systeem naar twee actief gebruikte systemen. Dat is opnieuw een toename.

3.4. Transmissie in een gemeenschap zorgt voor standaardisering

In deze sectie bespreken we deze studies met de volgende onderzoeksvragen:

- Fagyal e.a. (2010)
Hoe verloopt de verspreiding van taalkundige varianten onder verschillende vormen van sociale structuren?
- Pierrehumbert, Stonedahl, en Daland (2014)
Hoe kan een taalkundige innovatie die ontstaat bij 'gewone' (niet-centrale) mensen toch verspreid raken?

De laatste onderzoekscategorie beslaat alle modellen die uitsluitend gericht zijn op het verspreiden van taalvarianten in een gemeenschap, en de eventuele standaardisering van die varianten. Fagyal e.a. (2010) en Pierrehumbert, Stonedahl, en Daland (2014) zijn voorbeelden van dit soort onderzoek. In het model van Fagyal e.a. (2010) wordt onderzocht hoe een innovatieve talige vorm zich verspreidt door een sociaal netwerk. Daarbij wordt gebruik gemaakt van de lokale populariteit van agents: hoeveel contacten hebben ze in vergelijking met anderen? De taal van populaire agents wordt daarbij sneller overgenomen dan die van onpopulaire agents. Pierrehumbert, Stonedahl, en Daland (2014) onderzoeken ook de verspreiding van talige vormen in een netwerk. In tegenstelling tot Fagyal e.a. (2010) proberen de onderzoekers hier de circulatie van een nieuwe taalvariant te modelleren zonder gebruik te maken van sociale status. In plaats daarvan laten ze agents de taalvarianten verkiezen van agents die dicht bij hen staan. Zoals kan worden afgeleid uit de voorbeelden is het onderzochte fenomeen in dit soort studies altijd ongeveer hetzelfde, maar het perspectief van waaruit het onderzocht wordt (en de bijbehorende implementatie) varieert sterk.

Het onderwerp van transmissie-onderzoek is ver verwijderd van betekenis. Het maakt niet uit wat de exacte betekenis is van de vorm die wordt doorgegeven van agent tot agent; het is eigenlijk zelfs niet belangrijk op welk taalkundig niveau de studie zich afspeelt. Of er nu een klank, een woord of een constructie doorgegeven wordt, het belangrijkste aspect van transmissie-onderzoek is dat de verspreiding van een vorm geheel ongerelateerd is aan taal zelf (taalinterne taalveranderingen behoren tot de andere drie categorieën). Zowel Fagyal e.a. (2010) als Pierrehumbert, Stonedahl, en Daland (2014) formaliseren betekenis niet.

Cognitie is ook niet belangrijk in deze vorm van onderzoek, wat betekent dat transmissiestudies beperkt zijn tot simple agents. In tegenstelling tot de

studies uit de andere categorieën is de gesimuleerde ‘conversatie’ tussen agents bij transmissiestudies relatief beperkt. In feite wordt de conversatie als dusdanig volledig overgeslagen, en ligt de focus volledig op het beoogde effect: de uitwisseling van talige vormen. Dat betekent dat er in sommige modellen geen opzet is met gebeurtenissen waarover moet worden gecommuniceerd. Die communicatie wordt simpelweg verondersteld. De eigenheid van de modellen ligt dus op het niveau van de implementatie van het mechanisme dat beslist welke agent een talige vorm doorgeeft aan de ander. We zagen in de bespreking van de voorbeeldstudies voor deze categorie al dat het mechanisme van Fagyal e.a. (2010) relatieve populariteit gebruikt. Bij Pierrehumbert, Stonedahl, en Daland (2014) is dat mechanisme afhankelijk van nabijheid. Agents hebben in die laatste studie ook wel de eigenschappen ‘temperature’ (hoe zeer ze geneigd zijn om willekeurige beslissingen te nemen) en ‘bias’ (hoe conservatief ze zijn), maar die eigenschappen kunnen we nauwelijks formalisaties van cognitie noemen. We noemen de agents hier dus simple agents.

Wat wél belangrijk is in dit soort onderzoek is de gestructureerde gemeenschap. Het zijn de subgemeenschappen en de subgemeenschappen alleen die de drijvende kracht zijn achter de gesimuleerde taalverandering in deze modellen, en het is ook dit aspect waar de onderzoeksvragen van deze categorie betrekking op hebben. In Fagyal e.a. (2010) verschilt het tempo waaraan standaardisering zich voltrekt aanzienlijk naargelang hoe de gemeenschap in de simulatie is georganiseerd. Soms zorgt de gemeenschapsvorm er zelfs voor dat er helemaal geen standaardisering optreedt (Fagyal e.a. 2010, 2071–72). Pierrehumbert, Stonedahl, en Daland (2014) zien ander emergent gedrag in hun model naargelang hoe de aantrekking tot taalkundige innovaties wordt verspreid. Het zijn dit soort alternanties die de kern vormen van taalkundig simulatieonderzoek.

Ten slotte focust transmissie-onderzoek zich op een afname in talige vormen, of in sommige gevallen louter op een vervanging. Dit soort onderzoek probeert namelijk uit te zoeken hoe een gemeenschap met verschillende concurrerende vormen uiteindelijk, met alleen sociale factoren, een variant kan uitkiezen als de standaard, wat een afname van vormdiversiteit betekent. Uiteraard is het ook mogelijk dat een bestaande standaard wordt verdrongen door een andere. Dan is er toename noch afname. Fagyal e.a. (2010) onderzoekt afname, Pierrehumbert, Stonedahl, en Daland (2014) onderzoeken vervanging.

4. GENERALISATIES, AMBIGUE STUDIES EN DE ROBUUSTHEID VAN DE CLASSIFICATIE

Wanneer we kijken naar de verschillen tussen de categorieën in Tabel 1, zien we dat de transmissiecategorie het hardste verschilt van alle andere categorieën (tussen de 2 en 3 verschillen met alle andere categorieën). Dat is niet

verwonderlijk, aangezien we in de bespreking van de categorieën zagen dat deze categorie het meeste over de taalkundige component abstraheert. Als we verder kijken naar de onderlinge verschillen (de transmissie categorie buiten beschouwing gelaten), is er het grootste verschil tussen de betekenis- en frequentie categorieën. Ook dit verschil is te verklaren in de verschillende doelen van het onderzoek uit deze categorieën. Waar de betekenis categorie zich vooral bezighoudt met het ontstaan van succesvolle informatie-uitwisseling, is de frequentie categorie meer gefocust op uitbreidingen *na* dit ontstaan.

Tegelijkertijd is het gepast om de robuustheid van de classificatie beter te bekijken, aangezien sommige simulatiestudies moeilijk te classificeren blijken. Ze lijken namelijk tot verschillende categorieën tegelijkertijd te behoren. De classificatie van dit soort onderzoeken hangt nog meer af van de exacte onderzoeksvraag en van de gebruikte technieken. We zagen deze ambiguïteit bijvoorbeeld bij het onderzoek van Stanford en Kenny (2013). Een gelijkaardig probleem zien we bij het onderzoek van Beuls en Steels (2013) over de oorsprong van grammaticale congruentie. Het is misschien verwonderlijk dat deze studie, die sterk lijkt op die van Lestrade (2016) over de oorsprong van argument marking, toch niet in dezelfde categorie staat. Dit verschil in classificatie komt voort uit een verschil in methodologie. Het model van Lestrade (2016) maakt uitgebreid gebruik van frequentie, waardoor het wordt geclassificeerd als een frequentiemodel. Beuls en Steels (2013) gebruiken geen frequentie an sich in hun model, maar wel ‘laterale inhibitie’. Dit is een systeem waarin, als een bepaalde vorm geassocieerd met een betekenis gebruikt wordt, de score van die vorm verhoogd wordt, en de score van alle andere vormen verlaagd (Beuls en Steels 2013, 4). Omdat dit mechanisme een *afgeleide* is van absolute frequentie, wordt het artikel als bij betekenisonderhandeling ingedeeld.

De ‘problemen’ om deze studies te classificeren leiden misschien tot vragen over de robuustheid van de classificatie. We mogen echter niet vergeten dat andere studies veel ondubbelzinniger in de gedefinieerde categorieën passen. Pijpops, Beuls, en Van de Velde (2015) is een voorbeeld van een dergelijke ondubbelzinnige studie. We zagen ook dat de classificatie in staat is om betekenisvolle generalisaties te maken, en dat de categorieën relatief onprobleematisch geassocieerd kunnen worden met bestaande taalkundige disciplines, wat vooral voor nieuwe onderzoekers binnen het veld erg handig is. Die eigenschappen van de classificatie versterken haar positie. Toch moeten we erkennen dat de kern van simulatieonderzoek erg veelzijdig is, en dat classificatie daarmee een moeilijke taak is. Dat wordt weerspiegeld in de ambigue studies die hierboven besproken werden. Tegelijkertijd hopen we met deze classificatie toch de krijtlijnen te hebben getrokken die kunnen dienen als wegwijzer om onderzoeken beter te kunnen plaatsen en vergelijken.

BIBLIOGRAFIE

- ABRAMS, Daniel M., en Steven H. STROGATZ. 2003. 'Modelling the Dynamics of Language Death'. *Nature* 424 (6951): 900–900. <https://doi.org/10.1038/424900a>.
- ABRAMS, Marshall. 2013. 'A Moderate Role for Cognitive Models in Agent-Based Modeling of Cultural Change'. *Complex Adaptive Systems Modeling* 1 (1): 1–33. <http://dx.doi.org/10.1186/2194-3206-1-16>.
- ALISHAHI, Afra. 2010. 'Computational Modeling of Human Language Acquisition'. *Synthesis Lectures on Human Language Technologies* 3 (1): 1–107. <https://doi.org/10.2200/S00304ED1V01Y201010HLT011>.
- BARABASI, Albert-Laszlo. 1999. 'Emergence of Scaling in Random Networks'. *ArXiv: Org* 286 (5439): 509–12. <https://doi.org/10.1126/science.286.5439.509>.
- BARONCHELLI, Andrea. 2014. 'Modeling Is a Tool, and Data Are Crucial: A Comment on "Modelling Language Evolution: Examples and Predictions" by Tao Gong et Al.' *Physics of Life Reviews* 11 (2): 317–18. <https://doi.org/10.1016/j.plrev.2014.01.014>.
- BEULS, Katrien, en Luc STEELS. 2013. 'Agent-Based Models of Strategies for the Emergence and Evolution of Grammatical Agreement'. *PloS One* 8 (3): e58960–e58960. <https://doi.org/10.1371/journal.pone.0058960>.
- BEULS, Katrien, en Remi VAN TRIJP. 2016. 'Computational Construction Grammar and Constructional Change'. *Belgian Journal of Linguistics* 30 (1): 1–13. <https://doi.org/10.1075/bjl.30.01beu>.
- BLOEM, Jelke. 2021. *Processing Verb Clusters*. NL: LOT. <https://doi.org/10.48273/LOT0586>.
- BYBEE, Joan. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511750526>.
- CIVICO, Marco. 2019. 'The Dynamics of Language Minorities: Evidence from an Agent-Based Model of Language Contact'. *Journal of Artificial Societies and Social Simulation* 22 (4): 27. <https://doi.org/10.18564/jasss.4097>.
- DEKKER, Peter, en Bart DE BOER. 2020. 'Neural Agent-based Models To Study Language Contact Using Linguistic Data: 4th NeurIPS Workshop on Emergent Communication'. In , 1–6. <https://sites.google.com/view/emecom2020/accepted-papers>.
- FAGYAL, Zsuzsanna, Samarth SWARUP, Anna María ESCOBAR, Les GASSER, en Kiran LAKKARAJU. 2010. 'Centers and Peripheries: Network Roles in Language Change'. *Lingua* 120 (8): 2061–79. <https://doi.org/10.1016/j.lingua.2010.02.001>.
- GARNER, Mark. 2017. 'Language Ecology as Linguistic Theory'. *Kajian Linguistik Dan Sastra* 17 (2): 91–101. <https://doi.org/10.23917/kl.v17i2.4485>.
- GEERAERTS, Dirk. 2005. 'Lectal Variation and Empirical Data in Cognitive Linguistics'. In *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction*, 32:163–89. Cognitive Linguistics Research. Berlin: Mouton de Gruyter. <https://lirias.kuleuven.be/retrieve/374363>.
- GONG, Tao, en Lan SHUAI. 2013. 'Computer Simulation as a Scientific Approach in Evolutionary Linguistics'. *Language Sciences* 40 (november): 12–23. <https://doi.org/10.1016/j.langsci.2013.04.002>.
- GRIFFITHS, Thomas L., en Michael L. KALISH. 2007. 'Language Evolution by Iterated Learning With Bayesian Agents'. *Cognitive Science* 31 (3): 441–80. <https://doi.org/10.1080/15326900701326576>.

- HAUGEN, Einar. 1972. 'The Ecology of Language'. In *The Ecology of Language: Essays by Einar Haugen. Selected and Introduced by Anwar S. Dil*, onder redactie van Anwar S. Dil, 325–39. Language Science and National Development Series. Stanford (Calif.): Stanford university press.
- KIRBY, Simon. 1999. *Function, Selection, and Innateness: The Emergence of Language Universals*. Oxford Linguistics. Oxford: university press.
- . 2000. 'Syntax Without Natural Selection: How Compositionality Emerges from Vocabulary in a Population of Learners'. In *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form*, onder redactie van Chris Knight, James Hurford, en Michael Studdert-Kennedy, 303–23. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511606441.019>.
- KIRBY, Simon, Hannah CORNISH, en Kenny SMITH. 2008. 'Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language'. *Proceedings of the National Academy of Sciences* 105 (31): 10681–86. <https://doi.org/10.1073/pnas.0707835105>.
- LESTRADE, Sander. 2016. 'The Emergence Of Argument Marking'. In *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANGX11)*, onder redactie van S. G. Roberts, C. Cuskley, L. McCrohon, L. Barceló-Coblijn, O. Fehér, en T. Verhoef. Online at <http://evolang.org/neworleans/papers/36.html>.
- MARTIN, Willy. 2001. 'Natiolectismen in het Nederlands en hun lexicografische beschrijving'. *Revue belge de Philologie et d'Histoire* 79 (3): 709–36. <https://doi.org/10.3406/rbph.2001.4544>.
- NOWAK, Martin A., Joshua B. PLOTKIN, en Vincent A. A. JANSEN. 2000. 'The Evolution of Syntactic Communication'. *Nature* 404 (6777): 495–98. <https://doi.org/10.1038/35006635>.
- PIERREHUMBERT, Janet B., Forrest STONEDAHL, en Robert DALAND. 2014. 'A model of grassroots changes in linguistic systems'. *arXiv:1408.1985 [nlin, physics:physics]*, augustus. <http://arxiv.org/abs/1408.1985>.
- PIJPOPS, Dirk. 2022. 'Lectal Contamination: Evidence from Corpora and from Agent-Based Simulation'. *International Journal of Corpus Linguistics* 27 (3): 259–90. <https://doi.org/10.1075/ijcl.20040.pij>.
- PIJPOPS, Dirk, en 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 (december): 5–23.
- PIJPOPS, Dirk, Katrien BEULS, en 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 (november): 81–102.
- SEVENANTS, Anthe, en Dirk SPEELMAN. 2021. 'Keeping up with the Neighbours - An Agent-Based Simulation of the Divergence of the Standard Dutch Pronunciations in the Netherlands and Belgium'. *Computational Linguistics in the Netherlands Journal* 11 (december): 5–26.
- SIJS, Noline van der. 2011. *Dialectatlas van het Nederlands*. Amsterdam: Bakker.
- SMITH, Andrew D. M. 2014. 'Models of Language Evolution and Change'. *WIREs Cognitive Science* 5 (3): 281–93. <https://doi.org/10.1002/wcs.1285>.
- SMITH, Kenny, Simon KIRBY, en Henry BRIGHTON. 2003. 'Iterated Learning: A Framework for the Emergence of Language'. *Artificial Life* 9 (4): 371–86. <https://doi.org/10.1162/106454603322694825>.

- SNELLER, Betsy, en Gareth ROBERTS. 2018. 'Why Some Behaviors Spread While Others Don't: A Laboratory Simulation of Dialect Contact'. *Cognition* 170 (januari): 298–311. <https://doi.org/10.1016/j.cognition.2017.10.014>.
- STANFORD, James N., en Laurence A. KENNY. 2013. 'Revisiting Transmission and Diffusion: An Agent-Based Model of Vowel Chain Shifts across Large Communities'. *Language Variation and Change* 25 (2): 119–53. <https://doi.org/10.1017/S0954394513000069>.
- STEELS, Luc. 1995. 'A Self-Organizing Spatial Vocabulary'. *Artificial Life* 2 (3): 319–32. <https://doi.org/10.1162/artl.1995.2.3.319>.
- STEELS, Luc, Martin LOETZSCH, en Michael SPRANGER. 2016. 'A Boy Named Sue: The Semiotic Dynamics of Naming and Identity'. *Belgian Journal of Linguistics* 30 (1): 147–69. <https://doi.org/10.1075/bjl.30.07ste>.
- VALENZUELA, Javier, Joseph HILFERTY, en Oscar VILARROYA. 2016. 'Why Are Embodied Experiments Relevant to Cognitive Linguistics?' *Belgian Journal of Linguistics* 30 (1): 265–86. <https://doi.org/10.1075/bjl.30.12val>.
- VAN EERTEN, Laura. 2007. 'Over het corpus gesproken nederlands'. *Nederlandse Taalkunde* 12 (3): 194–215.