



Universiteit
Leiden
The Netherlands

Constructions emerging : a usage-based model of the acquisition of grammar

Beekhuizen, B.F.

Citation

Beekhuizen, B. F. (2015, September 22). *Constructions emerging : a usage-based model of the acquisition of grammar*. *LOT dissertation series*. LOT, Utrecht. Retrieved from <https://hdl.handle.net/1887/35460>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/35460>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/35460> holds various files of this Leiden University dissertation

Author: Beekhuizen, Barend

Title: Constructions emerging : a usage-based model of the acquisition of grammar

Issue Date: 2015-09-22

Samenvatting

Hoe kinderen binnen zo'n korte tijd de taal van hun gemeenschap verwerven, is een van de centrale vragen in de taalkunde. De gebruiksgebaseerde, constructivistische benadering van taalverwerving stelt dat kinderen dit doen door gebruik te maken van domein-algemene leermechanismes zoals sociale cognitie en patroonherkenning. Het computationeel modelleren hiervan (dus: het nabootsen van het gedrag van een kind door belangrijke aspecten van de gebruiksgebaseerde hypothese te formaliseren en implementeren als computerprogramma's) wordt een steeds belangrijker methode in het veld van de kindertaalverwerving. Dit proefschrift snijdt vier centrale zaken in het computationeel modelleren van de gebruiksgebaseerde benadering aan:

- Het bereiken van een grotere omvattendheid van de computationele cognitieve modellen: het model moet het hele proces, van geen tot veel taalkennis, zowel in productie als begrip, kunnen uitvoeren.
- Het bereiken van een naturalistischere manier van het modelleren van betekenisverwerving: de informatie die een kind tot haar beschikking heeft, moet zo getrouw mogelijk aan het model gegeven worden.
- Een herwaardering van de 'begin-klein' hypothese binnen de gebruiksgebaseerde theorie: kinderen breken niet alleen grotere gehelen op in hun delen (zoals de 'begin-groot' hypothese stelt), maar leren ook grotere gehelen te vormen op basis van de kleine delen.
- Een nader onderzoek naar de aard en conceptualisatie van de leermechanismes en de algoritmes die deze mechanismes instantiëren in het modelleren ervan: veel leermechanismes worden nog steeds beschreven in deductivistische of rationalistische termen, terwijl deze twee perspectieven geen natuurlijke denkwijzen binnen de gebruiksgebaseerde benadering zijn.

Naast deze specifieke zaken, bespreek ik in hoofdstuk 2 een verzameling algemene theoretische desiderata en empirische explananda waaraan een computationeel model zou moeten voldoen. Eerdere modellen hebben belangrijke voortgang op deze desiderata en explananda gemaakt, en een belangrijk doel van deze dissertatie is het bijeenbrengen van deze inzichten.

Als we een omvattend model willen ontwikkelen (d.w.z. een model dat zowel taalbegrip als -productie nabootst), hebben we een antwoord nodig op de vraag hoe kinderen de communicatieve intentie van de spreker begrijpen zonder de talige elementen die deze uitdrukken te begrijpen. Computationele modellen die zich met betekenisverwerving bezig houden, nemen meestal aan dat het kind beschikt over een verzameling conceptualisaties van de situatie waar de taaluiting over zou kunnen gaan. In hoofdstuk 4, heb ik de aanneemelijkheid van deze aanname onderzocht. Ik bevond dat de niveau's van 'ruis' (de afwezigheid van in de zin aanwezige betekenselementen uit de situationele context) en 'onzekerheid' (de mate van aanwezigheid van in de situationele context aanwezige mogelijke betekenselementen die niet in de zin worden uitgedrukt) in de meeste computationele modellen laag worden ingeschat vergeleken met de waardes die we in eigenlijke ouder-kindinteractie aantreffen. Deze conclusie bereikte ik door middel van een studie van een grote verzameling ouder-kindinteracties waarin alle conceptuele elementen in de aandacht van het kind en de ouder en alle taaluitingen precies beschreven zijn. Een verder inzicht uit deze studie was dat opeenvolgende gebeurtenissen in hoge mate afhankelijk van elkaar zijn: als de ouder op het ene moment een handeling op een bal uitvoert, dan is het zeer waarschijnlijk dat deze nog een handeling op die bal zal uitvoeren erna (i.p.v. op een ander object).

Aangezien het met de hand beschrijven van de data een tijdrovend proces is, is deze methode niet op te schalen naar de hoeveelheid data die een computationeel model nodig heeft. De studie naar de eigenschappen van ouder-kindinteractie 'in het wild' gaf ons evenwel wel een mogelijkheid om de methode van Alishahi & Stevenson (2010) om kunstmatig input voor het model te genereren, aan te passen. Deze aangepaste procedure genereert input items, paren van een taaluiting en een situationele context op basis van een realistische inschatting van de waarden van 'ruis' en 'onzekerheid'. Daarnaast vormt de situationele context een keten: welke situatie de volgende zal zijn, hangt af van de huidige situatie.

In hoofdstuk 3 formaliseer ik het model: de Syntagmatisch-Paradigmatische Leerder (SPL). Het model begint zonder enige talige kennisinhouden, en leert geleidelijk zinnen te begrijpen en te produceren. In het proberen zinnen in situationele contexten te verwerken, bouwt SPL een constructicon op, een inventaris van zowel lexicale als grammaticale constructies. De leermechanismen van SPL kunnen het best gezien worden als de sporen van het verwerken van taal (i.p.v. als hypothese-testende operaties, zoals taalverwerving vaak, op deductivistische wijze, voorgesteld wordt). De constructies die SPL leert zijn paren van signifiërende elementen (fonologische en conceptuele structuren) en gesignifiëerde conceptuele

structuur.

Het model verwerkt elk input-item met het constructicon dat het op dat moment heeft. Deze verwerking geschiedt zonder een optimalisatie over de hele zin: SPL verwerkt de zin lineair en onthoudt alleen de meest waarschijnlijke analyse na elk woord. De analyse die hier uitkomt vormt de input voor de leerprocedure. Door middel van een verzameling leermechanismes vormt SPL nieuwe constructies en versterkt het bestaande. De leermechanismes vormen de centrale theoretische vernieuwing van het model door nauwgezet de gebruiksgebaseerde benadering van Langacker (1988) te operationaliseren en implementeren. Het model bereikt dit doel doordat alle leermechanismes, mogelijk met uitzondering van het cross-situationele leren (zie beneden), online leermechanismes zijn. Dat wil zeggen: het zijn geen post-hoc, of achteraf plaatsvindende, operaties op het constructicon die het constructicon reorganiseren. De leermechanismes kunnen gezien worden als de sporen die het verwerken van de zinnen in hun situationele contexten achterlaten in de geest van de taalgebruiker. Deze sporen kunnen op verschillende niveau's gevonden worden.

Ten eerste laten de concreetste representaties van de uitingen die verwerkt worden een spoor achter op het representatieve systeem. Deze operatie heeft het effect dat zeer concrete constructies, als ze maar vaak genoeg gebruikt worden, met de tijd representatief sterker kunnen worden. We kunnen dit effect interpreteren als de formatie van prototypes van categorieën: de vaak versterkte, zeer concrete representaties zijn hogelijk toegankelijk voor het model in het analyseren en produceren van uitingen.

Ten tweede worden de concreetste *gebruikte* constructies versterkt. Dit zijn de constructies die het model heeft gebruikt in de analyse, maar die mogelijk abstracter zijn dan de gebruikssituatie zelf. De versterking van deze representatie staat het model toe om vaak gebruikte abstractere constructies in representatieve kracht te laten groeien. Het potentieel van een constructie hangt daarmee dus af van de frequentie van gebruik ervan. Dit kenmerk van SPL instantieert de notie van het effect van typefrequentie, zoals Bybee (2006) die bespreekt. Een abstracte constructie zal normaliter alleen één maal versterkt worden voor elk uniek gebruiksgeval waar het voor gebruikt wordt. Wanneer het model datzelfde gebruiksgeval opnieuw tegenkomt, zal het model immers een concretere constructie hebben (door het eerste leermechanisme) die dat gebruiksgeval beter dekt. Dit mechanisme instantieert niet alleen het effect van typefrequentie, maar ook van tokenfrequentie: als een constructie frequent gebruikt wordt, zal zijn representatieve kracht toenemen en een cognitieve routine gaan vormen.

Ten derde bouwt het model toenemend lange constructies op door middel van de syntagmatizatie-operatie. Syntagmatizatie is het spoor dat achtergelaten wordt in de representaties wanneer er meerdere, kortere, constructies naast elkaar verwerkt worden. Deze kortere constructies vormen dan tezamen de constituenten van een langere constructie. Syntagmatizatie is een bottom-up leerprocedure waarmee SPL grotere representatieve eenheden opbouwt.

Ten slotte vormt het model abstracties met de paradigmatisatie-operatie. Deze operatie staat het model toe taalkennis op ongeziene uitingen toe te passen. Paradigmatizatie neemt de gemeenschappelijke elementen van twee constructies en 'onttrekt' deze om er een nieuwe constructie van te vormen. Deze abstracties zijn evenwel alleen onttrokken in de implementatie. Er vindt immers geen selectie over de abstracties plaats en derhalve kunnen ze als 'immanent' in de concretere constructies worden gezien. Door de versterking van de concreetste gebruikte constructie (het tweede leermechanisme) kunnen ze zelf ook representatieve kracht opdoen (cf. de beschrijving van Langacker (2009) van hoe abstracties de status van een 'unit' kunnen bereiken). Op deze manier vindt de selectie van 'goede' of 'bruikbare' abstracties plaats zonder dat het model een globale evaluatie van het constructicon uitvoert.

De eerste representaties van het model komen tot stand door cross-situationeel leren. Dit mechanisme vergelijkt recente gebruikgevallen met elkaar en onttrekt alle overlappen in uiting en situatie tussen deze als initiële lexicaal constructies. Een tweede manier om nieuwe constructies te ontdekken is door het gebruik van de 'bootstrapping'-operator. Bootstrapping is het vermogen van het model om een niet-fonologisch gevulde constituent van een constructie van toepassing te laten zijn op een woord, zonder dat het dat woord kent.

Beide leermechanismen zorgen ervoor dat het model 'chunks' kan extraheren: intern niet geanalyseerde lexicaal constructies die groter zijn dan een enkel woord in de taal van een volwassene. Deze chunks worden evenwel niet in hun delen opgebroken door de paradigmatisatie-operatie. Dit zou immers veronderstellen dat het model ze achteraf, en dus niet on-line herinterpreteert, en dat is een operatie die ik wilde voorkomen om het leren daadwerkelijk een neveneffect van het verwerken te laten zijn.

In hoofdstuk 3 betoog ik dat dit model behoorlijk goed de desiderata instantieert. Voor zover ik weet, is het het eerste gebruikgebaseerde computationele model dat zowel zinnen kan begrijpen als produceren zonder dat het met enige representatieve kennis van de taal begint. SPL instantieert verder belangrijke aspecten van de gebruikgebaseerde theorie: de representaties zijn zowel kwalitatief als kwantitatief gegrond in de talige gebruikgevallen, hun representatieve kracht hangt af van de frequentie van gebruik, de geleerde abstracties zijn immanent, en SPL verwerkt zinnen op een redelijk realistische wijze (lineair en zonder alle mogelijke analyses bij te houden).

SPL's gedrag wordt vervolgens geëvalueerd in hoofdstukken 5 (begrip) en 7 (productie). In het begripexperiment onderzoek ik hoe goed het model de juiste interpretatie aan een zin kan geven door het te laten kiezen uit een aantal mogelijke interpretaties. Daarnaast bekijk ik hoe goed SPL de zin en de geïnterpreteerde situatie dekt met de beschikbare taalkennis. Op alle drie de vlakken zien we dat SPL een steeds competentere taalgebruiker wordt. In de productie-experimenten zien we dat, gegeven een situatie die uitgedrukt moet worden, SPL steeds langere en adequatere zinnen vormt. De fouten die het model maakt zijn voornamelijk fouten van weglating (het niet produceren

van woorden die een volwassene wel zou produceren) en niet van toevoeging (het wel produceren van woorden die een volwassene niet zou produceren. Dit is in overeenstemming met wat kinderen doen.

Daarnaast observeer ik in beide hoofdstukken dat het model robuust is. De parameters voor ruis en onzekerheid zijn voor de hierboven besproken experimenten ingesteld op grond van het onderzoek in hoofdstuk 4, maar we kunnen ons afvragen hoe het model presteert als we deze waarden hoger leggen. SPL blijkt relatief goed hogere niveaus van ruis en onzekerheid aan te kunnen, als de situaties maar een keten vormen. Zodra de 'coherentie' van de situatie wegvalt, gaat de prestatie beduidend achteruit. Deze bevinding suggereert dat de coherentie van de situatie een belangrijke rol speelt. De reden hiervoor zou kunnen zijn dat zelfs bij de misidentificatie van de situatie, er nog steeds relatief veel elementen in die verkeerd geïdentificeerde situatie zijn die wel correct zijn.

Naast deze algemene tendensen kijk ik in de hoofdstukken 5 en 7 ook op een gedetailleerder niveau naar het gedrag van het model. In de productie-experimenten, bijvoorbeeld, observeerde ik dat het aantal uitgedrukte argumenten groeide als een effect van het toenemend aantal grammaticale constructies, en niet per se als gevolg van een toenemend aantal woorden. In veel gevallen had het model de juiste woorden wel, maar produceerde het deze toch niet, omdat het geen grammaticale constructies had om deze woorden mee te combineren. Het bekende effect dat vooral grammaticale onderwerpen worden weggelaten kon ik niet nabootsen, maar dit is, m.i., te verklaren doordat het model geen pragmatische kennis meeneemt en doordat het geen focus op de rechterkant van de zin heeft, twee fenomenen waarvan we weten dat ze van invloed zijn op het weglaten van onderwerpen.

Een centrale vraag in de kindertaalverwerving is waarom kinderen soms argumentstructuurconstructies gebruiken waar volwassenen deze niet zouden gebruiken en hoe ze dit gedrag 'afleren'. Deze twee fenomenen zijn gemodelleerd in hoofdstuk 7. Het antwoord dat SPL biedt is dat het vrij snel een inventaris van abstracte, en dus generaliseerbare constructies opbouwt en die vrij vroeg combineert met werkwoorden waar deze constructies niet mee gecombineerd kunnen worden (bv. *you fall ball* in een situatie waarin iemand een bal laat vallen). De aanwezigheid van alternatieve constructies (bv. *you drop ball*) voorkomt, even later, dat deze combinaties nog gebruikt worden. Ik betoog in dat hoofdstuk dat het blokkeren van de 'foute' constructie op twee manieren gebeurt. Ten eerste is de mate van representatieve kracht van de alternatieve constructie van belang: hoe sterker deze is, hoe sneller het model de overgeneralisatie 'afleert'. Ten tweede vinden we het effect dat hoe vaker het model zinnen als *ball fall* tegenkomt, hoe waarschijnlijker de blokkade van *you fall ball* wordt. Dit laatste effect is een geval van latente blokkade: we zien het niet terug in het gedrag, aangezien het model niet *ball fall* zal produceren (omdat dit minder expressief is dan, bv. *you fall ball* of *you drop ball*, waar telkens de agens genoemd is).

Een interessante eigenschap van computermodellen is dat we, in tegenstelling tot bij kinderen, hun interne representaties kunnen bestuderen los van het gedrag dat het model vertoont. Dit doe ik in hoofdstuk 6. Een eerste bevinding hier is dat, hoewel alle leermechanismes te allen tijd beschikbaar zijn voor het model, ze met een variabele frequentie gebruikt worden. Voor de verwerving van lexicale constructies zien we dat de naïeve methode van het cross-situationeel leren slechts in de allereerste fases gebruikt wordt, waarna het model een inventaris van half-open en open grammaticale constructies opgebouwd heeft waarmee het de betekenis van onbekende woorden kan 'bootstrappen'. De paradigmatisatie-operatie, op het grammaticale vlak, vertoont verder interessante uitbarstingen van activiteit over de ontwikkelingstijd, wat betekent dat het model niet gradueel tot nieuwe abstracties komt, maar dat het gebruiksgevallen tegenkomt die nieuwe deelruimtes van de mogelijkhedenruimte van grammaticale constructies ontsluit.

De abstracties die SPL op deze manier leert, vertonen verder de interessante eigenschap dat ze niet direct te observeren hoeven te zijn in het gedrag van het model in begrip en productie. Als we niet 'onder de motorkap' hadden gekeken, zouden we tot de onjuiste conclusie kunnen komen dat het representatieve systeem hogelijk concreet is. Deze redenering klopt evenwel niet: uitgaand van de gebruiksgebaseerde stelling dat taalgebruikers in het gebruik concretere constructies de voorkeur geven boven abstractere, valt het te verwachten dat in het gebruik voornamelijk die concrete constructies zullen opduiken. Tegelijkertijd heeft het model echter een sterker potentieel om te generaliseren. Sterker nog: dit potentieel komt zeer vroeg op en het model leert daarna voornamelijk concretere constructies erbij die deze abstracties verder blokkeren. Dit is te verwachten: zodra abstracties beschikbaar zijn, zal het model deze gebruiken om expressief te zijn, tenzij het een concretere constructie beschikbaar heeft die minstens even expressief is.

Een verdere interessante eigenschap van de abstracties die we in het model aantreffen, is dat ze direct de typefrequentie van de meer concrete patronen die er in voorkomen, weerspiegelen: de transitiefconstructie, bij voorbeeld, is representationeel sterk met een open werkwoordspositie omdat er veel verschillende werkwoorden in voorkomen, terwijl de veroorzaakte-verplaatsingsconstructies met slechts twee werkwoorden gezien wordt door het model, en daarom als twee werkwoordsspecifieke constructies wordt opgeslagen.

Als we het perspectief omdraaien (kijkend vanuit de woorden in constructies i.p.v. vanuit de constructies), zien we dat sommige woorden als onafhankelijke lexicale constructies geleerd worden, terwijl andere woorden voornamelijk als constituent van een grotere constructie worden geleerd. Woorden die naar entiteiten verwijzen (zelfstandig naamwoorden) worden typisch geleerd als onafhankelijke lexicale constructies. Voor andere woordsoorten nemen we meer variatie waar, zowel tussen woorden als tussen verschillende simulatierondes. Persoonlijk voornaamwoorden worden in veel en diverse contexten gebruikt, wat zou moeten leiden tot een verwerving als on-

afhankelijke lexicale constructies, maar ze worden ook zeer frequent *binnen* die constructies gebruikt, wat zou moeten leiden tot een opslag als deel van de grotere constructie. Voor deze voornaamwoorden, maar ook voor voorzetsels en werkwoorden, vinden we dat ze in sommige simulatierondes onafhankelijk verworven worden, terwijl ze in andere als deel van een groter patroon geleerd worden. Er zijn drie factoren te identificeren voor de neiging van een woord om onafhankelijk verworven te worden. Ten eerste: hoe meer verschillende elementen er op die plek voorkomen, hoe hoger de waarschijnlijkheid dat het woord onafhankelijk geleerd wordt. Ten tweede: hoe hoger de frequentie van dat woord op specifieke posities van grammaticale constructies, hoe lager de waarschijnlijkheid dat het onafhankelijk geleerd wordt. Ten derde: hoe meer verschillende constructies er zijn waarin dat woord voorkomt, hoe hoger de waarschijnlijkheid dat het onafhankelijk geleerd wordt.

Ook op andere aspecten van de representaties vinden we individuele variatie tussen de simulatierondes. De gemiddelde abstractie van de representatie varieert ook tussen verschillende simulatierondes. Dit is interessant, aangezien de diverse simulatierondes wel hetzelfde gedrag in productie en begrip vertonen – ze presteren even goed op de verschillende taken.

