Zijn sommige getalwoorden semantische virussen?

Bachelorscriptie taalwetenschap

Sebastiaan Jans
begeleider Rick Nouwen
Universiteit Utrecht

mei 2021

Samenvatting

Ik bekijk in dit eindwerkstuk woorden voor „vreemde” getallen, zoals 0, en rationale, irrationale, en negatieve getallen. De semantiek van deze getallen is minder uitgewerkt dan die van de „normale” natuurlijke getallen $\{1, 2, 3, \dots\}$. Specifiek bekijk ik of deze getallen te zien zijn als *virussen*, indringers in de taal die niet echt tot het taalsysteem behoren. Er vallen hier moeilijk eenduidige conclusies over te trekken omdat de virustheorie hiervoor nog te weinig is uitgewerkt. Aan de hand van de uitwerkingen van verschillende auteurs zouden verschillende conclusies getrokken worden. Ik geef een suggestie van een onderscheid tussen *virus* en *extragrammaticaal* die bij kan dragen aan verdere uitwerking van de virustheorie. Sommige soorten getallen lijken in sommige contexten virussen te zijn, of misschien eerder extragrammaticaal.

Inhoudsopgave

0 Inleiding

Getallen worden in zeer veel moderne grote talen gebruikt, vaak met het ons bekende decimale systeem (Comrie, 2013). Vooral over de natuurlijke getallen $\{1, 2, 3, \ldots\}$ is veel geschreven. De morfologie en semantiek van deze getallen is een centraal onderdeel van de formeel-semantische literatuur. Er zijn echter ook andere soorten getallen. Neem bijvoorbeeld 0: de hoeveelheid dingen als je niks hebt. Dit lijkt nu een best normaal concept, maar is in de westerse wereld relatief nog best nieuw met slechts enkele eeuwen (Kaplan, 1999). Ook hebben we natuurlijk de reële getallen, rationaal en irrationaal, die niet discrete objecten tellen, maar eerder dingen meten. Wat is dan de semantiek van die getallen? Er bestaan ook nog exotischere soorten getallen, zoals de complexe getallen of zelfs de quaternionen. Hier zouden we kunnen vragen wat de semantiek van die getallen is, maar eigenlijk is er nog een eerdere vraag: „Hebben deze getallen wel een semantiek?" Het spreekt bepaald niet voor zich dat deze getallen deel zijn van het taalsysteem. En mochten deze getallen niet deel zijn van het taalsysteem, zijn de reële getallen of 0 dat dan wel? Kunnen we dan een grens vaststellen?

Morzycki (2017) oppert dat 0 als in „Er staan nul emails in mijn inbox" wel eens semantisch gezien een speciaal geval zou kunnen zijn: een *virus*. Hij doet hetzelfde voor negatieve getallen. Een virus zou een fenomeen zijn dat gebruikt wordt in de taal, maar er niet echt helemaal thuishoort. Dit idee komt van Emonds (1986), die dit suggereert voor bepaalde regels in de syntaxis die niet echt echte regels lijken te zijn. Sobin (1997) werkt hiervan de details verder uit.

De hoofdvraag die ik probeer te beantwoorden is de volgende:

- Kunnen we „vreemde"[1] getallen analyseren als normale semantische fenomenen, of kunnen we ze beter zien als virussen, zoals Morzyki voorstelt voor 0 en negatieve getallen?

Ik behandel hierbij verschillende soorten getallen die potentieel virussen zijn afzonderlijk. Als 0 bijvoorbeeld een virus is betekent dat namelijk nog niet dat negatieve getallen dat ook zijn, of vice versa. Ook maakt het uit op welke manier de getallen gebruikt worden. Omdat de virustheorie nog weinig is uitgewerkt moet eerst nog een methodische vraag beantwoord worden:

[1] Morzycki (2017) gebruikt de term *weird*.

- Hoe kunnen we bepalen of een fenomeen een virus is?

Ik gebruik bewust de term *getalwoord*, en niet de gebruikelijke term *telwoord*, omdat die laatste term impliceert dat het alleen maar over tellen gaat, terwijl ik het juist veel heb over woorden die niet gebruikt worden om mee te tellen, maar om mee te meten of wiskundige uitspraken mee te doen. Het valt bijvoorbeeld al te betwisten of je kan tellen tot tweeëneenhalf, laat staan tot 6.28 of π.[2] Met de term *getalwoord* bedoel ik dus woorden voor getallen, wat voor soort getallen dat ook zijn. Ik gebruik ook wel *getaluitdrukking* als het morfologisch niet duidelijk om een enkel woord gaat, zoals bijvoorbeeld bij het uitdrukken van een complex getal $2i + 3.4$ als *twee i plus drie komma vier*.[3]

In Sectie 1 bespreek ik de virustheorie zoals voorgesteld in de (schaarse) literatuur. In Sectie 2 geef ik een beknopt overzicht van de semantiek van getalwoorden zoals die nu is en wat haar beperkingen zijn. Ik bespreek daar ook wat meer exotische soorten getallen dan waar meestal over geschreven wordt, zoals 0, reële getallen en negatieve getallen. In Sectie 3 ga ik de verschillende exotische soorten getalwoorden langs en beoordeel ik of het zinnig is om deze als virus te classificeren. Ik toets deze daarbij aan de virustheorie zoals beschreven in de literatuur, en mijn eigen inzichten hierover. Ik voeg ook nog mijn eigen suggestie voor een onderscheid tussen *virus* en simpelweg *extragrammaticaal* toe.

1 Virustheorie

In deze sectie illustreer ik wat een taalkundig *virus* precies inhoudt, zowel in de originele syntactische betekenis als in de nieuwere semantische zin.

1.1 Syntactische virussen

Emonds (1986) merkt op dat sommige structuren in het prestigedialect van (Amerikaans) Engels—vanaf nu afgekort als PE—zich niet als consequente grammaticaregels gedragen. Specifiek heeft Emonds het over de voorschriften van de handboeken *Harbrace College Handbook* (*HCH*) en *The Careful Writer* (*CW*). In (1) worden de nominatieve persoonlijk voornaamwoorden voorgeschreven. In het dagelijks natuurlijk taalgebruik (normaal Engels, NE) zijn de accusatieve opties echter gebruikelijk. De nominatieve variant in (1b) klinkt zelfs zo onnatuurlijk dat het enigszins komisch is.

[2] Dat men niet tot π kan tellen is duidelijk omdat π een irrationaal getal is. $2\frac{1}{2}$ en 6.28 zijn allebei rationale getallen $\left(6.28 = 6\frac{28}{100}\right)$, dus wat dat betreft verschillen ze niet. Echter, zoals ik bespreek in Sectie 2.2.2 stelt Nicolas (2015) dat bij getalwoorden als *tweeëneenhalf* nog geteld wordt: er worden helften geteld. Het lijkt mij lastiger vol te houden dat bij 6.28 achtentwintig honderdsten geteld worden, laat staan zeshonderachtentwintig hondersten.

[3] De Nederlandse conventie is om een komma te gebruiken om decimalen te noteren. Ik gebruik een punt om conflicten met andere wiskundige notatie te vermijden. Ook in alledaags Nederlands wordt de punt wel gebruikt, en zou $2i + 3.4$ dus ook als *twee i plus drie punt vier* uitgesproken kunnen worden.

(1) a. Our landlord and we /us very often disagree.
Onze huisbaas en wij.NOM /wij.ACC heel vaak het.oneens.zijn.
'We zijn het vaak oneens met onze huisbaas.'

b. It is I /me.
Het is ik.NOM /ik.ACC
'Ik ben het.'

Het lijkt hier dat er een sociolinguïstisch verschil is tussen PE en NE, namelijk dat NE in deze posities accusatief verwacht, terwijl PE nominatief verwacht. Dit klinkt redelijk, neem bijvoorbeeld Nederlands, waar de nominatief in vergelijkbare zinnen als (2a) en (2b) wel natuurlijk is.[4]

(2) a. Jij en ik/*mij zijn het vaak oneens.
b. Dat was ik/*mij.

Het normale gebruik wordt volgens Emonds informeel beschreven in regel (3). Deze regel staat in (1a) en (1b) geen nominatief toe. De nominatief zou dan wellicht ouder taalgebruik kunnen zijn. Emonds stelt echter dat dat niet het geval is. Hij stelt dat PE niet een normaal grammaticaal verschijnsel is: het houdt zich niet aan wat we begrijpen van de Universele Grammatica (modern vertegenwoordigd in bijvoorbeeld Chomsky, 2014). Emonds ziet dit als een extragrammaticaal fenomeen: de normale grammatica genereert de accusatief, zoals in NE, en die wordt vervolgens bewust naar de PE-versie „gecorrigeerd" op basis van wat de spreker zich herinnert uit handboeken als *CW* en *HCH*.

(3) De nominatieve persoonlijk voornaamwoorden (*I*, *we*, *he*, *she* en *they*) worden gebruikt als *noun phrase* (NP) als en alleen als de *phrase* een directe constituent van een zin (S) is die een vervoegd werkwoordelijk element bevat.

Op de precieze syntactische redenering zal ik hier niet in gaan, maar ook aan de hand van voorbeelden is aan te tonen dat PE door de elite-doelgroep zelf niet consequent gebruikt wordt. Ten eerste worden de regels lang niet altijd nageleefd door de doelgroep, wat we wel zouden verwachten als het echte regels zouden zijn. Ten tweede worden nominatiefvormen door de doelgroep vaak gebruikt waar het ook volgens de PE-regels niet zou moeten. In (4a) en (4b) zijn *they* en *he* lijdend voorwerp, maar staan ze zijn toch nominatief. Dit is *hypercorrectie*. Ook is inconsequent gebruik van PE niet nieuw: er zijn al voorbeelden van enkele eeuwen oud. Emonds (1986) citeert voorbeelden van onder andere Shakespeare.

[4] Het moet opgemerkt worden dat ook in het Nederlans wel een vergelijkbaar onderscheid bestaat. In vergelijkende zinnen als (i) is de nominatief *ik* de voorgeschreven standaard, maar komt de accusatief *mij* ook veel voor.

(i) Hij is groter dan ik/mij.

(4) a. *For they to be understood correct-ly, ...
voor zij.NOM.3PL om.te zijn begrepen correct-ADV
'(Voor hen) om correct begrepen te worden, ...'

b. *Who, if not he, could we have hired?
wie als niet hij.NOM konden we hebben ingehuurd
'Wie hadden we kunnen inhuren, als niet hij?'

De handboeken definiëren ook geen eenduidige regels, maar geven vooral voorbeelden van wat ze goed of fout vinden (omdat er geen eenduidige regel te definiëren is). Waar de PE-„regels" constructies genereren die zó onnatuurlijk zijn dat ze duidelijk niet als echt correct Engels gezien kunnen worden, zoals (5a) en (5b), geven de handboeken geen oplossing, maar zeggen ze simpelweg „vermijd de constructie".

(5) a. *Bill is quite confused; now he 's sure that John 's I.
Bill is behoorlijk verward nu hij is zeker dat John is ik.NOM
'Bill is behoorlijk verward; nu is hij er zeker van dat John mij is.'

b. *Othello is a wonderful role; I should be he in
Othello is een geweldige rol ik zou.moeten zijn hij.NOM in
your production. I 'm never he when we 're on tour.
jouw productie. I ben nooit hij.NOM wanneer we zijn op tournee
'Othello is een geweldige rol; ik zou hem moeten zijn in jouw productie. Ik ben hem nooit als we op tournee zijn.'

Emonds maakt hier een duidelijk sociolinguïstisch punt van. Hij vindt dat mechanismes die klasseverschillen in stand houden waar mogelijk geëlimineerd moeten worden. De virusregels van PE zijn volgens hem zulke mechanismes. De elite houdt opzettelijk onleerbare regels in stand. Ze kunnen dan zelf—met de nodige hulp van goed onderwijs en secretarissen of redacteuren—nog vaak wel redelijk die regels volgen, in ieder geval als ze schrijven. Voor mensen met minder middelen is het echter heel lastig om PE te verwerven.[5] De elite kan dan makkelijk iedereen die niet PE gebruikt onterecht wegzetten als „niet slim genoeg" en zo het klasseverschil in stand houden.

Emonds begrijpt dat het zeer lastig is om zulke schadelijke opvattingen weg te krijgen, maar roept toch stellig tot actie op, vooral van taalwetenschappers, die toch de enigen zullen zijn die zijn artikel lezen. Hij roept op om in ieder geval zelf niet de valse PE-regels te gebruiken en zeker niet bij te dragen aan het stigmatiseren van het gebruik van NE. Ook roept hij op voor onderwijs dat gelijke aandacht schenkt aan verschillende variëteiten van het Engels, die de verschillen ertussen expliciet behandelt als verschillen, zonder dat één variëteit beter is dan de ander.

Sobin (1997) voegt getalcongruentie in zogenaamde *expletive constructions* (expletiefconstructies) toe als virus. In (6a) zien we meervoudscongruentie van

[5] Met de huidige ontwikkelingen in automatische grammaticacontrole wordt dit verschil misschien wat kleiner, al blijft de vraag of we ons met kunstmatige hulpmiddelen aan kunstmatige, onnatuurlijke regels moeten willen houden.

het onderwerp *books* (boeken) en de persoonsvorm *are* (zijn). In de expletiefconstructie (6b) schrijft PE ook congruentie met het onderwerp voor, de persoonsvorm zou dus *are* moeten zijn. We zien hier echter vaak enkelvoudscongruentie, met de persoonsvorm *is* of *'s*. Sobin beargumenteert dat de voorgeschreven meervoudscongruentie tegen de syntactische theorie ingaat, en dat het alle door hem gegeven kenmerken van een virus heeft.

(6) a. Books are on the table.
boeken zijn op de tafel
'Boeken liggen/staan op (de) tafel.'

b. There *are* /-*'s* books on the table.
er zijn /is boeken op de tafel
'Er liggen/staan boeken op (de) tafel.'

Lasnik en Sobin (2000) voegen ook nog *whom* toe als virus. Dit zou de accusatief zijn van *who* (wie) volgens de PE-regels. Dit is echter niet acceptabel als accusatiefvorm in zinnen als (7).

(7) *Whom was it?
wie.ACC was het
'Wie was het?'

Net als Emonds erkent Sobin dat virussen duidelijk gedefiniëerd moeten worden, anders verwordt het tot een vage klasse van onverklaarbare fenomenen. Hij geeft daarom enkele concrete kenmerken, vooral geïllustreerd aan de hand van de „*... and I ...*"-regel. Om de kenmerken te verantwoorden gebruikt hij data van een eigen onderzoek naar de acceptatie van de constructies bij studenten, een groep die over het algemeen PE gebruikt, of probeert te gebruiken. Daarnaast gebruikt hij data van Quattlebaum (1994), die sprekers vroeg om een prestigevorm te kiezen uit twee vormen.

Lexicale Specificiteit De prestigeregels worden afhankelijk van specifieke woorden in de context meer of minder gebruikt. „*... and I ...*" wordt bijvoorbeeld aanzienlijk vaker gebruikt dan „*...and he ...*" in dezelfde constructie. We zien bijvoorbeeld de nominatief vaker in zinnen als (8a) dan in zinnen als (8b). Bij echte grammaticale regels verwachten we dat ze consequent toegepast worden ongeacht welke specifieke woorden voorkomen in de context.

(8) a. John and I/me went to the cinema last night.
John en ik.NOM/ik.ACC gingen naar de bioscoop laatste nacht
'John en ik zijn naar de bioscoop geweest gisteravond.'

b. John and he/him went to the cinema last night.
John en hij.NOM/hij.ACC gingen naar de bioscoop laatste nacht
'John en ik zijn naar de bioscoop geweest gisteravond.'

Directionaliteit en naastliggendheid De „*... and I ...*”-regel is sterker voor de conjuncten aan de rechterkant dan aan de linkerkant. In zinnen als (9a) wordt rechts van *and* vaker de nominatief geselecteerd in Quattlebaum (1994), terwijl in zinnen als (9b) links van *and* juist vaker de accusatief geselecteerd wordt, die hier in lijdendvoorwerppositie zowel normaal als PE is. Sobin noemt deze eigenschap *directionaliteit*.

(9) a. They fired the other employees and I/me.
zij.3PL ontsloegen de andere werknemers en ik.NOM/ik.ACC
‘Ze hebben mij en de andere werknemers ontslagen’

b. They fired I/me and the other employees.
zij.3PL ontsloegen ik.NOM/ik.ACC en de andere werknemers
‘Ze hebben mij en de andere werknemers ontslagen’

Ook wordt de prestigenominatief vaker geselecteerd als deze naast de *complementizer* ligt in de recursieve syntactische structuur (als ze dezelfde *ouder* hebben in de syntaxisboom). Zo wordt in zinnen als (10a) vaker de nominatief geselecteerd dan in zinnen als (10b). Dit noemt Sobin *adjacency* (naastliggendheid).

(10) a. She /her and NP ...
zij.1sg.NOM /zij.1sg.ACC en NP ...
‘Zij en ...’

b. NP and he /him ...
NP en hij.NOM /hij.ACC ...
‘...en hij ...’

Overextensie Net als Emonds (1986) vindt Sobin *overextensie* (of *hypercorrectie*) een essentieel kenmerk van een virus. We zien dit bijvoorbeeld in (11), waar de accusatief NE en PE is omdat het na een voorzetsel komt, maar vaak de nominatief geselecteerd wordt.

(11) ...between you and I/me ...
...tussen jou en ik.NOM/ik.ACC ...
‘...tussen jou en mij ...’

Onderextensie We zien bij virussen ook het tegenovergestelde effect van overextensie, namelijk *onderextensie*. Virusregels worden vaak niet toegepast waar in situaties waar ze volgens de regel wel toegepast zouden moeten worden. Zo wordt bijvoorbeeld in (1a) en (1b) vaak de accusatief gebruikt, ook door sprekers van PE.

Virussen worden dus algeheel inconsequent toegepast: ze worden wel gebruikt terwijl ze niet gebruikt horen te worden (*overextensie*) en ze worden niet gebruikt terwijl ze wel gebruikt horen te worden (*onderextensie*).

Nonlokaliteit: ongevoeligheid voor non-lexicale hiërarchische constituenten Virusregels worden soms over de grenzen van constituënten uitgevoerd, wat we van normale grammaticale regels niet verwachten. We zien bijvoorbeeld nominatiefvormen in zinnen als (12a). Dit is een toepassing van de „*… and I …*"-regel die over de grens van een *small clause* (SC) heen gaat. Pas als de *head* lexicaal wordt (expliciet wordt uitgesproken), zoals in (12b) wordt het ook voor een PE-spreker duidelijk dat een nominatief hier niet past.

(12) a. For Mary to be the winner and [$_{\text{SC}}$ I the loser] is
voor Mary te zijn de winnaar en [$_{\text{SC}}$ ik.NOM de verliezer] is
unfair.
oneerlijk
'Het is oneerlijk als Mary wint en ik verlies'

b. For Mary to be the winner and [$_{\text{IP}}$ me /??I to be
voor Mary te zijn de winnaar en [$_{\text{IP}}$ ik.ACC /??ik.NOM te zijn
the loser] is unfair.
de verliezer] is oneerlijk
'Het is oneerlijk als Mary wint en ik verlies'

De regels van PE worden door *HCH* en *CW* niet duidelijk gedefiniëerd, maar aan de hand van wat we hier gezien hebben kunnen we wel proberen een regel op te stellen die het daadwerkelijke gebruik van PE-sprekers reflecteert. Sobin geeft daarvoor de regel in (13). Deze regel kan ook toegepast worden op andere persoonlijk voornaamwoorden, maar dat gebeurt minder dan met *I*: lexicale specificiteit. De regel kan ook links van *and*, of niet direct ernaast, toegepast worden, maar dat gebeurt minder door directionaliteit en naastliggendheid.

(13) Als het eerste persoon enkelvoud persoonlijk voornaamwoord rechts naast het voegwoord *and* staat (direct na *and* komt), krijgt deze de nominatiefvorm *I*.

Emonds en Sobin concluderen dat PE-sprekers eerst de normale grammaticale zin genereren, met bijvoorbeeld „*… and me …*", en dan voordat ze die uitspreken nog even—buiten de grammatica om—de lineaire structuur van de zin (de volgorde waarin woorden uitgesproken worden, geen rekening houdend met hiërarchische structuur) checken of er een „*… and me …*" in voorkomt en dat „corrigeren" voordat ze de zin uitspreken.

1.2 Semantische virussen

Morzycki (2017) merkt een aantal semantische fenomenen op die ook wat van virussen weg hebben. Hij stelt de vraag in hoeverre we moeten proberen om dit soort structuren te analyseren als deel van het menselijke taalvermogen. Voor sommige van deze fenomenen zouden we behoorlijke veranderingen moeten aanbrengen in de huidige analyse, terwijl ze eigenlijk weinig voorkomen. Het is dan wellicht beter om deze semantische fenomenen ook te zien als virussen, en de huidige analyse te behouden. Daarvoor zou de virustheorie dus uitgebreid moeten worden naar de semantiek.

Sobins kenmerken voor virussen zijn duidelijker dan de beschrijving van Emonds, maar ze zijn erg specifiek. Directionaliteit, naastliggendheid en non-lokaliteit zijn in ieder geval alleen toepasbaar op syntactische regels. Morzycki herformuleert de kenmerken van virussen die Sobin (1997) daarvoor naar een algemere versie, die niet over specifieke syntactische constructies gaan. Hij vermeldt hierbij niet dat deze kenmerken aanzienlijk anders zijn dan door Sobin voorgesteld werd.

1. Ze komen vooral voor in de prestigevariant van de taal.
2. Ze zijn moeilijk te verwerven en worden relatief laat verworven.
3. Ze kunnen wat Emonds (1986) „tutorial support" noemt vereisen: expliciete instructie.
4. Sprekers beheersen ze niet goed, ze worden vaak „fout" gebruikt of gehypercorrigeerd.
5. Ze worden bewust toegepast.

Aan de hand van deze definitie bespreekt Morzycki de volgende fenomenen als potentiële virussen.

Respectievelijk en vergelijkbaren Het Engelse *respectively*, of het Nederlandse equivalent *respectievelijk*, zoals in zinnen als (14a), is problematisch. De gangbare analyses (Morzycki gaat vooral uit van Gawron en Kehler, 2004) gaan ervan uit dat het *respectievelijk* de argumenten in de lineaire structuur van een zin opzoekt. Dit gaat in tegen de aannames van UG, er zou namelijk alleen gezocht kunnen worden in de recursieve hiërarchische structuur. Het feit dat volgens de regels van *respectievelijk* de zin in (14b) iets anders betekent dan die in (14a) impliceert dat *Alpo en Whiskas* iets anders betekent dan *Whiskas en Alpo*.

(14) a. Fido en Felix eten respectievelijk Alpo en Whiskas.
b. Fido en Felix eten respectievelijk Whiskas en Alpo.

Het is niet ondenkbaar om zoiets uit te werken, maar het geeft een enkel bijwoord wel heel veel kracht, en een soort kracht dat we in andere bijwoorden eigenlijk niet zien. Als dit normaal zou zijn zouden volgens Morzycki ook dingen als een bijwoord dat de laatste twee werkwoorden verwisselt normaal zijn. Hij kiest er daarom liever voor om dit niet als normaal te zien, maar het als virus te beschouwen. Of dit expliciete instructie vereist is niet helemaal duidelijk, maar het heeft volgens Morzycki wel duidelijk alle andere hierboven genoemde kenmerken van virussen. Het bijvoeglijk naamwoord *respective* en *vice versa* werken op een vergelijkbare manier.

***Nul* en negatieve getallen** Bylinina en Nouwen (2018) geven een semantiek van het gebruik van *nul* als telwoord in zinnen als (15).

(15) Er staan nul emails in m'n inbox.

Volgens hun analyse verschilt *nul* feitelijk niet van andere telwoorden als *één*, *drie*, *vijfhonderdzevenentachtig*. Ze nemen aan dat er lege pluraliteiten bestaan, en dat is mogelijk problematisch. Het is wellicht wenselijker om dit als virus te bestempelen dan om hun semantiek aan te nemen. Morzycki stelt dat negatieve getallen dan wellicht ook virussen zijn. Dat impliceert dat dat vergelijkbaar is met het voorstel van Bylinina en Nouwen over *nul*, maar het moet opgemerkt worden dat dit niet echt vergelijkbaar is. Ik ga uitgebreider in op de semantiek die Bylinina en Nouwen (2018) voorstellen in Sectie 2.2.1 en in Sectie 3 bespreek ik of deze fenomenen virussen zijn.

Factor phrases Wellicht de meest onzekere virussugestie die Morzycki maakt gaat over *factor phrases*, zoals in (16a) en (16b).

(16) a. Ik verdien nog niet half zo veel als zij.
b. Claas skeelerde vandaag twee keer harder als/dan anders.

Omdat analyses van *degrees* (graden) veel uiteenlopen, en ook erg lijken te verschillen tussen talen, is het mogelijk beter om in plaats van een complexe semantiek hiervan uit te werken het als virus te beschouwen. Morzycki erkent wel dat dit misschien wat ver gaat. Ik bespreek factor phrases niet verder, maar in Sectie 3.5 betrek ik het argument dat Morzycki maakt voor de virusstatus van factor phrases op de radicale suggestie dat *graden* (*degrees*) virussen zouden zijn.

1.3 Andersoortige virussen

Het is ook goed om op te merken dat de algemenere definitie van een virus van Morzycki (2017) ruimte biedt voor andere fenomenen die wellicht ook zinnig geanalyseerd kunnen worden als virussen. Neem bijvoorbeeld het Nederlandse onderscheid tussen *hen* en *hun*. Volgens de voorschriften van de Algemene Nederlandse Spraakkunst (ANS, te lezen in „Gebruik van *hen* en *hun*") zou *hen* gebruikt moeten worden als lijdend en oorzakelijk voorwerp en na voorzetsels, en *hun* als meewerkend voorwerp (zonder voorzetsel) en als ondervindend voorwerp.[6] In (17) zijn de persoonlijk voornaamwoorden die niet aan deze regel voldoen gemarkeerd met een #-teken.

(17) a. Hij heeft hen /#hun geslagen.
b. Hij heeft #hen /hun een boek gegeven.
c. Hij heeft een boek aan hen /#hun gegeven.

[6]Ik heb het hier—voor de duidelijkheid—niet over de bezittelijke *hun*, of *hun* als onderwerp.

Dit zou een verschil zijn tussen datief (*hun*) en accusatief (*hen*). Ramaker (g.d.) laat echter zien dat dat verschil in modern Nederlands niet bestaat. Het is in 1652 verzonnen door Christiaan van Heule (in zijn boek *De Nederduytsche Grammatica ofte Spraec-konst*), die het Latijnse systeem van naamvallen probeerde toepassen op het Nederlands, waar dat niet past. Zelfs de ANS erkent dat het onderscheid niet gangbaar is in spreektaal, wat dit geval iets anders maakt dan de „*… and I …*"-regel. Toch waart de „regel" nog rond onder de sprekers en is er nog steeds een kans dat je „verbeterd" wordt als je het „fout" doet, hoewel het weinig zal voorkomen.

Uit eigen ervaring zou ik zeggen dat het onderscheid tussen *hen* en *hun* redelijk aan Morzycki's eisen voldoet. Voor zover het voorkomt in spraak komt het alleen voor in de prestigevariant. Het is moeilijk te verwerven: ik zou nog steeds de regel moeten opzoeken als ik hem toe zou willen passen. Ik heb de regel aangeleerd gekregen op de middelbare school, niet van huis uit: *tutorial support*. Onderextensie en overextensie komen volgens mij ook vaak voor. Ik zou in ieder geval niet opkijken als iemand de met een # gemarkeerde vormen in (17) gebruikt. Als de regel toegepast wordt lijkt me dat zeker bewust. Om zeker te zijn moet dit natuurlijk empirisch onderzocht worden, maar het lijkt mij in ieder geval aannemelijk dat het onderscheid een virus is volgens Morzycki.

Het voldoet echter duidelijk niet aan de eisen van Sobin (1997). Er is hier waarschijnlijk sprake van zowel overextensie als onderextensie, en lexicale specificiteit is niet ondenkbaar, dat zou empirisch onderzocht moeten worden. Directionaliteit, naastliggendheid en nonlokaliteit daarentegen zijn duidelijk niet van toepassing hier. Deze kenmerken gaan over syntactische structuur, terwijl dit niet over structuur gaat, het gaat simpelweg om het al dan niet bestaan van een naamval in een taal. Toch lijkt het mij zinnig om dit wel als virus te beschouwen. Morzycki suggereert ook dat virussen wellicht ook in de fonologie gevonden zouden kunnen worden, dit zou met deze algemenere kenmerken ook kunnen.

Nog een interessant idee dat Morzycki bijdraagt aan de virustheorie is het idee dat virussen na verloop van tijd via grammaticalisatie opgenomen kunnen worden in de normale taal. Dit zou niet kunnen met de regels die Emonds (1986) beschrijft als „onleerbaar", maar misschien met andere regels—zoals bijvoorbeeld de legepluraliteitssemantiek van *nul*—wel. Omdat er ook veel talen zijn die niet een uitgebreider getallensysteem hebben dan tellen tot twee of drie (Calude, 2021), is het ook mogelijk dat getallen überhaupt als virus origineel binnengedrongen zijn in de taal. In de vergelijking met biologische virussen zou dit overeen kunnen komen met het opnemen van viraal DNA (deoxyribonucleic acid) in het genoom van geïnfecteerde organismes.

1.4 Virussen samengevat

Emonds (1986) behandelt syntactische virussen, met veel focus op de sociolinguïstische aspecten. Sobin (1997) werkt de syntactische details meer uit. Morzycki (2017) neemt virussen juist weer breder dan Sobin, en geeft specifiek voorstellen van mogelijke virussen in de semantiek, maar focust net als Sobin vooral veel op theorie en weinig op de sociale aspecten. Lasnik en Sobin (2000) geven een virus (*whom*) dat waarschijnlijk een archaïsch overblijfsel is. Morzycki stelt voor dat juist het tegenovergestelde misschien ook mogelijk is: het opnemen van virussen in de normale taal.

De beschrijvingen van virussen lopen dus nogal uiteen, en de virustheorie zal voor een algemener gebruik duidelijker en eenduidiger gedefinieerd moeten worden. In Sectie 3.6 draag ik daarvoor mijn eigen suggestie van *extragrammaticaliteit* bij.

2 Getalwoordensemantiek

In dit hoofdstuk geef ik een (kort) overzicht van de hedendaagse analyse van de semantiek van getalwoorden. Ik begin bij de strict-positieve natuurlijke getallen ($\mathbb{N}^+ = \mathbb{N} - \{0\}$), waar het meest over geschreven is. Ik ga hierbij uit van het verschil tussen *tellen* en *meten* zoals beschreven door Rothstein (2017). Vervolgens behandel ik de semantiek van de getallen die buiten $\mathbb{N}^+$ vallen, dus de getallen die ofwel niet geheel, ofwel niet positief zijn, of beide.

Het is relevant om op te merken dat de morfologie van verschillende soorten getalwoorden of getaluitdrukkingen ook uit te werken is (zie bijvoorbeeld van Drie, 2015), maar ik ga hier helaas niet verder op in.

2.1 Traditionele telwoordensemantiek

Hier behandel ik traditionele telwoordensemantiek van de positieve natuurlijke getallen, te beginnen met *tellen*, maar ook hun gebruik voor *meten*.

2.1.1 Tellen

Meervoudsbetekenis met pluraliteiten Voor de semantiek van tellen is de betekenis van meervouden erg belangrijk. Ik geef daarom eerst een korte, informele uitleg van *pluraliteiten*. In de algemene ontologie van taal nemen we aan dat er *entiteiten* zijn: mensen, dieren, dingen. Soms hebben we het over meerdere entiteiten tegelijk, dit is over het algemeen wanneer we een meervoudsvorm gebruiken. Link (1983) stelt dat zo een verzameling van entiteiten ook een entiteit op zich vormt: een *pluraliteit*. Met de *join*-operatie $\sqcup$ kunnen entiteiten (en ook pluraliteiten) samengevoegd worden in een pluraliteit. Zo is de persoon *Ada* een entiteit, en de persoon *Bob* ook, en is $Ada \sqcup Bob$ de pluraliteit die zowel *Ada* als *Bob* bevat, en niets meer. Met $\sqsubseteq$ geven we *inclusie* aan: dat een entiteit lid is van een pluraliteit. We hebben bijvoorbeeld $Ada \sqsubseteq (Ada \sqcup Bob)$ en $(Ada \sqcup Bob) \sqsubseteq (Ada \sqcup Bob \sqcup Carel)$, maar ook $Ada \sqsubseteq Ada$: de inclusierelatie is *reflexief*. De $\sqsubseteq$-operatie is een partiële ordening op de verzameling pluraliteiten en maakt deze daarmee een *join semi-lattice*.[7]

[7] Mijn notatie van pluraliteiten is grotendeels ontleend aan Nouwen (2016). Voor de inclusierelatie gebruik ik wel $\sqsubseteq$ in plaats van $\sqsubset$, zoals Rothstein (2017) doet, om de reflexiviteit te benadrukken.

Met de *-operatie laten we predicaten werken over pluraliteiten. Dat is nuttig om de semantiek van zinnen als (18a) uit te drukken. Dit valt zonder pluraliteiten uit te drukken als in (18b), maar dat reflecteert de structuur van de zin en de ontologie niet goed. Omdat de atomaire entiteiten *Ada* en *Bob* ook pluraliteiten zijn (met kardinaliteit 1) kunnen we dit ook uitdrukken als in (18c) met het pluraliteitspredicaat $^*Boos$. We geven dan met datzelfde $^*Boos$-predicaat de equivalente semantiek in (18d).

(18) a. Ada en Bob zijn boos.
b. $Boos(Ada) \wedge Boos(Bob)$
c. $^*Boos(Ada) \wedge {}^*Boos(Bob)$
d. $^*Boos(Ada \sqcup Bob)$

De semantiek van tellen Rothstein (2017) geeft een uitwerking van de semantiek van getalwoorden (*numericals* in haar termen). Ik begin bij de getalwoorden die gebruikt worden om te tellen: de woorden waar de klassieke term *telwoorden* op slaat.

Rothstein merkt op dat telwoorden ambigu zijn, met twee mogelijke betekenissen. Vergelijk (19a), waar *vier* gebruikt wordt om het aantal brieven te tellen, met (19b), waar vier als entiteit op zich gebruikt wordt, om een predicaat op toe te passen.

(19) a. Ik heb vier brieven ontvangen.
b. Vier is een even getal.

In (19a) zien we dat *vier* fungeert als een soort bijvoeglijk naamwoord, terwijl het in (19b) als zelfstandig naamwoord fungeert. Voor het gebruik in (19a) geeft Rothstein de semantiek in (20a), met *vier* als predicaat van type $\langle e, t\rangle$. Hier geeft $|x|$ de kardinaliteit van een pluraliteit x: het aantal atomaire entiteiten dat het bevat.

Voor *vier* als entiteit op zich, zoals in (19b) geeft ze de semantiek in (20b), met het speciale getaltype n. De $^\cap$-operator verlaagt getallen van type $\langle e, t\rangle$ naar dit type n. De $^\cup$-operator doet het tegenovergestelde.

(20) a. $[\![\, vier_{\langle e,t\rangle} \,]\!] = \lambda x\,[\,|x| = 4\,]$
b. $n = {}^\cap\lambda x\,[\,|x| = n\,]$

Dit type n kan gebruikt worden om eigenschappen van getallen te definiëren. Zo kan de eigenschap *even* gedefiniëerd worden als in (21a) en *groter dan* als in (21b).

(21) a. $[\![\, even_{\langle n,t\rangle} \,]\!] = \lambda x\,[\,x \mod 2 = 0\,]$
b. $[\![\, groter\ dan_{\langle n,\langle n,t\rangle\rangle} \,]\!] = \lambda y \lambda x\,[\,x > y\,]$

2.1.2 Meten

Rothstein ziet meten als een wezenlijk ander ding dan tellen. De type-n-eigenschappen kunnen daarbij gebruikt worden. Ze illustreert dit aan de hand van de uitdrukking *twee glazen water*.

(22) a. Marie, neem twee glazen water mee voor onze gasten!
b. Voeg twee glazen water toe aan de soep!

In (22a) zien we een *individuërende* lezing, waar het duidelijk om twee individuële onderscheidbare glazen gaat, die water bevatten. In (22b) hoeft er echter geen sprake te zijn van glazen, het gaat hier om het meten van een bepaald volume water wat ongeveer gelijk staat aan het volume van een glas, zeg ongeveer 200 milliliter. Als die 200 milliliter wordt afgemeten met een maatbeker wordt de instructie in (22b) nog steeds succesvol gevolgd. Als Marie echter met de instructie in (22a) aan komt zetten met een maatbeker die tot 400 milliliter gevuld is, in de instructie niet juist uitgevoerd.

Dit verschil werkt Rothstein als volgt uit. De zin in (22a) krijgt de telsemantiek als in (20a), omdat het gaat over individuele glazen, dit zijn gewoon entiteiten. De zin in (22b) daarentegen gaat niet over entiteiten. Deze gebruikt een meetpredicaat $\mu_{\textsc{grootheid, eenheid}}$ zoals in (23a) waarvan de grootheid en eenheid contextueel bepaald worden om de semantiek in (23b) te verkrijgen.[8]

(23) a. $\lambda n.\lambda x\,[\,\mu_{\textsc{grootheid, eenheid}}(x) = n\,]$
b. $[\![\, \textit{drie glazen water} \,]\!] = \lambda x\,[\,Water(x) \wedge \mu_{\textsc{volume, glazen}}(x) = 3\,]$

2.2 De semantiek van potentiële virusgetalwoorden

Hier behandel ik de semantiek van de verschillende getallen buiten $\mathbb{N}^+$ die virussen zouden kunnen zijn.

2.2.1 Nul

Bylinina en Nouwen (2018) stellen een uitbreiding van het telsysteem voor om *nul* erin op te nemen, zoals het gebruikt wordt in (15). Ze nemen 0 dus niet als wiskundig getal, maar echt als een discreet aantal atomaire entiteiten waarmee geteld kan worden, als kardinaliteit, net als 1 en 2. Ze beargumenteren dat *nul* ook geen kwantor is—het is niet hetzelfde als *geen* (of in dit geval het Engelse *no*)—aan de hand van verschillen in distributie en het al dan niet toestaan van *negative polarity items* (*NPI licensing*).

Bylinina en Nouwen stellen voor om de pluralisatieoperatie * uit te breiden naar een $^\times$-operator die ook lege pluraliteiten genereert. De lege pluraliteit noteren ze als $\perp$.

Een formelere definitie van de verschillende pluralisatieoperatoren vereist eerst een definitie van de *som*, die van een verzameling entiteiten een pluraliteit maakt. De som $\bigsqcup X$ van een verzameling entiteiten X geeft de pluraliteit die

[8] Notatie naar Snyder (2020).

bestaat uit alle elementen van X, en—omdat die entiteiten ook pluraliteiten kunnen zijn—de entiteiten die deel zijn van die elementen. De pluralisatieoperator * is dan zoals in (24a). Het geeft uit een verzameling Z alle pluraliteiten die gemaakt worden uit elementen van Z. Dit werkt op predicaten omdat deze wiskundig gedefiniëerd zijn als de verzameling van objecten waarvoor ze gelden. Het predicaat *blauw* is bijvoorbeeld de verzameling van alle blauwe dingen. Het enige wat anders is aan de nieuwe $^\times$-operator—gedefinieerd in (24b)—is dat deze pluraliteiten ook uit een lege deelverzameling van Z (simpelweg de lege verzameling) kunnen komen, en dus leeg kunnen zijn: $\bigsqcup \varnothing = \bot$.

(24) a. $^*Z = \{ \bigsqcup X \mid X \subseteq Z \wedge X \neq \varnothing \}$
b. $^\times Z = \{ \bigsqcup X \mid X \subseteq Z \}$

Met deze aanname kunnen we *nul emails* simpelweg representeren als in (25). De zin in (15) krijgt dan de analyse in (26)

(25) $[\![\, \textit{nul emails} \,]\!] = \lambda x \, [\, ^\times Email(x) \wedge |x| = 0 \,]$

(26) $\exists x \, [\, ^\times Email(x) \wedge {}^\times InMijnInbox(x) \wedge |x| = 0 \,]$

Dit werkt, maar levert nog wel een probleem op in normale meervouden zonder getallen. Neem bijvoorbeeld (27a). De analyse in (27b) geeft duidelijk de juiste waarheidsvoorwaarden.[9] Het geeft aan dat er een pluraliteit van typfouten bestaat die in de tekst staat. Echter, volgens (27c) kan die pluraliteit ook leeg zijn, en dus geen typfouten bevatten. Dan betekent (27a) dus eigenlijk „Er staan nul of meer typfouten in deze scriptie". Dat kan niet onwaar zijn: het is een tautologie. Het is ook duidelijk niet wat bedoeld wordt.

(27) a. Er staan typfouten in deze scriptie.
b. $\exists x \, [\, ^*Typfout(x) \wedge {}^*InDezeScriptie(x) \,]$
c. $\exists x \, [\, ^\times Typfout(x) \wedge {}^\times InDezeScriptie(x) \,]$

Dit probleem lossen Bylinina en Nouwen op door te stellen dat *nul* als enige telwoord een *precies*-lezing afdwingt. Ze introduceren hiervoor ook een nieuwe variant op de existentiële kwantor, **E**, die niet-leegheid aangeeft, gedefinieerd als in (28). Dit is mogelijk minder problematisch dan de aanname van lege pluraliteiten, al doet het misschien wel wat af aan het idee dat 0 net als alle andere telwoorden is.

(28) $\mathbf{E} \, [\, \varphi \,] \Leftrightarrow \exists x \, [\, |x| > 0 \wedge \varphi \,]$

2.2.2 Reële getallen: rationaal en irrationaal

We hebben nu de semantiek van gehele getallen gezien. We zien in dagelijks taalgebruik echter ook getallen die preciezer zijn dan gehelen: de reële getallen

[9] De details van de semantiek van in een scriptie staan vergeten we hier even, dit is niet relevant.

($\mathbb{R}$), bestaande uit de rationale getallen ($\mathbb{Q}$) en de irrationale getallen ($\mathbb{R} - \mathbb{Q}$).[10] Of het onderscheid tussen rationale en reële getallen belangrijk is in de taal, is niet heel duidelijk. Veel auteurs (bijvoorbeeld Wągiel, 2018, Snyder, 2020) hebben het simpelweg over reële getallen en benoemen het onderscheid niet. Ik zal nu proberen een schets te geven van wat mogelijk is in de semantische analyse van deze reële, al dan niet rationale getallen.

Liebesman (2015) beweert dat we niet tellen met identiteit. Sterker nog, hij beweert dat we nooit echt tellen, en dat alle getaluitdrukkingen eigenlijk vormen van meten zijn. Nicolas (2015) geeft daarvoor tegenargumenten, en geeft daarbij ook een analyse van rationale uitdrukkingen als *anderhalf*. Hij ziet dit als meten, maar stelt dat hierbij ook nog geteld wordt: er worden gehelen en helften geteld.

Snyder (2020) geeft de semantiek in (29) voor voor *half/halve*, in navolging van Ionin, Matushansky en Ruys (2006).

(29) $[\![\, half/halve \,]\!] = \lambda P.\lambda x.\exists y.$
   a. $\wedge \; P(y) \wedge \forall z \, [\, P(z) \rightarrow z \sqsubseteq y \,]$
   b. $\wedge \; \exists s \; [\, \Pi(S)(y) \wedge |S| = 2 \wedge x \in S$
   c. $\wedge \; \exists \mu \, [\, \mu \in M \wedge \forall s_1, s_2 \, [\, (s_1 \in S \wedge s_2 \in S) \rightarrow (\mu(s_1) = \mu(s_2)) \,]\,]\,]$

Regel (23a) geeft aan dat *half* het argument x verdeelt in maximale delen. Regel (23b) geeft aan dat een verzameling S de partitie is van die twee delen. Volgens de derde regel (23c) hebben deze delen dezelfde waarde met een contextueel meetpredicaat μ.

Hij stelt ook dat dit makkelijk uit te breiden is naar andere fracties van de vorm $\frac{1}{m}$ met n een natuurlijk getal. Dit zou kunnen zoals in (30)

(30) $[\![\, \text{m-}de \,]\!] = \lambda P.\lambda m.\lambda x.\exists y.$
   a. $\wedge \; P(y) \wedge \forall z \, [\, P(z) \rightarrow z \sqsubseteq y \,]$
   b. $\wedge \; \exists s [\, \Pi(S)(y) \wedge |S| = m \wedge x \in S$
   c. $\wedge \; \exists \mu \, [\, \mu \in M \wedge \forall s_1, s_2 \, [\, (s_1 \in S \wedge s_2 \in S) \rightarrow (\mu(s_1) = \mu(s_2)) \,]\,]\,]$

Dit combineert met de interpretatie van telwoorden als in (20a) naar een algemene semantiek voor rationale getallen van de vorm $\frac{n}{m}$. Zo is vier vijfde $\left(\frac{4}{5}\right)$ bijvoorbeeld zoals in (31).

(31) $[\![\, vier \; vijfde \,]\!] = [\![\, vier \,]\!] \, ([\![\, vijfde \,]\!])$
   a. $= \big(\big| \lambda P.\lambda x.\exists y.$
   b. $\wedge \; P(y) \wedge \forall z \, [\, P(z) \rightarrow z \sqsubseteq y \,]$

[10] Rationale getallen zijn wiskundig gezien getallen die uit te drukken zijn als een deling (ratio) $\frac{n}{m}$ van twee gehele getallen n en m, waarbij m niet 0 is. Ook kommagetallen met eindige decimalen, of oneindige decimalen die een herhalend patroon vormen, zijn rationaal. Bijvoorbeeld 6.28 $\left(= 6\frac{28}{100}\right)$ en 0.666... $\left(= \frac{2}{3}\right)$. De reële getallen zijn alle getallen die zich op de getallenlijn uit te drukken zijn. Dit zijn de rationale getallen, maar ook de irrationale getallen, die niet als deling uit te drukken zijn. Deze hebben oneindige decimalen die geen patroon vormen. Bekende voorbeelden zijn $\pi = 3.1415\ldots$ en $\sqrt{2} = 1.4142\ldots$.

c. $\wedge\ \exists s[\,\Pi(S)(y) \wedge |S| = 5 \wedge x \in S$

d. $\wedge\ \exists\mu\,[\,\mu \in M \wedge \forall s_1, s_2\,[\,(s_1 \in S \wedge s_2 \in S) \rightarrow (\mu(s_1) = \mu(s_2))\,]\,]\,]\big|$

e. $= 4\big)$

Dat rationale getallen in de wiskunde uitgedrukt worden als fracties/ratios, betekent niet dat ze in het taalsysteem ook altijd zo werken. Misschien geldt dit voor bijvoorbeeld *twee derde* wel, maar voor 6.28 waarschijnlijk niet. Ook de morfologie en semantiek van kommagetallen die met een expliciete „komma" uitgesproken worden zal uit te werken zijn, maar zal waarschijnlijk anders zijn dan Snyders analyse van rationale getallen.

Voor irrationale getallen werkt Snyders analyse in ieder geval niet. Toch is het denkbaar—hoewel twijfelachtig—dat deze deel zouden zijn van het taalsysteem voor meetuitdrukkingen. Een zin als (32a) is duidelijk wiskundig van aard. Het is in feite een vertaling van de (niet geheel correcte) wiskundige uitdrukking $\big(x^2 = 2\big) \therefore \big(x = \sqrt{2}\big)$. Het koppelwerkwoord *is* fungeert hier als vertaling van de gelijkheidsrelatie (=). De variabele x is een abstract wiskundig object, het wordt niet geteld of gemeten. In (32b) zien we echter iets anders. Ook deze uitspraak vereist wat wiskundige kennis: men moet bekend zijn met de stelling van Pythagoras. Toch is dit niet puur wiskundig, het gaat hier namelijk om het meten van een afstand in de echte wereld, uit te drukken met het meetpredicaat als $\mu_{\text{AFSTAND, KILOMETERS}}\big(\sqrt{2}\big)$. Een koppelwerkwoord is hier dan ook niet nodig. In (32c) zien we een vergelijkbaar gebruik van π.

(32) a. We hebben x kwadraat is 2 $\big(x^2 = 2\big)$, dus x is wortel twee $\big(x = \sqrt{2}\big)$.

b. Het is een kilometer naar het oosten en een kilometer naar het noorden. Als je door dat weiland afsnijdt fiets je dus wortel twee $\big(\sqrt{2}\big)$ kilometer.

c. Deze rondweg heeft een diameter van een kilometer, dus als je rondom wil lopen moet je π kilometer lopen.

2.2.3 Negatieve getallen

De semantiek van negatieve getallen is onduidelijk. Morzycki (2017) geeft de voorbeelden in (33a) en (33b).

(33) a. ??Het werd min tien graden warmer.

b. ??Min zes mensen arriveerden.

(33a) is vreemd, maar heeft nog een redelijke interpretatie—„Het werd tien graden kouder."—maar (33b) kan niet betekenen dat er zes mensen weggingen. Morzycki noemt de voorbeelden in (33) nerderige halfgrappen. Tellen met negatieve getallen lijkt dus zeker niet plausibel. Zelf hoorde ik laatst (34), waar spreker B er niet in slaagt een zinnig gebruik van een negatief getal te produceren.

(34) A: Wordt het niet eens tijd voor een VVD zónder Rutte?
B: ??Dat wordt het al min vier jaar lang tijd.

Voor meten is misschien nog wel enige ruimte, zie bijvoorbeeld (35a). Een medewerker van NASA (National Aeronautics and Space Administration) heeft het hier over het moment 20 minuten voor de lancering van Marsrover Perseverance. Dit is in feite het uitspreken van de wiskundige uitdrukking $\ell - 20$. Een natuurlijker voorbeeld is (35b), waar een negatief getal gebruikt wordt om temperatuur te meten op de intervalschaal van graden celsius.[11] Omdat het hier om een intervalschaal gaat kan wel beargumenteerd worden dat dit niet echt negatieve betekenis is. Ook moet opgemerkt worden dat dit alleen werkt met een koppelwerkwoord: (33a) werkt niet. Dit koppelwerkwoord kan dus weer een soort vertaling van de gelijkheidsrelatie zijn, zoals in (32a).

(35) a. It started with an earthquake in this very room on launch day at ℓ minus twenty minutes.
'Het begon met een aardbeving in deze ruimte op de dag van de lancering op ℓ min twintig minuten.'

b. Het is min vijf graden buiten.

Het lijkt erop dat negatieve getallen niet een productief deel uitmaken van het taalsysteem. Er bestaan ook andere getallen die nog wiskundiger van aard zijn, zoals de complexe getallen. Omdat hiervoor geen semantiek uitgewerkt is behandel ik deze in Sectie 3.4.

3 Zijn sommige getalwoorden virussen?

In deze sectie bespreek ik de verschillende soorten getalwoorden die potentieel als virus geanalyseerd kunnen worden. Zoals besproken in Sectie 1 werken de definities van virussen die Emonds (1986) en Sobin (1997) geven niet goed voor semantische virussen. Om te analyseren of verschillende soorten getaluitdrukkingen virussen zijn toets ik ze daarom aan de algemenere kenmerken die Morzycki (2017) geeft. Aan het einde van deze sectie opper ik nog een mogelijk zinnig onderscheid: virus versus *extragrammaticaal*.

3.1 Nul

Morzycki's belangrijkste argument voor het analyseren van *nul* als virus is dat het aannemen dat lege pluraliteiten bestaan om *nul* te verklaren, zoals Bylinina en Nouwen (2018) doen, veel kracht geeft aan een fenomeen dat eigenlijk weinig voorkomt, net als voor *respectievelijk*. Als lege pluraliteiten normaal zouden zijn zouden we ze ook tegen moeten komen in andere talen, zelfs talen die geen concept van *nul* hebben, en dat is volgens Morzycki niet het geval.

[11] Een intervalschaal is een schaal waarop het nulpunt niet echt nul betekent: nul graden celsius betekent niet nul warmte. Dit in tegenstelling tot een ratioschaal als de Kelvinschaal: nul Kelvin betekent geen beweging van deeltjes, minder dan geen beweging bestaat niet, dus de Kelvinschaal gaat niet onder nul.

Morzycki stelt ook terloops dat *nul* alle kenmerken van virussen heeft. Ik vind het echter niet voor alle kenmerken even duidelijk dat *nul* ze heeft, en zal daarom hier per kenmerk mijn commentaar toevoegen:

1. Nul *wordt geassocieerd met een formeel register.* Dat kan wel kloppen.

2. *Kleine kinderen gebruiken* nul *niet vaak.* Ik vraag me daarbij wel af hoe veel meer kleine kinderen positieve natuurlijke getallen gebruiken, zeker grote getallen.

3. *Het vereist langdurig onderwijs.* Daarbij is wederom de vraag hoeveel meer onderwijs *nul* vereist dan *zeventien* of *achteneenhalf duizend en vier.* Bovendien, hoewel *nul* als wiskundig object duidelijk aangeleerd wordt, lijkt het mij niet dat het grammaticale gebruik ervan als in *nul emails* op school aangeleerd wordt.

4. Morzycki heeft het niet over of *nul* vaak onjuist gebruikt wordt.

5. *Het wordt bewust toegepast.* Dat kan, maar ik vind dit niet zo voor zich spreken als Morzycki het voordoet.

Het is zeker mogelijk dat *nul* een virus is, maar voordat dit definitief geconcludeerd kan worden moet er nog wel het een en ander onderzocht worden. Er zou bijvoorbeeld gezocht kunnen worden naar evidentie van lege pluraliteit in andere talen, en sommige van de kenmerken zouden empirisch gecontroleerd moeten worden.

Vervolgonderzoek zou ook kunnen kijken naar nog een ander grammaticaal gebruik van *nul* wat Bylinina en Nouwen (2018) en Morzycki (2017) niet bespreken. *Nul* wordt soms gebruikt zoals in (36a) in plaats van het neutralere *geen*, waarschijnlijk om extra nadruk te geven. Dit gaat niet over het tellen van *zinnen* (dan zou het een meervoudsvorm zijn) maar lijkt te gaan om het meten van *zin* als een soort grootheid en eenheid op zich. Ook interessant is dat omdat *zin* niet daadwerkelijk iets meetbaars is, elke hoeveelheid *zin* groter dan 0 niet werkt. Voorbeeld (36b) is net zo goed een nerderige halfgrap als de zinnen in (33). Of misschien werkt het met *nul* omdat 0 (op een ratioschaal) geen eenheid nodig heeft. *Zin* is al niet duidelijk een grootheid, maar *afstand* wel. 0 kilometer is hetzelfde als 0 mijl, of 0 van welke afstandseenheid dan ook, dus *nul afstand* is duidelijk, hoewel misschien wat vreemd. Zodra er een daadwerkelijke gemeten hoeveelheid aan te pas komt werkt dat niet: *zeven afstand* betekent niks. Het moet wel opgemerkt worden dat het zou kunnen dat (36a) expres vreemd is om nadruk te geven.

(36) a. Ik heb daar echt nul zin in, dus ik ga niet.
b. ??Ik heb er twaalf zin in, dus ik ga wel.

3.2 Reële getallen

Volgens de viruskenmerken van Morzycki is er niet echt reden om positieve rationale getallen in meetuitdrukkingen (of het tellen van delen zoals Nicolas, 2015 suggereert) als virus te zien.

1. Rationale getallen komen voldoende voor in niet-prestigetaal.
2. Rationale getallen als *anderhalf* zullen niet veel later verworven worden dan positieve natuurlijke getallen.
3. Ze worden aangeleerd op school, maar wel relatief jong, en hetzelfde geldt voor natuurlijke getallen.
4. Fouten in het gebruik komen voor, maar het lijkt mij niet dat ze zoveel voorkomen als bij de virussen die we gezien hebben.
5. Het gebruik van simpele rationale getallen lijkt mij niet erg bewust.

De normale semantiek zoals bijvoorbeeld Snyder (2020) geeft lijkt dus adequaat.

Irrationale getallen zijn misschien iets anders. Hoewel ze ook in meetuitdrukkingen kunnen voorkomen, zoals in (32b) en (32c), hebben ze wellicht wat meer prestige, worden ze later en expliciet aangeleerd, zullen ze vaker onjuist gebruikt worden, en zal het gebruik bewuster zijn. Hoe deze getallen gevormd zouden worden binnen het taalsysteem is ook niet duidelijk. Het valt dus te beargumenteren dat irrationale getallen in taal virussen zijn.

3.3 Negatieve getallen

Negatieve getallen zijn mogelijk problematischer dan reële getallen. Negatieve metingen zijn niet duidelijk aanwezig in de echte wereld. Er zijn geen linealen die negatieve lengte meten, want negatieve lengte bestaat alleen wiskundig. Negatieve waarden kunnen hooguit in vergelijkingen gemeten worden.

Morzycki (2017) noemt negatieve getallen in dezelfde sectie als waar hij ook de semantiek van *nul* behandelt. Dit is opmerkelijk, omdat de semantiek die Bylinina en Nouwen (2018) uitwerken duidelijk gaat over niet-wiskundige pluraliteiten die aanwezig zijn in de echte wereld, alleen met nul leden. Ook kan 0 gemeten worden op ratioschalen. Om de lineaalvergelijking door te zetten, een lineaal begint wel bij 0. Hierdoor is 0 wellicht iets duidelijker aanwezig in de ontologie dan negatieve getallen, maar waarschijnlijk toch niet zo duidelijk als positieve getallen.

Morzycki somt ook de viruskenmerken van negatieve getallen op samen met die van „nul”. Het ligt echter voor negatieve getallen wat meer voor de hand dat de kenmerken gelden.

1. Negatieve getallen lijken me inderdaad vrij formeel.
2. Kleine kinderen zullen niet vaak negatieve getallen gebruiken.

3. Net als nul moeten ze onderwezen worden.
4. Hoewel Morzycki het niet noemt, worden negatieve getallen in ieder geval soms onjuist gebruikt, zoals in (34).
5. Het lijkt me ook redelijk om te zeggen dat ze bewust toegepast worden.

Toch zou het wel waardevol zijn om dit empirisch te toetsen.

3.4 Andersoortige getallen

Nog andere getallen zijn echt duidelijk puur wiskundig. Neem bijvoorbeeld de complexe getallen ($\mathbb{C}$). Dit zijn getallen die bestaan uit de som van een reëel deel en een *imaginair*[12] deel. Dat imaginair deel is een veelvoud van het getal i, wat gedefinieerd is als het getal dat in het kwadraat -1 geeft, oftewel $i^2 = -1$, dus $i = \sqrt{-1}$. Dit valt met normale vermenigvuldiging—vergelijkbaar met de vermenigvuldiging die wordt gebruikt in de morfologie van getalwoorden (onder andere Rothstein, 2017; van Drie, 2015)—niet uit te rekenen: kwadraten zijn altijd positief. Toch is het wiskundig nuttig, en kunnen we daarom zeggen dat complexe getallen wiskundig wel *echt* zijn. Echter, omdat $\sqrt{-1}$ geen zinnige tel- of meetinterpretatie heeft, lijkt het mij redelijk om aan te nemen dat de structuur en betekenis van complexe getallen niet deel is van het taalsysteem.

(37) Ik heb $4i + \sqrt{7}$ mannelijke vrienden. Relatiestatus: Complex.

Over nog minder bekende getalsoorten als de quaternionen of p-adische getallen kunnen we hetzelfde zeggen, al is de kans daar kleiner dat er überhaupt grappen over gemaakt zullen worden door de nog veel kleinere bekendheid.

Hoewel er geen tel- of meetinterpretatie bestaat van deze getallen, zouden ze nog een plek kunnen hebben in het taalsysteem in het getaltype n van Rothstein (2017). Ze zouden dan exclusief type n zijn, en dus niet kunnen worden verhoogd naar type $\langle e, t\rangle$ met de $^{\cup}$-operator. Ze zouden daardoor alleen in wiskundige uitdrukkingen gebruikt kunnen worden.

We kunnen de vraag stellen of deze andersoortige getallen virussen zijn, maar misschien kunnen we beter vragen of het type n, zoals ik het voorstel met complexe en andersoortige getallen of zelfs zonder die uitbreiding, en daarmee wiskundige uitdrukkingen, virussen zijn. Ze lijken wel de kenmerken te hebben:

1. Wiskundige taal heeft zeker een soort prestige.
2. en 3. Wiskunde wordt pas op de middelbare school aangeleerd. Kleine kinderen zeggen geen dingen als „Vier is een even getal".
4. Het lijkt me duidelijk dat wiskundige fouten vaak gemaakt worden.
5. Bij wiskundige uitdrukkingen wordt bewust rekening gehouden met de regels van de wiskunde.

[12] Niet iedereen vindt *imaginair* een goede term. Omdat er een duidelijke wiskundige definitie is, bestaat het wiskundig wel degelijk.

Volgens de virustheorie van Morzycki zijn wiskundige uitdrukkingen en het type n volgens mij dus virussen. Het argument dat Morzycki buiten de kenmerken om maakt over dat virussen meer kracht hebben dan we ze zouden willen geven gaat ook op: we willen niet de kracht van de hele wiskunde aan het taalsysteem toekennen als er vele talen bestaan die nauwelijks een concept van getallen hebben, laat staan wiskunde.

3.5 *Graden* en meten

Ik besteed nog wat tijd aan Morzycki's suggestie dat meetuitdrukkingen, of zelfs *graden* (*degrees*) in het algemeen, virussen zijn. Het moet wel benadrukt worden dat dit wel erg ver gaat, Morzycki noemt het zelf ook radicaal. Ik zal er dan ook niet al te uitgebreid op in gaan.

Morzycki (2017) bespreekt factor phrases als (16a) en (16b) als potentiële virussen. Dit zou zijn omdat ze niet universeel zijn in talen, maar ook omdat de analyse van *graden* (*degrees*) niet erg universeel lijkt. Bochnak (2013) suggereert zelfs dat er eigenlijk überhaupt weinig bewijs voor graden is. Het zou vreemd zijn als de ontologie van getallen sterk verschilt tussen talen, we verwachten dat de ontologie van alle mensen in essentie hetzelfde is. We zouden dus de radicale suggestie kunnen maken dat de gehele literatuur over meten (zoals Rothstein) graden (onder vele andere Cresswell, 1976) eigenlijk virussen beschrijft. Wat betreft de kenmerken, graden—uitgedrukt als getallen—lijken me niet op zich enorm prestigieus, en zo bewust en foutgevoelig lijken ze me ook niet, maar ze worden wel grotendeels pas op ongeveer zesjarige leeftijd bewust aangeleerd. Uiteindelijk is dit waarschijnlijk wel te radicaal, maar het is de moeite waard om even bij de mogelijkheid stil te staan.

3.6 Extragrammaticaliteit

Tot slot geef ik nog een eigen suggestie voor een mogelijk zinnig onderscheid: ik stel voor om een expliciet onderscheid te maken tussen *virus* en *extragrammaticaal*. Ik begin daarbij met een voorbeeld.

Spelers van het van origine Oost-Aziatische spel *baduk* (beter maar verwarrender bekend als *go*, soms ook *weiqi*) kunnen onderling gesprekken hebben als in (38), doorspekt met (grotendeels uit het Japans geleend) jargon. Een normaal mens snapt hier natuurlijk niets van. Op een vergelijkbare manier kunnen wiskundigen dingen zeggen als in (39). Ook dit zal de gemiddelde persoon niet begrijpen zonder expliciete instructie.

(38) A: Je groep daar staat in feite in *atari* door die *snapback*, als je die mist wordt het in die hoek *seki* en wint wit op de *komi*.
B: Ja, maar ik kan niet ontsnappen, want ik zit in een *trap* die ik niet kan breken in *sente*.

(39) Volgens mij ligt $0.3i - 0.919$ niet in de Mandelbrotverzameling.

Misschien is de reden dat de gemiddelde persoon deze twee verschillende dingen niet begrijpt eigenlijk niet zo verschillend. Het jargon van de badukspelers ontstaat voor een deel uit Japanse leenwoorden (atari, komi), en voor een deel uit Nederlandse (of Engelse) termen die een specifiekere betekenis krijgen in een niche-context (trap, snapback). Bij sommigen zijn deze woorden bekend, bij de meesten niet. De woorden maken dus deel uit van het taalsysteem van sommigen, net als andere woorden, maar hebben geen interne structuur.

Wiskundige uitdrukkingen als $0.3i - 0.919$ (*nul komma drie i min nul komma negen één negen*) of $\Theta\left(2^n\right)$ (*grote thèta twee tot de n-de*) zijn op een vergelijkbare manier deel van de taal van sommigen. In tegenstelling tot de baduktermen worden deze wiskundige uitdrukkingen ook begrepen als ze nooit eerder specifiek zo gehoord zijn. Ze ontstaan namelijk (wederom in tegenstelling tot de baduktermen) wél uit een systeem, een soort grammatica: het systeem van de wiskunde. Zoals eerder besproken is dit systeem niet deel van het taalsysteem, maar het is wel een systeem. Misschien kunnen we dit zien als een buiten de taal staand systeem dat nieuwe woorden genereert die het taalsysteem dan kan gebruiken. De „woorden” worden dus niet geleend of aangepast van andere woorden, maar gegenereerd door een wiskundig systeem elders in het brein. Op het niveau van het taalsysteem is de reden dat (38) en (39) wel of niet begrepen worden dan dus lexicaal, en niet grammaticaal.

Het is in ieder geval een verzameling regels die net als PE-regels die Emonds (1986) beschrijft buiten het taalsysteem leven, maar toch invloed hebben op de taal. In de analyse van Morzycki (2017) zou dit misschien als virus gezien kunnen worden, maar op die manier kunnen we alle systemen met uitspreekbare uitkomst (binair tellen, hardop spellen, etc.) wel een virus noemen. Het mist de prescriptivistische kant die volgens Emonds zo belangrijk is, want dit soort taalgebruik wordt niet echt voorgeschreven voor alledaags gebruik. Daarom zou ik voorstellen om drie verschillende statussen toe te kennen aan fenomenen die in taal gebruikt worden:

1. *Grammaticaal*: Dit zijn alle normale grammaticale fenomenen.

2. *Virus*: Virussen staan buiten de grammatica, ze hebben alle kenmerken van virussen volgens Morzycki, plus het onderdrukkende sociolinguïstische prescriptivisme waar Emonds op focust. Voorbeelden zijn de „*… and I …*”-regel, het *hen*/*hun*-onderscheid, en misschien *respectievelijk*.

3. *Extragrammaticaal*: Dit zijn fenomenen die niet inherent deel zijn van het taalsysteem, maar er wel in gebruikt worden. Ze hebben wel de kenmerken van virussen volgens Morzycki, maar worden niet onderdrukkend voorgeschreven zoals virussen. In deze categorie zou de legepluraliteitssemantiek van *nul* kunnen vallen als die inderdaad voldoet aan de eisen van Morzycki. Ook wiskundige uitdrukkingen die het type *n* gebruiken zouden hier passen.

4 Conclusie

In dit eindwerkstuk heb ik de theorie van talige virussen beschreven, en aan de hand daarvan bekeken of sommige getallen volgens die theorie virussen zijn. Emonds (1986) en Sobin (1997) beschrijven de virustheorie als iets puur syntactisch. Emonds legt daarbij veel nadruk op het bevechten van prescriptivisme. Morzycki (2017) breidt de virustheorie uit naar de semantiek. Omdat ik getallen hier als semantische fenomenen behandel, focus ik op de theorie zoals Morzycki die beschrijft. Volgens die theorie lijken verschillende soorten getallen—te veel om hier nog eens op te noemen—in sommige contexten inderdaad virussen te zijn.

De theorie van Morzycki mist toch wel een deel van de sociolinguïstische lading die Emonds belangrijk vindt. Om dit verschil zinnig te bespreken heb ik mijn eigen voorstel bijgedragen: het idee om een onderscheid te maken tussen een echt *virus* en een slechts *extragrammaticaal* fenomeen. Dit lost nog niet op dat de virustheorie niet erg duidelijk uitgewerkt is. Er zal dus meer werk nodig zijn voordat de theorie meer toegepast kan worden.

Referenties

Bernstein, T. (1965). *The Careful Writer: A Modern Guide to English Usage.* New York: Atheneum Press.

Bochnak, M. R. (2013). *Cross-linguistic variation in the semantics of comparatives* (proefschrift, Universiteit van Chicago).

Bylinina, L. & Nouwen, R. (2018). On "zero" and semantic plurality. *Glossa: a journal of general linguistics*, *3*(1). doi:10.5334/gjgl.441

Calude, A. S. (2021). The history of number words in the world's languages—what have we learnt so far? *Philosophical Transactions of the Royal Society B*, *376*(1824), 20200206.

Chomsky, N. (2014). *Aspects of the Theory of Syntax*. MIT press.

Comrie, B. (2013). Numeral Bases. In M. S. Dryer & M. Haspelmath (Red.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Verkregen van https://wals.info/chapter/131

Cresswell, M. J. (1976). The semantics of degree. In *Montague grammar* (pp. 261–292). Elsevier.

Emonds, J. (1986). Grammatically deviant prestige constructions. In M. Brame, H. Contreras & F. Newmeyer (Red.), *A festschrift for Sol Saporta* (pp. 98–129). Seattle, Washington: Noit Amrofer.

Gawron, J. M. & Kehler, A. (2004). The semantics of respective readings, conjunction, and filler-gap dependencies. *Linguistics and Philosophy*, *27*(2), 169–207.

Gebruik van *hen* en *hun*. (g.d.). Verkregen 7 maart 2021, van http://ans.ruhosting.nl/e-ans/05/02/05/02/03/body.html

Hodges, J. & Whitten, M. (1977). *Harbrace College Handbook* (8ste ed.). New York: Harcourt Brace Jovanovich.

Ionin, T., Matushansky, O. & Ruys, E. G. (2006). Parts of speech: Toward a unified semantics for partitives. In C. Davis, A. R. Deal & Y. Zabbal (Red.), *Proceedings of the 36th Annual Meeting of the North East Linguistic Society (NELS 36)* (Deel 1, pp. 357–370). Amherst, Massachusetts: GLSA.

Kaplan, R. (1999). *The nothing that is: A natural history of zero*. Oxford University Press.

Lasnik, H. & Sobin, N. (2000). The who/whom puzzle: On the preservation of an archaic feature. *Natural Language & Linguistic Theory*, *18*(2), 343–371.

Liebesman, D. (2015). We do not count by identity. *Australasian Journal of Philosophy*, *93*(1), 21–42.

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. *Formal semantics: The essential readings*, *127*, 147.

Morzycki, M. (2017). Some viruses in the semantics. In N. LaCara, K. Moulton & A.-M. Tessier (Red.), *A Schrift to Fest Kyle Johnson* (pp. 281–291). Amherst, Massachussetts: Linguistics Open Access Publications.

Nicolas, D. (2015). Counting and measuring–a reply to Liebesman. ijn_01433033.

Nouwen, R. (2016). Plurality. In M. Aloni & P. Dekker (Red.), *The Cambridge Handbook of Formal Semantics* (pp. 267–284). doi:10.1017/CBO9781139236157.010

Quattlebaum, J. A. (1994). A Study of Case Assignment in Coordinate Noun Phrases. *Language Quarterly*, *32*, 131–47.

Ramaker, E. (g.d.). Over ‘hen’ en ‘hun’. Verkregen 7 maart 2021, van https://www.ernieramaker.nl/schrijfsels.php?tekst=henhun

Rothstein, S. (2017). *Semantics for counting and measuring*. Cambridge University Press.

Snyder, E. (2020). Counting, measuring, and the fractional cardinalities puzzle. *Linguistics and Philosophy*, 1–38. doi:10.1007/s10988-020-09297-5

Sobin, N. (1997). Agreement, default rules, and grammatical viruses. *Linguistic inquiry*, *28*(2), 318–343.

van Drie, E. G. A. (2015). *The Morphological Derivation of Numerals* (bacheloreindwerkstuk, Universiteit Utrecht).

van Heule, C. (1625). *De Nederduytsche Grammatica ofte Spraec-konst*. Leiden.

Wągiel, M. (2018). *Subatomic Quantification* (proefschrift, Masaryk-Universiteit in Brno).