

1

Wat is syntaxis?

1.1 *Wat begrippen en misverstanden*

1.1.1 **Waar gaat het om bij de syntaxis?**

Dit boek behandelt een eigenschap van de menselijke taal die bekend staat onder de naam 'syntaxis'. 'Syntaxis' betekent 'zinsbouw': de manier waarop woorden samen woordgroepen en zinnen vormen. Syntaxis is een onderdeel van de GRAMMATICA. De grammatica van een taal bevat verder informatie over het klanksysteem, de vorm van de woorden, de betekenis, enzovoorts.

De term 'syntaxis' wordt ook gebruikt voor de *bestudering* van de syntactische eigenschappen van talen: net zoals 'stilistiek' gebruikt wordt voor de bestudering van literaire stijl. We zullen gaan bestuderen hoe talen hun zinsbouw organiseren; daaronder valt de indeling van de woorden in woordklassen, de volgorde van de woorden binnen woordgroepen en zinnen, de structuur van de woordgroepen en de zinnen, en de verschillende typen zinnen die in talen gebruikt worden. Wij willen je laten begrijpen hoe de syntaxis werkt, en daarvoor heb je een aantal belangrijke begrippen en technische termen nodig die we dus moeten introduceren. Daarom zullen we grammaticale termen tegenkomen zoals 'naamwoord', 'werkwoord', 'voorzetsel', 'relatiefzin', 'subject', 'nominatief', 'conjunctief', en 'passief'. Wij gaan er niet van uit dat je deze termen nu al kent. We geven ook niet altijd meteen een definitie van zo'n term, maar we proberen wel te laten zien wat hij betekent, zodat je het concept begrijpt en klaar bent voor de bespreking die dan volgt.

We gaan in dit boek voorbeelden bekijken van de zinsbouw van veel verschillende talen; sommige daarvan zijn verwant aan het Nederlands, maar lang niet allemaal. Syntaxis zit in alle talen, en 'primitieve' samenlevingen hebben net zo complexe talen als de meest 'geciviliseerde' industriële samenlevingen. Voor een taalkundige is dat duidelijk, omdat (zoals we zullen zien) dezelfde geavanceerde eigenschappen in allerlei talen terugkomen en omdat sprekers van alle talen dezelfde gedachten en begrippen

kunnen uitdrukken, dezelfde discussies kunnen voeren, en dezelfde ingewikkelde redeneringen kunnen opzetten. Zelfs de stapjes die kinderen zetten bij het leren van hun moedertaal zijn opmerkelijk gelijk van cultuur tot cultuur, en in alle culturen doen ze er ongeveer even lang over. Er is geen enkele taal die zo ingewikkeld is dat hij niet vloeiend gesproken wordt door iedere puber! Als kinderen alle talen even makkelijk kunnen leren, dan moeten alle talen even moeilijk zijn.

Nu zul je misschien tegenwerpen dat sommige talen beslist moeilijker zijn dan andere. Misschien was je op school goed in Engels, maar heb je nu grote problemen met je avondcursus Arabisch. Laten we eens kijken wat dat zegt. Heel belangrijk is dat, als het gaat om het leren van een taal, volwassenen erg in het nadeel zijn ten opzichte van kinderen. Tegen de tijd dat we de puberteit naderen raakt ons vermogen talen te leren in verval. Zelfs als we totaal ondergedompeld worden in een nieuwe cultuur blijft het leren van de taal een heel karwei. Misschien kunnen we hem op een gegeven moment vloeiend spreken, maar we worden toch nooit moedertaalsprekers — we hebben niet hetzelfde gevoel voor de taal als iemand die hem als kind heeft leren spreken.

Daarnaast is het altijd gemakkelijker om een taal te leren spreken die aan je eigen taal verwant is of die er *TYPOLOGISCH* op lijkt — dat wil zeggen, een taal die voor bepaalde kenmerken tot hetzelfde type behoort, bijvoorbeeld omdat hij dezelfde woordvolgorde heeft. Dus als je een taal moeilijk noemt, zeg je in feite dat die taal een aantal eigenschappen heeft waar je niet vertrouwd mee bent. Arabisch is niet intrinsiek moeilijk, net zo min als Swahili of Mohawk of enige andere taal, ook al zijn sommige eigenschappen beslist lastiger dan andere, ook voor kinderen die hem als hun moedertaal leren.

Om te laten zien wat de bestudering van de syntaxis eigenlijk is moeten we het eerst hebben over wat het *niet* is. Als je leest dat de 'syntaxis' een onderdeel is van de 'grammatica', dan denk je misschien aan bepaalde zaken die wij in dit boek niet voor ogen hebben. Ten eerste: ook al hebben we het over grammatica, dit is niet een *BESCHRIJVENDE GRAMMATICA* van het Nederlands of van enige andere taal. Zulke boeken bestaan ook (bijvoorbeeld de *Algemene Nederlandse Spraakkunst* of *ANS*), maar gewoonlijk bespreken die de eigenaardigheden van *één* taal in plaats van de principes die verantwoordelijk zijn voor de organisatie van taal in zijn algemeenheid. Ten tweede: we gaan niet proberen je beheersing van de grammatica van het Nederlands te verbeteren. Een *PRESCRIPTIEVE GRAMMATICA* (één die voorschrijft hoe de auteurs menen dat je moet praten) zal je misschien proberen te vertellen wanneer je *hen* moet zeggen en wanneer *hun*; of dat je *heeft gelopen* moet schrijven en niet *gelopen heeft*; of dat het absoluut fout is om te zeggen *groter als* in plaats van *groter dan*; of dat passieve

werkwoorden zoveel mogelijk vermeden dienen te worden. Wij doen daar niet aan mee (dus als je niet weet hoe het hoort, maakt niet uit).

Taalkundigen zien het zo: iedere moedertaalspreker van het Nederlands kent de grammatica van het Nederlands perfect (maakt niet uit welk dialect je spreekt). En iedere moedertaalspreker van een andere taal kent de grammatica van die taal perfect. Daarmee bedoelen we niet dat je (bewust) een paar regeltjes kent, zoals die uit de vorige alinea, maar dat je (onbewust) de veel indrukwekkendere mentale grammatica kent van je eigen taal — zoals alle moedertaalsprekers. Die grammatica hebben we allemaal geleerd, maar het is kennis die ons nooit onderwezen is, en het is ook kennis die we er niet uit kunnen halen om te onderzoeken. Rond hun zevende jaar hebben kinderen een tamelijk complete kennis van de grammatica van hun moedertaal, en wat er daarna gebeurt is voornamelijk het uitbouwen van hun woordenschat. Een grammatica die volledig verworven is kan niet meer verbeterd worden. Het gaat net zo als bij het 'leren' lopen. Je kunt het lopen niet onderwijzen; we doen het allemaal vanzelf zodra we er klaar voor zijn. Zelfs als we volledig op de hoogte zijn van de spierbewegingen en hersenactiviteiten die er bij betrokken zijn, dan nog 'weten' we niet hoe we lopen. Het leren van je moedertaal is net zo: het gebeurt zonder tussenkomst van buitenaf, en de kennis die het oplevert blijft ontoegankelijk voor ons.

Maar, werp je dan misschien tegen, ik heb wel degelijk onderwijs gekregen in mijn moedertaal. Misschien denk je dat je ouders hem je geleerd hebben, of anders dat je het op school geleerd hebt. Maar dat is een misverstand. Alle normale kinderen in alle culturen leren hun moedertaal (of moedertalen) perfect zonder enig formeel onderricht. Het enige wat je nodig hebt is blootstelling aan normale menselijke interactie. Om dit te testen hoeven we maar te vragen of alle culturen hun kinderen 'grammatica' onderwijzen. Nee dus. Daarom mogen we concluderen dat alle kinderen in staat zijn een mentale grammatica van hun moedertaal te construeren zonder enige formele instructie. De meeste taalkundigen geloven inmiddels dat kinderen dit kunnen omdat ze voorgeprogrammeerd zijn voor het leren van taal, net zoals we voorgeprogrammeerd zijn om te leren lopen.

Maar als je de grammatica van je moedertaal niet van iemand geleerd hebt, wat probeerden ze je dan te leren toen je ouders zeiden dat je niet *Hij hep* of *Hun zeggen* moest zeggen, of toen je op school te horen kreeg dat *Hij schaamt z'n eigen* niet goed is? Denk opnieuw aan de parallel met het leren lopen. Kinderen leren dat perfect zonder ouderlijke instructie, maar hun ouders vinden het misschien maar niks dat hun kind zo slentert, of met zijn tenen over de grond sleept. Dan vertellen ze hun kind misschien wel dat hij rechtop moet lopen en zijn voeten op moet tillen. Dat wil niet zeggen dat het kind niet goed kan lopen, hooguit dat zijn manier van lopen niet past bij

bepaalde ideeën over esthetiek of stijl. Precies zo hebben sommige mensen het idee dat bepaalde taalvormen mooier of stijlvoller of gewoonweg correct zijn. Maar het idee dat sommige taalvormen beter dan andere zijn heeft geen taalkundige grondslag; het gaat in wezen om sociale factoren. Omdat we onze oordelen over mensen vaak baseren op hun accent of dialect, bestaat de neiging die oordelen over te dragen op hun taalvormen. Dan denken we al gauw dat bepaalde vormen onwenselijk zijn, dat sommige 'goed' zijn en andere 'fout'. Maar voor een taalkundige is er geen enkele moedertaalspreker die 'taalfouten' produceert.

Wellicht werp je nu tegen dat SUBSTANDAARD Nederlands zoals de voorbeelden uit de vorige alinea of zoiets als *Heus echt waar niet* slordig taalgebruik zijn, of zelfs in strijd met de logica. Dit beroep op logica en exactheid wekt de indruk dat de prescriptieve grammatica zich beweegt in hogere sferen, en dat er meer aan de hand is dan alleen sociale vooringenomenheid. Laten we daarom het logica-argument nader onderzoeken en kijken of het overeind blijft. 'Twee keer min is plus', hoor je vaak zeggen, zodat zinnen als in (1) 'eigenlijk' betekenen *Ik heb altijd iets gedaan*:

(1) Ik heb nooit niks gedaan

Maar dat is natuurlijk niet zo. Ten eerste bedoelt een spreker die (1) zegt helemaal niet *Ik heb altijd iets gedaan*. Net zo min als iemand die (1) hoort denkt dat het betekent *Ik heb altijd iets gedaan*, ook al is hij nog zo gebrand op dubbele ontkenningen. Ten tweede zijn er talen, zoals het Frans en het Bretons, waarbij de dubbele ontkenning niet iets is van een dialect, maar van de STANDAARD, zoals geïllustreerd in (2).¹

(2) Je **ne** mange **jamais** de viande (Frans)
 ik NEGATIVE eet nooit van vlees
 'Ik eet nooit vlees.'

Voorbeeld (2) laat zien dat de ontkenning in het Standaard Frans twee onderdelen heeft: naast het kleine ontkenningswoordje *ne* is er nog een ander ontkenningswoord *jamais* 'nooit'. Trouwens, het Middelnederlands (het Nederlands uit de jaren 1100-1500) had ook een dubbele ontkenning. Ironisch genoeg voor het 'logica'-argument is de variëteit van het Frans die een dubbele ontkenning heeft juist de meest formele en

¹ Paragraaf 1.2.2.1 legt precies uit hoe je taalkundige voorbeelden dient te lezen. Je hoeft geen Frans te kennen om het punt te begrijpen dat aan de hand van voorbeeld (2) gemaakt wordt: de methode die je moet volgen is om eerst de Nederlandse vertaling te lezen, en dan de tweede regel, een letterlijke vertaling van de voorbeeldzin.

prestigieuze variëteit, terwijl in het gesproken Frans het kleine ontkenningwoordje meestal wegvalt.

Dialecten en andere vormen van gesproken Nederlands vertonen wel vaker trekken die voor de prescriptieve grammatica niet door de beugel kunnen. Een voorbeeld is geïllustreerd in (3):

(3) Wij gaan drie **maand** op vakantie

Hier gaat het logica-argument als volgt: het aantal maanden dat jullie op vakantie gaan is meer dan één, dus moet je de meervoudsvorm *maanden* gebruiken. Maar dit argument deugt niet. Ook in het algemeen Nederlands komt een constructie als (3) voor, alleen niet met *maand* maar met *uur* en *jaar*.

(4) a. Wij waren drie **uur** onderweg
b. Kim is al drie **jaar**

De standaardtaal is dus niet logischer dan het dialect, en de schrijftaal niet logischer dan de spreektaal. Variëteiten van een taal kunnen natuurlijk verschillen in de keuzes die ze maken: voor het standaard-Nederlands zijn *jaar* en *uur* anders dan *maand* en *dag*, voor de dialecten waar (3) mogelijk is hoort *maand* gewoon in het groepje van *uur* en *jaar*.

Het is dus duidelijk dat sociaal gestigmatiseerde taalvormen niet meer of minder logisch zijn dan het standaard Nederlands. Sprekers van substandaarddialecten volgen natuurlijk net zoals de sprekers van het meest prestigieuze dialect een verzameling mentale regels. Eigenlijk hebben de verschillende variëteiten van een taal de overgrote meerderheid van de regels van die taal gemeen, en verschillen ze maar op een paar puntjes, maar de mate waarin wordt overschat omdat de verschillen zulke sterke reacties oproepen. Kortom, sprekers van een prestige-dialect mogen dan wel denken dat zij de enige 'grammaticaal correcte' variëteit van het Nederlands spreken, die visie kan niet verdedigd worden met logische of taalkundige argumenten.

Aan de andere kant, stel dat een spreker van het Nederlands zinnen als (5) zou produceren.

(5) *Ik niets heb gedaan nooit.
*Nooit niets heb gedaan ik.

Dan zouden we met recht kunnen claimen dat zijn spraak ongrammaticaal is. Zulke voorbeelden gaan volledig tegen de mentale regels van alle dialecten van het Nederlands in. Daarover is iedereen het eens, en toch hebben sprekers van het Nederlands nooit te horen gekregen dat de zinnen in (5) fout zijn. Onze oordelen moeten daarom wel onderdeel uitmaken van de mentale grammatica van het Nederlands die we allemaal gemeen hebben.

De meeste regels van deze mentale grammatica kom je nooit tegen in prescriptieve grammatica's of schoolgrammatica's. Bijvoorbeeld, geen enkele grammatica van het Nederlands zou ooit de moeite nemen uit te leggen dat we wel (6a) en (6b) kunnen zeggen, maar niet een vraag als (7).

- (6) a. Wat eten ze?
b. Ze eten eieren en frites.
- (7) *Wat eten ze eieren en ____ ?

De regels die (7) onmogelijk maken zijn zo onveranderlijk en fundamenteel dat ze nauwelijks voor bespreking in aanmerking lijken te komen. Voorbeelden als (7) zijn niet alleen ongrammaticaal in het Nederlands (d.w.z., moedertaalsprekers vinden ze onmogelijk), maar ook in het Welsh, zoals in (8).²

- (8) *Beth maen nhw yn bwyta wyau a ____ ?
wat zijn zij in eten eieren en
['*]Wat eten zij eieren en ____ ?

Het equivalent van (8) is zelfs in alle menselijke talen onmogelijk. Het lijkt er dan ook op dat veel van de regels van de mentale grammatica die we onbewust 'kennen' als sprekers van individuele talen als het Nederlands en het Welsh in feite UNIVERSEEL zijn — d.w.z. dat ze voor alle talen gelden.

Het feit dat bepaalde regels en principes die de taal vormgeven universeel zijn leidt voor veel taalkundigen tot de conclusie dat mensen een AANGEBOREN taalvermogen hebben. We kunnen de inhoud van dat taalvermogen niet rechtstreeks bestuderen, maar wel zijn output — namelijk de taalbouwsels. Door naar de syntaxis te kijken hopen we de eigenschappen te ontdekken die talen gemeen hebben, en uiteindelijk misschien zelfs iets te ontdekken over de manier waarop het menselijk brein werkt.

² Het sterretje tussen vierkante haakjes [*] gebruiken we wanneer de vertaling van de voorbeeldzin zelf (ook) ongrammaticaal is.

Nu gaan we terug naar (9). Als *Hasan* gepromoveerd is zoals in (9b) kunnen we het nog verder promoveren naar de voorste plaats in de zin:

(11) Hasan kreeg van Ali ___ de brief opgestuurd

De plek waar *Hasan* vandaan komt geven we aan met een liggend streepje (___). In (11) is *opsturen* veranderd in *opgestuurd krijgen*, een signaal dat aangeeft dat *Hasan* gepromoveerd is. Dat zo'n signaal nuttig is wordt duidelijk als je even nadenkt over het verschil in betekenis tussen *Hasan heeft de brief opgestuurd* en *Hasan kreeg de brief opgestuurd*.

Dezelfde promotie van *Hasan*, uitgaande van (9a), waar *Hasan* nog niet eerder gepromoveerd is, zou (12) opleveren: we geven de positie waar *Hasan* vandaan komt weer aan met een liggend streepje. Maar (12) is geen mogelijke Nederlandse zin (zoals het sterretje aangeeft):

(12) *Hasan kreeg van Ali de brief naar ___ opgestuurd

Dus als het woord *Hasan* al gepromoveerd is, kan het nog verder opschuiven, maar zo niet, dan is verdere promotie onmogelijk.

Misschien denk je dat dat komt doordat er in (12) *naar* staat: als we *naar* schrappen is de zin weer goed. Maar als we kijken naar (13) en (14), de Indonesische tegenhangers van (11) en (12), krijgen we een sterke aanwijzing dat het daar niet aan ligt (de verandering van *meng-kirim* in (10) in *di-kirim* in (13) en (14) is vergelijkbaar met de verandering van *opsturen* naar *opgestuurd krijgen*):

(13) Hasan di-kirim-**kan** surat itu oleh Ali
 Hasan kreeg gestuurd brief de door Ali
 'Hasan kreeg van Ali de brief opgestuurd.'

(14) *Hasan di-kirim surat itu (kepada) oleh Ali
 Hasan kreeg opgestuurd brief de naar door Ali
 * 'Hasan kreeg de brief van Ali naar opgestuurd.'

Net als in het Nederlands is de ene zin prima, en de andere onmogelijk. Wat is het verschil? Aan het Indonesisch kunnen we zien dat het niets te maken heeft met het woord voor 'naar' (*kepada*), want (14) is onmogelijk met of zonder dat woord — de haakjes betekenen dat het niet uitmaakt of het woord er nu wel of niet staat, de zin blijft ongrammaticaal. De reden dat alleen (13) acceptabel is, is dat we moeten beginnen

- (16) a. **Niemand en** spaerde anderen daer
 ‘**Niemand** spaerde daar de ander.’
- b. Dit **en** is **niet** geloghen
 ‘Dit is **niet** gelogen.’

Omdat negatie hier dubbel uitgedrukt werd kon het zwakke element *en/ne* vervolgens verdwijnen, althans in het algemeen Nederlands. In sommige dialecten bleef het element *ne/en* als negatie in stand (bijvoorbeeld het Westvlaams). Een dergelijke ontwikkelingsgang in de uitdrukking van de ontkenning is in veel talen waargenomen. Het is blijkbaar een natuurlijke gang van zaken, en het is niet duidelijk in welke zin het Nederlands er beter of slechter op geworden is.

Een ander voorbeeld van een ontwikkeling in de grammatica van het Nederlands heeft betrekking op het grammaticaal GESLACHT (Latijn ‘genus’, Engels ‘gender’). Veel talen verdelen hun naamwoorden in klassen, die we ‘geslachten’ noemen. Die term heeft waarschijnlijk hiermee te maken dat de woorden voor ‘man’ en ‘vrouw’ vaak in verschillende woordklassen zitten, maar het is van belang om in het oog te houden dat biologisch geslacht (‘sexe’, Latijn ‘sexus’) er niets mee te maken heeft. Het aantal geslachten kan in de loop van de geschiedenis van een taal veranderen. Zo had het Middelnederlands drie geslachten (mannelijk, vrouwelijk, en onzijdig), en heeft het huidige Nederlands er maar twee (mannelijk/vrouwelijk enerzijds en onzijdig anderzijds).

Misschien ben je gewend te denken dat ook het Nederlands van nu drie geslachten heeft. In de woordenboeken staat immers nog steeds een aanduiding ‘m’, ‘v’, of ‘onz’ voor ‘mannelijk’, ‘vrouwelijk’ en ‘onzijdig’. Maar dat is helaas een hardnekkig archaïsme dat niets te maken heeft met de grammatica van het Nederlands. Een aanwijzing daarvoor is dat niemand weet welke woorden precies mannelijk of vrouwelijk zijn (daarom moet het vermeld worden in het woordenboek!). Maar welke woorden onzijdig zijn weet iedereen precies. Dat laatste is de normale situatie bij geslachten. Moedertaalsprekers weten feilloos van welk geslacht de woorden van hun taal zijn, ook als de taal niet twee of drie maar wel zestien geslachten heeft, zoals de Bantu-talen. Dat Nederlanders zelfs niet bij benadering weten welke woorden mannelijk zijn en welke vrouwelijk geeft aan dat het onderscheid mannelijk-vrouwelijk in het Nederlands niet meer bestaat.

Hoe weet je nu bij welk geslacht een naamwoord hoort? Dat is eenvoudig vast te stellen aan de hand van de vorm van de woordjes die in de omgeving van dat naamwoord staan. Het Duits bijvoorbeeld heeft drie geslachten (mannelijk, vrouwelijk, en onzijdig), en dat is onmiddellijk af te lezen aan de vorm van de LIDWOORDEN die bij de

naamwoorden optreden (MASC = mannelijk, Engels 'masculine'; FEM = vrouwelijk, Engels 'feminine'; NTR = onzijdig, Engels 'neuter'):

- (17) a. **der** Mann
de:MASC man 'de man'
- b. **die** Frau
de:FEM vrouw 'de vrouw'
- c. **das** Kind
de:NTR kind 'het kind'

Hier verschillen de vormen van het lidwoord afhankelijk van het geslacht: *der* is voor naamwoorden van het mannelijk geslacht, *die* voor naamwoorden van het vrouwelijk geslacht, en *das* voor naamwoorden van het onzijdig geslacht. In het Nederlands hebben we maar twee lidwoorden, één voor het onzijdig (*het*) en één voor de rest (*de*), zoals in de vertaling van (17) te zien is.

Ook aan de vorm van adjectieven die bij het naamwoord optreden vallen de verschillende geslachten af te lezen. Dit zien we het beste wanneer de naamwoordgroepen INDEFINIET zijn, d.w.z. dat ze als lidwoord niet *de/het* hebben, maar *een*. Opnieuw zien we in het Duits dan drie verschillende vormen van het adjectief:

- (18) a. ein alt-**er** Mann
een oud-MASC man 'een oud-**e** man'
- b. ein-e alt-**e** Frau
een-FEM oud-FEM vrouw 'een oud-**e** vrouw'
- c. ein alt-**es** Kind
een oud-NTR kind 'een oud kind'

Zoals we in de Nederlandse vertaling in (18) kunnen zien is er één vorm voor het adjectief dat bij een onzijdig woord hoort, *oud*, en een andere vorm voor de rest, *oude*. Tenslotte zien we hetzelfde patroon bij de vorm van de RELATIEFPRONOMINA (betrekkelijk voornaamwoorden) die bij een naamwoord aansluiten. Dit zijn woordjes als *die* in *de man die ik zie*, die een bijzin inleiden die iets over het naamwoord zegt (een relatiefzin of betrekkelijke bijzin). Het Duits heeft opnieuw drie verschillende vormen, het Nederlands maar twee:

- (19) a. der Mann **der** das sagt
 de:MASC man REL:MASC dat zegt
 'de man **die** dat zegt'
- b. die Frau **die** das sagt
 de:FEM vrouw REL:FEM dat zegt
 'de vrouw **die** dat zegt'
- d. das Kind **das** das sagt
 de:NTR kind REL:NTR dat zegt
 'het kind **dat** dat zegt'

In het Nederlands is er één relatiefpronomen, *dat*, voor bij onzijdige naamwoorden, en een ander, *die*, voor bij mannelijke en vrouwelijke naamwoorden. Telkens zien we dat waar het Duits een indeling in drieën maakt, het Nederlands maar een indeling in tweeën gebruikt, en telkens vallen bij die indeling in tweeën de zgn. mannelijke en vrouwelijke woorden in één en dezelfde groep. Er is maar één verschijnsel in het Nederlands dat lijkt te wijzen op een onderscheid tussen mannelijke en vrouwelijke woorden, en dat is de vorm van een pronomen dat terugverwijst naar een naamwoord, zoals in *Kim en haar vriend*. Hier verwijst *haar* terug naar *Kim*, en omdat 'Kim' in al onze voorbeelden een meisje is, is het pronomen van het vrouwelijk geslacht, *haar*. Ook bij deze ANAFORISCHE pronomina zijn er maar twee vormen, *zijn* en *haar*, maar nu vallen de mannelijke en de onzijdige woorden samen, en staan de vrouwelijke naamwoorden apart:

- (20) a. de man en **zijn** vriend
 b. de vrouw en **haar** vriend
 c. het huis en **zijn** ligging

Maar de vorm van het anaforisch pronomen vertelt ons niets over het grammaticaal geslacht, maar over het biologisch geslacht. Dat blijkt uit het voorbeeld in (21): *meisje* en *jongetje* zijn allebei grammaticaal onzijdig, zoals je kunt zien aan het lidwoord *het*, maar biologisch respectievelijk vrouwelijk en mannelijk. Het anaforisch pronomen moet dan bij *meisje* vrouwelijk zijn, *haar*, en niet *zijn*, wat voor mannelijke en onzijdige naamwoorden wordt gebruikt:

- (21) a. Het meisje en **haar**/*zijn vriend
 b. Het jongetje en **zijn**/*haar vriend

Het Nederlands heeft dus in zijn huidige stadium maar twee geslachten, het onzijdig en wat we noemen het COMMON GENDER (gemeenschappelijk geslacht, ook wel ‘zijdig’ of ‘nonneuter’ genoemd). Maar dat is niet altijd zo geweest. In het Middel-Nederlands zijn er drie geslachten aan te wijzen, en ook deze situatie leeft in sommige dialecten nog altijd voort. In (22) staan de vormen van het lidwoord en het adjectief bij naamwoorden van de drie geslachten op een rijtje. Er is een aparte vorm voor de naamwoordgroep die als object gebruikt wordt (zoals ook het huidige Nederlands een aparte vorm van het persoonlijk voornaamwoord heeft dat als object gebruikt wordt, vgl. *Ik zie jou* en *Jij ziet mij*).

- (22) a. SUBJECT **die** goed-e gast
 OBJECT **die-n** goed-e-n gast
- b. SUBJECT **die** goed-e daet
 OBJECT **die** goed-e daet
- c. SUBJECT **dat** goed-e hof
 OBJECT **dat** goed-e hof

In de subjectvorm zijn er in het Middelnederlands, zoals in het huidige Nederlands, maar twee vormen te onderscheiden: *die* voor mannelijk *gast* en voor vrouwelijk *daet*, en *dat* voor onzijdig *hof*. Maar de objectsvorm ziet er voor het lidwoord bij mannelijke naamwoorden anders uit dan voor het lidwoord bij vrouwelijke naamwoorden: *dien* vs. *die*. Ook het adjectief heeft een extra uitgang *-n* bij mannelijke naamwoorden. Deze situatie lijkt dus meer op die van het huidige Duits dan op die van het huidige Nederlands. De objectsvorm raakte in het Nederlands — afgezien van de pronomina — al vrij snel buiten gebruik (behalve in de schrijftaal, waar hij pas in het midden van de twintigste eeuw afgeschaft werd), en het is duidelijk dat bij het wegvallen van de objectsvorm er een systeem ontstaat waarbij de mannelijke en vrouwelijke vormen identiek zijn.

De veranderingen in het Nederlands die we nu gezien hebben deden zich binnen een periode van enkele eeuwen voor. Het is moeilijk te zeggen dat het Nederlands door deze veranderingen verloederd is: zo'n term is als het gaat over taalverandering feitelijk niet gedefinieerd. Een taal schuift deels door toevallige omstandigheden op van het ene stadium naar het andere, maar elk stadium is een volwaardige taal die sprekers in staat stelt hun gedachten tot in de perfectie uit te drukken.

Ingrijpende veranderingen doen zich vaak voor onder invloed van contact met andere talen. Ook hier is niet duidelijk sprake van achteruitgang, maar de vrees dat contact met een vreemde taal de moedertaal besmet is er niet minder om. Hier is het goed om te weten dat taalcontact meestal alleen een uitwerking heeft op de woordenschat: het systeem van een taal, zijn grammatica, wordt alleen onder extreme omstandigheden aangetast. De huidige invloed van het Engels op het Nederlands heeft een stroom van nieuwe woorden opgeleverd, maar de syntaxis van het Nederlands is daardoor niet meer op die van het Engels gaan lijken. Om te waarderen hoe de syntaxis van het Nederlands en het Engels van elkaar verschillen moeten we eerst kennismaken met alle verschijnselen en processen in de syntaxis die we in dit boek zullen bespreken. Maar we kunnen al wel illustreren dat woorden die uit een andere taal ontleend worden in principe moeiteloos ingepast worden in het systeem van de taal. Zo heeft het Engels ons het woord *stewardess* gegeven, en wij maken daar een Nederlandse meervoudsvorm bij (*stewardessen* tegenover het Engelse *stewardesses*) en een Nederlandse verkleinvorm (*stewardessje*) die in het Engels niet bestaat. Ook als we een werkwoord aan het Engels ontlenen, zoals *mailen*, geven we dat werkwoord Nederlandse uitgangen (*ik mail, jij/hij/zij mail-t, wij/jullie/zij mail-en, ik/jij/hij/zij mail-de, wij/jullie/zij mail-den, ik heb ge-mail-d, al mail-end*) die afwijken van de Engelse. Een leenwoord wordt dus in principe ingepast in het systeem van de Nederlandse grammatica. Pas wanneer sprekers in een gemeenschap intensief twee talen naast elkaar gebruiken, en de nieuwe taal duidelijk dominant is, kunnen we meer ingrijpende vormen van taalverandering door taalcontact verwachten. Maar ook dan is de term ‘verloedering’ niet op zijn plaats, omdat taalverandering iets is van alle talen en van alle eeuwen.

Samengevat, in paragraaf 1.1 hebben we betoogd dat alle moedertaalsprekers van een taal een interne grammatica gemeen hebben, waarvan ze de regels kennen zonder dat ze er les in gehad hebben. Dat wordt gesuggereerd door het feit dat we het in grote lijnen eens zijn over wat wel en niet een zin van onze taal kan zijn, ook al verschillen sprekers meestal wel in de mate waarin ze elementen van de gesproken taal en van dialecten accepteren. Bovendien hebben talen die niet aan elkaar verwant zijn veel eigenschappen en constructies gemeen, wat suggereert dat mensen een aangeboren taalvermogen hebben. We hebben ook beargumenteerd dat talen niet ingedeeld kunnen worden in ‘moeilijke’ en ‘gemakkelijke’ talen, omdat kinderen van elke taal de grammatica even gemakkelijk kunnen leren. Tenslotte hebben we gezien dat talen steeds veranderen, en hebben we een onderscheid gemaakt tussen ontlening, wat geen effect heeft op het grammaticale systeem, en verandering van het grammaticale systeem zelf. Nu gaan we laten zien hoe we om moeten gaan met voorbeelden uit andere talen.

1.2 Voorbeeldzinnen gebruiken

1.2.1 Waarom niet gewoon alleen Nederlandse voorbeelden?

In dit boek vind je voorbeelden uit een grote verscheidenheid aan talen, waaronder het Nederlands. Misschien vind je het in het begin lastig om voorbeelden uit exotische talen te analyseren en vraag je je af waarom we het niet gewoon bij Nederlandse voorbeelden laten. Er zijn twee belangrijke redenen om ook voorbeelden uit vreemde talen te gebruiken: zij leren ons enerzijds over de verschillen, en anderzijds over de overeenkomsten tussen talen.

Om te beginnen zien niet alle talen er hetzelfde uit, en als je alleen naar onze eigen taal en een paar nauw verwante talen kijkt, mis je de enorme verschillen die er tussen talen kunnen bestaan. Stel dat je naast het Nederlands alleen kennis gemaakt hebt met het Duits, een nauw verwante noordeuropese Germaanse taal. Voorbeeld (23), uit het Duits, is een woordelijke vertaling van het Nederlands.

(23) der schöne Wasserfall (Duits)
de mooie waterval

Je zou kunnen gaan denken dat deze woordgroep er in alle talen hetzelfde uitziet: eerst een woordje voor 'de', dan een woord voor 'mooie', en dan een woord voor 'waterval'. Maar zo is het niet. In het Spaans bijvoorbeeld krijgen we (24):

(24) la cascada hermosa (Spaans)
de waterval mooie
'de mooie waterval'

Hier verschilt de woordvolgorde op één punt: het woord voor 'mooie' komt na het woord voor 'waterval'. Voor de rest is het Spaans niet zo heel anders dan het Nederlands: het gebruikt precies dezelfde drie woorden en een woord voor 'de' in dezelfde positie. Dat is niet zo verrassend, want het Spaans is ook aan het Nederlands verwant, zij het minder nauw dan het Duits. Maar in andere talen komt het equivalent van 'de' helemaal achteraan, zoals geïllustreerd in (10), of is er helemaal geen equivalent voor 'de', zoals in het Japans en het Chinees, of zelfs geen directe vertaling voor het bijvoeglijk naamwoord 'mooie'.

De talen van de wereld bezitten veel interessante en belangrijke syntactische eigenschappen waar we graag willen dat jullie over weten. Het Nederlands heeft er daar

een paar van, maar lang niet allemaal, zodat je een hoop zou missen als we alleen maar naar het Nederlands keken. In (25) zien we één voorbeeld, uit het Spaans:

(25) Es nuevo (Spaans)
 is nieuw
 'Het is nieuw.'

Voorbeeld (25) heeft geen woordje voor 'het'; letterlijk staat er 'Is nieuw' — een onmogelijke zin in het Nederlands. Veel talen hebben vergelijkbare voorbeelden, maar als we de discussie tot het Nederlands beperkt hielden zouden we daar nooit achter komen. In weer andere talen, zoals het Arabisch en het Indonesisch, wordt de zin van drie woorden *Het is nieuw* vertaald als *Het nieuw* (dit wordt geïllustreerd in hoofdstuk 2). Deze eenvoudige voorbeelden laten zien dat we er niet van uit kunnen gaan dat zinnen in andere talen woord-voor-woord vertalingen zijn van Nederlandse zinnen. Daarom bestuderen we andere talen om het hele scala van constructies en kenmerken te ontdekken — om meer te weten te komen over TAALDIVERSITEIT.

De tweede reden dat we naar voorbeelden van andere talen kijken is dat taalkundigen willen ontdekken welke eigenschappen talen gemeen hebben — hun HOMOGENITEIT of gelijkheid. Eén van de belangrijkste ontdekkingen van de moderne taalkunde is dat talen niet willekeurig van elkaar verschillen, maar opmerkelijke gelijkenissen vertonen. Sommige kenmerken komen in alle talen voor. Bijvoorbeeld, elke taal maakt onderscheid tussen de klasse van NAAMWOORDEN (zoals *boom*, *vloeistof*, *uitdrukking* en *student*) en de klasse van WERKWOORDEN (zoals *liquideren*, *leren*, *lijken* en *groeien*), ook al zijn er talen die geen verdere woordklassen hebben. (In hoofdstuk 2 gaan we nader op de woordklassen in.) Om dit soort informatie te achterhalen moeten taalkundigen naar een representatieve doorsnede van de talen van de wereld kijken, talen uit verschillende taalfamilies en uit verschillende delen van de wereld.

De meeste taalkundigen willen de kernpatronen ontdekken die alle talen gemeen hebben. Ook al zijn bepaalde constructies niet universeel (= van alle talen), talen gebruiken toch allemaal dezelfde grammaticale basiselementen. Elke taal heeft een woordenlijst of LEXICON die alle sprekers gemeen hebben, en die woordenlijst bevat altijd woorden van verschillende woordklassen. Alle talen combineren deze woorden tot woordgroepen en zinnen, en kunnen de volgorde van de woordgroepen voor verschillende doeleinden manipuleren — wellicht om vragen te stellen, of om verschillende onderdelen van de zin nadruk te geven, of om te laten zien wie wat met wie doet. Dat is syntaxis, en het vormt de stof van de hoofdstukken die voor ons liggen.

1.2.2 Hoe je voorbeeldzinnen moet lezen

1.2.2.1 De vormgeving van de voorbeeldzinnen

Het eerste wat je als syntacticus te doen staat is het leren gebruik maken van voorbeelden uit andere talen. Dit boek bevat voorbeelden uit ca. 80 verschillende talen. Uiteraard kunnen wij de meeste van die talen niet zelf spreken — we krijgen de voorbeelden van andere taalkundigen, of van moedertaalsprekers van de taal in kwestie (die soms zelf taalkundigen zijn). Maar we kunnen die voorbeelden gebruiken omdat taalkundigen ze presenteren op een speciale manier voor mensen die de taal niet spreken maar hem wel willen leren of onderzoeken.

Voorbeelden van die speciale vormgeving zagen we al in de Spaanse voorbeelden (24) en (25). Elk voorbeeld bestaat uit drie regels. De bovenste regel komt uit de taal waar we het over hebben. De onderste regel is een vertaling van de bovenste regel in het Nederlands. Die heb je nodig om te weten wat de oorspronkelijke voorbeeldzin betekent, maar het is niet het belangrijkste onderdeel van de presentatie van het voorbeeld, want hij zegt alleen maar iets over het Nederlands — hij zegt niets over de andere taal. De regel waar het om draait is de middelste, die we GLOSSE noemen. De glosse is een letterlijke vertaling van het oorspronkelijke voorbeeld; elk betekenisdragend onderdeel van het origineel wordt vertaald, of het nu precies met een Nederlands woord correspondeert of niet. Neem zin (2) er nog eens bij: het Franse woordje *ne* is eenvoudig GEGLOST (vertaald) als NEGATIEF omdat er geen Nederlands woord is dat er rechtstreeks mee correspondeert.³

Als je wilt weten waarom de glosse zo belangrijk is, kijk dan eens naar (26) en (27), uit het Japans en het Welsh. De regel met glossen hebben we weggelaten. De twee voorbeelden betekenen hetzelfde, in die zin dat ze dezelfde Nederlandse vertaling krijgen:

- | | | |
|------|---|----------|
| (26) | Sensei ga gakusei ni tegami o kaita
'De leraar schreef de student een brief.' | (Japans) |
| (27) | Sgwenodd yr athro lythyr at y myfyriwr
'De leraar schreef de student een brief.' | (Welsh) |

³ De functionele informatie zullen we in de regel met de Engelse termen aangeven. Een lijst met afkortingen die we in de glossen gebruiken staat voor in dit boek.

- (31) De student vraag-t om die boek-en
 DEF student vragen-3SG om DEM:PL boek-PL
 'De student vraagt om die boeken.'

Aan de hand van dit voorbeeld kunnen we de taalkundige conventies illustreren. Sommige woorden zijn gewoon vertaald, zoals 'student', 'vragen', 'om', en 'boek', maar bijvoorbeeld *de* hebben we zo geglosst dat alleen de functionele informatie die het overbrengt weergegeven wordt: het is een bepaald ('DEFINIET') lidwoord en staat als zodanig tegenover het onbepaalde ('INDEFINIETE') lidwoord *een*. We hadden dat ook anders kunnen doen, bijvoorbeeld door *de* gewoon te glossen als 'de', en meestal zullen we dat ook zo doen (zoals bijv. in (30) met *te*).

Naast de lexicale en functionele informatie staan er ook nog bepaalde tekens in de glossen: de dubbele punt (:) en het streepje (-). We hebben *die* geglosst als DEM:PL: *die* is een aanwijzend ('DEMONSTRATIEF') woordje, net als *dit*, *dat*, *deze*; het is ook meervoud ('PLURALIS'), en daarom kan het gebruikt worden in combinatie met een meervoudsvorm als *boeken*. De dubbele punt in de glosse DEM:PL betekent dat je binnen het woord niet iets kan aanwijzen wat 'meervoud' betekent: *die* is tegelijkertijd in zijn geheel 'aanwijzend' en 'meervoud'; sommige auteurs gebruiken hiervoor een plusteken (+) of een punt (.) in plaats van een dubbele punt.

Een streepje (-) voor of na een stukje functionele informatie in de glosse betekent dat die informatie aan het woord vastgeplakt zit; je kunt dan binnen het woord duidelijk aanwijzen waar de functionele informatie zich bevindt. Bijvoorbeeld, in *boeken*, geglosst als *boek-PL*, zit het element 'meervoud' hem in de uitgang *-en*. In *vraagt*, geglosst als *vraag-3SG*, brengt de uitgang *-t* over dat het werkwoord de kenmerken 'derde persoon enkelvoud (SINGULARIS)' heeft. We hebben het streepje in (31) ook gebruikt in de voorbeeldzin zelf, om aan te geven hoe in de oorspronkelijke taal de lexicale en de functionele elementen gescheiden kunnen worden, maar dat zullen we niet het hele boek door blijven doen.

Functionele informatie die aan het begin of het eind van een woord vastgeplakt zit noemen we een AFFIX. Meestal geeft een streepje dus aan dat er een affix staat, zoals bijvoorbeeld de meervoudsuitgang *-en*. Er zijn twee soorten grammaticale affixen: suffixen en prefixen. In het Nederlands zijn het element voor meervoud *-en* en het element voor 'derde persoon enkelvoud' SUFFIXEN: zij worden aan het eind van een woord vastgeplakt. Wordt iets aan het begin van een woord vastgeplakt dan noemen we dat een PREFIX: voorbeelden in het Nederlands zijn *on-* in *onecht* en *ge-* in *gedaan*.

1.2.2.3 De categorieën persoon en getal

In deze paragraaf bespreken we de standaardmethode om de functionele categorieën **PERSOON** (PERSON) en **GETAL** (NUMBER) weer te geven, aan de hand van voorbeelden uit het Frans en het Kwamera (een taal uit Vanuatu in de Stille Oceaan).

Als je ooit iets aan een vreemde taal gedaan hebt ken je waarschijnlijk wel die rijtjes met werkwoordsvormen zoals in (32), uit het Frans:

(32) Tegenwoordige tijd (presens) van *parler* ‘praten’

	enkelvoud (singularis)	meervoud (pluralis)
1e	je parle	nous parlons
2e	tu parles	vous parlez
3e	il/elle parle	ils/elles parlent

Zulke rijtjes, die **PARADIGMA’S** heten, laten zien welke vormen een bepaald woord kan aannemen. Het voorbeeld in (32) toont de tegenwoordige tijd (**PRESENS**) van het werkwoord *parler* ‘praten’: de vormen betekenen, van boven naar beneden, in het eerste rijtje ‘ik praat, jij praat, hij/zij praat’ en in het tweede rijtje ‘wij praten, jullie praten, zij praten (resp. mannelijk en vrouwelijk)’. De labels 1e (**EERSTE**), 2e (**TWEEDE**), en 3e (**DERDE**) in de linkerkolom geven de functionele categorie **PERSOON** (PERSON) aan. Eerste persoon is de spreker, of een groep mensen waar de spreker bij hoort: de vormen met ‘ik’ en ‘wij’. Tweede persoon is de aangesprokene: de vormen met ‘jij’ en ‘jullie’. Derde persoon is een ‘derde’, een persoon of groep die noch spreker, noch aangesprokene is: de vormen met ‘hij/zij/het’ en ‘zij’.

De categorie **GETAL** (NUMBER) heeft betrekking op het onderscheid tussen enkelvoud (**SINGULARIS**, d.w.z. één persoon) en meervoud (**PLURALIS**, d.w.z. meer dan één persoon). In het Frans is het getal, net als in de meeste Europese talen, ofwel ‘enkelvoud’ ofwel ‘meervoud’. (Maar merk op dat de tweede persoon meervoud, *vous* ook gebruikt wordt als beleefdheidsvorm, ‘u’, en dan is de vorm van het werkwoord 2e persoon meervoud, hoewel de aangesprokene er maar één kan zijn.) Maar in sommige talen zijn er meer vormen van meervoud, bijvoorbeeld één die verwijst naar twee personen (**DUALIS**) (of drie, ‘trialis’), naast een normale meervoudsvorm die verwijst naar meer dan twee (of drie) personen. Bijvoorbeeld de Polynesische taal Kwamera heeft precies zo’n systeem.

Kwamera heeft meer PERSOONsonderscheidingen dan we kennen uit de Indo-Europese talen. Er zijn twee vormen voor de eerste persoon: INCLUSIEF en EXCLUSIEF. Deze maken onderscheid tussen ‘wij (=spreker + aangesprokene)’ en ‘wij (=spreker + derde, dus *niet* de aangesprokene)’. Stel, je zegt tegen iemand ‘We kunnen vanavond wel een filmpje gaan pakken’. Diegene kan dan antwoorden ‘Wij? Bedoel je ons (*inclusief ‘wij’*) of jullie (*exclusief ‘wij’*)?’ Het Nederlands heeft geen twee vormen voor ‘wij’ die die informatie geven, maar het Kwamera wel:

- (33) a. **sa-ha-akw** (Kwamera)
 1INC-PLURAL-opstappen
 ‘Wij stappen op.’ (inclusief ‘wij’)
 b. **ia-ha-vehe**
 1EXC-PLURAL-komen
 ‘Wij kwamen.’ (exclusief ‘wij’)

Het prefix *ia-* in (33b) betekent niet ‘ik’ of ‘wij’ maar gewoon eerste persoon: het wordt pas ‘wij’ wanneer het meervoudsaffix er achter staat. Daarom kan dezelfde vorm *ia-* ook ‘ik’ betekenen:

- (34) **ia-pkagkiari-mha**
 1EXC-praten-NEG
 ‘Ik heb niet gepraat.’

In de rest van het boek glossen we persoon en getal zoals in (35), tenzij de taal in kwestie speciale inclusieve en exclusieve vormen heeft zoals het Kwamera. De eerste twee rijtjes geven de glossen en hun betekenis — dit is de grammaticale informatie — en de derde kolom geeft de voornaamwoorden (PRONOMINA) die in het Nederlands met die grammaticale informatie overeenkomen, als geheugensteuntje.

- (35) Glossen voor persoon en getal

glosse	betekenis	Nederlands pronomen
1SG	eerste persoon enkelvoud	<i>ik</i>
2SG	tweede persoon enkelvoud	<i>jij</i>
3SG	derde persoon enkelvoud	<i>hij/zij</i>
1PL	eerste persoon meervoud	<i>wij</i>
2PL	tweede persoon meervoud	<i>jullie</i>
3PL	derde persoon meervoud	<i>zij</i>

Als de glosse alleen de persoon geeft, dus '1', '2', of '3', maar niet zegt of het enkel- of meervoud is, dan betekent dat dat het woord in kwestie geen onderscheid naar getal maakt.

1.2.2.4 *Schriftsystemen en glossen*

Niet alle talen gebruiken het romeinse alfabet (wat je nu leest). Bijvoorbeeld, het Russisch gebruikt het cyrillische alfabet, en het Chinees en het Japans gebruiken geen alfabet maar een schriftsysteem dat gebaseerd is op karakters. Maar meestal bestaan er afspraken over hoe die talen in het romeinse alfabet geschreven worden, en daardoor kunnen taalkundigen voorbeelden uit die talen ook gebruiken. We volgen op dat punt de publicaties waar we onze gegevens vandaan hebben, al passen we soms de glossen aan het hier gebruikte systeem aan. Daarnaast standaardiseren we de glossen die teveel of juist te weinig informatie geven. Hier en daar vereenvoudigen we door een bepaald element ongeglost te laten, met name als we de grammaticale term in kwestie nog niet gehad hebben. Tonen in voorbeelden uit talen als het Chinees laten we weg: ze hebben betrekking op de uitspraak en zijn niet van levensbelang voor de bespreking van de syntaxis. Tenslotte moet er op gewezen worden dat sommige talen helemaal geen schriftsysteem hebben, omdat ze nooit opgeschreven zijn. In dat geval geven taalkundigen in principe een fonetische representatie van de oorspronkelijke taal; het fonetische alfabet volgt niet de conventies van een schriftsysteem.

1.3 *Waarom hebben talen syntaxis?*

Wat voor zaken drukt de syntaxis uit? Deze paragraaf geeft een eerste indruk van de eigenschappen van talen die typisch te maken hebben met zinsbouw en toont aan dat talen werkelijk een syntactische structuur hebben.

1.3.1 **Woordvolgorde**

De bouw van een zin wordt in de eerste plaats zichtbaar in de woordvolgorde (hoewel we zullen zien, in hoofdstuk 5, dat zinnen strikt genomen niet opgebouwd zijn uit WOORDEN maar uit WOORDGROEPEN). Met de drie woorden *Teun*, *thee*, en *dronk* kunnen we, door de volgorde van de woorden te variëren, zes verschillende zinnen maken, die niet allemaal mogelijk zijn in het Nederlands:

- (36) a. *Teun thee dronk
 b. Teun dronk thee
 c. Dronk Teun thee
 d. *Thee Teun dronk
 e. Thee dronk Teun
 f. *Dronk thee Teun

De meest normale Nederlandse zin is (36b). (36c) en (36e) zijn alleen onder speciale omstandigheden mogelijk. Maar voor andere talen ligt dat anders. Elk van de zes volgordes kan in één of meer talen worden aangetroffen als de meest normale volgorde, hoewel sommige volgordes veel gebruikelijker zijn dan andere (zie hoofdstuk 6). De drie meest voorkomende 'neutrale volgordes' in andere talen dan het Nederlands zijn die in (36a)-(36c). We zagen al in paragraaf 1.2.2 dat het Japans als neutrale volgorde (36a) heeft, en het Welsh (36c). Het Malagasy, de taal van Madagascar, heeft (36f) als neutrale volgorde. De twee volgordes in (36d) en (36e) zijn het meest zeldzaam als neutrale volgorde, maar ze worden aangetroffen in de Carib-talen die gesproken worden in het Amazone-gebied: het Hixkaryana bijvoorbeeld heeft (36e) als neutrale volgorde.

We kunnen dus voor elk van de talen van de wereld een NEUTRALE VOLGORDE vaststellen. Maar talen verschillen ook enorm ten aanzien van de vrijheid die ze hebben om van hun neutrale volgorde af te wijken, en ten aanzien van de functie die zo'n afwijkende woordvolgorde heeft in het gebruik. In het Nederlands bijvoorbeeld is (36c) alleen mogelijk als een vraag, terwijl (36e) gebruikt kan worden om de nadruk (FOCUS) te leggen op wat Teun drinkt. In beide zinnen staat het woord *dronk* vóór het woord *Teun*, terwijl het in de neutrale volgorde erachter staat. In het laatste geval, (36e), is ook het woord *thee* ten opzichte van zijn neutrale positie (de positie die het inneemt in de neutrale volgorde (36b)) naar voren verplaatst. Hiervoor is een bepaalde context nodig: stel je voor dat Teun op bezoek gaat bij een excentrieke tante die zelf thee en bier vervaardigt op basis van de kruiden uit haar tuin. Dan kun je zeggen: *De thee dronk Teun __ (wel), maar het bier weigerde zij __*. In technische termen: hier is OBJECT ('lijdend voorwerp') van de werkwoorden *dronk* en *weigerde* vooropgeplaatst; het liggende streepje (__) geeft aan waar het zou staan in de neutrale volgorde. Omdat voor deze volgorde een bepaalde context nodig is, wordt hij GEMARKEERD genoemd, terwijl de neutrale volgorde ONGEMARKEERD is.

Er kan ook een stylistische reden zijn om de volgorde om te gooien. In (37) bijvoorbeeld is het vooropgeplaatste element aan de lange kant:

Dat de keus tussen positie 2 en 3 bepaald wordt door grammaticale factoren blijkt ook uit zinnen van het type *En hij maar lachen* of *Hij een huis kopen? (Kom nou!)*. Hier valt op dat het werkwoord *lachen/kopen* geen enkele functionele kenmerken vertoont (zoals bijvoorbeeld de kenmerken PAST/NONPAST, PERSON, NUMBER die we tegengekomen zijn in paragraaf 1.2.2.2 en 1.2.2.3), in tegenstelling tot bijvoorbeeld *dronk* in (36) en (39), dat de kenmerken PAST (verleden tijd) en SG (enkelvoud) draagt. Op de vraag hoe we weten dat deze kenmerken aan- of afwezig zijn komen we in paragraaf 2.3.2.2 terug. Op dit moment is van belang dat zo'n werkwoord zonder functionele kenmerken alleen maar in positie 3 kan voorkomen:

- (42) a. En Teun maar thee drinken
 b. *En Teun drinken maar thee
 c. *En drinken Teun maar thee

Dat geeft opnieuw aan dat de positie van het werkwoord in het Nederlands iets te maken heeft met de grammatica, en niet met de context of met stijl.

Op de technische termen die we bij deze voorbeelden gebruikt hebben komen we in de rest van het boek nog uitgebreid terug. Ook het aspect van de woordvolgorde behandelen we nog een keer apart (in hoofdstuk 6). Voor nu is van belang dat talen verschillen in de keus van wat hun neutrale woordvolgorde is, en in de vrijheid die ze hebben om van die neutrale volgorde af te wijken. Beide aspecten van de woordvolgorde worden in de syntaxis zo goed mogelijk beschreven en verklaard.

1.3.2 Promotie en degradatie

De syntactische variatie in paragraaf 1.3.1 bestond er simpelweg in dat we de volgorde van de zinsdelen veranderden. Maar een verandering van de zinsbouw kan ook veel ingrijpender consequenties hebben. In paragraaf 1.1 hebben we het al gehad over promotie- en degradatieprocessen. Nu gaan we het hebben over een andere constructie die het gevolg is van zulke processen: de PASSIEF. Onze voorbeelden komen uit het Nederlands en het Japans.

- (43) De natuurlijke beweging wordt afgeremd door een weerstandbiedend medium
 (uit *En toch beweegt zij* Dick Stafleu: Boom, 1992)

- (44) Hendrik de Vries werd als opgroeiende jongeman nog even heftig door erotische driften beheerst als toen hij kind of puber was

(uit *Hendrik de Vries (1896-1989)* Jan van der Vegt: De Prom, 1996)

- (48) a. Sensei wa John o sikat-ta (actief) (Japans)
 leraar TOPIC John OBJECT berispen-PAST
 ‘De leraar berispte John.’
- b. John wa sensei ni sikar-are-ta (passief)
 John TOPIC leraar door berispen-PASS-PAST
 ‘John werd door de leraar berispt.’

In (48a) is ‘de leraar’ het subject en tevens de ‘topic’ (gespreksonderwerp) van de zin, en ‘John’ is gemarkeerd als het object van ‘schelden’ door middel van het elementje *o*. In (48b) is ‘John’ naar de subjectpositie gepromoveerd en ‘de leraar’ is gedegradeerd; het staat nu in de *door*-groep *sensei ni* ‘leraar door’ — merk op dat de volgorde van de woorden in deze groep in het Japans anders is dan in het Nederlands. Het werkwoord heeft ook een speciaal suffix, *-are*.⁶

In hoofdstuk 7 worden passieven en andere promotie/degradatie-constructies in detail besproken.

1.3.3 Alle talen hebben structuur

Syntactische structuur is aanwezig in alle menselijke talen, dood of levend, inclusief (uiteraard) gebarentalen. Dat wil zeggen dat een taal niet gewoon bestaat uit rijtjes woorden, maar dat de woorden zich groeperen in woordgroepen, en de woordgroepen in grotere woordgroepen en uiteindelijk in zinnen.

Dat blijkt al uit het feit dat alle talen INGEBEDDE zinnen hebben — zinnen die in een andere zin zitten — zoals *Kim zei tegen Teun dat Daan niet kon zwemmen*, waar (*dat*) *Daan niet kon zwemmen* de ingebedde zin is. In (49) staan nog meer voorbeelden hiervan uit het Nederlands: de ingebedde zinnen staan tussen vierkante haakjes.

- (49) Ik vraag me af [of Kim niet te laat komt]
 Jouw stelling [dat Kim Teun niet mag] verbaast mij enorm
 [Zodra Kim en Teun er zijn] gaan we er van door
 [Dat de koffie op is] is mijn schuld niet

⁶ De stam van het werkwoord (het gedeelte vóór de suffixen) is *sikar*, maar voor het verleden-tijdssuffix *-ta* verandert hij in *sikat*.

Dat de groepen woorden tussen haakjes in (49) echt zinnen zijn is misschien niet onmiddellijk duidelijk. Je ziet het pas als je telkens het eerste woordje (*of, dat, zodra*) weghaalt en het werkwoord (*komt, mag, zijn, is*) naar voren schuift: het resultaat is dan een zin die volledig op zichzelf kan staan (bijvoorbeeld: *dat de koffie op is* -> (eerste woordje weg) *de koffie op is* -> (werkwoord naar voren) *de koffie is op*). In hoofdstuk drie zullen we zien wat precies de definitie is van 'zin'.

We kunnen zinnen tot in het oneindige in elkaar inbedden. Bijvoorbeeld, van *Kim kan niet zwemmen* kunnen we maken *Teun dacht dat Kim niet kon zwemmen*, en daarna *Daan zei dat Teun dacht dat Kim niet kon zwemmen*, enzovoorts. Deze eigenschap heet RECURSIE, en ook die eigenschap is in alle talen aanwezig. Dat betekent dat het nooit mogelijk is, in welke taal dan ook, om een 'langste zin' te maken.

We sluiten dit hoofdstuk af met een paar eenvoudige voorbeelden die het bestaan van syntactische structuur demonstreren. Kijk eerst eens naar de woordgroepen in (50):

- (50) a. geachte dames en heren
b. zwangere vrouwen en partners

Op het eerste gezicht hebben deze woordgroepen dezelfde opbouw: in beide zinnen staan dezelfde (soorten) woorden in dezelfde volgorde. Maar het is duidelijk dat (in de meest voor de hand liggende interpretatie) *geachte* in (50a) slaat op zowel *dames* als *heren*, terwijl in (50b) *zwangere* alleen betrekking heeft op *vrouwen*. Deze interpretatie geven we weer door de woorden in (50a) en (50b) op verschillende manieren te groeperen, wat we opnieuw aangeven met vierkante haakjes:

- (51) a. geachte [dames en heren]
b. [zwangere vrouwen] en partners

Deze notatie geeft aan dat er in (50a) een woordgroep *dames en heren* is, maar dat de corresponderende reeks woorden in (50b), *vrouwen en partners* geen woordgroep vormt: *vrouwen* zit immers in (50b) met *zwangere* in een woordgroep. Omgekeerd zit *dames* niet (althans niet direct) met *geachte* in een woordgroep in (50a) (wel indirect, omdat de groep *dames en heren* als geheel weer wel samen met *geachte* in een groep zit).

De indeling in groepen in (51) voorspelt nu dat we *dames en heren* in (50a) kunnen vervangen door iets anders zonder dat de woordgroep ongrammaticaal wordt, maar dat dat met *vrouwen en partners* in (50b) niet kan. Dat kunnen we dus testen door met de woordgroepen in (50) te gaan spelen:

- (52) a. geachte [dames en heren] → geachte aanwezigen
 b. [zwangere vrouwen] en partners ↗ zwangere deelnemers

In (52b) kan *deelnemers* niet gezien worden als vervanging van *vrouwen en partners*. Omgekeerd kunnen we in (50b) wel *zwangere vrouwen* vervangen door iets anders, maar *geachte dames* in (50a) niet:

- (53) a. geachte [dames en heren] ↗ aanwezigen en heren
 b. [zwangere vrouwen] en partners → cursisten en partners

Deze tests in (52) en (53) maken gebruik van de aanname dat een reeks woorden die je door een andere reeks woorden kunt vervangen een *eenheid* is. De tests laten dus zien dat de reeksen woorden in (50a) en (50b) opgebouwd zijn uit verschillende eenheden, met andere woorden, dat ze een verschillende structuur hebben.

Een ander voorbeeld dat het bestaan van syntactische structuur aantoont staat in (54) (deze voorbeelden moeten zo gelezen worden, dat de grootgeschreven woorden de nadruk krijgen):

- (54) a. Hij bleef de hele reis DOOR zeuren
 b. Hij bleef de hele reis door ZEUREN

Opnieuw lijkt het alsof de twee zinnen dezelfde opbouw hebben: hier zijn zelfs de woorden en de woordvolgorde exact hetzelfde. Toch kun je laten zien dat de syntactische structuur van de zinnen verschillend is. In (54a) kun je *de hele reis* vervangen door *drie kwartier*, maar in (54b) niet (let op dat je de zinnen uitspreekt met de juiste nadruk!):

- (55) a. Hij bleef drie kwartier DOOR zeuren
 b. *Hij bleef drie kwartier door ZEUREN

En in (54b) kun je *zeuren* vervangen door *op alle slakken zout leggen*, maar in (54a) niet:

- (56) a. *Hij bleef de hele reis DOOR op alle slakken zout leggen
 b. Hij bleef de hele reis door op alle slakken ZOUT leggen

Deze tests suggereren dat in (54a) *de hele reis* een woordgroep is, waar *door* niet bijhoort: daarom kunnen we in (55a) *de hele reis* vervangen door *drie kwartier*, terwijl we *door zeuren* gewoon laten staan. *Door zeuren* moet dus een eigen groep vormen

in (54a), maar in (54b) niet: daar moet *zeuren* in zijn eentje een groep vormen, en *door* moet bij *de hele reis* horen. Daarom kunnen we *zeuren* in (55b) vervangen door *op alle slakken zout leggen* terwijl we *de hele reis door* gewoon laten staan. De zinnen in (54) verschillen dus in hun structuur, zoals de haakjes in (57) laten zien:

- (57) a. Hij bleef [de hele reis] [DOOR zeuren]
 b. Hij bleef [de hele reis door] [ZEUREN]

Het mooie is dat we op nog veel meer manieren kunnen laten zien dat in (54a) *de hele reis* een eenheid is, en in (54b) *de hele reis door*. Bijvoorbeeld, in (58a) hebben we *de hele reis* naar voren verplaatst, en in (58b) *de hele reis door*.

- (58) a. [De hele reis] bleef hij DOOR zeuren
 b. [De hele reis door] bleef hij ZEUREN

Als we *door* niet mee vooropplaatsen, zoals in (58a), moet het de nadruk krijgen, en correspondeert de zin dus met (54a). Als we *door* wel mee vooropzetten, zoals in (58b), dan valt de nadruk op *zeuren*, en correspondeert de zin dus met (54b). Deze voorbeelden laten zien dat verplaatsing, net als vervanging, gebruik maakt van de syntactische eenheden (de woordgroepen) en dus inzicht geeft in de syntactische structuur.

Zulke demonstraties kunnen we geven voor alle talen van de wereld: de regels van de syntaxis van alle talen maken gebruik van structuur (zijn STRUCTUURAFHANKELIJK). Daarom maken talen geen gebruik van regels die bijvoorbeeld een mededeling omzetten in een vraag door de woorden van de zin van achteren naar voren te lezen — zo'n regel zou geen gebruik maken van de structuur van de zin, en zou dus niet werken. Als wij, als taalkundigen, iets te weten willen komen over de structuur van de zin van een bepaalde taal, dan voeren we tests uit (zoals hierboven geïllustreerd) waarbij we steunen op de manier waarop moedertaalsprekers de resultaten beoordelen. Zij vinden bijvoorbeeld een bepaalde vervanging of verplaatsing mogelijk of onmogelijk (in het laatste geval krijgt de zin een sterretje). Deze GRAMMATICALITEITSOORDELEN vormen de gegevens waar de taalwetenschap mee werkt. Moedertaalsprekers hoeven niet te kunnen aangeven *waarom* een zin voor hun goed of slecht is; de intuïties die ze hebben laten zien dat er tussen zinnen structurele verschillen en overeenkomsten bestaan.