

Marijke De Belder
Carl von Ossietzky Universität Oldenburg
Contact: marijke.de.belder@uol.de

Nederlandse woordsoortspecifieke lettergreepstructuren

1. Inleiding

In de jaren 1980 merkte Mieke Trommelen op dat Nederlandse werkwoorden minder en eenvoudigere lettergrepen hebben dan Nederlandse zelfstandige naamwoorden (Trommelen 1989). Met die observatie is decennialang niet veel gedaan. Nochtans zijn woordsoortspecifieke lettergreepstructuren gekend uit andere talen en belangrijk voor de taalkundige theorievorming. Met hedendaags computationeel en experimen-

teel onderzoek hebben we aangetoond dat Trommelen overschot van gelijk had: het Nederlands heeft inderdaad woordsoortspecifieke lettergreepstructuren. Die kennis behoort bovendien tot de intuïties van de moedertaalspreker. Theoretisch volgen er twee belangrijke vragen uit dit onderzoek:

1. Hoe kan het dat een syntactische eigenschap (zoals woordsoort) in verband staat met een fonologische eigenschap?
2. Waarom zou dat verband kunnen bestaan?

2. Andere talen dan het Nederlands

Uit een typologisch overzichtartikel (Smith 2011) blijkt dat fonologische eigenschappen die het niveau van het foneem overstijgen in heel wat talen woordsoortspecifieke eigenschappen vertonen. Meer bepaald gaat het om ‘lettergreepstructuren’, ‘klemtoonpatronen’ (zoals in het Engelse *rebel* dat een andere klemtoon krijgt als zelfstandig naamwoord dan als werkwoord) en om ‘toon’ voor toontalen zoals het Mandarijns. Het is over talen heen typisch zo dat zelfstandige naamwoorden de meest permissieve woordsoort zijn: ze laten meer atypische patronen toe dan werkwoorden. Bijvoeglijke naamwoorden vallen tussen de zelfstandige naamwoorden en werkwoorden in. Deze observaties stemmen overeen met Trommelens observaties. Het Nederlands zou dan een van de talen in de wereld zijn waarin er een relatie bestaat tussen woordsoort en fonologie. Het spijtige is dat Trommelen haar methodologie niet beschreef. Het was dus moeilijk om te bepalen hoe betrouwbaar haar conclusies waren.

3. Computationeel onderzoek

Met hedendaagse middelen hebben we het onderzoek heropend. We hebben eerst en vooral gebruikgemaakt van een kunstmatige leerder (Hayes & Wilson 2007). De software is zo ontworpen dat die beperkingen kan meten op hoe klinkers en medeklinkers in een taal gecombineerd kunnen worden. De software meet hiervoor gegevens uit lijsten van woorden en hun frequenties. Voor de woordenlijsten maakten we gebruik van het ondertitelingscorpus SUBTLEX (Keuleers, Brysbaert & New 2007). We gaven twee aparte lijsten van eenlettergrepige, morfologisch ongelede woorden in: een lijst van werkwoordstammen en een lijst van zelfstandige naamwoorden. Dat stond ons toe om voor de twee woordsoorten apart beperkingen op lettergrepen te meten.

Uit deze methode bleek dat werkwoorden meestal bestaan uit de meest prototypische Nederlandse syllaben, terwijl zelfstandige naamwoorden ook vaker minder typische lettergreepstructuren kunnen hebben. De meest prototypische lettergrepen hebben een medeklinker in de aanzet en bestaan uit een korte klinker, gevolgd door een of twee medeklinkers (*kam*, *werk*) of een lange klinker of tweeklank, gevolgd door een

medeklinker (*maan, dijk*). Voor andere lettergreetypes (*hoorn, herfst, ambt*) maten we zeer duidelijke verschillen tussen zelfstandige naamwoorden en werkwoordstammen: werkwoorden hebben zonder enige twijfel een eenvoudiger lettergreepstructuur dan zelfstandige naamwoorden.

We hebben ook gemeten of morfologisch ongelede werkwoorden vaker monosyllabisch zijn dan zelfstandige naamwoorden. Ook daar meet je inderdaad duidelijke verschillen. Eigenlijk zie je het al meteen als je simpelweg telt. De volgende tabel toont hoeveel lettergrepen de woorden in het corpus hadden. De tabel behandelt enkel morfologisch ongelede woorden waarvan geen enkele lettergreep een sjwalettergreep is. Voor de werkwoorden tellen we de lettergrepen van de stam, niet van de infinitief:

	totaal	1 l		2 l		3 l		4 l		5 l	
		n	%	n	%	n	%	n	%	n	%
werkw.	951	872	91,69	79	8,31						
zelfst. nw.	2.212	1.207	54,57	691	31,24	259	11,71	51	2,31	4	0,18

(werk w = werkwoord, zelfst. nw. = zelfstandig naamwoord, l = lettergreep, n = absolute aantal)

Tabel 1 – Verdeling van werkwoorden en zelfstandige naamwoorden volgens hoeveelheid aan lettergrepen.

Je ziet onmiddellijk dat zelfstandige naamwoorden veel langer kunnen zijn dan werkwoorden en dat de overgrote meerderheid van de werkwoorden monosyllabisch is (*zing, vecht, huil, lach, werk*). Er is slechts een handvol tweelettergrepige werkwoorden zonder sjwa (*neurie, boycot, ravot*). Kortom: als je naar de inventaris van Nederlandse werkwoorden en zelfstandige naamwoorden kijkt, meet je overduidelijke verschillen in lettergreepstructuur. (Voor het Duits geldt overigens hetzelfde.)

4. Experimenteel onderzoek

Het is niet omdat je kunt aantonen dat iets geldt voor het Nederlandse lexicon, dat dat feit ook een rol speelt voor de moedertaalspreker. Is het verschil in lettergreepstructuur over woordsoorten heen deel van de talige competentie van de spreker? Gebruikt de moedertaalspreker deze kennis onderbewust? Om dat te meten, hebben we een grootschalig online experiment uitgevoerd (zie ook: Don & Erkelens 2006). We hebben aan moedertaalsprekers gevraagd of ze een intuïtie hebben over de woordsoort van pseudowoorden door ze een voorkeurscontext te laten aanvinken, zoals in de volgende voorbeelden:

blech	mireun
<input type="checkbox"/> ik blech	<input type="checkbox"/> een mireun
<input type="checkbox"/> een blech	<input type="checkbox"/> ik mireun

Pseudowoorden zijn woorden die volledig gehoorzamen aan de Nederlandse fonologie en dus Nederlandse woorden hadden kunnen zijn. Ze bestaan gewoon toevallig niet. We hebben voor 38 soorten lettergreepstructuren telkens twintig pseudowoorden ontworpen. Elk pseudowoord is door 60 moedertaalsprekers beoordeeld geworden. We hadden dus voor elk type lettergreep 1.200 oordelen.

Opnieuw werd duidelijk dat er een groot contrast is tussen zelfstandige naamwoorden en werkwoorden. Tweelettergrepige pseudowoorden worden overwegend als zelfstandig naamwoord aangevoeld, eenlettergrepige woorden met een complexere lettergreepstructuur ook. Interessant genoeg zijn er voor sommige types verschillen tussen wat we computationeel en experimenteel gemeten hebben. Dat betekent dat de intuïties van de moedertaalsprekers niet puur een reflectie zijn van onderbewuste statistische kennis van het lexicon. Het lijkt erop dat er een specifiek mechanisme aan het werk is in de talige competentie van de moedertaalspreker dat een verband legt tussen lettergreepstructuur en woordsoort.

5. Functiewoorden en taalverwerking

We weten natuurlijk al langer dat functiewoorden een aparte fonologie hebben (Zec 2005). De Nederlandse lidwoorden en voornaamwoorden hebben duidelijk andere lettergreep eigenschappen dan de inhoudswoorden. Onze lidwoorden, bijvoorbeeld, hebben een sjwa als enige klinker, wat onmogelijk is voor inhoudswoorden. Bij woorden uit gesloten klassen ontbreken ook vaker medeklinkercontouren: ze hebben opvallend vaak lege aanzetten (*een, ik, iets*) of het zijn open lettergrepen (*die, hij*).

Er is in het verleden succesvol aangetoond dat pasgeborenen (in een Engelstalige omgeving) het fonologische verschil opmerken tussen functiewoorden en inhoudswoorden (Shi, Werker & Morgan 1999). Vermoed wordt dat dit een duidelijk voordeel oplevert bij taalverwerving. Door simpelweg te luisteren naar de lettergreepstructuur in de zin kun je als kind heel wat syntactische informatie bijeen puzzelen: de fonologie helpt je om te bepalen welk woord in de zin welke functie kan hebben. Dat principe kun je doordenken voor de verschillen tussen de inhoudswoorden onderling. Een kind kan dan raden waar de zelfstandige naamwoorden en de werkwoorden staan op basis van de lettergrepen.

Als een kind dat kan, dan kan een volwassene dat ook. We vermoeden dat talen in de wereld woordsoortgebaseerde fonologische verschillen kunnen hebben omdat dat een voordeel oplevert bij de mentale verwerking van gesproken taal. Puur op basis van wat

je hoort, kun je dan al grotendeels de syntactische structuur van de zin ontrafelen, zelfs zonder dat je alle woorden moet kennen of moet nadenken over de betekenis van wat je hebt gehoord. Deze hypothese willen we in de toekomst verder uitwerken. We doen momenteel ook onderzoek naar de gevolgen van het verband tussen woordsoortinformatie en lettergreepstructuren voor Duitse kinderen met een gehoorstoornis. Als kinderen niet alle details van de lettergreep goed horen, kunnen ze misschien ook wel gehinderd zijn bij de verwerving van de zinsstructuur.

Referenties

- Don, J. & M. Erkelens (2006). "Vorm en Categorie". In: *Taal & Tongval*, 19, p. 40-54.
- Hayes, B. & C. Wilson (2007). "A maximum entropy model of phonotactics and phonotactic Learning". In: *Linguistic Inquiry*, 39 (3), p. 379-440.
- Keuleers, E., M. Brysbaert & B. New (2010). "SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles". In: *Behavior Research Methods*, 42 (3), p. 643-650.
- Shi, R., J.F. Werker & J. Morgan (1999). "Newborn infants' sensitivity to perceptual cues to lexical and grammatical words". In: *Cognition*, 72 (2), p. 11-21.
- Smith, J.L. (2011). 'Category-specific effects'. In: M. Van Oostendorp et al. (eds.). *The Blackwell Companion to Phonology*. NY: John Wiley & Sons Ltd.
- Trommelen, M. (1989). "Lettergreepstructuur en woordcategorie". In: *De Nieuwe Taalgids* 82, 64-85.
- Zec, D. (2005). "Prosodic differences among function words". In: *Phonology*, 22 (1), p. 77-112.