

# 'Leenwoordfonologie' in het Esperanto

Marc van Oostendorp, Universiteit van Amsterdam

Verschenen in Gramma 7.2:137-148.

## 0. Inleiding

Op veel fonologische generalisaties zijn uitzonderingen te bedenken. Soms worden deze uitzonderingen in voetnoten bij taalkundige publicaties afgedaan als 'leenwoorden' en niet verder behandeld, maar gezien de frequentie waarmee het probleem zich voordoet, kan dit nauwelijks een bevredigende oplossing worden genoemd.<sup>1</sup>

Een voorbeeld van dat soort uitzonderingen vinden we in het Nederlands. Het betreft de structuur van de lettergreep rond de sjwa (Van Oostendorp 1995). Normaal gesproken kunnen aan een sjwa in het Nederlands niet twee medeklinkers voorafgaan binnen dezelfde lettergreep (1a). Als er twee medeklinkers samen met een sjwa in een lettergreep staan, volgt een van die twee medeklinkers op de sjwa:

### (1) Franse leenwoorden in het Nederlands

- a. wat<sub>ɔ</sub>r, heng<sub>ɔ</sub>l, tunn<sub>ɔ</sub>l, ad<sub>ɔ</sub>m  
\*[watr<sub>ɔ</sub>], \*[h<sub>ɔ</sub>ŋl<sub>ɔ</sub>], \*[t<sub>ɥ</sub>nl<sub>ɔ</sub>], \*[adm<sub>ɔ</sub>]
- b. oeuvre, timbre

De uitzonderingen zijn hier enkele leenwoorden in het Frans (1b), waarin de sjwa wel degelijk wordt voorafgegaan door een complexe aanzet. In de discussie over de sjwa spelen de feiten in (1b) nauwelijks een rol (zie Van Oostendorp 1995 voor een overzicht van deze discussie). Deze uitzondering geldt ook niet voor alle leenwoorden: sommige daarvan hebben zich wel aan het Nederlands aangepast. *Papaver*, van het Franse *papavre*, is daar misschien een voorbeeld van. Het feit dat de woorden in (1b) allebei een 'leenfoneem' in de beklemtoonde klinkerpositie hebben houdt misschien verband met de uitzonderlijke structuur waarin de sjwa hier verschijnt. Hierbij moet ik aantekenen dat de klinkerkwaliteit in de Nederlandse woorden in (1b) waarschijnlijk fonetisch niet precies dezelfde is als die in hun Franse tegenhangers: deze woorden hebben zich dus op zijn minst aangepast aan de Nederlandse fonetiek. We kunnen daarom niet zeggen dat een spreker even helemaal overschakelt op het Frans als hij de woorden in (1b) uitspreekt.

Het is ook bekend dat de lexicons van veel natuurlijke talen onderverdeeld kunnen worden in zogeheten 'lexicale strata' met verschillend morfologisch en fonologisch gedrag. Gewoonlijk vormen deze strata een synchronische neerslag van historisch taalcontact. Het feit dat het Turks een Perzisch stratum heeft, en het Engels een Romaans stratum, maar geen omvangrijk Perzisch stratum, moet worden begrepen in het licht van de geschiedenis van de Turks- en Engelstalige gebieden.

Het is soms moeilijk om een grens te trekken tussen 'uitzonderlijke leenwoorden' aan de ene kant, en lexicale strata aan de andere. In informeel spraakgebruik zou je kunnen spreken over 'leenwoorden' als het aantal uitzonderingen klein is en over

---

<sup>1</sup> Fragmentjes van dit artikel verschenen in het Engels in mijn bijdrage aan Arends en Bruyn (1998). Ik dank Peter-Arno Coppen, Paul Hopkins en Wouter Kusters voor commentaar op een eerdere versie en discussies over dit onderwerp.

`strata' als dat aantal groter is, maar omdat `groot' en `klein' relatieve begrippen zijn is dit geen hard criterium.

Een ander criterium lijkt veelbelovender: meer dan één fonologisch proces is van toepassing binnen hetzelfde stratum. Woorden die gevoelig zijn voor fonologische regel R, zijn altijd ook gevoelig voor S, terwijl woorden die niet gevoelig zijn voor R ook S niet ondergaan. Nu doet zich echter de theoretische mogelijkheid voor dat op bepaalde woorden R noch S kan worden toegepast, enkel en alleen omdat het woord min of meer toevallig niet door de structurele descriptie van die regels gedekt wordt. Zo'n vorm kan echter, bijvoorbeeld om redenen van theoretische elegantie, toch tot het stratum gerekend worden, al voldoet het niet aan het formele criterium. Ook dit criterium is daarmee niet erg hard.

Het feit dat er lexicale uitzonderingen bestaan, dat deze uitzonderingen zich soms in hun gedrag lijken te groeperen, en dat regels naar uitzonderingen en klassen van uitzonderingen kunnen verwijzen, impliceert dat een adequate theorie over de synchrone grammatica en het synchrone lexicon ruimte moet bieden aan dergelijke uitzonderingen.

Gewoonlijk wordt aangenomen dat een moedertaalspreker over allerlei synchrone kennis over de woorden van zijn taal beschikt, maar niet over hun geschiedenis. In dit artikel laat ik nieuwe evidentie zien voor deze veronderstelling. Ik doe dit door enkele aspecten van het verschil tussen `leenwoordfonologie' en `inheemse' fonologie in het Esperanto te belichten.<sup>2</sup> In een etymologische zin zijn vrijwel alle Esperantowoorden `leenwoorden': de morfeeminventaris van de taal is voor het grootste deel afgeleid van bestaande talen, meestal van Indo-europese oorsprong. Slechts een paar procent van de Esperantowoorden is authentiek: deze zijn echt `bedacht' of hebben een onduidelijke oorsprong; voorbeelden zijn kial `waarom', tiam `zoveel' en edzo `echtgenoot'. Uit een strikt historisch oogpunt bekeken zijn dit dus de echte `inheemse' woorden van de taal. Ik heb echter geen enkele fonologische test kunnen vinden waardoor we deze woorden zouden kunnen onderscheiden van de andere woorden uit het Esperanto.

Ik hoop te laten zien dat er desondanks redenen zijn om verschil te maken tussen minstens twee categorieën van Esperantowoorden, met een verschillend fonologisch gedrag. De scheidslijn bevindt zich daarbij niet tussen `inheemse' of `uitheemse' woorden, maar tussen `alledaagse' en min of meer `geleerde' woorden. In andere talen loopt dit onderscheid min of meer parallel met dat tussen `inheemse' en `uitheemse' woorden. Leenwoorden worden vaak gebruikt voor min of meer onalledaagse concepten en zodra een woord alledaags wordt, past het zich snel aan de fonologie van de ontvangende taal aan. We kunnen het verschil tussen `geleerd' en `alledaags' bijvoorbeeld ook gebruiken om onderscheid te maken tussen de woorden in (1a) in (1b). Het bijzondere van de Esperantofeiten is dat ze ondubbelzinnig laten zien dat `geleerd' - `alledaags' het cruciale criterium is.

Een fonologische theorie die niet alleen wil verklaren hoe regelmatige gevallen en

---

<sup>2</sup> Het Esperanto is een goeddeels geplande taal; dat wil zeggen dat de oorsprong ervan duidelijk is aan te wijzen: deze ligt in een in 1887 door de Warschause oogarts Elejzer Zamenhof gepubliceerde brochure. Tegenwoordig heeft de taal naar schatting 1000 moedertaalsprekers die verspreid over de wereld wonen, met name in Europa, en een onbekend aantal tweedetaalsprekers.

uitzonderingen zich gedragen, maar die ook de manier wil beschrijven waarop in een taalsysteem besloten wordt of een bepaald woord zich uitzonderlijk kan gedragen, moet over een theoretisch apparaat beschikken dat een dergelijke verklaring mogelijk maakt. In dit artikel betoog ik dat betrekkelijk recente ideeën binnen de optimaliteitstheorie een dergelijk apparaat leveren: zij maken het mogelijk om het verschil in gedrag tussen 'geleerde' en 'alledaagse' woorden formeel te koppelen aan het verschil tussen 'formele' en 'informele' spraak. Ik betoog dat deze ideeën daarom te verkiezen zijn boven een aanpak die werkt met kenmerken zoals [+/- native] of [+/- geleerd]. Ik ga hierbij als volgt te werk. In de volgende paragraaf geef ik een kort overzicht van enkele theorieën over 'leenwoordfonologie' zoals deze binnen de (generatieve) fonologie ontwikkeld zijn. In paragraaf 2 geef ik enkele feiten uit het Esperanto die het volgens mij mogelijk maken verschil te maken tussen de verschillende theorieën. Paragraaf 3 geeft een conclusie.

### 1. Inheemse woorden en leenwoorden

In de door Chomsky en Halle (1968) ontwikkelde generatieve fonologie worden strata meestal verantwoord met behulp van een diakritisch morfologisch kenmerk, bijvoorbeeld [+/- native] (zie ook Zonneveld 1978 voor een verhelderende discussie). Fonologische en morfologische regels kunnen in deze visie verwijzen naar de waarden van dit kenmerk. Per definitie werkt deze benadering alleen voor strata, dat wil zeggen voor grotere groepen van woorden met een afwijkend fonologisch gedrag, zoals Perzische leenwoorden in het Turks.

Voor de grote differentiatie die er bestaat tussen min of meer aangepaste leenwoorden biedt een aanpak met kenmerken nauwelijks een oplossing. In de uitspraak van veel Nederlanders heeft het eerste segment van het woord *garage* zich aangepast aan het Nederlandse klanksysteem. Het wordt daarom als een [ɣ] of een [g] uitgesproken. Voor een deel van deze sprekers geldt datzelfde niet voor woorden als *goal* en *gate*, die een [g] krijgen. We kunnen nu echter er niet mee volstaan *garage* [+ native] te verklaren en *goal* [- native], want ook het eerste woord heeft zich nog niet op alle manieren aangepast aan de fonologie van het Nederlands. Dat blijkt bijvoorbeeld uit de uitspraak van de tweede <g> als een [ɣ], eveneens een 'leenfonem' in het Nederlands. Een uitgebreider kenmerksysteem is hier dus nodig, en het gevaar bestaat dat we op deze manier eindigen met een grote verzameling kenmerken die een even grote verzameling 'strata' definiëren met ieder slechts een of twee woorden.

Een tweede probleem is dat sommige generalisaties helemaal niet in dit mechanisme gevat kunnen worden. We vinden bijvoorbeeld meestal complexere structuren -- complexere lettergrepen, complexere segmenten ('leenfonemen'), complexere klemtoonpatronen ('lexicale klemtoon') -- in leenwoorden dan in inheemse morfemen. De laatste klasse is met andere woorden gewoonlijk restrictiever in haar fonotaxis dan de eerste. Als dit een geldige generalisatie is, moet ze uit de theorie volgen, maar in een theorie die uitgaat van het kenmerk [+/- native] is dit niet het geval. In zo'n theorie kan het inheemse stratum in een bepaalde taal best juist ingewikkelder lettergrepen hebben dan uitheemse strata.

De optimaliteitstheorie geeft ons nu de middelen om het soort uitzonderingen dat leenwoorden op fonotactische generalisaties geven beter te benaderen. Zoals

bekend is er in deze theorie een universele verzameling met fonologische condities, CON. De verschillen tussen talen worden enkel en alleen beschreven als verschillen in de ordening van deze condities.

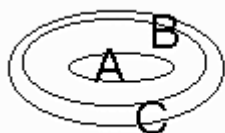
Van belang is daarbij vooral het onderscheid dat wel gemaakt wordt tussen twee soorten condities: welgevormdheidscondities en condities op respect.

Welgevormdheidscondities maken onderscheid tussen 'welgevormde' en 'niet-welgevormde' oppervlaktestructuren. ONSET van Prince en Smolensky (1993) is een voorbeeld van een welgevormdheidsconditie. Deze conditie bepaalt dat alle lettergrepen een aanzet moeten hebben; dat wil zeggen dat ze moeten beginnen met een medeklinker. We hoeven alleen te bezien of alle lettergrepen in de oppervlaktevorm een aanzet hebben om na te kunnen gaan of ze aan deze conditie al dan niet voldoen. De invoer is hierbij volkomen irrelevant.

Respectcondities ('faithfulness constraints') vergelijken invoer met uitvoer. In essentie stellen zij dat de uitvoer gelijk moet zijn aan de invoer: er mag geen materiaal verwijderd worden, en er mag geen materiaal worden toegevoegd. In een op de optimaliteitstheorie gebaseerde analyse van taalverandering hebben we welgevormdheidscondities nodig om te verklaren waarom vormen soms veranderen; we hebben respectcondities nodig om te verklaren waarom niet alle woorden in alle talen uitgesproken worden in een 'optimale' vorm als [tata] of [bi] (Chomsky 1995).

Volgens sommige aanhangers van de optimaliteitstheorie moet het verschil tussen strata uiteindelijk een verschil tussen grammatica's zijn. Itô en Mester (1995) stellen bijvoorbeeld dat de grammatica's van inheemse en uitheemse strata op slechts een manier kunnen verschillen: welgevormdheidscondities zijn altijd hoger geordend in de eerste dan in de tweede. Op deze manier krijgen we een gelaagde stratumstructuur:

(2)



De welgevormdheidscondities zijn het sterkst in het inheemse stratum A. Omdat welgevormdheidscondities een voorkeur hebben voor eenvoudige segmenten, lettergrepen en voeten, kunnen alleen deze aan de oppervlakte komen. In de strata B en C zijn steeds complexere structuren toegestaan. In deze strata wordt het belangrijker om de invoer te respecteren dan om een welgevormde fonologische structuur te krijgen.<sup>3</sup> Een van de interessante aspecten van het voorstel van Itô and

<sup>3</sup> Het gaat hier alleen om strata en leenwoorden *binnen* een taalsysteem. Codewisseling moet op een andere manier behandeld worden. Ervan uitgaande dat Turkse woorden klinkerharmonie ondergaan en Nederlandse woorden klinkerreductie, moeten we concluderen dat een spreker die van het Turks naar het Nederlands overstapt, overstapt tussen verschillende soorten welgevormdheidseisen. Overigens is de grens soms lastig te trekken, want men zou kunnen beargumenteren dat de stap van 'Nederlands' *water* naar 'Frans' *oeuvre*

Mester is dat dit het ons mogelijk maakt te voorspellen welk van twee gegeven strata in een verder onbekende taal inheems is en welk uitheems. Itô and Mester's theorie is aan de andere kant ook weer duidelijk ingesteld op strata; individuele leenwoorden zijn moeilijker te vangen. Waarschijnlijk komt het woord een lexicon binnen via de buitenste rand en dringt het langzaam maar zeker door in de binnenste ringen, naarmate het langer in de taal blijft. De manier waarop dit precies werkt, is echter niet uitgewerkt.

Er is een voor de hand liggend alternatief dat op een subtiële manier verschilt van het voorstel van Itô and Mester. We kunnen zeggen dat een taal slechts één grammatica, dat wil zeggen slechts één ordening van condities, heeft, die hetzelfde is voor alle woorden in het lexicon, onafhankelijk van de vraag of het inheemse woorden zijn, woorden in een uitheems stratum, of leenwoorden. In deze unieke ordening staan de respectcondities relatief hoog -- ruwweg op hetzelfde niveau waarop ze bij Itô and Mester staan voor de buitenste schil van leenwoorden. De verschillen tussen strata zijn dus geen verschillen in conditieordering. Het verschil is wél de hoeveelheid en de aard van de structuur die gespecificeerd is: in leenwoorden is veel gemarkeerde structuur aangebracht die in inheemse woorden in de loop der tijd om zo te zeggen geërodeerd is.

We moeten dan wel begrijpen hoe deze lexicale erosie precies werkt. Als een woord een complexe structuur heeft, moet deze structuur aan de oppervlakte komen omwille van de respectcondities. Maar in zo'n geval zal de vorm vrijwel per definitie een aantal welgevormdheidscondities schenden. In de loop der tijd kan zo'n vorm dan proberen wat van het aanstootgevende onderliggende materiaal kwijt te raken, zo dat het op een dusdanige wijze aan de oppervlakte kan komen dat het aan zowel welgevormdheidscondities als respectcondities kan voldoen. Een vorm die onderliggend al welgevormd is komt veel minder geschonden aan de oppervlakte dan een vorm die dit niet is. In elke taal is dus permanent een proces van 'lexicale optimalisatie' aan de gang in de zin van Prince en Smolensky (1993). Dit proces kan geïnterpreteerd worden als een beweging in de richting van vereenvoudiging van de onderliggende structuren.

Het is moeilijk om goede empirische tests te bedenken die verschil zouden kunnen maken tussen deze twee analyses. De hier gepresenteerde interpretatie heeft in ieder geval een aantal conceptuele voordelen boven die van Itô and Mester. In de eerste plaats kan het klassen van uitzonderingen beter verklaren, zonder hiervoor het theoretisch construct van strata te postulieren. Het is bijvoorbeeld ongewenst om een aparte grammatica te moeten opstellen voor twee uitzonderingen in een taal. In de tweede plaats biedt de zojuist geschetste interpretatie aanknopingspunten voor een verklaring waarom sommige woorden zich sneller inheems gedragen dan andere.

Een factor die hierbij waarschijnlijk een rol speelt is gebruiksfrequentie: hoe vaker een woord gebruik wordt, des te vaker wordt het door de verzameling condities geëvalueerd, en des te eerder zal het dus aan lexicale erosie onderhevig zijn. Een andere factor die een rol speelt is waarschijnlijk het register waarin een woord staat. In Van Oostendorp (1997) heb ik beargumenteerd dat het verschil tussen formelere en informelere registers zich laat karakteriseren met de ordening van

---

(in *waterverfoevre*) ook een stap is van de ene vorm van welgevormdheid naar de andere.

respectcondities: hoe formeler het register, des te hoger de respectcondities geordend zijn. Het ligt dan ook voor de hand dat de woorden in de `formelere' registers -- de zogeheten `moeilijke' of `geleerde' woorden -- minder snel aan erosie onderhevig zullen zijn dan die in de meer informele registers.

Er zijn nu nog veel vragen onbeantwoord en zoals gezegd is het moeilijk een strikt empirisch verschil te maken tussen de twee theorieën. In ieder geval bieden zowel het voorstel van Itô and Mester (1995) als het zojuist geschetste alternatief aanknopingspunten voor een theorie over het uitzonderlijk gedrag van `geleerde' leenwoorden; een theorie die niet verwijst naar de taalgeschiedenis. Het kan daarom zelfs van toepassing worden verklaard op gevallen waar de diachronie van de talen in kwestie onduidelijk of -- in het uiterste geval -- afwezig is. Dit is nuttig in het geval van een taal als het Esperanto. Zoals ik hierboven al heb opgemerkt is het immers betrekkelijk zinloos om onderscheid te maken tussen inheemse en uitheemse woorden in deze taal: historisch gezien zijn vrijwel alle woorden uitheems. Een analyse volgens Chomsky en Halle (1968) is voor feiten als deze minder aantrekkelijk. Een verdeling van het lexicon met morfologische kenmerken is per definitie arbitrair is daardoor niet in staat de parallellie tussen de dichotomieën geleerd-alledaags en formeel-informeel op een adequate manier te beschrijven.

## 2. Leenwoordfonologie in het Esperanto

De lettergreepstructuur van het Esperanto is vrijwel even gecompliceerd als die van bijvoorbeeld het Nederlands of het Engels:

- (3) a. trin-ki `drinken', on-klo `oom', bran-cho `tak'  
b. a-ta-ki `aanvallen', o-be-i `gehoorzamen', ka-o-so `chaos'  
c. skri-bi `schrijven', shtrum-pon `sok-ACC'

De meeste grafemen in deze voorbeelden corresponderen met IPA-symbolen: de digraaf <ch> geeft echter een stemloze palatale affricaat weer, en de digraaf <sh> een palatale fricatief. De eerste lettergreep in elk van de drie woorden in (3a) is gesloten en de tweede open. In (3b) zien we verder daar ook lettergrepen zonder aanzet zijn toegestaan, zowel aan het begin als binnenin het woord. De woorden in (3c) tonen dat we tamelijk ingewikkelde aanzetclusters hebben als het eerste segment een coronale fricatief is. Alles bij elkaar kunnen we dus het volgende templaaf opstellen voor de lettergreep van het Esperanto, waarin facultatieve elementen tussen haakjes staan:<sup>4</sup>

- (3) (s/sh) (C<sub>i</sub>)(C<sub>j</sub>) V (C<sub>k</sub>)

De nucleus van de lettergreep kan in het Esperanto alleen een klinker zijn. Syllabische consonanten kent de taal niet. C<sub>i</sub> en C<sub>j</sub> moeten in een stijgende sonoriteitscurve staan, en C<sub>k</sub> mag alleen een niet-coronale obstruent zijn als hij gevolgd wordt door een andere obstruent in hetzelfde woord. (Woorden mogen niet in een obstruent eindigen, maar ze kunnen eindigen op een nasaal, een liquida of een stemloze coronale fricatief.)

Dit alles mag nauwelijks een verrassing heten. Gezien het feit dat alle woorden in het Esperanto geleend zijn uit andere talen, mogen we verwachten dat de eisen op

---

<sup>4</sup> Een uitgebreidere beschrijving van de lettergreepstructuur van het Esperanto is te vinden in Van Oostendorp (1999).

lettergreepstructuur nogal ontspannen zijn. Bovendien zijn de meeste morfemen van Indo-europese oorsprong en de talen uit deze familie hebben zelf een structuur die min of meer overeenkomt met het templaat in (3).

Daarbij moet dan wel worden opgemerkt dat zelfs een oppervlakkige studie van enkele andere plantalen laat zien dat deze twee factoren op zich niet voldoende zijn om een taal een complexe lettergreepstructuur te bezorgen. De taal Volapük (in 1879 gecreëerd door J.M. Schleyer) heeft bijvoorbeeld een veel beperktere lettergreepstructuur. Alle woorden in deze taal hebben de vorm (CV)\*C; alle lettergrepen hebben een enkelvoudige aanzet en zijn open, behalve de laatste lettergreep van het woord die wel een enkelvoudige aanzet heeft, maar gesloten is. Op deze manier verandert bijvoorbeeld het Engelse compositum 'world speech' in 'volapük'. Maar ook in het Volapük zijn alle morfemen geleend uit Indo-europese talen, zij het dat de meeste natuurlijk moeilijk te herkennen zijn nadat ze aan het systeem zijn aangepast.

Het systeem dat ten grondslag ligt aan het Volapük kan goed dienen om de wijze waarop de standaardwelgevormdheidscondities op lettergreepstructuur binnen de optimaliteitstheorie werken te illustreren. Er zijn condities die zeggen dat lettergrepen een aanzet moeten hebben en dat alle lettergrepen een nucleus hebben met een klinker erin. Er is ook een conditie die eist dat de aanzet van de lettergreep niet complex mag zijn. Dan is er ook nog een conditie die lettergreepcoda's verbiedt en een conditie die zegt dat elk woord moet eindigen in een medeklinker.<sup>5</sup> De laatste twee condities conflicteren: ze kunnen niet allebei tegelijk waar zijn voor de laatste lettergreep van het woord. In dat geval wint de laatstgenoemde conditie in het Volapük. Al deze condities zijn aanwezig in het werk van Prince en Smolensky (1993); het enige bijzondere aan het Volapük is dat in deze taal alle condities op respect gedomineerd worden door welgevormdheidscondities.<sup>6</sup>

---


<sup>5</sup> De laatste conditie lijkt misschien enigszins *ad hoc*, maar hij wordt wel degelijk door Prince en Smolensky (1993) voorgesteld voor de beschrijving van de fonologie van natuurlijke talen.

<sup>6</sup> Ik vereenvoudig de analyse enigszins door aan te nemen dat de [a] van *volapük* het resultaat van fonologische epenthese is, en niet een apart morfeem. Deze aanname is controversieel en waarschijnlijk onjuist; waarschijnlijk is de [a] in ieder geval eerder een soort genitiefmarkeerder. Ik ga hier ook niet verder in op de fonologische veranderingen /w/>[v], /i/>[ü] of de keuze welke consonant aan de oppervlakte moet komen aan het eind van /world/> [vol].

Verder suggereert (6) misschien dat ik aanneem dat de 'oorspronkelijke' Engelse vormen altijd onderliggend zijn voor de spreker van het Volapük. Dit is echter niet juist. Een standaardaanname binnen de optimaliteitstheorie is *Rijkdom van de Basis* (Prince en Smolensky 1993): de grammatica van elke taal moet een in die taal welgevormde output geven voor *elke* logisch denkbare input, al is het maar door het grootste deel van die input weg te gooien. In het onderhavige geval dient (6) er dus toe om te illustreren hoe een willekeurige input -- bijvoorbeeld een Engels woord -- door de grammatica op een dusdanige wijze wordt geamendeerd dat het een output oplevert die in het rigide templaat van Volapükwoorden past.

- (5) ONSET: Alle lettergrepen moeten een aanzet hebben  
 NUCLEUS: Alle lettergrepen moeten een nucleus hebben.  
 \*COMPLEX: CC-clusters zijn niet toegestaan.  
 NOCODA: Consonants mogen niet in de coda staan  
 FINALC: Elk woord moet eindigen op een consonant.

(6)

/world spik /	ONSET	NUCLEUS	*COMPLEX	FINALC	NOCODA	Faith
a. vorld-spük			*! **		*	
b. vor-l-spük		*!			**	d
c. vo-la-spük			*!		*	rda
d.  vo-la-pük					*	rds <u>a</u>
e. vo-la-pü				*!		rds <u>ka</u>
f. o-la-pük	*!				*	vrds <u>a</u>
g. volük					*	rdsp! <sup>7</sup>

De vormen in (6a-c) laten zien dat we consonantdeletie en klinkerepenthese nodig hebben om een welgevormde uitvoer af te leiden. (6f) toont dat de toepassing van deletie en epenthese door dezelfde condities ook weer wordt ingeperkt. (6g) laat zien dat 'onnodige' deleties geblokkeerd worden. Door (6d) te vergelijken met (6e) zien we hoe twee welgevormdheidscondities met elkaar in interactie treden.

In het Esperanto is het templaar voor welgevormde woorden niet zo strikt; complexe aanzetten, lettergrepen zonder aanzet en woordinterne lettergrepen met een aanzet zijn in deze taal toegestaan. We kunnen dit beschrijven door aan te nemen dat de condities op respect in deze taal een stuk hoger geordend zijn dan de hier behandelde welgevormdheidscondities.<sup>8</sup> De enige welgevormdheidsconditie in de lijst in (4) die in het Esperanto nooit geschonden wordt is NUCLEUS: alle lettergrepen hebben in deze taal op zijn minst één klinker. Toch zijn er nog steeds bepaalde fonologische beperkingen op de vorm van de woorden aan te wijzen die niet onmiddellijk in verband kunnen worden gebracht met de beperkingen die de brontalen opleggen. Een belangrijke restrictie is bijvoorbeeld dat lange klinkers en gegemineerde medeklinkers niet zijn toegestaan.<sup>9</sup> Als woorden met geminaten

<sup>7</sup> Het is kennelijk slechter om een consonant te verwijderen dan om een klinker in te voegen. We kunnen dit effect verkrijgen door de respectcondities tegen deletie te scheiden van die tegen insertie, zoals in de optimaliteitstheorie gebruikelijk is.

<sup>8</sup> Het Volapük is misschien nooit ver genoeg ontwikkeld om een klasse van 'geleerde' woorden of 'leenwoorden' te laten ontstaan. We zouden kunnen speculeren dat de relatief lage ordening van de respectcondities hiermee samenhangt.

<sup>9</sup> Mogelijk worden beklemtoonde klinkers fonetisch iets verlengd (klemtoon ligt op de

geleend worden, worden deze verkort. Dit zien we bijvoorbeeld in vormen als *betulo* `berk' (< Lat. *betulla*, It. *betulla*)<sup>10</sup>, *peki* `zondigen' (< Lat. *peccare*, It. *peccare*).

Dit punt is iets moeilijker te maken in het geval van de klinkers, omdat het niet gemakkelijk is om woorden te vinden waarin alle mogelijke brontalen onmiskenbaar een lange klinker hebben. Een voorbeeld is misschien het woord voor *boot*. Het morfeem voor dit begrip is afgeleid van de Germaanse talen, waarin het een lange klinker heeft (Eng. *boat*, Du. *Boot*, Nl. *boot*). Het woord in het Esperanto is *boato* /bo-a-to/. De oorspronkelijke lange klinker is in dit geval over twee lettergrepen verdeeld, waarbij de uitspraak van de tweede lettergreep geïnspireerd lijkt door de spelling van het Engelse woord. En zelfs als voorbeelden als deze niet voldoende overtuigen, kunnen we in ieder geval observeren dat er eenvoudigweg geen woorden in het Esperantolexicon zijn met een lange klinker: er is bijvoorbeeld geen minimaal paar /bo-to/ - /boo-to/.

Alles bijeengenomen kunnen we nu de volgende conditieschikking opstellen: `NOLONGSEGMENTS >> FAITH`. Er is dus op zijn minst één plaats waar de kern van het Esperantolexicon aan strengere wetten gehoorzaamt dan welke gelden voor andere talen.

Een vergelijkbaar punt is de toelaatbaarheid van zogenaamde `superzware lettergrepen'. In het Latijn en de Germaanse talen kunnen woorden lettergrepen bevatten die op meer dan twee medeklinkers eindigen. Zulke lettergrepen staan dan meestal aan het eind van het woord. Alle Esperantowortels worden gevolgd door een klinker die de woordcategorie bepaalt<sup>11</sup>. Op die manier zijn veel van de oorspronkelijke superzware lettergrepen verdwenen: het Duitse *Bank* is bijvoorbeeld veranderd in *ban-ko*. We kunnen hieruit echter geen conclusies trekken omdat de -o om morfologische redenen is ingevoegd en niet om fonologische. Toch zijn er in sommige talen woorden die eindigen op lettergrepen die nog langer zijn dan superzwaar. In dergelijke gevallen zien we wel degelijk aan het licht komen dat het Esperanto te zware lettergrepen probeert te vermijden.

Een voorbeeld hiervan is het Esperantowoord voor *korps* (En. *corps*, Du. *Korps*, Fr. *corps*). Dit woord is niet *\*korp-so* maar *kor-pu-so*. De reden dat we hier een -u- vinden en niet een andere klinker, is natuurlijk dat het Latijnse woord waarvan de woorden in de Europese talen afgeleid zijn *corpus* is. Toch is deze -u- in geen enkele (andere) brontaal van het Esperanto terug te vinden. Bovendien hebben de meeste talen nog een ander woord dat is afgeleid van dezelfde Latijnse bron, namelijk het woord

---

voorlaatste lettergreep in het woord), maar zelfs als dat het geval is, speelt deze lengte geen rol in de fonologie: ze is niet distinctief.

<sup>10</sup> De dubbele medeklinkers corresponderen in ieder geval in het Italiaans met lange medeklinkers en de studie van de Latijnse klemtoon geeft ons alle aanleiding om aan te nemen dat dit in het Latijn ook geldt. Dit woord heeft een korte /l/ in het Portugees (*bétula*), maar het is niet waarschijnlijk dat deze taal hier enige invloed heeft gehad. De /o/ in de Esperantovorm is het gevolg van de morfologie van deze taal waarin alle zelfstandig naamwoorden op een -o eindigen.

<sup>11</sup> Alle nomina eindigen op -o, alle adjectieven op -a, alle verba op een -i in de infinitief, -as in de tegenwoordige tijd, -is in het praeteritum en -os in het futurum. Tegenwoordige deelwoorden hebben -ant, verleden deelwoorden -int, toekomstige deelwoorden -ont; voor het passief bestaat er een parallel rijtje -at, -it, -ot.

voor een organisch lichaam of een lijk (En. *corpse*, Du. *Körper*, Fr. *corps*). Waarschijnlijk om homoniemen te vermijden is in het Esperanto hier een ander vorm gekozen, maar weer is deze vorm niet \**korps*. Deze keer is in plaats daarvan gekozen voor *korpo*. In beide gevallen is er dus een fonotactische reden waarom precies deze keus gemaakt is. Het belangrijkste punt is dat vormen als \**korps* bij de planning van de taal vrij consistent vermeden zijn. Ook in deze gevallen lijken de respectcondities minder sterk te zijn dan welgevormdheidscondities.

Op beide hiervoor gemaakte generalisaties nu kunnen we uitzonderingen vinden. In de eerste plaats is er een klein aantal woorden dat bij uitzondering een lange medeklinker bevat: *getto* (idem) en *finno* `fin'. Het laatste woord vormt een minimaal paar met *fino* `einde'. Het aantal uitzonderingen is klein; de zojuist gegeven voorbeelden lijken de enige, en dan kan *getto* ook nog gerealiseerd worden als *geto* en wordt in plaats van *finno* ook wel *suomo* gebruikt. Onder de analyse van Itô en Mester (1995) zouden we zeggen dat respect in deze gevallen belangrijker is dan welgevormdheid. Onder de hierboven gepresenteerde analyse zouden we ook kunnen zeggen dat deze woorden als enige onderliggend een lange medeklinker hebben en dat respect in het hele Esperantolexicon een rol speelt.

De conditie tegen superzwarte lettergrepen kent ook een paar uitzonderingen: *interpunkcio*, *arkta* (arctisch), *planktono*<sup>12</sup> en natuurlijk hebben we hier dezelfde oplossing tot onze beschikking. Hierbij valt dan nog aan te tekenen dat in dit geval de meeste woorden een tamelijk gespecialiseerd en `geleerd' karakter hebben, zoals we in het geval van *strata* ook zouden verwachten, zoals ik hierboven uiteengezet heb.

In Van Oostendorp (1998) wordt opgemerkt dat bepaalde woorden verkort worden om redenen van `representatieve zuinigheid'. Het woord *vibracio* `trilling' kan bijvoorbeeld verkort worden tot *vibro*. De sequentie *-aci-* kan geanalyseerd worden als een overbodige nominalisatiemarkeerder, omdat de finale *-o* de nominale categorie al markeert. Toch ondergaan lang niet alle vormen altijd dit soort verkorting. De informanten van Van Oostendorp (1998) rapporteerden voor veel van de paren van lange en verkorte vormen dat ze vermoedden dat de langere vormen een gespecialiseerde en `geleerde' betekenis hadden. *Vibro* betekende dan 'trilling' in een alledaagse context, terwijl *vibracio* een gespecialiseerd soort vibratie is, zoals beschreven in een handboek voor de natuurkunde. `Moeilijke' woorden krijgen dus een `moeilijke' fonologie. Misschien zien we hier de eerste tekenen van het ontstaan van een stratum, in plaats van een handjevol geïsoleerde leenwoorden. Maar zelfs in dat geval is het mogelijk om de verkorte vorm een `inheemse' ingang in het lexicon te geven, zodat we geen verkorting hoeven te postuleren.

Een laatste vorm van uitzonderlijke lettergreepstructuur betreft de aanzet. Er zijn aanvullende restricties op de sequentie  $C_i C_j$  in (4). Bijvoorbeeld moet er sprake zijn van een voldoende steile sonoriteitscurve over de twee medeklinkers. De `beste' complexe aanzet is er een waarin de eerste medeklinker een plosief is en de tweede een liquida. Bij nadere beschouwing vinden we echter een gat in dit systeem van `beste' aanzetclusters. Hoewel [pr] (*problemo*), [gr] (*granda*), [kl] (*klara*), [bl] (*blua*), [tr]

---

<sup>12</sup> Al deze voorbeelden bevatten een /k/ als de `extra' medeklinker, met een volgende coronale obstruent. Meer onderzoek is nodig om te zien of dit een toevallig resultaat is van de weinig systematische zoekmethode die ten grondslag ligt aan deze observatie.

(*trajno* 'trein') en *dr* (*droni* 'verdrinken') heel gewone woorden zijn, ontbreken de clusters [tl] en [dl]. Dit is natuurlijk naar alle waarschijnlijkheid een overgeërfd eigenschap van de brontalen van het Esperanto, waar dergelijke clusters ook ontbreken of erg zeldzaam zijn. In al deze talen zou de restrictie een bijzonder geval kunnen zijn van de zogenaamde *Obligatory Contour Principle*, het principe dat adjacentie van gelijkkluidende elementen verbiedt: twee coronale consonanten mogen nooit naast elkaar staan. In Waringhien e.a. (1987) vinden we één tegenvoorbeeld tegen deze generalisatie: *tlaspə*. Dit woord betekent 'thlaspi' en is zonder twijfel te classificeren als een 'geleerd' woord.

### 3. Conclusie

In dit artikel heb ik evidentie aangedragen dat het nuttig is om 'regelmatige' en 'onregelmatige' woorden te onderscheiden in het Esperanto. Zelfs in een taal als deze -- waarin vrijwel alle woorden uit andere talen stammen en waarin per definitie geen enkel woord ouder is dan ongeveer 110 jaar -- blijkt het mogelijk onderscheid te maken tussen inheemse en uitheemse woorden. Er is nog niet veel evidentie dat de leenwoorden in een apart stratum georganiseerd zijn. Dit laat echter onverlet dat dergelijke strata wel degelijk zouden kunnen gaan bestaan: vooral als de eisen op het inheemse lexicon sterker worden, zouden sommige fonologische eigenaardigheden heel wel kunnen beginnen te clusteren.

Het geval van het Esperanto is interessant omdat het aantoont dat het niet zozeer de oorsprong van de woorden is die hun fonologisch afwijkende gedrag verklaart, maar eerder hun gespecialiseerde karakter. Omdat een optimaliteitsanalyse die uitgaat van de hier beschreven gemodificeerde versie van Itô en Mester (1995) het verband tussen uitzonderlijke woorden en stijlniveaus kan beschrijven, terwijl kenmerktheorieën zoals Chomsky en Halle (1968) hier in de problemen raken, geven de Esperantofeiten een argument voor het eerste soort analyse en tegen het tweede.

### Bibliografie

- Arends, Jacques en Adriëne Bruyn (red.) (1998) *Mengelwerk voor Muysken*. Amsterdam, Universiteit van Amsterdam.
- Bastien, Louis (1950) *Naŭlingva etimologia leksikono* [Negentalig etymologisch lexicon]. Leicester, The Esperanto Publishing Company Ltd.
- Chomsky, N. (1995) *The Minimalist Program*. Cambridge, Mass: MIT Press.
- Chomsky, N. en M. Halle (1968) *The Sound Pattern of English*. Cambridge, Mass.: The MIT Press.
- Itô, J. & A. Mester (1995) 'Japanese Phonology.' In: J. Goldsmith, ed., *Handbook of Phonological Theory*. Basil Blackwell, Oxford. pp. 817-838.
- Oostendorp, M. van (1995) *Vowel quality and phonological projection*. Dissertatie, Katholieke Universiteit Brabant.
- Oostendorp, M. van (1997) 'Style levels in conflict resolution,' in: F. Hinskens, R. van Hout, L. Wetzels, red., *Language variation and linguistic theory*. John Benjamins, Amsterdam.
- Oostendorp, M. van (1998) 'Economy of representation in the Esperanto word', in: R.

van Bezooijen and R. Kager, eds., *Linguistics in the Netherlands 1998*. Amsterdam, Benjamins, 175-187.

Oostendorp, M. van (1999) 'Syllable structure in Esperanto as an instantiation of universal phonology'. To appear in *Esperantologio/Esperanto Studies*.

<http://www.math.uu.se/esperanto/>

Prince, A. and P. Smolensky (1993) 'Optimality Theory; Constraint Internaction and Satisfaction in Generative Grammar.' Manuscript, Rutgers University & University of Colorado, Boulder.

Waringhien, G. et al. (1987) *Plena ilustrita vortaro de Esperanto*. [Volledig, geïllustreerd woordenboek van het Esperanto] Paris, Sennacieca Asocio Tutmonda.

Zonneveld, W. (1978) *A formal theory of exceptions*. Lisse, Peter de Ridder.