

Inhoudsopgave

Inleiding	2
1. Een Friestalige spraaksynthesizer	3
1.1 Het Fries	3
1.2 De mogelijkheden voor een Friestalige spraaksynthesizer	3
2. Festival	5
2.1 Festival in het kort	5
2.2 Hoe is Festival opgebouwd?	5
3. MBROLA	8
3.1 Wat is MBROLA?	8
3.2 Gebruik van MBROLA	8
4. Het bouwen van een nieuwe stem	10
4.1 Ingrediënten voor het bouwen van een Friestalige stem in een spraaksynthesizer	10
4.1.1 Friese foneemset	10
4.1.2 Overige benodigdheden	12
4.2 Werkwijze	14
4.2.1 Algemeen	14
4.2.2 Opnames	14
4.2.3 Aanleggen van een difoonset	15
4.3 Het omzetten van de bestanden	15
4.3.1 Bouw van de stem (an_mbrola.scm)	15
4.3.2 Foneemset (mrpa_phones.scm)	16
4.3.3 Lexicon (frislex.scm)	16
4.3.4 Tokenisatie (fristoken.scm)	18
4.3.5 Intonatie en duur (frisint.scm, mrpa_durs.scm, gswdurtreeZ.scm)	18
4.3.6 Koppeling naar MBROLA (mbrola.scm)	19
5. Evaluatie van de spraaksynthesizer	20
6. Conclusie en suggesties	22
Geraadpleegde literatuur	23

Bijlagen

bijlage 1:	Worldbet voor het Fries
bijlage 2:	an_mbrola.scm
bijlage 3:	frislex.scm
bijlage 4:	fristoken.scm
bijlage 5:	frisint.scm
bijlage 6:	mrpa_phones.scm
bijlage 7:	mbrola.scm
bijlage 8:	mrpa_durs.scm
bijlage 9:	gswdurtreeZ.scm
bijlage 10:	anmrpa
bijlage 11:	temp.pho
bijlage 12:	opbouw directories

Inleiding

Voor het vak 'spraaktechnologie' van de opleiding 'Sprakcommunicatie en spraaktechnologie' aan de UvA te Amsterdam schrijf ik een verslag over het ontwikkelen van een spraaksynthesizer voor het Fries. Het idee is ontwikkeld vanuit mijn interesse voor deze minderheidstaal. Het is namelijk mijn moedertaal.

Op de Fryske Akademy, het wetenschappelijk onderzoekscentrum voor het Fries, is men momenteel bezig met het verzamelen van een Korpus Sprutsen Frysk (Corpus Gesproken Fries) waarna men nog plannen heeft voor het ontwikkelen van een Friestalige spraakherkenner. Het ontwikkelen van een Friestalige spraaksynthesizer lijkt mij een interessante aanvulling, al is het alleen al op papier. Hierbij wordt het multilinguale spraaksysteem van Festival als basis gebruikt, aangezien het relatief eenvoudig is om modules van dit systeem te veranderen, zonder dat men een geheel nieuw systeem moet bouwen.

In dit werkstuk wordt eerst informatie gegeven over het Fries in het algemeen en worden enkele toepassingen van een Friestalige spraaksynthesizer gegeven. Hierna wordt kort de bouw van het Festival-systeem besproken, waarna wordt ingegaan op de benodigheden voor een nieuwe stem. De werkwijze om deze nieuwe stem te bouwen wordt gepresenteerd en er worden mogelijkheden voor evaluatie aangereikt. Tot slot worden een conclusie en suggesties voor verder onderzoek gegeven.

1. Een Friestalige spraaksynthesizer

1.1 Het Fries

Met het Fries wordt in dit werkstuk het Westerlauwers Fries bedoeld, voornamelijk gesproken in de provincie Fryslân. Fries is erkend als tweede rijkstaal van Nederland. Fryslân is een tweetalige provincie; er wordt daar door de meeste inwoners zowel Nederlands als Fries gesproken. In de praktijk blijkt dat alle inwoners het Nederlands op alle modaliteiten in meer of mindere mate beheersen. Dit komt onder andere doordat Nederlands als onderwijstaal gebruikt wordt. Uit het meest recente taalonderzoek uit 1994 bleek dat 74% van de populatie Fries spreekt. Van dit aantal spreekt een groot gedeelte het Fries als tweede taal. Meer dan de helft van de populatie gebruikt het Fries als thuistaal. Verder verstaat 94% van de populatie het Fries, 65% kan het Fries lezen en 17% kan het schrijven (Gorter e.a. (1995):8). Het Fries wordt het meest gesproken op het platteland (Gorter (1997):1154).

Wanneer men het Fries vergelijkt met het Nederlands, ziet men dat het Fries fonologisch gezien meer vocalen heeft dan Nederlands. Ook zijn er duidelijk meer diftongen aanwezig. Op morfo-syntactisch niveau vinden we alleen verschillen in de affixen, diminutieven en in de volgorde van werkwoorden in bijzinnen en infinitieve constructies. Het Fries en het Nederlands verschillen het meest op lexicaal niveau. Echter, op dit niveau vinden we ook de meeste interferenties. Nederlandse woorden worden vaak in het Fries opgenomen om nieuwe concepten aan te duiden of juist ter vervanging van oudere Friese woorden (Gorter (1997):1156). Er is een groot verschil tussen het gesproken Fries en het geschreven Fries. De schrijftaal is conservatiever en staat dichter bij de taalnormen.

Ook het Fries kent allerlei dialecten. Onderling zijn deze dialecten echter goed verstaanbaar. Er zijn drie hoofddialecten: het Klaaifrysk (Kleifries), Wâldfrysk (Woudfries) en het Súd-Westhoeks (Zuidwesthoeks). In het noordoosten van de provincie, rondom Kollum, wordt Noard-Easthoeks (Noordoosthoeks) gesproken. Deze naam omvat diverse dialecten uit die streek en men vindt er Saksische invloeden in terug.

Het Klaaifrysk wordt, ruwweg gezien, gesproken in de (noord)westelijke helft van de provincie. Het Wâldfrysk, gesproken in het oostelijke gedeelte van de provincie, onderscheidt zich door de typische uitspraak van pronouns: 'hy' (/hEi/) en 'my' (/mEi/) als respectievelijk /hi/ en /mi/. Het Súd-Westhoeks, dat het zuidelijke gedeelte van de provincie voor zich neemt, onderscheidt zich grotendeels door de afwezigheid van het fonologische proces dat 'breking' (zie 'stijgende diftongen' paragraaf 4.1.1.) wordt genoemd. Het standaardfries baseert zich voornamelijk op de Klaaifryse dialecten. In enkele steden wordt Stedsfrysk (Stadsfries) gesproken, een mengvorm van Nederlands en Fries. In de gemeente It Bilt (Het Bildt), ten noorden van Fryslân aan de Waddenzee gelegen, wordt door de bevolking Bilts gesproken. Dit land is rond 1500 ingepolderd door boeren uit het tegenwoordige Zuid-Holland en men vindt dan ook enkele Zuid-Hollandse invloeden in dit dialect terug. In twee gemeenten in het zuidoosten van de provincie, Oost- en West-Stellingwerf, wordt door ongeveer eenderde van de bevolking Stellingwerfs gesproken. Deze streektaal kent ook Saksische invloeden. Op de eilanden leven voornamelijk dialecten die bestaan uit mengvormen van het Nederlands en het Fries. Alleen Schiermonnikoog kent nog een Fries dialect, maar dit wordt ernstig met uitsterven bedreigd (Gorter (1997):1154; Tiersma (1985):4-6).

1.2 De mogelijkheden voor een Friestalige spraaksynthesizer

Veel minderheidstalen hebben het zwaar. Sommige worden zelf met uitsterven bedreigd. Om zich toch naast de meerderheidstaal te kunnen handhaven is het belangrijk om het

aantal sprekers in stand te houden. Een van de middelen om dit te doen is om mee te gaan met de technologische ontwikkelingen op spraak- en taalgebied. Wanneer sprekers van een minderheidstaal namelijk bij het werken met commerciële spraak- en/of taaltechnologische producten als een spraaksynthesizer of spraakherkenner gedwongen worden om in de meerderheidstaal te spreken, zal dit juist alleen maar meer leiden tot een afname van het gebruik van de minderheidstaal waardoor de kans op uitsterven groter wordt (Nadeu e.a. (2001):1). Vandaar dat dit project belangrijk is voor het voortbestaan van het Fries. Door een Friestalige spraaksynthesizer te ontwikkelen, biedt men het Fries een stukje extra zekerheid voor de toekomst.

Mijn doel is om een Friestalige spraaksynthesizer te ontwikkelen met behulp van het bestaande spraak synthese systeem Festival. De taak van een spraaksynthesizer is om geschreven tekst om te zetten in spraak. Wat zou nu een efficiënte toepassing van zo'n synthesizer in Fryslân zijn? Een spraaksynthesizer kan worden gebruikt als informatie-systeem, als leesmachine voor blinden en slechtzienden of als communicatiehulpmiddel voor mensen met een communicatiehandicap (Holmes & Holmes (2001):67).

Laten we deze mogelijkheden eens bij langs lopen en beginnen bij de toepassing als communicatiehulpmiddel voor mensen met een communicatiehandicap. Elke Fries beheerst in meer of mindere mate het Nederlands als spreektaal. Wanneer men een spraaksynthesizer als communicatiehulpmiddel bij een handicap gebruikt, zou een Nederlandstalige spraaksynthesizer dus in principe niet tot een grote communicatienood lijden. Maar de gebruiker van deze apparatuur kan het als een grote handicap ervaren dat hij/zij niet met intimi in de moedertaal kan spreken. Vandaar dat dit een reële toepassing kan zijn.

Wanneer we de mogelijkheden voor een Friestalige leesmachine voor blinden en slechtzienden aftasten, moeten we kijken naar de hoeveelheid geschreven materiaal in het Fries. Zo'n leesmachine zou met name van belang kunnen zijn bij de geschreven media. Wanneer we echter kijken naar de twee grote dagbladen in Fryslân, dan zien we dat er slechts minder dan 3% per dag aan Friestalige teksten in de artikelen wordt gebruikt (Gorter (1997):1156). Bekend is dat beide kranten eens per week een speciale Friestalige pagina, de "Fryske side" (Friese pagina), publiceren. In het percentage is het aantal Friestalige familieberichten en advertenties niet meegerekend. Verder zijn enkele literaire tijdschriften volledig Friestalig en verschijnen ieder jaar verschillende nieuwe boeken, voornamelijk verhalen en poëzie (Gorter (1997):1156).

Bij het gebruik van de computer, bijvoorbeeld bij tekstverwerken (horen wat men heeft geschreven) en e-mailen of websurfen (het aantal Friestalige internetsites wordt steeds meer uitgebreid) zou men ook gebruik kunnen maken van een Friestalige leesmachine.

Wil men juist een zo groot mogelijke groep met deze spraaksynthesizer in aanraking te laten komen, dan kan men het beste voor een toepassing als een (telefonische) informatiesysteem kiezen. Men zou kunnen denken aan het informatiesysteem van een bank, VVV of een spraakportaal met de meest bezienswaardige objecten in een stad. Hiervoor kan men beter multilinguale systemen gebruiken, aangezien ook niet-Frietaligen deze bronnen kunnen raadplegen. Bij een speciale reserveringslijn voor de professionele Friestalige toneelgroep Tryater zou men kunnen voldoen met een eentalig systeem.

2. Festival

2.1 Festival in het kort

Festival is het multilinguale spraaksynthese systeem dat in het CSTR, het Centre for Speech Technology Research, aan de universiteit van Edinburgh ontwikkeld is door Alan W. Black, Paul Taylor en Richard Caley. Het systeem is opgebouwd uit modules. De filosofie achter dit systeem is, dat men met Festival gemakkelijk aan een kleine deel van de spraaksynthesizer kan werken om het geheel te verbeteren. Het systeem is speciaal gemaakt om nieuwe modules toe te laten zodat de ontwikkeling op dit gebied door kan gaan (Black e.a. (1999):11). Hierdoor is het ook relatief gemakkelijk om een nieuwe taal te implementeren. Er zijn al verschillende talen en stemmen geïmplementeerd in Festival: Amerikaans- en Brits-Engels, Castiliaans- en Mexicaans-Spaans, Duits, Pools, Grieks, Welsh Gaelic en Baskisch (Black e.a. (2000):1).

Bij deze opdracht wordt een (gedeeltelijk) Spaanstalig systeem omgebouwd naar een Friestalig systeem. Er is niet specifiek voor deze taal gekozen, dit is gewoon zo tot stand gekomen. Wanneer men een taal bewust gekozen zou hebben, dan zou men juist eerder voor een taal kiezen die uit dezelfde familie als het Fries komt.

Omdat snelheid belangrijk is in onderzoek is de code van Festival geschreven in een gecompileerde 'low-level' taal, namelijk C++. Deze taal is veel sneller dan alternatieve 'high-level' talen. Het hele systeem kan echter niet op zo'n gecompileerde taal lopen, omdat het dan erg veel werk is om bijvoorbeeld algoritmes te veranderen. Daarom wordt naast C++ ook Scheme gebruikt (Taylor e.a. (1998):4).

Scheme is een dialect van Lisp, beter bekend als de taal met veel haakjes. Meer informatie over het gebruik van Scheme in Festival is te vinden in Black e.a. ((2000):17-28)

2.2 Hoe is Festival opgebouwd?

Festival heeft drie hoofdstructuren (Taylor e.a. (1998):2-4):

- Item

Een item is één enkele linguïstische eenheid, zoals bijvoorbeeld een foneem, woord, syllabe, intonatiefrase, etc. Elk item in Festival heeft een verzameling (set) kenmerken (features) met bijbehorende eigenschappen. Dit kunnen zowel waarden als functies zijn. Naast kenmerken bevatten items ook een verzameling van verbindingen naar zogenaamde knopen (nodes) in relaties (relations). Elk item kan met een onbepaald aantal relaties verbonden worden.

- Relatie (relation)

Een relatie verbindt de linguïstische types aan elkaar. Dit gebeurt door een verzameling knopen met elkaar te verbinden. Elke knoop heeft één verbinding naar een item. Anders gezegd: Men kan bij een relatie spreken van een gegevensstructuur die gebruikt wordt om linguïstische items in linguïstische structuren te organiseren zoals boomstructuren of lijsten. Er bestaan bijvoorbeeld woord-, foneem-, syntax- of syllaberelaties. Een lijst is een lineaire relatie waarbij elke knoop een vorige en een volgende verbinding heeft. Een voorbeeld is een relatie die alle woorden in een uiting aan elkaar verbindt in de volgorde van voorkomen. In boomstructuren hebben knopen ook links en rechts verbindingen. Een boomstructuur is bijvoorbeeld een syllaberelatie met onset, coda, nucleus en rhyme.

Belangrijk bij Festival is dat items in meerdere relaties kunnen voorkomen.

- Uiting (utterance)

Er is nu een ongeordende verzameling items, waarbij elk item ook nog eens een verzameling kenmerken bevat. De relaties zijn structuren van knopen die allemaal naar een item gaan. De uitingstructuur verzamelt deze en maakt er één object van. Een uiting (utterance) bevat dus een lijst van alle relaties.

De uittingsstructuur ligt in het centrum van Festival. Deze structuur zorgt voor synthese van tekst naar spraak. Dit hoeft niet eens beslist een zin te zijn; vaak bedraagt het een andere vorm dan wat linguïstisch gezien een zin zou kunnen zijn (Black e.a. (1999):63).

Aan de hand van deze structuren onderneemt het systeem bepaalde stappen om van geschreven tekst naar spraakgeluid te komen. Deze stappen zijn:

- tokenisatie
- token identificatie
- token naar woord
- Part of Speech tagging
- prosodische frasering
- lexicon
- intonationale accenten
- duur toekennen
- F0-contour genereren
- geluidsgolf weergeven

Deze stappen zijn verwerkt in de modulen van Festival. Elke module zorgt voor een stukje nieuwe informatie over de structuur van de uiting. Het hangt af van de geselecteerde stem en het type uiting welke modulen precies worden doorlopen (Black e.a. (1999):64-65; Rog (1999):4).

Festival bevat onder andere de volgende modulen (Holmes & Holmes (2001):93-118; Black e.a. (1999):47-106; Rog (1999):4-12). De modulen worden in het kort toegelicht:

- Tekst analyse (Text analysis)

Hier vindt tokenisatie plaats: de leestekens wordt eruit gehaald en de tokens (getallen, afkortingen, etc.) worden omgezet in woorden. Tevens wordt rekening gehouden met de homografische disambigüiteit, bijvoorbeeld dat een jaartal anders wordt uitgesproken dan een gewoon getal. Een voorbeeld van deze module is in dit werkstuk te vinden in bijlage 4, bestand fristoken.scm.

- Foneemset (Phone set)

Een foneemset is een verzameling symbolen die verder gedefinieerd worden in kenmerken (features) zoals vocaal/consonant, plaats van articulatie, etc. Deze kenmerken zijn weer gedefinieerd in waarden. Een voorbeeld van deze module is in dit werkstuk te vinden in bijlage 6, bestand mrpa_phones.scm.

- Lexicon

Hier wordt de uitspraak van de woorden bepaald. Dit kan volgens een addenda (korte lijst van met de hand toegevoegde woorden), een gecompileerd lexicon (groot uitspraakwoordenboek dat tienduizenden woorden kan bedragen) of een methode voor woorden die niet in beide lexicons voorkomen (grafeem-naar-foneem regels). Een voorbeeld van deze module is in dit werkstuk te vinden in bijlage 3, bestand frislex.scm.

- Uiting (Utterance)

Deze module zorgt voor de synthese. Het type uiting bepaalt welke modulen wel geactiveerd worden en welke niet. Deze modulen zullen de benodigde relaties en items verzamelen zodat uiteindelijk een golfvorm geactiveerd kan worden.

- Part Of Speech tagging

Syntactische analyse: woordsoorten worden aan woorden toegekend aan de hand van een Hidden Markov Model.

- Frasering (Phrase breaks)

Hierbij vindt identificatie plaats van de frasen in een uiting. Dit kan op twee manieren. De ene gaat via een zogenaamde CART-boom (met behulp van een test wordt van elk woord de kans berekend dat het aan het einde van een prosodische frase staat). Er is ook een makkelijkere versie die een pauze zet na het laatste teken van het woord dat gevolgd wordt door een teken van punctuatie.

- Intonatie (Intonation)

Er zijn verschillende intonatiemodules in Festival aanwezig:

- standaard intonatie (geeft alleen begin- en eind-F0 waarde)
- gemakkelijke intonatie (voorspelling per syllabe m.b.v. CART-boom)
- boom intonatie (twee CART-bomen voor accenten en eindwaarde, hier is het bijvoorbeeld mogelijk om ToBI (Tones and Break Indices) te implementeren)
- Tilt intonatie (via getraind model voorspellen van accenten)

Een voorbeeld van deze module is in dit werkstuk te vinden in bijlage 5, bestand frisint.scm.

- Duur (Duration)

Er zijn verschillende duurmodules in Festival aanwezig:

- standaard duur (alle segmenten 100 ms)
- gemiddelde duur (aparte lijst)
- Klatt duur (regels van Klatt)
- CART duur (duur van klank of duur m.b.v. z-scores)

- Synthesizer

Er zijn verschillende synthesizers werkzaam in Festival, onder andere:

- de UniSyn-synthesizer, standaard als module aanwezig in Festival
- de oudere difoon synthesizer (vroeger standaard aanwezig in Festival)
- LPC difoon synthesizer
- MBROLA

Omdat de makers van de Nederlandstalige versie van Festival dit met behulp van MBROLA doen, zullen wij deze synthesizer ook gebruiken (zie hoofdstuk 3).

- Geluidsgolf (Audio output)

De opgevraagde difonen worden geconcateneerd en afgespeeld. Men hoort de ingevoerde tekst.

Het syllabeonderdeel wordt niet als aparte module in de handleiding van Festival genoemd. Toch denk ik dat dit onderdeel ook belangrijk is aangezien hier de morfologische analyse plaatsvindt. Elk syllabe-item is in de zogenaamde SylStructure-relatie aanwezig. In deze boomstructuur zijn de ouders de woorden en de dochters de segmenten. Syllaben bevinden zich tevens in de intonatielatie om een verbintenis te vormen met de betreffende intonatieaccenten.

3. MBROLA

3.1 Wat is MBROLA?

Het MBROLA-project is ontwikkeld in het TCTS Lab van de Faculté Polytechnique de Mons in België. MBROLA staat voor Multi Band Resynthesis OverLap Add (Dutoit e.a. (1996)). Het patent op deze techniek is in handen van de Faculté Polytechnique de Mons. In het kort komt het erop neer dat MBROLA een synthesizer is waarbij gebruik wordt gemaakt van concatenatie van difonen. Net als bij de PSOLA-methoden voegt MBROLA direct overlappende frames in het tijdsdomein toe. Spectrale oneffenheden worden afgevlakt, wat de kwaliteit van de gesynthetiseerde spraak verhoogt.

De input bij MBROLA bestaat uit een lijst met fonemen en prosodische informatie (duur en pitch). De output bevat 16 bits-spraak, in .raw-, .wav-, .au- of .aiff-formaat.

Omdat men geen ruwe tekst als input kan gebruiken is het MBROLA-systeem geen spraaksynthesizer. Men kan MBROLA gebruiken in combinatie met systemen als Festival waarbij de fonetische en prosodische gegevens verkregen worden.

Het doel van MBROLA is om spraaksynthesizers te creëren voor zoveel mogelijk stemmen, talen en dialecten. Deze databases zijn vrij in gebruik voor niet-commerciële doeleinden. Het uiteindelijke doel van het MBROLA-project is het bevorderen van het wetenschappelijk onderzoek, vooral op het gebied van de prosodie. Meer informatie over MBROLA is te vinden op hun internetsite: <http://tcts.fpms.ac.be/synthesis/mbrola.html>

Wanneer men het programma niet voor commerciële of militaire doeleinden gebruikt, is op deze site is het MBROLA-systeem gratis te downloaden, evenals diverse databases van stemmen in talen als bijvoorbeeld Amerikaans-Engels, Arabisch, Braziliaans-Portugees, Bretons, Brits-Engels, Kroatisch, Nederlands, Frans, Duits, Grieks, Mexicaans-Spaans, Roemeens, Spaans en Zweeds. Dutoit e.a. (1996) spreken in hun artikel over MBROLA versie 2.00. Tegenwoordig kan men versie 3.01 van internet downloaden.

3.2 Gebruik van MBROLA

Wanneer men een nieuwe database wil maken, zal men alvorens het aanleveren van een difoon database (geluidsbestanden en segmentatie-informatie), een officiële overeenkomst met MBROLA aan moeten gaan. Copyrights van de verkregen MBROLA-database liggen bij Faculté Polytechnique de Mons – T. Dutoit. De medewerkers zullen de difoon database vernietigen. De MBROLA-database wordt ook binnen het MBROLA-project gebruikt en zal vrij via internet te verkrijgen zijn. De maker krijgt licentie van de gebruiksrechten van de nieuwe database, op voorwaarde dat de database alleen met het MBROLA-systeem gebruikt wordt (Dutoit e.a. (1996)). Ons contactpersoon binnen MBROLA is Baris Bozkurt (bozkurt@tcts.fpms.ac.be).

Middels een .pho-file kan men de input van het MBROLA-systeem verzorgen. Een voorbeeld hiervan is temp.pho in bijlage 11. Het eerste teken op elke regel is de naam van het foneem met erachter de duur in milliseconden. Optioneel kan deze duur gevolgd worden door waarden die pitchpunten markeren. Elk pitchpunt bestaat uit twee waarden: de eerste geeft de relatieve positie waarop het pitchpunt ligt aan in percentage van de foneemduur, de tweede waarde is de pitchwaarde in Hertz. Wanneer we bijvoorbeeld naar bijlage 11 kijken, zien we dat het eerste pitchpunt bij /Y/ ligt op 30% van de duur (75 ms) met een waarde van 145 Hz. Het tweede pitchpunt ligt bij de /a:/ op 45% van de duur (170 ms) met een waarde van 175 Hz.

De verkregen MBROLA-database kan men (in Linux) testen met het volgende commando:
\$../mbrola-linux-i386 temp.dat TEST/temp.pho temp.wav

Hierbij is “mbrola-linux-i386” het MBROLA-systeem. Dit programma ligt één directory hoger dan het opgegeven pad, vandaar “./”. Het bestand “temp.dat” is de database. “Temp.pho” is het input-bestand dat één directory (TEST) dieper ligt. “Temp.wav” is het, middels MBROLA, verkregen synthese-fragment.

De file mbrola.scm binnen Festival zorgt voor koppeling van de fonetische en prosodische informatie (de input, te vergelijken met het .pho-bestand) naar het MBROLA-systeem, dat zorgt voor de output, de gesynthetiseerde spraak. Omdat niet alle SAMPA en Worldbetsymbolen gelezen kunnen worden door MBROLA worden deze symbolen op het laatste moment met behulp van het anmrpa-bestand (bijlage 10) opnieuw benoemd. Zo leest MBROLA ‘#’ niet als stiltefoneem. Hiervoor gebruikt het ‘_’. Het is mij onbekend wat bedoeld wordt met het commando FLUSH.

4. Het bouwen van een nieuwe stem

4.1 Ingrediënten voor een Friestalige stem in een spraaksynthesizer

4.1.1 Friese foneemset

Natuurlijk is voor het maken van een Friestalige spraaksynthesizer een foneemstelsel van het Fries nodig. Opvallend was dat hierover weinig in de literatuur en op internet te vinden was. Na contact te hebben opgenomen met de Fryske Akademy bleek dat Hoekstra en Siebenga (2001) recentelijk een artikel over het Friese foneemstelsel met bijbehorende SAMPA symbolen hebben geschreven. Dit symboolstelsel gebruiken zij bij hun project Korpus Sprutsen Frysk. Samen met het boek van Cohen e.a. (1961) en het foneemstelsel uit het Fries-Nederlands woordenboek (Zantema (1984)) als uitgangspunten heb ik met hulp van Paul Boersma (UvA) getracht een overzicht van het Friese foneemstelsel in Worldbet te maken.

Worldbet is een fonetisch alfabet dat een overzicht geeft van alle talen van de wereld, waar SAMPA voornamelijk gericht is op Indo-europese talen en dus symbolen voor andere talen mist. Voor SAMPA werd in eerste instantie namelijk verondersteld dat het gebruikt zou worden voor transcriptie binnen één taal, niet tussen talen waardoor soms één symbool gebruikt wordt voor verschillende geluiden (Hieronymus (1994):1-2).

Het resultaat van onze bevindingen is te vinden in bijlage 1. De voorbeelden die in deze bijlage genoemd worden zijn bijna allemaal afkomstig uit het artikel van Hoekstra en Siebenga (2001). Ik heb bij het opstellen van deze lijst het Klaaifrysk (Kleifries) als uitgangspunt genomen. Ten eerste omdat het Klaaifrysk een grote bijdrage heeft geleverd aan het standaardfries en tevens de basis vormt van de literatuurtaal (Cohen e.a. (1961):112). Een andere reden voor deze keuze is dat het Klaaifrysk mijn moedertaal is. Bij twijfel heb ik dan ook enkele keren mijn eigen spraak geanalyseerd met het programma 'Praat' van Paul Boersma en David Weenink. Dit wil echter niet zeggen dat mijn uitspraak altijd garant staat voor het Klaaifrysk.

Enkele opmerkingen naar aanleiding van deze foneemset:

Consonanten:

De plosief /g/ komt meestal aan het begin van een beklemtoonde syllabe voor (Hoekstra e.a. (2001):1). Dit in tegenstelling tot de stemloze fricatief /x/ die vrijwel nooit aan het begin van een syllabe staat. De /G/ staat vaak in mediale positie van een woord, bijvoorbeeld tussen twee vocalen als in 'drage' /dra:G&/ (dragen). Er bestaat een fonologische tegenstelling tussen het werkwoord 'bargje' /barGj&/ (morsen) en het zelfstandig naamwoord 'bargje' /barxj&/ (varkentje) (Cohen e.a. (1961):123; Hoekstra e.a. (2001):5). De stemhebbende fricatieven /G/, /v/, en /z/ komen nooit aan het begin van een woord voor (Hoekstra e.a. (2001):2).

Het woordenboek van Zantema (1984) stelt de 'w' van 'wetter' (water) gelijk aan een /v/. Deze 'w' wordt echter eerder uitgesproken dan een approximant, of /v/ zonder frictie (Hoekstra e.a. (2001):5). Bovendien komt de /v/ niet aan het begin van een woord voor (zie bovenstaand). Ook kan er in het Fries geen stemhebbende fricatief volgen op een stemloze plosief. Dus 'kwart' (kwart) kan niet worden getranscribeerd als /kvart/. Ik ben van mening dat ik, bij het uitspreken van diverse voorbeelden, een /v/ zonder frictie fabriceer in deze gevallen. Net als Cohen e.a. ((1961):125) onderscheid ik een labiodentale 'w' of als IPA-symbool: [ʋ]. In Worldbet wordt dit symbool als /V/ gegenereerd. In Festival wordt heb ik er /V|/ van gemaakt, omdat bij feature-aanduidingen met de letter V alle vocalen worden aangeduid.

Waar ik het wel met Hoekstra e.a. ((2001):5) eens ben is dat de 'w'-klank als eerste element in een diftong een bilabiale /w/ is.

Bij de /r/ heb ik gekozen voor de tongpunt-r, aangezien deze door de meeste Friezen wordt gebruikt (Cohen e.a. (1961):124).

Zowel Hoekstra e.a. ((2001):4) als Cohen e.a. ((1961):XIII) erkennen de /S/, zij het als leenklank. Als je de /S/ erkent, zou je naar mijn mening ook de /Z/ moeten erkennen. Hierin word ik gesteund door Cohen e.a. ((1961):XIII).

Vocalen:

Om misverstanden te voorkomen noteer ik hier dat, evenals in het bovenstaande, met de gebruikte symbolen Worldbetsymbolen bedoeld worden.

Hoekstra e.a. (2001) gebruiken de kardinale klinkers /e/, /ɪ/, /o/ (in IPA respectievelijk [e], [ø] en [o]) en de speciale 'lax'-klinkersymbolen (gesloten-centraal) /I/, /Y/ en /U/ (in IPA respectievelijk [ɪ], [ʏ] en [u]) door elkaar. Dit is erg ongewoon.

Net als Cohen e.a. (1961) gebruik ik een ander symbool voor 'rom' (ruim) dan voor 'rook' (rook), want de vocalen in 'ring' (ring) worden ook anders geschreven dan in 'reed' (oprit). Evenzo worden de vocalen in 'nut' (nut) anders geschreven dan in 'deun' (zeer dicht). De vocalen in 'ring', 'nut' en 'rom' worden even hoog uitgesproken. Vandaar dat ik hier voor de speciale 'lax'-klinkersymbolen heb gekozen, dus /I/, /Y/ en /U/. Wel gebruik ik van de kardinale klinkers de /e:/ voor de vocaal in 'reed' en /ɪ/ in 'deun' aangezien deze ook even hoog worden uitgesproken.

Worldbet laat maar 2 karakters per klank toe (Hieronymus (1994):3). Vandaar dat ik bij de diftongen het kenmerk 'lang' niet heb toegekend aangezien dit niet tot verwarring leidt. Bovendien is de lengte van het eerste element uit een diftong soms moeilijk te bepalen (Hoekstra e.a. (2001):6).

Diftongen:

De Friese diftongen eindigen allemaal op /i/, /u/ of /ɔ/ (Hoekstra e.a. (2001):3). Net als Hoekstra en Siebenga laat ik de (betreffende) diftongen op /i/ of /u/ eindigen in plaats van op /j/ of /w/.

De diftong /Yɔ/ heeft als eerste element /Y/ en geen /ɔ/ (Hoekstra e.a. (2001):6; Cohen (1961): 113-116). Deze diftong komt alleen voor wanneer zij voor de consonanten /n/ en /r/ staat: 'gleon' (gloeiend), 'kleur' (kleur).

Evenals Hoekstra en Siebenga ((2001):6) hoor ik geen verschil tussen de diftong in 'blau' (blauw) en 'au' (auw). Ditzelfde geldt voor de diftong in 'goeie' (dag) en 'bloei' (bloei). Iets wat Zantema (1984) in beide gevallen wel doet. Vandaar dat beide diftongen in deze foneemset respectievelijk als /Au/ en /ui/ getranscribeerd worden.

Door Hoekstra e.a. ((2001):3) wordt de diftong in 'laitsje' (lachen) als /ai/ getranscribeerd. In het Kleifries wordt deze klank echter meestal als />i/ uitgesproken (Dijkstra (1991):59). Vandaar dat ik deze mogelijkheid heb toegevoegd en als eerste keus heb aangenomen. Ik hoor geen verschil in een lange /a:i/-klank en een korte /ai/-klank, in bijvoorbeeld respectievelijk 'aai' (ei) en 'aike' (eitje). Vandaar dat ik dit als een klank heb beschouwd. Ook Hoekstra e.a. (2001) en Cohen e.a. (1961) maken dit onderscheid niet.

De diftong /Ui/ komt slechts in beperkt aantal woorden voor (Visser in Hoekstra e.a. (2001):7).

De diftong in het woord 'moai' (mooi) die Hoekstra e.a. ((2001):3,7) transcriberen als /o:i/, wordt in mijn Kleifries uitgesproken als /U>i/. Dit lijkt meer op een triftong. Worldbet laat echter geen 3 karakters toe, dus dit zou waarschijnlijk genoteerd moeten worden als /U>_i/.

Hoekstra e.a. ((2001):4,7) gebruiken net als Zantema (1984) acht stijgende diftongen (zogenaamde brekingen), vanuit de /j/ of /w/ gemaakt. Men zou zo'n klank eventueel ook als twee aparte klanken kunnen karakteriseren in plaats van een diftong. Cohen e.a. ((1961):119) pleiten er echter voor om ze deze "gebroken klanken" als zodanig te erkennen en "niet als varianten (allofonen) van de 'ongebroke"ne"', omdat er soms verschil is tussen de ongebroken klanken en de gebroken in dezelfde positie. Kijk bijvoorbeeld naar 'gien' /gi&n/ (gegaan) tegenover 'gjin' /gjln/ (geen).

Bij de diftongen /i&/, l&/, /u&/ en /U&/ komt deze zogenaamde breking voor wanneer het hoofdwoord dat deze diftong bevat, verlengd wordt. Hierbij gaat het om respectievelijk /j/, /jE/, /wU/ en /wa/. Andere stijgende diftongen zijn: /jo/ en /j>/ in respectievelijk 'rjocht' /rjox/& (recht) en 'ljocht' /lj>xt/ (licht) en /jY/ in 'slurkje' /sljUrkJ&/ (zachtjes stromen). Dit laatste voorbeeld komt uit Cohen e.a. ((1961):119).

Het woordenboek van Zantema (1984) noemt nog zeven triftongen voor het Fries. Hoekstra e.a. (2001) behandelen deze triftongen niet expliciet in SAMPA omdat deze combinaties al besproken zijn. Evenals Hoekstra e.a. ((2001):7) noem ik ze hier. Zoals is te zien, verglijden deze triftongen in een /i/ of /u/, in plaats van een /j/ of een /w/.

<i>worldbetsymbool</i>	<i>voorbeeld</i>	<i>transcriptie</i>
wU_i	muoike	m wU_i k e
wa_i	moaist	m wa_i s t
wa_:i	koai	k wa_:i
jy_u	bliuw	b l jy_u
ju_:	priuwe	p r ju_: &
jU_:u	bleau	b l jU_:u
j_Au	fjouwer	f j_Au & r

Aan deze tabel zou mijn uitspraak van 'moai' als /mU>i/ nog aan toegevoegd kunnen worden, al gaat deze triftong niet uit van een /j/ of /w/.

Omdat hier wel verwarring zou kunnen ontstaan over het kenmerk 'lang' heb ik hier wel het diacritische teken laten staan. Al ken ik zelf bijna geen woorden met de triftong /wa_:i/. Het voorbeeld gebruik ik altijd met een andere betekenis, waar ook een andere uitspraak aan vast zit, namelijk /U>_i/. Gevolg is dat elke triftong een verlenging met de underdog (_) kent.

Dijkstra ((1991):71-73) schrijft dat de uitspraak van 'iuw' soms gelijk kan zijn aan de uitspraak van 'eau'. Deze uitspraak ligt echter niet vast. Zo kan de uitspraak van 'eau' in 'leau' ((ik) geloof) en 'iuw' in 'skriuwe' ((ik) schrijf) bijna gelijk zijn. Woorden met 'eau' kunnen echter niet als 'iuw' uitgesproken worden. Andersom kan dit wel. Ikzelf spreek alle drie de vormen in 'bliuw', 'priuwe' en 'bleau' uit als /jo_:u/.

Een vocaal wordt genasaleerd uitgesproken als erna een cluster voorkomt gevormd door een /n/ gevolgd door een /s/, /f/, /v/, /j/, /w/, /l/, of /r/ (Hoekstra e.a. (2001):3,6-7; Cohen e.a. (1961):117). Het nasale kenmerk wordt dan achter de vocaal aangeduid met het diacritische teken /-/. Wanneer een vocaal echter al twee karakters bevat, kan het kenmerk nasaliteit met een underdog worden toegevoegd: /_~/.

4.1.2 Overige benodigdheden

Alle hier genoemde projectleiders zijn werkzaam op de Fryske Akademy te Leeuwarden.

Uitspraakwoordenboek en informatie over woordaccent:

Wurdboek fan de Fryske Taal:

Een woordenboek dat volledig en alleen over de uitspraak van het Fries gaat bestaat niet. Wel is er een digitaal bestand van het Wurdboek fan de Fryske Taal (Woordenboek der Friese Taal), hierna afgekort als WFT. Hierin wordt alleen bij de stamwoorden de uitspraak aangegeven. Tevens is het woordaccent bij deze woorden aangegeven.

De uitspraak is echter niet gegeven in fonemen die het IPA gebruikt. De fonemen zouden dus (sowieso) moeten worden omgezet naar Worldbetsymbolen. Het WFT is echter pas (redactioneel) in 2003 klaar.

Projectleider is Piter Boersma.

Frysktalich hânwurdboek:

Ook is er een digitale versie van het Frysktalich hânwurdboek (Fries-Fries handwoordenboek). Dit woordenboek zal in 2002 uitkomen. Hierin staan ook bij samenstellingen de uitspraak en woordaccent aangegeven. Waarschijnlijk wordt in dit woordenboek ook niet het IPA-model gebruikt dus ook hier zouden de fonemen (eigenlijk sowieso al) omgezet moeten worden naar Worldbetsymbolen.

Projectleiders zijn Rienk de Haan en Hindrik Sijens.

Of alle nieuwe woorden (denk aan "sms", "mobieltje") ook in deze woordenboeken zullen staan is niet duidelijk. Aanvullend kan men eventueel het boek van Rienk de Haan gebruiken, genaamd "Terminology fan 'e saakfakken" (terminologie van de zaakvakken). Dit boek is geschreven om de Friese terminologie voor het voortgezet onderwijs duidelijk op papier te krijgen. Allerlei schoolvakken passeren de revue, waaronder bijvoorbeeld ook informatica. Mij is onbekend of hierin ook aandacht wordt besteed aan uitspraak en woordaccent.

Op computertechnisch gebied heeft de heer Cor Jousma ook een lijst met computertermen gemaakt. Deze is gepubliceerd op de internetsite van de Fryske Beweging. Woordaccent en uitspraak worden hier niet in aangegeven. De heer Jousma is tevens werkzaam op de Fryske Akademy.

POS-tagging:

In het Korpus Sprutsen Frysk (Corpus Gesproken Fries) zal een gedeelte van het corpus 'getagged' worden voor POS (Part Of Speech). Hierbij wordt gebruik gemaakt van het systeem dat gebruikt wordt bij het WFT (zie boven) (Hoekstra (2001)). Voor POS-tagging bij deze spraaksynthesizer kan misschien ook gebruik worden gemaakt van dit systeem, al weet ik niet of dit mogelijk is. Projectleider van het Korpus Sprutsen Frysk is Eric Hoekstra. Volgens Black e.a. ((1999):79) is POS-tagging echter niet taalspecifiek. Dit zou betekenen dat wij de al in Festival aanwezige module zouden kunnen gebruiken.

Intonatie:

Volgens Cohen e.a. (1961: 147) komt zowel woordaccent (enkele uitzonderingen daargelaten) als zinsaccent van het Fries overeen met die van het Nederlands. Hier is echter nooit onderzoek naar gedaan. Ook in de literatuur is betrekkelijk weinig te vinden over de intonatie in het Fries. Een enkeling heeft een paragraaf(je) aan dit onderwerp gewijd. Tiersma ((1985):139-142) noemt bijvoorbeeld dat de meeste inhoudwoorden meer accent krijgen dan functiewoorden. Verder schrijft hij dat prefixen die ook op zichzelf kunnen staan eerste accent in een woord kunnen krijgen. En ook (inhouds)woorden die nieuwe informatie met zich meebrengen of een tegenstelling uitdrukken krijgen een accent. Ja/nee-vragen hebben een stijgende intonatie op het laatste inhoudswoord. Dit zijn vrij algemene regels die ook voor het Nederlands gelden.

Het enige onderzoek met betrekking tot de intonatie van het Fries is gedaan door Jarich Hoekstra (1991). Hieruit kwam naar voren dat voorzetsels in het Fries relatief vaker accent

krijgen dan in het Nederlands (en minder dan in het Engels). Het gaat hierbij om lexicale voorzetsels en in bijzondere gevallen functionele voorzetsels.

Misschien dat er nog nieuwe informatie in het Handbuch des Friesischen (Munske e.a. (2001)) staat. Dit boek is echter nog steeds in bestelling op het moment van schrijven.

Voor de intonatie bij de Friestalige spraaksynthesizer zou men gebruik kunnen maken van een van de al aanwezige modules. Maar als de intonatie, en eventueel ook de prosodie, inderdaad als zodanig overeenkomt met die van het Nederlands zou men kunnen overwegen om ToDI (Transcription of Intonation in Dutch) te gebruiken. ToDI is ontwikkeld door Carlos Gussenhoven (KUN), Toni Rietveld (KUN) en Jacques Terken (TUE). Meer informatie over ToDI is te vinden op de webpagina <http://lands.let.kun.nl/todi/todi/home.htm>. ToDI vereist veel specialistische kennis, dus waarschijnlijk is verstandig een expert te raadplegen.

4.2 Werkwijze

4.2.1 Algemeen

Wanneer men een nieuwe stem bouwt, zal men, naast het verzamelen van foneemset en andere benodigdheden, ook een difoonschema op moeten stellen van alle mogelijke difonen in die taal. Difonen zijn eenheden van twee opeenvolgende spraaksegmenten die beginnen in het midden van het ene foneem en eindigen in het midden van de andere (Olive (2001):5). Deze difonen worden opgenomen en opgeslagen. De verkregen difoon database wordt na ondertekening van een officiële overeenkomst opgestuurd naar MBROLA (zie paragraaf 3.2), alwaar de bestanden worden omgezet naar MBROLA. Festival zal met behulp van deze MBROLA-database de spraaksynthese tot stand brengen.

4.2.2 Opnames

Voor de opnames is een proefpersoon met heldere, vaste stem vereist. De persoon mag niet mompelen of een andere spraakafwijking hebben. Verder moet men er voor zorg dragen dat de persoon niet verkouden is of een kater heeft. Bij voorkeur worden de opnames in een sessie gedaan en in de ochtend zodat de stem het helderst is. Wanneer men toch nieuwe opnames moet maken, dan op hetzelfde tijdstip als de eerdere opnames. Kwaliteit is van belang bij de opnames. Daarom wordt voorkeur gegeven aan opname in een stille kamer (Black e.a. (2000):127).

De difonen worden middels een draagzin opgenomen. Deze draagzin is zo gekozen dat de buurklanken het difoon zo min mogelijk beïnvloeden. Wegens tijdgebrek is het voor mij alleen mogelijk één woord door om de synthesizer te laten synthetiseren. Alleen die difonen die voor dit woord noodzakelijk zijn, worden door mij opgenomen. Voor het woord 'bjusterbaarlik' (wonderbaarlijk) zijn de volgende draagzinnen gebruikt:

bjusterbaarlik
bə (stiltfoneem-b)
pə-pə-bjə-bjə-bjə-bjə-pə-pə
pə-pə-jɣ-jɣ-jɣ-jɣ-pə-pə
pə-pə-pɣs-ɣs-ɣs-ɣs-pə-pə
pə-pə-stə-stə-stə-stə-pə-pə
pə-pə-tə-tə-tə-tə-pə-pə
pə-pə-pər-ər-ər-ər-pə-pə
pə-pə-rbə-rbə-rbə-rbə-pə-pə

pə-pə-ba:-ba:-ba:-ba:-pə-pə
pə-pə-pa:r-a:r-a:r-a:r-pə-pə
pə-pə-rlə-rlə-rlə-rlə-pə-pə
pə-pə-lə-lə-lə-lə-pə-pə
pə-pə-pək-ək-ək-ək-pə-pə
ək (k-stiltefoneem)

Belangrijk is om de draagzinnen monotoon uit te spreken. Voor de opnames van de Friese difonen is mijn eigen stem gebruikt. De opnames vonden plaats in de stille kamer van het Fonetisch Instituut van de UvA.

4.2.3 Aanleggen van een difoonset

Met behulp van 'Praat' van Paul Boersma en David Weenink, zijn de difonen uit het signaal gehaald door dat gedeelte uit het signaal te extraheren dat begint in het midden van het ene foneem en eindigt in het midden van het tweede foneem. Rondom de difonen is nog 50 ms spraak gelaten. Tevens is in een segmentatie-file vastgelegd wat het begin-, eind- en middelpunt van het difoon per signaal is.

De difoon database en de segmentatiefile worden na ondertekening van de officiële overeenkomst opgestuurd naar MBROLA (zie paragraaf 3.2), waar de geluidsfragmenten worden omgezet naar MBROLA. De verkregen database wordt opgeslagen in Festival.

4.3 Het omzetten van de bestanden

Zoals eerder vermeldt is een gedeeltelijk Spaanstalig systeem omgezet naar een Fries systeem. Dit betrof de volgende bestanden: het lexicon (frislex.scm), de tokenisatie (fristoken.scm) en de intonatie (frisint.scm). De laatste versies van de betreffende bestanden zijn in de bijlagen terug te vinden. Hieronder wordt een korte toelichting gegeven. Vanuit oorspronkelijk Engelstalige bestanden die te maken hebben met MBROLA (de Spaanstalige versie van Festival had deze niet) zijn de volgende bestanden gecreëerd: de stemopbouw (an_mbrola.scm), de foneemset (mrpa_phones.scm), de koppelingsbestand naar MBROLA (mbrola.scm) en twee duur-bestanden (mrpa_durs.scm en gswdurtreeZ.scm).

4.3.1 Bouw van de stem (an_mbrola.scm)

Oorspronkelijk is de stem opgebouwd vanuit een bestand uit de Spaanstalige versie van Festival (el_diphone.scm, na bewerking an_diphone.scm genoemd). Echter, omdat later besloten is om met MBROLA te werken is overgegaan op het bestand en1_mbrola.scm uit de Engelstalige versie. Bestand an_diphone.scm is hierin volledig gekopieerd, waarna het bestand en1_mbrola.scm de naam an_mbrola.scm kreeg. Toen dit niet goed bleek te werken (er was sprake van /r/-deletie, terwijl hiertoe nergens opdracht was gegeven) is het de stemopbouw uit het oorspronkelijke en1_mbrola.scm-bestand gekopieerd naar het oorspronkelijke an_diphone.scm-bestand. Dit nieuw verkregen an_mbrola.scm-bestand werkt wel goed, vermoedelijk omdat de stemopbouw nu in dezelfde directory staat als het lexicon.

An_mbrola.scm is als volgt opgebouwd:

Eerst worden de juiste directory en de benodigde files geladen.

Hierna wordt een begin gemaakt met POS-tagging. In een zin komen accenten voor. Deze worden veelal op inhoudswoorden gemaakt. Om dit enigszins na te kunnen bouwen zal het systeem functiewoorden en inhoudswoorden moeten herkennen. Dit gebeurt in de

zogenaamde gpos-functie die de woordenlijst in de `frisian_guess_pos`-set gebruikt. Elk woord dat niet in deze set voorkomt wordt gedefinieerd als een inhoudswoord (Black e.a. (2000):86). Vermoedelijk zullen de (lexicale) voorzetsels uit deze lijst moeten worden gehaald, omdat zij in het Fries relatief meer accent krijgen.

Dan wordt over gegaan op de frasering (Phrase Breaks). Dit bestand bevat een CART-boom, zie ook 'frasering' paragraaf 2.2 en zorgt ervoor dat achter het laatste woord waarna een leesteken (de tekens die tussen haakjes staan) volgt, een pauze (oftewel grens van frase) komt. De eerste regel verzekert dat alleen achter het laatste woord een pauze komt. De BB staat voor Big Break, voor aan het einde van de uiting. De B staat voor een kleinere Break (Black e.a. (2000):85). De oorspronkelijke boom bevatte een regel die een pauze inlaste na alle aangegeven leestekens. De boom maakt nu onderscheid tussen eind van een uiting/zin en overige leestekens.

De accenten worden gemaakt via de lijst `frisian_accent_cart_tree`. Elk woord voor een functiewoord wordt hier als inhoudswoord gezien (mits dit woord dus geen functiewoord is). Hier krijgt de syllabe waar een 1 achter staat (dit is de syllabe met woordaccent), het daadwerkelijke accent (Black e.a. (2000):90).

In meeste talen duren fonemen langer op het eind van de frase dan aan het begin. Om dit in de synthese te bereiken kan men de duur met een factor vermenigvuldigen. Dit gebeurt in de volgende boomstructuur: `frisian_dur_tree`. In aanvulling hierop worden de gemiddelde duren van ieder foneem in de foneemset genoemd in `frisian_an_phone_data`. Hierbij is een schatting gemaakt van de gemiddelde duur door vocalen langer dan consonanten te laten duren en plosieven de kortste duur toe te schrijven (Black e.a. (2000):96-97). Als basis is de duur van de Spaanse fonemen genomen.

Met de overige opdrachten wordt de stem opgebouwd. Hierbij kunnen zowel de commando's uit de Spaanse stemopbouw als uit de Engelse stemopbouw gebruikt worden. Vooralsnog wordt bij intonatie gebruik gemaakt van de Engelstalige versie en duur van de Spaanstalige versie. Hier moet nog uitgebreid naar gekeken worden.

4.3.2 Foneemset (`mrpa_phones.scm`)

De set heeft een unieke naam (`frisian`). Hierna volgt een featurelijst ('featuredefs') waarin alle mogelijke kenmerken van de klanken worden gedefinieerd en hun mogelijke waarden worden weergegeven. In de 'phonedefs' worden alle vocalen en consonanten van de betreffende taal gedefinieerd en een waarde aan elk feature gegeven op de volgorde zoals de features gedefinieerd zijn in de 'featuredefs'. Er wordt tevens een definitie voor het stiltefoneem toegevoegd.

In de Spaanstalige versie worden vocalen genoemd met accent (de vocalen met een 1 erachter). Deze heb ik ook voor het Fries overgenomen, echter, ze worden nergens opgeroepen. Aangezien het kenmerk nasaliteit in het Fries mijns inziens veel voorkomt, is van alle vocalen een nasale versie aangegeven.

4.3.3 Lexicon (`frislex.scm`)

Wat in het lexicon aanwezig moet zijn, is de uitspraak van de afzonderlijke letters. Verder zijn nog enkele tekens (zoals sterretje, procent, etc.) en leestekens verwoord. In de downcase krijgen alle letters hun evenbeeld toegekend in kleine letters, dat wil zeggen dat bijvoorbeeld de hoofdletter "A" nu de kleine letter "a" krijgt toegekend.

Hierna komen de belangrijkste grafeem-naar-foneem regels (letter-to-sound rules). De moeilijkheid hierbij was dat in het Fries op veel vocalen een 'dakje' voorkomt. Deze heb ik aangemerkt door "^" voor het betreffende vocaal te zetten. Ook andere accenten zijn op

deze manier aangegeven. Hier zal nog aandacht aan geschonken moeten worden, ook bij keuze van de “konsole” in Linux. Niet elke “konsole” accepteert deze accenten.

Hier wordt begonnen met de definitie van enkele kenmerken (features) te definiëren. Zo zorgen de consonanten in het woord “soldaten” zorgen vaak voor een andere uitspraak als zij op een “a” volgen, deze wordt dan “>”. En wanneer zij op een “r” volgen, vervalt deze in de uitspraak (Dijkstra (1991):23,84; Cohen e.a. (1961):129).

Het stilte-foneem en de syllabe-streep worden gedefinieerd. En dan wordt er een begin gemaakt met de grafeem-naar-foneem regels. Zo wordt aandacht geschonken aan de syllabificatie van de /n/ en de uitspraak van lange en korte vocalen. Verder worden vocalen, consonanten, diftongen en triftong(en) gedefinieerd en wordt er aandacht besteed aan nasaliteit, zie voor deze regels paragraaf 3.1.1.

Enkele uitspraakregels:

-Bij een gesloten lettergreep heb je een korte klinker (Dijkstra (1991):119-121)

-Als een woord onder andere eindigt op -nen, -mem, -ken dan worden de eindsyllaben als één foneem uitgesproken, ook wel sonorant genoemd (Hoekstra e.a. (2001):2; Cohen e.a. (1961):124).

-Wanneer de laatste syllabe (suffix) een ‘e’ bevat en wel volledig wordt uitgesproken, wordt deze ‘e’ in veel gevallen een schwa of /&/. Ook in affixen met een ‘e’ verandert de ‘e’ in een /&/.

Naast enkele grafemen worden wanneer noodzakelijk ook grafeemcombinaties omgezet naar foneem(combinaties). Enkele regels worden hieronder toegelicht met betreffende referentie:

-Voor de /r/ klinkt de ‘eu’ als /Y&/: zie paragraaf 4.1.1.

-De lettercombinatie ‘uu’ komt alleen in leenwoorden voor (Dijkstra (1991):50). In alle tot dusver bekeken voorbeelden met deze combinatie wordt dit uitgesproken als /yj&/, vandaar dat ik deze combinatie aan de grafeem-naar-foneem regels heb toegevoegd.

-Voor de uitspraak van de ‘g’, zie paragraaf 4.1.1.

-Wanneer de combinatie ‘hj’ aan het begin van een woord staat, dan wordt de /h/ niet uitgesproken.

-De ‘ng’-grafemen staan voor de klank /N/. Het kan echter voorkomen dat deze combinatie gescheiden wordt door een syllabebreuk. In dit geval worden de letters afzonderlijk van elkaar uitgesproken. Dit gebeurt voornamelijk in woorden die beginnen met ‘ûng-’

Na de belangrijkste grafeem-naar-foneem regels zien we in het bestand de syllabificatie staan. Allereerst zien we de meest voorkomende clusters (van 2 karakters). Ik wil niet de indruk wekken dat deze lijst compleet is. De meeste clusters zijn afkomstig van de distributielijsten in Cohen e.a. ((1961):132-147) en het woordenboek van Zantema (1984).

Hierna volgen de gebroken vocaalgroepen. Dit gebeurt in het Fries meestal met behulp van een trema op of afbrekingsstreep voor de betreffende vocaal. Bij de (/mijn) uitspraak wordt echter vaak /j/ als overgangsemivocaal toegevoegd.

Het onderdeel frisian.stress heeft nog speciale aandacht nodig. Er staan enkele ‘uitprobeersels’ genoteerd, maar deze werken echter niet.

In het oorspronkelijke Spaanstalige bestand werd bij de opbouw van de grafeem-naar-foneem-regels gesproken van spanish_weak_vowels. Aangezien Fries dit onderscheid bij mijn weten niet (zo sterk) heeft is eerst geprobeerd deze toepassing te verwijderen. Dit lukte echter niet. Vandaar dat variabele frisian_weak_vowels gedefinieerd staat als alle consonanten en vocalen tot zichzelf. Elke variabele wordt trouwens hiermee afgesloten, anders worden de niet-genoemde klanken niet herkend en kunnen zij niet ‘door’ naar de volgende variabele.

Bij het testen van deze database kwam ik nog enkele obstakels tegen. De meeste heb ik op kunnen lossen, maar enkele zijn blijven zitten. Dit staat achter de betreffende regel/opdracht cursief aangegeven.

Tot slot wordt het lexicon afgesloten door ingangen te creëren voor het uitspraakwoordenboek en relaties te ondergaan met andere modulen. Aanvankelijk was het zo dat bij de 'opbouw' van de regels in het lexicon de uiting eerst 'door' de variabele `frisian_downcase` gehaald werd, dan werden de grafeem-naar-foneem regels toegepast en dan pas vond syllabificatie plaats. Omdat de syllabificatie niet werkte is besloten om na de `frisian_downcase` over te gaan tot syllabificatie en dan pas door te gaan met de grafeem-naar-foneem regels. Nu werkt alleen de variabele `frisian.stress` helaas (nog) niet.

Alle woorden die niet met de grafeem-naar-foneem regels kunnen worden opgebouwd krijgen een plekje in het uitspraakwoordenboek. In het uitspraakwoordenboek voor het Fries zouden onder andere de volgende woorden een plaats krijgen:

- functiewoorden als 'it', 'in', 'ris', 'jin' (i = /&/)
- alle woorden met open lettergreep eindigend op 'i', uitgesproken als /i/
- alle woorden waarbij de 'oe' als /u/ uitgesproken wordt (dus niet als de diftong /u&/)
- alle woorden waarbij de 'û' als diftong /u&/ wordt uitgesproken
- alle woorden waarbij de 'ú' wordt uitgesproken als /y:/
- alle woorden waarbij de 'u' (in open lettergreep) wordt uitgesproken als /y:/
- alle woorden waarbij 'ië' als /i j E/ wordt uitgesproken
- alle woorden waarbij 'ei' los wordt uitgesproken als /e:/ en //
- alle woorden waarbij de 'iuw' als /iu/ wordt uitgesproken
- alle woorden waarbij 'eau' als /Au/ wordt uitgesproken
- werkwoorden waarbij de uitspraak afwijkt
- woorden als 'jier', waar de /j/ niet wordt uitgesproken
- diminutieven waarbij het stamwoord eindigt op /st/, bij het diminutief vervalt de /t/
- dagen van de week en 'middei': 'dei' wordt /di/
- leenwoorden die niet kunnen worden opgebouwd uit de grafeem-naar-foneem regels

NB Dit overzicht is niet compleet.

4.3.4 Tokenisatie (`fristoken.scm`)

Deze module bepaalt voornamelijk de uitspraak van de nummers. De uitspraak van deze nummers is veranderd naar het Fries. Hoewel mijn kennis van het Spaans nihil is, begrijp ik dat getallen met twee cijfers of meer (boven de achttien) als "tien acht", "zestig twee", "honderd veertig vijf", etc. wordt uitgesproken. Ik heb getracht dit te veranderen in de Friese uitspraak: respectievelijk "achtjtjn", "twaënsechstich", "hûndertfiifenfjirtich".

Het vreemde bij de honderdtallen was dat, naast honderd, alleen vijf-, zeven- en negenhonderd waren uitgeschreven. Ik heb de honderdtallen van honderd tot negenhonderd uitgeschreven.

De duizendtallen gaan in de synthese waarschijnlijk verkeerd. Als ik het goed begrepen heb, zegt men in het Spaans eerst het duizendtal, dan het honderdtal etc. In het Nederlands en het Fries telt men door in honderdtallen: 2573 wordt bijvoorbeeld vijfentwintighonderd drieënzeventig. Ik bezit niet genoeg kennis van Scheme om dit nu te veranderen in de juiste uitspraak.

4.3.5 Intonatie en duur (`frisint.scm`, `mrpa_durs.scm`, `gswdurtreeZ.scm`)

Voor de intonatie is vooralsnog gekozen voor de Engelstalige versie. Het bestand `frisint.scm` uit de Spaanstalige versie is echter ook een optie, nadat de pitch aangepast is aan mijn stem.

Bij de duur wordt in `an_mbrola.scm` vooralsnog teruggeschakeld op `frisian_dur_tree` en `frisian_an_phone_data` (beide beschreven in `an_mbrola.scm`). Maar men kan ook de bestanden `mrpa_durs.scm` en `gswdurtreeZ.scm` uit de Engelstalige versie gebruiken. Omdat ik de precieze functies in deze twee bestanden nog niet precies begrijp, ben ik hier nog niet op overgeschakeld.

Kortom, intonatie en duur zullen nog de nodige aandacht moeten krijgen!

4.3.6 Koppeling naar MBROLA (`mbrola.scm`)

Zoals eerder gezegd, zorgt dit bestand voor koppeling van de fonetische en prosodische informatie uit Festival aan de synthesizer van MBROLA. Vanuit het bestand `an_mbrola.scm` wordt de MBROLA-Synthesizer aangeroepen. Deze wordt in `mbrola.scm` gedefinieerd. De uiting wordt als `.au`-bestand opgebouwd. Na afspelen wordt dit geluidsfragment weer verwijderd.

5. Evaluatie van de spraaksynthesizer

Zolang er een verschil hoorbaar is tussen synthetische spraak en menselijke spraak is een goede evaluatie van belang. Omdat ontwerpers van TTS-systemen met hun systeem meegroeien, zullen zij de output van het systeem eerder begrijpen dan mensen die het systeem voor het eerst gebruiken. Met een goede evaluatie kan men ontdekken hoe goed een systeem presteert. Bovendien kan men dan meerdere systemen met elkaar vergelijken (Van Heuven e.a. (1995):709-710). Black e.a. ((2000):141) hameren erop dat het evalueren echter geen taak op zichzelf is. Men kan de testdata bijvoorbeeld overtrainen wat weer een onbetrouwbare uitkomst tot gevolg heeft.

Voordat men tot de echte evaluatie overgaat, kan het geen kwaad de afzonderlijke delen even te testen (of ze überhaupt werken). Wanneer blijkt dat elk onderdeel het doet, raden Black e.a. ((2000):141-143) aan om 100-500 zinnen te verzamelen en door het systeem te halen alvorens te testen met een echt tekstfragment. Dit fragment kan bijvoorbeeld het nieuws van internet (www.omroepfryslan.nl; of de column 'dwers' uit de Leeuwarder Courant van Pieter de Groot (www.leeuwardercourant.nl)) of een e-mail zijn. Hierdoor kan men direct ontdekken waar de knelpunten zitten en kijken of deze problemen het waard zijn om op te lossen. Men stuit bijvoorbeeld op missende woorden in het lexicon. Als men een lijn kan ontdekken waarbij het steeds fout gaat bij bepaalde woorden, kan men dit oplossen door nieuwe letter-to-sound-rules toe te voegen. Anders is het opnemen van deze woorden in het uitspraakwoordenboek een oplossing.

In Van Heuven e.a. (1995) passeren veel evaluatietesten de revue. Ook op de website van Disk2 is een overzicht van bestaande testen te lezen. Veel testen zijn echter irrelevant, aangezien ze taalspecifiek zijn en voor het Fries niet bestaan. De enige linguïstische test die in aanmerking zou kunnen komen voor deze synthesizer is de Semantically Unpredictable Sentences (SUS). Deze test wordt ook door Black e.a. ((2000):143) genoemd, al bestaat er enige discussie over. De SUS is ontwikkeld door SAM. Het is een gefixeerde verzameling van 5 syntactische structuren die veel voorkomen in de meeste West-Europese talen. Deze zinnen bevatten lexicale open ruimtes en die kunnen worden opgevuld met hoogfrequente woorden van uit de taalspecifieke lexica (Van Heuven e.a. (1995):725). De 5 structuren zijn:

Subject – Verb – Adverbial : *The table walked through the blue truth.*

Subject – Verb – Direct Object: *The strong way drank the day.*

Adverbial – Transitive Verb – Direct Object (imperative): *Never draw the house and the fact.*

O-Word – Transitive Verb – Subject – Direct Object: *How does the day love the bright word?*

Subject – Verb – Complex Direct Object: *The place closed the fish that lived.*
(Disk2 (1999/2000)).

Deze structuren komen ook in het Fries voor (Tiersma (1985):103-139), al heb ik geen duidelijk bewijs kunnen vinden voor de derde structuur. Bij deze structuur vind ik alleen zinnen met het Verb op het eind.

Het wordt aanbevolen om 50 zinnen, dat wil zeggen 10 zinnen per structuur, bij deze test door het systeem te halen (Disk2 (1999/2000)).

Er zijn ondertussen studies gedaan in verschillende talen met positief resultaat. De SUS kan worden gebruikt als een betrouwbare vergelijking tussen synthesizers, mits de richtlijnen goed worden opgevolgd (Benoit e.a. (1996)). Het exacte doel van deze test is mij echter nog steeds wat onduidelijk. Ik kan er echter niet in slagen meer doeltreffendere informatie over SUS te vinden.

Het meest efficiënt is om met een algemene test de output van het systeem te testen. Van Heuven e.a. ((1995):32) raden aan om het testen zoveel mogelijk in het veld te doen. Dat wil zeggen dat men zoveel mogelijk het systeem in de toepassing moet testen: in dezelfde

setting, en de doelgroep als proefpersonen gebruiken. Testen met proefpersonen geeft de beste resultaten aangezien zij representatief zijn voor de gebruikers van het systeem. Het leereffect treedt binnen een paar minuten op (Van Heuven e.a. (1995):718-719). Men zou kunnen denken aan de SAM Overall Quality Test waarbij beoordeeld wordt op acceptatie, verstaanbaarheid en natuurlijkheid. De testen van de laatste twee kwaliteiten zijn gebaseerd op paren van zinnen (Van Heuven e.a. (1995):173). Om deze test te gebruiken zullen deze gefixeerde lijsten naar het Fries moeten worden vertaald. Ik weet echter niet hoever dit de betrouwbaarheid van de testen schaadt.

Hoewel het testen met proefpersonen mij meest effectief lijkt, is het een erg kostbare en tijdrovende onderneming. En bovendien is nooit zeker in hoeverre proefpersonen de instructies navolgen, wat de betrouwbaarheid van de uitkomsten niet ten goed komt. Wanneer men deze spraaksynthesizer toepast in als telefonisch informatiesysteem zullen proefpersonen moeten worden gevraagd een (gratis) nummer te bellen en informatie op te vragen. Men kan deze dienst bijvoorbeeld naast een gelijke, al bestaande dienst laten draaien, bijvoorbeeld een Nederlandstalig informatiesysteem of een telefoniste die de klant van informatie kan voorzien.

Bij toepassing van het informatiesysteem als leesmachine voor blinden en slechtzienden zal men blinden en slechtzienden als proefpersonen nodig hebben. Deze kan men oproepen door de NVBS (Nederlandse Vereniging voor Blinden en Slechtzienden), regio Friesland in te schakelen. Bij de toepassing als hulpmiddel bij mensen met een communicatiestoornis is waarschijnlijk persoonlijke aanpassing noodzakelijk. Evaluatie zal men hier per persoon moeten bekijken.

6. Conclusie en suggesties

In dit werkstuk is geprobeerd een overzicht te geven hoe men een Friestalige spraaksynthesizer bouwt. Men heeft een foneemset nodig, een uitspraakwoordenboek, etc. Men moet difonen opnemen en verwerken en gegevens in bestanden aanpassen aan het Fries. Spijtig is dat veel benodigdheden zoals het uitspraakwoordenboek en gegevens over woordaccent en woordsoort nog niet volledig digitaal ontwikkeld zijn. Men zou het restant allemaal met de hand in moeten voeren, een tijdrovend karwei.

Belangrijk is ook de evaluatie, want men moet ook naar kwaliteit streven. Echte evaluatietesten zijn er (nog) niet op het gebied van het Fries dus, evalueren met hulp van proefpersonen is misschien nog het meest effectief.

Toch is het de moeite waard om een Friestalige synthesizer te bouwen. Het is belangrijk voor een minderheidstaal om met de technologische ontwikkelingen mee te gaan. Als sprekers bij gebruik van spraaktechnologische middelen gedwongen zijn om de meerderheidstaal te gebruiken, zal dit uiteindelijk leiden tot een afname van het aantal sprekers van de minderheidstaal, zoals de SALTMIL, de Special Interest Group van de ESCA voorspelt (Nadeu e.a. (2001:1).

Waar in Fryslân gebruik zou kunnen worden gemaakt van Nederlandse spraaksynthese, kan ook gebruik worden gemaakt van Friestalige synthese. Bijvoorbeeld als leesmachine voor blinden en slechtzienden, als hulpmiddel voor mensen met een handicap in de communicatie of als informatiesysteem bij bank, VVV of reserveringslijn voor Tryater of bioscoop.

Omdat dit werkstuk voornamelijk op papier de praktische kant belicht en er alleen naar de uitvoer van één Fries woord wordt gekeken, is het aan te raden om de synthesizer af te bouwen en geschikt te maken voor grotere teksten. Als deze synthesizer dan naar tevredenheid werkt, zou men, naast bovengenoemde toepassingen, met behulp van de synthesizer ook veel onderzoek kunnen doen. Bijvoorbeeld onderzoek op het gebied van de intonatie van het Fries, een achtergesteld gebied.

Een Friestalige spraaksynthesizer is van groot belang. Niet alleen vanwege de vele toepassingen en het belang voor het Fries, maar ook vanwege verdere mogelijkheden op het gebied van onderzoek.

Geraadpleegde literatuur

Benoit, C.; Grice, M.; Hazan, V.

The SUS test: a method for the assesment of text-to-speech synthesis intelligibility using Semantically Unpredictable Sentences.

Abstract, 1996

<http://www.phon.ucl.ac.uk/cgi-bin/w3-mysql/pub/detail.htm?PID=75&function =Load>

Black, A.W.; Lenzo, K.A.

Building Voices in the Festival Speech Synthesis System

Processes and issues in building speech synthesis voices

Edition 1.2:beta, for Festival Version 1.4.1

10th July 2000

http://festvox.org/festvox/festvox_toc.html

Black, A.W.; Taylor, P.; Caley, R.

The Festival Speech Synthesis System

System documentation

Edition 1.4, for Festival Version 1.4.0

17th June 1999

http://www.cstr.ed.ac.uk/projects/festival/manual/festival_toc.html

Cohen, A.; Ebeling, C.L.; Fokkema, K.; Holk, A.G.F., van

Fonologie van het Nederlands en het Fries

Inleiding tot de moderne klankleer

Martinus Nijhoff, 's-Gravenhage, 1961

Dijkstra, F.B., drs.

Stavering

Oersjoch fan de staveringsregels yn it Frysk

Afûk, Ljouwert, 1991

Disc2: <http://www.disc2.dk/tools/Sgsurvey.html>

A survey of existing methods and tools for developing and evaluation of speech synthesis and of commercial speech synthesis systems

1999/2000

Dutoit, T.; Pagel, V.; Pierret, N.; Bataille, F.; Vrecken, O., van der

"The MBROLA Project: Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes"

Proc. ICSLP'96, Philadelphia, vol. 3, pp. 1393-1396

Fryske Akademy

Het wetenschappelijk onderzoekscentrum voor het Fries

Informatie over het WFT en het Frysktalich hânwurdbboek.

www.fa.knaw.nl

Fryske Beweging

Een organisatie die zich bezig houdt met het in stand houden van het Fries

Friestalige computerterminologie.

www.fryskebeweging.nl

Gorter, D.

Dutch - West Frisian

uit: Contact Linguistics - An international Handbook of Contemporary Research

Goebel, H.; Nelde, P.H.; Starý, Z.; Wölck, W.

p. 1152-1157, Volume 2

Walter de Gruyter, Berlin, New York, 1997

Gorter, D.; Jonkman, R.J.

Taal yn Fryslân op 'e nij besjoen

Fryske Akademy, Ljouwert, 1995

Haan, R., de
Terminology fan 'e saakfakken
Taalbuero Fryske Akademy, 2000

Heuven, V.J., van; Bezooijen, R., van
Quality Evaluation of Synthesized Speech
Uit: Speech Coding and Synthesis
Kleijn, W.B.; Paliwal, K.K.
Elsevier Science BV, Amsterdam, 1995

Hieronymus, J.L.
ASCII Phonetic Symbols for the World's Languages: Worldbet
AT&T Bell Laboratories, Murray Hill, USA, 1994
<ftp://speech.cse.ogi.edu/pub/docs/worldbet.ps>

Hoekstra, E.
Korpus Sprutsen Frysk
Fryske Akademy, Ljouwert, 2001
www.fa.knaw.nl

Hoekstra, E.; Siebenga, S.
SAMPA for Frisian and Comments
Fryske Akademy, Ljouwert, 2001

Hoekstra, J.
Oer it beklamjen fan ferhâldingswurden yn it Frysk, it Hollânsk en it Ingelsk
Us Wurk, Volume 40, p. 67-103
Fryske Akademy, 1991

Holmes, J.; Holmes, W.
Speech Synthesis and Recognition, 2nd edition
Taylor and Francis, London, 2001

Munske H.H., Århammer, N.
Handbuch des Friesischen/Handbook of Frisian Studies
Niemeyer, Tübingen, 2001

Nadeu, C.; Ó'Cróinín, D.; Petek, B.; Sarasola, K.; Williams, B.
ISCA SALT MIL SIG: Speech and Language Technology for Minority Languages
Proc. Eurospeech 2001, Alborg Dinamarca, september 2001
<http://gps-tsc.upc.es/veu/research/pubs/download/Nad01b.pdf>

Olive, J.P.
The Talking Computer: Text to Speech Synthesis
HAL's Legacy: 2001's Computer as Dream and Reality, 2001
Urbana, Illinois, USA, 2001
<http://mitpress.mit.edu/e-books/Hal/chap6/six1.htm>

Rog, Judith
Het Festival Spraak Synthese Systeem
Spraaktechnologie, UvA, 1999

Taylor, P.; Black, A.W.; Caley, R.
The Architecture of the Festival Speech Synthesis System
ESCA Workshop on Speech Synthesis, Jenolan Caves, NSW, Australia, 1998

Tiersma, P.M.
Frisian Reference grammar
Fryske Akademy
Dordrecht Foris Publications, 1985

Zanema, J.W.
Frysk Wurdboek Frysk-Nederlânsk, 12^e druk
Fryske Akademy, Ljouwert, 1992
O.J. Osinga Uitgeverij, Drachten/Ljouwert