

Diftongen van dove en horende sprekers: productie en perceptieve beoordeling

E. Gerrits¹, B.A.G. Elsendoorn² en W.J.M. Peeters¹

¹*Utrecht Institute of Linguistics OTS, Universiteit Utrecht* en ²*Instituut voor Doven, Sint-Michiëlsgestel*

In dit onderzoek zijn de uitspraak en de perceptieve beoordeling bestudeerd van de diftongen [ɔu], [ɛi] en [ʌy] van dove sprekers, logopedisten en ongetrainde sprekers. Wat betreft de uitspraak zijn de duur en de formantpatronen van de diftongen van deze groepen sprekers vergeleken. De resultaten laten zien dat de duur van de diftongen van dove sprekers langer is dan die van logopedisten en ongetrainde sprekers. Er zijn echter ook duurverschillen tussen de diftongen van de horende sprekers: de diftongen van de logopedisten zijn langer dan die van de ongetrainde sprekers. Het effect van de sprekers komt terug in de verschillende formantpatronen. De tweede formant van de diftongen van de horende sprekers bevat een snelle stijging die ontbreekt in de tweede formant van de diftongen van de dove sprekers. Het ontbreken van de verglijding in de tweede formant duidt er op dat dove sprekers bij de uitspraak van diftongen vooral moeite hebben met de voor-achterwaartse tongbeweging die deze verglijding bepaalt. Dit gegeven suggereert dat aandacht voor deze tongbeweging in logopedische therapie de uitspraak van diftongen door dove sprekers zou kunnen verbeteren.

In een serie luisterexperimenten werd de invloed van een variërende duur en/of formantfrequenties op de perceptie van de diftongen bestudeerd. Met behulp van spraaksynthese werden standaard diftongen gegenereerd waarin de duurwaarden en/of de formantfrequenties aangepast werden aan de data uit het productie-experiment. Getrainde en ongetrainde luisteraars gaven een oordeel over de acceptabiliteit van deze synthetische diftongen. De diftongen met sprekersspecifieke waarden van de dove sprekers werden als minst acceptabel beschouwd. Dit wordt met name veroorzaakt door een afwijkende spectrale compositie. Daarnaast hadden de luisteraars meer voorkeur voor de diftongen van de ongetrainde sprekers dan voor die van de logopedisten.

Inleiding

Een opmerking die vaak over dovenspraak gemaakt wordt is dat deze onnatuurlijk

Correspondentieadres: Ellen Gerrits, Utrecht Institute of Linguistics OTS, Universiteit Utrecht, Trans 10, 3512 JK, Utrecht. E-mail adres: Ellen.Gerrits@let.uu.nl.

klinkt en slecht verstaanbaar is. Om de uitspraak te verbeteren zijn er diverse onderzoeken gedaan, waarin is getracht de fouten in dovenspraak in kaart te brengen. Deze studies richtten zich onder meer op de gemiddelde verstaanbaarheid van dovenspraak en op de inventarisatie van herkenbare foutpatronen in dovenspraak. Het spraakmateriaal dat voor de foutpatronen bestudeerd is bestond meestal uit een selectie van fonemen, bijvoorbeeld de vocalen of de consonanten. Bij beide groepen klanken blijken afwijkingen in de dovenspraak voor te komen in vergelijking met normale spraak. Een aparte groep klanken binnen de vocalen zijn de diftongen (tweeklanken). De Nederlandse taal kent drie diftongen: de [au], [ʌy] en [ɛi] (Collier, Bell-Berti & Raphael, 1982). Uit onderzoek naar diftongen in dovenspraak blijkt dat deze vaak nog meer vervormd worden dan de overige vocalen (Markides, 1970). Deze vervorming wordt veroorzaakt door de complexe articulatie van de diftong: een beweging van een beginpositie naar een eindpositie. Hoewel diftongen minder vaak in spraak voorkomen dan consonanten heeft juist de onjuiste uitspraak van diftongen en de overige vocalen een grote negatieve invloed op de verstaanbaarheid van dovenspraak (Smith, 1975).

De praktijkervaring van logopedisten van het Instituut voor Doven in St. Michielsgestel is in overeenstemming met de bevindingen van Markides (1970). Het aanleren van een goede uitspraak van diftongen aan de dove leerlingen levert veel moeilijkheden op en de diftong blijft doorgaans afwijkend klinken. Het leek daarom wenselijk om met behulp van objectieve registratietechnieken vast te stellen op welke manier diftongen door dove sprekers geproduceerd worden en deze te vergelijken met de productie van horende sprekers. Dit vormt de basis voor de hoofdvraag in dit onderzoek waarin een vergelijking wordt gemaakt tussen de diftongen [au], [ʌy] en [ɛi], uitgesproken door dove sprekers, ongetrainde horende sprekers en logopedisten. Zowel de productie als de perceptie van de diftongen is bestudeerd.

In dit artikel zal in het eerste deel 'productie' een analyse van de duur en formantfrequenties van de diftongen worden besproken. Het tweede deel, 'perceptieve beoordeling', beschrijft de vergelijking van synthetische versies van de diftongen van de drie spreker groepen door middel van acceptabiliteitsoordelen van getrainde en ongetrainde luisteraars.

Deel 1: productie

Uit onderzoek naar de foutpatronen in dovenspraak komt naar voren dat consonanten vaak worden weggelaten (91% van alle fouten) of vervangen (Markides, 1970; Smith, 1975). In de productie van vocalen worden 16% minder fouten gemaakt dan in de productie van consonanten (Markides, 1970), maar de onjuiste uitspraak van de vocalen heeft een grotere negatieve invloed op de verstaanbaarheid van dovenspraak dan de fouten in de uitspraak van de consonanten (Smith, 1975). De meest voorkomende fouten in de productie van vocalen zijn:

- neutralisatie;
- het diftongeren van monoftongen;
- substitutie van de scherpe naar de doffe vocaal, bijvoorbeeld [i] wordt [ɪ];

– verlenging van de vocaal.

(Markides, 1970; Smith, 1975; Monsen, 1978; Geffner, 1980; Levitt, 1980; McGarr, 1983; Gulian, 1983; Maassen, 1985; Tye-Murray & Iler Kirk, 1993)

De articulatiefouten komen voor in zowel het temporele als in het spectrale domein. De temporele verschillen tussen de uitspraak van de [i] en de [ɪ] door dove en horende sprekers werd bestudeerd door Monsen (1974). In tegenstelling tot de, onder invloed van de context, relatieve duurverschillen tussen de [i] en de [ɪ] in de spraak van de horenden bleek dat er in dovenspraak constante duurverschillen optraden tussen deze klanken.

Wat het spectrale domein betreft blijkt dat in het algemeen vocalen van dove sprekers een beperkte variabiliteit in formantfrequenties bevatten, veroorzaakt door een beperkte tongbeweging (Angelocci, 1964; Monsen, 1976; Suonpää, 1981; McGarr & Gelfer, 1983; Gulian, 1983). Geffner (1980), Smith (1975) en Tye-Murray & Iler-Kirk (1993) vermelden dat de vocalen met lage formantfrequenties, de zogenaamde achterklinkers, het beste worden geproduceerd door dove sprekers. Als verklaring hiervoor wordt genoemd dat deze klanken vaak nog enigszins hoorbaar of voelbaar zijn. Dat restgehoor belangrijk is wordt bevestigd door een cinefluografisch onderzoek naar articulatiebewegingen van Tye-Murray (1991). Voor de implantatie van een binnenoorprothese maakten dove sprekers dezelfde tongbewegingen voor de productie van verschillende vocalen. Hierna trad er een verschuiving op van de productie van vooral centrale klinkers naar een breed scala aan voorklinkers, centrale klinkers en achterklinkers (Tye-Murray, 1991). Dat achterklinkers het beste worden geproduceerd door dove sprekers wordt echter tegengesproken door Suonpää (1981), McGarr & Gelfer (1983) en Gulian (1983). Zij beweren dat de hoge- ofwel voorklinkers het beste worden uitgesproken. McGarr & Gelfer (1983) geven als verklaring voor de betere uitspraak van hoge klinkers door dove sprekers dat de articulatiebeweging bij deze klanken beter zichtbaar is. Ook Geffner (1980), Suonpää (1981) en Mencke (1985) noemen de positieve relatie tussen visuele spraakperceptie (spraakafzien) en de fooneemproductie van dove sprekers. Binnen de groep vocalen is de uitspraak van diftongen het meest problematisch voor dove sprekers. De fouten die genoemd worden zijn:

- substitutie van de diftong door de neutrale schwa;
 - omissie van de tweede component;
 - overmatige verlenging waardoor de diftong opgesplitst wordt in twee klanken.
- (Markides, 1970; Levitt, 1980; Tye-Murray & Iler Kirk, 1993)

Uit een analyse van spontane spraak door Tye-Murray & Iler Kirk (1993) bleek dat slechts 25% van het aantal diftongen dat voorkwam correct werd uitgesproken door dove sprekers. Monsen (1978) onderzocht een aantal akoestische eigenschappen van de Engelse diftong [ai], uitgesproken door dove sprekers. De frequentieverandering in de eerste formant (F1) van de diftong was voor horende en dove sprekers gelijk, maar de snelle stijging in formantfrequentie in het verloop van de tweede formant (F2), die duidelijk zichtbaar was in het formantspoor van de diftongen van de horende sprekers, ontbrak in de diftongen van de dove sprekers. Juist de variabiliteit van de tweede for-

mant had een hogere correlatie met de verstaanbaarheid van dovenspraak dan de variabiliteit van de eerste formant (Monsen, 1978). Het ontbreken van verschillen in het verloop van de F1 wordt verklaard doordat de eerste formant (lage frequenties) beter hoorbaar is voor de dove luisteraar, omdat de gehoorverliezen van de dove sprekers meestal bij de hoge frequenties ernstiger zijn dan bij de lage frequenties (Monsen, 1978; Gulian, 1983, Tye-Murray & Iler Kirk, 1993). Dit laatste argument komt overeen met de interpretatie van Geffner (1980) voor wat betreft de duidelijke productie van lage vocalen door dove sprekers. De tweede formant van de diftong met de hogere frequenties is dus minder hoorbaar dan de eerste formant. Daarnaast vereist de snelle verandering van de frequenties van de tweede formant snelle articulatiebewegingen, met name snelle voor- en achterwaartse tongbewegingen. Omdat deze tongbewegingen nauwelijks zichtbaar zijn bij andere sprekers, is het moeilijk voor dove sprekers deze bewegingen goed te imiteren. De hoog-laag tongbewegingen en de kaakbewegingen die voor een groot deel bepalend zijn voor de F1 zijn beter zichtbaar en kunnen daarom ook beter geïmiteerd worden door dove sprekers.

De diftong blijkt niet alleen voor dove sprekers complex te zijn, ook als concept is de diftong complex gezien de verschillende termen en definities die worden gebruikt om deze klank te beschrijven. Termen die voorkomen zijn ‘double-target’ versus ‘single target’, ‘phonemical’ versus ‘phonetic’, maar er is geen duidelijk diftongconcept (Lehiste & Peterson, 1961; Collier, Bell-Berti & Raphael, 1982; Bond, 1982; Neary & Assmann, 1986). Collier, et al (1982) en Peeters (1991) maken onderscheid tussen ‘echte’ en ‘pseudo’ diftongen. Bij een ‘echte’ diftong vormen de articulatiebewegingen een eenheid. Bij de ‘pseudo’ diftong kan de articulatiebeweging omschreven worden als een opeenvolging van componenten. Volgens Collier et al (1982) zijn de ‘echte’ diftongen in het Nederlands de [au], [ɛi] en [ʌy] (in dit artikel wordt het woord diftong gebruikt wanneer deze drie klanken bedoeld worden).

In de literatuur worden de akoestische kenmerken van diftongen beschreven volgens de ‘Target’-theorie en/of de ‘Trajectory’-theorie. In de Target-theorie worden de begin- en eindfrequenties van de eerste drie formanten van de diftongen als belangrijkste distinctief kenmerk beschouwd (Bond, 1982; Bladon, 1985; Pols, 1977; Nabelek, Czyzewski, & Crowley, 1994; Neary & Assmann, 1986; Lehiste & Peterson, 1961). De ‘Trajectory’-theorie legt daarentegen veel nadruk op de richting en de snelheid van de transitie tussen deze begin- en eindposities (Nabelek, Czyzewski, & Crowley, 1994; Gerber, 1971; Pols, 1977; Neary & Assmann, 1986; Gay, 1968). Deze theorieën richtten zich beide vooral op de componenten van de formanten. Nadat er in eerste instantie veel onderzoek werd uitgevoerd naar de afzonderlijke componenten van de formanten van de diftongen is er sinds vijftien jaar meer aandacht voor het totale formantpatroon (Toledo, 1987; Yang, 1987; Peeters, 1991; Stollwerck, 1991; Cao, 1991; Lindau, Norlin, & Svantesson, 1990). Peeters (1991) onderscheidt vier mogelijke formantpatronen:

- steady-state / transitie / steady-state
- steady-state / transitie
- transitie / steady-state
- transitie

Volgens Peeters (1991) zijn juist de onderlinge verhoudingen van de componenten in het formantspoor van diftongen karakteristiek voor deze klanken en is het temporele verloop van de formanten taalspecifiek.

Onderzoeksvragen

In dit onderzoek zijn verschillen tussen spraak van dove en horende sprekers bestudeerd. Omdat diftongen specifieke uitspraakproblemen geven is er een vergelijking gemaakt tussen de uitspraak van deze klanken door dove sprekers en logopedisten van het Instituut voor Doven in St. Michielsgestel. De uitspraak van de diftongen van deze sprekers is ook vergeleken met die van ongetrainde horende sprekers. Deze groep werd toegevoegd om te zien of er verschillen waren tussen getrainde en ongetrainde sprekers. De vergelijking van de diftongen bestond uit een analyse van de duur en de formantpatronen. De onderzoeksvragen in dit experiment waren:

- 1 Zijn er significante verschillen in de duur van diftongen, uitgesproken door ongetrainde sprekers, logopedisten en dove sprekers?
- 2 Zijn er significante verschillen in de formantpatronen van diftongen, uitgesproken door ongetrainde sprekers, logopedisten en dove sprekers?

Methode

Sprekers

Er waren drie groepen mannelijke sprekers. De groep dove sprekers bestond uit vier dove leerlingen (oraal-auraal onderwijs) van het Instituut voor Doven in St. Michielsgestel. De sprekers waren prélinguaal doof, het gemiddelde gehoorverlies was 103 dB (FI). De gemiddelde leeftijd was 16 jaar. Tijdens de opnamen volgden zij het voortgezet speciaal onderwijs van het Instituut voor Doven. De tweede groep bestond uit vier logopedisten van het Instituut voor Doven te St. Michielsgestel. De gemiddelde leeftijd was 45 jaar. Zij hadden allen een ruime ervaring in het geven van spreeklessen aan dove kinderen. De derde groep, de ongetrainde sprekers, werd gevormd door vier studenten van de Katholieke Universiteit in Nijmegen. Ze hadden een gemiddelde leeftijd van 25 jaar. De studenten kwamen oorspronkelijk uit verschillende delen van Nederland, maar waren al enige jaren woonachtig in Nijmegen. Zij hadden allen het Nederlands als moedertaal.

Spraakmateriaal

Het spraakmateriaal bestond uit de diftongen [au], [ɛi] en [ʌy], uitgesproken in woorden en in zinnen (zie appendix). De woorden bestonden uit de volgende consonant-vocaal combinaties: CV (open context), CVC, CVCV en CVCVC (gesloten context). De eerste vocaal was de diftong, de laatste een schwa. De consonantcontext bestond uit [p, t, l]. Voor de diftong op zinsniveau werden bij de bovenstaande woorden zinnen gevormd. Het woord met de diftong werd in het midden van de zin uitge-

sproken. In totaal bestond het spraakmateriaal uit vijf woorden en vijf zinnen per diftong.

Van elk van de 12 sprekers werden opnamen gemaakt van 15 woorden en 15 zinnen, uitgesproken op een normaal leestempo. De logopedisten en ongetrainde sprekers lazen het materiaal nog een tweede keer voor, met als doel een uitspraakvariant te krijgen die model staat voor de dove sprekers. Om dit te bewerkstelligen werd aan de logopedisten en de ongetrainde sprekers gevraagd de diftongen te produceren, alsof ze tegen een dove persoon spraken. Door deze extra spreekconditie toe te voegen werd het mogelijk te onderzoeken of de manier waarop logopedisten tegen dove luisteraars spreken afwijkt van de manier waarop ongetrainde sprekers dit doen. De extra spreekconditie wordt aangeduid als de conditie 'met nadruk' (MN). Het spraakmateriaal wordt voortaan als volgt omschreven: dove sprekers, logopedisten, logopedisten (MN), ongetrainde sprekers, ongetrainde sprekers (MN).

Procedure

De geluidsopnamen van de logopedisten en dove sprekers vonden plaats in een geluidsarme ruimte op het Instituut voor Doven. De apparatuur die gebruikt werd bestond uit een digitale taperecorder en microfoon. De opnamen van de ongetrainde sprekers gebeurden in de geluidsstudio van de Vakgroep Taal en Spraak van de Katholieke Universiteit te Nijmegen. Hiervoor werd gebruik gemaakt van een dat-recorder, een Studer mengtafel (type 089) en een condensatormicrofoon (Sennheiser MKH 415.T) met een super-carioïde richtingskarakteristiek. De bemonsteringsfrequentie van de digitale opname was 48 kHz.

Het stimulusmateriaal werd aangeboden met behulp van kaartjes. Op elk kaartje stond een woord of een zin gedrukt. De spreker las deze een voor een voor. De eerste vier uitgesproken woorden en zinnen werden uitgesloten van de data-analyse. De tijd die de spreker nodig had om deze kaartjes voor te lezen werd gebruikt om de opname-apparatuur af te stellen. Voor elke diftong werden eerst de vijf kaartjes met de woorden aangeboden, gevolgd door de vijf kaartjes met de zinnen. Bij de dove sprekers bleken er veel stoorgeluiden op te treden tijdens het zelfstandig hanteren van de kaartjes. Dit werd opgelost door deze via een derde persoon aan te bieden.

Akoestisch-dynamische analyse

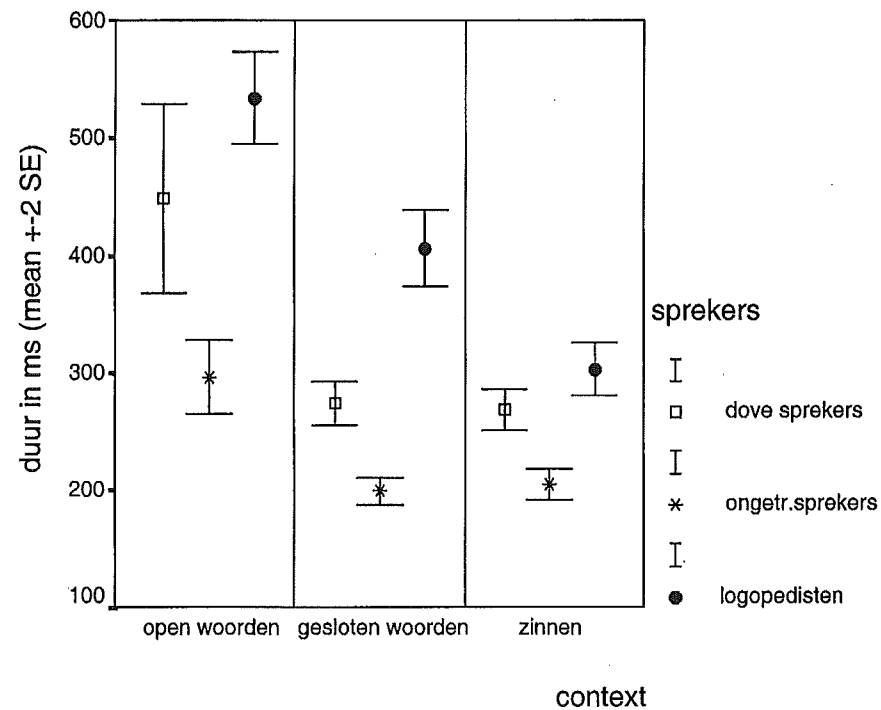
Om de duur en de formantfrequenties te kunnen bepalen werd het spraakmateriaal gedigitaliseerd met een bemonsteringsfrequentie van 10 kHz. De diftongen werden uit de woorden en zinnen gesegmenteerd met behulp van auditieve en visuele feedback. Met behulp van een signaaleditor werd de duur van de diftongen bepaald. Met behulp van LPC-technieken werd voor elk frame (10ms) de frequenties bepaald van de eerste vijf formanten met de bijbehorende bandbreedtes. De meetgegevens van de eerste, tweede en derde formant en de bijbehorende bandbreedtes werden gebruikt als data. Om het temporele patroon van de formanten te kunnen bestuderen, werd een tijdsnormalisatie van de spectrale data uitgevoerd. Dit betekende dat in plaats van te volstaan met een gemiddelde frequentie voor elke formant, nu voor elke formant de begin- en eindformantfrequentie werd bepaald, alsmede de frequenties op acht tijdsequidistante

meetpunten hier tussenin (Holbrooks & Fairbanks, in: Gottfried, 1993; Clermont, 1993; Stollwerck, 1991). Met behulp van een analyse op tien meetpunten bleef het dynamische karakter van de diftong bewaard. Daarnaast was het aantal meetpunten voor elke diftong gelijk en was de transitie in het formantspoor herkenbaar.

Resultaten

Duur

In figuur 1 zijn de duurverschillen tussen de diftongen van de drie groepen sprekers weergegeven. De resultaten zijn verdeeld over de drie contexten: open woorden, gesloten woorden en zinnen. Om een eerlijke vergelijking te maken tussen de uitspraak van de dove en horende sprekers werden de diftongen van de dove sprekers vergeleken met de diftongen van de groep logopedisten en ongetrainde sprekers die werden uitgesproken in de spreekconditie 'met nadruk' (MN). Deze diftongen zijn het 'model' voor de dove sprekers. De diftongen van de horende sprekers die in de 'normale' spreekconditie waren uitgesproken waren korter dan die in de conditie MN, maar de verhouding van de duren tussen de sprekergruppen was in beide condities hetzelfde. Er waren



Figuur 1 De gemiddelde duur (ms) van de diftongen [au], [ay] en [ei], geproduceerd in drie stimulus contexten door dove sprekers, ongetrainde sprekers (MN) en logopedisten (MN)

geen significante verschillen tussen de duren van de [au], [ɛi] en [ʌy], zodat gemiddelde duur van deze drie diftongen is weergegeven.

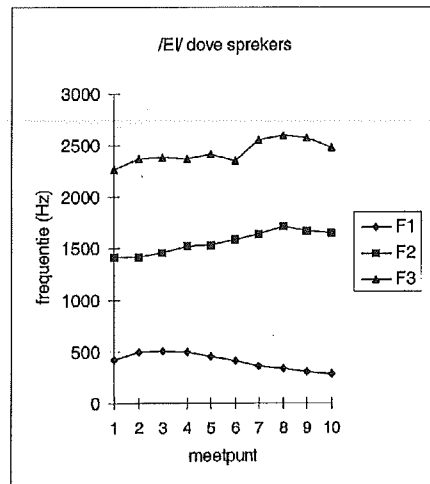
In de figuur is te zien dat verschillen in duur optreden als functie van de sprekers en de context. De verschillen tussen de groepen sprekers blijken uit de uitzonderlijke lengte van de diftongen van de groep logopedisten (MN), gevolgd door de diftongen van de dove sprekers en die van de ongetrainde sprekers (MN). Dit patroon geldt voor alle contexten, met een algemene afname in duur van de open context naar de context zinnen. Een variantie-analyse bevestigt deze bevindingen. In de open context werd het effect gemeten van de factoren Spreker (3) en Diftong (3). De factor Sprekers had een significant effect, $F(2,8)=15,95$ ($p < 0,01$ tenzij anders vermeld). De verschillen in de gesloten context werden getest met behulp van de factoren Spreker (3), Diftong (3) en Woord (4). Er was een hoofdeffect voor Spreker en Woord (resp. $F(2,35)=90,82$ en $F(3,35)=7,16$). De verschillen in de zinnencontext zijn kleiner maar niettemin is er een significant effect voor Spreker (3), $F(2,44)=33,40$ en Woord (5), $F(4,44)=12,73$. Een posthoc Newman-Keuls-test met de factor Spreker bevestigt de volgorde van de duurvverschillen tussen de sprekergroepen. Een posthoc test met factor Woord laat zien dat in de gesloten woordencontext de diftongen uit CVCVC woorden het kortst zijn (dit klopt met de verwachtingen volgens Nooteboom, 1972). In de zinnencontext zijn de verschillen tussen de gesloten woorden niet meer significant en is alleen de diftong uit het CV-woord significant langer.

Formanten

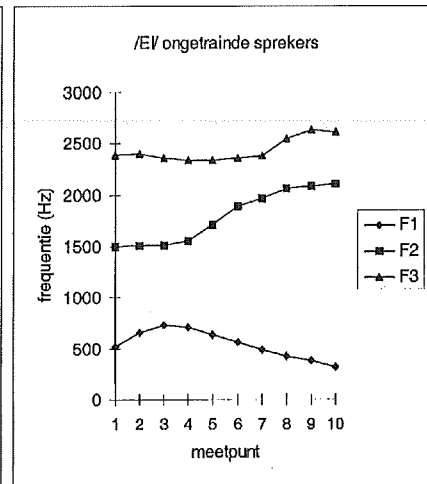
De resultaten van de formantenanalyse zijn weergegeven in Figuur 2. In de figuren staan de patronen van de eerste drie formanten van de [ɛi] (gesloten context) van de dove sprekers, de ongetrainde sprekers (MN) en de logopedisten (MN). De patronen voor de [au] en [ʌy] zijn niet weergegeven, omdat de verschillen tussen de sprekers vergelijkbaar zijn met die van de [ɛi]. De formantsporen van de diftongen van de horende sprekers hadden in de twee spreekcondities, 'normaal' en 'met nadruk', hetzelfde patroon.

De figuren 2a, 2b, 2c laten een duidelijk beeld zien van het verloop van de formanten van de diftongen van de drie groepen sprekers. Wanneer de formanten van de dove en horende sprekers vergeleken worden, valt op dat de eerste formant van de diftongen van dove sprekers eenzelfde verglijding vertoont als die van de horende sprekers. Het verloop van de tweede formant is echter verschillend. De tweede formant van de logopedisten en ongetrainde sprekers bevat een transitie van een beginpositie naar een eindpositie. Deze snelle formantbeweging ontbreekt echter in het F2-patroon van de diftongen van de dove sprekers. Naast verschillen tussen de formantpatronen van de diftongen van de dove en horende sprekers zijn er verschillen tussen de formanten van de logopedisten en ongetrainde sprekers. De F1-frequenties van de diftongen van de logopedisten liggen lager, en de F2-frequenties liggen hoger dan die van de ongetrainde sprekers.

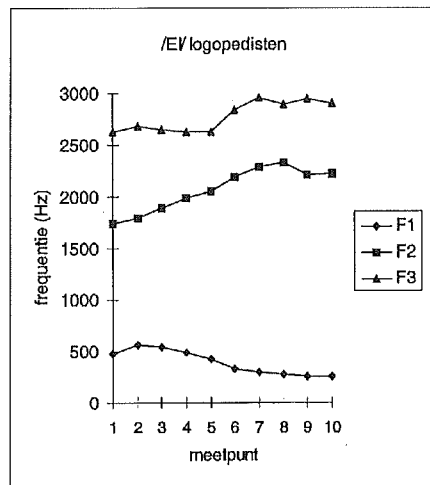
De verschillen tussen de eerste, tweede en derde formantpatronen van de diftongen van de verschillende sprekergroepen (context gesloten woorden) werd getest met behulp van een variantie-analyse met de faktor 'Spreker' (3). Deze faktor heeft een signi-



Figuur 2a Formantsporen [ei] dove sprekers



Figuur 2b Formantsporen [ei] ongetrainde sprekers



Figuur 2c Formantsporen [ei] logopedisten

ficant ($p < 0,01$) effect op de formantpatronen van alle diftongen. Een variantie-analyse op de frequenties binnen een formantspoor met als faktor 'Meetpunt' (10) laat zien dat de transities in de eerste en tweede formant significant zijn. De derde formant bestaat bij alle sprekers uit een steady-state: er zijn geen significante verschillen tussen de tien formantfrequenties binnen deze formant.

Discussie

De duur van de diftongen van de dove sprekers blijkt langer te zijn dan die van onge-

trainde horende sprekers. Opmerkelijk is de lange duur van de diftongen van de logopedisten. In de gesloten woorden context zijn de diftongen van de logopedisten (MN) langer dan de diftongen van de dove sprekers. Daarnaast is de duur van de diftongen van de logopedisten zelfs in de conditie 'normaal spreken' aanmerkelijk langer dan die van de ongetrainde sprekers in deze spreekconditie. Een verklaring hiervoor zou kunnen zijn dat de manier van spreken tijdens de spreeklessen het normale spreektempo van de logopedisten heeft vertraagd. De logopedisten die aan dit onderzoek hebben meegewerkt waren reeds geruime tijd werkzaam op het Instituut voor Doven.

Naast de verschillen in duur zijn er ook verschillen tussen de formantpatronen. Het verloop van de F1 van de dove sprekers is vergelijkbaar met dat van de horende sprekers, maar de snelle stijging in het verloop van de tweede formant ontbreekt in de diftongen van de dove sprekers. Deze resultaten met Nederlandse diftongen zijn in overeenstemming met de resultaten van Engelse diftongen (Monsen, 1978). Volgens Monsen (1978) en Tye-Murray (1991) hebben dove sprekers moeite met het produceren van voldoende variatie van de tweede formant voor het onderscheid van verschillende vocalen en is het nog moeilijker om een intrinsieke vergelijding te maken. Een kenmerk van diftongen is juist een snelle verandering van de frequenties van de eerste en tweede formant. De productie van een vergelijding van de tweede formant is extra moeilijk, omdat de tongbewegingen die kenmerkend zijn voor deze formant vrijwel onzichtbaar zijn en daardoor moeilijk te imiteren door dove sprekers. Een afwijking van de tweede formant is belangrijk, omdat de variabiliteit van deze formant een hogere correlatie met de verstaanbaarheid van dovenspraak heeft dan de variabiliteit van de eerste formant (Monsen, 1978). De beweging van de eerste formant is meestal wel aanwezig, omdat de eerste formant (lage frequenties) beter hoorbaar is voor de dove luisteraar (Monsen, 1978). De gehoorverliezen van dove sprekers zijn vaak bij hoge frequenties ernstiger dan bij lage frequenties (Monsen, 1978; Gulian, 1983; Tye-Murray & Iler Kirk, 1993). Daarnaast wordt de eerste formant voor een gedeelte bepaald door de grootte van de mondopening die duidelijk zichtbaar is en makkelijk te imiteren door dove sprekers. De resultaten doen vermoeden dat aandacht voor de tongbeweging (met name de voor- en achterwaartse beweging) tijdens logopedische therapie de uitspraak van de diftongen door dove sprekers zou kunnen verbeteren.

Het belangrijkste verschil tussen de formanten van de horende sprekers is de verschillende begin- en eindfrequenties van de F1 en F2. De formanten van de logopedisten (MN) lijken extremere waarden te bevatten waardoor de afstand tussen de F1 en F2 bij deze diftongen groter is dan tussen de formanten van de ongetrainde sprekers (MN). De grotere afstand tussen de formanten van de logopedisten (MN) kan mogelijk verklaard worden door het feit dat deze sprekers ook grotere articulatiebewegingen maakten tijdens het produceren van de diftongen in vergelijking met de articulatiebewegingen van de ongetrainde sprekers (MN). Dit is in dit onderzoek echter niet nader onderzocht. De verschillende begin- en eindfrequenties van de diftongen van de logopedisten en ongetrainde sprekers leiden niet tot verschillen in diftongidentiteit. Dit bevestigt het idee dat het totale temporele patroon van de formanten bepalend is voor de diftongidentiteit en niet de exacte begin- en eindfrequenties van de formanten.

Deel 2: Perceptieve beoordeling

Zoals gezegd zijn de diftongen van de dove sprekers langer dan die van de ongetrainde sprekers. De gemiddelde lengte van de diftongen van de ongetrainde sprekers in de normale spreekconditie was in de context gesloten woorden 162 ms. De diftongen uit deze context van de dove sprekers hadden een gemiddelde lengte van 274 ms. Dit is aanmerkelijk langer. Wordt de diftong onverstaanbaar door deze verlenging? In Monsen (1978) werd de invloed van de variabele spreektempo op de verstaanbaarheid van dove sprekers bepaald. Er bleek geen duidelijke correlatie te zijn tussen spreektempo en verstaanbaarheid. Enkele goed verstaanbare dove sprekers hadden een laag spreektempo en enkele slecht verstaanbare dove sprekers spraken vrijwel net zo snel als horende sprekers. Deze resultaten hebben echter betrekking op de invloed van het spreektempo op de totale verstaanbaarheid van dovenspraak. Omdat temporele aspecten juist bij de productie van de diftongen erg belangrijk zijn (Peeters, 1991), kan niet met zekerheid gesteld worden dat de conclusies van Monsen ook voor deze klanken gelden. Daarom is in het tweede gedeelte van dit onderzoek het effect van de variërende duur en formanten van de diftongen van de dove sprekers op de perceptie bestudeerd.

In onderzoek naar de verstaanbaarheid van dovenspraak wordt meestal gebruik gemaakt van een luisterpanel dat een cijfer geeft voor de verstaanbaarheid of een transcriptie maakt van wat er gezegd is (Monsen, 1974; Monsen, 1978; Markides, 1970; McGarr, 1983; Mencke & Ochsner, 1983; Mencke, Ochsner & Testut, 1985; Suonpää & Aaltonen, 1981). Er zijn echter een aantal nadelen aan deze methode verbonden. Ten eerste geeft een verstaanbaarheidspercentage van dovenspraak wel een indicatie van de hoeveelheid verstaanbare spraak, maar wordt niet duidelijk waarom de spraak (on)verstaanbaar is (Monsen, 1978). Een tweede nadeel van het gebruik van verstaanbaarheidsscores is dat deze geen consistent beeld opleveren. In de literatuur lopen de verstaanbaarheidsscores van dovenspraak uiteen van 19% tot 76% (Markides, 1970; Smith, 1975; Monsen, 1974). Een oorzaak van deze verschillen ligt in de aard van het stimulusmateriaal dat werd geanalyseerd. Monsen (1974) gebruikte monosyllabische woorden in korte zinnen, terwijl het stimulusmateriaal van Smith (1975) bestond uit lange zinnen met een relatief complexe grammaticale structuur. Naast de verschillen in het stimulusmateriaal spelen factoren als leeftijd van de spreker, mate en type gehoorverlies, opnamekwaliteit en samenstelling van het luisterpanel een rol. Een derde nadeel van het gebruik van verstaanbaarheidsscores is dat de scores niets zeggen over de acceptabiliteit van de spraak. Dit is van belang aangezien spraak verstaanbaar kan zijn, maar tegelijkertijd niet acceptabel. De omgekeerde relatie bestaat niet: onverstaanbare spraak zal nooit als acceptabel beoordeeld worden.

Vanwege de nadelen die kleven aan het gebruik van verstaanbaarheidsscores worden in de hier beschreven luisterexperimenten acceptabiliteitsoordelen toegepast. Om te achterhalen of luisteraars een voorkeur hebben voor diftongen uitgesproken door dove of horende sprekers, werd aan hen gevraagd van paarsgewijs aangeboden diftongen de meest acceptabele diftong aan te geven. Een subvraag bij deze experimenten was of er een verschillend oordeel zou zijn afhankelijk van de ervarenheid van de

luisteraars. In onderzoek naar de verstaanbaarheid van dovenspraak wordt regelmatig opgemerkt dat de verstaanbaarheidspercentages sterk afhankelijk zijn van de mate van ervaring die de luisteraar heeft met dovenspraak. De getrainde luisteraar verstaat significant meer dan de ongetrainde luisteraar (Mencke & Ochsner, 1983; Mencke, Ochsner & Testut, 1985). Dit bleek onder meer uit het onderzoek van Markides (1970) waar de getrainde luisteraars 31% van de woorden begrepen die dove sprekers uitgesproken hadden versus de 19% die door de ongetrainde luisteraars verstaan waren. Volgens Monsen (1978) worden de verschillen tussen de verstaanbaarheidsscores van getrainde en ongetrainde luisteraars echter kleiner naarmate de ongetrainde luisteraars meer gewend raken aan dovenspraak. In zijn onderzoek daalde het verschil van 14% naar 5% na het derde experiment.

McGarr (1983) veronderstelt dat er twee factoren in het voordeel van de getrainde luisteraars zijn: het beter gebruik maken van informatie in de context en de familiariteit met het decoderen van dovenspraak. Uit een toetsing van deze hypothese bleek echter dat verschillen tussen de getrainde en ongetrainde luisteraars niet verklaard konden worden door bovenstaande factoren. Er was geen interactie tussen mate van ervarenheid van de luisteraar en de voorspelde woordverstaanbaarheid. McGarr concludeerde dat de verschillen tussen de resultaten van ongetrainde en getrainde luisteraars blijikbaar veroorzaakt worden door een complex aantal factoren, waaronder bekendheid met de spreker en het optimaal gebruik maken van visuele informatie. Deze conclusie wordt bevestigd door de bevindingen van Mencke, Ochsner & Testut (1985). Zij vermelden dat de getrainde luisteraars het slechts in één stimulusconditie beter deden. In deze conditie werd het materiaal zowel auditief als visueel aangeboden. Op basis van de resultaten van Mencke & Ochsner (1983), McGarr (1983) en Monsen (1978) wordt verwacht dat er in dit onderzoek verschillen zullen optreden in de acceptabiliteitsoordelen van getrainde en ongetrainde luisteraars. Er zijn echter verschillen in het stimulusmateriaal die deze verwachting afzwakken. In de beschreven literatuur bestond het stimulusmateriaal uit woorden en/of zinnen, terwijl nu geïsoleerde diftongen worden aangeboden. Volgens Maassen (1985) is het, zowel voor getrainde als voor ongetrainde luisteraars, niet mogelijk aan te geven of ze dovenspraak of spraak van horenden waarnemen, wanneer de geluidsstimuli bestaan uit geïsoleerde vocalen. Daarvoor zijn langere stukken spraak nodig. Een tweede verschil met het stimulusmateriaal uit eerder onderzoek is dat nu geen gebruik wordt gemaakt van natuurlijke uitingen, maar van gesynthetiseerde spraak. De synthetische diftongen werden gegenereerd met behulp van de duur- en formantwaarden van de originele diftongproducties. Uit onderzoek is gebleken dat ongetrainde luisteraars moeiteloos synthetische diftongen kunnen herkennen (Bond, 1982; Gay, 1968; Peeters, 1991). In een van de luisterexperimenten van Peeters (1991) werden synthetische diftongen beoordeeld op hun acceptabiliteit door ongetrainde luisteraars. De luisteraars bleken in staat voorkeurscores te geven aan deze stimuli, waarvan onder meer de duur en de formanten waren gemanipuleerd (Peeters, 1991). Op basis van deze resultaten wordt verwacht dat, ook in het huidige onderzoek, luisteraars in staat zullen zijn een oordeel te vellen over de acceptabiliteit van in isolatie aangeboden, synthetische diftongen.

Onderzoeksvragen

Om te achterhalen hoe luisteraars diftongen beoordelen die uitgesproken zijn door dove sprekers, logopedisten en ongetrainde sprekers werden twee luisterexperimenten uitgevoerd met als onderzoeksvragen:

- 1 Hebben luisteraars een voorkeur voor diftongen van dove sprekers, logopedisten of ongetrainde sprekers?
- 2 Wordt het oordeel van luisteraars beïnvloed door variatie van de duur of van de formanten in de diftongen?
- 3 Zijn er significante verschillen tussen de beoordelingen van ongetrainde of getrainde luisteraars?

Methode

Stimulusmateriaal

In plaats van de natuurlijke uitingen werd gebruik gemaakt van gesynthetiseerde spraak. Dit had als voordeel dat de onderzoeksvariabelen controleerbaar waren. De spraaksynthese werd gegenereerd met behulp van de duur en formantfrequenties van de originele diftongproducties. Op deze manier bleven de sprekerspecifieke kenmerken in de diftongen bewaard. Als originele diftongen werden de opnamen gebruikt van de diftongen in /pauw/ en /lui/, geproduceerd door de drie groepen sprekers in het productie-experiment. De diftong [ei] die ook in het productie-experiment vermeld wordt, werd niet voor de perceptie-experimenten gebruikt. Tijdens een proefsynthese werd namelijk duidelijk dat de synthetische diftong [ei] op het gehoor vaak klonk als de klank [ʌy], hoewel de ingevoerde formantfrequenties van de oorspronkelijke [ei] afkomstig waren. Deze overlapping tussen de diftong [ei] en [ʌy] werd al eerder opgemerkt door Pols (1977). Het zal duidelijk zijn dat van een luisteraar niet gevraagd kan worden een oordeel te geven over de [ei] wanneer een [ʌy] waargenomen wordt.

Synthese

De synthese van de diftongen gebeurde met de difoon-regelsynthesator van de Vakgroep Taal en Spraak van de Universiteit Nijmegen. Deze synthese bevat een veelvoud van parameters die aangepast kunnen worden. De diftong werd gerepresenteerd als twee segmenten. Het eerste segment bestond uit waarden voor de vocaal die als uitgangspunt wordt genomen en het tweede segment is de doelvocaal. Het formantpatroon bestond uit drie componenten: steady-state/transitie/steady-state. Tijdens de synthese werden twee akoestische parameters gevarieerd: duur en formantfrequentie. Hiervoor werden de sprekerspecifieke waarden ingevoerd die waren geanalyseerd in het productie-experiment. Er werden drie verschillende typen stimuli gegenereerd:

1. diftongen met sprekerspecifieke duur;
2. diftongen met sprekerspecifieke formantfrequenties;
3. diftongen met sprekerspecifieke duur en sprekerspecifieke formantfrequenties.

Voor elke diftong ([au] en [ʌy]) werden 60 verschillende stimuli gevormd: 5 (groepen sprekers) x 4 (sprekers per groep) x 3 (typen synthese).

Diftongen met sprekersspecifieke duur. Voor de analyse van de duur van de componenten van elke diftong werd gebruik gemaakt van de figuren van de formantpatronen. De formantpatronen van de eerste twee formanten werden met de hand gestileerd om het steady-state/transitie/steady-state verloop te benadrukken. Hierna werd de formanten in twee stukken verdeeld: de twee segmenten. Van elke diftong was reeds de totale duur bepaald. Nu werd er bekeken hoe lang de twee segmenten duurden en wat de lengte was van de steady-state/transitie/steady-state in de formantsporen.

Voor de constante formantfrequenties en de bandbreedtes van de eerste twee formanten werden de gegevens gebruikt uit de experimenten van Peeters (1991). Deze weken in geringe mate af van de standaardwaarden van de synthetisator, maar gaven een natuurlijker resultaat. In de standaardsynthese was voor diftongen een monotone intonatie ingesteld op 110 Hz. Deze werd gewijzigd, omdat geïsoleerde klanken met een intonatiepatroon natuurlijker klinken (Peeters, 1991). De standaard intonatiecontour voor de [ʌy] werd vervangen door een contour die begon op 130 Hz en eindigde op 95 Hz. De contour voor de [au] begon op 135 Hz en eindigde op 110 Hz (Peeters, 1991). Voor alle overige parameters (onder meer grondfrequentie en amplitude) werden de waarden van het synthese programma gebruikt.

Diftongen met sprekersspecifieke formantfrequenties. Voor de sprekersspecifieke formantwaarden werd van elke steady-state van de eerste drie formanten de gemiddelde formantfrequentie bepaald. Het syntheseprogramma bepaalde de frequenties van de transitie door lineaire interpolatie. De constante waarde voor de duur werd bepaald door te middelen over de duur van de diftongen van de ongetrainde sprekers uit de context open woorden. Dit is in tegenstelling met wat misschien verwacht zou worden, namelijk dat de duur bepaald zou worden door de lengte van de diftongen van de getrainde sprekers, de logopedisten. Echter, tijdens de analyse van de duur van de diftongen van de verschillende sprekers bleek dat de diftongen van de logopedisten opvallend lang waren. Daarom werd als uitgangspunt voor de duur van het luistermateriaal de spraak van de ongetrainde sprekers genomen. De totale duur van zowel de [au] als de [ʌy] werd gesteld op 235 ms. De standaardwaarde voor de duur in het syntheseprogramma was 130 ms. De synthetisator gaat echter uit van een duur voor de klanken in lopende spraak. Omdat in dit onderzoek geïsoleerde klanken beoordeeld werden, is gekozen voor een langere duur. Voor de overige parameters werden de waarden gebruikt als beschreven bij het vorige stimulustype.

Diftongen met sprekersspecifieke duur plus formantfrequenties. Deze synthetische diftong is een combinatie van de twee genoemde syntheses. De sprekersspecifieke data voor de duur en formantfrequenties waren variabel. De belangrijkste constante parameters waren de grondfrequentie, bandbreedte, intonatiecontour en de amplitude.

Design experiment I

Als eerste wordt het experiment beschreven waarin werd onderzocht of luisteraars een voorkeur hebben voor diftongen van dove sprekers, ongetrainde sprekers, of logope-

disten. Hierna volgt een weergave van experiment II waarin de tweede onderzoeksvraag werd beantwoord over de invloed van de variatie van de duur of van de formanten in de diftongen op de luisteraarsoordelen. De derde onderzoeksvraag, over de verschillen tussen getrainde en ongetrainde luisteraars werd in beide experimenten onderzocht.

Met behulp van een computerprogramma van Van Hessen (persoonlijke communicatie, april 1994) werd het luisterexperiment opgezet en uitgevoerd. Met behulp van dit programma werden de diftongen van hetzelfde type (dus type 1, 2 of 3) at random aan elkaar gekoppeld, met de restrictie dat er geen paren mochten voorkomen van diftongen van sprekers uit dezelfde sprekergroep. Een stimuluspaar met diftongen uitgesproken door dove sprekers en ongetrainde sprekers was dus wel geoorloofd, maar een combinatie van diftongen van twee dove sprekers niet. Alle paren werden in twee verschillende volgorden aangeboden: volgorde (AB) en (BA). Daarnaast werd ieder paar één keer herhaald. In totaal kregen de luisteraars in elk experiment 640 paren te beoordelen.

Luisteraars experiment I

Aan dit experiment hebben 10 luisteraars meegewerkt: een vijftal getrainde (vijf vrouwen) en vijf ongetrainde luisteraars (vier vrouwen en een man). De getrainde luisteraars waren studenten Spraak & Taalpathologie met een diploma logopedie. Als onderdeel van de studie Spraak & Taalpathologie hadden zij trainingen gevolgd in het analytisch luisteren naar normale en pathologische spraak (waaronder dovenspraak). De groep ongetrainde luisteraars bestond ook uit studenten van de faculteit Letteren van de Universiteit Nijmegen. Zij waren echter niet bekend met dovenspraak en ongetraind in het analytisch luisteren naar spraakmateriaal.

Procedure experiment I

De luisteraars deden mee aan zes luisterexperimenten: drie (duur, frequentie, duur plus frequentie) voor de synthetische [au] en drie (duur, frequentie en duur plus frequentie) voor de synthetische [ay]. De [au]- en [ay]-experimenten werden afwisselend aangeboden. De samenstelling van de paren en de volgorde van de paren varieerde in elk experiment voor elke luisteraar. De responsietijd voor de luisteraar was 2 seconden. Het interstimulusinterval bedroeg 300 ms. De stimulusparen werden aangeboden via een hoofdtelefoon. Na elk stimuluspaar werd van de luisteraar gevraagd een voorkeur aan te geven voor één van de twee diftongen. De scoring gebeurde door het aanklikken van de linker- of rechtermuisknop (de linkermuisknop werd gebruikt, wanneer de eerste diftong het beste werd gevonden en de rechtermuisknop, wanneer de tweede diftong als beste werd beoordeeld). Voorafgaande aan het experiment kreeg de luisteraar een instructie te lezen. Deze werd mondeling toegelicht, waarbij werd benadrukt dat elk stimuluspaar beoordeeld moest worden. Na een korte uitleg over de scoringswijze volgde zes oefenparen. Deze oefenparen stelde de luisteraar in staat te wennen aan het hanteren van de muis en aan de 'computerachtige' kwaliteit van het spraakmateriaal.

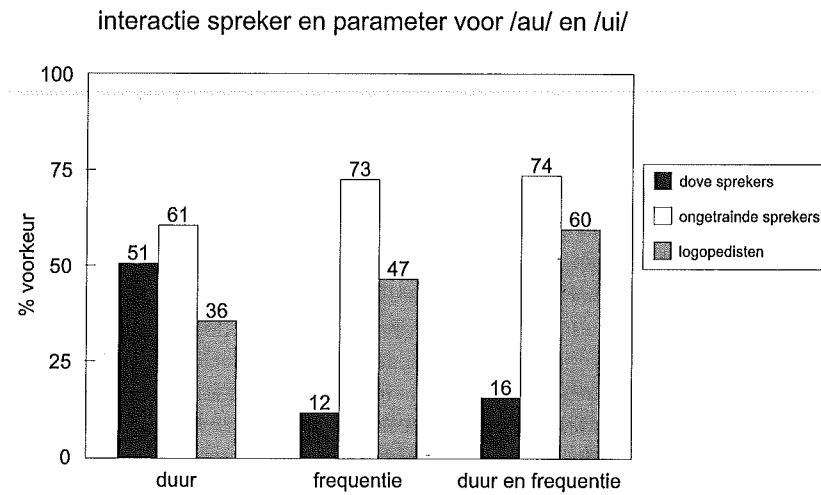
De experimenten werden uitgevoerd in een geluidsarme luistercabine van de Ka-

tholieke Universiteit Nijmegen. Elk experiment duurde een uur. Iedere luisteraar voerde de zes experimenten uit binnen een tijdsbestek van 10 dagen. Het maximum aantal experimenten per luisteraar was drie per dag. Wanneer er door een luisteraar meerdere experimenten op een dag werden verricht, zat er minimaal één uur pauze tussen de experimenten.

Resultaten experiment I

De resultaten van experiment I zijn weergegeven in Figuur 3. In deze figuur staat de percentages waarmee de luisteraars hun voorkeur hebben aangegeven voor een stimulus. De percentages zijn berekend door de werkelijke voorkeur die de luisteraars hadden voor de diftongen van een bepaalde groep sprekers te delen door de maximaal te behalen score (256 per sprekergroep). De percentages zijn verdeeld over drie typen: duur, formantfrequentie en de combinatie duur+formantfrequentie. Er was geen verschil tussen de luisteraarsoordelen als een functie van de diftong ([au] en [ʌy]) en daarom zijn deze scores samen genomen en gemiddeld. Omdat er geen verschil was tussen de voorkeurscores voor ongetrainde sprekers vs. ongetrainde sprekers (MN) en logopedisten vs. logopedisten (MN) zijn deze scores samengevoegd.

In Figuur 3 is te zien dat de luisteraars een voorkeur hebben voor de synthetische diftongen met de sprekersspecifieke gegevens van de ongetrainde sprekers. Deze diftongen klinken blijkbaar het beste, gevolgd door de diftongen van de logopedisten. De diftongen met de waarden van de dove sprekers klinken het slechtst. Wanneer gekeken wordt naar de percentages voor de drie typen stimuli, blijkt dat de voorkeur van de luisteraars variabel is onder invloed van de parameter die de sprekersspecifieke waarden bevat. Bij de diftongen met sprekersspecifieke waarden van de dove sprekers vinden de luisteraars de diftongen waarin de duur is aangepast het meest acceptabel, terwijl de diftongen met formantfrequenties van dove sprekers heel laag scoren. Opmerkelijk is dat de stimuli met de duur van de diftongen van de dove en ongetrainde sprekers gelijk scoren, terwijl die van de logopedisten minder acceptabel zijn. Uit Figuur 3 komt ook naar voren dat de verdeling van de voorkeurscores over de sprekergruppen gelijk is voor de stimuli met manipulatie van de formantfrequentie alleen en de stimuli met de combinatie van de duur plus formantfrequentie. De luisteraars hebben hier weer de voorkeur voor de diftongen van de ongetrainde sprekers, gevolgd door die van de logopedisten en de diftongen van de dove sprekers. De resultaten worden bevestigd door een variantie-analyse (op de ruwe scores) met als onafhankelijke factoren Spreker (5) Parameter (3) en Luisteraar (2). De verschillen in voorkeur als een effect van de sprekers zijn significant ($p < 0,01$) ([au]: $F(4, 14) = 231,59$; [ʌy]: $F(4, 14) = 101,94$). De factor Parameter is niet significant als hoofdeffect. Er is een interactie-effect van Sprekers met Parameter, [au] $F(8, 14) = 55,52$; [ʌy] $F(8, 14) = 20,31$. Dat betekent dat de luisteraars zich, binnen stimuli van een sprekergroep, in verschillende mate laten beïnvloeden door de variatie van de duur- en de formantfrequenties of de combinatie van deze twee. De voorkeurscores van de ongetrainde en de getrainde luisteraars verschillen niet significant.



Figuur 3 Het percentage voorkeursoordelen van de luisteraars voor de diftongen (per stimulus type) van de dove sprekers, ongetrainde sprekers en logopedisten

Design experiment II

Experiment II werd uitgevoerd om een antwoord te krijgen op de vraag of de luisteraars op een verschillende manier werden beïnvloed door de manipulaties van duur of formantwaarden tijdens beoordeling van de diftongen. Het design van dit experiment is vrijwel identiek aan dat van experiment I. Hetzelfde stimulusmateriaal werd gebruikt, met als verschil dat in dit experiment alleen de diftongen voorkwamen die in de spreekconditie 'met nadruk' waren uitgesproken. De drie groepen sprekers worden dan: dove sprekers, ongetrainde sprekers (MN) en logopedisten (MN).

Met de sprekerspecifieke duren en formantfrequenties van de [au] en de [ʌy] van elke groep sprekers werden in totaal 36 synthetische diftongen gevormd (3 groepen sprekers x 4 sprekers x 3 typen stimuli). Er werden stimulusparen gevormd van diftongen van dezelfde sprekergroep met als restrictie dat er geen paren geoorloofd waren waarin dezelfde parameter gemanipuleerd werd. Een stimulus met de duurwaarden van diftongen van dove sprekers kon dus gekoppeld worden aan een stimulus met de formanten van de diftongen van dove sprekers, maar twee stimuli van hetzelfde type, bijvoorbeeld met duurmanipulaties, mochten niet gematched worden. Alle paren werden in twee verschillende volgorden aangeboden: volgorde (AB) en (BA). Daarnaast werd ieder paar één keer herhaald. In totaal kregen de luisteraars in elk experiment 192 paren te beoordelen.

Luisteraars experiment II

Aan het experiment hebben 5 getrainde en 5 ongetrainde luisteraars meegedaan (6 vrouwen en 4 mannen). De getrainde luisteraars waren medewerkers van de afdeling

Research, Development & Support van het Instituut voor Doven in St. Michielsgestel. Zij waren bekend met dovenspraak en getraind in het analytisch luisteren naar spraakmateriaal. De ongetrainde luisteraars waren drie stagiaires en twee administratief medewerkers van de afdeling Research, Development & Support van het Instituut voor Doven in St. Michielsgestel. Zij waren niet bekend met dovenspraak en ongetraind in het luisteren naar spraakmateriaal.

Procedure experiment II

De procedure voor experiment II was identiek aan die van experiment I, behalve dat dit experiment niet op de universiteit werd uitgevoerd, maar in de experimenteerruimte van de afdeling Research en Development van het Instituut voor Doven in St. Michielsgestel. Elke luisteraar werd twee keer uitgenodigd voor het uitvoeren van drie experimenten (een experiment duurde 17 minuten).

Resultaten experiment II

De resultaten van experiment II staan in Figuur 4. In de figuren staat het effect van de duur- en formantmanipulaties op de luisteraarsoordelen. De percentages zijn berekend door de werkelijke voorkeur die de luisteraars hadden voor de diftongen met een bepaalde manipulatie te delen door de maximaal te behalen score (128 per diftongtype). De percentages zijn verdeeld over drie groepen sprekers: dove sprekers, ongetrainde sprekers (MN) en logopedisten (MN). Er zijn geen verschillen in de beoordeling van de diftongen door de ongetrainde of de getrainde luisteraars.

De resultaten van het tweede experiment laten een verschil zien in het effect van de akoestische manipulaties op de beoordeling van de [Ay] (Figuur 4a) en de [au] (Figuur 4b). Bij de [Ay] is er een groot effect van de duur op de luisteraarsoordelen. Uit de resultaten van de beoordeling van de [Au] blijkt dat de luisteraars alle drie typen stimuli gelijk beoordelen. Dit wordt bevestigd door de variantie-analyse (op de ruwe scores) met als onafhankelijke factoren waren Parameter (3), Spreker (3) en Luisteraar (2). Alleen voor de [Ay] is er een significant ($p < 0,01$) hoofdeffect van Parameter, $F(2,8)=338,16$. Daarnaast is er, zowel bij de beoordeling van de [au] als de [Ay] een interactie tussen Parameter en Spreker ([au], $F(4,8)=53,81$; [Ay], $F(4,8)=16,07$).

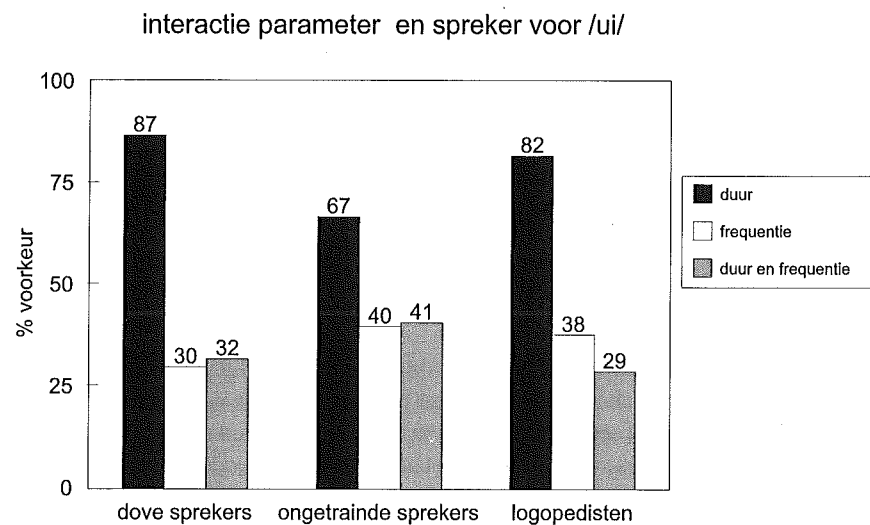
Discussie

Uit de resultaten van experiment I blijkt dat de luisteraars de diftongen van de dove sprekers het minst acceptabel vinden. Dit ligt niet aan de duur van deze diftongen, want de luisteraars hebben een gelijke voorkeur voor de diftongen met de duur van de dove sprekers en de ongetrainde sprekers. Het acceptabiliteitsoordeel verandert wanneer stimuli met sprekerspecifieke formanten worden aangeboden. De luisteraars hebben dan een duidelijke voorkeur voor de diftongen van de ongetrainde sprekers. De formanten van de diftongen van de dove sprekers worden als het minst acceptabel beschouwd. Dit betekent dat van de verschillen in duur en formantfrequenties die werden gevonden in het productie-experiment, vooral de afwijkende formantpatronen een

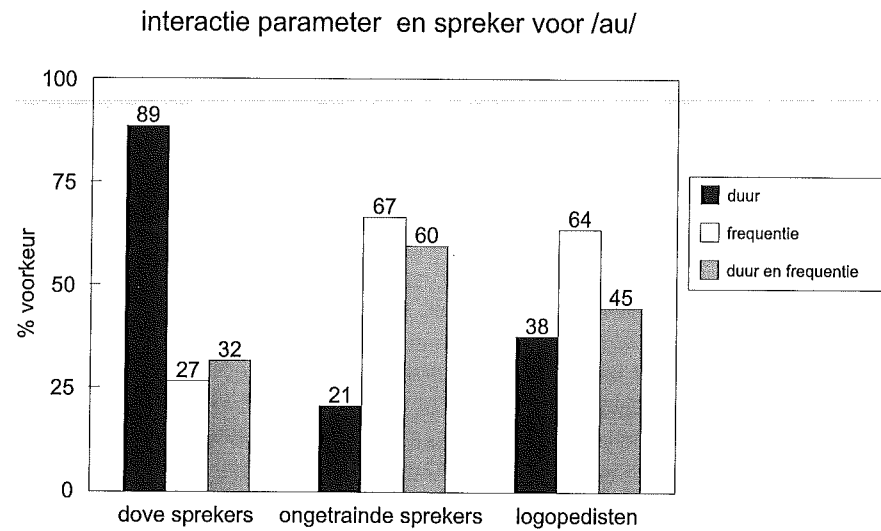
grote invloed hebben op de perceptie. Dit bevestigt het idee van Peeters (1991) dat het totale formantpatroon bepalend is voor de identiteit van de diftong. De vervorming van de diftongen in de uitspraak van dove sprekers zal de verstaanbaarheid van dovenspraak verslechteren. Hoe groot deze negatieve invloed is, is onduidelijk omdat diftongen minder vaak in spraak voorkomen dan consonanten en de overige vocalen. Volgens Smith (1975) echter heeft juist de onjuiste uitspraak diftongen en vocalen een grotere negatieve invloed op de verstaanbaarheid van dovenspraak dan die van de veelvuldig voorkomende consonanten.

Een opvallend resultaat is dat de duur en de formantwaarden van de diftongen van de logopedisten minder acceptabel worden gevonden dan die van de ongetrainde sprekers. De verwachting was dat de logopedisten, de getrainde sprekers, de meest acceptabele diftongen zouden produceren. Blijkbaar hebben zowel de extreme duur als de extreme formantwaarden van deze diftongen een negatief effect op de perceptie.

In het tweede luisterexperiment werd nagegaan of er een effect was van het type stimulus op de beoordeling van de luisteraars. Uit de resultaten blijkt dat bij de beoordeling van de [ʌy], de duur een cruciale rol speelt, terwijl deze parameter geen speciale betekenis heeft bij de beoordeling van de [au]. Dit verschil wordt waarschijnlijk veroorzaakt door de consonant die aan de diftongen vooraf ging. Bij de [au] was dit de [p] en bij de [ʌy] was dit de [l]. Deze consonanten waren niet gelijk als gevolg van de keuze om betekenisvolle woorden als stimulusmateriaal te gebruiken. De [p] en [l] consonanten hebben onvermijdelijk mede het eerste gedeelte van het formantspoor van de diftongen bepaald, waarbij verwacht kan worden dat de [l] een grotere verkleuring van de diftong tot gevolg heeft dan een stemloze plosief. Deze verkleuring kan geleid heb-



Figuur 4a Het effect van de stimulusvariabelen Parameter en Spreker op de beoordeling van de [au]



Figuur 4b Het effect van de stimulusvariabelen Parameter en Spreker op de beoordeling van de [ʌy]

ben tot minder acceptabele formantwaarden voor een geïsoleerde diftong, wat de lage voorkeurscores verklaart.

Uit de interactie tussen Parameter en Spreker bij de beoordeling van de [au] blijkt wederom dat bij de diftongen van de dove sprekers alleen de duur acceptabel is. Net als in het eerste luisterexperiment levert voor de logopedisten de formantfrequenties in combinatie met een standaard duur van 235 ms de beste scores op. Blijkbaar is de werkelijke duur van de diftongen van de logopedisten te lang, waardoor de voorkeurscores omlaag gaan voor de stimuli duur+formantfrequentie. De luisteraars beoordelen de stimuli met de formantfrequenties en de combinatie duur+formantfrequenties van de diftongen van de ongetrainde sprekers vrijwel gelijk. Dit werd verwacht aangezien de gemiddelde duur van de diftongen van de ongetrainde sprekers ook als standaardwaarde voor synthese werd gebruikt.

Uit de resultaten van beide luisterexperimenten is gebleken dat er geen verschillen optreden tussen de beoordeling van ongetrainde en getrainde luisteraars. Het ontbreken van een significant verschil in de beoordeling van de ongetrainde luisteraars en die van de getrainde luisteraars is in tegenspraak met de resultaten van ander onderzoek waar wel een significant verschil optrad (Markides, 1970; Monsen, 1978; McGarr, 1983; Mencke, 1985). Het ontbreken van een verschil in de huidige studie kan mogelijk verklaard worden aan de hand van het stimulusmateriaal: geïsoleerde klanken. Door het gebrek aan context in het luistermateriaal hebben de getrainde luisteraars geen profijt van de informatie in de context. McGarr (1983) noemt het beter gebruik maken van de informatie in de context een van de factoren die in het voordeel van getrainde luisteraars kunnen zijn. Volgens Maassen (1985) kunnen luisteraars geen goed

oordeel geven over korte geïsoleerde vocalen en zijn langere stukken spraak nodig. Dit wordt tegengesproken door de huidige resultaten. In dit onderzoek werd niet expliciet van de luisteraars gevraagd het luistermateriaal te identificeren als normale spraak of dovenspraak. Uit de beoordeling van de diftongen van de verschillende groepen sprekers blijkt echter dat de luisteraars een voorkeur aangeven voor de stimuli van één bepaalde groep. Dus, ondanks de korte duur van de stimuli, kunnen de luisteraars onderscheid maken tussen vocalen van de verschillende spreker-groepen.

Conclusie

Uit de resultaten van het productie-experiment kan geconcludeerd worden dat diftongen op verschillende manieren gerealiseerd worden door dove sprekers, ongetrainde sprekers en logopedisten. Uit duurmetingen blijkt dat de diftongen van de logopedisten en de dove sprekers significant langer zijn dan de diftongen van de ongetrainde sprekers. De analyse van de formanten van de diftongen laat zien dat de snelle stijging of daling in de tweede formant van de diftongen van de horende sprekers ontbreekt in die van de dove sprekers. De gevolgen van de spectro-temporele afwijkingen worden duidelijk in de perceptieve beoordeling van deze diftongen. Getrainde en ongetrainde luisteraars beoordelen de diftongen van de dove sprekers als het minst acceptabel, wat vooral te wijten is aan de afwijkende formantpatronen van deze diftongen. In tegenstelling tot wat werd verwacht zijn de hoogste voorkeurscores niet voor de diftongen van de logopedisten, maar voor die van de ongetrainde sprekers.

Referenties

- Angelocci, A. Kopp, G. & Holbrook, A. (1964). The vowel formants of deaf and normal hearing eleven-to-fourteen old boys. *Journal of the Acoustical Society of America*, 29, 156-170.
- Bladon, A. (1985). Diphthongs: A case study of dynamic auditory processing. *Speech Communication*, 4, 145-154.
- Bond, Z.S. (1982). Experiments with synthetic diphthongs. *Journal of Phonetics*, 10, 259-264.
- Cao, J. (1991). Temporal structure of diphthongal finals in standard chinese. In: *Report of phonetic research. Institute of Linguistics, Chinese Academy of Social Sciences*, 39-54.
- Clermont, F. (1993). A Spectro-temporal description of diphthongs in F_1 - F_2 - F_3 space. *Speech Communication*, 13, 377-390.
- Collier, R., Bell-Berti, F. & Raphael, L.J. (1982). Some acoustic and physiological observations on diphthongs. *Language and Speech*, 25, 305-323.
- Dolan, W.B. & Mimori, Y. (1986). Rate-dependent variability in English and Japanese complex vowel F2 transitions. *UCLA Working Papers in Phonetics*, 63, 125-153.
- Ferguson, G.A. & Takane, Y. (1989). *Statistical analysis in psychology and education*. McGraw-Hill Book Company, New York.
- McGarr, N.S. (1981). The effect of context on the intelligibility of hearing and deaf children's speech. *Language and Speech*, 24 (3), 255-264.
- McGarr, N.S. (1983). The intelligibility of deaf speech to experienced and inexperienced liste-

- ners. *Journal of Speech and Hearing Research*, 26, 451-458.
- McGarr, N.S. & Gelfer, C.E. (1983). Simultaneous measurements of vowels produced by a hearing-impaired speaker. *Language and Speech*, 26 (3), 233-246.
- Gay, T. (1968). Effect of speaking rate on diphthong formant movements. *Journal of the Acoustical Society of America*, 44 (6), 1570-1573.
- Gay, T. (1970). A perceptual study of American English diphthongs, *Language and Speech*, 13, 65-88.
- Geffner, D. (1980). Feature characteristics of spontaneous speech production in young deaf children. *Journal of Communication Disorders*, 13, 443-454.
- Gerber, S.E. (1971). Perception of segmental diphthongs. Proceedings of the 7th International Congress on Phonetic Sciences, Montreal, 479-492.
- Gottfried, M., Miller, J.D. & Meyer, D.J. (1993). Three approaches to the classification of American-English diphthongs. *Journal of Phonetics*, 21, 205-229.
- Gulian, E., Hinds, P., Fallside, F. & Brooks, S. (1983). Vowel learning and the vowel system of deaf children: Age and feedback related differences. *Journal of Communication Disorders*, 16, 449-469.
- Hessen Van, H. (1994). *Persoonlijke communicatie*. Onderzoeksinstituut voor Taal & Spraak, Utrecht.
- Kerckhoff, J. & Boves, L. (1993). Designing control rules for a serial pole-zero vocal tract model. Eurospeech '93. Proceedings of the 3rd European Conference on Speech, Communication and Technology, Berlin, 1705-1708.
- Kirk, R.E. (1982). *Experimental Designs: Procedures for the Behavioral Sciences*. Belmont: Brooks/Cole, 249-281.
- Lehiste, I. & Peterson, G.E. (1961). Transitions, glides and diphthongs. *Journal of the Acoustical Society of America*, 33, 268-277.
- Levitt, H., Stromberg, C., Smith, C. & Gold, T. (1980). The structure of segmental errors in the speech of deaf children. *Journal of Communication Disorders*, 13, 419-441.
- Lindau, M., Norlin, K. & Svantesson, J-O. (1990). Some cross-linguistic differences in diphthongs. *Journal of the International Phonetic Association*, 20 (1), 10-14.
- Maassen, B. (1985). *Artificial corrections to deaf speech*. Doctoral Diss. Enschede: Sneldruk.
- Markides, A. (1970). The speech of deaf and partially hearing impaired children with special reference to factors affecting intelligibility. *British Journal of Disorders of Communication*, 5, 126-140.
- Mencke, E.O. & Ochsner, G.J. (1983). Listener judges and the speech intelligibility of deaf children. *Journal of Communication Disorders*, 16, 175-180.
- Mencke E.O., Ochsner, G.J. & Testut, E.W. (1985). Distinctive-feature analyses of the speech of deaf children. *The Journal of Auditory Research*, 25, 191-200.
- Monsen, R.B. (1974). Dural aspects of vowel production in the speech of deaf children. *Journal of Speech and Hearing Research*, 17, 386-398.
- Monsen, R.B. (1978). Toward measuring how well hearing-impaired children speak. *Journal of Speech and Hearing Research*, 21, 197-219.
- Nabelek, A.N., Czyzewski, Z. & Crowley, H. (1994). Cues for perception of the diphthong /aɪ/ in either noise or reverberation. Part I. Duration of the transition. *Journal of the Acoustical Society of America*, 95(5), 2681-2693.
- Nearey, T.M. & Assmann, P.F. (1986). Modelling the role of inherent spectral change in vowel identification. *Journal of the Acoustical Society of America*, 80 II, 1297-1308.
- Nooteboom, S.G. (1972). *Production and perception of vowel duration*. Doctoral Dissertation. University of Utrecht, the Netherlands.

- Peeters, W.J.M. (1991). Diphthong Dynamics: A cross-linguistic perceptual analysis of temporal patterns in Dutch, English and German. Doctoral Dissertation. University of Utrecht, the Netherlands, Kampen: Mondiss.
- Pols, L.C.W. (1977) Spectral analysis and identification of Dutch vowels in monosyllabic words. Doctoral Dissertation. University of Amsterdam, the Netherlands, Amsterdam: Academische Pers B.V.
- Smith, C.R. (1975). Residual hearing and speech production in deaf children. *Journal of Speech and Hearing Research*, 18, 795-811.
- Stollwerck, L.E. (1991). A contrastive study of the dynamic nature of the diphthongs /au/ and /ai/ in the production of English and German native speakers. MA thesis. University College London, dept. of Phonetics and Linguistics.
- Suonpää, J. & Aaltonen, O. (1981). Intelligibility of vowels in words uttered by profoundly hearing-impaired children. *Journal of Phonetics*, 9, 445-450
- Toledo, G.A. (1987) The influence of speaking rate in Spanish diphthongs. Proceedings of the 11th International Congress of Phonetic Sciences, Tallin, 3, 125-128.
- Tye-Murray, N. (1991). The establishment of open articulatory postures by deaf and hearing talkers. *Journal of Speech and Hearing Research*, 34, 485-459.
- Tye-Murray, N. & Iler Kirk, K. (1993). Vowel and diphthong production by young users of cochlear implants and the relationship between Phonetic Level Evaluation and spontaneous speech. *Journal of Speech and Hearing Research*, 36, 488-502.
- Yang, S. (1987). An articulatory dynamic model for diphthongs in Chinese. Proceedings of the 11th International Congress of Phonetic Sciences, Tallin, 1, 239-242.

APPENDIX: Stimulusmateriaal productie-experiment*Woorden*

[pauw]	[lɔy]	[pɛi]
[kaut]	[lɔyt]	[pɛip]
[pauwən]	[lɔyjə]	[pɛijən]
[kautən]	[lɔytən]	[pɛipən]
[kaudə]	[lɔydə]	[lɛidə]

Woorden in draagzinnen

De mooie [pauw]loopt trots voorbij.

De soep was [kaut] toen we gingen eten.

De staartveren van [pauwən] worden vaak als versiering gebruikt.

Het woord [kautən] betekent gezellig praten.

Er stond een [kaudə] wind tijdens de fietstocht.

Na het werk gingen we [lɔy] op de bank zitten.

Moeder riep [lɔyt] dat het eten klaar was.

Op de vensterbank lag een [lɔyjə] poes te zonnen.

De banjo's en [lɔytən] zijn kleine snaarinstrumenten.

Met een [lɔydə] knal stopte de auto.

De lange bruine [pɛi] wordt door priesters gedragen.

Voor sommige mensen is [pɛip] roken een echte hobby.

In de pastorie hangen de [pɛijən] netjes op een rij.

In een tabakszaak liggen [pɛipən] die erg kostbaar zijn.

De gids [lɛidə] het gezelschap rond.