Fluisterspraak in cochleaire implantaten

Een studie naar toonhoogtewaarneming in CI-gesimuleerde spraak.

Bachelorscriptie

Elise Catelijne Frederieke Prins
Studentnummer: 4022599
e.c.f.prins@students.uu.nl

Bachelor Taalwetenschap
Faculteit Geesteswetenschappen
Universiteit Utrecht
Blok 4, Juni 2015

Begeleiding: dr. W.F.L. Heeren
Universiteit Utrecht

Tweede lezer: dr. H.R. Terband
Universiteit Utrecht

Universiteit Utrecht

Samenvatting

In dit onderzoek staat toonhoogtewaarneming in verschillende spraaktypen in cochleaire implantaten centraal. Er wordt onderzocht of er verschil is in toonhoogtewaarneming wanneer proefpersonen naar gefluisterde versus normaal gesproken stimuli luisteren, als stimuli door een CI-simulator gehaald zijn. Eerder onderzoek (vb. Heeren, 2014) liet zien dat sprekers in fluisterspraak compenseren voor het feit dat zij geen F0 produceren. Dit staat tevens in lijn met de Hypo-en Hyperspeechtheorie die stelt dat sprekers zich aanpassen aan de luisteraar (Lindblom, 1990). Wellicht is deze compensatie voordelig voor CI-gebruikers omdat cochleair implantaten vaak de lage frequenties in spraak niet goed doorgeven, waardoor informatie over de F0 mist. Door 24 proefpersonen een discriminatietaak te laten maken, waarin zij geacht werden te letten op de toonhoogtes van stimuliparen, is onderzocht of het percentage correct herkende toonhoogtepatronen hoger lag wanneer zij naar gefluisterde stimuli luisterden. Alle stimuli samengenomen, bleek er echter geen significant verschil aanwezig te zijn tussen de scores bij het luisteren naar gefluisterde stimuli en de score bij het luisteren naar normaal gefoneerde spraak in het algemeen. Wel is er gevonden dat de stimuli met daarin de klinker /a/ significant beter gemaakt werden dan stimuli met als klinker /u/. Tevens bleek dat binnen de stimuli met als klinker /a/, de toonhoogtepatronen in de gefluisterde woordparen wél significant beter herkend werden dan die uit de normale, gefoneerde woordparen. In de andere twee klinkers is geen effect gevonden.

Sleutelwoorden: fluisterspraak, toonhoogtewaarneming, cochleair implantaat

Inhoudsopgave

	Pagina
Inleiding	4
Theoretisch kader	4
Onderzoeksvraag en hypothese	7
Methode	8
Analyse en Resultaten	10
Discussie en conclusie	13
Literatuur	15

Inleiding

Cochleaire implantaten (CI's) hebben het nadeel dat vooral de lage frequenties in spraak soms niet goed doorgegeven worden aan de elektroden, waardoor informatie over periodiciteit of de F0 uit het spraaksignaal dan ook miniem is. In fluisterspraak trillen de stembanden niet en mist daardoor óók de F0. Echter: fluisterspraak kan ondanks de afwezigheid van de F0, toch informatie bevatten over toonhoogte, stemhebbendheid en emotie (vb. Tartter, 1989; Tartter & Braun, 1994; Fónagy, 1969). Deze informatie komt tot stand door secundaire cues, informatie die wordt toegevoegd in zowel normale spraak als fluisterspraak; maar ook door eventuele compensatie van de spreker om de luisteraar het ongemak van de missende F0 weg te nemen, die enkel in fluisterspraak voorkomt. Deze compensatie kan wellicht voordelig zijn voor CI-gebruikers. Zou het zo kunnen zijn dat fluisterspraak akoestisch meer informatie bevat dan normale spraak met betrekking tot bijvoorbeeld toonhoogte door compensatie? Bosker et al. (2010) onderzochten of fluisterspraak inderdaad voordeligere input kan zijn voor CI-gebruikers, maar vonden geen resultaten die deze hypothese bevestigden. Dit zou kunnen betekenen dat fluisterspraak inderdaad niet voordeliger is, maar wellicht heeft het materiaal dat gebruikt werd in het experiment ook een rol aangezien niet onderzocht was of er wel compensatoire informatie in het materiaal aanwezig was. Materiaal waarvan vastgesteld is dat het inderdaad compensatoire informatie bevat, en een makkelijkere taak, aangezien de moeilijkheidsgraad een discussiepunt was bij Bosker et al. (2010), zouden wellicht tot een andere conclusie kunnen leiden. Presteren luisteraars luisterend naar CI-gesimuleerde spraak beter op een toonhoogte-discriminatietaak wanneer er fluisterspraak met daarin compensatoire informatie aangeboden wordt in plaats van normale spraak?

Theoretisch kader

Een cochleair implantaat is een gehoorprothese die bestaat uit een uitwendig en een inwendig gedeelte. Geluid wordt opgevangen via een microfoon in het uitwendige gedeelte en doorgestuurd naar een spraakprocessor. Vervolgens wordt het geluid omgezet in elektrische pulsen (Loizou, 2006). Deze pulsen worden doorgegeven aan elektroden in de cochlea, en via daar aan de zenuwuiteinden van de gehoorzenuw. De gehoorzenuw geeft het signaal door aan de hersenen. Ondanks de technologische vooruitgang is de CI nog geen vervanging van een goed werkend oor. Een goed werkend oor beschikt over ca. 3000 zintuigcellen en 30.000 zenuwuiteinden om de geluidssignalen aan door te geven; bij een cochleair implantaat wordt de functie van de 3000 zintuigcellen overgenomen door maximaal 22 of 24 elektroden. De elektroden kunnen wel afzonderlijk van elkaar

geactiveerd worden, waardoor zowel hoge als lage frequenties gestimuleerd worden. De elektrodes worden vooral geplaatst op plekken waar frequenties van ca. 250 tot 6000 Hz. gestimuleerd worden, omdat dit ook ongeveer het frequentiegebied van spraak is. Toch blijft het lastig voor cochleaire implantaten om met dit bereik de laagste frequenties over te brengen, mede door het beperkte aantal electroden. Informatie over periodiciteit en de F0 (grondfrequentie) is daardoor lastig te verkrijgen door CI-gebruikers. Naast het beperkte aanbod van de laagste frequenties uit het frequentiegebied, speelt ook de resolutie waarmee subtiele verschillen kunnen worden gecodeerd een rol. De frequenties worden ingedeeld in banden (in dit experiment acht banden) en de variatie binnen één bandbreedte is niet makkelijk hoorbaar. Doordat de informatie over onder andere periodiciteit en de grondfrequentie mist, mist er soms ook informatie over prosodie, klemtoon, het geslacht van de spreker, of over bijvoorbeeld stemhebbendheid.

CI-gebruikers hebben, zoals gezegd, vaak moeite met het horen van een F0, en juist die F0 is in gefluisterde spraak óók afwezig. Tijdens het fluisteren trillen de stembanden niet. Lange tijd werd gedacht dat het dan ook vrijwel onmogelijk was om verschil te maken in stemhebbende en stemloze klanken tijdens het fluisteren. Onderzoek van Tartter (1989) wees echter het tegengestelde uit. Ondanks het niet trillen van de stembanden, blijken mensen toch manieren te vinden om te compenseren voor het gebrek aan fonatie. Ze zorgen er op andere manieren voor dat de luisteraar doorheeft of een klank bijvoorbeeld stemhebbend of stemloos is. Ook in normale, gefoneerde spraak hoort men niet alléén aan de stemgeving of een klank stemhebbend of stemloos is, maar speelt andere informatie ook een rol. Het zou hier dus om secundaire cues kunnen gaan. In Tartters onderzoek (1989) werd ongetrainde luisteraars gevraagd naar consonanten aan het begin van een gefluisterde nonsens-syllabe, en het bleek dat zij de consonanten significant beter dan kansniveau konden herkennen. Dit terwijl men van gedachte was dat fluisterspraak voor sterke achteruitgang zorgde in spraakverstaan. Tartter (1989) gaf zelf aan dat de sprekers overcompenseren door verandering in gebruik van aspiratie, de plof, de duur van de fricatieven en de intensiteit ervan, en een eventuele verlaging van de eerste formant. Zij kwam niet alleen tot de conclusie dat het verschil tussen stemhebbend-stemloos toch te onderscheiden is in fluisterspraak, ook kunnen luisteraars het verschil tussen verdrietig en neutraal, of blij en neutraal horen in gefluisterde spraak (Tartter & Braun, 1994). In normale spraak verschaft de F0 veel informatie over emoties. Emoties met een hoge prikkeling (angst, woede, blijheid) zorgen voor een stijging van de gemiddelde F0 en emoties met weinig prikkeling (verdriet, verveling) zorgen voor een daling van de F0 (Klasmeyer et al., 1990). Zelfs zonder de F0 in fluisterspraak waren mensen in staat de emoties te herkennen. Luisteraars halen dus uit fluisterspraak informatie over stemhebbendheid en emotie. Higashikawa et al. (1996) vonden bovendien dat luisteraars de toonhoogte van een gefluisterde /a/ correct konden identificeren als normaal, hoog of laag. Ook het verschil tussen vraag en mededeling in zinnen waar de prosodie (en

niet de woordvolgorde) dit verschil maakt, blijkt goed herkend te worden in fluisterspraak (Fónagy, 1969; Heeren en van Heuven, 2009).

Het lijkt erop dat sprekers in de hogere frequenties compenseren om de luisteraars het ongemak van de missende F0 weg te kunnen nemen. Met compenseren wordt bedoeld dat er informatie toegevoegd wordt in het spraaksignaal die de spreker niet zou toevoegen op die plek wanneer hij in normale spraak spreekt. Compensatoire cues verschillen daarmee van secundaire cues, informatie voor opheldering voor de luisteraar in fluisterspraak die óók aanwezig is in normale spraak. In dit geval gaat het om compensatie in de fluisterspraak, omdat de spreker informatie toevoegt in fluisterspraak, wat hij niet ook zou doen in normale spraak op dat moment. Het idee dat een spreker compenseert om het voor de luisteraar makkelijker te maken past goed binnen de hyper-and hypo-speech-theorie van Lindblom (1990). De theorie stelt dat sprekers hun spraak aanpassen aan de behoeftes van de luisteraar. Wanneer men bijvoorbeeld tegen een slechthorend persoon spreekt of spreekt tegen luisteraars in een rumoerige setting, past men de spraak aan (Picheny et al., 1986). Dit zou ook het geval kunnen zijn in fluisterspraak, waar men misschien meer hyper-speech gaat gebruiken om de intonatie duidelijker te maken voor de luisteraar. Het onderzoek van Bosker et al. (2010) richtte zich ook op dit idee. Het idee dat sprekers compenseerden in fluisterspraak, leidde tot de hypothese dat dit wellicht voordelig zou zijn voor CI-gebruikers. De onderzoekers testten onder andere de hypothese dat gefluisterde spraak bijkomende ‘cues’ voor toonhoogte bevat om het gebrek aan F0 te compenseren, waardoor onder andere zinsprosodie, geslacht van de spreker en emotie beter waargenomen worden in gefluisterde spraak dan in normale (gefoneerde) spraak voor CI-gebruikers. De studie onderzocht dus in feite of fluisterspraak of normale spraak een betere input is voor CI-gebruikers. Hun hypothese was dat fluisterspraak voordeliger was, vanwege de compensatie die sprekers toepassen en vanwege de luidheid van de consonanten ten opzichte van de klinkers in fluisterspraak. Deze hypothese heeft echter in het onderzoek geen ondersteuning gekregen. Dit resultaat gaat in tegen de hypo-and-hyperspeech theorie van Lindblom (1990), waarin in de resultaten hyperspeech verwacht zou worden. De vraag is echter of de conclusie daarmee getrokken mag worden dat fluisterspraak inderdaad geen voordeel heeft ten opzichte van normale spraak in CI-gesimuleerde spraak. Een onderzoek met beter gecontroleerd materiaal en een andere taak kan opheldering geven.

Onderzoeksvraag en hypothese

Fluisterspraak zou in theorie voordeel kunnen hebben boven normale gefoneerde spraak voor CI-gebruikers. In het onderzoek van Bosker et al. (2010) kwam dit echter niet naar voren. Wellicht heeft dit met het materiaal en de gebruikte taak van Bosker et al. (2010) te maken. Er is in het onderzoek niet onderzocht welke akoestische correlaten er aanwezig waren in het materiaal en of er dus wel compensatie door de sprekers in te vinden was. Wellicht dat wanneer er materiaal wordt gebruikt waarvan al vastgesteld is dat sprekers inderdaad compenserende informatie produceren (Heeren 2014), de resultaten anders zullen zijn. Compenserende informatie kan bijvoorbeeld te zien zijn in de klinkerduur of duur van fricatieven in fluisterspraak. In het gebruikte materiaal bleek bijvoorbeeld dat sprekers moeite deden in fluisterspraak om de intensiteit van de fricatieven constant te houden gedurende de hele fricatief voor de luisteraar; iets dat sprekers niet deden in normale spraak.

Naast beter gecontroleerd materiaal is een ander soort taak wellicht een uitkomst, omdat in de discussie van Bosker et al. (2010) naar voren kwam dat de proefpersonen de taak te moeilijk vonden. Hier wordt dan ook gekozen voor een ander soort taak: een discriminatietaak voor toonhoogtepatronen.

De onderzoeksvraag is in dit onderzoek dan ook:
Is fluisterspraak voordeliger als input voor cochleaire implantaten met betrekking tot het herkennen van toonhoogtepatronen?

In dit onderzoek is onderzocht of er een verschil te zien is in score voor het waarnemen van verschillen in opeenvolgend hoog-laag of laag-hoog uitgesproken paren nonsenswoorden, als deze in gefluisterde spraak werden aangeboden versus normale spraak. Specifieker: Horen proefpersonen in CI-gesimuleerde spraak of twee nonsenswoorden een dalende of stijgende sprong in toonhoogte vertonen ten opzichte van elkaar en gaat dit beter in fluisterspraak?
Omdat er al een duidelijke voorspelling is van de richting die de resultaten op kunnen gaan, is ervoor gekozen om een eenzijdige alternatieve hypothese op te stellen.

H0: Er is geen significant verschil te vinden in prestatieniveau voor toonhoogtewaarneming tussen fluisterspraak en normale spraak.

H1: Met fluisterspraak als input voor CI's presteert men beter in toonhoogtewaarneming dan met normale spraak als input.

Methode

Materiaal

Het materiaal dat voor dit onderzoek is gebruikt is het materiaal van Heeren (2014). Het bestaat uit nonsens syllabes opgebouwd uit steeds klinker-consonant-klinker (V_iCV_i). De consonant is ofwel /s/ ofwel /f/ en de klinkers zijn /i/, /a/ of /u/. Dit resulteerde in de stimuli: *isi, ifi, asa, afa, usu* en *ufu*. De sprekers hebben elke stimulus vier keer achter elkaar uitgesproken in ofwel een stijgend, ofwel een dalend toonhoogtepatroon. Alleen de eerste drie realisaties zijn meegenomen (hoog, midden, laag). De vierde realisatie is weggelaten om het effect van (zins-)finale verlenging tegen te gaan. In totaal waren er 6 mannelijke sprekers en 6 vrouwelijke sprekers, die elk normale en gefluisterde spraak produceerden.

In dit onderzoek zijn alleen de hoge en de lage realisaties meegenomen. Door voor een hoge en een lage toon te kiezen, wordt het verschil zo groot mogelijk. Aan de ene kant moet in dat geval gelet worden op de ecologische validiteit, omdat zo'n groot verschil in toonhoogte achter elkaar niet vaak voorkomt in spontane spraak. Aan de andere kant, en vandaar ook de keuze, is het zo dat áls er in deze situatie gevonden wordt dat fluisterspraak geen voordelen heeft ten opzichte van normale spraak, dat mogelijk nergens te vinden zal zijn.

Alle stimuli verkregen van het onderzoek van Heeren (2014) zijn door een CI-simulator gehaald om na te bootsen hoe een CI-gebruiker de stimuli ongeveer zou horen. De instellingen voor deze CI-simulator zijn gelijk gesteld aan de instellingen die Bosker e.a. (2010) gebruikt hebben voor de simulator TIGERCIS (http://www.tigerspeech.com/. De stimuli zijn door een noise vocoder met acht kanalen geleid. Band-passfilters zijn gebruikt om het signaal in acht kanalen te verdelen tussen 351 en 8333 Hz, met als instellingen voor de kanalen: (1) 351-496 Hz, (2) 496-699 Hz, (3) 699-987 Hz, (4) 987-1393 Hz, (5) 1393-1965 Hz, (6) 1965-2773 Hz, (7) 2773-3913 Hz, (8) 3913-8333 Hz (Bosker et al., 2010).

De data bevat gesproken en gefluisterde spraak van 12 deelnemers. Deze hebben in elk spraaktype allen in twee volgordes opgenomen, namelijk *hoog, midden, laag* versus *laag, midden, hoog* (dalend versus stijgend). Verder kwamen de stimuli in het experiment in twee presentatievolgordes voor (eerst de hoge toon en dan de lage, versus eerst laag, dan hoog). 12 sprekers x 2 opnamevolgordes x 3 klinkers x 2 consonant contexten x 2 presentatievolgordes: leidt tot 288 stimuli per spraaktype (gefluisterd of gesproken). Deze stimuli zijn opgedeeld in twee groepen van 144 stimuli. Elke proefpersoon luistert dus naar 144 stimuli per spraaktype. De taak voor de participanten was het vergelijken van telkens een set van twee stimuli, die enkel verschilden in toonhoogte. De spreker, het spraaktype, en de volgorde van opnemen zijn alle gelijk binnen één paar. Er moest bij elk paar van stimuli beantwoord worden of de tweede CVC hoger of lager klonk ten

opzichte van de eerste CVC voor die proefpersoon. Het experiment is gemaakt in het programma Praat (Boersma & Weenink, 2015). Om een eventueel effect van een voorkeurshand te vermijden is ervoor gekozen om bij de helft van de proefpersonen de antwoordoptie 'hoger' rechts in beeld neer te zetten, en bij de andere helft links in beeld. De volgorde van de twee knoppen is ook gematcht met de volgorde van die woorden in de uitleg en de tekst bovenin het scherm. De helft van de proefpersonen heeft eerst alle gesproken stimuli gehoord, en daarna alle gefluisterde stimuli, en de andere helft van de proefpersonen heeft de spraaktypes in omgekeerde volgorde gekregen. Alle stimuliparen zijn binnen elk spraaktype gerandomiseerd aangeboden door het programma Praat. Op vaste plekken in het experiment en tevens tussen de twee helften was er de mogelijkheid om pauze te houden voor de participant.

Alle proefpersonen maakten de taak op een Samsung NP535U3C notebook (2014, Windows 8.1) met dezelfde koptelefoon zodat een eventueel effect van verschillende luisterapparatuur bij de proefpersonen thuis is weggenomen.

Proefpersonen

Er hebben in totaal 21 vrouwelijke en 3 mannelijke proefpersonen geparticipeerd (n=24). Zij voldeden aan de volgende eisen:

- Tussen de 18 en 30 jaar oud
- Niet dyslectisch
- Niet meertalig opgevoed
- Geen gehoorstoornis

De proefpersonen hebben allen vrijwillig geparticipeerd en ontvingen geen vergoeding.

Procedure

De test werd afgenomen in een rustige ruimte van een universiteitsgebouw in de binnenstad van Utrecht. De proefpersonen maakten voorafgaand aan het werkelijke onderzoek eerst een oefentest, met daarin vier voorbeelden. Zo kon men wennen aan het geluid van een simulatie van een cochleair implantaat en kon de onderzoeker testen of men de taak begrijpt. In de oefenronde kreeg men feedback door het tonen van een groen ofwel rood scherm na het gegeven antwoord. Er werd vier maal een set van twee stimuli aangeboden, beide uit hetzelfde spraaktype en dezelfde spreker: een hoge variant en een lage variant. In beeld stond voor de proefpersonen de vraag 'was de tweede hoger of lager?' en men had vervolgens de keuze tussen 'hoger' of 'lager'. Na de oefenronde was er nog gelegenheid voor vragen en daarna begon het echte experiment. In het experiment werd weer telkens een set van twee stimuli aangeboden, beide uit hetzelfde spraaktype (of allebei gefluisterd, of allebei gesproken) en van dezelfde spreker. De stimuli verschilden dus enkel in de toonhoogte waarop

zij waren uitgesproken. De vraag voor de proefpersonen was: werd het tweede woord (CVC) uitgesproken op een hogere of lagere toonhoogte dan het eerste? Een discussiepunt bij Bosker e.a. (2010) was dat de taak wellicht te moeilijk was voor de proefpersonen. In dit experiment lag de moeilijkheidsgraad lager. Proefpersonen hoefden enkel twee nonsenswoorden met elkaar te vergelijken op hun toonhoogte en daar een uitspraak over te doen, en dat is makkelijker dan een herkenningstaak. Wanneer een proefpersoon de taak af had, werden de resultaten opgeslagen door de examinator.

Analyse en Resultaten

Er zijn van 24 proefpersonen resultaten; iedere proefpersoon heeft de taak voltooid. De gegeven antwoorden zijn in Excel ingevoerd en omgerekend naar percentage correct gegeven antwoorden per persoon. Deze percentages zijn ingevoerd in SPSS (IBM, versie 22) om de analyses uit te kunnen voeren.

Allereerst is de variabele 'volgorde' onderzocht, om de resultaten te testen op ongewenste effecten. Het zou zo kunnen zijn dat proefpersonen het gefluisterde deel heel anders maken als dit deel als eerst aan bod komt dan wanneer hij als tweede aan de beurt is. Het zou ook zo kunnen zijn dat men standaard de tweede helft beter maakt. Dit bleek niet zo te zijn; er was geen significant verschil in score te vinden tussen de proefpersonen die volgorde 1 (eerst gesproken, dan gefluisterd), of volgorde 2 (eerst gefluisterd, dan gesproken) hadden gemaakt ($t(22)=0.975$, $p=0,340$).

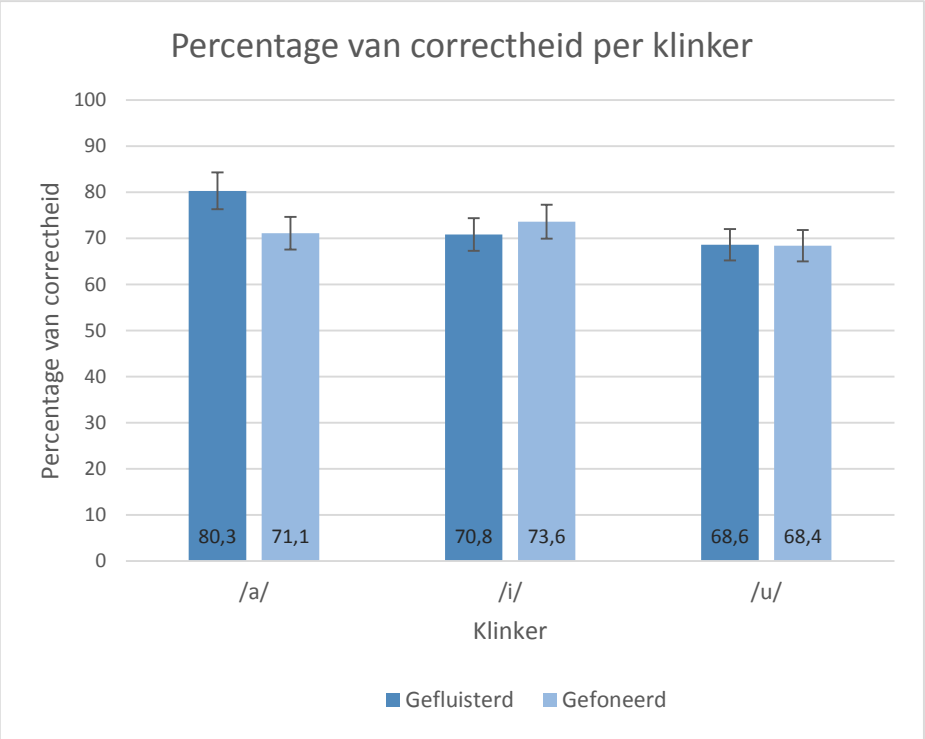
Na de controle op de data kon de onderzoeksvraag onderzocht worden. Was fluisterspraak voordeliger als input voor cochleaire implantaten met betrekking tot het herkennen van toonhoogtepatronen?

In overeenstemming met deze onderzoeksvraag en hypothese is gekeken of het percentage correct gegeven antwoorden hoger was wanneer er twee gefluisterde VCV's elkaar volgden, dan wanneer twee normaal gesproken VCV's elkaar volgden. Tevens is onderzocht of de klinker in de desbetreffende stimulus een effect kon hebben. De hypothese dat de uitgesproken klinker een effect kan hebben op het resultaat, komt voort uit eerder onderzoek van bv. Heeren & Lorenzi (2014). In de perceptie van prosodie spelen de formanten een significante rol en aangezien de locatie en sterkte van de formanten per klinker verschillen; kunnen er perceptieve verschillen zijn tussen klinkers. In de resultaten van Heeren & Lorenzi (2014) werd dan ook verschil gevonden tussen klinkers. Er werden in dit onderzoek meer correcte antwoorden gegeven voor de klinkers /a/, /e/ en /o/ dan voor de klinkers /i/ en /u/. Dit is te verklaren met het feit dat de open klinkers een voordeel hebben qua intrinsieke

intensiteit in vergelijking met de gesloten klinkers. Vooral de open klinker /a/ had een groot voordeel ten opzichte van de gesloten klinkers (/i/,/u/) door een grotere intensiteit. De middenklinkers tot open klinker leken betere dragers van informatie over intonatie te zijn dan de gesloten klinkers. Er is in dit onderzoek dan ook voor gekozen om tevens te kijken naar de resultaten van alle stimuli gesorteerd per klinker.

Om alle resultaten te analyseren is gebruik gemaakt van een tweeweg ANOVA met als afhankelijke variabele de score (in percentage van correctheid) en als onafhankelijke variabelen het spraaktype (gefluisterd of gesproken) en klinker (/a/,/i/ of /u/).
Er werd een effect gevonden voor *klinker* ($F(1,46)=6.21$; $p=0,016$) en een effect voor de interactie *klinker*speechmode* ($F(1, 46)=7.32$; $p=0,0095$). Om de details van deze interactie te onderzoeken is een post-hoctest uitgevoerd (Tukey).

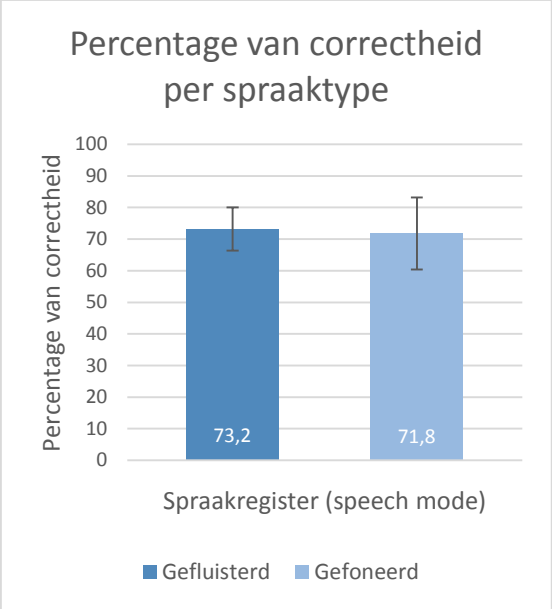
In de normaal gesproken woorden met daarin de klinker /a/ was er een correctheidspercentage van 71,1 %, terwijl dit in de gefluisterde woorden met /a/ 80,3 % was (Figuur 1). Bij de /i/ is te zien dat juist de gesproken stimuli beter gemaakt werden, en bij de /u/ maakte het voor het resultaat vrijwel niet uit of er gefluisterd of gesproken werd. De interactie klinker*speechmode bleek significant te zijn bij een 95 %- betrouwbaarheidsinterval ($F(1, 46)=7.32$; $p=0,0095$) en zelfs ook nog wanneer als alpha-waarde 0,01 genomen wordt in plaats van 0,05. Na een posthoctest bleek dat dit verschil te vinden was in de stimuli met klinker /a/: De scores van enerzijds de gefluisterde en anderzijds de gesproken stimuli in de woorden met /a/ bleken significant van elkaar te verschillen. De verschillen gevonden in de woorden met /i/ en /u/ zijn niet significant.



Figuur 1. Percentage van correctheid per klinker. Foutbalken: 95% betrouwbaarheidsinterval.

Verder werd er zoals gezegd een effect gevonden voor klinker. Dit is tevens met een posthoctest onderzocht, en het bleek te gaan om het verschil in algemene score tussen de woorden met een /a/ en met een /u/ erin. Er was sprake van een correctheidpercentage van 75,7% correct in stimuli met een /a/ versus 68,5% correct in stimuli met /u/. Dit verschil bleek significant: de toonhoogtewaarneming in woorden met een /a/ was significant beter dan de score voor de woorden met de /u/ (p=0,016).

In overeenstemming met de hypothese is ook het effect van spraaktype onderzocht. Het gemiddelde percentage van correctheid was in de fluisterconditie **73,2%** en in de gesproken conditie **71,8 %** (Figuur 2). De medianen waren respectievelijk 75,4 % en 71,2 %. In beide condities is er ver boven kansniveau geantwoord. De ANOVA gaf geen significant effect aan voor spraaktype tussen deze twee gemiddeldes van de correctheidspercentages ($F(1,46)=0.95$; $p=0,340$). Er kan dus niet gezegd worden dat het spraaktype invloed had op de score.



Figuur 2. Percentage van correctheid per spraaktype. De foutbalken geven de standaarddeviatie weer.

Hiermee is de onderzoeksvraag beantwoord: fluisterspraak had gemiddeld genomen geen voordeel in deze taak ten opzichte van normale spraak. Dit effect werd alleen gevonden in de stimuli met een /a/.

Dan is er nog als extra deelvraag te onderzoeken of proefpersonen die over het algemeen moeite hebben met het herkennen van toonhoogtepatronen wél baat kunnen hebben bij gefluisterde spraak. Om dit te onderzoeken zijn de resultaten opgesplitst in twee groepen: de resultaten van personen die boven het algemene gemiddelde scoorden en de personen die daaronder scoorden. Uit de resultaten bleek dat in de groep proefpersonen die onder het gemiddelde scoorden, in de

gesproken conditie het percentage goed beantwoorde stimuli 62,6 % was, waar dit in de fluisterconditie 67,9 % was. De proefpersonen die meer moeite dan gemiddeld hadden met de taak leken het dus beter te doen met de fluisterspraak (een verschil van 5,3 %), maar deze percentages bleken na analyse echter niet significant van elkaar te verschillen ($t(22)=1.29$; $p=0,211$).

Verder was de spreiding in allebei de condities interessant. In de fluisterconditie was de standaarddeviatie 6,81 en in de spraakconditie bedroeg deze 11,39. De steekproefvariantie in de fluisterconditie was 46,42 en in de normaal gesproken conditie 129,75. Het lijkt er dus op dat men in fluisterspraak consistenter is wat betreft het juiste antwoord geven. Deze getallen zijn echter nog niet statistisch met elkaar vergeleken. Na analyse daarvan zou onderzocht kunnen worden of hier toch nog een klein voordeel zou kunnen zijn voor fluisterspraak wat betreft consistentie. Men doet het dus over het algemeen niet beter in fluisterspraak, maar er zijn wel minder grote missers dan in de normaal gesproken conditie.

Discussie en Conclusie

In dit onderzoek is bekeken of gefluisterde spraak een voordeel kan bieden in een toonhoogtewaarnemingstaak ten opzichte van normale, gefoneerde spraak als deze spraak door een simulator voor een cochleair implantaat is gehaald en als stimuli gebruikt zijn waarvan bekend is dat er compensatie in de fluisterspraak aanwezig is. Gesproken en gefluisterde stimuli zijn verwerkt door een noise vocoder om de stimuluskwaliteit van de spraak na te bootsen zoals een CI-gebruiker die zou horen. Participanten luisterden telkens naar een paar van twee nonsenswoorden en gaven daarbij aan of zij een dalende of stijgende sprong waarnamen ten opzichte van elkaar. Daarover kan het volgende geconcludeerd worden. Normaal horende proefpersonen horen in CI-gesimuleerde spraak of er een dalende of stijgende sprong in twee opeenvolgde nonsenswoorden zit; aangezien de gemiddelden boven kansniveau liggen. Of deze stimuli gefluisterd of normaal gesproken worden aangeboden geeft geen significant effect op de score van correctheid. Wat betreft de hypothese kan daarmee geconcludeerd worden dat de alternatieve hypothese niet wordt geaccepteerd en de nulhypothese niet verworpen blijft.

Wel is te zien dat de spreiding in de fluisterspraak minder groot is dan in de normaal gesproken stimuli, wat wellicht kan duiden op dat proefpersonen consistenter zijn wanneer zij fluisterspraak aangeboden krijgen. Dit opvallende verschil in spreiding zou nog verder onderzocht kunnen worden in een vervolgonderzoek.

Heeren en Lorenzi (2014) leverden motivatie om de resultaten per klinker te bekijken.

Aangezien verschillende klinkers verschillende formanteigenschappen hebben en er verschil kan zijn in intensiteit van de stimulus doordat een klinker bijvoorbeeld meer open is dan een andere klinker, kunnen er verschillen in de waarneming zijn tussen verschillende klinkers. De formanten van klinkers zijn in fluisterspraak altijd iets hoger dan in normale spraak. Het was dan ook nuttig om de resultaten van dit onderzoek ook uiteen te zetten per klinker. Het bleek dat in de woorden met als klinker /a/, er wel een significant effect gevonden is voor het verschil in score voor fluisterspraak versus normale spraak. Bovendien werden deze woorden, in lijn met de resultaten van Heeren & Lorenzi (2014), significant beter gemaakt dan de woorden met een /u/ erin. Dit verschijnsel wordt waarschijnlijk veroorzaakt door een grote intensiteit in de open klinker wat leidt tot een betere verstaanbaarheid. Het lijkt erop dat sommige klinkercontexten betere dragers zijn voor informatie over intonatie dan andere.

Wat betreft de interpretatie van de resultaten moeten enkele discussiepunten meegenomen worden. Het feit dat er geen verschil werd gevonden tussen de prestaties in fluisterspraak en normale spraak was in lijn met het onderzoek van Bosker et al. (2010). Het zou inderdaad zo kunnen zijn dat het voordeel voor fluisterspraak niet bestaat. Toch werd er wel enige evidentie gevonden in de stimuli met klinker /a/.

Ook moet stilgestaan worden bij het feit dat een CI-simulator niet per se gespiegeld mag worden met hoe een echte CI-gebruiker deze taak zou maken. Met duidelijke uitspraken over CI-gebruikers moet dan ook voorzichtig worden omgesprongen. Verder is de ecologische validiteit van de stimuli een punt om over na te denken. Het komt niet vaak voor dat een luisteraar zo vaak naar twee nonsenswoorden achter elkaar luistert in deze context. Bovendien hoort hij niet vaak zulke sprongen in toonhoogte waarop deze nonsenswoorden zijn uitgesproken. Tevens zijn hier woorden in isolatie bekeken, wat andere resultaten geeft dan woorden in een zinscontext.

In een vervolgonderzoek zou het ook belangrijk zijn om evenveel vrouwen als mannen mee te nemen in het onderzoek. In deze studie is dat wegens tijdsredenen niet gelukt. Er zou dan gekeken kunnen worden of mannen een ander patroon vertonen dan vrouwen in toonhoogtewaarneming en welke invloed het spraaktype hier op heeft per geslacht. In Heeren & Lorenzi (2014) werd al evidentie gevonden dat vrouwelijke en mannelijke sprekers verschil vertoonden in score op woorden met een /a/ of /i/.

Tevens zou een vervolgonderzoek opgezet kunnen worden waarin er een extra conditie toe wordt gevoegd. Naast normale spraak en gefluisterde spraak zou er gefluisterde spraak specifiek met een bereik tussen de 1.5-2 kHz aangeboden kunnen worden. Informatie over prosodie is namelijk met name in die (hogere) regio aanwezig, dankzij de afwezigheid van harmonischen in fluisterspraak (Heeren & Lorenzi, 2014). In fluisterspraak blijkt de regio met de hoge frequenties het meest informatief te zijn voor luisteraars, waar dit in normale spraak de lage frequenties zijn (Heeren &

Lorenzi, 2014). Er zou in dit geval gekeken kunnen worden of fluisterspraak met een bepaald bereik voordeel kan hebben boven normale spraak en normale fluisterspraak.

Al met al zijn de algemene resultaten voor fluisterspraak of normale spraak in lijn met het onderzoek van Bosker et al. (2010), in de zin van dat een echt significant effect voor spraaktype niet is gevonden. Toch werd er in dit onderzoek wel een kleine evidentie voor een voordeel voor fluisterspraak gevonden in de stimuli met klinker /a/. Het is nuttig gebleken om de resultaten per klinker te analyseren. Tevens zou het interessant zijn om de spreiding van de resultaten in beide spraaktypen nader te onderzoeken.

Literatuurlijst

Boersma, P., & Weenink, D. (2015). Praat: doing phonetics by computer [Computerprogramma], Versie 5.4.09. Gebruikt mei 2015 via http://www.praat.org/.

Bosker H.R., Briaire J.J., Heeren W.F.L., Heuven V.J.J.P. van & Jongman S. (2010). Whispered speech as input for cochlear implants. *In: Kampen J. van, Nouwen R. (Eds.) Linguistics in The Netherlands 2010. Amsterdam: John Benjamins. 1-15*.

Fónagy, J. (1969). Accent et intonation dans la parole chuchotée. *In: Phonetica 20, pp. 177–192. Engels vertaald.*

Heeren, W.F.L. (2014). Acoustic Correlates of Height in the Production of Whispered Fricatives. *Leiden University Centre for Linguistics, Leiden Institute for Brain and Cognition Leiden University.*

Heeren, W.F.L. (2014). Fricative cues to pitch in whispered versus normal speech. *In: Fuchs S., Grice M., Hermes A., Lancia L., Muecke D. (Eds.) Proceedings of the 10th international Seminar on Speech Production, Cologne, May 5-8 2014*. 182-185.

Heeren, W.F.L. & Heuven, V.J.J.P. van (2009). Perception and production of boundary tones in whispered Dutch. *In: Proceedings of Interspeech, September 6-10, 2009. Brighton UK. 2411-2414.*

Heeren, W.F.L. & Lorenzi, C. (2014). Perception of prosody in whispered French. *Journal of the Acoustical Society of America. 135*, 2026–2040.

Higashikawa, M., Nakai, K., Sakakura, A, & Takahashi, H. (1996). Perceived pitch of whispered vowels – relationship with formant frequencies: a preliminary study. *Journal of Voice* 10.155–158.

Klasmeyer, G. & Sendlmeier, W.F. (1999). Voice and Emotional States. Ed. Kent, R. D. and Ball, M. J. Voice Quality Measurement, Singular Publishing Group, 339- 359, 1999

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H and H theory. *Hardcastle W, Marchal A. Speech production and speech modeling. Kluwer; Dordrecht: 1990. pp. 403–439.*

Loizou, P.C. (2006). Speech processing in vocoder-centric cochlear implants. *Cochlear and Brainstem Implants ed. by Aage R. Møller,* 109-143. Basel: Karger. Engels vertaald.

Picheny, M.A., Durlach, N.I. & Braida, L.D. (1986). Speaking clearly for the hard of hearing II: acoustic characteristics of clear and conversational speech. *In: Journal of Speech and Hearing Research 29, pp. 434–446.*

Tartter, V. (1989). What's in a whisper? *Journal of the Acoustical Society of America* 86.1678–1683

Tartter, V. & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *Journal of the Acoustical Society of America* 96.2101-2107.