



Universiteit  
Leiden  
The Netherlands

## **The processing of Dutch prosody with cochlear implants and vocoder simulations**

Velde, D.J. van de

### **Citation**

Velde, D. J. van de. (2017, July 5). *The processing of Dutch prosody with cochlear implants and vocoder simulations*. LOT dissertation series. LOT, Utrecht. Retrieved from <https://hdl.handle.net/1887/50406>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/50406>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/50406> holds various files of this Leiden University dissertation.

**Author:** Velde, D.J. van de

**Title:** The processing of Dutch prosody with cochlear implants and vocoder simulations

**Issue Date:** 2017-07-05

---

## Samenvatting in het Nederlands

---

In dit proefschrift is de verwerking van prosodie door gebruikers van cochleair implantaten (CI) onderzocht. Prosodie is spraakinformatie die niet gereduceerd kan worden tot individuele segmenten en opeenvolgingen van segmenten. De voornaamste dimensies waarin het varieert zijn de fundamentele frequentie (F0), intensiteit en duur van delen van uitingen. Het geeft onder meer informatie over informatiestructuur (zoals de het markeren van nieuwe informatie, ook wel focus genoemd), de manier waarop zinnen in de uitspraak in woordgroepen worden onderverdeeld (frasering), zinstype (vraag of mededeling), alsook de emotie en attitude waarmee de spreker een uiting realiseert. Prosodie is tegelijkertijd belangrijk voor spraakbegrip en (wat betreft bepaalde aspecten ervan) vormt een berucht struikelblok voor CI-gebruikers. Daarmee is het een belangrijk onderwerp voor onderzoek naar die populatie. In de onderzoeken voor dit proefschrift zijn drie soorten deelnemers onderzocht, namelijk kinderen met CI's, normaalhorende (NH) kinderen zonder CI's en NH volwassenen die luisteren naar simulaties (vocoders) van het horen met een CI (alsook, als controleconditie, naar niet-gevocoderde stimuli). Het onderzoeksonderwerp van dit proefschrift is benaderd vanuit vijf perspectieven: (1) taalkundige versus emotionele prosodie, (2) de perceptie en de productie van prosodie, (3) prosodie en muziek, (4) de weging van cues bij de waarneming van prosodie en (5) de ontwikkeling van prosodie bij kinderen. Deze vijf perspectieven zijn

verdeeld over vijf onderzoeken en worden behandeld in vijf onderzoekshoofdstukken. Elk van die hoofdstukken wordt hieronder samengevat.

In Hoofdstuk 2 zijn de eventuele verschillen tussen basale prosodische maten op het gebied van F0- en duurvariaties, in de spontane spraak van vroeg- en geïmplanteerde kinderen vergeleken met die van NH kinderen, van wie de hoorleeftijd correspondeerde met die van de klinische groep. Metingen hebben plaatsgevonden op drie tijdstippen – 18, 24 en 30 maanden – na implantatie (voor de CI-groep), dan wel na geboorte (voor de controlegroep). De hypothesen waren (1) dat afwijkingen in de prosodische F0-maten bij CI-kinderen groter zouden zijn dan die in duurmaten, (2) dat de spraak van laatgeïmplanteerde sterkere afwijkingen zou vertonen dan die van vroeggeïmplanteerde kinderen en (3) dat afwijkingen zouden afnemen als functie van de hoorleeftijd. De eerste twee hypothesen werden niet ondersteund door de resultaten, omdat geen systematische verschillen in afwijkingen waren geobserveerd tussen prosodische maten noch tussen de twee klinische groepen. De resultaten veronderstelden echter dat CI-kinderen meer afwijkingen vertoonden als het ging om parameters die een grotere mate van beheersing vragen van de uitspraak van prosodie ten opzichte van maten die kunnen worden beschouwd als automatische neveneffecten van spraak. Dit zou een reflectie kunnen zijn van waarnemingsproblemen. De derde hypothese werd ondersteund door de resultaten, omdat de prestaties van de meeste maten minder afwijkend werden op latere testmomenten.

In Hoofdstuk 3 wordt verslag gedaan van een onderzoek waarbij de perceptie van gevocoderde intonatiecontouren door NH volwassenen is getest. De stimuli bestonden uit varianten als natuurlijke uitingen opgenomen korte Nederlandse frases, zoals *een agenda*, die alleen door de F0-contour van elkaar te onderscheiden waren. De F0-contouren waren gestileerde versies van varianten de frases waarin respectievelijk verrassing, nieuws en teleurstelling werd uitgedrukt. Vervolgens waren de stimuli gevocoderd met filterhellingen van 20 dB/octaaf en 40 dB/octaaf. Participanten werd

in drie verschillende (de twee filterhellingcondities en een conditie zonder vocoding) gevraagd om aan te geven welk van die types was uitgedrukt. De prestaties in de gevocoderde condities waren lager (op kansniveau) dan die in de niet-gevocoderde conditie (rond de 90% correct), maar er was geen verschil tussen de twee filterhellingcondities onderling. Deze resultaten tonen aan dat dit type vocoding de waarneming van prosodische contrasten op basis van alleen de F0, onmogelijk maakte maar dat voor scores boven kansniveau hoogstwaarschijnlijk scherpere filterhellingen nodig zouden zijn.

Het onderzoek dat in Hoofdstuk 4 is beschreven is een uitbreiding op het onderzoek uit Hoofdstuk 4. In plaats van slechts twee filterhellingen te testen (20 en 40 dB/octaaf) zijn nu vijf hellingen getest (5, 20, 80, 120, en 160 dB/octaaf). De stimuli bestonden uit korte frases van de vorm 'LIDWOORD-BIJVOEGLIJK NAAMWOORD-ZELFSTANDIG NAAMWOORD' (zoals *een blauwe bal*), opgenomen in vijf varianten, te weten twee emoties (verdrietig en blij), twee focusposities (op het bijvoeglijk naamwoord en op het zelfstandig naamwoord) en een neutrale variant (zo veel mogelijk zonder specifieke emotie en met gelijkwaardige focus op het bijvoeglijk naamwoord en het zelfstandig naamwoord). Deze stimuli waren opgenomen als natuurlijke uitingen en vervolgens was ofwel de F0-contour, waren de segmentduren of waren beide types informatie van de neutrale variant vervangen door die van de niet-neutrale varianten. Zo waren voor gebruik van elk deexperiment (een emotietest en een focustest) drie nieuwe half-synthetische varianten gecreëerd. In elk van beide testen was de enige beschikbare informatie om de emoties (in de ene test) dan wel de focusposities (in de andere test) van elkaar te onderscheiden de vervangen cue. Als laatste stap waren de stimuli gevocoderd met een 15-kanaals-ruisvocoder. Deelnemers werd gevraagd om voor elk van zes condities bestaande uit vijf filterhellingen en een conditie zonder vocoding aan te geven welke emotie dan wel, in een aparte test, welke focuspositie ze dachten dat was uitgedrukt. De prestaties in de conditie zonder

vocoding waren tegen het plafondniveau aan, wat aantoont dat de spreker de emoties en focusposities succesvol had uitgedrukt. In de vocodercondities varieerden de scores van nabij het kansniveau in het geval van het minst scherpe filter (5 dB/octaaf) tot hoge scores in het geval van filters van 120 dB/octaaf; scores waren in het algemeen in de focustest echter lager dan in de emotietest. De scores bij 160 dB/octaaf waren vergelijkbaar met die bij 80 dB/octaaf, beide lager dan bij 120 dB/octaaf. In de emotietest was het scorepatroon in de conditie met beide cues beschikbaar het gelijkst aan die met alleen F0-informatie beschikbaar, terwijl die in de focustest het gelijkst was aan die met alleen duurinformatie beschikbaar. Samengenomen tonen deze resultaten aan dat verscherping van de filters een positief effect heeft op prosodiewaarneming tot de extreme waarde van 120 dB/octaaf, maar dat dit effect (gegeven de stimuli die in dit experiment waren gebruikt) sterker is voor focuspositiewaarneming dan voor emotiewaarneming, omdat de F0-cues voor emoties informatiever zijn dan voor focuspositie. De filterhelling van 120 dB/octaaf kan als theoretische doelwaarde worden gebruikt voor spraakverwerkingsalgoritmes voor CI's in de toekomst.

Hoofdstuk 5 bevat het verslag van een onderzoek waarbij NH volwassenen een korte training van 45 minuten krijgen om ofwel temporele (één groep) ofwel melodische (een tweede groep) contrasten beter leerden waarnemen in gevocoderde muzikale stimuli. Het doel van dit onderzoek was om te testen of zulke cue-specifieke training een groter vertrouwen op de getrainde cue dan op de niet-getrainde cue zou teweegbrengen. De antwoorden op een vooraf ingevulde vragenlijsten lieten zien dat de muzikale achtergronden van de groepen niet van elkaar verschilden. Na de training voerden de deelnemers de testen naar de waarneming van emotie en focustest uit die in Hoofdstuk 4 zijn beschreven, alsook een herkenningstesten van bekende muzikale melodieën met ofwel duur- ofwel F0- ofwel beide cues beschikbaar, en als derde een test waarin werd onderzocht of ze een voorkeur hadden om sequenties van vier noten met ambigue startposities (dat wil zeggen, op de luidste of op de hoogste noot) op

basis van ritme of op basis van melodie te segmenteren. In dit onderzoek zijn geen trainingseffecten van het ene domein naar het andere (van muziek naar prosodie) noch van de ene cue naar de andere (van durcues naar F0- of melodische cues of andersom) gevonden, hoewel er wel een tendens van een trainingseffect voor de getrainde cue bij de bekende-melodieëntest is geobserveerd. Ook is een segmentatievoorkeur in de ambigue-melodieëntest geconstateerd die overeenkwam met de getrainde cue. Tot slot zijn correlaties op het gebied van de individuele deelnemers gevonden voor prestaties bij beschikbaarheid van ongelijke cues en van ongelijk domein. Deze resultaten laten bij elkaar zien dat langere cue-specifieke training een positief effect zou kunnen hebben op effecten binnen en tussen cues en domeinen waarbij de perceptie van muzikale melodieën en prosodie wordt verbeterd.

In Hoofdstuk 6 komen vier van de vijf perspectieven van het proefschrift samen. Van zes tot twaalf jaar oude kinderen met CI's en in hoorleeftijd overeenkomende NH kinderen werd het gebruik van cues onderzocht in vier verschillende computertesten op het gebied de perceptie en productie van taalkundige en emotionele prosodie, alle gebruikmakend van in hoge mate overeenkomstige stimuli. Behalve het kwartet aan kerntesten (de perceptie en productie van zowel taalkundige als emotionele prosodie), is hun algemene nonverbale emotionele en taalkundige ontwikkelingsniveau getest door middel van respectievelijk zinnen en situaties die bepaalde emoties uitdrukken en nonwoordherhaling. De prestaties verschilden bij deze testen niet significant tussen de deelnemersgroepen, wat aangeeft dat ze vergelijkbare basisontwikkelingen hadden. Voorafgaand aan de kerntesten werden kinderen bekendgemaakt met de procedure en de stimuli door middel van simpele benoemings- en identificatietaken; beide groepen scoorden op of nabij het plafondniveau. Van de kerntesten waren die van de perceptie van emotie en focuspositie gelijk aan die uit Hoofdstuk 4, inclusief de exacte stimuli en beschikbaarheid van verschillende cues. In de focusproductietest werd deelnemers gevraagd antwoord te geven op een vraag van de vorm *Is*

*dit een blauwe bal?*, waarbij ofwel het bijvoeglijk naamwoord, ofwel het zelfstandig naamwoord of beide contrasteerde met een afbeelding op het scherm. In de emotieproductietest werd kinderen gevraagd om een afbeelding te beschrijven (bijvoorbeeld van een rode stoel) en dat op een verdrietige of blijde manier te zeggen, afhankelijk van de emotie van een gezichtje dat erbij op het scherm stond. De geproduceerde emoties en focusposities werden naderhand als zodanig beoordeeld door een onafhankelijk panel van tien Nederlandse volwassenen. De resultaten lieten geen verschil in cue-wegingstrategie tussen de twee groepen zien, noch in de effectiviteit van de emoties en focusposities in hun producties (dit geldt met name voor emotionele prosodie, omdat de resultaten van de focusperceptietest niet konden worden geanalyseerd). Er zijn echter zwakke correlaties gevonden tussen emotionele-prosodieperceptie en -productie alsook tussen, aan de ene kant, emotionele-prosodieperceptie en -productie en, aan de andere kant, nonverbale-emotiebegrip in de CI-groep maar niet, of in mindere mate, in de NH groep. Tot slot bleek hoorleeftijd in beide groepen de scores in de emotieproductietest maar niet die in de emotieperceptietest in beperkte mate te voorspellen. Samengenomen veronderstellen deze resultaten dat CI-kinderen op deze leeftijd, ondanks de problemen en uitstel die ze kunnen ervaren door hun nadelige gehoorsituatie, de achterstand op het gebied van de perceptie en productie van prosodie op hun leeftijdsgenoten hebben ingehaald.