



Universiteit  
Leiden  
The Netherlands

## Speech across species : on the mechanistic fundamentals of vocal production and perception

Ohms, V.R.

### Citation

Ohms, V. R. (2011, May 3). *Speech across species : on the mechanistic fundamentals of vocal production and perception*. Retrieved from <https://hdl.handle.net/1887/17608>

Version: Not Applicable (or Unknown)  
License: [Leiden University Non-exclusive license](#)  
Downloaded from: <https://hdl.handle.net/1887/17608>

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

## **Nederlandse samenvatting**

Spraak bij verschillende soorten:

Mechanistische grondlagen van vocale productie en perceptie

*Dit is een vertaling van hoofdstuk 1. De literatuurverwijzingen zijn weggelaten om de leesbaarheid te verbeteren. De figuur is terug te vinden in hoofdstuk 1.*

## **Algemene inleiding**

Menselijke taal is één van de meest complexe gedragingen die wij kennen. Zij maakt het mogelijk oneindig veel ideeën te ontwikkelen en te communiceren. Hoewel het belang van het menselijke vermogen tot taal evident is, is tot dusver onduidelijk, en is er veel discussie over, hoe taal evolutionair ontstaan is en welke specifieke eigenschappen menselijke taal uniek maken.

Spraak is het primaire fysieke fenomeen waarmee taal wordt overgedragen. Een beperkt aantal betekenisloze geluiden wordt gecombineerd om een potentieel oneindige set van betekenisvolle woorden te vormen. De onderliggende productie- en perceptiemechanismen kunnen met behulp van akoestische, fysiologische, anatomische en neurobiologische methoden worden bestudeerd. Helaas kunnen dit soort studies ons echter weinig leren over de evolutie van spraak. Ook fossiele vondsten, voor zover voorhanden, leveren te weinig informatie om de evolutie van spraak in kaart te kunnen brengen. Het bestuderen van vocale communicatie bij dieren maakt het echter wel mogelijk om mechanismen op te sporen die homoloog of door convergentie zijn ontstaan, en kan daarom helpen onderliggende evolutionaire processen te identificeren. Deze vergelijkende benadering is één van de belangrijkste methoden van een nog jong wetenschapsgebied dat biolinguïstiek wordt genoemd.

### *Bevindingen van de vergelijkende benadering*

In de laatste jaren zijn in een groeiend aantal studies mogelijke overeenkomsten tussen menselijke en dierlijke vocale communicatie onderzocht. Eén van de belangrijkste eigenschappen van spraak is dat het geleerd wordt door imitatie, iets wat zeldzaam is in de ontwikkeling van vocalisatie bij andere dieren. Binnen de zoogdieren komt vocaal leren, behalve bij de mens, slechts in een beperkt aantal niet-verwante groepen voor, waaronder zeehonden, walvissen, vleermuizen en olifanten. Opmerkelijk is dat vocaal leren bij onze naaste verwanten, mensapen en andere primaten, niet voorkomt.

Behalve bij zoogdieren komt vocaal leren ook voor in drie vogelgroepen, namelijk de zangvogels, papegaaien en kolibries. Deze en andere parallellen tussen menselijke spraak en vogelvocalisatie hebben ertoe geleid dat vogelzang als het meest geschikte diermodel wordt beschouwd om onderliggende mechanismen van spraakontwikkeling, -productie en -perceptie te bestuderen.

Zowel mensen als vogels vertonen vroeg in het leven een sensitieve fase waar vocaal leren wordt gefaciliteerd. Auditieve feedback door het horen van eigen geproduceerde

vocalisaties en vocalisaties van anderen is hierbij cruciaal. Er wordt ook gedacht dat aangeboren predisposities in mensen en vogels bepalen welke geluiden worden geleerd. Sinds kort is het dankzij de ontwikkeling van nieuwe moleculaire methoden ook mogelijk om de genetische basis van vocaal gedrag te onderzoeken. Er is een genmutatie gevonden die een spraakstoornis bij mensen veroorzaakt en er bij zebrevinken toe leidt dat ze de zang van hun tutor onnauwkeurig kopiëren.

De hierboven beschreven parallellen hebben vooral betrekking op vocale ontwikkeling en leren. Ook op het gebied van vocale productie en perceptie lijken echter overeenkomsten tussen vogels en mensen te bestaan. Vogelzang en spraak worden beide door een geluidsbron geproduceerd en daarna in het spraakkanaal gefilterd, hoewel over de details hiervan bij vogels minder bekend is. Wat betreft vocale perceptie is nog onduidelijk of er speciale perceptuele capaciteiten nodig zijn voor het waarnemen en onderscheiden van spraak, en of zangvogels deze capaciteiten ook hebben. In dit proefschrift onderzoek ik zowel vocale productie- als perceptiemechanismen in (zang)vogels en vergelijk ik deze met die van mensen om erachter te komen welke overeenkomsten en verschillen er zijn.

### *Formanten en hun relevantie in vocale communicatie*

Menselijke spraak wordt gekarakteriseerd door een breed spectrum van verschillende frequenties. Stemhebbende geluiden worden in de larynx geproduceerd door het vibreren van de stembanden. Daardoor ontstaat een fundamentele frequentie met harmonische boventonen. Dit geluid wordt in het spraakkanaal, bestaande uit keel-, mond- en neusholtes, gefilterd. Afhankelijk van de positie van de articulatoren, zoals tong en lippen, worden hierbij sommige frequenties versterkt. Deze frequenties worden 'spraakkanaalresonanties' ofwel 'formanten' genoemd. Formanten zijn in een spectrogram als zwarte banden herkenbaar (Fig. 1.1). Formantpatronen worden onafhankelijk van de stembanden gevormd, en spelen vooral bij de productie en perceptie van klinkers een belangrijke rol. Het verschil tussen 'wit' en 'wet', bijvoorbeeld, is gebaseerd op verschillende formantwaarden, vooral met betrekking tot de twee laagste formanten. Deze worden veroorzaakt door resonanties van de keel- (F1) en mondholte (F2).

Tijdens de menselijke ontogenese daalt de larynx, waardoor de tong zowel horizontaal als verticaal kan bewegen. Dit maakt het mogelijk om de geometrie van het spraakkanaal op complexe manieren te moduleren, waardoor resonantiefrequenties veranderen en complexe formantenpatronen kunnen ontstaan (Fig. 1.1).

In het algemeen wordt aangenomen dat het dalen van de larynx en het verlies van luchtzakken rond de larynx een noodzakelijke voorwaarde is geweest voor de evolutie

van spraak. Het is echter niet waarschijnlijk dat spraakevolutie een veroorzakende kracht is geweest bij het ontstaan van deze anatomische configuratie, gezien het feit dat ook andere diersoorten een gedaalde larynx vertonen. Zo verlagen edelherten en damherten hun larynx tijdens hun burlen, en hebben sommige vogelsoorten verlengde luchtpijpen. In beide gevallen leidt dit tot lagere resonantiefrequenties van het spraakkanaal, waardoor deze individuen groter klinken. Het is dus waarschijnlijk dat het zakken van de larynx door seksuele selectie werd veroorzaakt, en daarmee een preadaptatie geworden is voor de evolutie van spraak. Bovendien hebben ook chimpansees een verlaagde larynx en wordt soms gespeculeerd dat het vervlakken van het gezicht belangrijker was voor het ontstaan van de karakteristieke configuratie van het spraakkanaal.

Over spraakkanaalresonanties in zangvogels is echter minder bekend. Het “spraakorgaan” van vogels, de *syrinx* genoemd, is veel gecompliceerder dan de menselijke larynx. Zangvogels bezitten twee paar vibrerende membranen die aan het eind van elke bronchus zitten, en die betrokken zijn bij geluidsproductie. Deze twee stembandsets kunnen onafhankelijk van elkaar worden gecontroleerd, waardoor de vogel met twee stemmen tegelijk kan zingen of kan afwisselen. Vanwege deze complexiteit op stemniveau verwachtte men aanvankelijk dat alle complexiteit in vogelzang door de geluidsbron wordt veroorzaakt en dat filtermechanismen zoals in menselijke spraak geen rol spelen. Recent onderzoek laat echter zien dat zang- en spraakproductie meer gemeen hebben dan oorspronkelijk werd aangenomen. Eén van de doelen van dit proefschrift is om een beter begrip te verkrijgen van potentiële articulators in vogelzang en hun invloed op het geproduceerde geluid.

Het tweede doel is om een beter begrip te verkrijgen van hoe formanten door vogels waargenomen worden. Mensen zijn erg gevoelig voor formantenpatronen en spraakvariatie. Het is echter niet duidelijk of deze sensitiviteit in samenhang met spraakproductie is geëvolueerd, en dus uniek is voor mensen, of dat deze berust op algemene eigenschappen van auditieve systemen. Sommige perceptuele fenomenen waarvan men dacht dat ze uniek waren voor de mens, zoals categorische perceptie, zijn ook bij zoogdieren en vogels gevonden. Een ander belangrijk kenmerk van spraak is dat wij woorden kunnen herkennen ondanks het feit dat deze akoestisch sterk kunnen verschillen wanneer ze door verschillende sprekers uitgesproken worden. Dit fenomeen wordt sprekernormalisatie genoemd. Het is onbekend of dit een fenomeen is dat alleen bij mensen voorkomt, en dus waarschijnlijk in samenhang met spraak is ontstaan, of dat het veroorzaakt wordt door algemene eigenschappen van het auditieve systeem.

### *Dit proefschrift*

In dit proefschrift worden vier experimenten beschreven, waarvan twee over productie- en twee over perceptiemechanismen van formanten bij vogels gaan. Ik heb vooral zebra-vinken voor mijn experimenten gebruikt omdat deze het meest bestudeerde modelsysteem zijn voor vocaal leren, en er desondanks weinig bekend is over geluidsproductie en -perceptie in deze soort. Daarnaast heb ik ook geluidsproductiemechanismen bij monniksparkieten onderzocht, omdat er indicaties voor belangrijke verschillen tussen zangvogels en papegaaien zijn.

In **hoofdstuk 2** beschrijf ik een experiment dat is uitgevoerd om potentiële articulatoren in zebra-vinken te identificeren en hun betrokkenheid bij zangproductie te evalueren. Eerst hebben wij röntgenfilms van zingende zebra-vinken gemaakt, waarvan de analyse heeft laten zien dat vooral snavelopening en het vergroten van de orofaryngeale-esofoageale holte (OEC) als articulatoren dienen. Dit resultaat is in overeenstemming met eerdere studies die positieve correlaties tussen snavelopening en frequentiepatronen in meerdere zangvogelsoorten, waaronder zebra-vinken, hebben aangetoond. Interessanter is de observatie dat zebra-vinken hun OEC sterk uitbreiden tijdens de zang. Dit is eerder in de rode kardinaal en de witkeelgors aangetoond. Deze twee soorten produceren echter allebei een redelijk eenvoudig, tonaal gezang met weinig energie in harmonische boventonen, en uitbreiding van de OEC bleek bij deze soorten te corresponderen met variatie in de fundamentele frequentie. Zebra-vinken produceren daarentegen een redelijk ingewikkelde zang met veel verschillende elementtypes. De meeste van deze elementen hebben een breed frequentiespectrum en gevarieerde amplitudepatronen. De complexiteit van zebra-vinkenzang maakt het lastig om duidelijke verbanden tussen articulatorenconfiguratie en geluidspatronen op te sporen. Desalniettemin hebben wij in het tweede deel van de studie, waarin we snavelopening en OEC variatie experimenteel hebben gemanipuleerd, gevonden dat de piekfrequentie zakt tijdens de OEC vergroting terwijl amplitude toeneemt, vooral tussen 1.5 en 4.5 kHz. Snavelopening blijkt frequenties rond de 5 kHz en daarboven te versterken. Deze resultaten tonen aan dat modulatie van het spraakkanaal, vooral door snavelopening en de uitbreiding van de OEC, een belangrijke rol speelt in het dynamisch filteren van zebra-vinkenzang, wat in ingewikkelde elementtypes met formantachtige patronen resulteert.

Papegaaien vormen een andere vogelgroep die complexe geluiden over een breed frequentiespectrum produceert; deze vogels staan bekend om hun vermogen tot het imiteren van menselijke spraak. In tegenstelling tot zangvogels hebben papegaaien een tong met veel intrinsieke spieren en een vlezig, flexibel oppervlak, die lijkt op de

menselijke tong. Om deze reden gaat men ervan uit dat de tong een veel belangrijkere rol speelt bij de geluidsproductie van papegaaien dan bij die van zangvogels. Observaties van een spraakimiterende papegaai en experimentele manipulatie van tongpositie bij monniksparkieten ondersteunen deze stelling. Tot nu toe waren echter nog geen observaties van tongbewegingen in natuurlijk communicerende papegaaien gedaan, noch was bekend welke andere articulatoren bij geluidsproductie van papegaaien betrokken zijn.

In **hoofdstuk 3** heb ik dit onderzocht door röntgenfilms van natuurlijk communicerende papegaaien te analyseren. Aan de hand van de video's konden wij drie typen articulatorische bewegingen identificeren: snavelopening, veranderingen in de tonghoogte en samentrekking van de luchtpijp. Hoewel eerdere studies het belang van de tong al hebben aangetoond, zijn in deze studies vooral effecten in de horizontale dimensie van de tongpositie gevonden terwijl de parkieten in onze studie vooral tonghoogte veranderden. Van de negen bekende vocalisatietypes die volwassen monniksparkieten produceren, produceerden onze vogels in het laboratorium er maar drie. Daarom is het mogelijk dat de horizontale tongpositie wel een belangrijke rol speelt in de vocalisaties die wij niet konden opnemen. Niettemin blijken veranderingen in de verticale tongdimensie van begroetingsroepen, die wij wel hebben opgenomen en die formantachtige patronen laten zien, belangrijker te zijn dan veranderingen in de horizontale dimensie. Interessant genoeg hebben wij ook bewijzen gevonden voor een samentrekking van de luchtpijp, terwijl een eerdere studie in zebra's geconcludeerd heeft dat veranderingen in de lengte van de luchtpijp te klein zijn om invloed te hebben op het geluid in deze soort. Verder vonden we positieve correlaties tussen amplitude in begroetingsroepen en schettergeluiden voor snavelopening, tonghoogte en luchtpijpcontractie in sommige van de parkieten. Omdat modulaties in de fundamentele frequentie (F0) redelijk snel zijn in monniksparkietroepen, maar articulatorische bewegingen relatief langzaam, is het waarschijnlijk dat veranderingen in F0 door de geluidsbron worden veroorzaakt. Formantpatronen zoals in begroetingsroepen ontstaan waarschijnlijk wel door de spraakkanaalfilter en worden in dat geval gemoduleerd door de werking van de articulatoren. Helaas was het niet mogelijk om duidelijke relaties tussen formantpatronen en articulatie te bepalen omdat de precieze eigenschappen van de geluidsbron grotendeels onbekend zijn. Derhalve moeten toekomstige studies de precieze aard van deze relaties bepalen, en zijn er meer gegevens nodig over de anatomie en fysiologie van het vocale systeem om een betrouwbaar model van geluidsproductie in deze vogels op te stellen.

In de tweede helft van dit proefschrift wordt formantperceptie bij vogels vergeleken met die bij mensen. Hoewel er verschillen zijn in vocale communicatie bij mensen, zangvogels en papegaaien, maken alle drie groepen gebruik van actieve geluidfiltermechanismen waardoor ze veel verschillende geluiden kunnen produceren. Dit leidt tot de vraag of de onderliggende mechanismen van formantperceptie ook vergelijkbaar zijn tussen mensen en vogels. Als dat het geval is, dan is het niet noodzakelijk om ervan uit te gaan dat speciale mechanismen voor formantperceptie in mensen zijn geëvolueerd door coëvolutie van spraakproductie en -perceptie. In plaats daarvan zouden algemene verwerkingsmechanismen van het auditieve systeem voldoende zijn om menselijke spraakgeluiden te onderscheiden.

**Hoofdstuk 4** beschrijft een studie naar sprekernormalisatie bij zebra-vinken. Hiervoor zijn natuurlijk geproduceerde menselijke woorden gebruikt, ingesproken door jongvolwassen Nederlandstalige sprekers. Eén van de belangrijkste aspecten van menselijke spraak is ons vermogen woorden te kunnen herkennen onafhankelijk van de spreker en ondanks veel variatie tussen sprekers onderling. Taalwetenschappers gaan ervan uit dat dit een gevolg is van het menselijke vermogen voor extrinsieke en intrinsieke sprekernormalisatie. Intrinsieke sprekernormalisatie verklaart waarom geluiden die als dezelfde fonemen worden waargenomen verschillende akoestische realisaties kunnen hebben door ervan uit te gaan dat elk spraakmonster door een normaliseringstransformatie wordt gecategoriseerd. Tegelijkertijd is bekend dat er een sprekereffect bestaat, wat in het begin spraakdiscriminatie tussen sprekers moeilijker maakt. Dit probleem wordt opgelost via een leerproces waarin een referentiefraam van verschillende spraakgeluidmonsters ontstaat. Wij hebben middels operante conditionering acht zebra-vinken getraind om twee woorden, 'wit' en 'wet', te onderscheiden en deze later te generaliseren naar onbekende sprekers van (1) hetzelfde geslacht en (2) het andere geslacht. De twee woorden verschillen voornamelijk in hun formantenpatroon. Alle acht vogels waren in staat de woorden te onderscheiden en te categoriseren onafhankelijk van de sprekers. Onze analyse heeft laten zien dat ze dit onderscheid aan de hand van de formantenpatronen hebben gemaakt. Bovendien gebruikten de vogels, net als mensen, hiervoor een combinatie van extrinsieke en intrinsieke sprekernormalisatie. Dit resultaat impliceert dat formantperceptie ofwel wijdverspreid is in het dierenrijk, ofwel convergent is ontstaan in mensen en vogels.

Het laatste hoofdstuk van dit proefschrift, **hoofdstuk 5**, beschrijft een directe vergelijking van akoestische parameterweging in klinkerperceptie tussen zebra-vinken en Nederlandse volwassenen. In het verleden is aangetoond dat zowel Canadees-Engelse als Zweedse baby's tussen drie en twaalf maanden oud gevoeliger zijn voor lage



frequenties (F1) wanneer ze woorden onderscheiden. Dit is verrassend omdat algemeen aangenomen wordt dat de taalomgeving een grote invloed heeft op klinkerperceptie vanaf een leeftijd van zes maanden. Dit kan ofwel een indicatie voor een algemene menselijke gevoeligheid voor lage frequenties in klinkerperceptie zijn, ofwel te maken hebben met een rijpingsproces van het auditieve systeem. Maar de vraag of deze gevoeligheden verbonden zijn aan spraakperceptie en dus een uniek menselijke eigenschap zijn, of juist een algemeen karakteristiek zijn voor auditorische perceptie, is nog niet onderzocht. In een eerste poging deze vraag te beantwoorden hebben wij zebra's en Nederlandse volwassenen mensen op dezelfde manier getest. Beide werden in een 'Go/NoGo'-procedure blootgesteld aan twee gesynthetiseerde woorden (dit en doet of dut en diet) die alleen in hun klinkers van elkaar verschilden en moesten leren tussen die twee woorden onderscheid te maken. De klinkers hadden verschillende F1 en F2 frequenties. In de volgende stap hebben wij naast de eerste twee woorden twee testwoorden gepresenteerd. Deze testwoorden werden nooit beloond en de reacties op de testwoorden liet zien hoe mensen en vogels deze hebben waargenomen. Eén testwoord had dezelfde F1 frequentie als het eerste originele woord en de dezelfde F2 frequentie als het tweede originele woord en andersom voor het tweede testwoord. De reacties van vogels en mensen op de testwoorden waren opvallend overeenkomstig: beide vertoonden ze een neiging om categorieën te maken gebaseerd op hoge frequenties (F2). Dit is precies andersom dan wat baby's doen en toont ten eerste aan dat akoestische parameterweging niet uniek is voor mensen en afhankelijk van spraakperceptie en ten tweede dat de maturatie van het auditieve systeem waarschijnlijk een rol zal spelen bij het ontstaan van een dergelijke parameterweging.

### *Discussie en conclusie*

In dit proefschrift heb ik laten zien dat zebra's en monniksparkieten, evenals mensen, gebruik maken van verschillende articulatoren om het geluid dat in de syrxin wordt geproduceerd te modificeren. Terwijl in zangvogels vooral snavelopening en OEC uitbreiding een belangrijke rol spelen blijken papegaaien voornamelijk de tong te gebruiken om het geluid te filteren. Dit is vergelijkbaar met menselijke spraakproductie en waarschijnlijk één van de redenen waarom papegaaien spraak zo goed kunnen imiteren. Gebaseerd op deze observaties kan geconcludeerd worden dat de geluidsproductiemechanismen in vogels en mensen meer gemeen hebben dan oorspronkelijk werd vermoed, wat convergentie in evolutionaire patronen suggereert. Mogelijk zijn sommige articulatoren die nu worden gebruikt voor geluidsverandering

oorspronkelijk als delen van het voedselverwerkingssysteem ontstaan. In dat geval zijn ecologische adaptaties voor verschillende voedselsoorten de drijvende kracht achter de evolutie van deze structuren geweest, en zijn deze structuren later door het communicatiesysteem gebruikt om vocale variatie te verhogen.

Met betrekking tot spraakperceptie heb ik laten zien dat zebra-vinken, net als mensen, formantpatronen gebruiken om op elkaar lijkende woorden te onderscheiden en tegelijkertijd extrinsieke en intrinsieke sprekernelnormalisatie toepassen om woorden onafhankelijk van de spreker te categoriseren. Dit heeft belangrijke implicaties voor de evolutie van formantperceptie. Zoals al eerder werd gespeculeerd lijkt formantperceptie in veel verschillende diersoorten te zijn geëvolueerd om informatie over geslacht, grootte en leeftijd van een individu beschikbaar te maken. Tijdens de evolutie van spraak zou deze capaciteit vervolgens kunnen hebben geleid tot een communicatiesysteem dat gebruik maakt van formanten om informatie te coderen. Het feit dat volwassen mensen en zebra-vinken dezelfde parameterweging in klinkerperceptie vertonen versterkt deze hypothese. Beide soorten zijn gevoeliger voor F2 frequenties, welke in het meest sensitieve bereik van het auditieve systeem vallen, bij het categoriseren van meerduidige klinkers. Menselijke baby's aan de andere kant blijken gevoeliger te zijn voor frequenties die sterker worden betoond, namelijk F1.

Samenvattend heb ik laten zien dat (zang)vogels in staat zijn tot formantenproductie en -perceptie en dat de onderliggende mechanismen meer overeenkomen in mensen en vogels dan eerder werd aangenomen. Zowel zangvogels als papegaaien kunnen als modelorganismen worden gebruikt om meer over de selectiedrukken, die tot zulk ingewikkelde communicatiesystemen hebben geleid, te weten te komen. Nu sommige van de onderliggende mechanismen voor formantenproductie en -perceptie zijn geïdentificeerd kan in toekomstige studies begonnen worden met het onderzoek naar de rol die formanten in soorteigen vogelvocalisaties spelen, en met het ontwikkelen van gedetailleerde modellen voor geluidsproductie en -perceptie.