



Universiteit
Leiden
The Netherlands

Sound of mind: electrophysiological and behavioural evidence for the role of context, variation and informativity in human speech processing
Nixon, J.S.

Citation

Nixon, J. S. (2014, October 14). *Sound of mind: electrophysiological and behavioural evidence for the role of context, variation and informativity in human speech processing*. Retrieved from <https://hdl.handle.net/1887/29299>

Version: Corrected Publisher's Version

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

Downloaded from: <https://hdl.handle.net/1887/29299>

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

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/29299> holds various files of this Leiden University dissertation.

Author: Nixon, Jessie Sophia

Title: Sound of mind: electrophysiological and behavioural evidence for the role of context, variation and informativity in human speech processing

Issue Date: 2014-10-14

Nederlandse samenvatting

In dit proefschrift wordt bewijsmateriaal gepresenteerd dat door middel van verscheidene empirische en statistische methoden is verkregen om te onderzoeken hoe gezonde volwassen moedertaalsprekers de intrinsiek ruizige akoestische informatie verwerken waar de gesproken communicatie uit bestaat. Daar waar psycholinguïstische modellen (met name die van spraakproductie en van hardop lezen) fonologische verwerking doorgaans in termen van foneemsequenties hebben gemeten, toont dit proefschrift verschillende soorten sub-fonemische verwerking aan zoals die in verscheidene experimentele taken plaatsvindt. Eén van de hoofdonderwerpen van dit proefschrift is de manier waarop sprekers en luisteraars fonetische variatie verwerken. Belangrijk is dat een onderscheid wordt gemaakt tussen de reguliere informatieve variatie, zoals bij allofonische varianten, enerzijds, en anderzijds onvoorspelbare ruis, die de informativiteit van akoestische cues beperkt. Verder laten we zien dat zelfs in een taal waarbij geen informatie over spraakcontrasten in de spelling wordt weerspiegeld, bij zowel de overte productie als de visuele verwerking van geschreven woorden sprake is van verwerking van contrastieve spraakklankcategorïen.

In hoofdstuk 2 hebben we gebruik gemaakt van de plaatje-woord-interferentietask om vast te stellen hoe toonvarianten worden verwerkt bij de spraakproductie van het Mandarijn. In het Peking-Mandarijn is het zo dat als Toon 3 wordt gevolgd door een karakter met nog eens Toon 3 (*sandhi* van toon 3), de eerste toon uiteindelijk als Toon 2 klinkt. Dit aspect van

het tonale systeem behelst dat sandhi-woorden fonologisch zowel gerelateerd zijn aan Toon 3 als aan Toon 2. Dat stelde ons er toe in staat om twee soorten fonologische informatie te manipuleren: de fonologische tooncategorie en de akoestische tooncontour. In hoofdstuk 2 zijn we ingegaan op de vraag welke van deze twee types van fonologische verbanden belangrijk is bij spraakproductie. Hoewel er een overvloed aan psycholinguïstisch onderzoek bestaat dat aantoont dat verwerking wordt ondersteund door de voor-activatie van congruente fonologische informatie, is de precieze aard van deze fonologische informatie nog niet opgehelderd. Gaat met spraakproductie het ophalen van de spraakklankcategorie gemoeid? Of gaat het om de activering van de eigenlijke akoestische realisatie? Hoofdstuk 2 laat zien dat spraakproductie een kwestie is van fonologische verwerking op meerdere niveaus (zie ook Nixon et al., 2014). De productie van afbeeldingsnamen in sandhi-conditie voltrok zich significant sneller wanneer de afleider en de doelafbeelding overeenkwamen in tooncategorie maar verschillende overte realisaties hadden *en ook* wanneer doelafbeelding en afleider in overte realisatie overeenkwamen maar onder andere tooncategorieën vielen (vergeleken met controleafleiders, die in toneem noch in contour met de doelafbeelding overeenkwamen). Interessant genoeg waren er verschillen in het tijdsverloop van de effecten. De tooncontour versnelde de productie zowel wanneer die gelijktijdig met de doelafbeelding werd gepresenteerd als wanneer die er 83 milliseconden achteraan kwam. De tooncategorie, echter, versnelde de productie alleen in het geval van gelijktijdige presentatie. Dit doet vermoeden dat de spraakcategorie in een vroeg stadium wordt geactiveerd (mogelijk op het moment van het ophalen van het woord), maar dat de daadwerkelijke akoestische realisatie in een later stadium van de verwerking van belang wordt (bijvoorbeeld bij de spraakvoorbereiding). Er zijn twee verklaringen denkbaar voor deze resultaten. Het kan zijn dat de gerealiseerde contour langer actief blijft dan de tooncategorie gedurende overte productie. Een andere mogelijkheid is dat de contour en de categorie beide actief blijven, terwijl de taak overgaat van lexicale retrieval naar de voorbereiding van de articulatie, maar dat alleen het samenkomen van articulatie en akoestische realisatie de productie ten goede komt. Er is meer onderzoek nodig om te be-

palen welk van deze twee mogelijkheden de voorkeur verdient. Een tweede vraag die zich opdringt naar aanleiding van de resultaten van Experiment 1 is of het verwerken van context-specifieke tooncontouren automatisch verloopt of dat het zich pas voordoet bij overte productie van spraakvarianten, in voorbereiding op de articulatie. Om te achterhalen wat voor soort fonologische informatie wordt geactiveerd wanneer sandhi-woorden worden verwerkt maar niet overt worden geproduceerd, zijn in Experiment 2 de prime- en doelcondities omgewisseld.

In Experiment 2 bestonden de primes uit visueel gepresenteerde aan sandhi onderhevige toonvarianten of afleiders als controlewoorden. Als doelwoorden dienden plaatjes met Toon 3 of Toon 2. Net zoals in Experiment 1 vond er versnelling plaats van zowel de doelwoorden van de tooncategorie als de doelwoorden van de tooncontour met sandhi-afleiders ten opzichte van de controlewoorden. Hieruit valt op te maken dat de verwerking van de context-specifieke verwezenlijking een automatisch proces is en niet simpelweg het gevolg van articulatievoorbereiding. Interessant genoeg verschilde het tijdsverloop van activatie met dat van Experiment 1. Wanneer de toonvarianten visueel werden verwerkt in plaats van overt geproduceerd was het de tooncategorie die werd versneld, zowel in het geval van de gelijktijdige (0 ms) als in het geval van de verlate (83 ms) aanbieding van het afleiderwoord. Versnelling van de tooncontour deed zich alleen voor bij de gelijktijdige aanbieding. Dit duidt erop dat de activatie van de context-specifieke contour gedurende de visuele verwerking van toon-sandhi varianten tijd kost in verhouding tot de spraakcategorie. Bij gelijktijdige presentatie van het afleiderwoord met het doelwoord wordt de contour op tijd geactiveerd om productie te kunnen versnellen. Wanneer de aanbieding echter later plaatsvindt kan het de productie van de doelcontour niet meer versnellen.

In hoofdstuk 3 hebben we met behulp van een onderzoek met event related potentials (ERP) de verwerking van sub-fonemische segmentele informatie in een typologisch andere taal bekeken. Er is erg weinig bekend over sub-fonemische verwerking bij het hardop lezen. In dit onderzoek hebben we ons bediend van gemaskeerde priming om na te gaan of en wanneer er fonetische

kenmerken worden verwerkt bij het hardop lezen in het Nederlands. De segmenten van de Nederlandse klankparen t-d en p-b hebben onderling dezelfde articulatieplaats (respectievelijk alveolair en bilabiaal), terwijl de segmenten van de paren t-p en d-b dezelfde stemhebbendheid hebben (respectievelijk stemloos en stemhebbend). We hebben EEG en reactietijden gemeten bij deelnemers die hardop bestaande Nederlandse woorden (bijv. *huid*) oplazen. Elk doelwoord werd voorafgegaan door een korte aanbidding van een gemaskeerde nonwoord-prime waarvan de laatste klank ofwel in stemhebbendheid overeenkwam (*huib*), dan wel in articulatieplaats (*huit*) of noch in stemhebbendheid noch in articulatieplaats (de controleconditie, *huip*). De reactietijden, geanalyseerd met linear mixed effects-regressiemodellen), waren significant korter bij primes die in stemhebbendheid overeenkwamen dan bij controleprimes. In lijn hiermee was de verminderde early negativity in de conditie met overeenkomstige stemhebbendheid vergeleken met de controleconditie. Deze resultaten wijzen op een snelle verwerking van sub-fonemische stemhebbendheid bij hardop lezen in het Nederlands. Deze versnelling kan niet het gevolg zijn van spellingsinformatie of informatie op foneemniveau aangezien de verschillen tussen het aantal letters en de fonemen van prime en doelwoord identiek waren in de condities met overeenkomstige en niet-overeenkomstige kenmerken: alle primes verschilden op precies één foneem en één letter van de doelwoorden. Alleen op het sub-fonemisch kenmerkenniveau was de overlap groter in de overeenkomstconditie, in vergelijking met de controleconditie.

Deze resultaten zeggen ook iets over de manier waarop sub-fonemische kenmerken gerepresenteerd worden. Nederlandse stemhebbende plosieven kennen door ‘final devoicing’ (stemloos worden van een finale stemhebbende klank) twee realisaties. Stemhebbende (bijv. /d/ en /b/) en stemloze (/t/ en /p/) plosieven worden in woord-initiale positie onderscheiden door voice onset time (VOT). Maar in woordfinale positie zitten de VOT-waardes van stemhebbende en stemloze plosieven heel dicht bij elkaar. Op het niveau van de articulatie zijn de beide primetypes daarom stemloos. De versnelling van verwerking in de conditie met overeenkomstige stemhebbendheid, echter, wijst erop dat het stemhebbendheidsonderscheid behouden is en wordt verwerkt

tijdens het hardop lezen. Dit duidt erop dat stemloze en stemloos geworden plosieven ondanks hun vergelijkbare overte realisatie als aparte categorieën worden verwerkt. Dit is in lijn met de data uit hoofdstuk 2 waaruit bleek dat gedurende spraakproductie en visuele woordverwerking spraakvarianten op het niveau van categorieën en in specifieke contexten een rol speelden. Ondanks het feit dat in dit experiment niet specifiek is getest of activatie van een context-specifieke variant plaatsvond, past het wel bij het scenario dat er zich ook op het kenmerkenniveau meerlagige verwerking voordoet.

Hoofdstuk 4 beschrijft een ERP-experiment waarin de vraag centraal staat hoe de fonologisch verwerking van hardop lezen afhangt van de context. We hebben van moedertaalsprekers van het Peking-Mandarijn ERP's en reactietijden geregistreerd terwijl ze hardop woorden oplazen die een lengte hadden van twee karakters en die vooraf werden gegaan door gemaskeerde primes. Het eerste karakter van alle kritieke primes was een karakter met Toon 3 en de beginkarakters waren identiek voor alle primingcondities. Alleen het tweede karakter verschilde tussen condities. De fonetische context die gecreëerd werd door de toon van het tweede karakter maakte uit of het eerste karakter een stijgende (sandhi) contour of de canonieke, lage contour had. De kritieke doelwoorden waren de woorden die begonnen met Toon 2, waarvan de contour ook stijgend is. Alle primes kwamen dus niet overeen in de tooncategorie, maar de sandhi-primes kwamen wel overeen in de tooncontour. Het best passende generalised additive mixed model (GAMM) dat we hebben toegepast had een volledige random effects-structuur en significante predictoren van conditie, trial en doelwoordfrequentie over de tijd, alsmede de interacties daartussen. Uit de predictor van condities over de tijd bleek dat er significante verschillen waren in ERP-amplitude binnen het tijdsvenster van 300-350 ms. Deze resultaten duiden erop dat de top-downverwerking van informatie vanuit de fonetische context de activatie van alternatieve realisaties van allofonische varianten beïnvloedt. Identieke beginkarakters (tussen condities) leidden tot verschillende amplitudes in de EEG afhankelijk van de tooncontext van het volgende karakter. Bovendien toonde de interactie tussen conditie en doelwoordfrequentie over de tijd aan dat de congruentie van de contour het ster-

kste was voor laagfrequente doelwoorden. Dit komt waarschijnlijk door de lagere resting state-activatie van laagfrequente woorden, die maakt dat ze ontvankelijker zijn voor primingeffecten. Gezien de bevinding dat doelwoordfrequentie in samenspel met andere predictoren een (deel van de) verklaring vormt voor de gevonden neurale activiteit heeft het meenemen van iteminformatie in ERP-onderzoek een toegevoegde waarde.

Waar de hoofdstukken tot nu toe ingingen op de verwerking van spraakcontrasten, contrastieve kenmerken en reguliere variatie, bogen we ons in Hoofdstuk 5 over de effecten van de hoeveelheid akoestische ruis op het verwerken van spraakcontrasten. We hebben de oogbewegingen gemeten van moedertaalsprekers van het Kantonees die een taak uitvoerden waarvoor ze naar plaatjes keken en naar akoestische stimuli luisterden. De plaatjes correspondeerden met geaspireerde en niet-geaspireerde woordparen en de akoestische stimuli bevatten ofwel een relatief grote hoeveelheid variatie (de conditie met de *wijde* distributie) ofwel een relatief kleine hoeveelheid variatie (de conditie met de *nauwe* distributie). Door gebruik te maken van generalised additive mixed modelling (GAMM) konden we complexe, niet-lineaire effecten en interacties modelleren zonder continue variabelen, zoals tijd en frequentie, discreet te hoeven maken. Uit het best passende model kwam een duidelijke vorm van de distributie naar voren in de conditie met de nauwe distributie (weinig variatie). Het kijkgedrag verschilde tussen gemiddeldes, grenzen en periferieën van categorieën. In de wijde conditie (veel variatie) was de verdeling platter, met name in het laatste stuk van de trial. Het effect van VOT was, met andere woorden, zwak in de wijde conditie, vooral na 600 ms, waar de verdeling erg plat was voor alle VOT-waarden. Wij maken uit deze resultaten op dat tijdens het online waarnemen van spraakcontrasten eerdere ervaring met de verdeling van een akoestische cue de mate waarin die cue wordt gebruikt om een talige uitkomst te voorspellen beïnvloedt. Anders gezegd zijn de subtiele verschillen in de mate van variatie van een bepaalde akoestische cue van belang voor de manier waarop die wordt waargenomen en gebruikt om spraakcontrasten van elkaar te onderscheiden.

Samengevat biedt dit proefschrift nieuwe inzichten in de manier waarop mensen contrastieve en niet-contrastieve akoestische in-

formatie verwerken bij verscheidene taalverwerkingstaken. Het lijkt erop dat sprekers verschillende soorten regelmatigigheden uit het spraaksignaal opvangen, zoals contrastieve spraakcategorieën, reguliere allofoonvarianten en sub-fonemische kenmerk-informatie. Deze regelmatigigheden spelen in het taalsysteem een rol bij overte spraakproductie, visuele verwerking van geschreven woorden die niet overert worden geproduceerd en hardop lezen. Verwerking gaat echter ook gepaard met de activatie van een context-specifieke realisatie. Activatie wordt beïnvloed door top-downverwerking van informatie over de fonetische context. Tot slot hangt de mate waarin akoestische informatie als cue houvast biedt aan het onderscheiden van klankcontrasten bij spraakperceptie af van de informativiteit—dat wil zeggen, de vorm van de statistische verdeling—van de akoestische cues.

