
Een akoestisch-perceptief onderzoek naar de Maastrichtse tonen

Charlotte Gooskens

Katholieke Universiteit Nijmegen

Toni Rietveld

In this article an investigation is presented which was set up in order to describe the acoustical properties of the two tones, *stoottoon* (hard tone) and *sleeptoon* (slurred tone), in the dialect of Maastricht. Minimal pairs with the two tones were placed in different linguistic contexts and pronounced by a native speaker of the Maastricht dialect. The tonal differences were described in terms of segment duration and pitch movements. Vowel duration turned out to be significantly longer in the words with a *sleeptoon* than in the words with a *stoottoon*. Furthermore, there is a clear difference between the pitch movements in the two tones, the *sleeptoon* having a falling-rising pattern and the *stoottoon* a rising-falling pattern during the vowel. On the basis of the results, 'standardized contours' were made for different linguistic contexts.

In order to check the correctness of the 'standardized contours' and in order to get an impression of the role which duration and pitch contour play in the perception of the tones, a perception experiment was carried out. Sentences were manipulated in such a way that these two parameters were varied systematically in sentences with different linguistic contexts. For each of the different versions, native speakers of the Maastricht dialect were asked to indicate which of the two tones they heard. The results showed that at least for one of the two minimal pairs the 'standardized contours' which had been developed can be considered correct. Furthermore, it could be concluded that listeners use both duration and pitch movements to distinguish between the two tones. The relative contribution of the two parameters may be different for different minimal pairs, depending on the duration of the segments and phonetic contexts.

1 Introductie

Het onderzoek dat hier wordt gepresenteerd sluit aan bij eerder onderzoek dat is gedaan naar toononderscheidingen in Limburgse dialecten (zie bijvoorbeeld De Bot, Cox & Weltens 1990 en Verhoeven 1992). Het maakt deel uit van een onderzoek naar intonatie en de rol daarvan bij de karakterisering van dialecten. Een van de te onderzoeken dialecten is het Maastrichts. Dit dialect kent net als andere Limburgse (en Nederduitse) dialecten twee lexicale tonen, namelijk de *stoottoon* en de *sleeptoon*. Elk Maastrichts inhoudswoord heeft in principe een van beide tonen en er zijn woorden die alleen door middel van deze tonen van elkaar worden onderscheiden (minimale paren). Een voorbeeld van een minimaal paar is *bâl* ('dansfeest') met een *stoottoon* en *bâl* ('rond voorwerp') met een *sleeptoon*. Dit *sleeptoon/stoottoon*-onderscheid wordt vaak als een van de meest karakteris-

de twee tonen in het Maastrichts. Deze twee akoestische parameters worden in de literatuur over het algemeen als de belangrijkste parameters voor het toononderscheid aangegeven. Tot nu toe zijn er slechts twee onderzoeken gedaan die specifiek betrekking hadden op het Maastrichts. Het ene is het onderzoek van De Bot et al. (1990). Op basis van hun resultaten concludeerden ze dat duur weinig bijdraagt aan de perceptie van het tooncontrast. Ze vinden wel een duidelijk verschil in toonhoogteverloop tussen de twee tonen. Het tweede is het impressionistische onderzoek van Van Buuren (1991). Deze neemt uitsluitend een duurcontrast waar en geen verschil in toonhoogteverloop tussen de twee tonen. De twee Maastrichtse onderzoeken spreken elkaar dus tegen en wij zullen in ons onderzoek een poging ondernemen om een duidelijker beeld te krijgen van de rol die duur en toonhoogteverloop spelen bij het toononderscheid in het Maastrichts. In dit onderzoek maken wij in termen van deze twee parameters eerst een akoestische beschrijving van de tonen in verschillende linguïstische contexten. Op basis hiervan ontwikkelen we 'standaardcontouren' voor de twee tonen in deze verschillende contexten (paragraaf 3).

Om het relatieve belang van de verschillende akoestische parameters voor de perceptie van de tonen te kunnen vaststellen is het van belang dat de verschillende parameters onafhankelijk van elkaar worden gevarieerd en getoetst. In het tweede deel van dit onderzoek manipuleren we de toonhoogteverlopen en duren van de tonen in een aantal zinnen onafhankelijk van elkaar. Deze verschillende versies bieden we aan aan Maastrichtse luisteraars om het relatieve belang van duur en toonhoogteverloop voor de perceptie van de tonen vast te kunnen stellen (paragraaf 4). De resultaten van de twee delen van het onderzoek worden besproken in paragraaf 5 en ten slotte wordt er in paragraaf 6 een aantal algemene conclusies getrokken. We beginnen in paragraaf 2 met het bespreken van het gebruikte materiaal.

2 Materiaal

2.1 Zinnen

Uit het boven besproken onderzoek naar tonen en intonatie in de Limburgse dialecten volgt dat de volgende linguïstische factoren relevant kunnen zijn voor een onderzoek naar de Maastrichtse tonen:

1. positie in het woord (finaal versus niet-finaal);
2. positie in de zin (finaal versus niet-finaal);
3. accent (geaccentueerd versus ongeaccentueerd);
4. zinstype (vraagwoordvraag, inversievraag, bewarend).

Het doel van het beschrijvende deel van het onderzoek was de invloed van deze vier linguïstische factoren op de fonetische realisatie van de tonen te onderzoeken. De testzinnen werden zo gekozen dat alle vier de factoren systematisch konden worden gevarieerd. De tonen zelf werden gevarieerd door woorden te kiezen die slechts in toon verschillen (minimale paren). Het voordeel van het gebruik van

ster was ook aanwezig tijdens de opnamen en alle communicatie ging via haar om te zorgen dat de sfeer zo Maastrichtse mogelijk werd gehouden.

De opnamen vonden plaats in een geluidsarme studio. De sprekers werd gevraagd iedere zin zo vaak uit te spreken totdat zij er zelf tevreden over waren.

2.3 *Keuze van spreker*

In het vervolg van het onderzoek wilden we spraakmateriaal gebruiken van de spreker met de meest typische Maastrichtse dialectuitspraak die tevens het duidelijkste onderscheid maakte tussen woorden met een stoottoon en woorden met een sleптоon. Om een keus te maken tussen de vier sprekers werd een perceptie-experiment uitgevoerd. Van iedere spreker werden 16 zinnen, verdeeld over alle zinstypen en doelwoorden in geaccentueerde positie, in random volgorde op een band gezet. Zes sprekers van het Maastrichtse werd gevraagd te beoordelen hoe Maastrichtse ieder zin klonk op een schaal van 1 (minst Maastrichtse) tot 7 (meest Maastrichtse). Hun werd ook gevraagd de juiste betekenis van het doelwoord aan te geven door een van twee Standaard-Nederlandse vertalingen te omcirkelen. Resultaten van een Tukey-post-hoc-analyse lieten zien dat spreker 3 significant 'Maastrichtser' werd gevonden dan spreker 4 in alle zinnen, $F(3,18) = 3.66$ ($p < .05$). Spreker 3 verschilde ook significant van de andere drie sprekers in zinnen met alleen sleптоonwoorden, $F(3,18) = 6.83$ ($p < .01$), maar niet in zinnen met stoottoonwoorden. Verder maakte spreker 3 een duidelijker toononderscheid dan de andere sprekers, in ieder geval wat de geaccentueerde doelwoorden betreft. De beoordelaars herkenden 80% van zijn geaccentueerde doelwoorden correct. Bij de andere sprekers was dat 58%, 55% en 65%. In het onderzoek van De Bot et al. zijn de correctscores net zo slecht (60 à 65%). Zij noemen als mogelijke oorzaak dat het sleптоon/stoottoon-onderscheid misschien een verliesgevoelig dialectkenmerk zou zijn. Op basis van deze resultaten werd besloten de opnamen van spreker 3 te gebruiken voor het verdere onderzoek. Alle 120 zinnen van spreker 3 werden in de computer opgeslagen (10 kHz, 12 bits, 4,8 kHz LP) en onderworpen aan een robuuste formantextractieprocedure volgens het Split-Levison algoritme 256 punten venster en 10 punten vensterverschuiving (Willems 1987). Vervolgens werden de toonhoogtecontouren bepaald.

3 **Fonetische beschrijving van de tonen**

Het doel van het tweede deel van het onderzoek was een beschrijving te geven van de tonen in verschillende linguïstische contexten in termen van duur en toonhoogtecontour. Op basis van deze beschrijving ontwikkelden we standaardcontouren voor ieder van de relevante linguïstische factoren (3.2).

3.1 *Duur*

Eerst werden alle lettergrepen met de te onderzoeken tonen gesegmenteerd in klinkers, voorafgaande medeklinkers en volgende medeklinkers. De segmentatie werd uitgevoerd zowel op het gehoor als op basis van visuele inspectie van de

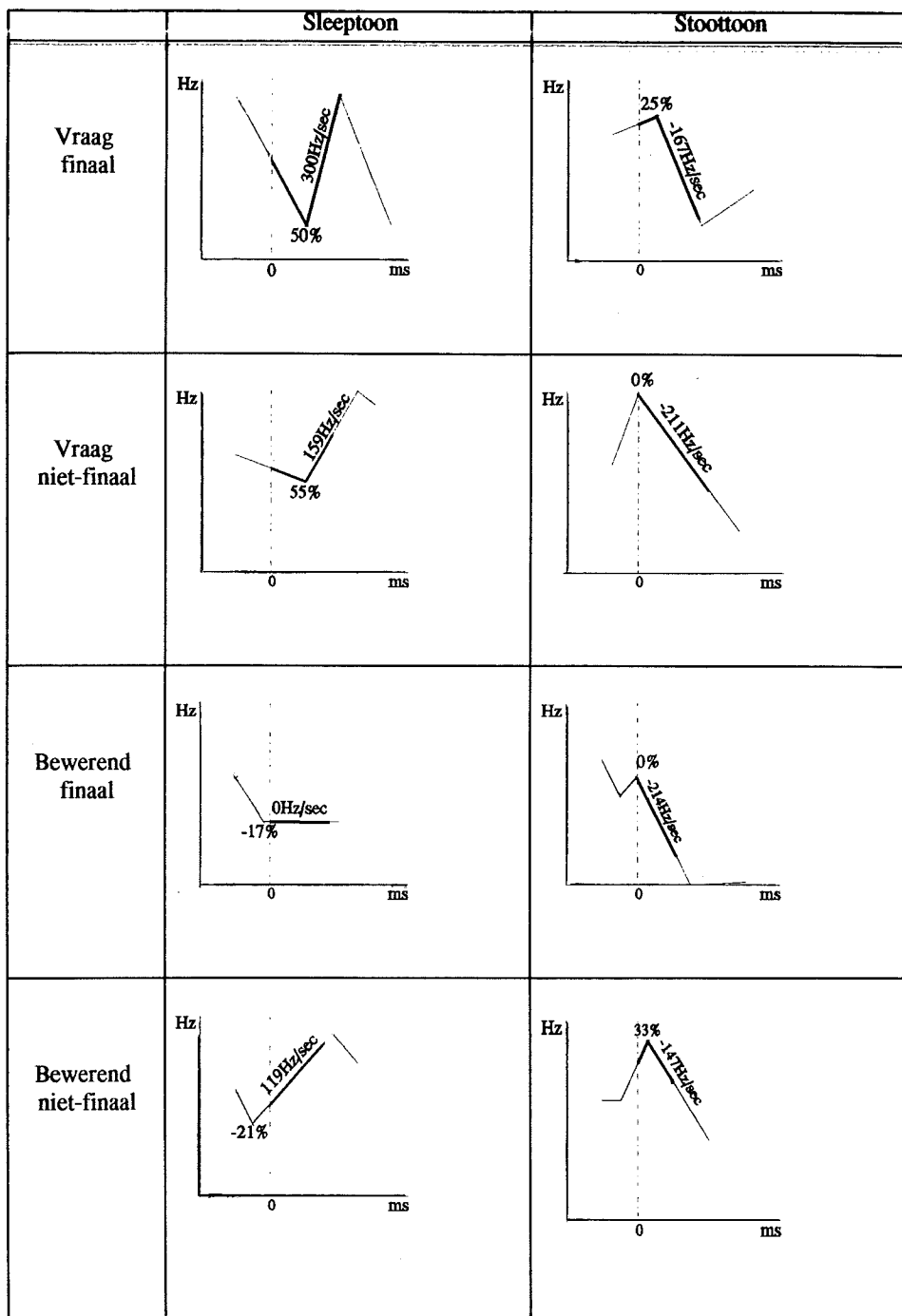
aangezien de duur van de klinkers een ondergrens kent (Allen, Hunnicut & Klatt 1987).

3.2 *Toonhoogtecontour*

De ruwe grondfrequentie (F0) werd per 10 ms bepaald volgens de methode van subharmonische sommatie (Hermes 1988). Er werd besloten om slechts de geaccentueerde tonen te beschrijven omdat de toonhoogtebewegingen op ongeaccentueerde woorden soms moeilijk herkenbaar en te beschrijven zijn. In de meeste gevallen waren er geen toonhoogtebewegingen waar te nemen op de ongeaccentueerde lettergrepen, zodat luisteraars waarschijnlijk uitsluitend met behulp van duurverschillen de tonen kunnen onderscheiden. Alle zinnen met geaccentueerde doelwoorden werden gestileerd volgens de IPO-methode ('t Hart, Collier & Cohen 1990). Het kenmerkende van deze methode is dat de toonhoogtecontour zo wordt gemanipuleerd dat hij bestaat uit zo weinig mogelijk, rechte lijnstukken (in een log-frequentie bij lineaire tijdplot). Er mag geen hoorbaar verschil zijn tussen de intonatie van de zin met de oorspronkelijke toonhoogtecontour en de intonatie van de gestileerde versie. Het voordeel van een gestileerde contour is dat men deze makkelijk kan beschrijven en vergelijken met andere gestileerde contouren.

De procedure die werd gevolgd bij de toonhoogtemetingen wordt geïllustreerd door middel van het schematisch overzicht in figuur 1. Op alle keerpunten van de toonhoogtecontour (de punten a en b in figuur 1) van de geaccentueerde lettergrepen werd de toonhoogte gemeten in Hertz. De plaats van deze keerpunten werd gemeten in milliseconden (ms) vanaf het begin van de lettergreep. Vervolgens werd de duur van iedere toonhoogtebeweging bepaald. Ten slotte werden de gemiddelden van de verschillende metingen uitgerekend voor beide tonen in ieder zinstype. Het keerpunt a in de klinker werd beschreven aan de hand van de plaats in de klinker in procenten vanaf de klinkerinzet. Als dit punt bijvoorbeeld na 30 ms werd bereikt in een klinker van 60 ms was het percentage 50. Doordat de knikpunten qua timing gedefinieerd zijn in relatieve zin (% van de klinkerduur) wordt er vanuit gegaan dat de tonen een vaste timing hebben in de lettergreep (dat wil zeggen samendrukbaar zijn en onafhankelijk van duur). Er is echter geen a priori noodzaak waarom dit zo zou zijn (zie bijvoorbeeld Caspers & Van Heuven 1993). Een voordeel van onze aanpak is dat er bij duurmanipulaties in ieder geval goed gefundeerde F0-bewegingen zijn. De gemiddelde hellingen van de toonhoogtebewegingen na deze punten werden beschreven in Hz/sec. Rietveld & Gussenhoven (1985) hebben aangetoond dat bij een beperking tot mannelijke sprekers een beschrijving in Hertz een meer adequate weergave geeft van hetgeen luisteraars horen dan een beschrijving in semitonen.

Er bleek geen aantoonbaar verschil te bestaan tussen de fonetische realisatie van de twee tonen in de twee verschillende soorten vraagzinnen. Daarom werd hierover gemiddeld. Evenmin werden, in tegenstelling tot Schmidt (1986), verschillen gevonden tussen korte en lange woorden, zodat ook hier het gemiddelde werd genomen. Op deze manier kregen we in totaal 4 standaardcontouren voor iedere toon (zie figuur 2). Op de klinker laat de sleeptoon duidelijk een (dalend-)stijgend patroon zien en de stoottoon een (stijgend-)dalend patroon. Een uitzondering hierop vormt de sleeptoon in finale positie in bewerende zinnen. Hier blijft



Figuur 2. Standaard-contouren van de sleeptoon en de stoottoon in vier verschillende linguïstische contexten. Iedere contour laat de hele geaccentueerde lettergreep zien. De dikke lijnen zijn de klinkers.

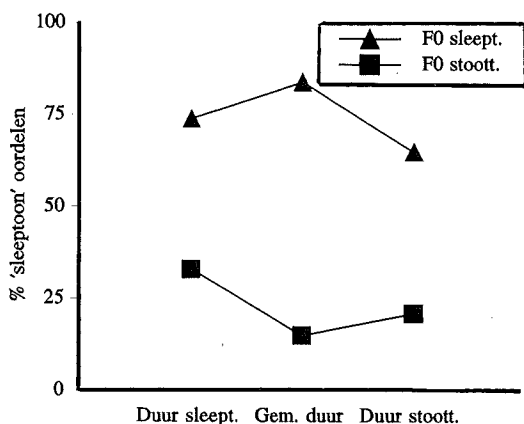
tussen de medeklinkerduren waren in de meeste condities niet significant, maar omdat er in één conditie (de volgende medeklinker in woorden met stoottoon in beklemtoonde en onbeklemtoonde lettergrepen, zie paragraaf 3.1) wel een significant verschil werd gevonden, werd besloten om ook de duren van de medeklinkers te manipuleren. Het is moeilijk om van sleptonen en stoottonen combinaties van kenmerken te maken, omdat duur en toonhoogtecontour niet onafhankelijk zijn van elkaar. Als de duur namelijk langer wordt, wordt de helling òf minder steil òf de plaats van de toonhoogtecontour in de klinker verandert. Wij volgden de beschrijvingen die werden gemaakt voor de standaardcontouren in paragraaf 3. Dit betekent dat het hoogste/laagste punt werd uitgerekend als een percentage vanaf de klinkerinzet. Verder werden de hellingen van de toonhoogtebewegingen gemaakt in overeenstemming met deze beschrijvingen. Dit betekent dat de inzet van de stijging/daling in ms vanaf klinkerinzet veranderde als de duur werd veranderd. Daarnaast werd het toonhoogtebereik constant gehouden. Om het eventuele effect van intensiteit, klankkleur en fonetische context op de herkenning van de woorden te balanceren, werden alle sleptonwoorden op zinnen gelegd die oorspronkelijk met een stoottoon werden uitgesproken, en andersom. De twee tonen zetten op verschillende hoogten in. Dit betekent dat de toonhoogtecontour van de geaccentueerde lettergreep in sommige gevallen moest worden gecoarticuleerd met de resterende toonhoogtecontour. Dit werd gedaan zonder de standaardcontour van de geaccentueerde lettergreep te beïnvloeden. Slechts de resterende toonhoogtecontour werd waar nodig aangepast. De manipulaties werden uitgevoerd met LVS (Vogten 1984).

De 48 verschillende versies van de vier zinnen werden in random volgorde op een band gezet. Alle zinnen werden uitgeschreven en de gemanipuleerde woorden werden onderstreept. Ernaast stonden de twee Nederlandse vertalingen van de onderstreepte woorden. Achttien luisteraars die geboren en getogen waren in Maastricht en het Maastrichts als moedertaal hadden, werd gevraagd de Nederlandse vertaling die zij het beste vonden overeenkomen met hetgeen ze hoorden te omcirkelen (gedwongen keuze). De stimuli werden via koptelefoons aangeboden. Voordat het experiment begon, kregen de luisteraars eerst zes oefenzinnen. Ze werden betaald voor hun medewerking.

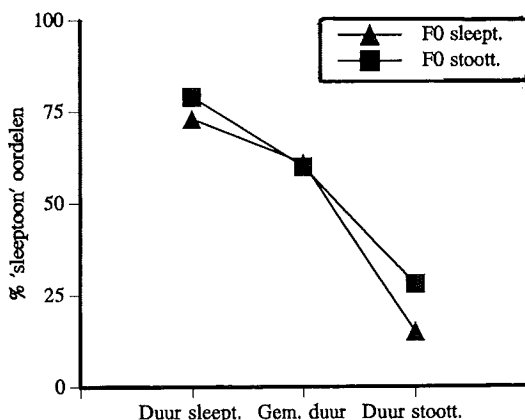
4.2 Resultaten

Om na te gaan of het luisterpanel een betrouwbaar meetinstrument vormde, is de alpha-coëfficiënt berekend. Deze bedroeg .85 ($F(47,799) = 6.71$ $p < 0.01$) wanneer de oordelen over alle versies gepooled waren, en .88 ($F(31,527) = 8.25$ $p < 0.01$), als de scores werden uitgerekend op basis van de versies waarin geen akoestische parameters gekruist zijn (de versies 1 en 6 in tabel 2). De betrouwbaarheid is dus redelijk hoog.

In figuur 3 worden de gemiddelde resultaten voor alle zinnen gegeven. Op de verticale as wordt het percentage sleptonherkenning voor alle versies vermeld. De resultaten laten zien dat zowel de toonhoogtecontour (F0) als de duur belangrijk zijn voor de herkenning van de tonen. Dit wordt ook bevestigd door de log-lineaire analyse met responsie, toonhoogtecontour (F0) en duur als variabelen (zie Rietveld & Van Hout 1993). Er werd tweezijdig getoetst op het 5%-niveau.



Figuur 4. Percentage 'sleeptoon'-oordelen voor *deur*, als functie van temporele structuur en toontype, gepooled over alle versies.



Figuur 5. Percentage 'sleeptoon'-oordelen voor *numme*, als functie van temporele structuur en toontype, gepooled over alle versies.

moeite gehad om informatie over de toonhoogtecontour te gebruiken en hebben bijna uitsluitend de duurverschillen gebruikt in hun beoordelingen. Er is een interactie tussen duur en toonhoogtecontour ($z = -1.69, 2.19$).

5 Discussie

In dit onderzoek werden duur- en toonhoogtemetingen uitgevoerd op minimale toonparen in zinnen waarvan een aantal linguïstische factoren systematisch waren gevarieerd. De resultaten van deze metingen lieten zien dat er behalve een significant verschil in duur ook duidelijke verschillen zijn in toonhoogtecontouren. De toonhoogtecontour van een klinker in een woord met stoottoon heeft een vorm die

Het tweede doel van het perceptie-experiment was om conclusies te kunnen trekken over de bijdragen die de twee akoestische parameters duur en toonhoogtecontour leveren aan de toonherkenning. Wat *deur* betreft is de toonhoogtecontour de belangrijkste parameter voor de luisteraars bij de herkenning van de twee tonen, hoewel de duur ook een bijdrage levert.

Numme wordt net zo goed herkend als *deur* (79% van de sleптоonwoorden en 85% van de stoottoonwoorden worden juist herkend). De toonhoogtecontouren van *numme* leveren echter slechts in enkele versies een bijdrage aan de herkenning van de toon. Er zijn hier drie mogelijke verklaringen voor te geven. Ten eerste zou het kunnen dat de standaardcontouren niet juist zijn. Deze standaardcontouren zijn echter gemiddeld over twee korte woorden (*deur* en *veer*) en twee lange woorden (*numme* en *boete*). Voor *deur* blijkt het wel de juiste standaardcontour te zijn. Om deze reden zou men verwachten dat de standaardcontour ook voor *numme* correct zou zijn. Een tweede verklaring zou te vinden kunnen zijn in de segmentele opbouw van het woord zelf. Het is al eerder gebleken dat toonhoogtebewegingen niet in alle segmenten even goed waarneembaar zijn. Zo is in de Nijmeegse synthese van grenssegmenten gebleken dat toonhoogtebewegingen niet goed waarneembaar zijn in bepaalde segmenten. Tijdens een nasaal moeten toonhoogtebewegingen bijvoorbeeld worden gerealiseerd met een extra grote range om ze hoorbaar te maken (Kerkhoff, Rietveld, Gussenhoven & Ehlich, te verschijnen). Een plausible verklaring is dat de mond tijdens het uitspreken van een nasaal dicht blijft, hetgeen een lage amplitude tot gevolg heeft. De toonhoogtecontouren zijn moeilijk waarneembaar als de amplitude laag is, en daarom worden de luisteraars gedwongen de duurverschillen te gebruiken om de verschillen te kunnen horen tussen de verschillende versies. Een derde verklaring zou kunnen liggen in het feit dat de duurverschillen tussen de klinkers van de stoottoon en de sleптоon groter zijn in *numme* (ongeveer 1:2) dan in *deur* (ongeveer 1:1,5). Het is dus makkelijker de duur te gebruiken voor de toonherkenning bij *numme* dan bij *deur*. Omdat we er in onze standardisaties vanuit gaan dat de gehele toonbeweging binnen één lettergreep valt is er bovendien ook minder tijd om een stoottoonbeweging te realiseren in *numme* (gemiddeld 83 ms) dan in *deur* (gemiddeld 119 ms). Om deze reden is het makkelijker een verschil in toonhoogteverloop te horen in *deur* dan in *numme* en is duur een belangrijker parameter in *numme* dan in *deur*. Het zou kunnen dat de toonhoogtebeweging in natuurlijke spraak overvloeit naar de volgende (onbeklemtoonde) lettergreep.

6 Conclusie

De akoestische metingen lieten zien dat er een verschil is tussen de klinkerduren van stoottonen en sleптоnen. Het toonhoogteverloop laat ook een duidelijk verschil zien. Het is mogelijk om voor elke toon een basisconfiguratie aan te nemen. De sleптоon heeft een (dalend-)stijgend patroon en de stoottoon een (stijgend-)dalend patroon. Het is onduidelijk hoeveel variatie er mogelijk is binnen de verschillende basisconfiguraties. Er werden geen duidelijke verschillen gevonden tussen verschillende soorten vraagzinnen. De plaats van de toon in het woord

- Hermes, D.J. (1988), Measurement of pitch by subharmonic summation, *Journal of the Acoustical Society of America* 83, 257-264.
- Jongen, R. (1972), *Rheinische Akzentuierung und sonstige prosodische Erscheinungen. Eine Beschreibung der suprasegmentalen Zeichenformdiakrise in der Moresneter Mundart*, Bonn: Ludwig Röhrscheid Verlag.
- Kerkhoff, J., T. Rietveld, C. Gussenhoven en L. Ehlich (te verschijnen), NIROS: The Nijmegen Interactive Rule Oriented Speech Synthesis System.
- Nooteboom, S.G. (1972), Production and perception of vowel duration. (Proefschrift Rijksuniversiteit Utrecht.)
- Notten, J.G.M. (1988), *De Chinezen van Nederland. Opstellen over Limburgse Dialecten en een Bibliografie*. Valkenburg aan de Geul: Valkdruk B.V.
- Rietveld, A.C.M. & C. Gussenhoven (1985), On the relation between pitch excursion size and prominence, *Journal of Phonetics* 13, 299-308.
- Rietveld, T. & R. van Hout (1993), *Statistical Techniques for the Study of Language and Language Behaviour*. New York: Mouton de Gruyter.
- Schmidt, J.E. (1986), *Die Mittelfränkischen Tonakzente (Rheinische Akzentuierung)*, Stuttgart: Franz Steiner Verlag Wiesbaden.
- Sluiter, A.M.C. & V.J. van Heuven (1995), Effects of focus distribution, pitch accent, and lexical stress on the temporal organisation of syllables in Dutch, *Phonetica* 52, 71-89.
- Stevens, A. (1955), Intonatieproblemen in en om West-Limburg, *Taal en Tongval* 7, 135-142.
- Verhoeven, J. (1992), Fonetische kenmerken van sleep- en stoottoon in het Dialect van Weert, *Taal en Tongval* 44, 140-155.
- Vogten, L.L.M. (1984), Analyse, zuinige codering en resynthese van spraakgeluid. (Proefschrift TH Eindhoven.)
- Welter, W. (1933), *Die Niederfränkischen Mundarten im Nordosten der Provinz Lüttich*, Den Haag: Nijhoff.
- Willems, L.F. (1987), Robust formant analysis for speech synthesis applications. In: *European Conference on Speech Technology*, 250-253. Edinburgh: CEP Consultants.