

# *Mondjesmaat*

Een onderzoek naar de beoordeling van regionale kleuring binnen het Standaardnederlands en een vergelijking tussen het hanteren van een Likertschaal met 6 of 7 antwoordmogelijkheden

Onder leiding van Hans Van de Velde en Anne-France Pinget

Eerste lezer: Hans Van de Velde

Tweede lezer: Sander van der Harst

## Inhoudsopgave

Inleiding .....	4
1. Voorafgaand onderzoek .....	6
2. Theoretisch kader .....	8
2.1 Standaardnederlands .....	8
2.2 Attitudeonderzoek op het gebied van standaard en non-standaard taalvariëteiten .....	9
2.3 Likertschalen.....	10
2.3.1 Likertschaal en Osgoodschaal .....	10
2.3.2 Discussie ideaal aantal Likertschaalcategorieën .....	10
2.3.3 Discussie even vs. oneven aantal Likertschaalcategorieën .....	11
3. Hypothesen .....	13
3.1 Hypothese Likertschaalonderzoek .....	13
3.2 Hypothesen attitudeonderzoek .....	13
4. Methode .....	16
4.1 De internetenquête .....	16
4.2 Stimuli/sprekers .....	16
4.3 Beoordelaars .....	17
4.4 Verwerking van de resultaten .....	18
5. Resultaten.....	20
5.1 Correlaties tussen de schalen en betrouwbaarheid.....	20
5.2 Resultaten factoren repeated measures-model .....	21
5.3 Resultaten Likertschaalonderzoek .....	24
5.4 Resultaten attitudeonderzoek.....	25
5.4.1 Standaard .....	27
5.4.2 Mooi .....	28
5.4.3 Niet-regionaal.....	30
5.4.4 Verzorgd .....	32

6. Discussie .....	35
6.1 De hypothesen herhaald .....	35
6.2 Discussie Likertschaalonderzoek .....	35
6.3 Discussie Attitudeonderzoek.....	36
7. Conclusie en vervolgonderzoek.....	39
7.1 Attitudeonderzoek .....	39
7.2 Likertschaalonderzoek.....	39
Bronnen .....	41

# Inleiding

Deze scriptie is een onderzoek naar hoe Standaardnederlands, gesproken door docenten uit verschillende Nederlandse regio's, wordt beoordeeld. Het onderzoek kan beschouwd worden als een attitudeonderzoek met daarin een methodologisch onderzoek naar de invloed van het hanteren van een 6- of 7-punts Likertschaal.

## **Het Attitudeonderzoek:**

In een enquête beoordeelden informanten hoe *standaard*, *regionaal*, *mooi* en *verzorgd* ze de spraak van 20 sprekers uit 4 verschillende regio's vonden. In de onderzoeksopzet zijn de beoordelaarsregio (Midden, Noord, Randstad, Zuid), sekse, leeftijd en opleidingsniveau als factor in een repeated measures-onderzoeksmodel opgenomen. Dit attitudeonderzoek kan gezien worden als een vervolgstudie op de onderzoeken van Pinget et al. (2013) en Grondelaers et al. (2010).

- De hoofdvraag van het attitudeonderzoek is: *“In hoeverre wordt het Standaardnederlands uit de vier gebieden als regionaal gekleurd beoordeeld en welke normoordelen verbinden de beoordelaars aan regionaal gekleurd Standaardnederlands?”*

## **Het Likertschaalonderzoek:**

Het Likertschaalonderzoek naar de invloed van het hanteren van 6 of 7 Likertschaalcategorieën is gedaan binnen het attitudeonderzoek. In het Likertschaalonderzoek hebben twee verschillende beoordelaarsgroepen een verschillende internetenquête ingevuld, waarbij het enige verschil was dat de ene groep de spraak op 6-puntsschalen beoordeelde, en de andere groep op 7-puntsschalen. De enige variabele bij het Likertschaalonderzoek was het aantal meetpunten (lengte) van de schalen.

- De hoofdvraag wat betreft het Likertschaalonderzoek is: *“Wat zijn de verschillen in de uitkomsten van een attitudeonderzoek met een 6-puntsschaal of een attitudeonderzoek met een 7-puntsschaal?”*

## **Opbouw van de scriptie:**

In het eerste hoofdstuk van deze scriptie worden de onderzoeken van Pinget et al. (2013) en Grondelaers et al. (2010) belicht. Hierbij wordt aangegeven in welke opzichten de onderzoeken van elkaar verschillen en hoe deze scriptie zich verhoudt tot de voorgaande onderzoeken.

In hoofdstuk 2 wordt een overzicht gegeven van de verschillende onderzoeken op het gebied van het Standaardnederlands. In dit hoofdstuk wordt het onderzoek behandeld dat gedaan is op het gebied van attitudeonderzoek naar accenten en non-standaardvariëteiten binnen (en eventueel tegenover) het Standaardnederlands. Daarna worden Likertschalen in het algemeen besproken, en wordt aandacht besteed aan onderzoek op het gebied van de invloed van de lengte van de schalen.

Het derde hoofdstuk bestaat uit de hypothesen voor zowel het attitudeonderzoek als het Likertschaalonderzoek, die aan de hand van het onderzoek worden getest.

In hoofdstuk 4 wordt de onderzoeksmethode besproken, alsmede de beoordelaars, de stimuli en de verwerking van de resultaten. Als eerste worden in het hoofdstuk daarna de resultaten weergegeven van het Likertschaalonderzoek, waarna het attitudeonderzoek per schaal wordt behandeld.

Vervolgens volgt hoofdstuk 6 “Discussie”, waarin de resultaten worden besproken aan de hand van de verschillende hypothesen.

In hoofdstuk 7 worden de hoofdvragen beantwoord, de conclusies van beide onderzoeken behandeld en worden eventuele vervolgstudies vermeld.

# 1. Voorafgaand onderzoek

Dit onderzoek kan gezien worden als een vervolgstudie op de onderzoeken van Pinget et al. (2013) en Grondelaers et al. (2010). Daarin werd onderzocht hoe sprekers van het Nederlands de spraak van leerkrachten van de middelbare school (uit het Vlaams-Nederlandse Leerkrachtencorpus, CGN) uit vier regio's van Nederland beoordeelden.

De vraagstelling en onderzoeksopzet van Pinget et al. (2013) zijn als uitgangspunt genomen voor dit scriptieonderzoek. Echter, de stimuli die worden beoordeeld komen meer overeen met de stimuli uit het onderzoek van Grondelaers et al. (2010). De overeenkomsten en verschillen tussen deze scriptie, Pinget et al. (2013) en Grondelaers et al. (2010), worden in dit hoofdstuk op een rijtje gezet. Ook is te zien waar dit onderhavige onderzoek een toevoeging kan zijn bij de voorgaande onderzoeken.

In de onderzoeken van Pinget et al. (2013) en Grondelaers et al. (2010) werd een attitudeonderzoek naar regionale kleuring in het Standaardnederlands gecombineerd met een classificatieonderzoek waarin de proefpersonen moesten inschatten uit welke regio (Midden, Noord, Randstad, Zuid) de spreker afkomstig was. In Pinget et al. (2013) werd naast het classificatieonderzoek, ook de invloed van spreekstijlen (spontaan vs. gelezen) onderzocht. In deze scriptie zijn dit classificatieonderdeel en de invloed van spreekstijlen niet opgenomen (met name i.v.m. de tijdsduur van het invullen van de enquête).

De twee hiervoor genoemde onderzoeken verschillen in opzet op een aantal belangrijke aspecten van elkaar.

- Pinget et al. (2013) hanteerde een directe vraagstelling, waar Grondelaers et al. (2010) een combinatie van een directe en een indirecte gebruikte. In deze scriptie wordt de directe vraagstelling uit Pinget et al. (2013) gevolgd.
- In Pinget et al. (2013) werd gebruikt gemaakt van spraakfragmenten van jongere sprekers uit het CGN. Grondelaers et al. (2010) werkte met een (kleinere) selectie van oudere sprekers. Wat dat betreft komt deze scriptie meer overeen met laatstgenoemde. De door de beoordelaars waargenomen leeftijd speelt een rol in de beoordeling van de spreker (ook al kent de beoordelaar de leeftijd van de spreker niet (Cargile et al. 1994)). Daarom vormt de leeftijd van de sprekers een interessant contrast tussen Pinget et al. (2013) en deze scriptie die de leeftijdscategorie van Grondelaers et al. (2010) volgt.
- De schalen in deze scriptie, waarop de beoordelaars de spraak beoordelen, zijn dezelfde als in Pinget et al. (2013). De gebruikte schaallabels zijn *standaard*, *regionaal*, *mooi* en *verzorgd*. In Grondelaers et al. (2010) beoordeelde men de sprekers uit de vier regio's aan de hand van andere maatstaven, namelijk: *norm*, *status*, *integriteit* en *welluidendheid*.

Wat betreft de resultaten zijn er zowel overeenkomsten als verschillen tussen de twee voorgaande onderzoeken:

- De sprekers uit de Randstad werden over het algemeen positiever beoordeeld dan de sprekers uit het Midden, het Noorden en het Zuiden van het land. Deze generalisatie gaat

niet op voor het kernmerk *integriteit* en *welluidendheid* in Grondelaers et al. (2010), waarbij de regio Zuid de hoogste scores kreeg.

- Een duidelijk verschil in de resultaten tussen beide onderzoeken was de factor “beoordelaarsregio” (en de andere demografische factoren van de beoordelaars, te weten “sekses” en “leeftijd”) die in Pinget et al. (2013) resulteerde in significante effecten, wat niet het geval was bij het onderzoek van Grondelaers et al. (2010) (er werd één effect bij de factor “leeftijd” gevonden, maar dit effect werd toegeschreven aan de grote variatie in de leeftijd van de beoordelaars). In het laatstgenoemde onderzoek werd geconcludeerd dat de herkomst van de beoordelaar nauwelijks invloed had op de beoordeling, en er een algemene standaardattitude bestaat tegenover sprekers uit de verschillende regio’s. Pinget et al. (2013) concludeerden juist dat er wel degelijk een effect van de beoordelaarsregio plaatsvond op de vier schalen. Wellicht geeft het onderzoek in deze scriptie uitsluitel.

## 2. Theoretisch kader

Nu de opzet van deze scriptie in het licht van de voorgaande onderzoeken duidelijk is, wordt in dit hoofdstuk het ruimere theoretische kader besproken. In de eerste paragraaf komt het Standaardnederlands aan de orde, wat men standaard vindt klinken en hoe luisteraars standaardtaal en standaardtaalsprekers evalueren. Daarna zal het theoretisch kader van de Likertschaal behandeld worden, met daarin de discussies over de lengte van de schalen en de discussie over een even vs. oneven aantal antwoordopties. In deze volgorde worden de onderwerpen in dit hoofdstuk uiteengezet.

### 2.1 Standaardnederlands

De uitspraak van een Standaardnederlandsspreker uit Limburg en die van een spreker uit Amsterdam zijn niet hetzelfde. Ook wordt Standaardnederlands uit verschillende regio's door de luisteraars anders beoordeeld. De luisteraar hoeft een taalvariëteit niet eens te herkennen om er toch een mening over te vormen. Wat wordt er precies onder Standaardnederlands verstaan?

“In de taal kan “standaard” omschreven worden als de taal, of uitspraak ervan, die het meest overeenkomt met de meest dominante linguïstische norm in een taalgemeenschap” (Smakman 2006). In Nederland wordt de Nederlandse spraaknorm voor het publieke domein, de media en het onderwijs, wel het “ABN” genoemd. Het Algemeen Beschaafd Nederlands. De meest algemene kritiek op deze term houdt verband met het waardeoordeel dat kleeft aan het woord “beschaafd” (Smakman 2006). De term “ABN” werd de laatste jaren door taalwetenschappers vervangen door de meer politiek correcte term “Standaardnederlands”. Deze modernere term heeft ook het voordeel dat deze rechtstreeks samenhangt met andere talen met vergelijkbare termen in hun eigen taalgemeenschap (Smakman 2006).

Hoewel er een norm is binnen de Nederlandse taalgemeenschap, betekent dat nog niet dat Standaardnederlands accentloos hoeft te zijn. Of een spreker als “standaardtalige” wordt geëvalueerd ligt niet alleen aan de mate waarin fonetische kenmerken worden gebruikt, maar ook aan het foneem in kwestie en de positie van het foneem (Smakman 2006). Verschillen in de uitspraak van het Nederlands zijn voor de luisteraar voornamelijk waar te nemen in de manier waarop klinkers en medeklinkers worden uitgesproken. Lexicale en grammaticale variatie is minder duidelijk of komt minder vaak voor (Smakman 2006).

De meest gangbare opvatting is dat de gemiddelde Nederlander vindt dat het meest standaard en accentloze Nederlands “ergens in het westen van het land wordt gesproken”. Veelal wordt “Haarlem” genoemd als de plaats van het Standaardnederlands, of de plaats waarvan de uitspraak het dichtst bij de standaardnorm ligt. Smakman (2006) concludeerde in zijn proefschrift, een sociolinguïstische en fonetische beschrijving van het Nederlands, inderdaad dat sprekers uit het westen van het land het meest werden geassocieerd met het Standaardnederlands. Dit komt ook naar voren in taalkundig onderzoek op het gebied van taalvarianten en verspreiding van fonetische kenmerken (Hoppenbrouwers 1990, Smakman & Van Bezooijen 1999) en onderzoek naar de perceptuele dialectafstand van het Standaardnederlands (Van Hout & Münstermann 1981).



## 2.2 Attitudeonderzoek op het gebied van standaard en non-standaard taalvariëteiten

Er is sinds de jaren '60, maar hoofdzakelijk vanaf de jaren '70, veel onderzoek gedaan naar beoordeling van spraak en taalattitudes. In 1960 werd het vroegste en één van de meest geciteerde standaardwerken op het gebied van taalattitudeonderzoek gepubliceerd door Lambert et al. (Cargile et al. 1994). Ook in het Nederlandse taalgebied is veel attitudeonderzoek gedaan. Het kan hierbij gaan om de attitudes wat betreft de esthetische waarde van spraak aan de hand van context, prosodische kenmerken en verzorgdheid bij de articulatie (Van Bezooijen 2002), maar ook bijvoorbeeld om attitudes over talen die de beoordelaar niet kent (Van Bezooijen 1998).

In dit onderzoek gaat het echter over wat luisteraars als Standaardnederlands ervaren en hoe regionale variatie in het Standaardnederlands wordt beoordeeld.

Er zijn verschillende manieren om de attitudes ten opzichte van spraak te onderzoeken. Een directe vraagstelling waarbij spraak wordt beoordeeld, of een indirecte vraagstelling waarbij de spreker wordt beoordeeld naar aanleiding van spraak. Beide vraagstellingen komen in attitudeonderzoek veelvuldig voor (voor directe vraagstelling bijv. Pinget et al. (2013), Van de Velde & Houtermans (1999), Smakman (2006) en voor indirecte vraagstelling bijv. Van Bezooijen (1988), Hout & Münstermann (1981) voor een combinatie van directe en indirecte vraagstelling bijv. Grondelaers et al. (2010), Knops (1988)).

In attitudeonderzoek is er traditioneel een aantal belangrijke variabelen: de achtergrond van de informant (socio-economische factoren, geboortestreek, etnische achtergrond) en de leeftijd van de informant (kinderen en tieners leren naarmate ze ouder worden steeds meer de sociale spraaknormen kennen, terwijl ouderen meer open staan voor non-standaard variatie van spraak (Cargile et al. 1994)). Daarnaast zijn er nog variabelen als opleidingsniveau en sekse die veelvuldig in attitudeonderzoek worden toegepast. Deze variabelen zullen ook in dit onderzoek aan worden gehouden.

Als taal door luisteraars wordt bestempeld als *standaard*, wordt deze ook vaak in verband gebracht met omschrijvingen als *formeel*, *correct* en *modern*. Uit onderzoeken naar de perceptie van het Standaardnederlands is gebleken dat de schalen *mooi*, *verzorgd* en *regionaal* voornamelijk correleren met de schaal *standaard* (Van de Velde & Houtermans (1999), Van Bezooijen (2002), Smakman (2006), Pinget et al. (2013)). Deze onderzoeken zijn gedaan aan de hand van gangbare Likertschalen met 7 of 10 antwoordmogelijkheden. Hoewel in de vorige paragraaf het westen van Nederland werd genoemd als de regio waar het meest standaard Nederlands wordt gesproken, wordt deze over het algemeen toch als niet-regionaal beoordeeld. Sterke correlaties tussen *niet-regionaal* en *standaard* werden in Pinget et al. (2013) voornamelijk bij de leesstijl gevonden. De sprekers van standaardtaal worden omschreven als *beschaafd*, *hoogopgeleid* en daarbij hebben ze een *hoge status* (indirecte vraagstelling). Deze positieve benamingen voor de sprekers en de uitspraak van het Standaardnederlands correleren sterk met het begrip *standaard*.

Uitgaande van deze gegevens wordt in hoofdstuk 4 een overzicht gegeven van de hypothesen van het attitudeonderzoek van deze scriptie.

## 2.3 Likertschalen

Voor het onderzoeken naar de houding of attitude van proefpersonen stelde Rensis Likert (1932) schalen op met vijf antwoordalternatieven. Deze waren gelabeld met aanduidingen van meer of minder mee eens of neutraal. Het aantal alternatieven kan volgens Likert (1932) ook gemanipuleerd worden. Een even of een oneven aantal schaalpunten kan bijvoorbeeld van invloed zijn op de resultaten, met name vanwege de (on)mogelijkheid van het geven van een neutraal antwoord. In de literatuur komen 4-tot 7-punts Likertschalen het meeste voor (Clason & Dormody 1994). Deze manier van onderzoek doen naar moeilijk te kwalificeren materie is voornamelijk in de psychologie, psychometrie en de sociale wetenschappen een belangrijk hulpmiddel, maar bijvoorbeeld ook in marketingonderzoek. Er is veel onderzoek gedaan op het gebied van Likertschalen en het te hanteren aantal Likertschaalcategorieën. In de psychologie en psychometrie o.a. Chang (1994), Finn (1972), Matell & Jacoby (1971), Wakita et al. (2012), in marketingonderzoek Dawes (2008), Kim (2010), Garland (1991), Worcester & Burns (1975) en in de sociale wetenschappen Carifio & Perla (2007).

### 2.3.1 Likertschaal en Osgoodschaal

De Likertschaal is niet de enige schaal die wordt gebruikt om moeilijk te kwalificeren gegevens te meten in de psychologie en de sociale wetenschappen: ook de Osgoodschaal wordt veelvuldig gebruikt. Bij zowel de Likertschaal als de Osgoodschaal geven de informanten aan welk punt op de schaal overeenkomt met hun attitudeoordeel.

De klassieke Likertschaal bestaat uit een aantal schaalpunten die elk zijn gelabeld met categorieën als “zeer mee eens” tot en met “zeer mee oneens” of “nooit” en “soms”. De vraagstelling bij een Likertschaal is over het algemeen een stelling.

Een Osgoodschaal, of een semantische differentiaal, maakt geen gebruik van deze labels. Dit type psychometrische schaal is genoemd naar C.E. Osgood, die in plaats van een label per antwoordmogelijkheid, twee tegengestelde adjectieven aan weerszijden van de schaal plaatste (Osgood et al. 1957). De vraagstelling van een Osgoodschaal is anders. De vraagstelling is meer open te noemen, omdat er gekozen moet worden tussen twee contrasterende adjectieven, of een punt hiertussen. Bij zowel een gelabelde als een ongelabelde schaal wordt aangenomen dat het oordeel door de beoordelaars proportioneel over elk verschillend alternatief wordt verdeeld (Kim 2010).

In het onderzoek van deze scriptie is de gehanteerde methode strikt gezien een variant op de klassieke Likertschaal, aangezien de antwoordopties niet alle werden gelabeld. De uiteinden van de schalen hadden echter wel de labels “helemaal mee eens” en “helemaal niet mee eens”. Aangezien de beoordelaars in het onderzoek zelf het attitudeoordeel verdeelden op het aantal schaalpunten, kunnen we zeggen dat de data ordinaal verdeeld zijn. Dit in tegenstelling tot een klassiek gelabelde Likertschaal, waarbij er nog steeds discussie bestaat in hoeverre Likertschaaldata geïnterpreteerd mogen of moeten worden als ordinale data (Norman 2010).

### 2.3.2 Discussie ideaal aantal Likertschaalcategorieën

Zoals eerder genoemd bestaat het Likertschaalonderzoek uit deze scriptie in feite uit twee verschillende attitudeonderzoeken waarbij het enige verschil is dat bij het ene onderzoek beoordeeld wordt op 6-punts Likertschalen en het andere op 7-punts Likertschalen.

Het aantal antwoordmogelijkheden kan effect hebben op een onderzoek. Te weinig mogelijkheden leiden tot een gecompliceerde dataverwerking van data zonder veel nuanceverschillen. Dat maakt het accuraat onderzoeken naar de attitude van de beoordelaar moeilijk (Kim 2010). Te veel antwoordmogelijkheden kunnen daarentegen zorgen voor onbetrouwbare waardeoordelen en veroorzaken sneller vermoeidheid. Daarbij is het de vraag of het onderscheidingsvermogen van de beoordelaars een groot aantal antwoordmogelijkheden aankan. Tegelijkertijd wordt wel aangenomen dat een Likertschaal betrouwbaarder wordt naarmate deze meer antwoordopties heeft (Matell & Jacoby 1971).

Er is veel onderzoek gedaan naar de optimale lengte van de Likertschaal. Dat wil zeggen, het optimaal te hanteren aantal antwoordopties op de schaal. Onderzoek op dit gebied spitte zich voornamelijk toe op de vraag bij welke Likertschaallengte de betrouwbaarheid en validiteit het grootst zijn. Er is enige variatie in het aantal categorieën, maar volgens het invloedrijke psychometrische boek "Psychometric Theory" van Nunnally (1978) lijkt een Likertschaal met 7 punten de hoogste mate van betrouwbaarheid te hebben. Recentere studies over de betrouwbaarheid en validiteit van Likertschalen bevestigen dit. Over het algemeen komen deze studies uit op een optimale schaal van 5 tot 7 categorieën (Kim 2010), of 6 tot 8 volgens Carifio & Perla (2007).

Een kortere of langere schaal lengte hoeft echter niet noodzakelijkerwijs de oorzaak te zijn van de mate van betrouwbaarheid en validiteit van een onderzoek. Bij verschillende schaal lengtes kan er namelijk ook sprake zijn van variatie in het betrouwbaarheidscoëfficiënt omdat hetgeen gemeten wordt niet past in een hoger of juist lager aantal schaal categorieën (Chang 1994, Kim 2010).

### **2.3.3 Discussie even vs. oneven aantal Likertschaal categorieën**

In dit onderzoek is gebruikgemaakt van 6- en 7-puntsschalen. De literatuur gaat er bijna unaniem vanuit dat de lengte van deze schalen binnen de ideale marge van Likertschalen vallen. In het voorliggende onderzoek wordt er vanuit gegaan dat wanneer er een significant verschil optreedt tussen de twee datasets, dit niet vanwege het aantal antwoordopties is. (Immers, zowel de 6- als 7-puntsschaal liggen binnen de "ideale" marge). Als er een significant verschil optreedt, dan is dit vanwege de mogelijkheid tot een neutraal antwoord of juist de onmogelijkheid hiervan.

Bij de discussies tussen het hanteren van even of oneven Likertschaal categorieën geldt over het algemeen het volgende: bij een 7-puntsschaal kan de beoordelaar een neutraal antwoord geven, of in ieder geval een antwoord geven waarbij er geen keuze wordt gemaakt in de richting van één van de twee uitersten. Bij een 6-puntsschaal (of een andere schaal met een even aantal categorieën) wordt de beoordelaar gedwongen om een voorkeur aan te geven voor één van de twee zijden van de schaal.

Het hanteren van een even of oneven aantal Likertschaal categorieën kan dan ook van invloed zijn op de uiteindelijke resultaten. Hoewel er in marktonderzoek en psychologie veel onderzoek is gedaan naar de invloed van het aantal Likertschaalpunten op de hoogte van de antwoordscores (o.a. Chang (1994), Dawes (2008), Finn (1972)), is dit in de sociolinguïstiek nog een onontgonnen gebied. Aangezien deze type schalen veelvuldig in sociolinguïstisch onderzoek worden gebruikt, (o.a. Pinget

et al. (2013), Van Bezooijen (2002) en Grondelaers et al. (2010), Smakman (2006)) is een onderzoek op dit gebied in een methodologisch kader erg interessant.

De keuze tussen het gebruiken van een even of een oneven aantal categorieën in onderzoeken hangt over het algemeen af van wat de onderzoeksvraag is. Het is mogelijk dat iets onderzocht wordt waarbij een neutraal antwoord of een categorie “ik weet het niet” wenselijk is. Bijvoorbeeld wanneer men verwacht dat er veel neutrale antwoorden zullen worden gegeven, of wanneer mensen zich schamen om een kant te kiezen (Grover & Vriens 2006). In attitudeonderzoek, zoals het in deze scriptie is uitgevoerd, is een antwoord als “neutrale mening” een minder bruikbaar antwoord. Maar bijvoorbeeld een evaluatiescore van 4 uit 7 voor een bepaalde perifere taalvariëteit is wel bruikbaar in vergelijking met een 2 uit 7 van de taalvariëteit uit het centrum.

Uit een marketingonderzoek bestaande uit een vergelijking van de data van een traditionele Likertschaal met de data van een lege schaal zonder punten (ook wel “visual analogue scale”) van Worcester & Burns (1975), werd geconcludeerd dat een even schaal (met 4 punten) zonder neutrale optie de beoordelaars naar meer positieve antwoorden dreef dan hetzelfde onderzoek met een 5-puntsschaal. Een herhaald marketingexperiment wat betreft deze onderzoeksmethode, door Garland (1991), resulteerde in het tegenovergestelde: de beoordelaars gaven negatievere scores op de 4-puntsschaal ten opzichte van de 5-puntsschaal. Garland veronderstelde dat dit verschil te verklaren was in de geteste materie en het relatief kleine aantal proefpersonen. Deze marktonderzoeken geven dus geen duidelijkheid in deze kwestie.

In een recent psychometrisch onderzoek werden de antwoorden vergeleken op 4-, 5- en 7-puntsschalen om de invloed van de verschillende schaallengtes te testen (Wakita et al. 2012). De onderzoekers concludeerden dat bij de 7-puntsschaal met neutrale optie meer negatieve antwoorden als “oneens” werden gegeven. Daarbij werd op dezelfde schaal het antwoorden op de extremen van de schalen gemeden. Hierbij moet aangetekend worden dat de negatieve vraagstelling (d.w.z. een vraagstelling met een ontkenkende inleidende vraag) wellicht invloed heeft gehad op de antwoorden. In een methodologisch marketingonderzoek van Kim (2010) werd onderzocht wat het verschil was tussen onderzoeken met 5-, 6- en 7-puntsschalen in betrouwbaarheid en validiteit. Ook werden andere effecten onderzocht, zoals de hoogte van de scores van de beoordelaars bij verschillende schaallengtes. Uit het onderzoek werd de tendens gevonden dat een oneven schaal zorgt voor lagere scores (minder positief). Daarnaast werd geconcludeerd dat naarmate er meer antwoordcategorieën waren, de scores positiever werden. Deze verschillen waren echter niet significant (Kim 2010).

De verschillende onderzoeken binnen zowel de psychologie als het marktonderzoek lijken elkaar tegen te spreken, en sterk af te hangen van de onderzoeksmaterie en het aantal beoordelaars. Over het algemeen zijn de verschillen klein. Er is eerder sprake van een tendens dan van duidelijke conclusies, zoals te zien is in onder andere de onderzoeken van Wakita et al. (2012) en Kim (2010).

# 3. Hypothesen

In dit hoofdstuk worden de hypothesen van het attitudeonderzoek en het Likertschaalonderzoek besproken aan de hand van de relevante literatuur en een korte uitleg. Eerst zal hypothese 1 betreft het Likertschaalonderzoek worden behandeld. Daarna volgen hypothesen 2 tot en met 5 die betrekking hebben op het attitudeonderzoek. De volgorde van de hypothesen zal aangehouden worden in de resultaten die worden behandeld in hoofdstuk 5.

## 3.1 Hypothese Likertschaalonderzoek

Bij het opstellen van de hypothese voor het Likertschaalonderzoek moet rekening gehouden worden met het feit dat er op het gebied van Likertschalen veel tegengestelde conclusies zijn (zie hoofdstuk 2.3.3). Onder andere over de vraag of een neutrale antwoordoptie hogere of lagere scores genereert. Daarnaast zijn de verschillen in scores tussen onderzoeken met 6- en 7-puntsschalen kleiner dan bijvoorbeeld de verschillen tussen onderzoeken met een 4- en 7-puntsschaal. De verschillen zijn (met genoeg proefpersonen) veelal niet significant, en het gaat hierbij vaker om zogenaamde “trends” of een tendens dan om duidelijke conclusies. De hypothese voor dit onderzoek is dat er geen significant verschil is tussen de beschrijvende statistieken aan de hand van de enquêtes met 6 of 7 antwoordopties.

De hypothese van het Likertschaalonderzoek is de volgende:

1. Er wordt geen significant verschil gevonden tussen de data van de 6-puntsschaal en de 7-puntsschaal.

## 3.2 Hypothesen attitudeonderzoek

De hypothesen voor het attitudeonderzoek zijn grotendeels dezelfde als die van Pinget et al. (2013). Aangezien de classificatietask en het verschil tussen spontane en minder spontane spraak in deze scriptie niet voorkomt, kan slechts een subset van de hypothesen getest worden. Ook worden er twee hypothesen uit Pinget et al. (2013) samengevoegd, waardoor er vier overblijven. Daarnaast zijn er nog factoren zoals de variabelen sekse, leeftijd en opleidingsniveau, die als laatste besproken worden.

2. *De Randstadshypothese* : De Nederlandse uitspraaknorm ligt in de Randstad.

Grootschalige taalveranderingen worden in gang gezet door een “externe, buitentalige factor”, dit gebeurt wanneer taalvormen “... een bepaalde prestige ontleen aan het feit dat ze gebruikt worden door de bewoners van een invloedrijk centrum.” (Hoppenbrouwers 1990). De Randstad is het centrum van waaruit taalinnovaties zich hebben verspreid en is de plek waar het meest standaard Nederlands wordt gesproken. De norm ligt in het centrum (de Randstad), de non-standaard variëteiten bevinden zich in de periferie, en het midden van Nederland is een overgangsgebied (Hoppenbrouwers 1990). Doorgaans hebben mensen uit de Randstad het gevoel dat ze Standaardnederlands spreken, waar ze ook enige trots aan ontleen (Smakman 2006). In Pinget et al. (2013) werd in de resultaten een patroon zichtbaar waarbij de spraak uit de Randstad als het meest *standaard* werd beoordeeld, daarna volgden achtereenvolgens de regio’s Midden, Noorden en Zuiden.

3. *De Vertrouwdheidshypothese*. Spraak uit de eigen regio wordt positiever beoordeeld.

De Vertrouwdheidshypothese in deze scriptie wijkt af van Pinget et al. (2013), aangezien in dit onderzoek twee hypothesen uit het voorgenoemde onderzoek samen zijn gevoegd. In Pinget et al. (2013) werden twee hypothesen opgesteld die de invloed van de factor “beoordelaarsregio” behandelden. Te weten:

- *De vertrouwheidshypothese*, die is gebaseerd op de vergelijkbare “familiarity driven hypothesis” opgesteld door Van Bezooijen (2002). Aangezien men meer vertrouwd is met de spraak uit de eigen regio, waardeert men die ook meer en zal men de spraak uit eigen regio ook hoger beoordelen op de schaal *mooi*.
- Daarnaast stelden Pinget et al. (2013) ook de volgende hypothese op waarbij de beoordelaarsregio een factor was: *De regionale identiteitshypothese*, die stelt dat als luisteraars uit een bepaalde regio een voorkeur vertonen voor spraak van de eigen regio, dit het duidelijkst/eerst merkbaar zal zijn op de schaal *mooi*. Door een hernieuwde interesse in de regionale taalvariëteiten en een opbloei van de regionale identiteit in een tijd van globalisatie, is men in de verschillende regio’s trotser op eigen regionaal accent. Deze veranderingen kunnen invloed hebben op de gegevens uit het attitudeonderzoek.

Aangezien beide hypothesen uit Pinget et al. (2013) uitsluitend de factor “beoordelaarsregio” behandelden, is ervoor gekozen om deze hypothesen samen te voegen.

Uit het onderzoek van Pinget et al. (2013) kwam naar voren dat er inderdaad een significant effect optrad bij de factor “beoordelaarsregio” op de schaal *mooi*. (Hetgeen Van Bezooijen (2002) niet kon aantonen in haar onderzoek). In Pinget et al. (2013) bleek dit voor de beoordelaarsregio Zuid het geval te zijn. Ook was er een effect van de beoordelaarsregio op de andere drie schalen, waarbij beoordelaars uit het Zuiden positiever waren ten opzichte van eigen spraak dan de beoordelaars uit de andere regio’s waren over dezelfde stimuli.

4. *De regionale stereotypenhypothese*: Sprekers uit geografisch perifere gebieden worden minder standaard gevonden dan sprekers uit niet-perifere gebieden buiten de Randstad.

Deze hypothese heet de regionale stereotypenhypothese omdat er in de Nederlandse taalgemeenschap voornamelijk stereotypen worden gevormd over geografisch perifere gebieden, zoals Brabant, Limburg, Twente, Friesland, Groningen, Drenthe en Zeeland. Door herhaaldelijk te worden geconfronteerd met de taalvariëteiten uit de perifere gebieden via reclame en films, worden deze variëteiten herkend en gekoppeld aan een negatief taaloordeel. Het ligt in lijn der verwachting dat het oordeel over niet-perifere gebieden zoals het midden van Nederland minder negatief zal zijn, aangezien er in mindere mate sprake is van stereotypevorming in bijvoorbeeld de media.

5. *De uniforme normhypothese*: Nederlanders hanteren dezelfde Nederlandse uitspraaknorm ongeacht hun regionale afkomst.

Deze hypothese lijkt op een tegenstelling van de eerder geformuleerde vertrouwdheidshypothese. In Pinget et al. (2013) wordt de hypothese nauwer gedefinieerd; slechts de hoogopgeleide Nederlanders hanteren deze Nederlandse uitspraaknorm. In het onderzoek van Grondelaers et al.

(2010) gaan de onderzoekers zover dat dezelfde uitspraaknorm geldt voor alle beoordelaars, ongeacht sekse, leeftijd en opleidingsniveau. Oftewel, de onderzoekers stellen dat er in de Nederlandse taalgemeenschap (in ieder geval vanaf een zekere leeftijd) een standaardattitude bestaat tegenover de uitspraak van sprekers uit verschillende regio's.

Naast deze 4 hypothesen is er nog een aantal factoren dat in het attitudeonderzoek in deze scriptie wordt getoetst. Dit zijn de externe variabelen leeftijd, sekse en opleidingsniveau. Aangenomen wordt dat oudere sprekers meer tolerant worden voor non-standaard varianten van hun taal ten opzichte van jongere sprekers (Cargile et al. 1994). Ook wordt aangenomen dat vrouwen beter zijn in het observeren van spraakstijlen en uitspraakverschillen (Smakman 2006). Het observeren van de verschillende kleuringen in het Standaardnederlands kan tot een afwijkende beoordeling leiden ten opzichte van de mannelijke beoordelaars. En aangezien lageropgeleiden vaker tot de lagere sociale klasse behoren, en de lagere sociale klasse vaker een non-standaardvariëteit spreekt (Smakman 2006), is de invloed van opleidingsniveau ook in dit onderzoek opgenomen. In Grondelaers et al. (2010) werd geconcludeerd dat de demografische achtergrond van de beoordelaars slechts kleine invloed heeft op de antwoorddata (hoewel de jongere beoordelaars meer negatief waren dan de oudere beoordelaars tegenover de sprekers uit de regio Zuid).

## 4. Methode

Zoals eerder genoemd is de onderzoeksmethode die in dit onderzoek is gebruikt, behoudens een klein aantal verschillen, vergelijkbaar met de methode van Pinget et al. (2013). De beoordeelde sprekers (stimuli) zijn vergelijkbaar met Grondelaers et al (2010). Dit hoofdstuk zal beginnen met het bespreken van de enquête en waarin deze verschilt van die van Pinget et al. (2013). Daarna worden de stimuli besproken, oftewel de sprekers van het Standaardnederlands die worden beoordeeld. In de paragraaf die daarop volgt zullen deze beoordelaars worden behandeld. Tot slot volgt de verwerking van de resultaten.

### 4.1 De internetenquête

In het onderzoek van deze scriptie is gebruik gemaakt van twee enquêtes. Het enige verschil tussen de twee enquêtes is dat de ene enquête gebruik maakte van 6-puntsschalen en de andere van 7-puntsschalen. Deze zijn gemaakt op het internet in het taalevaluatieprogramma CLEO en ze bestaan uit persoonsvragen en beoordelingsvragen. De persoonsvragen waren vragen naar leeftijd, geslacht, opleidingsniveau, herkomst en wat de moedertaal van de beoordelaar was. Daarna kregen de beoordelaars twee oefenvragen om te beantwoorden, de resultaten hiervan werden niet meegeteld. Vervolgens kregen de beoordelaars tijdens het officiële attitudeonderzoek in een willekeurige volgorde 20 stimuli te horen, waarbij de stimuli moesten worden beoordeeld op schalen met de volgende labels: *standaard*, *mooi*, *regionaal* en *verzorgd*. De vraagstelling was direct. De beoordelaars werden net als de sprekers in vier regio's verdeeld: Noord, Midden, Randstad en Zuid. (De exacte verdeling van de vier regio's wordt in de paragraaf 4.3 "Beoordelaars" nog uitgebreider besproken).

### 4.2 Stimuli/sprekers

De beoordelaars krijgen tijdens de beoordelingsvragen in een willekeurige volgorde 20 stimuli (5 sprekers per regio) van ongeveer 20 seconden te horen die spontane spraak bevatte van docenten uit het CGN. De beoordelaars beantwoordden vragen over oudere, mannelijke sprekers. De sprekers waren hoogopgeleide leerkrachten Nederlands uit de sociale middenklasse. Ze zijn vóór hun achtste levensjaar in de desbetreffende regio komen wonen, waar ze opgroeiden tot minimaal hun achttiende. De sprekers werkten op een middelbare school, in dezelfde middelgrote gemeente waar ze ook woonden (of ze woonden in de buurt, maar wel in dezelfde regio).

De reden waarom oudere sprekers zijn gekozen is voornamelijk om te onderzoeken wat de invloed van de variabele "leeftijd" is op de resultaten en om ze op deze manier te kunnen vergelijken met Pinget et al. (2013). Er werden slechts mannelijke sprekers gekozen, in verband met de consistentie met de andere onderzoeken. De stimuli die in dit onderzoek zijn gebruikt, komen meer overeen met het onderzoek van Grondelaers et al. (2010). Echter werd er in het onderzoek van Grondelaers et al. (2010) gebruik gemaakt van twee sprekers, en in onderhavige scriptie gebruikt gemaakt van vijf sprekers. Pinget et al. (2013) maakten ook gebruik van vijf sprekers per regio, en noemen het kleine aantal sprekers in het onderzoek van Grondelaers et al. (2010) als mogelijke reden voor de verschillen tussen de beide onderzoeken.



### 4.3 Beoordelaars

Er zijn bij de beide enquêtes beoordelaars geweest die uit de antwoorddata verwijderd moesten worden vanwege het incompleet invullen van de test of omdat er soms op elke vraag éénzelfde antwoord werd gegeven, wat duidde op het niet serieus nemen van het onderzoek. De enquêtes zijn correct ingevuld door 199 Personen, waarvan 99 personen de 6-puntschaalenquête hebben ingevuld, en 100 personen de 7-puntsschaalenquête. In tabel 1 worden de aantallen beoordelaars weergegeven per herkomstregio en per enquête.

**Tabel 1.** *Kruistabel voor de regio's Midden, Noord, Randstad, Zuid en totaal met variabelen sekse en leeftijd (2 missing values voor "leeftijd" in de 7-puntsschaal)*

<b>6-Puntsschaal</b>		Midden	Noord	Randstad	Zuid	Totaal
Man	<30 jaar	2	6	1	1	10
	>30 jaar	2	19	5	4	30
Vrouw	<30 jaar	4	11	8	11	34
	>30 jaar	4	10	2	9	25
Totaal		12	46	16	25	99
<b>7-Puntsschaal</b>		Midden	Noord	Randstad	Zuid	Totaal
Man	<30 jaar	1	4	3	0	8
	>30 jaar	11	12	3	6	32
Vrouw	<30 jaar	6	8	6	5	25
	>30 jaar	4	15	10	4	33
Totaal		22	39	22	15	98

De enquêtes werden verspreid door middel van e-mail en sociale media (Facebook). Hierbij moest met een aantal zaken rekening gehouden worden. De verspreiding van de enquêtes moest proportioneel goed verdeeld zijn, i.e. de beide enquêtes moesten ongeveer evenveel beoordelaars hebben en de enquêtes moesten een degelijke spreiding vertonen wat betreft de herkomstregio, sekse, leeftijd en opleidingsniveau.

Om de verspreiding zo goed mogelijk te laten gebeuren, is een persoonlijk e-mailbestand van potentiële beoordelaars (ongeveer 400 mailadressen van mannen en vrouwen onder de 30 jaar, zowel werkenden als studerende) in tweeën gedeeld. Daarnaast zijn de twee enquêtes via de sociale media naar verschillende potentiële beoordelaars gestuurd. Het werven via de sociale media werd gedaan via een Facebook-evenement, waarbij de deelnemers werden aangespoord andere mensen uit te nodigen mee te doen. Als laatste is de enquête doorgestuurd aan 200 studenten van de Universiteit Utrecht om extra beoordelaars uit voornamelijk het midden van het land te werven.

Uiteindelijk bleek uit de data dat een overgrote meerderheid van de beoordelaars hoger opgeleid was (149 hoogopgeleid tegenover 45 laagopgeleid, 5 missing values). De externe factor "opleidingsniveau" kon door dit grote verschil statistisch niet verdedigd worden als voldoende betrouwbaar voor verdere analyse en is daarom uit de resultaten weggelaten.

In tabel 1 is te zien dat de cellen van de regio Noord het beste zijn gevuld wat betreft de opdeling in leeftijd en sekse. In het onderzoek zijn de factoren "leeftijd" en "sekse" dan ook slechts bij de regio Noord getest, om de statistische betrouwbaarheid van dit onderzoek te waarborgen. In de resultaten in hoofdstuk 5 wordt er verder ingegaan op de verwerking van het

repeated measures-onderzoeksmoedel waarbij de factoren "leeftijd" en "seks" uitsluitend worden behandeld voor de regio Noord.

De beoordelaars zijn ingedeeld in vier herkomstregio's, overeenkomstig de postcode, in overeenstemming met Grondelaers et al. (2010) en Pinget et al. (2013). De postcodes zijn als volgt gecodeerd:

- Midden: 3700-4299, 5300-5399, 6500-6999, 7200-7239, 7300-7399, 8200-8399
- Noord: 7000-7199, 7240-7299, 7400-8199, 8400-9999
- Randstad: 1000-3699
- Zuid: 4300-5299, 5400-6499

#### 4.4 Verwerking van de resultaten

De verwerking van de resultaten is gedaan in SPSS. De data van beide enquêtes zijn niet normaal verdeeld, maar worden in dit onderzoek wel met behulp van non-parametrische statistische tests behandeld. Deze werkwijze met non-parametrische statistische tests is mogelijk, in tegenstelling tot het vaak beweerd argument dat hierdoor de uitkomsten onbetrouwbaarder en ongeldig zouden zijn (Carifio & Perla 2007). "Voor Likertschalen met 5 categorieën hebben de t-test en de Mann-Whitney u-test een vergelijkbare power en onderzoekers hoeven zich geen zorgen te maken over het vinden van verschillen als deze zich ook niet in de populatie voordoen." (De Winter & Dodou 2010).

In de data heeft twee keer hercodering plaatsgevonden. Allereerst zijn de data van de schaal *regionaal* omgedraaid in beide datasets. Aangezien een hogere score bij de andere schalen een positievere attitude betekende, was een hercodering nodig. De schaal *regionaal* is vanaf nu *niet-regionaal*. De inversie van de schaal betekent dat een score 6 op de 6-puntsschaal in de data een 1 werd, en de score 5 een 2, etc. zoals hieronder weergegeven.

6-puntsschaal: 1 / 2 / 3 / 4 / 5 / 6 ==> 6 / 5 / 4 / 3 / 2 / 1

7 puntsschaal: 1 / 2 / 3 / 4 / 5 / 6 / 7 ==> 7 / 6 / 5 / 4 / 3 / 2 / 1

Aangezien de 6- en 7-puntsschalen moesten worden vergeleken, is er nogmaals gebruik gemaakt van een hercodering van de data. De beide schalen zijn opgedeeld rondom het midden, oftewel de 0, met aan beide kanten uitersten tot -3 en 3. Dit is in feite een standaardisering van beide datasets rondom de 0, zoals bij een z-score ook het geval is. De hercodering is hieronder te zien.

6 puntsschaal: 1 / 2 / 3 / 4 / 5 / 6 ==> -3 / -1,8 / -0,6 / 0,6 / 1,8 / 3

7 puntsschaal: 1 / 2 / 3 / 4 / 5 / 6 / 7 ==> -3 / -2 / -1 / 0 / 1 / 2 / 3

Deze hercodering is gedaan in navolging van een betrouwbaarheidsonderzoek wat betreft Likertschalen met verschillende aantallen Likertschaalcategorieën. Daaruit bleek dat een hercodering naar z-scores een even grote mate van constantheid en betrouwbaarheid opleverde als meer gecompliceerde statistische analyses, zoals power functies en third-order polynomials (Colman et al. 1997).

De hercodeerde data hebben dezelfde beschrijvende statistieken als de ruwe data (data vóór hercodering), en daarmee dezelfde implicaties (standard error, skewness en kurtosis zijn identiek).

De hercodering naar de schaal van -3 tot en met 3 heeft (naast de mogelijkheid tot vergelijken) als voordeel dat de twee datasets gecombineerd kunnen worden en er generalisaties gemaakt kunnen worden over deze gecombineerde data. Een ander voordeel is de eenvoudige leesbaarheid (>0 is positief oordeel, <0 is negatief). De verwerking van de resultaten voor zowel het attitudeonderzoek als het Likertschaalonderzoek is gedaan aan de hand van deze hercodering.

## 5. Resultaten

In dit hoofdstuk worden de resultaten van het onderzoek behandeld. Als eerste worden de correlaties en de betrouwbaarheid van het onderzoek besproken. Daarna volgen de resultaten van het repeated measures design, waarbij de significante effecten van de factoren op de 6-puntsschaaldata (6-P), 7-puntsschaaldata (7-P) en de gecombineerde data (GD = 6-P+7-P) worden getoond. Zowel de 6-P en 7-P zijn nodig om te zien of er verschillen zijn tussen de twee. Daarom zijn de gecombineerde data de enige data waarbij de factor “6- of 7-puntsschaal” getest kan worden; een factor die slechts het Likertschaalonderzoek betreft. Uit deze data volgen eerst de resultaten van het Likertschaalonderzoek. Hierna worden de resultaten van het attitudeonderzoek behandeld, waarbij de 6- en 7-puntsschaal en de gecombineerde data afzonderlijk van elkaar behandeld worden.

### 5.1 Correlaties tussen de schalen en betrouwbaarheid

In hoofdstuk 2 werd al vermeld dat de schalen *mooi*, *niet-regionaal* en *verzorgd* sterk met de schaal *standaard* correleren. Dit is uit dit onderzoek nogmaals gebleken aan de hand van een Pearson’s test. Ook de Pearson correlaties tussen de andere schalen zijn in tabel 2 te vinden. De correlatiescores tussen *niet-regionaal* en *verzorgd* zijn laag, maar wel significant. De correlatiescores tussen *mooi* en *verzorgd* zijn opvallend hoog. (In alle gevallen  $p < 0,00$ ).

**Tabel 2.** Pearson Correlatiescores tussen de schalen *standaard*, *mooi*, *niet-regionaal* en *verzorgd*

Pearson Correlaties	6-puntsschaal	7-puntsschaal	Gecombineerde data
	r	r	r
Standaard + Mooi	0,533	0,497	0,515
Standaard + Niet-regio	0,491	0,434	0,462
Standaard + Verzorgd	0,413	0,396	0,404
Mooi + Niet-regio	0,288	0,257	0,272
Mooi + Verzorgd	0,665	0,652	0,658
Niet-regio + Verzorgd	0,182	0,134	0,158

De betrouwbaarheid van de vier schalen bij de 6-P, 7-P en GD is getoetst door middel van Cronbach’s  $\alpha$ . Bij de 6-puntsschaal waren de waarden tussen de 0,732 en de 0,874. Bij de 7-puntsschaal tussen de 0,774 en de 0,871. En bij de gecombineerde data lagen de waarden tussen de 0,755 en de 0,872. Een Cronbachs  $\alpha$  van 7 wordt over het algemeen beschouwd als voldoende hoog voor verdere statistische analyse. De waardes van de betrouwbaarheid staan weergegeven in tabel 3.

**Tabel 3.** Cronbach’s Alfa van de 6-P, 7-P en GD

Cronbach’s Alfa	6-puntsschaal	7-puntsschaal	Gecombineerde data
Standaard	0,874	0,871	0,872
Mooi	0,801	0,844	0,825
Niet-regionaal	0,732	0,774	0,755
Verzorgd	0,846	0,857	0,851

## 5.2 Resultaten factoren repeated measures-model

De statistische analyse is opgesplitst in twee statistische repeated measures-tests. Deze twee repeated measures-onderzoeksmoedellen worden in deze paragraaf besproken.

- Eerst worden de resultaten van de factoren "beoordelaarsregio" en "6- of 7-puntsschaal" getoond van het repeated measures-model uitgevoerd op alle regio's. Hierin worden deze factoren als between-factoren behandeld en de "sprekersregio" als de repeated within-factor (er waren immers 5 verschillende sprekers per regio). De factor "6- of 7-puntsschaal" is slechts getest in de gecombineerde data, aangezien het hier gaat om een factor die de verschillen in de data van beide enquêtes moest onderzoeken.

- Daarna volgt de bespreking van het repeated measures-model dat op slechts de regio Noord is uitgevoerd. Zoals eerder is genoemd, is deze keuze gemaakt vanwege de goede spreiding bij deze regio vergeleken met de andere regio's bij de variabelen "leeftijd" en "seks". De andere regio's, die een minder goede spreiding kenden, konden de betrouwbaarheid van het onderzoek niet garanderen. De factoren die op de data van de regio Noord worden getest zijn de between-factoren "leeftijd" en "seks".

In tabel 4 zijn alle resultaten te zien uit het eerste gedeelte van de statistische analyse van het repeated measures-model waarbij de factoren "beoordelaarsregio" en "6-of 7-puntsschaal" werden getest ( $p < 0,05$ , Huynh-Feldt-test).

**Tabel 4.** Resultaten van het eerste gedeelte van de statistische analyse. Het gaat hierbij om een repeated measures-model met de within-factor "sprekersregio" en de between-factoren "6-of 7-puntsschaal" en "beoordelaarsregio". De significante effecten zijn aangegeven met een grijze rij.

	df <sub>1</sub>	df <sub>2</sub>	F	p		df <sub>1</sub>	df <sub>2</sub>	F	p
<b>Sprekersregio</b>									
Standaard					Niet-Regionaal				
6-punts	2,683	154,668	67,714	0,000	6-punts	2,948	304,593	84,384	0,000
7-punts	3,000	163,440	68,387	0,000	7-punts	2,838	305,425	101,143	0,000
GD	2,874	324,071	140,616	0,000	GD	2,880	609,490	182,861	0,000
Mooi					Verzorgd				
6-punts	2,349	30,124	8,759	0,000	6-punts	2,682	24,238	10,508	0,000
7-punts	2,484	23,456	7,788	0,000	7-punts	2,867	19,614	9,858	0,000
GD	2,410	53,080	16,444	0,000	GD	2,834	42,259	20,152	0,000
<b>Sprekersregio*6- of 7-puntsschaal</b>									
Standaard					Niet-Regionaal				
GD	2,874	2,544	1,104	0,346	GD	2,880	3,962	1,189	0,313
Mooi					Verzorgd				
GD	2,410	0,824	0,255	0,815	GD	2,834	0,728	0,347	0,780
<b>Sprekersregio*Beoordelaarsregio</b>									
Standaard					Niet-Regionaal				
6-punts	8,049	2,317	1,014	0,426	6-punts	8,844	7,351	2,037	0,036
7-punts	9,000	2,538	1,062	0,391	7-punts	8,514	5,320	1,762	0,080
GD	8,622	1,804	0,783	0,627	GD	8,641	1,305	0,391	0,935
Mooi					Verzorgd				
6-punts	7,048	3,280	0,954	0,466	6-punts	8,045	2,250	0,975	0,456
7-punts	7,453	3,272	1,086	0,373	7-punts	8,601	3,307	1,662	0,102
GD	7,230	2,465	0,764	0,622	GD	8,503	3,878	1,849	0,061
<b>Beoordelaarsregio</b>									
Standaard					Niet-Regionaal				
6-punts	3,000	29,732	2,913	0,038	6-punts	3,000	4,799	0,551	0,649
7-punts	3,000	2,931	0,247	0,863	7-punts	3,000	35,291	3,766	0,013
GD	3,000	17,676	1,579	0,196	GD	3,000	16,734	1,851	0,139
Mooi					Verzorgd				
6-punts	3,000	16,302	2,140	0,100	6-punts	3,000	33,702	3,741	0,014
7-punts	3,000	6,424	0,641	0,590	7-punts	3,000	8,064	0,755	0,522
GD	3,000	14,698	1,666	0,176	GD	3,000	28,059	2,849	0,039
<b>6- of 7-puntsschaal</b>									
Standaard					Niet-Regionaal				
GD	1,000	16,361	1,461	0,228	GD	1,000	0,056	0,006	0,937
Mooi					Verzorgd				
GD	1,000	0,373	0,042	0,837	GD	1,000	1,849	0,188	0,665

- De within-factor "sprekersregio" is een significant effect bij alle schalen in alle datasets.
- De factor "6- of 7-puntsschaal" vertoonde bij geen enkele schaal een significant effect.
- De interactie "sprekersregio\*beoordelaarsregio" zorgde bij de 6-puntsschaaldata voor een significant effect op de schaal *niet-regionaal*.

Tabel 5 laat de resultaten zien uit het repeated measures-model waarbij de between-factoren "sekse" en "leeftijd" op de data van de regio Noord werden getest ( $p < 0,05$ , Huynh-Feldt-test). Gezien de spreiding van de leeftijdsgroepen en de spreiding van de mannelijke en vrouwelijke beoordelaars (zie ook tabel 1), zijn de gecombineerde data het meest betrouwbaar. De 6-P en 7-P van de regio Noord bevatten te weinig beoordelaars per cel om betrouwbare conclusies op te baseren, daarom wordt er uitsluitend gebruik gemaakt van de gecombineerde data in de verdere data-analyse wat betreft de factoren "sekse" en "leeftijd".

**Tabel 5.** Resultaten van het tweede gedeelte van de statistische analyse. Het gaat hierbij om een repeated measures-model uitgevoerd op de regio Noord met de within-factor "sprekersregio" en de between-factoren "leeftijd", "seks" en "6- of 7-puntsschaal". De significante effecten zijn aangegeven met een grijze rij.

	df <sub>1</sub>	df <sub>2</sub>	F	P		df <sub>1</sub>	df <sub>2</sub>	F	p
<b>Sprekersregio</b>									
Standaard					Niet-Regionaal				
6-punts	3,000	52,011	27,191	0,000	6-punts	3,000	165,327	45,945	0,000
7-punts	3,000	64,515	23,765	0,000	7-punts	2,841	99,984	25,496	0,000
GD	3,000	114,092	50,100	0,000	GD	2,885	250,354	65,936	0,000
Mooi					Verzorgd				
6-punts	2,407	11,724	3,317	0,032	6-Punts	2,831	13,285	5,239	0,002
7-punts	2,999	15,098	5,608	0,001	7-Punts	3,000	3,217	1,751	0,161
GD	2,661	21,187	6,785	0,000	GD	3,000	12,543	5,861	0,001
<b>Sprekersregio*Leeftijd</b>									
Standaard					Niet-Regionaal				
6-punts	3,000	3,228	1,687	0,173	6-punts	3,000	7,620	2,118	0,101
7-punts	3,000	2,147	0,791	0,502	7-punts	2,841	2,610	0,666	0,567
GD	3,000	0,570	0,250	0,861	GD	2,885	2,708	0,713	0,540
Mooi					Verzorgd				
6-punts	2,407	11,712	3,314	0,032	6-punts	2,831	7,750	3,056	0,034
7-punts	2,999	4,161	1,546	0,207	7-punts	3,000	1,738	0,946	0,421
GD	2,661	14,073	4,507	0,006	GD	3,000	7,525	3,516	0,016
<b>Sprekersregio*Sekse</b>									
Standaard					Niet-Regionaal				
6-punts	3,000	5,480	2,865	0,039	6-punts	3,000	4,854	1,349	0,262
7-punts	3,000	1,366	0,503	0,681	7-punts	2,841	1,699	0,433	0,719
GD	3,000	3,925	1,723	0,163	GD	2,885	4,360	1,148	0,330
Mooi					Verzorgd				
6-punts	2,407	2,642	0,748	0,499	6-punts	2,831	5,261	2,075	0,111
7-punts	2,999	1,507	0,560	0,643	7-punts	3,000	2,693	1,466	0,228
GD	2,661	2,575	0,825	0,469	GD	3,000	0,765	0,357	0,784
<b>Sprekersregio*Leeftijd*Sekse</b>									
Standaard					Niet-Regionaal				
6-punts	3,000	0,689	0,360	0,782	6-punts	3,000	5,755	1,599	0,193
7-punts	3,000	8,193	3,018	0,033	7-punts	2,841	1,099	0,280	0,829
GD	3,000	7,022	3,083	0,028	GD	2,885	1,527	0,402	0,744
Mooi					Verzorgd				
6-punts	2,407	0,487	0,138	0,905	6-punts	2,831	1,557	0,614	0,598
7-punts	2,999	7,657	2,844	0,041	7-punts	3,000	6,599	3,592	0,016
GD	2,661	4,294	1,375	0,253	GD	3,000	1,776	0,830	0,479
<b>Sprekersregio* 6- of 7-puntsschaal</b>									
Standaard					Niet-Regionaal				
GD	3,000	3,969	1,743	0,159	GD	2,885	11,045	2,909	0,037
Mooi					Verzorgd				
GD	2,661	7,218	2,312	0,085	GD	3,000	2,066	0,965	0,410
<b>6- of 7-puntsschaal</b>									
Standaard					Niet-Regionaal				
GD	1,000	0,792	0,069	0,794	GD	1,000	40,506	4,096	0,046
Mooi					Verzorgd				
GD	1,000	5,787	0,743	0,391	GD	1,000	5,254	0,581	0,448
<b>Leeftijd</b>									
Standaard					Niet-Regionaal				
6-punts	1,000	9,209	0,852	0,361	6-punts	1,000	10,025	1,016	0,319
7-punts	1,000	12,000	0,972	0,331	7-punts	1,000	8,032	0,810	0,374
GD	1,000	0,343	0,030	0,863	GD	1,000	0,001	0,000	0,993
Mooi					Verzorgd				
6-punts	1,000	2,286	0,408	0,526	6-punts	1,000	0,020	0,003	0,958
7-punts	1,000	38,765	3,722	0,062	7-punts	1,000	7,003	0,620	0,436
GD	1,000	32,106	4,122	0,046	GD	1,000	4,309	0,476	0,492
<b>Sekse</b>									
Standaard					Niet-Regionaal				
6-punts	1,000	4,312	0,399	0,531	6-punts	1,000	0,365	0,037	0,848
7-punts	1,000	0,457	0,037	0,848	7-punts	1,000	50,308	5,076	0,031
GD	1,000	3,542	0,308	0,581	GD	1,000	32,656	3,302	0,073
Mooi					Verzorgd				
6-punts	1,000	1,302 <sup>F</sup> -005	0,000	0,999	6-punts	1,000	0,412	0,057	0,812
7-punts	1,000	1,455	0,140	0,711	7-punts	1,000	0,852	0,076	0,785
GD	1,000	0,821	0,105	0,746	GD	1,000	1,247	0,138	0,712

Onderstaande punten gelden alleen voor de regio Noord:

- De within-factor “sprekersregio” is een significant effect bij alle datasets, behalve bij de 7-P op de schaal *verzorgd*.
- De factor “leeftijd” geeft in de interactie met de within-factor “sprekersregio” een significant effect bij de schalen *mooi* en *verzorgd* bij de GD.
- De interactie tussen “sprekersregio”, “leeftijd” en “seks” geeft een significant effect op de schaal *standaard* bij de resultaten van de gecombineerde data.
- De factor “6- of 7-puntsschaal” zorgt bij de beoordelaarsregio Noord voor een significant effect bij de schaal niet-regionaal. Dit is een hoofdeffect voor deze factor. Dezelfde schaal geeft ook een significant effect van de interactie tussen “6- of 7-puntsschaal” en “sprekersregio”. Het zelfde effect werd niet gevonden bij tabel 4 met de data van de eerste repeated measures-test, gemeten over de gehele beoordelaarsgroep.
- Er is een hoofdeffect voor de factor “leeftijd” bij de schaal *mooi* in de gecombineerde data.

### 5.3 Resultaten Likertschaalonderzoek

In deze paragraaf zullen de resultaten in tabel 4 en 5 die betrekking hebben op het Likertschaalonderzoek en hypothese 1, worden behandeld. In de vorige paragraaf was in tabel 4 te zien dat er uit de repeated measures-test geen significant effect optrad bij de interactie tussen de within-factor “sprekersregio” en de between-factor “6- of 7-puntsschaal”. Dit betekent dat er in dit onderzoeksmodel geen significante verschillen tussen de 6- en de 7-puntsschaal zijn aangetoond.

In tabel 6 zijn de gemiddelden op de schalen *standaard*, *mooi*, *niet-regionaal* en *verzorgd* van beide datasets weergegeven.

**Tabel 6.** Gemiddelde scores 6-P en 7-P per schaal

Gemiddelden van de schaal:	6-puntsschaal	7-puntsschaal
Standaard	-0,067	-0,098
Mooi	-0,061	-0,026
Niet-regionaal	-0,457	-0,360
Verzorgd	0,385	0,392

Het tweede repeated measures-model, waarbij de factoren “seks” en “leeftijd” zijn getest, resulteerde wel in een hoofdeffect voor de factor “6- of 7-puntsschaal” en een significant effect bij de interactie tussen “sprekersregio” en “6- of 7-puntsschaal”. Dit gebeurde slechts op de schaal *niet-regionaal* en werd getest op de beoordelaarsregio Noord (zie tabel 5). Hierbij moet aangetekend worden dat het hier gaat om een kleinere subgroep van de totale groep beoordelaars.



## 5.4 Resultaten attitudeonderzoek

In deze paragraaf worden alle resultaten van het attitudeonderzoek besproken. Eerst wordt in deze paragraaf een overzichtstabel gegeven met daarin de gemiddelden van alle resultaten op alle vier schalen van de drie verschillende datasets. Daarna volgt een sectie waarbij alle attitudeonderzoeksresultaten per schaal worden behandeld alsmede de significante effecten die in tabel 4 en 5 zijn weergegeven, met behulp van Tukey's range tests en gepaarde t-tests.

In tabel 8 zijn de resultaten te zien van de gemiddelden van de schalen *standaard*, *mooi*, *niet-regionaal* en *verzorgd* per regio, per herkomst van de beoordelaar en per 6-,7-puntsschaal en de gecombineerde data.

**Tabel 8.** Resultaten gemiddelden 6-P, 7-P en GD op schalen standaard, mooi, niet-regionaal en verzorgd per beoordelaarsregio

		6-puntsschaal (N=99)					7-puntsschaal (N=100)					Beide schalen (N=199)						
	Beoordelaars						Beoordelaars						Beoordelaars					
	Sprekers Aantal:	M	N	R	Z	GmS	M	N	R	Z	GmS	M	N	R	Z	GmS		
<b>Standaard</b>	M	12	46	16	25		23	40	22	15		35	86	38	40			
	M	0,54	-0,22	0,08	-0,11	-0,05	0,04	0,11	-0,38	-0,05	-0,04	0,21	-0,07	-0,19	-0,09	-0,05		
	N	0,56	-0,07	0,03	0,17	0,08	0,13	0,33	0,04	0,19	0,20	0,28	0,12	0,03	0,18	0,14		
	R	1,24	0,36	0,77	0,48	0,56	0,44	0,36	0,52	0,45	0,43	0,72	0,36	0,62	0,47	0,50		
	Z	-0,40	-0,85	-0,83	-1,12	-0,86	-0,95	-1,04	-1,05	-0,79	-0,98	-0,76	-0,94	-0,96	-0,99	-0,92		
	GmB	0,49	-0,20	0,01	-0,15	-0,07	-0,08	-0,06	-0,22	-0,05	-0,10	0,11	-0,13	-0,12	-0,11	-0,08		
<b>Mooi</b>	Sprekers	M	N	R	Z	GmS	M	N	R	Z	GmS	M	N	R	Z	GmS		
	M	0,16	-0,33	-0,12	0,02	-0,15	-0,17	0,06	-0,27	0,23	-0,04	-0,05	-0,15	-0,21	0,10	-0,09		
	N	0,16	-0,42	0,11	0,02	-0,15	-0,13	-0,04	-0,25	0,35	-0,05	-0,03	-0,24	-0,10	0,15	-0,10		
	R	0,74	0,07	0,29	0,37	0,26	0,40	0,22	0,11	0,35	0,26	0,52	0,14	0,18	0,36	0,26		
	Z	-0,16	-0,19	-0,26	-0,24	-0,21	-0,17	-0,40	-0,18	-0,21	-0,27	-0,16	-0,29	-0,21	-0,23	-0,24		
	GmB	0,23	-0,22	0,00	0,05	-0,06	-0,02	-0,04	-0,15	0,18	-0,03	0,07	-0,13	-0,09	0,09	0,04		
<b>Niet-Regionaal</b>	Sprekers	M	N	R	Z	GmS	M	N	R	Z	GmS	M	N	R	Z	GmS		
	M	-0,18	-0,65	-0,24	-0,23	-0,42	-0,70	0,09	-0,92	-0,41	-0,39	-0,52	-0,31	-0,63	-0,30	-0,40		
	N	-0,02	-0,43	-0,40	-0,01	-0,27	-0,29	0,22	-0,17	-0,36	-0,07	-0,20	-0,13	-0,27	-0,14	-0,17		
	R	0,70	0,67	0,23	0,30	0,51	0,46	0,55	0,30	0,48	0,46	0,54	0,61	0,27	0,37	0,48		
	Z	-1,42	-1,54	-1,64	-1,95	-1,64	-1,39	-1,23	-1,88	-1,43	-1,44	-1,40	-1,40	-1,78	-1,76	-1,54		
	GmB	-0,23	-0,49	-0,51	-0,47	-0,46	-0,48	-0,09	-0,67	-0,34	-0,36	-0,39	-0,31	-0,60	-0,46	-0,41		
<b>Verzorgd</b>	Sprekers	M	N	R	Z	GmS	M	N	R	Z	GmS	M	N	R	Z	GmS		
	M	0,22	-0,09	0,39	0,22	0,10	0,11	0,25	0,02	0,37	0,18	0,15	0,07	0,17	0,27	0,14		
	N	0,38	-0,07	0,72	0,48	0,25	0,29	0,25	0,14	0,73	0,31	0,32	0,08	0,38	0,57	0,28		
	R	0,80	0,41	0,69	0,95	0,64	0,77	0,52	0,43	0,99	0,62	0,78	0,46	0,54	0,96	0,63		
	Z	0,42	0,36	1,01	0,67	0,55	0,50	0,31	0,66	0,45	0,45	0,48	0,34	0,81	0,59	0,50		
	GmB	0,45	0,15	0,70	0,58	0,38	0,42	0,33	0,31	0,64	0,39	0,43	0,24	0,48	0,60	0,39		

GmB = Gemiddelden beoordelaarsregio, GmS= Gemiddelden Sprekersregio

- De scores die de beoordelaars uit de regio Midden gaven op de 6-puntsschaalenquête waren opvallend hoog vergeleken met de andere beoordelaarsregio's (behalve op de schaal *verzorgd*). Dit kan te maken hebben met de ongelijke verdeling van de beoordelaars uit de verschillende regio's (zie tabel 1). Bij de paragraaf 5.4.1 "Standaard" wordt hier nog verder op ingegaan.
- Alle beoordelaars gaven gemiddeld de spraak uit de regio Noord op de schalen *standaard*, *niet-regionaal* en *verzorgd* een positievere waardering dan de spraak uit de regio Midden.
- De spraak uit de Randstad wordt op alle schalen het meest positief beoordeeld. De regio Zuid het minst positief. Alleen bij de schaal *verzorgd* scoorde de sprekersregio Zuid de op één na hoogste waardering, direct achter de regio Randstad.

### 5.4.1 Standaard

De spraak van de sprekers uit de Randstad (GD: 0,50) wordt het meest standaard gevonden, gevolgd door de regio's Noord (GD: 0,14) en vervolgens de regio Midden (GD: -0,05). De spraak uit de regio Zuid (GD: -0,92) wordt het minst standaard beoordeeld.

Aan de hand van gepaarde t-tests (met Bonferroni-correctie  $p < 0,017$ ) zijn de verschillen getoetst tussen de gemiddelde scores die gegeven zijn aan de vier afzonderlijke sprekersregio's. Hieruit blijkt dat alle sprekersregio's significant van elkaar verschillen in de GD op de schaal *standaard*. Dit is te zien in tabel 9. De significante verschillen zijn aangegeven door middel van een grijze rij.

**Tabel 9.** Data van gepaarde t-tests. Gepaarde sprekersregio's bij schaal *standaard*. De standaarddeviatie is die van de verschillscore

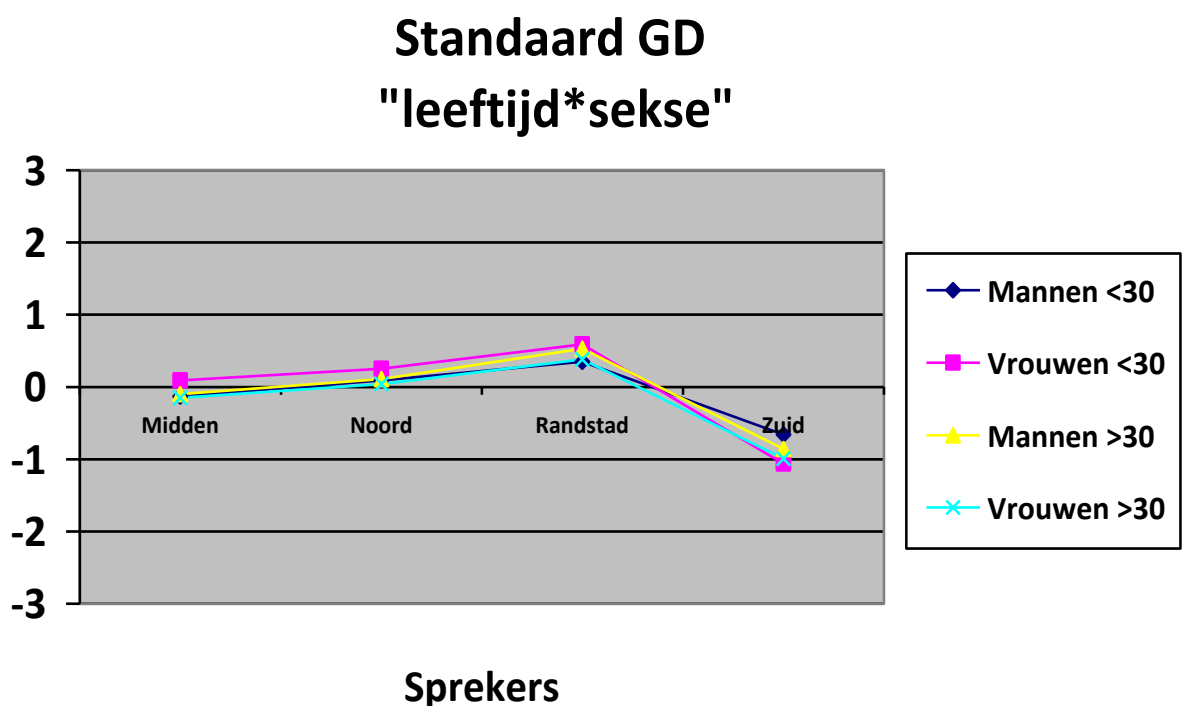
<b>Standaard</b>				
Paar	Gemiddelde verschillscore	Std. Deviatie	t	p
<b>6-puntsschaal</b>				
Midden - Noord	-,13818	,72819	-1,888	,062
Midden – Randstad	-,61576	,76062	-8,055	,000
Midden – Zuid	,80485	,93098	8,602	,000
Noord – Randstad	-,47758	,80151	-5,929	,000
Noord – Zuid	,94303	1,08000	8,688	,000
Randstad – Zuid	1,42061	1,03692	13,632	,000
<b>7-puntsschaal</b>				
Midden - Noord	-,23600	,84776	-2,784	,006
Midden – Randstad	-,46600	,92728	-5,025	,000
Midden – Zuid	,94800	,99559	9,522	,000
Noord – Randstad	-,23000	,95869	-2,399	,018
Noord – Zuid	1,18400	1,03580	11,431	,000
Randstad – Zuid	1,41400	1,08851	12,990	,000
<b>Gecombineerde data</b>				
Midden - Noord	-,18734	,79006	-3,345	,001
Midden – Randstad	-,54050	,84965	-8,974	,000
Midden – Zuid	,87678	,96423	12,827	,000
Noord – Randstad	-,35317	,89045	-5,595	,000
Noord – Zuid	1,06412	1,06223	14,132	,000
Randstad – Zuid	1,41729	1,06048	18,853	,000

Uit de resultaten van het eerste gedeelte van het repeated measures-model (zie tabel 4) bleek dat er een hoofdeffect bij de factor "beoordelaarsregio" was bij de 6-P, voor de schaal *standaard*. Het blijkt dat beoordelaars uit de regio Midden (0,49) significant hogere scores gaven op de schaal *standaard*

dan de beoordelaars uit de regio Noord (-0,20) (Tukey's range test: mean difference=0,664; std. error= 0,2315; p=0,026). Een mogelijke verklaring voor dit effect is het relatief kleine aantal beoordelaars uit de regio Midden (n=12) ten opzichte van het Noorden (n=46).

Bij de Noordelijke beoordelaars trad er een significant effect (zie tabel 5) op bij de schaal *standaard* bij de gecombineerde data met betrekking tot de interacties tussen de within-factor "sprekersregio" en de factoren "seks" en "leeftijd". Hierbij moet wel gemeld worden dat er bij de mannen onder de 30 jaar slechts weinig beoordelaars waren (n=10) ten opzichte van de andere groepen (n>19). De interactie is weergegeven in figuur 1.

**Figuur 1.** Gemiddelde scores gegeven aan de 4 sprekersregio's op de schaal *standaard*. De vier lijnen tonen mannen onder de 30 jaar, vrouwen onder de 30 jaar, mannen boven de 30 jaar en vrouwen boven de 30 jaar uit de beoordelaarsregio Noord.



In figuur 1 is te zien dat de Noordelijke mannen onder de 30 jaar, de regio's Midden, Noord en Randstad relatief laag beoordelen. Ze beoordelen de sprekersregio Zuid relatief hoog. De jonge vrouwen uit het Noorden geven het tegengestelde patroon weer. De jonge Noordelijke vrouwen waarderen het Zuiden met de laagste score van alle beoordelaarsgroepen, terwijl de rest van de sprekersregio's juist de hoogste waardering kreeg van deze beoordelaarsgroep van jonge vrouwen.

#### 5.4.2 Mooi

De spraak uit de Randstad (GD: 0,26) wordt het meest mooi gevonden, gevolgd door de spraak uit de regio's Midden (GD: -0,09) en Noord (GD: -0,10), die bijna identiek worden beoordeeld. De spraak uit de regio Zuid (GD: -0,24) wordt door de beoordelaars het minst mooi gevonden.

Aan de hand van gepaarde t-tests (met Bonferroni-correctie  $p < 0,017$ ) zijn de verschillen getoetst tussen de gemiddelde scores die gegeven zijn aan de vier afzonderlijke sprekersregio's. De verschillen

tussen de vier regio's zijn vrij klein, wat duidelijk is te zien in tabel 10. De significante verschillen zijn aangegeven door middel van een grijze rij. De regio's Midden, Noord en Zuid verschillen niet significant van elkaar, maar ze verschillen wel allemaal significant van de sprekersregio Randstad in alle datasets.

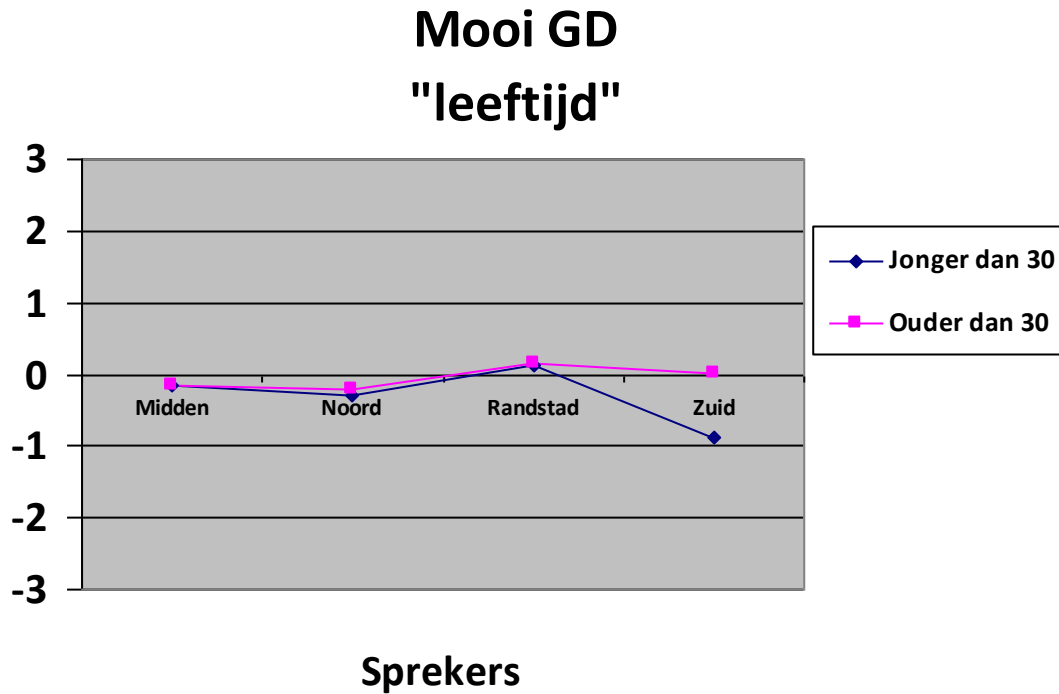
**Tabel 10.** Data van gepaarde t-tests. Gepaarde sprekersregio's bij schaal mooi. De standaarddeviatie is die van de verschillscore

<b>Mooi</b>				
Paar	Gemiddelde verschillscore	Std. Deviatie	t	p
<b>6-puntsschaal</b>				
Midden – Noord	,00485	,71425	,068	,946
Midden – Randstad	-,40970	,73074	-5,579	,000
Midden – Zuid	,06303	1,14207	,549	,584
Noord – Randstad	-,41455	,94880	-4,347	,000
Noord – Zuid	,05818	1,25276	,462	,645
Randstad – Zuid	,47273	1,27939	3,676	,000
<b>7-puntsschaal</b>				
Midden – Noord	,00800	,73191	,109	,913
Midden – Randstad	-,29600	,80024	-3,699	,000
Midden – Zuid	,23000	1,00760	2,283	,025
Noord – Randstad	-,30400	,93829	-3,240	,002
Noord – Zuid	,22200	1,19287	1,861	,066
Randstad – Zuid	,52600	1,22751	4,285	,000
<b>Gecombineerde data</b>				
Midden – Noord	,00643	,72135	,126	,900
Midden – Randstad	-,35256	,76664	-6,487	,000
Midden – Zuid	,14693	1,07713	1,924	,056
Noord – Randstad	-,35899	,94278	-5,372	,000
Noord – Zuid	,14050	1,22270	1,621	,107
Randstad – Zuid	,49950	1,25071	5,634	,000

Er is een hoofdeffect van de factor "leeftijd" op de schaal *mooi* bij de GD van de regio Noord (zie tabel 5). Jongere beoordelaars waren over het algemeen significant negatiever (-0,35) dan de oudere beoordelaars (-0,04) (gepaarde t-test: mean difference=-0,306; std. error=0,151; p=0,046).

In figuur 2 is het significante effect van de interactie tussen "sprekersregio" en "leeftijd" te zien in de GD. Het gaat hier om de factor "leeftijd", en heeft dus alleen betrekking op de Noordelijke beoordelaars.

**Figuur 2.** Gemiddelde scores gegeven aan de 4 sprekersregio's op de schaal mooi. De twee lijnen geven de oudere en jongere leeftijdsgroep weer uit de beoordelaarsregio Noord.



In figuur 2 is te zien dat de jongere en oudere beoordelaars uit de regio Noord de verschillende sprekersregio's vrijwel even mooi beoordelen. Dit geldt echter niet voor de regio Zuid, die door de jongere beoordelaars minder mooi wordt bevonden.

#### 5.4.3 Niet-regionaal

Sprekers uit de Randstad (GD: 0,48) worden het meest niet-regionaal beoordeeld, gevolgd door de sprekers uit de regio Noord (GD: -0,17) en vervolgens Midden (GD: -0,40). De spraak uit de regio Zuid (GD: -1,54) wordt het meest regionaal gevonden.

Aan de hand van gepaarde t-tests (met Bonferroni-correctie  $p < 0,017$ ) zijn de verschillen getoetst tussen de gemiddelde scores die gegeven zijn aan de vier afzonderlijke sprekersregio's. In de data van de 7-puntsschaal en de gecombineerde data verschillen alle regio's significant van elkaar op de schaal *niet-regionaal*. Dit is te zien in tabel 11.

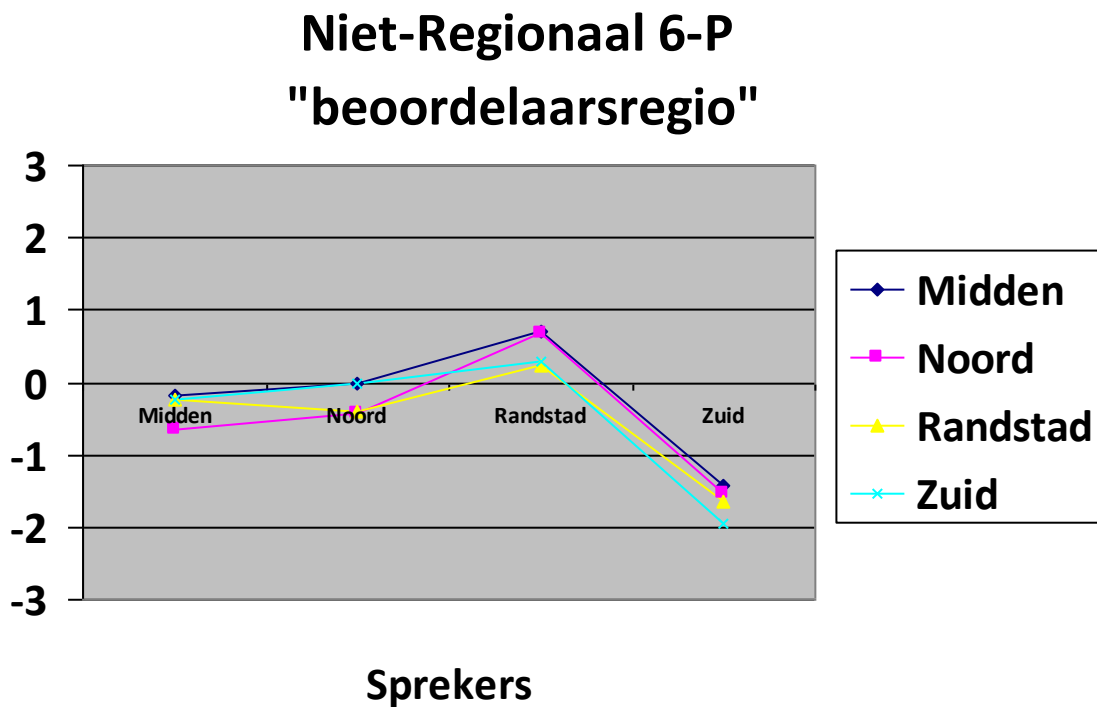
**Tabel 11.** Data van gepaarde t-tests. Gepaarde sprekersregio's bij schaal niet-regionaal. De standaarddeviatie is die van de verschilscore

<b>Niet-regionaal</b>				
Paar	Gemiddelde verschilscore	Std. Deviatie	t	p
<b>6-puntsschaal</b>				
Midden – Noord	-,14545	1,02389	-1,413	,161
Midden – Randstad	-,92606	1,20733	-7,632	,000
Midden – Zuid	1,22667	1,11304	10,966	,000
Noord – Randstad	-,78061	1,27474	-6,093	,000
Noord – Zuid	1,37212	1,23335	11,069	,000
Randstad – Zuid	2,15273	1,37547	15,572	,000
<b>7-puntsschaal</b>				
Midden – Noord	-,32000	,94324	-3,393	,001
Midden – Randstad	-,85200	1,00970	-8,438	,000
Midden – Zuid	1,05000	,99387	10,565	,000
Noord – Randstad	-,53200	1,07166	-4,964	,000
Noord – Zuid	1,37000	1,16424	11,767	,000
Randstad – Zuid	1,90200	1,27017	14,974	,000
<b>Gecombineerde data</b>				
Midden - Noord	-,23317	,98559	-3,337	,001
Midden – Randstad	-,88884	1,11022	-11,294	,000
Midden – Zuid	1,13789	1,05589	15,202	,000
Noord – Randstad	-,65568	1,18069	-7,834	,000
Noord – Zuid	1,37106	1,19609	16,170	,000
Randstad – Zuid	2,02673	1,32622	21,558	,000

Uit de resultaten van het eerste gedeelte van het repeated measures-model (zie tabel 4) bleek dat er een hoofdeffect bij de factor “beoordelaarsregio” was bij de 7-P, voor de schaal *niet-regionaal*. Het blijkt dat beoordelaars uit de regio Noord (-0,09) significant hogere scores gaven op deze schaal dan de beoordelaars uit de regio Randstad (-0,67) (Tukey's range test: mean difference=0,57; std. error= 0,182; p=0,011).

Er werd een significant effect gevonden bij de interactie tussen de factor “beoordelaarsregio” en de within-factor “sprekersregio” op de schaal *niet-regionaal*. Deze interactie van de 6-P is te zien in figuur 3. Dit is het enige significante effect van de interactie tussen de sprekersregio's en de beoordelaarsregio's wat betreft alle schalen en datasets.

**Figuur 3.** Gemiddelde scores gegeven aan de 4 sprekersregio's op de schaal niet-regionaal per beoordelaarsregio van de 6-P.



Bij de significante interactie tussen de sprekersregio en de beoordelaarsregio in figuur 3 zijn twee effecten af te lezen. De beoordelaars uit de regio's Midden (0,70) en Noord (0,67) beoordelen de sprekers uit de regio Randstad minder regionaal dan de beoordelaars uit de regio's Randstad (0,23) en Zuid (0,30). De beoordelaars uit de regio Noord (-0,43) en Randstad (-0,40) beoordelen de regio Noord laag, vergeleken met de beoordelaars uit de regio's Midden (-0,02) en Zuid (-0,01). Opvallend is dat de beoordelaars de sprekers uit de eigen regio niet positiever beoordelen dan de andere beoordelaars dat doen. Zowel de beoordelaars uit het Noorden, de Randstad en het Zuiden beoordelen de sprekers uit de eigen regio als het meest regionaal.

#### 5.4.4 Verzorgd

Alle sprekersregio's ontvingen een gemiddelde score boven de 0 in zowel de 6- en 7-puntsschalen als de gecombineerde data. Het is opvallend dat de regio Zuid bij de schaal *verzorgd* bijna net zo hoog scoort als de Randstad. Sprekers uit de Randstad (GD: 0,63) spreken het meest verzorgd, gevolgd door de sprekers uit de regio Zuid (GD: 0,50) en vervolgens Noord (GD: 0,28). De spraak van de sprekers uit de regio Midden (GD: 0,14) wordt het minst verzorgd gevonden.

Aan de hand van gepaarde t-tests (met Bonferroni-correctie  $p < 0,017$ ) zijn de verschillen getoetst tussen de gemiddelde scores die gegeven zijn aan de vier afzonderlijke sprekersregio's. De resultaten hiervan zijn weergegeven in tabel 12.



**Tabel 12.** Data van gepaarde t-tests. Gepaarde sprekersregio's bij schaal verzorgd. De standaarddeviatie is die van de verschillscore

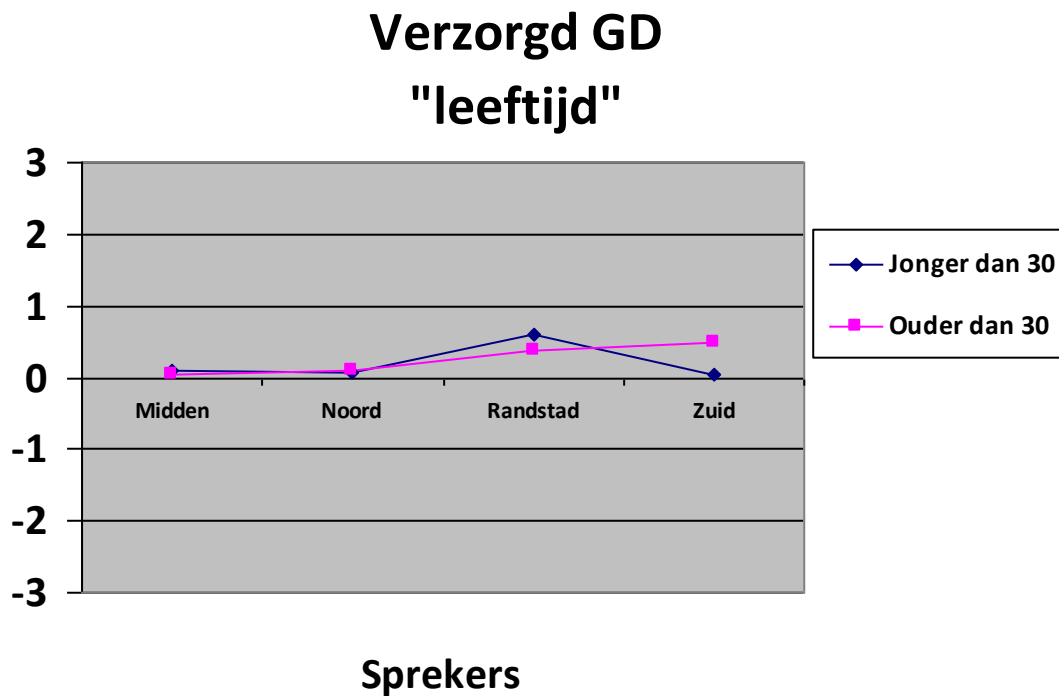
<b>Verzorgd</b>				
Paren	Gemiddelde verschillscore	Std. Deviatie	t	P
<b>6-puntsschaal</b>				
Midden - Noord	-,14788	,73989	-1,989	,050
Midden – Randstad	-,53333	,73499	-7,220	,000
Midden – Zuid	-,44606	,99488	-4,461	,000
Noord – Randstad	-,38545	,81122	-4,728	,000
Noord – Zuid	-,29818	1,04225	-2,847	,005
Randstad – Zuid	,08727	1,05979	,819	,415
<b>7-puntsschaal</b>				
Midden - Noord	-,12200	,74449	-1,639	,104
Midden – Randstad	-,44000	,81749	-5,382	,000
Midden – Zuid	-,27000	,80221	-3,366	,001
Noord – Randstad	-,31800	,91192	-3,487	,001
Noord – Zuid	-,14800	1,03655	-1,428	,156
Randstad – Zuid	,17000	,93954	1,809	,073
<b>Gecombineerde data</b>				
Midden - Noord	-,13487	,74044	-2,570	,011
Midden – Randstad	-,48643	,77699	-8,832	,000
Midden – Zuid	-,35759	,90524	-5,572	,000
Noord – Randstad	-,35156	,86177	-5,755	,000
Noord – Zuid	-,22271	1,03949	-3,022	,003
Randstad – Zuid	,12884	,99949	1,818	,070

Uit de resultaten van het eerste gedeelte van het repeated measures-model (zie tabel 4) bleek dat er een hoofdeffect bij de factor “beoordelaarsregio” was bij de 6-P en de GD, voor de schaal *verzorgd*. Het blijkt dat bij de 6-P de beoordelaars uit de regio Randstad (0,70) significant hogere scores gaven op de schaal *verzorgd* dan de beoordelaars uit de regio Noord (0,15) (Tukey’s range test: mean difference=0,549; std. error=0,1948; p=0,030).

Het effect op de GD op de schaal *verzorgd* trad op door een significant verschil tussen de beoordelingen van de regio Zuid (0,60), die significant hogere scores gaven dan de beoordelaars uit de regio Noord (0,24) (Tukey’s range test: mean difference=0,364; std. error=0,1348; p=0,038).

Er is een significant effect bij de schaal *verzorgd* in de GD wat betreft de interactie tussen de leeftijd van de beoordelaar en de sprekersregio. Het gaat hier wederom over de beoordelaars uit de regio Noord. Die interactie is weergegeven in figuur 4.

**Figuur 4.** Gemiddelde scores gegeven aan de 4 sprekersregio's op de schaal verzorgd. De twee lijnen geven de oudere en jongere leeftijdsgroep weer uit de beoordelaarsregio Noord.



De jongere beoordelaars uit de regio Noord beoordelen de sprekers uit het Zuiden minder verzorgd dan de oudere beoordelaars uit de regio Noord. De grafiek in figuur 4 lijkt sterk op de grafiek in figuur 2 die eenzelfde patroon weergeeft, waarbij de jongere beoordelaars strenger zijn ten opzichte van de spraak uit de regio Zuid.

## 6. Discussie

In dit hoofdstuk worden de resultaten besproken aan de hand van de hypothesen. Eerst worden de hypothesen nog een keer kort herhaald. Daarna volgt een gedeelte waarin de implicaties van de resultaten worden behandeld aan de hand van de hypothesen. Daarin komen ook verschillen en overeenkomsten met eerder onderzoek aan bod, met name de onderzoeken van Pinget et al. (2013) en Grondelaers et al. (2010). Bij elke hypothese wordt vermeld of deze aan de hand van de zojuist gepresenteerde data uit dit onderzoek is bevestigd of niet.

### 6.1 De hypothesen herhaald

De verschillende hypothesen die aan de hand van dit onderzoek worden getest zijn:

1. Er wordt geen significant verschil gevonden tussen de data van de 6-puntsschaal en de 7-puntsschaal.
2. De Randstadshypothese : De Nederlandse uitspraaknorm ligt in de Randstad.
3. De Vertrouwdheidshypothese. Spraak uit de eigen regio wordt positiever beoordeeld.
4. De regionale stereotypenhypothese: Sprekers uit geografisch perifere gebieden worden minder standaard gevonden dan sprekers uit niet-perifere gebieden buiten de Randstad.
5. De uniforme normhypothese: Nederlanders hanteren dezelfde Nederlandse uitspraaknorm ongeacht hun regionale afkomst.

### 6.2 Discussie Likertschaalonderzoek

In deze paragraaf volgt de discussie over het Likertschaalonderzoek naar aanleiding van de resultaten in het vorige hoofdstuk. In hoofdstuk 3 "Hypothesen" werd de volgende hypothese opgesteld:

1. Er wordt geen significant verschil gevonden tussen de data van de 6-puntsschaal en de 7-puntsschaal.

Uit de resultaten van het Likertschaalonderzoek is gebleken dat er geen significant verschil is tussen de data van de 6-puntsschaal en de 7-puntsschaal wat betreft de hele groep beoordelaars. Dit geldt voor alle schalen, *standaard*, *mooi*, *niet-regionaal* en *verzorgd*.

Echter, uit de repeated measures-test betreffende de beoordelaarsregio Noord kwam een significant effect naar voren bij de schaal *niet-regionaal*. Het gaat hier echter over een subset van de gehele groep beoordelaars en deze data zijn in die zin minder compleet en minder betrouwbaar. Daarnaast is het onwaarschijnlijk dat de Noordelijke beoordelaars anders zouden reageren op even of oneven Likertschalen dan beoordelaars uit andere regio's. Het is aannemelijk dat de verschillen die optraden bij de regio Noord in een onderzoek met meer beoordelaars zouden vervagen.

In de literatuur werden wel verschillen gevonden tussen onderzoeken met even of oneven schaalengtes, maar men kan hier beter spreken van "tendensen" dan van echte significante verschillen. In dit onderzoek gaven de beoordelaars van de 7-P over het algemeen hogere scores dan

de beoordelaars van de 6-P, behalve op de schaal *standaard*. Een duidelijke tendens is in deze resultaten dan ook niet te zien. Hypothese 1 is in dit onderzoek bevestigd.

### 6.3 Discussie Attitudeonderzoek

In deze paragraaf volgt de discussie over het attitudeonderzoek naar aanleiding van de resultaten in het vorige hoofdstuk. De hypothesen worden in deze paragraaf getest met de resultaten.

2. De Randstadshypothese : De Nederlandse uitspraaknorm ligt in de Randstad.

In de resultaten is te zien dat de spraak uit de regio Randstad het hoogst werd beoordeeld op de schaal *standaard*. De beoordelaars afkomstig uit alle vier de regio's van Nederland beoordeelden de spraak uit de Randstad als het meest standaardtalig. Het verschil met de andere regio's was significant. De Randstadshypothese wordt in dit onderzoek bevestigd.

Deze scriptie komt wat betreft deze hypothese grotendeels overeen met de onderzoeken van Pinget et al. (2013) en Grondelaers et al. (2010). De Nederlandse uitspraaknorm ligt in de Randstad. Dit scriptieonderzoek verschilt echter wel in de evaluatie van de andere sprekersregio's voor wat betreft de schaal *standaard*. Waarbij in de voorgaande onderzoeken de regio Midden een duidelijke overgangsregio bleek (niet alleen bij de schaal *standaard*, maar ook op de andere schalen), laten de resultaten van dit onderzoek juist zien dat er een "middenmoot" is met daarin de sprekersregio's Noord en Midden, waarbij de sprekersregio Noord zelfs iets hoger scoorde op *standaard*.

3. De Vertrouwdheidshypothese. Spraak uit de eigen regio wordt positiever beoordeeld.

Er is in dit onderzoek geen vertrouwdheidseffect gevonden, dat wil zeggen, er is geen significant effect opgetreden op de schaal *mooi* tussen de herkomstregio van de beoordelaars en de regio van de sprekers. De schaal *mooi* was de schaal met de kleinste verschillen in de gemiddelde scores van de oordelen per regio. De data laten zien dat de beoordelaars de spraak uit de eigen regio over het algemeen minder positief waardeerden dan de andere beoordelaars dat deden. Slechts bij één van de datasets (6-P) werd een significant effect gevonden in de interactie tussen de sprekersregio en de beoordelaarsregio, bij de schaal *niet-regionaal*. Dit significante effect kwam echter niet terug in de 7-P en de gecombineerde data. Daarnaast is deze uitkomst een tegenovergesteld effect van de verwachting. In Pinget et al. (2013) werd een significant effect gevonden voor de beoordelaarsregio Zuid, die de sprekers uit de eigen regio positiever beoordeelde op deze schaal. De vertrouwdheidshypothese wordt in dit onderzoek niet bevestigd.

4. De regionale stereotypenhypothese: Sprekers uit geografisch perifere gebieden worden minder standaard gevonden dan sprekers uit niet-perifere gebieden buiten de Randstad.

In voorgaande onderzoeken werd de spraak uit de perifere gebieden Noord en Zuid als minder standaard beoordeeld dan de regio Midden. De volgorde van positieve scores naar negatieve scores was over het algemeen: Randstad, Midden, Noord en Zuid. Dit onderzoek onderschrijft de resultaten van de voorgaande onderzoeken (Pinget et al. (2013) en Grondelaers et al. (2010)) slechts voor de regio Zuid. De regio Noord wordt relatief gezien niet zo non-standaard bevonden als Zuid, en scoort in dit onderzoek zelfs hoger dan de regio Midden. De verschillen tussen deze twee regio's was in

zowel de 7-P als in de gecombineerde data significant. De regionale stereotypenhypothese is in dit onderzoek bevestigd voor de regio Zuid en niet bevestigd voor de regio Noord.

Wat hier aan ten grondslag ligt is moeilijk te zeggen, aangezien dit onderzoek (in tegenstelling tot de voorgaande onderzoeken), geen classificatieonderzoek bevat. Mogelijke oorzaken zijn het classificeren van sprekers uit de regio Midden in een perifeer gebied, fonetische 'triggers' van sommige sprekers uit het Midden van het land (zoals de zachte g), en de verschillen in onderzoeksopzet tussen deze scriptie en Pinget et al. (2013) en Grondelaers et al. (2010). De verschillen in de onderzoeksopzet variëren van de gebruikte sprekers, oud versus jong, in vergelijking met Pinget et al. (2013) of het aantal gebruikte sprekersstimuli in vergelijking met Grondelaers et al. (2010).

5. De uniforme normhypothese: Nederlanders hanteren dezelfde Nederlandse uitspraaknorm ongeacht hun regionale afkomst.

De resultaten uit deze scriptie laten zien dat er geen effect is van "beoordelaarsregio" op de schaal *standaard*. De hypothese uit Pinget et al. (2013) waarbij slechts de hoogopgeleide beoordelaars deze uitspraaknorm zouden hanteren was vanwege het kleine aantal laagopgeleiden in dit onderzoek niet mogelijk statistisch betrouwbaar te onderzoeken. In Grondelaers et al. (2010) werd geconcludeerd dat de factoren "beoordelaarsregio", "seks", "opleidingsniveau" en in mindere mate "leeftijd", geen effect hadden op de attitude-profielen. Op het punt van de beoordelaarsregio lijkt deze scriptie de data van Grondelaers et al. (2010) gedeeltelijk te ondersteunen, hoewel er wel een effect van de beoordelaarsregio werd gevonden. Dit effect kwam alleen voor in de data van de 6-puntsschaal enquête op de schaal *niet-regionaal*, maar echter niet in de 7-P en de GD. De overeenkomsten met Grondelaers et al (2010) kan mogelijk veroorzaakt zijn door het gebruik van vergelijkbare (oudere) sprekers uit het corpus. Er was wel een aantal effecten te zien bij het tweede gedeelte van het repeated measures-model (met betrekking tot de regio Noord) bij de factoren "leeftijd", iets wat in Grondelaers et al. (2010) slechts bij de dimensie "welluidendheid" werd aangetroffen. (De invloed van de factoren "seks" en "leeftijd" worden in de volgende alinea's nog uitgebreider besproken.) Of de uniforme normhypothese bevestigd kan worden, heeft nog nader onderzoek.

Naast de vijf hypothesen werd ook de invloed van seks en leeftijd van de beoordelaars als factor in de data-analyse meeberekend. Dit gold echter alleen voor de regio Noord vanwege de spreiding van de beoordelaars.

- De factor "seks" genereerde geen significante effecten (wel in de interactie met de factor "leeftijd"). De in hoofdstuk 3 genoemde aanname dat de seks van de beoordelaar van invloed zou kunnen zijn, aangezien vrouwen beter zijn in het onderscheiden van spreekstijlen en spraakkenmerken, is uit de onderzoeksdata van de beoordelaarsregio Noord niet eenduidig naar voren gekomen.

- De factor "leeftijd" zorgde wel voor significante effecten. Op de schaal *mooi* werd een hoofdeffect van de factor "leeftijd" gevonden waarbij de beoordelaars onder de 30 jaar de spraak gemiddeld lager beoordeelden dan de beoordelaars boven de 30 jaar. Bij de schalen *mooi* en *verzorgd* speelde de leeftijd van de beoordelaars ook een rol bij het evalueren van de verschillende sprekersregio's. Zo gaven de Noordelijke beoordelaars onder de 30 jaar gemiddeld lagere scores aan

de sprekers uit de regio Zuid. Dit is een vergelijkbare uitkomst als in het onderzoek van Grondelaers et al. (2010). Er is geen beoordelingsverschil tussen de jongere en oudere groep voor de andere sprekersregio's. De Noordelijke beoordelingen van dit onderzoek kunnen duiden op het eerder genoemde feit dat is gebleken dat oudere proefpersonen meer tolerant zouden zijn voor non-standaardvariëteiten. Echter moet hierbij vermeld worden dat de uitkomsten van de Noordelijke beoordelaars niet per se geëxtrapoleerd kunnen worden over de gehele populatie inclusief de andere regio's.

## 7. Conclusie en vervolgonderzoek

In dit hoofdstuk worden de hoofdvragen herhaald en beantwoord. Ook wordt er een aantal mogelijkheden tot vervolgonderzoeken gegeven.

### 7.1 Attitudeonderzoek

De hoofdvraag voor het attitudeonderzoek luidde: *“In hoeverre wordt het Standaardnederlands uit de vier gebieden als regionaal gekleurd beoordeeld en welke normoorden verbinden de beoordelaars aan regionaal gekleurd Standaardnederlands?”*

Het Standaardnederlands uit de regio Randstad wordt het meest standaard bevonden en het minst regionaal. Daarna vormen de regio's Noord en Midden samen de middenmoot wat betreft de schaal *standaard* en *regionaal*. De regio Zuid wordt het minst standaard bevonden en het meest regionaal gekleurd. De spraak uit de regio Randstad wordt als het meest verzorgd beoordeeld, met opvallend genoeg op de tweede plek de regio Zuid. Daarna volgt Noord en als laatste de regio Midden. De hoge score voor de regio Zuid kan ook deels de lage Pearson's correlatiescores tussen de schaal *niet-regionaal* en *verzorgd* verklaren (hoewel ze wel significant waren). De schaal *mooi* zorgde voor weinig grote verschillen in de sprekersregio's. De sprekers uit de Randstad spreken het mooiste Standaardnederlands, gevolgd op kleine achterstand door het Midden, het Noorden en het Zuiden.

Aangezien deze scriptie geen classificatieonderzoek bevat, is er bij de bespreking van de resultaten een aantal vragen blijven hangen over de beoordelingen. Zo kan een classificatietask wellicht de relatief hoge scores voor de beoordelaarsregio Noord ten opzichte van de beoordelaarsregio Midden verklaren. In een vervolgonderzoek zou een classificatietask deze onbeantwoorde vragen kunnen beantwoorden.

### 7.2 Likertschaalonderzoek

De hoofdvraag voor het Likertschaalonderzoek luidde: *“Wat zijn de verschillen in de uitkomsten van een attitudeonderzoek met een 6-puntsschaal of een attitudeonderzoek met een 7-puntsschaal?”*

De verschillen in de uitkomsten van de beide enquêtes waren klein en niet-significant. De subset van slechts Noordelijke beoordelaars gaf dan wel een significant effect bij één schaal, maar het ging hierbij om een kleiner aantal beoordelaars, wat een invloed zou kunnen hebben gehad in de resultaten. Het antwoord op de hoofdvraag is dan ook dat er geen verschillen zijn in de uitkomsten tussen een taalattitudeonderzoek met een 6- of een 7-puntsschaal.

In deze scriptie trad er geen significant verschil op tussen de data van de 6- en de 7-puntsschaalenquête. Een onderzoek waarbij het verschil tussen het aantal Likertschaalcategorieën groter is, zou ook grotere verschillen kunnen opleveren. Deze grotere verschillen zijn ook te zien in de literatuur over verschillen in Likertschaalscores. Een onderzoek waarbij een 3-of 4-puntsschaalenquête met een 6- of 7-puntsschaalenquête wordt vergeleken, kan een waardevol vervolgonderzoek zijn.





# Bronnen

**Cargile, A.C., Giles, H., Ryan, E.B. & J.J. Bradac (1994).** Language attitudes as a social process: A conceptual model and new directions. *Language & Communication*, Vol. 14, No. 3, p 211-216.

**Carifio, J. & R.J. Perla (2007).** Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes. *Journal of Social Sciences*, Vol. 3, No. 3, p 106-116.

**Chang, L. (1994).** A Psychometric Evaluation of 4-Point and 6-Point Likert-Type Scales in Relation to Reliability and Validity. *Applied Psychological Measurement*, Vol. 18, No. 3, p 205-215.

**Clason, D.L. & T.J. Dormody (1994).** Analyzing Data Measured by Individual Likert-Type Items. *Journal of Agricultural Education*, Vol. 35, No. 4, p 31-35.

**Colman, A. M., Norris, C. E., & C.C. Preston (1997).** Comparing rating scales of different lengths: Equivalence of scores from 5-point and 7-point scales. *Psychological Reports*, Vol. 80, p 355-362.

**Dawes, J. (2008).** Do data characteristics change according to the number of scale points used ? An experiment using 5 point, 7 point and 10 point scales. *International Journal of Market Research*, Vol. 50, No. 1.

**De Winter, C.F. & D. Didou (2010).** Five-Point Likert Items: t test versus Mann-Whitney-Wilcoxon, Practical Assesment, *Research & Evaluation*, Vol. 15, No. 11, p 1-6.

**Finn, R.H. (1972).** Effects of Some Variations in Rating Scale Characteristics on the Means and Reliabilities of Ratings. *Educational and Psychological Measurement*, Vol. 32, p 255-265.

**Garland, R. (1991).** The Mid-point on a Rating Scale: Is it Desirable? *Marketing Bulletin*, Vol. 2, p 66-70.

**Grover, R & M. Vriens (2006).** The Handbook of Marketing Research: Uses, Misuses, and Future Advances, Thousand Oaks, CA: Sage Publications, p 88-89.

**Grondelaers, S.A., R. van Hout & M. Steegs (2010).** Evaluating Regional Accent Variation in Standard Dutch. *Journal of Language and Social Psychology*, Vol. 29(1), p 101-116.

**Hoppenbrouwers, C. (1990).** *Het regiolect. Van dialect tot Algemeen Nederlands*. Muiderberg: Coutinho. p 20-27, 29-30.

**Kim, S. (2010).** The Influence of Likert Scale Format on Response Result, Validity, and Reliability of Scale -Using Scales Measuring Economic Shopping Orientation. *Journal of the Korean Society of Clothing and Textiles*, Vol. 34, No. 6 p 913-927.

**Knops, U. (1988).** Attitudes towards Regional Variation in Dutch Pronunciation. In: R. van Hout & U. Knops, *Language attitudes in the Dutch Language Area*. Dordrecht: Floris, p 105-120.

**Matell, M. S., & J. Jacoby (1971).** Is there an optimal number of alternatives for likert scales items? Study I: Reliability and validity. *Educational and Psychological Measurement*, Vol. 31, p 657-674.

**Norman, G. (2010).** Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Science Education*. Vol. 15, p 625-632.

**Nunnally J.C. (1978).** *Psychometric theory*. New York: McGraw Hill.

**Osgood, C.E., Suci, G. & P. Tannenbaum (1957).** *The Measurement of Meaning*. University of Illinois Press, 1957.

**Pinget, A., Rotteveel, M. & H. Van de Velde (2013, te verschijnen).** *Standaardnederlands met een accent: herkenning en evaluatie van regionaal gekleurd Standaardnederlands in Nederland*, Utrecht Institute of Linguistics OTS, Utrecht.

**Smakman, D. & R. Van Bezooijen (1999).** De uitspraak van het Standaardnederlands in Nederland. Een evaluatief en descriptief onderzoek. *Artikelen van de Derde Sociolinguïstische Conferentie*. Lunteren, p 367-378.

**Smakman, D. (2006).** *Standard Dutch in the Netherlands. A sociolinguistic and phonetic description*. Dissertation University of Nijmegen. p 5-6, 32-33, 38, 197, 227.

**Van Bezooijen, R. (1988).** The relative importance of pronunciation, prosody, and voice quality for the attribution of social status and personality characteristics. *Language Attitudes in the Dutch Language Area*, Dordrecht/Cinnaminson: Foris, p 85-103.

**Van Bezooijen, R. (2002).** Aesthetic evaluation of Dutch. Comparisons across dialects, accents and languages. *Perceptual Dialectology*, vol. 2., Amsterdam / Philadelphia: Benjamins, p 13-30.

**Van de Velde, H. & M. Houtermans (1999).** Vlamingen en Nederlanders over de uitspraak van nieuwslezers. In: Huls, E. & B. Weltens (red.). *Artikelen van de Derde Sociolinguïstische Conferentie*. Delft: Eburon, p 451-462.

**Van Hout, R. & H. Münstermann (1981).** Linguïstische afstand, dialect en attitude. *Gramma* Vol. 5, p 101-123.

**Wakita, T., Ueshima, N. & H. Noguchi (2012).** Psychological distance between categories in the Likert scale comparing different numbers of options. *Educational and Psychological measurement*, Vol. 72, No. 4, p 533-546.

**Worcester, R.M. & T.R. Burns (1975).** A statistical examination of the relative precision of verbal scales. *Journal of Market Research Society*, Vol. 17, p 181-197.