

“Eerste stappen richting validiteitsonderzoek en betrouwbaarheidsonderzoek van de ANTAT-CU4”

Onderzoek naar de test-hertest sensibiteit en betrouwbaarheid van de ANTAT-CU4.



**“Eerste stappen richting validiteitsonderzoek en betrouwbaarheidsonderzoek
van de ANTAT-CU4.”**

Onderzoek naar de test-hertest sensibiteit en betrouwbaarheid van de ANTAT-CU4.

Nijmegen, april 2016

Radboud Universiteit Nijmegen

Faculteit der Letteren, Taal- en spraakpathologie

Masterscriptie

Auteur: Sabrina Dassek

Studentnummer: s4141210

Begeleiders: Dr. M. B. Ruiter, logopedist en taal- en spraakpatholoog,
Sint Maartenskliniek.
E. Lotgering, MA, logopedist en taal- en spraakpatholoog.

© April 2016, Radboud Universiteit, Nijmegen.

Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of op enige andere manier zonder voorafgaande schriftelijke toestemming van de Radboud Universiteit Nijmegen.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form by any means, electronics, mechanical, photocopying, recording or otherwise, without the written permission of the publisher.

Voorwoord

Deze masterscriptie is geschreven in het kader van mijn afstuderen voor de opleiding Taal- en Spraakpathologie aan de Radboud Universiteit Nijmegen in de periode juni 2015 tot april 2016. Vanuit mijn beroep als logopedist ben ik al bekend met mensen met afasie, zowel in de acute, subacute en chronische fase en daarom sprak dit onderwerp mij aan voor mijn masterscriptie.

Gezien de veranderingen in de gezondheidszorg, worden de indelingen van het ICF-model (International Classification of Functioning, Disability and Health; World Health Organization, 2005) steeds belangrijker. Het ICF-model classificeert mensen met gezondheidsproblemen vanuit functie-, activiteiten- en participatieniveau uniform middels coderingen. Alleen als alle vier onderdelen tijdens de diagnostiek en behandeling in acht worden genomen is het naar mijns inziens mogelijk om patiëntgericht te werken. Vanuit die reden was ik heel erg gemotiveerd deel uit te maken van de verdere ontwikkeling van de ANTAT-CU.

De ANTAT-CU is ontwikkeld om de functionele communicatievaardigheden van mensen met afasie in kaart te brengen. Het kwantitatief en objectief meten van de functionele communicatie is een uitdaging en kan alleen verder ontwikkeld worden als dit onderzoek een vervolg krijgt.

Deze masterscriptie kon alleen geschreven worden, met hulp en steun van de mensen om mij heen. Grote dank wil ik daarom uitspreken aan mijn begeleiders vanuit de Radboud Universiteit, namelijk dr. Marina Ruiters voor het meedenken over mijn onderzoeksopzet, het geven van structuur en inhoudelijke feedback. Daarnaast wil ik ook Erica Lotgering (MA) bedanken voor de prettige feedback en voor de inhoudelijke tips voor het toepassen van de CU-analyse. Vervolgens wil ik Maartje Giessen en Sharon aan de Stegge bedanken voor het vooronderzoek en de audiodata van de proefpersonen zonder afasie.

Speciale dank gaat uit naar de Nederlandse en Duitse proefpersonen die aan dit onderzoek hebben deelgenomen. Dat is niet vanzelfsprekend en heeft deze masterscriptie mogelijk gemaakt. Het verzamelen van Nederlandse proefpersonen was een grote uitdaging voor mij. Zonder de hulp van mijn tante, was het niet mogelijk geweest zo veel Nederlandse proefpersonen te vinden. Daarom gaat een bijzondere dank uit naar haar!

Vervolgens wil ik mijn ouders bedanken. Zij hebben het mogelijk gemaakt deze opleiding te volgen en zij hebben mij steeds steun gegeven.

Ik wil me bijzonder bedanken bij mijn vrienden die altijd een luisterend oor voor me hebben gehad. Veel steun heb ik gehad aan Stefanie Pruys, Aniek Rolink, Hilde Vloet en Nils Gerritzen! Zonder jullie allemaal was het niet mogelijk geweest!

Veel plezier met het lezen van deze masterscriptie!

Sabrina Dassek, Nijmegen, April 2016

Inhoudsopgave

Samenvatting	
Abstract	
Afkortingen	
1. Inleiding	1
1.1. Indeling	2
2. Theoretische achtergrond	3
2.1. Communicatieve effectiviteit en efficiëntie	3
2.2. Van kwalitatief naar kwantitatief	5
2.3. Psychometrische kwaliteiten van een test.....	7
2.3.1. Normen	7
2.3.2. Componenten van betrouwbaarheid	9
2.3.3. Componenten van validiteit	13
2.3.4. Sensitiviteit	16
2.4. Doelstelling	17
2.5. Onderzoekshypothesen.....	18
3. Methode	19
3.1. Het meetinstrument ANTAT-CU4 en de Content Unit analyse	19
3.1.1. De ANTAT-CU4.....	21
3.1.2. Ontwikkeling van het CU4-schema	22
3.2. Test-hertest sensibilliteits- en betrouwbaarheidsonderzoek	26
3.2.1. Nederlandstalige proefpersonen zonder afasie	26
3.2.2. Duitstalige proefpersonen met afasie	29
3.3. Duitstalig CU-schema	33
3.4. Advies voor twee parallele testversies	34
4. Resultaten	35
4.1. Descriptieve statistiek – Veranderingen in de functionele communicatie	35
4.2. Inductieve statistiek – Test-hertest stabiliteit van de ANTAT-CU4	37
4.3. Producten	38
5. Discussie en conclusie	40
5.1. Normen van de originele ANTAT	40
5.2. Hypothesen toetsing.....	41
5.3. De ANTAT-CU4.....	43
5.4. Het CU-schema	44
5.5. Responsietijden, effectiviteit en efficiëntie.....	46
5.6. Doelgroep van de ANTAT-CU4.....	48
5.7. Conclusie en aanbevelingen	49
Referenties	50
Bijlagen	Fout! Bladwijzer niet gedefinieerd.

Bijlage A: Scenario keuze oud-nieuw voor de ANTAT-CU4 **Fout! Bladwijzer niet gedefinieerd.**

Bijlage B: Afbeeldingen voor de Nederlandse ANTAT-CU4 **Fout! Bladwijzer niet gedefinieerd.**

Bijlage C: Overzicht uiteindelijke scenario's ANTAT-CU4 **Fout! Bladwijzer niet gedefinieerd.**

Bijlage D: Overzicht drempelwaarden voor alle scenario's uit de ANTAT-CU4 **Fout! Bladwijzer niet gedefinieerd.**

Bijlage E: Drempelwaarden per aanleiding en verzoek **Fout! Bladwijzer niet gedefinieerd.**

Bijlage F: Drempelwaarden per propositie **Fout! Bladwijzer niet gedefinieerd.**

Bijlage G: Data proefpersonen met en zonder afasie **Fout! Bladwijzer niet gedefinieerd.**

Bijlage H: Informatiebrief en toestemmingsformulier proefpersonen zonder afasie (NL) **Fout! Bladwijzer niet gedefinieerd.**

Bijlage I: CU-schema voor de ANTAT-CU4 **Fout! Bladwijzer niet gedefinieerd.**

Bijlage J: Informatiebrief en toestemmingsformulier proefpersonen met afasie (D) **Fout! Bladwijzer niet gedefinieerd.**

Bijlage K: Afbeeldingen voor de Duitstalige ANTAT **Fout! Bladwijzer niet gedefinieerd.**

Bijlage L: Ruwe scores op de ANTAT-CU4 van de proefpersonen zonder en met afasie **Fout! Bladwijzer niet gedefinieerd.**

 Nederlandstalige proefpersonen zonder afasie ($n=30$) **Fout! Bladwijzer niet gedefinieerd.**

 Duitstalige proefpersonen met chronische afasie ($n=8$) **Fout! Bladwijzer niet gedefinieerd.**

Bijlage M: Het maximale aantal CU per scenario **Fout! Bladwijzer niet gedefinieerd.**

Bijlage N: Duitse ANTAT met nieuwe scenario's (vertaling van de ANTAT-CU4) **Fout! Bladwijzer niet gedefinieerd.**

Bijlage O: Vragenlijst proefpersonen **Fout! Bladwijzer niet gedefinieerd.**

Samenvatting

Inleiding: De laatste normering van de Nederlandstalige ANTAT en de Duitstalige ANELT dateert uit 1994. Bij de beoordeling door de COTAN werd het onderdeel *normen* met onvoldoende beoordeeld. In de laatste jaren is veel onderzoek gedaan om de kwalitatieve beoordelingsmethode aan te passen naar een kwantitatieve beoordeling en de scenario's van de ANTAT te testen op de itemcriteria: actualiteit, inleefbaarheid en rolverdeling. De scenario's van de ANTAT voldoen niet meer aan deze itemcriteria. Giessen (2015) en Aan de Stegge (2015) hebben onderzocht welke scenario's aangepast of vervangen moeten worden. Zij hebben nieuwe scenario's ontwikkeld. Binnen deze masterscriptie worden de scenario's deels aangepast of vervangen. Nieuwe scenario's vragen naar een nieuw CU-schema.

Om de functionele communicatie goed te kunnen beoordelen moeten de maten communicatieve efficiëntie en communicatieve effectiviteit meetbaar gemaakt worden. Eerder onderzoek heeft aangetoond dat de kwantitatieve beoordelingsmethode (middels CU-schema) sensitiever meet, dan de oude kwalitatieve beoordelingsmethode van de ANTAT (middels 5-puntschaal). Door het toepassen van een kwantitatieve beoordelingsmethode verbeterde ook de construct validiteit van de ANTAT-CU. Deze masterscriptie is gericht op het meten van de communicatieve effectiviteit middels een nieuw CU-schema, welke binnen deze masterscriptie ontwikkeld wordt.

Methode: Ten eerste zijn de scenario's van de ANTAT aangepast of vervangen, zodat de ANTAT-CU4 is ontstaan. Vervolgens zijn de responsies op de gekozen scenario's van 60 proefpersonen zonder afasie (data verzameld door Giessen, 2015 en Aan de Stegge, 2015) gebruikt om een nieuw CU-schema te ontwikkelen. De ANTAT-CU4 werd vervolgens middels een test-hertest procedure bij 30 proefpersonen zonder afasie afgenomen om de teststabiliteit van de ANTAT-CU4 nader te onderzoeken. Vervolgens zijn uit deze groep, zes proefpersonen random gekozen om de responsies met behulp van het CU-schema te scoren en met een gepaarde t-toets te analyseren.

De ANTAT-CU4 werd middels een equivalentieonderzoek naar het Duits vertaald en vervolgens bij een groep Duitstalige proefpersonen met afasie ($n=10$) middels een test-hertest procedure afgenomen, om de functionele communicatie nader te onderzoeken. De responsies van acht proefpersonen met afasie zijn vervolgens middels de oude kwalitatieve beoordelingsmethode gescoord, omdat er nog geen Duitstalig CU-schema voor de ANTAT bestaat. De resultaten zijn middels descriptieve statistiek beschreven. Om in de toekomst een Duitstalig CU-schema verder te kunnen ontwikkelen, is de Duitstalige ANTAT met nieuwe scenario's bij 30 Duitstalige proefpersonen zonder afasie eenmalig afgenomen.

Resultaten: Binnen de groep proefpersonen zonder afasie ($n=6$) was er geen significant verschil tussen de ANTAT-CU4 scores op meetmoment 1 en de ANTAT-CU4 scores op meetmoment 2. De test-hertest resultaten van de proefpersonen met afasie ($n=8$) lieten zowel positieve als ook negatieve veranderingen in de functionele communicatie zien.

Discussie en conclusie: De ANTAT-CU4 lijkt stabiel te zijn, maar om een betere uitspraak erover te kunnen doen moeten in een vervolgonderzoek de data van alle 30 proefpersonen zonder afasie op de test-hertest stabiliteit onderzocht worden.

De testcores van de proefpersonen met afasie ($n=8$) moeten voorzichtig geïnterpreteerd worden, omdat deze alleen met de kwalitatieve beoordelingsmethode zijn onderzocht. Om de responsies van de proefpersonen met afasie betrouwbaar te scoren moet in de toekomst, met behulp van de responsies van de 30 Duitstalige proefpersonen zonder afasie, een Duits CU-schema ontwikkeld worden.

Voor toekomstig onderzoek zal het noodzakelijk zijn middels een grotere normgroep het nieuwe CU-schema van de ANTAT-CU4 aan te passen en de ANTAT-CU4 te normeren en te standaardiseren. Bovendien zal onderzocht moeten worden of de ANTAT-CU4 sensitief genoeg is, om veranderingen in de functionele communicatievaardigheden te kunnen meten. Op deze manier zal onderzocht moeten worden, of het mogelijk is met de ANTAT-CU4 een klinisch oordeel te kunnen geven.

Abstract

Introduction: The last standardization of the Dutch ANTAT and the German ANELT was done in 1994. The subsequent test assessment by COTAN showed that the evaluation component *norms* were insufficient. To assess verbal functional communication, the dimensions communicative effectiveness and communicative efficiency must be measurable. Previous research has shown, that a quantitative scoring method, i.e. a CU-scheme, measures verbal effectiveness over time more sensitively than the old qualitative scoring method (scoring by a 5-point scale). Additionally, the ANTAT scenarios were tested on following item criteria: object appropriateness, plausibility and imaginability. The outcome shows that the scenarios of the ANTAT do no longer meet those item criteria. Based on this Giessen (2015) and Aan de Stegge (2015) developed new scenarios which they then examined together with the old ones in order to find out which scenarios needed to be adapted or to be replaced. In this thesis, the scenarios are adapted according to their advice.

This thesis focusses on developing a new CU-scheme based on the chosen scenarios in order to allow a quantitative scoring of the ANTAT. The result of this scoring measures the effectiveness of functional communication.

Method: First, the scenarios of the ANTAT are adjusted so that the ANTAT-CU4 could be used in the thesis. Subsequently, responses to the chosen scenarios of the ANTAT of 60 test subjects without aphasia were used to develop a new CU-scheme. The ANTAT-CU4 was then used in a test-retest procedure with 30 test subjects without aphasia. The responses of 6 randomly selected test subjects out of this test group are then scored by using the CU-scheme and analyzed with a paired sampled t-test. The ANTAT-CU4 was translated into German by an equivalence analysis. The translated ANTAT was further used in a test-retest procedure with a group of German-speaking test subjects with aphasia (n=10). This was done to investigate the functional communication skills of persons with aphasia. The responses of eight subjects with aphasia were then scored by using the old qualitative scoring method, because no German CU-scheme exists for the translated ANTAT. The results are analyzed by using descriptive statistic. In order to develop a German CU-scheme in the future, the translated German ANTAT was also tested with 30 German test subjects without aphasia.

Results: Within the group of test subjects without aphasia (n=6), there was no significant difference between the initial ANTAT-CU4 scores and the retest scores. The test-retest results of the test subjects with aphasia (n=8) showed both positive and negative changes in the functional communication skills.

Discussion and conclusion: The ANTAT-CU4 seems to be stable, but follow-up research where all 30 test subjects without aphasia are further analyzed is needed. The test scores of the subjects with aphasia (n=8) should be interpreted with caution, because the responses were only scored by the qualitative assessment method. To score the responses of all subjects with aphasia (n=8) reliably, a German CU-scheme has to be further developed, using the responses of the 30 German test subjects without aphasia. For future research it will be necessary to standardize and normalize the CU-scheme and the ANTAT-CU4 through a larger group of test subjects. Furthermore, the ANTAT-CU4 has to be examined if it is sensitive enough to measure changes in functional communication skills. Only then it will be possible to decide if the ANTAT-CU4 can provide a clinical judgement about the functional communication of a patient with aphasia.

Afkortingen

AAT	Akense Afasie Test (Nederlandse versie)
AAT	Aachener Aphasie Test (Duitse versie)
ANTAT	Amsterdam Nijmegen Test voor Alledaagse Taalvaardigheden
ANELT	Amsterdam Nijmegen Everyday Language Test
APV	Afasie Partner Vragenlijst
ASHA	American Speech-Language-Hearing Association
BDAE	Boston Diagnostic Aphasia Examination
CETI	Communicative Effectiveness Index
CIU	Correct Information Unit
COTAN	Commissie Testaangelegenheden Nederland
CU	Content Unit
CU-schema	Content Unit schema
ICC	Interclasscorrelatiecoëfficiënt
ICF	International Classification of Functioning, Disability and Health
NIP	Nederlands Instituut van Psychologen
NT2	Nederlands als tweede taal
WHO	World Health Organization

1. Inleiding

De ANTAT (Amsterdam Nijmegen Test voor Alledaagse Taalvaardigheden; Blomert, Koster, & Kean, 1994) heeft als doel de functionele communicatie bij mensen met afasie op participatieniveau betrouwbaar en valide te testen. De laatste normering van de Nederlandstalige ANTAT en de Duitstalige ANELT (Amsterdam Nijmegen Everyday Language Test; Blomert & Buslach, 1994) is uit de jaren 1994. Voor deze laatste normering werden de Duitstalige en de Nederlandstalige normgroep samengevoegd. In 2008 is een nieuwe handleiding (Blomert & Kean, 2008) voor de ANTAT verschenen, maar worden de Duits-Nederlandse normen uit de jaren 1994 verder gebruikt.

De communicatieve effectiviteit werd in de originele ANTAT beoordeeld door de begrijpelijkheid en de verstaanbaarheid van de antwoorden kwalitatief te beoordelen op een 5-puntschaal. Deze methode is betrouwbaar, maar wel subjectief (Lotgering, 2009). In de laatste jaren is veel onderzoek gedaan om de kwalitatieve beoordelingsmethode aan te passen naar een kwantitatieve beoordeling en de scenario's van de ANTAT te testen op actualiteit, inleefbaarheid en rolverdeling (Dijkstra, 2006; Lotgering, 2009; Filipinski, 2014; Ruiters, Kolk, Rietveld, Dijkstra, & Lotgering, 2011).

Eerder onderzoek heeft aangetoond, dat de ANTAT zowel de communicatieve effectiviteit als ook de efficiëntie moet kunnen meten om een oordeel over de functionele communicatievaardigheden van mensen met afasie te kunnen geven (Dijkstra, 2006; Filipinski, 2014; Lotgering, 2009; Ruiters et al., 2011). De term functionele communicatie speelt binnen deze masterscriptie een essentiële rol. Functionele communicatie is volgens Frattali (1992): "the ability to receive a message or to convey a message, regardless of the mode, to communicate effectively and independently in a given environment" (Frattali, 1992; geciteerd door Ruiters et al., 2011). De communicatieve effectiviteit en de communicatieve efficiëntie zijn volgens Yorkston en Beukelman (1980) twee onderliggende dimensies van de functionele communicatie. De communicatieve effectiviteit wordt bepaald door de begrijpelijkheid van de communicatieve boodschap en de communicatieve efficiëntie wordt bepaald door de snelheid waarmee de essentiële delen uit de boodschap overgedragen worden (Lotgering, 2009). Een verbale boodschap is dus alleen efficiënt als deze voldoende snel en vloeiend overgebracht wordt en tevens essentiële informatie bevat. Essentiële informatie bestaat uit woorden die noodzakelijk zijn om een boodschap te kunnen begrijpen. Om in het alledaagse leven communicatief zelfredzaam te zijn, is het voor mensen met afasie dus noodzakelijk dat de functionele communicatie zowel effectief als ook efficiënt verloopt. Om deze onderdelen met de ANTAT kwantitatief te kunnen meten, moet de kwalitatieve beoordelingsmethode dus herzien worden.

Lotgering (2009) heeft een maat gevonden om de communicatieve efficiëntie bij afasiepatiënten kwantitatief te bepalen met behulp van Content Units (CU; Yorkston & Beukelman, 1980). CU's zijn als volgt gedefinieerd: "...a grouping of information that was always expressed as a unit by normal speakers" (Yorkston & Beukelman, 1980, p.30). De CU's geven een weerspiegeling van de essentiële informatie die geproduceerd wordt door gezonde sprekers. In de laatste jaren zijn drie verschillende CU-schema's voor de kwantitatieve beoordeling van de scenario's in de ANTAT ontwikkeld, namelijk de ANTAT-CU1, de ANTAT-CU2 en de ANTAT-CU3 (Lotgering, 2009; Ruiters et al., 2011; Filipinski, 2014). Ruiters et al. (2011) hebben vastgesteld, dat door het kwantitatieve beoordelingssysteem middels

CU-schema de psychometrische kwaliteiten van de ANTAT verbeterden. Met name de begripsvaliditeit en de sensitiviteit voor de meting van de verbale effectiviteit waren verbeterd.

Filipinski (2014) heeft in haar onderzoek de scenario's van de originele ANTAT getest op de itemcriteria van Blomert, Koster en Kean (1995), te weten: actualiteit, inleefbaarheid en duidelijkheid. Er kwam naar voren, dat 19 van de 20 scenario's niet aan een of meerdere criteria van Blomert et al. (1995) voldeden. Giessen (2015) en Aan de Stegge (2015) hebben de aanbevelingen van Filipinski (2014) gebruikt om 15 nieuwe scenario's te ontwikkelen. De originele en de nieuwe scenario's zijn door een groep proefpersonen zonder afasie ($n=60$) beoordeeld op rolverdeling, inleefbaarheid en actualiteit. Van de 15 nieuwe scenario's zijn zes scenario's (namelijk scenario 4, 5, 16, 17, 19 en 20) op inleefbaarheid en actualiteit als significant beter beoordeeld dan de originele scenario's (Giessen, 2015 en Aan de Stegge, 2015). Er zijn geen significante verschillen gevonden tussen originele en nieuwe scenario's wat betreft de beoordeling op rolverdeling. Met behulp van de uitkomsten van Giessen (2015) en Aan de Stegge (2015) wordt binnen deze masterscriptie een beslissing gemaakt welke scenario's (oud of nieuw) opgenomen worden in de ANTAT-CU4.

Het boven genoemde onderzoek van Giessen (2015) en aan de Stegge (2015) bestond nog uit een tweede onderdeel. Ze hebben niet alleen maar de scenario's laten beoordelen, maar ze hebben ook de originele en de nieuwe scenario's van de ANTAT bij de groep proefpersonen zonder afasie ($n=60$) afgenomen. De responsen zijn orthografisch getranscribeerd en worden binnen deze masterscriptie gebruikt om het nieuwe CU-schema voor de ANTAT-CU4 op te stellen.

De onderzoeken die tot nu toe beschreven zijn, hebben zich vooral beziggehouden met de Nederlandstalige ANTAT. In Duitsland wordt tot nu toe de ANELT gebruikt. De ANTAT-CU4 wordt binnen deze masterscriptie naar het Duits vertaald, om in een vervolgonderzoek voor de Duitstalige ANTAT met nieuwe scenario's ook een CU-schema te kunnen ontwikkelen.

1.1. Indeling

In hoofdstuk 2 wordt de theoretische achtergrond van de geschiedenis van het huidige CU-schema en de psychometrische kwaliteiten van testen beschreven. Vervolgens worden de doelstellingen en de hypothesen besproken. De methode die in dit scriptieonderzoek is gebruikt, wordt in hoofdstuk 3 nader toegelicht. Hierbij wordt de ANTAT-CU4 beschreven, de analyse van de Content Units (CU's) en het opstellen van het nieuwe CU-schema. Vervolgens wordt het onderzoeksontwerp voor de Nederlandstalige proefpersonen zonder afasie ($n=30$), de Duitstalige proefpersonen met afasie ($n=10$) en de Duitstalige proefpersonen zonder afasie ($n=30$) beschreven. In hoofdstuk 4 worden de resultaten van de test-hertest procedure en de producten die binnen deze masterscriptie zijn ontstaan besproken. De masterscriptie wordt afgerond met de discussie en conclusie in hoofdstuk 5. Hierin worden ten slotte ook de aanbevelingen en een uitzicht voor de toekomst beschreven.

2. Theoretische achtergrond

2.1. Communicatieve effectiviteit en efficiëntie

De ANTAT beoogt de communicatieve effectiviteit en efficiëntie te meten. Daarom is het belangrijk erachter te komen wat deze begrippen voor het huidige onderzoek betekenen. De communicatieve effectiviteit wordt bepaald door de begrijpelijkheid van de communicatieve boodschap (Lotgering, 2009). De communicatieve efficiëntie wordt bepaald door de snelheid waarmee de boodschap overgebracht wordt (Lotgering, 2009). Voorbeeld (1) (Filipinski, 2014, p. 13) zal de begrippen duidelijk maken:

(1) Spreker 1: *“Retourtje Zwolle, alstublieft.”*

Spreker 2: *“Goedemiddag mevrouw, ik wil graag met de trein naar Zwolle. Zou ik van u een retourtje Zwolle mogen hebben, alstublieft?”*

Voorbeeld (1) van Filipinski (2014, p. 13) laat zien, dat spreker 1 weinig woorden en tijd gebruikt om zijn communicatief doel te bereiken. De boodschap is dus efficiënt en effectief. Efficiënt, omdat hij alleen relevante woorden gebruikt in een korte tijd. De boodschap is effectief omdat het communicatieve doel - namelijk treinkaarten naar Zwolle te kopen - bereikt is. Spreker 2 bereikt ook zijn communicatief doel, dus hij is effectief, maar hij heeft meer woorden en tijd nodig om zijn communicatief doel te bereiken. Hij is daardoor minder efficiënt. Voorbeeld (2) zal een indruk geven van mensen met afasie. Het communicatieve doel is nog steeds een treinkaartje naar Zwolle te kopen.

(2) Spreker 3: *“Fiets...eh...morgen...nee...koffie.”*

Spreker 4: *“Goedemiddag mevrouw, ik wil graag met de fiets, ja of anders om naar, ja dat weet ik niet meer met de koffie. Zou ik dan dit mogen hebben, alstublieft?”*

Voorbeeld (2) laat zien, dat spreker 3 nog steeds efficiënt is, omdat hij weinig woorden en weinig tijd nodig heeft om te spreken. Maar met de gekozen woorden kan het communicatieve doel, de trein kaarten naar Zwolle te kopen, niet bereikt worden, waardoor de spreker niet effectief is. Spreker 4 bereikt zijn communicatief doel niet en is daardoor niet effectief. Bovendien is hij niet efficiënt, omdat hij veel woorden en tijd nodig heeft om te spreken.

Tabel 1 laat de verhouding tussen communicatieve efficiëntie en effectiviteit zien. Het is mogelijk het communicatieve doel met behulp van weinig woorden in een korte tijd te bereiken (efficiënt en effectief) of met veel woorden in een wat langere tijd (niet efficiënt, wel effectief). Aan de andere kant is het ook mogelijk, zijn communicatief doel niet te bereiken door weinig woorden (wel efficiënt, niet effectief) of veel woorden (niet efficiënt, niet effectief) te gebruiken die buiten de context vallen en niet relevant zijn.

Tabel 1: *Verhouding efficiëntie (CU's per minuut) en effectiviteit. Belangrijk te vermelden is, dat spreker 3 tenminste één CU moet produceren om efficiënt te kunnen zijn.*

Spreker	Efficiënt	Effectief
Spreker 1	+	+
Spreker 2	-	+
Spreker 3	+	-
Spreker 4	-	-

Om de communicatieve efficiëntie te kunnen meten is het volgens Lotgering (2009) noodzakelijk de uitgelokte response in telbare eenheden te verdelen en kwantificeerbaar te maken. Blomert, Koster, Van Mier en Kean (1987) onderscheidde elementen binnen een uiting als noodzakelijk (essentieel) of relevant. Relevante elementen zijn bijvoorbeeld begroetingen. Een begroeting is wel relevant binnen de communicatie maar niet noodzakelijk om het communicatieve doel te bereiken. Het is belangrijk aan de hand van proefpersonen zonder afasie vast te stellen, welke informatie-eenheden noodzakelijk (essentieel) zijn om een bepaald communicatief doel te bereiken.

Bij de communicatieve efficiëntie gaat het om de snelheid, waarmee een spreker de essentiële informatie overbrengt (Yorkston & Beukelman, 1980; Blomert et al., 1994). Volgens Yorkston en Beukelman (1980) kan het aantal informatie-eenheden (woorden/CU's/syllabes/lettergrepen) per minuut een uitspraak doen over de communicatieve efficiëntie.

De vraag is of men efficiënt kan zijn zonder effectief te zijn. Als men ervan uitgaat dat de efficiëntie afhankelijk is van de geproduceerde essentiële informatie-eenheden (in ons geval de CU's/min), dan heeft men tenminste 1 CU nodig om überhaupt van efficiëntie te kunnen spreken. Dat betekent voor spreker 3 in tabel 1, dat er tenminste 1 CU door spreker 3 geproduceerd moet zijn, om van efficiëntie te kunnen spreken. De informatie-eenheden die geteld worden moeten bestaan uit essentiële informatie (bijvoorbeeld de CU's) om efficiënt te zijn.

Uit onderzoek van Yorkston en Beukelman (1980) kwam naar voren dat de proefpersonen met afasie langzamer spraken en minder efficiënt waren, dan proefpersonen zonder afasie. Yorkston en Beukelman (1980) hebben beschreven dat patiënten die wellicht op stoornisniveau niet meer vooruitgaan, wel nog vooruitgaan op de communicatieve efficiëntie en effectiviteit, waarbij de efficiëntie meestal nog meer vooruit gaat dan de effectiviteit. Daarom is het voor de ANTAT-CU4 van belang de onderdelen communicatieve efficiëntie en communicatieve effectiviteit in de beoordelingsmethode mee te nemen.

2.2. Van kwalitatief naar kwantitatief

Een *kwalitatieve beoordelingsmethode* is een inhoudelijke analyse waarmee een beschrijving van een bepaald linguïstisch niveau gegeven wordt (Boxum, Van der Scheer, & Zwaga, 2013). Doordat er alleen een beschrijving gegeven wordt en niet met telbare eenheden gewerkt wordt, heeft de kwalitatieve beoordelingsmethode het nadeel dat een objectieve beoordeling niet mogelijk is.

Het te meten construct in de ANTAT is de *verbale communicatie*, om dit construct te beoordelen is het door Blomert et al. (1995) in twee onderconstructies verdeelt, namelijk de *begrijpelijkheid* en de *verstaanbaarheid*. De kwalitatieve beoordeling van de ANTAT, bestaat uit twee 5-punt beoordelingsschalen. De A-schaal voor de begrijpelijkheid van de te communiceren boodschap en een de B-schaal voor de verstaanbaarheid van de uiting op zich, onafhankelijk van de betekenis van deze uiting (Blomert et al., 1995). De begrippen *begrijpelijkheid* en *verstaanbaarheid* worden middels een citaat verduidelijkt: “als de patiënt een neologisme produceert dat de luisteraar zonder enige moeite kan herhalen of transcriberen, dan is de uiting goed verstaanbaar, maar in het geheel niet begrijpelijk.” (Blomert et al., 2008, p.15).

Het vaardigheidsniveau van de *begrijpelijkheid* en *verstaanbaarheid* wordt in de ANTAT per scenario gescoord door de 5-puntschaal met niveaus: 'geheel niet', 'enigszins', 'middelmatig', 'redelijk' en 'goed' (Blomert et al, 1995). Maar hoe kan een hoge mate van overeenstemming in de beoordeling tussen beoordelaars gegarandeerd worden, als de beoordeling uitsluitend een uitdrukking is “(...) van de mate waarin de luisteraar de boodschap begrijpelijk vindt” (Blomert et al., 1995, p. 52). Een objectieve en betrouwbare beoordeling is niet mogelijk, omdat “(...) no external criteria (i.e., the amount of essential information to be conveyed) are provided for assigning scores on the (ordinal) 5-point Comprehensibility scale.” (Ruiter et al., 2011, p. 6). De beoordelaar moet zijn beoordeling baseren op eigen kennis en ervaring en wordt niet gesteund door een extern criterium. Deze kennis en ervaring zijn bij iedere beoordelaar verschillend, waardoor de motieven (namelijk de redenen waarom een bepaalde beoordeling gegeven wordt) tussen de beoordelaars zal verschillen. Een extern criterium zou hier kunnen helpen. Zoals al door Ruiter al. (2011) is vastgesteld “(...) it is not indicated in the test manual *which* elements of meaning are considered essential in order to achieve the communicative goal.” (Ruiter et al., 2011, p. 6). Het is wenselijk om een kwantitatieve beoordelingsmethode te gebruiken, zodat objectief de communicatieve efficiëntie en tevens de communicatieve effectiviteit gemeten kan worden (Yorkston & Beukelman, 1980; Ruiter et al., 2011).

Een *kwantitatieve beoordelingsmethode* houdt een analyse in die telbare maten gebruikt, waardoor de beoordelingsmethode objectief is (Boxum et al., 2013). Hierbij vormen gegevens van gezonde sprekers de norm (Boxum et al., 2013). De *Content-Unit Analysis* (Yorkston & Beukelman, 1980) is de kwantitatieve beoordelingsmethode die voor de ANTAT-CU4 gebruikt wordt. De CU-analyse gebruikt een extern criterium en is in staat de communicatieve effectiviteit te meten (Craig et al., 2007). Eerdere onderzoeken (Lotgering, 2009; Ruiter et al., 2011; Filipinski, 2014) hebben de ANTAT bij een groep proefpersonen zonder afasie afgenomen en de responsen orthografisch getranscribeerd. Vervolgens was het mogelijk vanuit deze responsen de Content Units te bepalen en een CU-schema op te stellen. Uiteindelijk zijn hieruit 3 verschillende CU-schema's ontstaan. De procedure van het opstellen van een CU-schema wordt in hoofdstuk 3.1.2. nader toegelicht.

Bij het meten van de functionele communicatie moet ook de communicatieve efficiëntie meetbaar gemaakt kunnen worden. De efficiëntie kan alleen objectief bepaald worden als de onderliggende informatie-eenheden van een uiting in CU's verdeeld zijn (Yorkston & Beukelman, 1980; Lotgering, 2009). Binnen deze masterscriptie wordt de communicatieve efficiëntie niet verder onderzocht.

Concluderend kan worden gesteld, dat door het bepalen van de CU's een kwantitatieve maat voor de communicatieve effectiviteit bestaat. Binnen deze masterscriptie wordt een nieuw CU-schema voor de ANTAT-CU4 opgesteld. Hiervoor hebben Giessen (2015) en Aan de Stegge (2015) de ANTAT-CU bij een groep van 60 proefpersonen zonder afasie afgenomen en orthografisch getranscribeerd. Aan de hand van de transcripties worden de CU's voor het nieuwe CU-schema voor de ANTAT-CU4 bepaald. De communicatieve efficiëntie wordt binnen deze masterscriptie niet geanalyseerd. Dit onderdeel zal in een vervolgonderzoek wel in de beoordelingsmethode van de ANTAT-CU4 moeten worden opgenomen.

2.3. Psychometrische kwaliteiten van een test

Voordat de ANTAT-CU4 gebruikt kan worden dient deze gestandaardiseerd en genormeerd te worden. Om te kijken welke psychometrische kwaliteiten belangrijk zijn voor het standaardiseren en normeren, worden de *richtlijnen voor kwaliteit* van de Commissie Testaangelegenheden Nederland (COTAN) van Evers, Lucassen, Meijer en Sijtsma (2010) gebruikt. Er wordt een overzicht gegeven van de psychometrische kwaliteiten die in toekomstig onderzoek geanalyseerd moeten worden.

Binnen de volgende secties wordt ook gekeken hoe de ANTAT (Blomert et al., 1995) en de Scenario Test (Van Der Meulen & Van De Sandt-Koenderman, 2013) gestandaardiseerd en genormeerd zijn. De Scenario Test wordt ook beschreven, omdat de ANTAT niet de enige test is die de functionele communicatievaardigheden meet. Op deze manier kan gekeken worden welke psychometrische kwaliteiten bij de Scenario Test van belang waren en hoe deze zijn geanalyseerd. Tussen de ANTAT en de Scenario Test bestaan een aantal belangrijke verschillen.

Het eerste verschil is dat de patiënt bij de Scenario Test zowel verbale als ook non-verbale communicatie mag gebruiken om zijn communicatief doel te bereiken. Een ander onderscheid is, dat de Scenario Test in dialogovorm met een supportieve partner (de testleider) wordt afgenomen (Van Der Meulen et al., 2013), terwijl de patiënt bij de afname met de ANTAT een monoloog houdt (Blomert et al., 1995). Hierdoor bestaat bij de Scenario Test een soort afhankelijkheid van de gesprekspartner, die bij de afname van de ANTAT niet bestaat. Een ander belangrijk verschil is, dat de Scenario Test bedoeld is voor mensen met een matige tot ernstige expressieve stoornis en voor mensen voor die helemaal geen verbale communicatie meer mogelijk is (Van Der Meulen et al., 2013). Deze mensen kunnen dan proberen middels non-verbale communicatie hun communicatief doel duidelijk te maken.

2.3.1. Normen

Een normering moet regelmatig herzien worden. De richtlijn van de COTAN (Evers et al., 2010) vereist dat zodra er wijzigingen in de test zelf of in de instructie (bijvoorbeeld het veranderen of vernieuwen van de scenario's in de ANTAT), de normering helemaal nieuw opgesteld moet worden. De ANTAT uit 1995 werd voor het laatst in 1996 door de COTAN beoordeeld (Egberink, Janssen, & Vermeulen, 2010). Het criterium *normen* is op dat moment (2016) van de COTAN als *onvoldoende* beoordeeld, omdat de normen vanwege veroudering niet meer bruikbaar zijn (Egberink et al., 2010). De laatste normering was in 1994 op een Nederlands - Duitse steekproef van 465 patiënten (ANTAT I) en 473 patiënten (ANTAT II). ANTAT I en ANTAT II zijn twee parallelle versies van de ANTAT, die ieder 10 scenario's bevatten. Op deze manier is het mogelijk de ANTAT binnen een korte tijd een tweede keer af te nemen. Volgens de COTAN zijn normen na 15 jaar zonder hernormering verouderd en na 20 jaar niet meer bruikbaar (Evers et al., 2010).

In het Duitse beoordelingssysteem voor de kwaliteit van tests van Kersting (2006) wordt zelfs een periode van acht jaar voor hernormering aangeraden. Volgens de APA - richtlijnen uit 1999 wordt geen vast termijn vastgelegd voor de hernormering: "... so long as the test remains in print, it is the publisher's responsibility to assure that the test is renormed with sufficient frequency to permit continued accurate and appropriate score interpretations" (APA; 1999, p. 59, Standard 4.18).

Om een test opnieuw te normeren zijn er richtlijnen nodig over de grootte van een normgroep. In de literatuur zijn er in beperkte mate aanbevelingen te vinden over de grootte van een normgroep of men vindt tegenstrijdige informatie (Evers et al., 2010). Het beoordelingssysteem van de COTAN geeft richtlijnen voor de steekproefgrootte per normgroep en ook voor de steekproefgrootte op groepsniveau. Hierbij is niet duidelijk waarop deze gegevens gebaseerd zijn, omdat er beschreven staat dat er geen aanbevelingen in de literatuur ter onderbouwing gevonden werden (Evers et al., 2010, p. 22).

Tests voor belangrijke* beslissingen op individueel niveau (bijvoorbeeld personeelsselectie, verwijzing naar speciaal onderwijs, opname/ontslag kliniek, certificering).	$N \geq 400$ $300 \leq N < 400$ $N < 300$	goed voldoende onvoldoende
Tests voor relatief minder belangrijke beslissingen op individueel niveau (bijvoorbeeld voortgangscntrole, in het algemeen beschrijvend gebruik, zoals bij beroepskeuzebegeleiding, therapie-indicatie).	$N \geq 300$ $200 \leq N < 300$ $N < 200$	goed voldoende onvoldoende
* Met belangrijke beslissingen wordt bedoeld: beslissingen die op basis van de testscores worden genomen, die in principe, of op korte termijn, onomkeerbaar zijn, en die voor een belangrijk deel buiten de geteste persoon om worden genomen.		

Figuur 1: Steekproefgrootte per normgroep (Evers et al., 2010).

Tests voor onderzoek op groepsniveau.	$K \geq 40$ $30 \leq K < 40$ $K < 30$	goed voldoende onvoldoende
---------------------------------------	---	----------------------------------

Figuur 2: Steekproefgrootte op groepsniveau (Evers et al., 2010).

Zoals in figuur 1 te zien is, zijn er twee mogelijke eenheden voor een steekproefgrootte. De eerste richt zich op tests voor belangrijke beslissingen en de tweede richt zich op tests voor minder belangrijke beslissingen. Voor de tests voor belangrijke beslissingen geldt, dat als de steekproefgrootte kleiner dan $n=300$ is het criterium normen van de COTAN met een onvoldoende beoordeeld wordt. Voor de tests voor minder belangrijke beslissingen ligt die grens lager, namelijk pas bij een steekproefgrootte van kleiner dan $n=200$ wordt de test voor het criterium normen door de COTAN als onvoldoende beoordeeld.

Om de ANTAT-CU4 te normeren is dus een steekproefgrootte van tenminste $n=200$ noodzakelijk. Deze steekproef moet alle hoofdsyndromen van afasie (afasie van Broca, afasie van Wernicke, globale afasie en amnestische afasie (Dharmaperwira-Prins & Maas, 2006)) en overige afasiesyndromen (conductie-afasie, transcorticaal-motorische afasie, transcorticaal-sensorische afasie en gemengde transcorticale afasie (Dharmaperwira-Prins & Maas, 2006)) bevatten. Dat zal voor toekomstig normeringsonderzoek van de ANTAT-CU4 betekenen, dat er per afasie syndroom steekproeven op groepsniveau getrokken moeten worden. Figuur 2 laat richtlijnen van de COTAN zien over de grootte van een steekproef op groepsniveau. Zo mag volgens de COTAN een steekproefgrootte op groepsniveau niet kleiner dan $n=30$ proefpersonen zijn, anders wordt het criterium normen van de COTAN als onvoldoende beoordeeld. Voor de steekproefgrootte voor de ANTAT-CU4 betekent dat, dat er per afasiesyndroom groepen van $n \geq 30$ gebruikt moeten worden.

2.3.2. Componenten van betrouwbaarheid

Volgens Field (2009) geeft de *betrouwbaarheid* aan in hoeverre een meting zonder meetfouten is. Betrouwbaarheid bestaat uit drie componenten. Ten eerste de *test-hertest betrouwbaarheid*, ten tweede de *interbeoordelaar betrouwbaarheid* en ten derde de *intrabeoordelaar betrouwbaarheid*. Deze componenten worden in de volgende secties verder toegelicht.

Als er gesproken wordt over betrouwbaarheid, is het belangrijk het verschil tussen *overeenstemming* (agreement) en *betrouwbaarheid* (reliability) duidelijk te maken, omdat "(...) reliability and agreement measure different aspects of a set of measuring instruments" (Rietveld, 2013, p. 13). Betrouwbaarheid is een relatieve maat, die vaak uitgedrukt wordt in een correlatie (Verhagen & Alessie, 2014). Overeenstemming heeft vooral betrekking op de absolute meetfout van de test (De Vet, Terwee, Knol, & Bouter, 2005). Zo is het voor het vervolg belangrijk te begrijpen, dat alle componenten van betrouwbaarheid alleen betrouwbaar zijn, als er een hoge mate van overeenkomst tussen bijvoorbeeld de test en de hertest, of tussen twee of meer beoordelaars, of binnen een en dezelfde beoordelaar bestaat (Verhagen & Alessie, 2014).

Betrouwbaarheid wordt als statistische maat met Cronbach's alfa uitgedrukt, wat volgens Field (2009) een goede coëfficiënt is. Deze betrouwbaarheidscoëfficiënt is een maat die de onderlinge samenhang van de indicatoren tot uitdrukking brengt. Cronbach's alfa is een goede maat voor de interne consistentie, hij laat zien hoe consistent de items gescoord zijn (Van der Meulen, Van De Sandt - Koenderman, Duivenvoorden, & Ribbers, 2010). Volgens Wijk (2010) zijn de criteria voor Cronbach's alfa als volgt te interpreteren: $\alpha > .80$ goed; $.70 - .80$ adequaat; $.60 - .70$ matig; $< .60$ onvoldoende (geen schaal).

Test - hertest betrouwbaarheid

In deze sectie wordt de *test-hertest betrouwbaarheid*, de *teststabiliteit* en de *testparalliteit* (Allen & Yen, 1979) nader toegelicht. Deze componenten van betrouwbaarheid maken deel uit van de gebruikte onderzoeksmethode binnen deze masterscriptie (zie hiervoor hoofdstuk 3).

Volgens de COTAN wordt de generaliseerbaarheid en de stabiliteit van een test bepaald door test-hertest correlaties (Evers et al., 2010). Het idee achter een test-hertest procedure is, dat de test bij dezelfde groep proefpersonen herhaald wordt, waarbij het tijdsinterval tussen de twee meetmomenten nauwkeurig vermeld wordt (Brinkmann, 2011). Men hoopt bij een test-hertest procedure dat de scores op meetmoment 1 en op meetmoment 2 (bijna) gelijk aan elkaar zijn, dus dat er een goede overeenstemming te zien is. Op het moment dat er een hoge correlatie (overeenstemming) tussen de twee meetmomenten bestaat, bestaat er een goede teststabiliteit en een hoge maat aan betrouwbaarheid (Field, 2009). Volgens Weiß (2010) is de correlatie tussen test en hertest goed, als de correlatiecoëfficiënt zo dicht mogelijk aan 1 ligt en een slechtere correlatie is te zien als de correlatiecoëfficiënt dicht bij 0 ligt.

Een belangrijke assumptie voor de test-hertest procedure is dat de variabelen van de proefpersoon niet veranderen tijdens het testinterval (Field, 2009). Dat houdt in dat er proefpersonen getest moeten worden die geen interventie tussen de meetmomenten hebben gevolgd.

In deze masterscriptie worden gezonde proefpersonen zonder afasie ($n=30$) binnen een test-hertest procedure getest die geen interventie tussen beide meetmomenten hebben gevolgd. Op deze manier kan in een vervolgonderzoek de teststabiliteit van de ANTAT-CU4 geanalyseerd worden. Men hoopt bij de groep proefpersonen zonder afasie ($n=30$) geen verschillen tussen de scores op meetmoment 1 en meetmoment 2 te vinden.

De proefpersonen met afasie ($n=10$) die in deze masterscriptie ook middels een test-hertest procedure getest worden, krijgen wel een interventie tussen de meetmomenten. Het verschil is dat er middels de proefpersonen met afasie ($n=10$) niet de teststabiliteit van de ANTAT-CU4 onderzocht wordt, maar of de ANTAT met de nieuwe scenario's in staat is veranderingen in de functionele communicatie waar te nemen. Men hoopt hier dus verschillende scores op de test en de hertest te vinden om iets over de veranderingen van de functionele communicatie te kunnen zeggen.

Een tweede assumptie van de test-hertest procedure is dat de items van een test voor iedere proefpersoon random aangeboden moeten worden (Brinkmann, 2011). Dat zou voor de ANTAT-CU4 betekenen dat de scenario's steeds bij iedere proefpersoon in een random volgorde aangeboden zouden moeten worden. Dit is bij de ANTAT-CU4 niet mogelijk, omdat de scenario's volgens een vaste volgorde afgenomen moeten worden (Blomert et al., 2008). Dus aan deze assumptie kan bij de ANTAT-CU4 niet voldaan worden.

Om te kunnen interpreteren hoe goed de test en de hertest correleren, moet een statistische test uitgevoerd worden. Blomert, Kean, Koster en Schokker (1994) hebben een factoranalyse (Jöreskog & Sörbom, 1988) en een t-toets voor gepaarde steekproeven voor de ANTAT test-hertest analyse (met een tussentijdse periode van 3 maanden) gebruikt. Er kwam naar voren, dat de ANTAT een stabiel meetinstrument is en daarmee geschikt voor het meten van veranderingen in de tijd (Blomert et al., 1995). Van der Meulen et al. (2010) hebben bij de normering en standaardisering van de Scenario Test bij de test-hertest procedure ook gebruik gemaakt een factoranalyse en van de *intra class correlatiecoëfficiënt* (ICC; Shrout & Fleiss, 1976). De test-hertest betrouwbaarheid voor de Scenario Test was hoog ($ICC=0.89$).

Concluderend kan worden gezegd, dat de ANTAT van Blomert et al. (1995) een goede teststabiliteit laat zien, maar dat de scenario's van de ANTAT vernieuwd moeten worden (zie Filipinski, 2014; Giessen, 2015; Aan de Stegge, 2015). Dat heeft tot gevolg dat de test-hertest betrouwbaarheid van de ANTAT met de nieuwe scenario's opnieuw onderzocht moet worden. Dat betekent voor de ANTAT-CU4 dat een test-hertest procedure gebruikt moet worden, om iets over de betrouwbaarheid en stabiliteit van de ANTAT-CU4 te kunnen zeggen. In deze masterscriptie wordt middels een kleine groep proefpersonen zonder afasie ($n=6$) een eerste indruk gegeven hoe de betrouwbaarheid en stabiliteit van de ANTAT-CU4 eruitziet (zie hoofdstuk 3).

De term *testparalleliteit* is voor de ANTAT-CU4 belangrijk, omdat de ANTAT oorspronkelijk uit twee parallelle versies bestond. Daardoor is het mogelijk de ANTAT binnen een korte tijd twee keer af te nemen, zodat er geen leereffect optreedt. De testparalleliteit houdt voor de ANTAT in, dat de ANTAT I (10 items) en de ANTAT II (10 items) parallelle versies van elkaar zijn. Dat betekent dat beide versies dezelfde resultaten zouden moeten laten zien, als zij bij een en dezelfde persoon afgenomen worden.

De test paralleliteit is voor de ANTAT-CU4 nog niet onderzocht. Binnen deze masterscriptie wordt een advies gegeven, hoe de scenario's van de ANTAT-CU4 eventueel onderverdeeld zouden kunnen worden op twee testversies (zie hiervoor hoofdstuk 4). De testparalleliteit moet in een vervolgonderzoek nader onderzocht worden.

Interbeoordelaars- overeenstemming en betrouwbaarheid

Komen de beoordelaars goed overeen met hun beoordelingen, dan stijgt de betrouwbaarheid (De Vet et al., 2006). De overeenstemming is een absolute maat voor de overeenstemming tussen twee beoordelaars gegeven een groep proefpersonen (Verhagen en Alessie, 2014). Voor de interbeoordelaars- overeenstemming tussen twee beoordelaars kan *Cohens Kappa* (Cohen, 1960) als statistische maat gebruikt worden. Kappa is een maat voor het percentage overeenstemming dat wordt gecorrigeerd voor de toevallige overeenstemming (Verhagen en Alessie, 2014). Dat betekent, dat Kappa voor toevallige overeenstemmingen tussen beoordelaars corrigeert. Voor meer dan twee beoordelaars wordt er gebruik gemaakt van de statistische maat *Krippendorff's alfa* (Krippendorff, 1970; Sage researchmethods, 2013). Cohens Kappa en Krippendorff's alfa kunnen een waarde tussen 0 en 1 aannemen, waarbij een waarde van 0.70 aanwezig moet zijn om van een goede overeenstemming te kunnen spreken (Weiß, 2010). Hierbij betekent een waarde van 1 dat er een complete overeenstemming bestaat en een waarde van 0 dat er helemaal geen overeenstemming in de beoordeling tussen de twee beoordelaars bestaat (Field, 2009).

De *intraclass correlatiecoëfficiënt* (ICC; Shrout & Fleiss, 1976) is ook een methode om de overeenstemming tussen meerdere beoordelaars te bepalen (Verhagen en Alessie, 2014). De ICC is gedefinieerd als het aandeel ware variantie ten opzichte van de totale variantie en is daarmee net als de kappa een relatieve maat (Shrout & Fleiss, 1979). De ICC kan een waarde tussen 0 en 1 aannemen. Als er maar kleine variaties bestaan tussen de scores die de beoordelaars per item hebben gegeven, zal de ICC een hoge waarde (<.70) aannemen.

Bij de originele ANTAT moet de beoordelaar de taaluitingen door middel van twee 5-punt schalen per scenario op *begrijpelijkheid* en *verstaanbaarheid* beoordelen (Blomert et al., 1995). Blomert (1988) concludeerde in zijn onderzoek, dat getrainde beoordelaars in hun beoordeling bij de ANTAT nauwelijks van elkaar afweken. In Blomert et al. (1995) staat beschreven, dat de interbeoordelaars-betrouwbaarheid "van een ruim voldoende niveau" (Blomert et al., 1995, p. 20) is. Ruiten et al. (2011) hebben echter aangetoond dat middels een kwantitatieve beoordelingsmethode vooral de begripsvaliditeit verder verbeterd kan worden.

Concluderend kan worden gezegd, dat de interbeoordelaars- betrouwbaarheid van de ANTAT (1995) van een ruim voldoende niveau is (Blomert et al., 1995), maar dat de beoordelingsmethode verder verbeterd zou kunnen worden. De beoordelingsmethode die door Blomert et al. (1995) gebruikt wordt is een kwalitatieve methode. Bij een schaalbeoordeling kan variabiliteit binnen de beoordelingen optreden, omdat iedere beoordelaar een eigen persoonlijk idee van de beoordeling van de verstaanbaarheid en de begrijpelijkheid heeft (Bühner, 2004). Dit kan ondervangen worden door de kwalitatieve beoordelingsmethode te vervangen door een kwantitatieve beoordelingsmethode (het CU-schema). De nieuwe kwantitatieve beoordelingsmethode middels CU-schema wordt binnen deze

masterscriptie voor de ANTAT-CU4 ontwikkeld. De interbeoordelaars- betrouwbaarheid wordt binnen deze masterscriptie niet nader onderzocht, maar moet binnen een vervolgonderzoek verder onderzocht worden.

Intrabeoordelaars- overeenstemming en betrouwbaarheid

Het oordeel van een en dezelfde persoon hoeft niet altijd consistent te zijn. Bij de intrabeoordelaars-overeenstemming wordt gekeken in hoe ver de beoordelaar het met zichzelf eens is. Komt een beoordelaar niet met zijn eigen mening overeen, dan is de overeenstemming en vervolgens ook de betrouwbaarheid verminderd (Verhagen & Alessie, 2014). De intrabeoordelaars- betrouwbaarheid laat de mate van overeenstemming tussen de uitkomsten van de test en hertest meeting, uitgevoerd door dezelfde beoordelaar zien (Brinkmann, 2011). Hiervoor kan men weer de statistische maat *Cohens Kappa* (Cohen, 1960) gebruiken.

Over de intraobservator- betrouwbaarheid konden geen informatie gevonden worden in de handleidingen van de ANTAT (Blomert et al., 1995; Blomert et al., 2008). Voor de ANTAT-CU4 geldt, dat in toekomstig onderzoek ook de intrabeoordelaars- betrouwbaarheid van de ANTAT-CU4 gecontroleerd zou moeten worden.

2.3.3. Componenten van validiteit

De klassieke driedeling voor validiteitsonderzoek, zoals die door de *Richtlijnen voor ontwikkelen en gebruik van psychologische testen en studietoetsen* (Evers, Caminada et al., 1988) zijn vastgelegd, zijn de *inhoudsvaliditeit*, de *begripsvaliditeit* en de *criteriumvaliditeit*. De COTAN beoordeelt alleen de begripsvaliditeit en de criteriumvaliditeit (Evers et al., 2010). Deze vormen van validiteit worden besproken in het volgende segment.

Inhoudsvaliditeit (content validiteit)

Inhoudsvaliditeit betekent, dat een meetinstrument alle aspecten van het te meten construct meet (Brinkmann, 2011). Het construct wat door de ANTAT gemeten moet worden is de *verbale communicatie*. Blomert et al. (2008) stellen in de handleiding van de ANTAT de volgende vraag voor de toetsing van de inhoudsvaliditeit: "In hoeverre zijn de testscenario's representatief voor de situaties die een afasiepatiënt meemaakt in het alledaagse leven?" (p.31). Uit deze vraag komt naar voren dat het gaat om de scenario's die getest worden. Deze moeten actueel en voor een groot aantal mensen representatief zijn (Blomert et al., 1996): "De scenario's moeten van een voldoende algemeen karakter zijn om voor iedere willekeurige volwassene herkenbaar en voorstelbaar te zijn, onafhankelijk van de unieke aspecten van de achtergrond en ervaring van een gegeven individu. Daarnaast moet het waarschijnlijk geacht worden dat dit individu gelijkwaardige, zo niet identieke, situaties meer dan eens daadwerkelijk heeft meegemaakt." (Blomert et al., 2008, p.31). In de jaren 1996 toen de ANTAT beoordeeld werd door de COTAN, voldeden de scenario's van de ANTAT aan alle criteria.

Anno 2016 voldoen sommige scenario's niet meer aan de itemcriteria van Blomert et al. (1995). Filipinski (2014) heeft de originele scenario's middels een vragenlijst laten beoordelen op de door Blomert et al. (1996) opgestelde criteria, te weten: actualiteit, inleefbaarheid en duidelijkheid. Hieruit kwam naar voren dat 19 van de 20 scenario's niet aan een of meerdere criteria van Blomert et al. (1996) voldeden.

Giessen (2015) en Aan de Stegge (2015) hebben nieuwe scenario's ontwikkeld. De originele en de nieuwe scenario's zijn door een groep proefpersonen zonder afasie ($n=60$) op de itemcriteria beoordeeld. Van de 15 nieuwe scenario's zijn zes scenario's (namelijk scenario 4, 5, 16, 17, 19 en 20) op inleefbaarheid en actualiteit als significant beter beoordeeld dan de originele scenario's (Giessen, 2015; Aan de Stegge, 2015). Er zijn geen significante verschillen gevonden tussen originele en nieuwe scenario's wat betreft de beoordeling op rolverdeling. Dat betekent, dat het criterium rolverdeling zowel bij de oude scenario's als ook bij de nieuwe scenario's door de testpersonen goed beoordeeld werd.

Om de verbale communicatie daadwerkelijk valide meetbaar te kunnen maken, moet gegarandeerd zijn dat de scenario's van de ANTAT-CU4 aan alle itemcriteria voldoen. Met behulp van de uitkomsten uit de analyse van Giessen (2015) en Aan de Stegge (2015) wordt binnen deze masterscriptie een beslissing gemaakt welke scenario's (origineel of nieuw) opgenomen worden in de ANTAT-CU4.

Begripsvaliditeit (constructvaliditeit)

De begripsvaliditeit is ook een criterium van het beoordelingssysteem van de COTAN (Evers et al., 2010). Volgens Van der Meulen et al. (2010) wordt de begripsvaliditeit als volgt gedefinieerd: "Construct validity refers to the extent to which a test measures what it claims to measure. A valid test allows appropriate and meaningful inferences." (p.428).

Het te meten construct van de ANTAT is de *verbale communicatie*. Blomert et al. (2008) stellen daarom de vraag: "In hoeverre kan de testscore beschouwd worden als een maat voor het begrip *verbaal communicatieve vaardigheid*, dat de test beoogt te meten?" (p.31). Het construct *verbale communicatie* bestaat uit twee delen, de *begrijpelijkheid* en de *verstaanbaarheid* (Blomert et al., 1994).

Begripsvaliditeit kan verder onderscheiden worden in *divergente validiteit* en *convergente validiteit*. *Divergente validiteit* is het vergelijken van twee testen die niet hetzelfde construct meten (Brinkmann, 2011). De divergente validiteit wordt als volgt geïnterpreteerd: naarmate de correlatie tussen de testen daalt, stijgt de validiteit, omdat ze niet hetzelfde construct meten.

De *convergente validiteit* wordt ook door Van der Meulen et al. (2010) en door Blomert et al. (1994) gebruikt. Hiervoor vergelijkt men twee testen die beide hetzelfde construct meten. Om de correlatie tussen de testen te controleren wordt de statistische maat Spearman's r correlatiecoëfficiënt (Spearman, 1910) gebruikt. Men hoopt een hoge correlatie tussen de testen te vinden. Een waarde van 1 is een indicatie voor een perfecte correlatie. Indien twee testen laag met elkaar correleren, meten ze twee verschillende constructen. Het kan zijn dat beide testen een totaal ander construct meten dan het beoogde construct, of één test het beoogde construct meet en de andere test een ander construct meet. Een waarde van 0 laat helemaal geen correlatie zien.

Blomert et al. (1995) hebben de originele ANTAT vergeleken met het *semigestructureerd interview* van de *Akense Afasie Test* (AAT Nederlandse versie; Graetz, De Bleser, & Willmes, 1992). Zij vonden een hoge correlatie, namelijk 0.81 (ANTAT I) en 0.83 (ANTAT II), het significantieniveau werd in de handleidingen (Blomert et al., 1995; Blomert et al., 2008) niet weergegeven. De hoge correlatie (>.70) laat zien dat deze twee tests bijna hetzelfde construct meten.

Bovendien relateerden Blomert et al. (1995) de ANTAT-score aan het oordeel van de partner van de patiënt. Omdat er toen die tijd nog geen vragenlijst voor de partner bestond, heeft Blomert (1993) zelf een vragenlijst ontwikkeld: de Afasie Partner Vragenlijst (APV; Blomert, 1993). De scores van de ANTAT zijn dan vergeleken met de scores van de vragenlijst APV (Ritzen, Spreuwenberg, & Sensesen, 2005). De resultaten lieten een sterke relatie (correlatie van .69) tussen de A-schaal (voor begrijpelijkheid) en de APV-vragenlijst zien. Dit laat volgens Blomert et al. (2008) zien, dat "De sterke samenhang tussen het oordeel van de partner over de daadwerkelijke effectiviteit van de verbale communicatie van de patiënt en de testprestatie van de patiënt bevestigt dat de ANTAT een valide maat is voor de verbale communicatieve vaardigheid van de patiënt in zijn dagelijkse omgeving." (Blomert et al., 2008, p. 35).

Van der Meulen et al. (2010) vergeleken de Scenario Test met de ANTAT en de Communicative Effectiveness Index (CETI; Lomas et al., 1989). De CETI is een vragenlijst en wordt gebruikt om de alledaagse communicatievaardigheden van mensen met afasie door naasten (de partner, de kinderen etc.) te laten beoordelen. Er kwam naar voren, dat de correlatie tussen de Scenario Test en de ANTAT

sterk was (0.85, $p=.001$). Ook de correlatie tussen de Scenario Test en het onderdeel semigestructureerd interview van de AAT was sterk, namelijk 0.79 ($p=.001$) (Van der Meulen et al., 2010). Dit laat zien, dat deze testen onderliggend bijna hetzelfde construct meten. De correlatie tussen de Scenario Test en de CETI was matig, (0.5), maar wel significant ($p=.001$) (Van der Meulen et al., 2010). De matige correlatie tussen de Scenario Test en de CETI is te verklaren door het feit dat de CETI geen direct meetinstrument is, maar een vragenlijst. Hierdoor is het moeilijk overeenkomsten tussen de testen te vinden (Van der Meulen et al., 2010).

Ruiter et al. (2011) hebben de begripsvaliditeit van de ANTAT verder verbeterd door de kwalitatieve beoordelingsmethode te vervangen door een kwantitatieve beoordelingsmethode. De scoring van de ANTAT-CU4 wordt volgens een kwantitatieve beoordelingsprocedure door middel van Content Units gedaan, zoals in 2011 door Ruiter et al. beschreven. Daarom wordt voor de ANTAT-CU4 binnen deze masterscriptie het nieuwe Content Unit schema (CU-schema) opgesteld. Voor toekomstig onderzoek is het noodzakelijk de ANTAT-CU4 voor het normerings- en standaardiseringsonderzoek met soortgelijke tests te vergelijken, om te onderzoeken of het onderliggende construct *verbale communicatie* daadwerkelijk gemeten wordt. Hiervoor zal bijvoorbeeld de Scenario Test of het semigestructureerd interview van de AAT gebruikt kunnen worden. Op deze manier kan in toekomst onderzoek de begripsvaliditeit van de ANTAT-CU4 geanalyseerd worden.

Criterionvaliditeit

Volgens Blomert et al. (1994) is de criteriumvaliditeit als volgt te beschrijven: "...the score of a given patient can be taken as an index of his verbal communicative abilities in daily life" (p.397). Vanuit deze definitie stellen Blomert et al. (2008) volgende vraag: "In hoeverre vertoont de ANTAT een samenhang met andere maten van de communicatieve vaardigheid van afasiepatiënten?" (p.31).

De COTAN heeft de *criterionvaliditeit* van de originele ANTAT uit de jaren 1995 met *voldoende* beoordeeld. Blomert et al. (1995) hebben voor het criterium validiteit de ANTAT met een soortgelijke test vergeleken. Ze gebruiken de *Akense Afasie Test* om te kijken hoe goed deze met de ANTAT correleert. Van der Meulen et al. (2010) passen deze procedure alleen voor de begripsvaliditeit toe en ze beschrijven niet of en hoe ze de criteriumvaliditeit hebben gemeten.

2.3.4. Sensitiviteit

De sensitiviteit van een meetinstrument geeft aan hoe gevoelig het meetinstrument kleine verschillen over de tijd kan vaststellen (Brinkmann, 2011). De vereiste gevoeligheid van een meetinstrument hangt van het doel van de test af. De ANTAT-CU4 moet gevoelig genoeg zijn, om mensen met en zonder afasie zinnig te kunnen onderscheiden. Dat betekent dus ook, dat een proefpersoon met afasie (positief) onderscheiden moet kunnen worden van een proefpersoon zonder afasie (negatief) (Weiß, 2010). Blomert beschrijft dit in 1994 als het *pathologie criterium*. Dit criterium veronderstelt, dat proefpersonen zonder afasie op de ANTAT de maximale testscore moeten behalen.

In Van der Meulen et al. (2010) wordt gezegd dat het bovendien belangrijk is ook kleine verschillen, dus lichte verbeteringen of verslechtingen, te kunnen meten. Ze hebben voor de sensitiviteit van de Scenario Test met de gepaarde t-toets gewerkt. Het is ook voor de ANTAT-CU-4 belangrijk dat hij sensitief genoeg is om verschillen over de tijd te kunnen meten. Alleen op deze manier zal het mogelijk zijn vooruitgang door therapie in de functionele communicatie te kunnen meten.

2.4. Doelstelling

Binnen dit onderzoek worden de scenario's van de ANTAT naar aanleiding van het onderzoek van Giessen (2015) en Aan de Stegge (2015) deels vervangen of geactualiseerd. Door de nieuwe samenstelling van de scenario's ontstaat de ANTAT-CU4. Van de gekozen scenario's worden de CU's opnieuw vastgesteld. Hiervoor worden de data van Giessen (2015) en Aan de Stegge (2015) van 60 proefpersonen zonder afasie gebruikt en zal een CU-schema voor de ANTAT-CU4 ontstaan.

- **Doel 1:** *Binnen deze masterscriptie wordt de ANTAT-CU4 opgesteld.*
- **Doel 2:** *Binnen deze masterscriptie wordt een nieuw CU-schema voor de ANTAT-CU4 ontwikkeld.*

De oorspronkelijke ANTAT bestaat uit twee parallelle testversies. De scenario's van de ANTAT-CU4 zijn nog niet verdeeld in twee testversies, omdat nog niet bekend is hoeveel CU's er per scenario in de ANTAT-CU4 zijn. Door het opstellen van het CU-schema voor de ANTAT-CU4 zal het achteraf mogelijk zijn een advies uit te spreken hoe de scenario's op twee testversies verdeeld zouden kunnen worden. Door een vervolgonderzoek zal dan de testparalleliteit van de ANTAT-CU4 nader onderzocht kunnen worden.

- **Doel 3:** *Naar aanleiding van deze masterscriptie wordt een advies gegeven, hoe de scenario's op twee ANTAT-CU4 testversies onderverdeeld zouden kunnen worden.*

De ANTAT-CU4 wordt middels een test-hertest procedure met een tussentijdse periode van 8 weken bij een Nederlandstalige groep proefpersonen ($n=30$) afgenomen om digitale audiodata te verzamelen. Deze audiodata zullen in een opvolgend onderzoek gebruikt worden om de test-hertest stabiliteit en betrouwbaarheid van de ANTAT-CU4 te bepalen.

- **Doel 4:** *In het kader van deze masterscriptie worden de data van Nederlandstalige proefpersonen zonder afasie ($n=30$) verzameld die in een opvolgend onderzoek gebruikt zullen worden om de test-hertest stabiliteit en betrouwbaarheid van de ANTAT-CU4 te bepalen.*

De Nederlandstalige ANTAT-CU4 wordt middels een equivalentieonderzoek (Evers et al., 2010) naar het Duits vertaald, zodat een Duitstalige ANTAT met nieuwe scenario's ontstaat die binnen deze masterscriptie gebruikt wordt.

- **Doel 5:** *Binnen deze scriptie wordt middels equivalentieonderzoek de Nederlandstalige ANTAT-CU4 vertaald naar het Duits, zodat een Duitstalige ANTAT met nieuwe scenario's ontstaat.*

De Duitstalige ANTAT met nieuwe scenario's wordt op twee verschillende meetmomenten met een tussentijdse periode van 8 weken bij een groep Duitstalige proefpersonen met afasie ($n=10$) afgenomen. Deze data worden verzameld om binnen een vervolgonderzoek nader onderzocht te kunnen worden. Voor de Duitstalige ANTAT met nieuwe scenario's bestaat nog geen CU-schema. Voor toekomstig

onderzoek zal het een meerwaarde hebben om in een vervolgonderzoek een Duitstalig CU-schema op te stellen en de data van de proefpersonen met afasie ($n=10$) nader te onderzoeken.

- **Doel 6:** *In het kader van deze masterscriptie worden de data van Duitstalige proefpersonen met afasie ($n=10$) verzameld die in een opvolgend onderzoek gebruikt zullen worden om met behulp van een Duitstalig CU-schema de veranderingen van de functionele communicatie nader te onderzoeken.*

Om in de toekomst wel een Duitstalig CU-schema te kunnen gebruiken, wordt de Duitstalige ANTAT met nieuwe scenario's bij een groep van 30 Duitstalige proefpersonen zonder afasie afgenomen. Met behulp van deze audiodata zal in een vervolgonderzoek een Duitstalig CU-schema opgesteld kunnen worden.

- **Doel 7:** *Binnen deze masterscriptie worden data van Duitstalige proefpersonen zonder afasie ($n=30$) verzameld die in een vervolgonderzoek gebruikt kunnen worden om een Duitstalig CU-schema te kunnen opstellen.*

2.5. Onderzoekshypothesen

Binnen het huidige onderzoek wordt de ANTAT-CU4 middels een test-hertest procedure bij een groep proefpersonen ($n=30$) zonder afasie afgenomen. Het nieuwe CU-schema wordt bij een kleine groep proefpersonen zonder afasie ($n=6$) gebruikt om de testafnames te scoren. De ANTAT-CU4 scores van de twee meetmomenten worden vervolgens met behulp van een gepaarde t-toets geanalyseerd. Dit wordt gedaan om te controleren of de ANTAT-CU4 tussen de meetmomenten stabiel blijft. Aan de hand daarvan kan een eerste inschatting over de stabiliteit van de ANTAT-CU4 gegeven worden.

- **Hypothese 1:** *De scores op de ANTAT-CU4 (met CU-schema) op meetmoment 1 en meetmoment 2 met een tussentijdse periode van 8 weken bij een Nederlandstalige groep proefpersonen zonder afasie ($n=6$) laten geen significante verschillen zien.*

De Duitstalige proefpersonen met afasie ($n=10$) die met de Duitstalige ANTAT met nieuwe scenario's getest zijn, worden ook binnen een test-hertest procedure onderzocht. De data van 8 proefpersonen met afasie (2 proefpersonen zijn van het onderzoek uitgesloten) wordt hiervoor gebruikt. Bij deze groep proefpersonen hoopt men wel verschillende ANTAT-testscores tussen meetmoment 1 en meetmoment 2 te vinden, omdat men verwacht veranderingen in de functionele communicatie te zien. Het is belangrijk te vermelden, dat de data niet geanalyseerd kunnen worden met een Duitstalig CU-schema. De beoordeling van de scenario's wordt gedaan middels de oude kwalitatieve beoordelingsmethode (middels 5-put schaal).

- **Hypothese 2:** *De Duitstalige ANTAT met nieuwe scenario's en oude kwalitatieve beoordelingsmethode, laat veranderingen in de functionele communicatie bij Duitstalige mensen met afasie ($n=8$) zien.*

3. Methode

In dit hoofdstuk wordt beschreven welke stappen ondernomen zijn om een eerste indruk te geven van de test-hertest stabiliteit van de Nederlandstalige ANTAT-CU4. Ten eerste wordt ingegaan op de ontwikkeling van de ANTAT-CU4 en de ontwikkeling van het nieuwe Content Unit schema. Vervolgens wordt het onderzoeksontwerp voor de Nederlandstalige proefpersonen zonder afasie ($n=30$), de Duitstalige proefpersonen met afasie ($n=10$) en de Duitstalige proefpersonen zonder afasie ($n=30$) beschreven.

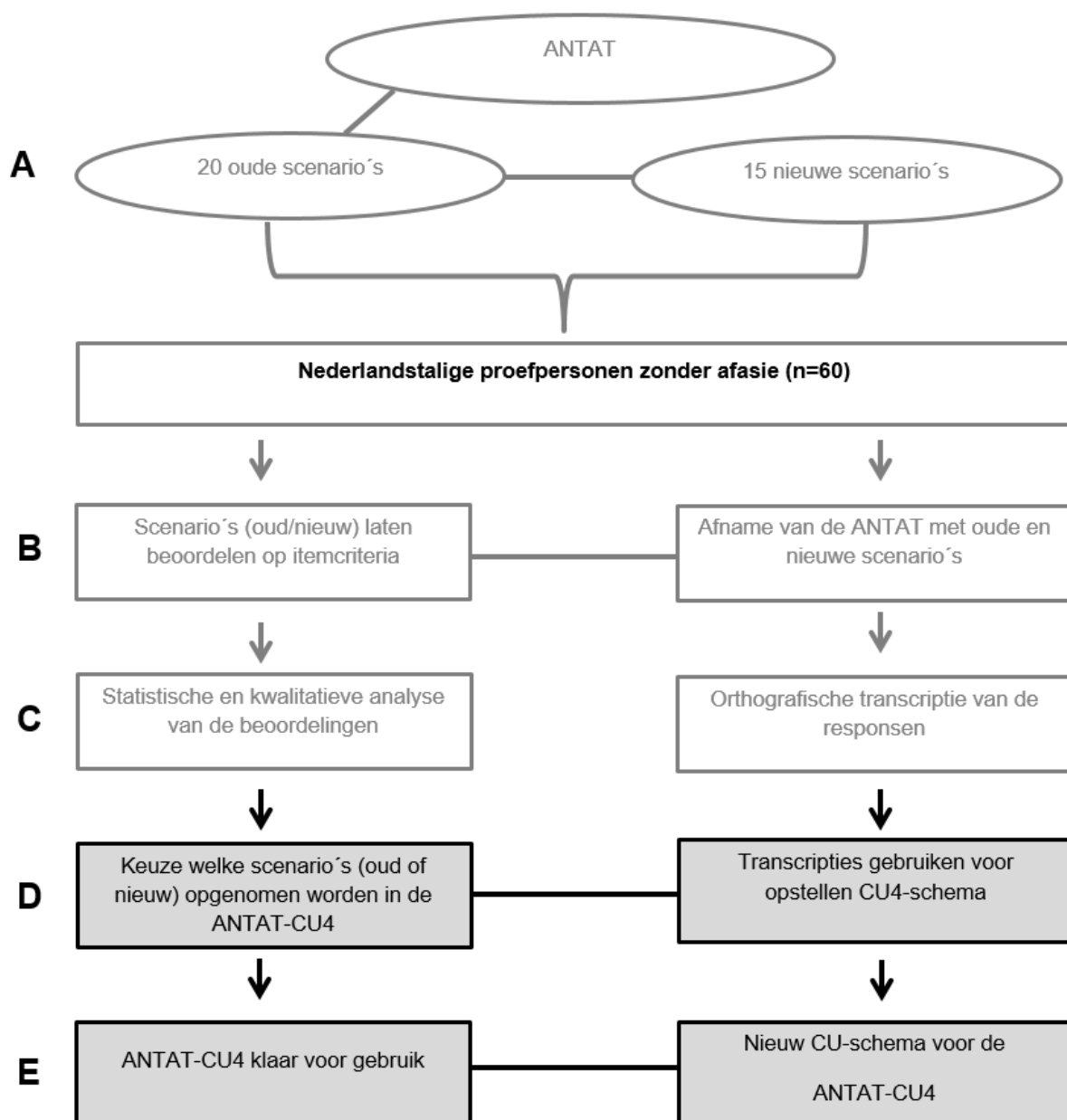
3.1. Het meetinstrument ANTAT-CU4 en de Content Unit analyse

In deze masterscriptie is gebruik gemaakt van de data van een groep proefpersonen zonder afasie ($n=60$) uit het onderzoek van Giessen (2015) en Aan de Stegge (2015). Onderdeel A van figuur 3 laat het vooronderzoek van deze masterscriptie zien (uitgevoerd door Giessen, 2015 en Aan de Stegge, 2015). Dit vooronderzoek was vanuit twee redenen belangrijk voor deze masterscriptie. Ten eerste voor de keuze van de scenario's voor de ANTAT-CU4 en ten tweede voor de Content Unit analyse die het mogelijk gemaakt heeft een CU-schema voor de ANTAT-CU4 te ontwikkelen.

Het onderzoek van Giessen (2015) en Aan de Stegge bestond uit twee delen (zie figuur 3 linker en rechter kant). Om de ANTAT-CU4 te kunnen ontwikkelen, was het noodzakelijk te onderzoeken welke scenario's nog aan de door Blomert et al. (1995) opgestelde itemcriteria voldeden en welke scenario's door een nieuw scenario vervangen moesten worden.

Figuur 3 onderdeel A laat zien, dat Giessen (2015) en Aan de Stegge (2015) de 20 oude scenario's van de ANTAT aangevuld hebben met 15 nieuwe scenario's. Zowel de oude als ook de nieuwe scenario's zijn vervolgens door de groep proefpersonen zonder afasie ($n=60$) beoordeeld op de itemcriteria van Blomert et al. (1995), namelijk rolverdeling, inleefbaarheid en actualiteit (zie figuur 3, onderdeel B (links)). Uit de kwalitatieve en kwantitatieve analyse (figuur 3, onderdeel C (links)) kwam naar voren, dat 6 van de 15 nieuwe scenario's (namelijk scenario 4, 5, 16, 17, 19 en 20) op inleefbaarheid en actualiteit als significant beter beoordeeld werden dan de originele scenario's (Giessen, 2015; Aan de Stegge, 2015). Er zijn geen significante verschillen gevonden tussen de originele en de nieuwe scenario's wat betreft de beoordeling op rolverdeling.

(Giessen, 2015 en Aan de Stegge, 2015)



Figuur 3: Overzicht proefpersonen zonder afasie ($n=60$). Het lichtgrijze gedeelte (A t/m C) laat het vooronderzoek van Giessen (2015) en Aan de Stegge (2015) zien. Het donkergrijze gedeelte (D t/m E) laat de stappen zien, die binnen deze masterscriptie zijn uitgevoerd.

Het tweede gedeelte van het onderzoek van Giessen (2015) en Aan de Stegge (2015) richtte zich op het verzamelen van data van proefpersonen zonder afasie. In figuur 3 onderdeel B (rechts) is te zien, dat zij zowel de 20 oude scenario's als ook de 15 nieuwe scenario's bij 60 proefpersonen zonder afasie hebben afgenomen. Alle responsen zijn op audio opgenomen en vervolgens middels het *Protocol voor Orthografische Transcriptie* (Corpus gesproken Nederlands, 2000) orthografisch getranscribeerd (figuur 3, onderdeel C (rechts)). Deze responsen zijn in deze masterscriptie belangrijk om voor de ANTAT-CU4 een nieuw CU-schema op te kunnen stellen.

3.1.1. De ANTAT-CU4

De onderdelen D en E in figuur 3 (links) laten zien, hoe met de data uit de kwalitatieve en kwantitatieve analyse van Giessen (2015) en Aan de Stegge (2015) binnen deze masterscriptie verder gewerkt werd. In een overleg samen met Dr. M. Ruiten en E. Lotgering (MA) werd besproken welke scenario's uiteindelijk in de ANTAT-CU4 opgenomen werden (figuur 3, onderdeel D, links). Hierbij zijn de aanbevelingen van Giessen (2015) en Aan de Stegge (2015) in de beslissing meegenomen en werd gekeken welke scenario's als significant beter beoordeeld werden. Bijlage A laat de gekozen scenario's zien en het bijhorende significantieniveau.

Er zijn uiteindelijk niet 20, maar 23 scenario's in de ANTAT-CU4 opgenomen. Dit werd veroorzaakt door het feit, dat de scenario's 11 (bril), 16 (gips) en 20 (gemeente) niet altijd goed begrepen werden door de proefpersonen zonder afasie ($n=60$). Daardoor zijn responsies gegeven die inhoudelijk niet passend waren bij de scenario's. Daarom werd ervoor gekozen om van deze drie scenario's een tweede versie in de ANTAT-CU4 op te nemen, om te bekijken welke versie van het scenario beter begrepen zal worden door de proefpersonen. De verschillen tussen de twee versies van de scenario's worden nu verder toegelicht.

Het verschil tussen scenario 11.1. en 11.2. is, dat de gebruikte afbeeldingen (zie bijlage B, afbeelding 4 en 9) van het scenario veranderd zijn. De afbeelding van Giessen (2015) en Aan de Stegge (2015) werd door sommige proefpersonen niet duidelijk herkend, daarom is een tweede afbeelding toegevoegd die een kapotte bril (bijlage B, afbeelding 9) laat zien, in plaats van een verbogen pootje (Giessen, 2015; Aan de Stegge, 2015). De formulering van scenario 11 is hetzelfde gebleven. In scenario 16 is de formulering van het scenario aangepast, omdat er uit de kwalitatieve analyse van Giessen (2015) en Aan de Stegge (2015) bleek dat niet altijd duidelijk was wat van de proefpersoon gevraagd werd. De oude formulering (16.2.) zag er als volgt uit: "U komt een oude kennis tegen op straat. Uw arm zit in het gips. Hij kijkt u verbaast aan. Leg aan hem uit wat er met u gebeurd is." De nieuwe formulering (16.1.) is licht veranderd: "U komt een oude kennis tegen op straat. Uw arm zit in het gips. Hij vraagt wat er gebeurd is, wat zegt u?". De nieuwe formulering vraagt expliciet wat er gebeurd is. Tenslotte is de formulering van scenario 20 aangepast. De formulering voor scenario 20.2 van Giessen (2015) en Aan de Stegge (2015) is: "U belt naar de gemeente. U wilt iemand van de gemeente spreken maar blijkbaar klopt het nummer niet. Per ongeluk krijgt u mij aan de lijn. Wat zegt u?". Uit de kwalitatieve analyse van Giessen (2015) en Aan de Stegge (2015) kwam naar voren, dat de proefpersonen vaak denken dat ze al met de gemeente verbonden zijn en alleen nog goed doorverbonden moeten worden. Maar eigenlijk is nog geen verbinding met de gemeente tot stand gekomen. Daarom is er een tweede formulering (20.1.) opgenomen, namelijk "U wilt iemand van de gemeente spreken. U belt, maar blijkbaar klopt het nummer niet. Per ongeluk krijgt u mij aan de lijn. Wat zegt u?".

Binnen het huidige scriptieonderzoek werd middels een nieuwe groep Nederlandstalige proefpersonen zonder afasie ($n=30$) gekeken welk van de twee scenario's beter begrepen werd. Op deze manier kon een advies uitgesproken worden welk van de boven beschreven scenario's beter geschikt is voor de ANTAT-CU4 (zie hiervoor hoofdstuk 5). Een gedetailleerd overzicht van alle scenario's is in bijlage C te vinden.

3.1.2. Ontwikkeling van het CU4-schema

De onderdelen D en E in figuur 3 (rechts) laten zien dat de orthografische transcripties van de 60 proefpersonen zonder afasie verder geanalyseerd zijn om het CU-schema op te stellen. De responsen zijn per scenario en per proefpersoon eerst onderverdeeld in de elementen *aanleiding* en *verzoek*. Voorbeeld (3) van een respons op scenario 3 (“We zijn in een winkel en u bent op zoek naar een nieuwe televisie. Ik ben de verkoopster daar. Kan ik u ergens mee helpen mevrouw/meneer?”) zal dit verduidelijken:

(3) Ik zoek een grote en platte televisie. Kunt u mij een advies geven?”

A

V

De reactie van voorbeeld (3) bestaat uit verschillende elementen, namelijk een aanleiding (A) en een verzoek (V). Een aanleiding is de reden waarom de conversatie begint en een verzoek is een poging het communicatieve doel te bereiken. Deze elementen (A en V) worden proposities genoemd en zijn als volgt gedefinieerd: “Propositions reflect preverbal messages as they stand for the structure of the lexical concepts. Although propositions correspond to words, one and the same proposition can be put into words in various manners” (Ruiter et al., 2011, blz. 966).

De proposities zijn verder op hun onderliggende elementen geanalyseerd, namelijk de Content Units (CU’s). De Content Units worden door Yorkston en Beukelman (1980, p.30) als volgt gedefinieerd: “...a grouping of information that was always expressed as a unit by normal speakers”. Het volgende voorbeeld (4) van scenario 3 laat de analyse stap van propositie naar CU zien:

(4) A:	<u>ik zoek een grote en platte televisie</u>	CU’s: ik – televisie – zoek – specificatie
V:	<u>u mij een advies geven</u>	CU’s: u – advies – geven

Voorbeeld (4) laat A en V van de spreker zien. De responsen zijn geanalyseerd op de onderliggende informatie-eenheden. Een CU kan uit een inhoudswoord (zelfstandige naamwoorden, zelfstandige werkwoorden, bijvoeglijke naamwoorden en bijwoorden) of uit een korte frase bestaan. Functiewoorden (hulpwerkwoorden, koppelwerkwoorden en voorzetsels) tellen niet als zelfstandige CU’s. Voor voorzetsels geldt een bijzondere regel. Volgens Lotgering (2009) bestaat een verschil tussen betekenisrijke en betekenisarme voorzetsels. Een betekenisarm voorzetsel is in de frase “op de koffie komen” (Lotgering, 2009, p. 46) te vinden. Als het voorzetsel weggelaten zou worden, zal het nog steeds mogelijk zijn de inhoud van de communicatieve boodschap te begrijpen. Een betekenisrijk voorzetsel is bijvoorbeeld de plaatsbepaling 'op tafel'. Op het moment dat men hier het voorzetsel weg zou laten, mist een belangrijk deel van de informatie en de communicatieve boodschap zal niet meer begrepen worden.

Het was vooral belangrijk de focus te leggen op de relevante (essentiële) informatie die belangrijk is om de inhoud van de verbale boodschap te begrijpen. Om relevante van noodzakelijke informatie te kunnen scheiden, is gebruik gemaakt van drempelwaarden. Er zijn ten eerste drempelwaarden gebruikt om de essentiële CU’s te onderscheiden van de niet essentiële CU’s (zie tabel 2). Vervolgens zijn drempelwaarden gehanteerd om te bekijken of er per scenario überhaupt een aanleiding of verzoek essentieel was (zie bijlage D en E). Hier kon het gebeuren dat sommige scenario’s

geen aanleiding of verzoek nodig hadden en daarom niet opgenomen werden in het CU-schema (zie bijlage E, tabel 12, scenario 16, 19 en 20). Bij sommige scenario's waren door de proefpersonen zonder afasie meerdere verschillende aanleidingen en verzoeken geproduceerd. Middels drempelwaarden werd ook hier per aanleiding en verzoek gekeken welke van deze aanleidingen en verzoeken essentieel waren of niet (zie bijlage F). Deze stappen worden nu nog verder beschreven.

Tabel 2 laat een voorbeeld van 10 proefpersonen zonder afasie zien, om duidelijk te maken welke stappen zijn ondernomen. Zoals in tabel 2 te zien is, kunnen proefpersonen meerdere van elkaar afwijkende responsies geven. Daarom werd per aanleiding en verzoek geprobeerd een overkoepelde term te vinden waarin de verschillende responsen van de proefpersonen terug te vinden zijn. In het geval van scenario 3 zal de overkoepelde term voor de aanleiding als volgt eruitzien: ik – televisie – zoek – specificatie x/y (CU1 – ik; CU2 – televisie; CU3 – zoek; CU4 – specificatie x, y).

Per CU werd vervolgens gekeken hoe vaak dit door de proefpersonen zonder afasie geproduceerd is (drempelwaarde CU1 = 80%; CU2 = 100%; CU3 = 100% en CU4 = 90%). Dat betekent, dat 80%, 100% en 90% van de proefpersonen deze CU's gebruikt hebben. Deze CU's zijn essentieel om de communicatieve boodschap te begrijpen. De drempelwaarde voor CU's was vastgezet bij 30%, om een keuze tussen essentiële en niet essentiële CU's te kunnen maken. Dat betekende dat op het moment dat een CU door 30% van de proefpersonen zonder afasie geproduceerd werd, dit CU meetelde als essentiële informatie en opgenomen werd in het CU-schema.

Tabel 2: Voorbeeld drempelwaarden per CU.

Proefpersonen n=10	CU1 ik	CU2 televisie	CU3 zoek	CU4 specificatie xy
pp 1	ik	televisie	zoek	groot en plat
pp 2	ik	televisie	zoek	patte
pp 3	ik	televisie	zoek	grote
pp 4	we	televisie	willen	-
pp 5	-	televisie	zoek	groot
pp 6	ik	televisie	zoek	groot
pp 7	ik	televisie	wil	groot
pp 8	we	televisie	zoeken	platte
pp 9	-	televisie	zoek	groot en plat
pp 10	ik	televisie	zoek	groot
Drempelwaarde voor n=10	80%	100%	100%	90%

Bijlage D laat een overzicht zien van alle scenario's en de procentuele verdeling van de hoeveelheid aanleidingen, verzoeken en proposities. De hoeveelheden zijn geteld aan de hand van de responsies van de proefpersonen zonder afasie ($n=60$). Een voorbeeld uit bijlage D is in tabel 3 voor scenario 3 weergegeven. De procentuele verdeling is weergegeven voor $n=58$, omdat van de 60 proefpersonen zonder afasie niet alle proefpersonen een responsie hebben gegeven.

Per scenario moest gekeken worden of er überhaupt een aanleiding of een verzoek opgenomen moest worden in het CU-schema. In de linker kolom in tabel 3 is weergegeven dat 56 (96,55%) van de 58 proefpersonen zonder afasie een aanleiding en 58 (100%) proefpersonen zonder afasie een verzoek

geproduceerd hebben. Binnen een persoonlijk overleg met Dr. M. Ruiter zijn de procentuele verdelingen van alle aanleidingen (zie bijlage E, tabel 11) en verzoeken (zie bijlage E, tabel 12) bekeken en werd besloten de grens voor het wel of niet opnemen van een aanleiding en/of verzoek bij een drempelwaarde van 50% vast te zetten. Met de gekozen drempelwaarde van 50% vallen er niet te veel aanleidingen en verzoeken buiten het CU-schema en worden niet essentiële aanleidingen en verzoeken niet opgenomen het CU-schema.

Tabel 3: Voorbeeld uit bijlage D, scenario 2. Overzicht van de procentuele verdelingen van de aanleidingen, de verzoeken en de proposities (CU-labels) van de proefpersonen (pp) zonder afasie (n=58).

		CU-label	Aantal personen n=58	A1+A2	Drempelwaarde
Aanleiding 56 pp (96,55%)	A1	ik-jullie-al vaker-zeggen	26	17 (29,31%)	44,82%
	A2	hier-niet voetballen-reden X,Y	30		51,72%
Verzoek 58 pp (100%)	V1	kinderen-ergens anders/plek X-voetballen/spelen	47	-	81,03%
	V2	ik-politie/ouders-bellen	9	-	15,51%
	V3	we-oplossing-vinden	3	-	5,17%

Binnen het persoonlijk overleg met Dr. M. Ruiter zijn ook de procentuele verdelingen van de verschillende aanleidingen (zie bijlage F, tabel 13) en verzoeken (zie bijlage F, tabel 14) binnen een scenario bekeken. Dit was belangrijk, omdat er per scenario deels meerdere verschillende aanleidingen en verzoeken door de proefpersonen zonder afasie geproduceerd werden. Dit is in tabel 3 in de kolom "CU-label" terug te zien. Binnen het scenario zijn verschillende aanleidingen en verzoeken mogelijk, maar niet alle geproduceerde aanleidingen en verzoeken zijn essentieel. Daarom werd een drempelwaarde van 10% werd gekozen. De beslissing is gebaseerd op de tabellen 13 en 14 in bijlage F. Deze tabellen makten duidelijk, dat als er een drempelwaarde van 10% gehanteerd wordt, 85% van de geproduceerde aanleidingen en verzoeken in het CU-schema opgenomen konden worden. De overige 15% waren niet essentiële aanleidingen en verzoeken en werden niet opgenomen in het CU-schema. In de kolom "drempelwaarde" in tabel 3 is procentueel het percentage voor de verschillende aanleidingen (A1 en A2) en verzoeken (V1, V2 en V3) weergegeven. Van de 56 proefpersonen die een aanleiding geproduceerd hebben, hebben 44,82% van de proefpersonen zonder afasie voor aanleiding A1 gekozen en 51,72% voor aanleiding A2. Het viel op dat 29,31% van de proefpersonen zonder afasie ervoor gekozen heeft, om zowel aanleiding 1 als ook aanleiding 2 te produceren. Van de 58 proefpersonen die een verzoek geproduceerd hebben, hebben er 81,03% voor verzoek V1 gekozen, 15,51% voor verzoek V2 en maar 5,17% voor verzoek V3.

Een drempelwaarde van 10% betekent voor tabel 3, dat A1, A2, V1 en V2 opgenomen werd in het CU-schema, maar V3 niet opgenomen werd. Daarom werd dit verzoek in tabel 3 roodgemaakt. Het verzoek V3 werd maar door 5,17% van de 58 proefpersonen zonder afasie geproduceerd. Deze waarde heeft de gehanteerde drempelwaarde van 10% niet behaald en hoort daarom niet bij de essentiële informatie.

Concluderend kan worden gezegd, dat in deze masterscriptie drie verschillende drempelwaardes gehanteerd zijn. Voor het vaststellen van essentiële en niet essentiële CU's werd een drempelwaarde van 30% gehanteerd. Voor het wel of niet opnemen van een aanleiding en/of een verzoek per scenario, werd een drempelwaarde van 50% gehanteerd en voor het scheiden van essentiële en niet essentiële aanleidingen en verzoeken binnen een scenario werd een drempelwaarde van 10% gehanteerd.

3.2. Test-hertest sensibiliteits- en betrouwbaarheidsonderzoek

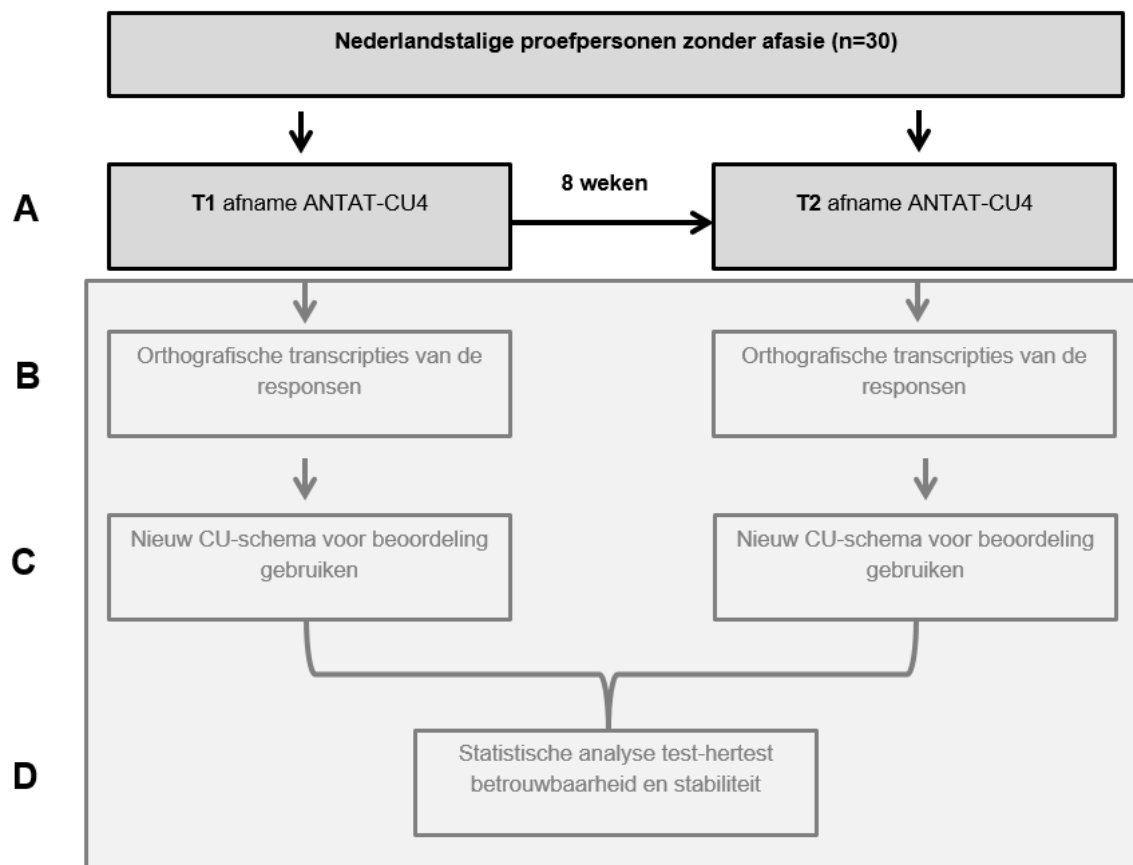
Om een eerste indruk te krijgen van de test-hertest betrouwbaarheid en stabiliteit is de ANTAT-CU4 bij een groep Nederlandstalige proefpersonen zonder afasie ($n=30$) (onderzoeksontwerp figuur 4) afgenomen. Bovendien is een onderzoeksontwerp opgenomen voor een groep Duitstalige proefpersonen met afasie ($n=10$) (onderzoeksontwerp figuur 5) en voor een groep Duitstalige proefpersonen zonder afasie ($n=30$) (zie hiervoor figuur 6). De verschillende proefpersoongroepen worden nu verder toegelicht en besproken.

3.2.1. Nederlandstalige proefpersonen zonder afasie

Om aan de test-hertest procedure deel te kunnen nemen moesten de proefpersonen zonder afasie aan bepaalde in- en exclusiecriteria voldoen. Dit waren: geen afasie, mannelijk/vrouwelijk vanaf het 18^{de} levensjaar tot het 90^{de} levensjaar, rechtshandig en het Nederlands als moedertaal. De proefpersonen mochten geen hersenbeschadiging hebben gehad of een aanwezigheid van mentale retardatie hebben. Het was bovendien belangrijk dat de proefpersonen geen (neuro) degeneratieve aandoeningen hadden. Ze mochten geen ernstige visusproblemen of gehoorproblemen (na correctie met een hulpmiddel) hebben en ook geen beperkingen in de spraakverstaanbaarheid.

In bijlage G (tabel 15, 16 en 17) is een overzicht van de steekproefverdeling van de proefpersonen zonder afasie (leeftijd, geslacht, hoogste voltooide opleiding en provincie) te vinden. Voor de test-hertest procedure met de ANTAT-CU4 konden 16 mannelijke en 14 vrouwelijke proefpersonen zonder afasie geworven worden die aan de in- en exclusiecriteria voldeden. De proefpersonen waren over de leeftijdscategorieën relatief gelijk verdeeld. De gemiddelde leeftijd van de proefpersonen zonder afasie lag bij 50 jaar (zie bijlage G, tabel 15). Tabel 16 in bijlage G laat een overzicht zien van het opleidingsniveau van de proefpersonen. Er zijn geen proefpersonen die “minder dan 6 klassen lager onderwijs” of “zes klassen lager onderwijs” hebben gevolgd. Merendeel van de proefpersonen heeft een VHMO-opleiding (30%) of een WO-opleiding (36%) gevolgd.

Voordat de ANTAT-CU4 afgenomen werd, zijn alle proefpersonen zonder afasie ($n=30$) over de procedure van de test-hertest mondeling en schriftelijk geïnformeerd en hebben alle proefpersonen een toestemmingsverklaring ondertekend (zie bijlage H).



Figuur 4: *Onderzoeksontwerp Nederlandstalige proefpersonen zonder afasie (n=30). Het donkergrijze gedeelte (onderdeel A) wordt binnen deze masterscriptie uitgevoerd en het lichtgrijze gedeelte (onderdeel A t/m D) wordt maar voor 6 proefpersonen binnen deze masterscriptie uitgevoerd. De onderdelen A t/m D moeten binnen een vervolgonderzoek voor alle proefpersonen zonder afasie (n=30) uitgevoerd worden.*

Figuur 4 laat de test-hertest procedure bij de Nederlandstalige proefpersonen zonder afasie ($n=30$) zien. Onderdeel A werd binnen deze masterscriptie uitgevoerd en laat zien, dat de ANTAT-CU4 per proefpersoon twee keer afgenomen werd, met een tussentijdse periode van 8 weken. Binnen de test-hertest procedure werd het eerste meetmoment T1 genoemd en het tweede meetmoment T2.

Bij elke proefpersoon werd de complete ANTAT-CU4 met alle 23 scenario's afgenomen. De 23 scenario's waren in twee versies verdeeld, waarbij versie 1 en versie 2 geen parallelle versies van elkaar waren. In bijlage C staat beschreven, hoe de scenario's over de twee versies verdeeld waren. Om te voorkomen dat de volgorde van de afname (eerst versie 1 dan versie 2) een invloed op de meetresultaten had, werd de groep proefpersonen zonder afasie ($n=30$) in tweeën gesplitst. Tabel 4 laat zien, dat de ene helft proefpersonen zonder afasie (groep 1) tijdens T1 eerst met versie 1 getest werd en vervolgens met versie 2. Tijdens T2 werd dit andersom gedaan, namelijk eerst versie 2 dan versie 1. De andere helft proefpersonen zonder afasie (groep 2) werd juist andersom getest, namelijk tijdens T1 eerst versie 2 dan versie 1 en tijdens T2 eerst versie 1 dan direct versie 2.

Tabel 4: Aangeboden testvolgorde voor de proefpersonen zonder afasie $n=30$.

$n=30$	T1	T2
Groep 1: 50% $n=15$	versie 1 – versie 2	versie 2 – versie 1
Groep 2: 50% $n=15$	versie 2 – versie 1	versie 1 – versie 2

De onderdelen B tot en met D in figuur 4 zijn licht grijs ingekleurd. Dit komt, omdat binnen deze masterscriptie niet alle 30 proefpersonen zonder afasie geanalyseerd konden worden. Uit iedere leeftijdscategorie werd 1 proefpersoon gekozen, behalve uit de leeftijdscategorie 70-80 omdat er geen proefpersonen binnen deze leeftijdscategorie hadden deelgenomen. Op deze manier konden 6 proefpersonen zonder afasie voor de analyse gekozen worden. De keuze voor deze 6 proefpersonen vond plaats door randomisatie. De responsen van deze 6 proefpersonen zonder afasie zijn vervolgens orthografisch getranscribeerd middels het *Protocol voor Orthografische Transcriptie* (Corpus gesproken Nederlands, 2000). Dat betekent, dat er voor 6 van de 30 proefpersonen zonder afasie de stappen B tot en met D binnen deze masterscriptie uitgevoerd zijn.

Het nieuwe CU-schema van de ANTAT-CU4 werd voor de beoordeling van de responsen van de proefpersonen zonder afasie ($n=6$) gebruikt. De transcripties zijn met behulp van het nieuwe CU-schema beoordeeld door de geproduceerde CU's per proefpersoon te tellen. Voor elk geproduceerde CU kreeg de proefpersoon 1 punt. Op deze manier kon door de proefpersonen zonder afasie een maximaal score van 118-120 punten bereikt worden (zie bijlage I).

De scores die uit het CU-schema kwamen zijn vervolgens voor de statistische analyse gebruikt (figuur 4, D) en met behulp van het programma SPSS 23 (Statistical Package of Social Sciences 23; IBM, 2015) statistisch onderzocht. Om een eerste indruk van de test-hertest stabiliteit van de ANTAT-CU4 te verkrijgen zijn de scores van de 6 proefpersonen zonder afasie met een gepaarde t-toets (tweezijdig getoetst) geanalyseerd. Middels de gepaarde t-toets kon bepaald worden of er bij de 6 proefpersonen zonder afasie sprake was van een significant verschil van de ANTAT-CU4 scores tussen meetmoment 1 (T1) en meetmoment 2 (T2). Voor de gepaarde t-toets werd een significantieniveau van 5% gehanteerd.

De stappen B tot en met D van figuur 4 (lichtgrijs gedeelte) moeten voor de complete Nederlandstalige proefpersoongroep zonder afasie ($n=30$) binnen een vervolgonderzoek nog een keer compleet uitgevoerd worden. Doordat alle testafnames op audio opgenomen zijn, kunnen de responsen van de proefpersonen binnen een vervolgonderzoek orthografisch getranscribeerd worden (figuur 4, B) middels het *Protocol voor Orthografische Transcriptie* (Corpus gesproken Nederlands, 2000). Op deze manier is het mogelijk de data op een later moment voor onderzoek te gebruiken.

3.2.2. Duitstalige proefpersonen met afasie

De Duitstalige proefpersonen met afasie moesten ook aan bepaalde in- en exclusiecriteria voldoen. Gezocht waren proefpersonen met een chronische afasie (mannelijk/vrouwelijk vanaf het 18^{de} levensjaar tot het 90^{de} levensjaar) handvoorkeur rechts, CVA van meer dan 1 jaar post onset en Duits als moedertaal.

De afasie van de proefpersonen moest met de Aachener Aphasie Test (AAT Duitse versie; Huber, Poeck, Weniger, & Willmes, 1983) vastgesteld zijn en mocht enkel ten gevolge van een cerebrovasculair accident (CVA) zijn. Ze mochten geen recidive hersenbeschadigingen hebben gehad of een aanwezigheid van mentale retardatie. Het was bovendien belangrijk dat de proefpersonen geen (neuro) degeneratieve aandoeningen hadden.

Het was belangrijk, dat de proefpersonen met afasie geen ernstige apraxie of dysartrie hadden, omdat dit de beoordeling van de spraakverstaanbaarheid negatief beïnvloedt (Liepold, Ziegler, & Brendel, 2003). Ruiters et al. (2011) hadden hiervoor vastgelegd dat de proefpersonen een verstaanbaarheid van minstens 95% moesten hebben. Binnen dit scriptieonderzoek werd ook een verstaanbaarheid van minstens 95% gehanteerd.

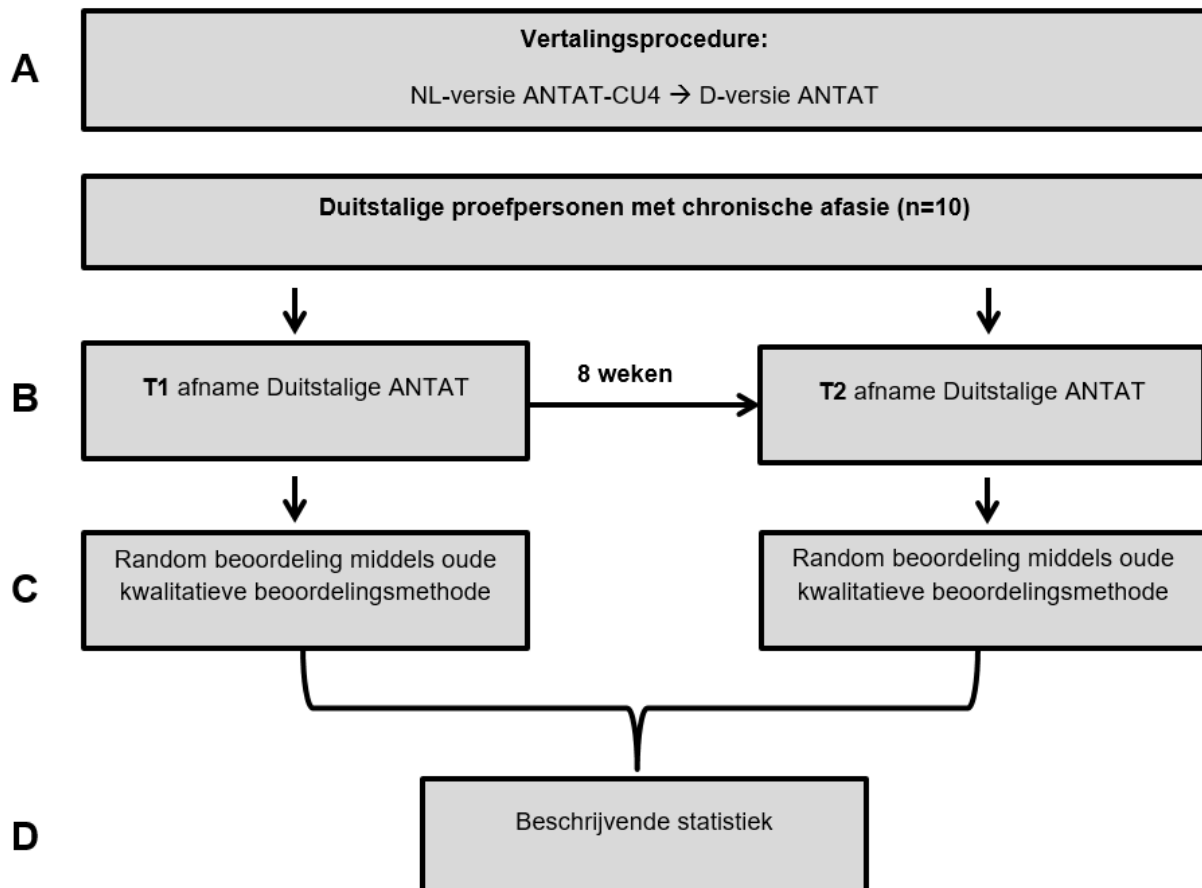
Het was bovendien van belang dat het taalbegrip op functieniveau niet ernstig gestoord was, omdat de testinstructies en de scenario's door de proefpersonen met afasie begrepen moesten worden om de Duitstalige ANTAT met nieuwe scenario's af te kunnen nemen. Om dit te controleren werd het onderdeel "taalbegrip" van de Aachener Aphasie Test (Huber et al., 1983) gebruikt. Hierbij mocht het onderdeel "taalbegrip" niet slechter gescoord zijn dan een stanine waarde (SN) van ≥ 4 (licht/matig gestoord; Huber et al., 1983).

De proefpersonen met afasie mochten geen ernstige concentratieproblemen, visusproblemen of gehoorproblemen (na correctie met een hulpmiddel) hebben. Proefpersonen met een neglect werden van het onderzoek uitgesloten, omdat bij een neglect problemen in de visuele waarneming bestaan die niet corrigeerbaar zijn (Elias & Saucier, 2006).

Het was wenselijk dat deze steekproef alle hoofdsyndromen van afasie (afasie van Broca, afasie van Wernicke, globale afasie en amnestische afasie (Dharmaperwira-Prins & Maas, 2006)) en overige afasiesyndromen (conductie-afasie, transcorticaal-motorische afasie, transcorticaal-sensorische afasie en gemengde transcorticale afasie (Dharmaperwira-Prins & Maas, 2006)) bevatte. Aan deze eis kon niet volledig voldaan worden.

Met hulp van een logopediepraktijk in Duitsland zijn 10 proefpersonen met chronische afasie gevonden die voldeden aan de in- en exclusiecriteria. In bijlage G (tabel 18 en 19) is een overzicht van de steekproefverdeling van de proefpersonen met afasie (leeftijd, geslacht, hoogste voltooide opleiding, tijd post onset, dysartrie/apraxie en het afasie syndroom) te vinden. In totaal hebben 6 mannelijke en 4 vrouwelijke proefpersonen met chronische afasie aan het onderzoek deelgenomen. De gemiddelde leeftijd van de proefpersonen lag bij 69 jaar (zie tabel 19). Alle proefpersonen kwamen uit NRW (Nordrhein Westfalen) en waren daar ook opgegroeid. Bij alle proefpersonen werd de AAT afgenomen, zodat bij alle proefpersonen een afasiesyndroom met behulp van het computerprogramma ALLOC toegekend kon worden (zie bijlage G, tabel 18). Van de 10 proefpersonen met afasie was de verdeling over de afasie syndromen als volgt: afasie van Broca (2), afasie van Wernicke (2), globale afasie (1),

anomische afasie (3), amnestische afasie (1) en restafasie (1). In totaal hadden 7 van de 10 proefpersonen met afasie een dysartrie en/of apraxie. Echter werd de spraakverstaanbaarheid bij deze proefpersonen niet negatief beïnvloed en had het dus geen effect op het onderzoek. Om de proefpersonen met afasie over het onderzoek en het doel van het onderzoek te informeren, werd met iedere proefpersoon een persoonlijke afspraak gemaakt om mondeling het onderzoek uit te leggen en de informatiebrief en de toestemmingsverklaring (zie bijlage J) te bespreken. Iedere proefpersoon met afasie ($n=10$) heeft de toestemmingsverklaring ondertekend.



Figuur 5: *Onderzoeksontwerp Duitstalige proefpersonen met afasie ($n=10$).*

Voordat de ANTAT bij de Duitstalige proefpersonen met afasie ($n=10$) afgenomen kon worden, moest de Nederlandstalige ANTAT-CU4 vertaald worden naar het Duits (figuur 5, A). Volgens de COTAN is voor het vertalen de volgende uitspraak belangrijk: “Wezenlijk daarbij is dat de vertaling niet letterlijk hoeft te zijn, maar dat het veel meer van belang is dat de teksten natuurlijk overkomen, omdat daardoor wordt bevorderd dat de betekenis van datgene wat wordt gevraagd, in beide talen hetzelfde is.” (Evers et al, 2010, p.8). Voor de vertalingsprocedure van de Nederlandstalige ANTAT-CU4 zijn de richtlijnen van de COTAN (Evers et al. 2010) en de richtlijnen van de “International Test Commission” (ITC, 2000) gebruikt.

Met behulp van equivalentieonderzoek moest het bewijs geleverd worden, dat met de Duitstalige ANTAT met nieuwe scenario's hetzelfde begrip gemeten wordt als met de Nederlandstalige ANTAT-CU4. De COTAN gaat ervan uit, dat als globaal dezelfde begrippen gemeten worden, men mag aannemen dat dezelfde theoretische uitgangspunten aan het instrument ten grondslag liggen (Evers et al., 2010). Het equivalentieonderzoek hield in, dat men de ANTAT-CU4 naar het Duits moest vertalen. Vervolgens moest men de Duitstalige ANTAT weer terugvertalen naar het Nederlands. De terugvertaalde Nederlandse versie van de ANTAT moest dan inhoudelijk vergeleken worden met de originele ANTAT-CU4. Op deze manier kon gecontroleerd worden of de vertaalde versie hetzelfde begrip meet als de originele Nederlandse ANTAT-CU4. Er stond in geen richtlijn beschreven welke kwalificatie de vertaler van de test moesten hebben.

Voor het equivalentieonderzoek in deze masterscriptie werd gewerkt met acht vertalers met Duits als moedertaal die de staatsexamens Nederlands als tweede taal (NT2, College voor Toetsen en Examens) acht jaar geleden gehaald hadden en regelmatig zowel Duits als Nederlands in het dagelijks leven gebruiken. De staatsexamens Nederlands als tweede taal zijn een indicatie dat de vertalers het Nederlands op niveau 3F beheersen. Alle acht vertalers zijn in Nederland afgestudeerd met een Bachelor logopedie. Bovendien hadden drie vertalers ook de Masteropleiding taal- en spraakpathologie afgerond. De Nederlandstalige ANTAT-CU4 werd door hen naar het Duits vertaald. Op deze manier waren acht Duitstalige ANTAT-versies met nieuwe scenario's ontstaan. De auteur van deze masterscriptie (ook Duits als moedertaal met NT2 diploma) heeft vervolgens van de acht vertaalde versies één Duitstalige ANTAT-versie met nieuwe scenario's ontworpen. Hierbij werd gekeken welke formulering in de vertalingen het meest gekozen werd door de acht vertalers. Deze formulering werd dan overgenomen in de vertaalde Duitse ANTAT met nieuwe scenario's. Deze vertaalde ANTAT met nieuwe scenario's werd door een tweetalig opgevoede Nederlander (Duits en Nederlands) weer terugvertaald naar het Nederlands. Dr. M. Ruiten heeft vervolgens de terugvertaalde ANTAT inhoudelijk vergeleken met de originele Nederlandstalige ANTAT-CU4. Op deze manier kon gekeken worden of de terug vertaalde Nederlandstalige ANTAT inhoudelijk hetzelfde was als de originele Nederlandstalige ANTAT-CU4. Beide versies werden als equivalent beoordeeld, waardoor de Duitstalige ANTAT met nieuwe scenario's in het onderzoek gebruikt mocht worden. Om de Duitstalige ANTAT met nieuwe scenario's af te kunnen nemen, zijn de afbeeldingen van scenario 13 naar het Duits vertaald. De aangepaste afbeeldingen zijn in bijlage K te vinden.

Figuur 5 onderdeel B laat de test-hertest procedure zien. De ANTAT-CU4 werd per proefpersoon twee keer afgenomen, met een tussentijdse periode van 8 weken (figuur 4, A). In totaal zijn op deze manier van de 10 proefpersonen met afasie, 20 audio opnames ontstaan. Voor de proefpersonen met afasie ($n=10$) werd de afname procedure op dezelfde manier gedaan, als bij de proefpersonen zonder afasie ($n=30$). Zoals in tabel 5 te zien, werd de groep proefpersonen met afasie ook in twee helften gesplitst, waarbij de ene helft tijdens T1 met de volgorde versie 1 - versie 2 getest werd en vervolgens tijdens T2 met versie 2 - versie 1. De andere helft kreeg de versies precies andersom aangeboden.

Tabel 5: *Aangeboden testvolgorde voor de proefpersonen met afasie n=10.*

n=10	T1	T2
Groep 1: 50% n=5	Versie 1 – versie 2	Versie 2 – versie 1
Groep 2: 50% n=5	Versie 2 – versie 1	Versie 1 – versie 2

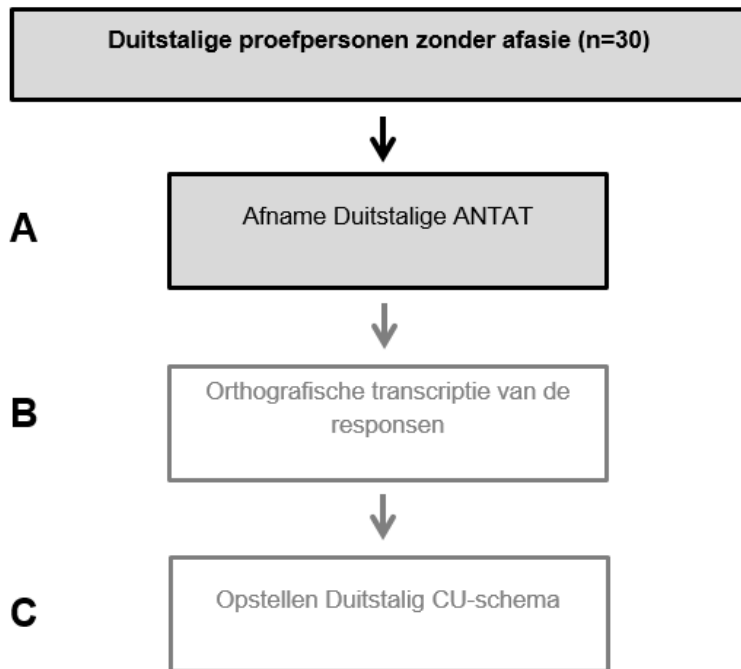
Alle testafnames zijn digitaal op audio opgenomen, zodat deze achteraf beluisterd en door de auteur van deze masterscriptie gescoord konden worden. De beoordeling van de responsen van de proefpersonen met afasie werd in een random volgorde gedaan om te voorkomen dat het achterelkaar beluisteren van T1 en T2 per proefpersoon de beoordeling subjectief zou kunnen beïnvloeden. Voor de beoordeling van de responsen werd de oude kwalitatieve beoordelingsmethode van de ANTAT van Blomert et al. (1995) gebruikt. De A-schaal voor begrijpelijkheid werd ter beoordeling van de responsen gehanteerd uit gebrek van een Duitstalig CU-schema. De B-schaal voor verstaanbaarheid werd binnen deze masterscriptie buiten beschouwing gelaten.

Voor de Duitstalige proefpersonen met afasie ($n=10$) werd alleen descriptieve statistiek gebruikt. Er werd geen gepaarde t-toets gebruikt, omdat iedere proefpersoon een ander afasie syndroom had met een andere mate van ernst. De resultaten van een gepaarde t-toets zijn een gemiddelde weerspiegeling voor de complete groep proefpersonen en daarom zou het gebruiken van een gepaarde t-toets geen inhoudelijke meerwaarde hebben. De informatie per proefpersoon op individueel niveau zou als het ware “weggemiddeld” worden en verloren gaan.

Met behulp van de beschrijvende statistiek kon op individueel niveau gekeken worden hoe de scores op de A-schaal voor begrijpelijkheid tussen de meetmomenten per proefpersoon met afasie zijn veranderd. Voor de Duitstalige proefpersonen met afasie zijn ook de responsietijden per proefpersoon middels de descriptieve statistiek beschreven. De resultaten hiervan zijn in hoofdstuk 4 weergegeven.

Uit de Duitstalige steekproef proefpersonen met afasie ($n=10$) zijn uiteindelijk maar acht proefpersonen met afasie voor de descriptieve statistiek gebruikt. De data van proefpersoon 7 kon niet gebruikt worden, omdat de audio opname tijdens meetmoment 1 niet te beluisteren was. Proefpersoon 9 kon niet voor de beschrijvende statistiek gebruikt worden, omdat het taalbegrip te ernstig gestoord was om de ANTAT-CU4 af te nemen. Daarom wordt in de volgende hoofdstukken vaak van $n=8$ gesproken in plaats van $n=10$. Dit wordt in hoofdstuk 5 nader besproken.

3.3. Duitstalig CU-schema



Figuur 6: Onderzoeksontwerp Duitstalige proefpersonen zonder afasie $n=30$. Het donkergrijze gedeelte (tot en met onderdeel A) werd binnen deze masterscriptie uitgevoerd. Het lichtgrijze gedeelte (onderdeel B tot en met C) moet binnen een vervolgonderzoek nader onderzocht worden.

Voor de Duitstalige ANTAT met nieuwe scenario's bestaat voor de beoordeling van de scenario's nog geen CU-schema. Om een Duitstalig CU-schema op te kunnen stellen was het noodzakelijk een Duitstalige normgroep van proefpersonen zonder afasie te werven. Het werven van de proefpersonen is binnen deze masterscriptie gedaan. Er zijn in totaal 30 Duitstalige proefpersonen zonder afasie geworven. De proefpersonen moesten aan dezelfde in- en exclusiecriteria voldoen, als de Nederlandstalige proefpersonen zonder afasie (zie bijlage G).

In totaal zijn 12 mannelijke en 18 vrouwelijke Duitstalige proefpersonen zonder afasie geworven met een gemiddelde leeftijd van 41,7 jaar. Het opleidingsniveau lag relatief hoog, namelijk 40% WO, 26,6% VHMO, 30% MULO en maar 3,3% met een gediplomeerde schoolopleiding lager dan mulo.

De Duitstalige ANTAT is bij alle 30 proefpersonen zonder afasie een enkele keer afgenomen en op audio opgenomen. Op deze manier is het voor een volgende onderzoeksgroep mogelijk de audiodata van alle proefpersonen zonder afasie ($n=30$) middels het *Protocol voor Orthografische Transcriptie* (Corpus gesproken Nederlands, 2000) orthografisch te transcriberen (figuur 6, B). Met de transcripties zal het dan mogelijk zijn een Duitstalig CU-schema voor de Duitstalige ANTAT met nieuwe scenario's op te stellen.

3.4. Advies voor twee parallele testversies

Het was helaas niet mogelijk het Nederlandstalige CU-schema voor de ANTAT-CU4 op te stellen voordat de ANTAT-CU4 bij de proefpersonen zonder afasie ($n=30$) werd afgenomen. Dit maakte het onmogelijk het advies voor de twee ANTAT-CU4 testversies al voor de testafname binnen deze masterscriptie te gebruiken.

Nadat het CU-schema voor de ANTAT-CU4 werd opgesteld kon gekeken worden hoe veel CU's per scenario aanwezig waren en welke communicatieve taak per scenario van de proefpersoon werd gevraagd (bijvoorbeeld een advies vragen, een afspraak maken of afzeggen etc.). Er waren in totaal drie inhoudelijke criteria die belangrijk waren voor het advies voor het opstellen van twee testversies:

1. Evenveel CU's per testversie
2. Dezelfde communicatieve taken
3. Evenveel scenario's met afbeeldingen

De twee testversies moesten inhoudelijk op elkaar lijken, zodat er binnen een vervolgonderzoek middels paralleliteitonderzoek vastgesteld kan worden of met beide testversies dezelfde resultaten verkregen kunnen worden.

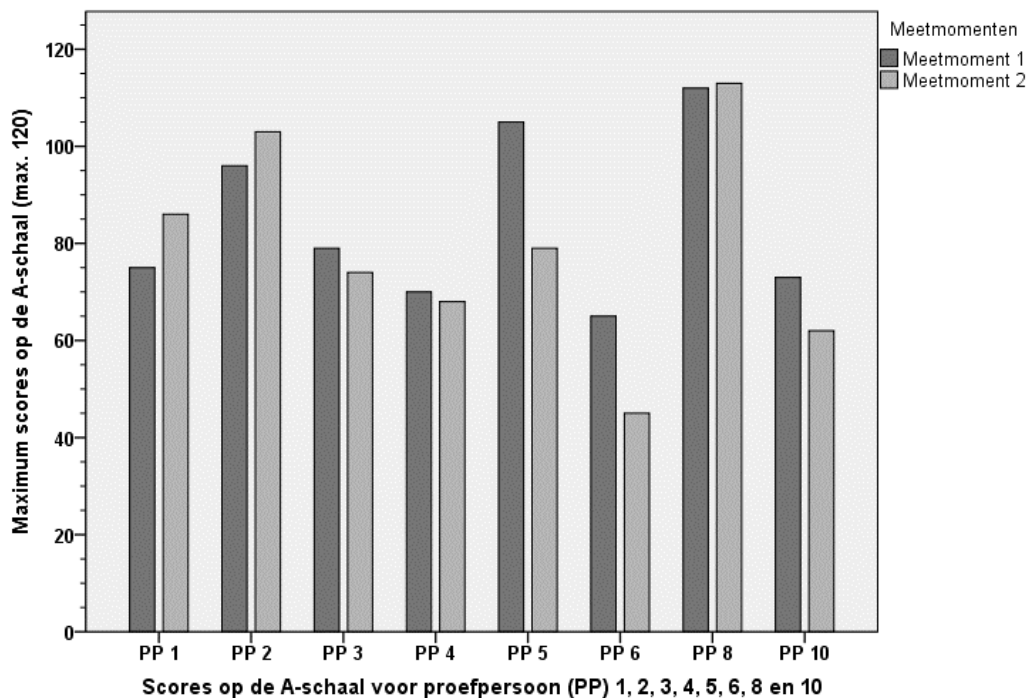
4. Resultaten

Binnen dit hoofdstuk zijn de resultaten van deze masterscriptie beschreven. Er wordt eerst ingegaan op de descriptieve statistiek, waar de testcores van de Duitstalige proefpersonen met afasie ($n=8$) besproken worden. Bovendien worden de responsietijden van de proefpersonen met afasie ($n=8$) beschreven. Vervolgens wordt ingegaan op de inductieve statistiek, waar de testresultaten van de Nederlandstalige proefpersonen zonder afasie ($n=6$) beschreven worden. Tot slot worden binnen dit hoofdstuk de producten gepresenteerd die binnen deze masterscriptie zijn ontstaan.

4.1. Descriptieve statistiek – Veranderingen in de functionele communicatie

A-schaal scores voor begrijpelijkheid

De verkregen data laten de ANTAT-scores op de A-schaal voor begrijpelijkheid van 8 proefpersonen met afasie op meetmoment 1 en op meetmoment 2 zien. Tabel 6 geeft een globaal overzicht over de ANTAT-scores van de proefpersonen met afasie ($n=8$). Er is weergegeven hoeveel punten op de A-schaal (voor begrijpelijkheid) gescoord zijn voor meetmoment 1 en meetmoment 2. In bijlage L, tabel 24 en 25 zijn de scores per scenario en per proefpersoon met afasie ($n=8$) terug te vinden.



Figuur 7: Maximum ANTAT-scores op de A-schaal (voor begrijpelijkheid) voor de proefpersonen met afasie ($n=8$) op meetmoment 1 en meetmoment 2.

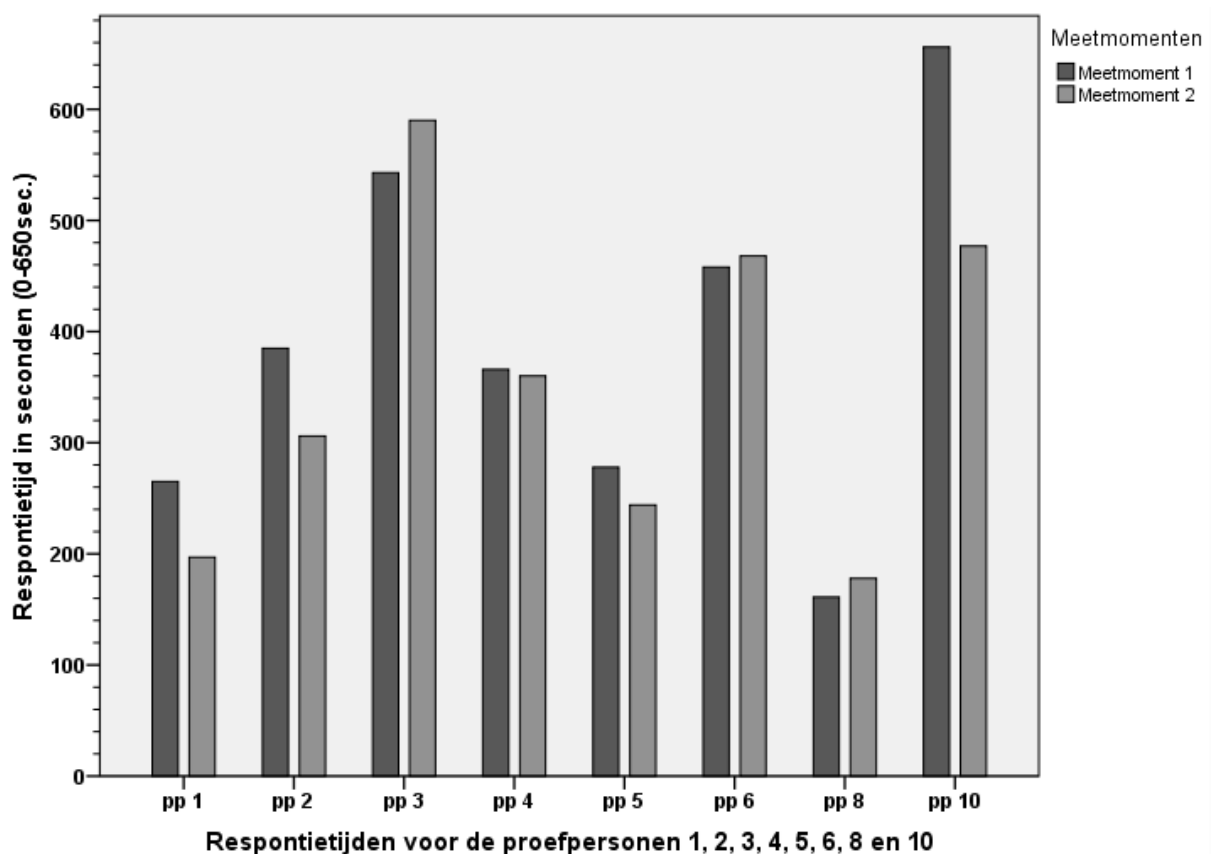
Figuur 7 laat zien hoe de ANTAT-scores op de A-schaal voor de proefpersonen met afasie ($n=8$) tussen meetmoment 1 en meetmoment 2 zijn veranderd. Zowel in tabel 6 als ook in figuur 7 is duidelijk te zien dat de ANTAT-scores op de A-schaal maar bij 3 proefpersonen met afasie (pp 1, 2 en 8) zijn toegenomen. Bij 5 proefpersonen met afasie (pp 3, 4, 5, 6 en 10) zijn de ANTAT-scores tijdens meetmoment 2 lager dan tijdens meetmoment 1.

Tabel 6: Compleet overzicht van de ANATAT-scores op de A-schaal (voor begrijpelijkheid) voor de proefpersonen met afasie ($n=8$) op meetmoment 1 en 2.

Proefpersonen ($n=8$)	T1	T2
	A-schaal max. score = 115	A-schaal max. score = 115
PP 1	75	86
PP 2	96	103
PP 3	79	74
PP 4	70	68
PP 5	105	79
PP 6	65	45
PP 8	112	113
PP 10	73	62

Responsietijden voor de proefpersonen met afasie

Per proefpersoon zijn voor ieder scenario de responsietijden beschreven. Per scenario werd de responsietijd in seconden (sec.) genoteerd. De tijden zijn voor alle proefpersonen met afasie ($n=8$) per scenario terug te vinden in bijlage L, tabel 26.



Figuur 8: Totale responsietijd (in seconden) voor de proefpersonen met afasie ($n=8$) per meetmoment.

Figuur 8 laat de totalen responsietijden van de ANATAT-afname van alle proefpersonen met afasie ($n=8$) tijdens meetmoment 1 en meetmoment 2 zien. Het viel op dat proefpersoon 1, 2, 4, 5 en 10 kortere reactietijden op meetmoment 2 hebben gehad, dan de proefpersonen 3, 6 en 8.

4.2. Inductieve statistiek – Test-herstest stabiliteit van de ANTAT-CU4

Een gepaarde t-toets werd gebruikt om te controleren of de ANTAT-CU4 scores op meetmoment 1 (T1) en op meetmoment 2 (T2) bij de proefpersonen zonder afasie ($n=6$) stabiel waren. De verwachting was geen significante veranderingen tussen de ANTAT-CU4 scores op meetmoment 1 en op meetmoment 2 te vinden.

Er was geen significant verschil tussen de ANTAT-CU4 scores op T1 ($M = 108.83$, $SD = 4.79$) en de ANTAT-CU4 scores op T2 ($M = 110$, $SD = 5.87$), $t(5) = -1.34$, $p > .05$. De correlatie tussen de ANTAT-CU4 score op meetmoment 1 en op meetmoment 2 was sterk, namelijk .939 bij een significantie van $p=.005$. Deze resultaten laten zien, dat de ANTAT-CU4 stabiel is, omdat de scores tussen de meetmomenten niet significant zijn veranderd.

4.3. Producten

De doelen die in het kader van deze masterscriptie in hoofdstuk 2.4 zijn opgesteld, hebben meerdere producten opgeleverd die nu beschreven worden.

Doel 1: Binnen deze masterscriptie wordt de ANTAT-CU4 opgesteld.

Product: De ANTAT-CU4 werd binnen deze masterscriptie opgesteld en is in bijlage C te vinden. De afbeeldingen voor de ANTAT-CU4 zijn in bijlage B weergegeven.

Doel 2: Binnen deze masterscriptie wordt een nieuw CU-schema voor de ANTAT-CU4 ontwikkeld.

Product: Het CU-schema voor de ANTAT-CU4 werd middels CU-analyse opgesteld en is in bijlage I te vinden. De maximale score op het CU-schema voor de ANTAT-CU4 ligt tussen 118 en 120 punten. Dit komt, omdat voor scenario 15 en 18 twee verzoeken mogelijk zijn waarbij of 5 CU's gescoord kunnen worden of 6 CU's gescoord kunnen worden. Het CU-schema voor de ANTAT-CU4 werd ook al binnen deze masterscriptie voor het scoren van responsies gebruikt.

Doel 3: Naar aanleiding van deze masterscriptie wordt een advies gegeven, hoe de scenario's op twee ANTAT-CU4 testversies onderverdeeld zouden kunnen worden.

Product: De scenario's van de ANTAT-CU4 waren nog niet onderverdeeld in twee parallele versies. Dit was nog niet mogelijk, omdat nog niet duidelijk was hoeveel CU's per scenario aanwezig waren. Het opstellen van het nieuwe CU-schema had als voordeel dat er per scenario naar het aantal CU's gekeken kon worden. In bijlage M, tabel 27 werd samengevat hoeveel CU's er per scenario zijn en in tabel 7 is vervolgens een voorstel gedaan hoe de scenario's op twee versies verdeeld zouden kunnen worden. Op deze manier werd gekeken, dat er evenveel CU's per testversie aanwezig waren. Bovendien werd gekeken, dat er per testversie evenveel scenario's werden opgenomen met dezelfde communicatieve taken en dat er evenveel scenario's met afbeeldingen per testversie zijn opgenomen. In tabel 7 is terug te zien, dat er per testversie 3 scenario's met afbeelding zijn opgenomen. Tabel 7 heeft bovendien geprobeerd aan de linker en rechter kant steeds twee scenario's tegenover te stellen die bijna dezelfde communicatieve taken van de proefpersoon vragen. Bij de scenario's 15 en 18 is het mogelijk 5 punten of 6 punten te halen. Dit heeft ermee te maken, dat er binnen deze scenario's twee verschillende verzoeken met 5 CU's of 6 CU's zijn opgenomen.

Tabel 7: Verdeling van alle scenario's op twee parallele versies voor de ANTAT-CU4.

Scenario's ANTAT-CU4 versie 1	Aantal CU's	Scenario's ANTAT-CU4 versie 2	Aantal CU's
Scenario 1: Stomerij (afbeelding)	6	Scenario 11: Bril (afbeelding)	5
Scenario 8: Hond	8	Scenario 2: Voetballen	7
Scenario 3: Televisie	7	Scenario 10: Bloemenzaak	8
Scenario 4: Schoenmaker (afbeelding)	6	Scenario 18: Kat	5/6
Scenario 5: Afspraak	5	Scenario 12: Verjaardag	6
Scenario 15: Dokter	5/6	Scenario 6: Slager (afbeelding)	7
Scenario 7: Vriend afspreken	7	Scenario 19: Etentje verjaardag	5
Scenario 9: Kennismaken	6	Scenario 17: Conducteur	8
Scenario 13: Betalingsherinnering (afbeelding)	6	Scenario 14: Wekkers (afbeelding)	7
Scenario 20: Gemeente	4	Scenario 16: Gips	2
Totaal:	60/61		60/61

Doel 4: In het kader van deze masterscriptie worden de data van Nederlandstalige proefpersonen zonder afasie ($n=30$) verzameld die in een opvolgend onderzoek gebruikt zullen worden om de test-hertest stabiliteit en betrouwbaarheid van de ANTAT-CU4 te bepalen.

Product: Een database is ontstaan van 30 Nederlandstalige proefpersonen zonder afasie die binnen een test-hertest procedure twee keer met de ANTAT-CU4 getest zijn. Op deze manier zijn 58 audio opnames voor vervolgonderzoek beschikbaar en kunnen in een vervolgonderzoek orthografisch getranscribeerd kunnen worden. Er konden maar 58 audio opnames opgenomen worden, omdat er bij twee proefpersonen de tweede audio opname ontbreekt. Middels het CU-schema voor de ANTAT-CU4 is het binnen een vervolgonderzoek mogelijk de responsies van de proefpersonen zonder afasie ($n=30$) te scoren en middels een gepaarde t-toets (tweezijdig getoetst) statistisch te analyseren. Op deze manier zal het binnen vervolgonderzoek mogelijk zijn een nog betere inschatting van de test-hertest stabiliteit en betrouwbaarheid te kunnen geven.

Doel 5: Binnen deze scriptie wordt middels equivalentieonderzoek de Nederlandstalige ANTAT-CU4 vertaald naar het Duits, zodat een Duitstalige ANTAT met nieuwe scenario's ontstaat.

Product: Een Duitstalige ANTAT met nieuwe scenario's werd ontwikkeld en is in bijlage N te vinden. De afbeeldingen voor deze testversie zijn in bijlage K te vinden.

Doel 6: In het kader van deze masterscriptie worden de data van Duitstalige proefpersonen met afasie ($n=10$) verzameld die in een opvolgend onderzoek gebruikt zullen worden om met behulp van een Duitstalig CU-schema de veranderingen van de functionele communicatie nader te onderzoeken.

Product: Een database met test-hertest audio opnames van 10 Duitstalige proefpersonen met afasie is ontstaan. De data van 8 proefpersonen zijn uiteindelijk bruikbaar voor vervolgonderzoek.

Doel 7: Binnen deze masterscriptie worden data van Duitstalige proefpersonen zonder afasie ($n=30$) verzameld die in een vervolgonderzoek gebruikt kunnen worden om een Duitstalig CU-schema te kunnen opstellen.

Product: In totaal zijn 30 Duitstalige proefpersonen zonder afasie één keer met de Duitstalige ANTAT getest. Gebaseerd op deze data is het mogelijk in een vervolgonderzoek middels CU-analyse een Duitstalig CU-schema voor de Duitstalige ANTAT met nieuwe scenario's op te stellen.

5. Discussie en conclusie

In dit hoofdstuk wordt kritisch ingegaan op de normen van de originele ANTAT (Blomert et al., 1994) en worden de hypothesen die binnen deze masterscriptie zijn opgesteld geëvalueerd. Vervolgens wordt ingegaan op de nieuwe ANTAT-CU4, het nieuwe CU-schema en de eerder opgestelde CU-schema's. Bovendien wordt duidelijk gemaakt waarom het meten van de responsietijden een onmisbaar onderdeel van het toekomstige CU-schema is en voor welke doelgroep de ANTAT-CU4 het beste geschikt is.

5.1. Normen van de originele ANTAT

In hoofdstuk 2.3. werd uitgebreid besproken welke onderdelen belangrijk zijn voor het normeren en standaardiseren van de ANTAT-CU4. Binnen dit onderdeel wordt vanuit het huidige standpunt nog eens naar de normen van de ANTAT (Blomert et al., 1995) gekeken.

De normen die in de ANTAT-handleiding uit 2008 te vinden zijn, zijn sinds 1994 niet meer aangepast of vernieuwd. Dat betekent dat de normen 21 jaar oud zijn. Het criterium *normen* is op dit moment (stand 2016) van de COTAN als *onvoldoende* beoordeeld, omdat de normen vanwege veroudering niet meer bruikbaar zijn (Egberink et al., 2010). Een hernormering is dus noodzakelijk. Bovendien is het onduidelijk uit welke proefpersonen de normgroep bestaat die door Blomert et al. (1995) ter normering gebruikt werd.

In de handleidingen van de ANTAT (Blomert et al., 1995; Blomert et al., 2008) staat beschreven, dat zij voor de normering van de ANTAT een samengevoegde steekproef uit Nederlandse ($n=268$) en Duitstalige ($n=205$) sprekers hebben gebruikt. Blomert et al. (2008) beschrijven in de handleiding: "een vergelijking van de Nederlandse en Duitse testversie toonde aan dat beide taalversies op geen enkel en relevant psychometrisch criterium verschilden" (p. 30). De *Adjusted Goodness of Fit Index* werd gebruikt, om te controleren of de A-schaal, de B-schaal, de gemiddelde somscore per schaal en per paralleltest en beide parallelversies verschillenden in de Nederlandstalige steekproef en de Duitstalige steekproef. Met een hoge betrouwbaarheid ($AGFI > .90$) kon geconcludeerd worden dat het samenvoegen van beide steekproeven geen probleem is.

Binnen deze masterscriptie werd begonnen met het verzamelen van data zowel bij Nederlandstalige als Duitstalige proefpersonen zonder afasie. De responsen van de Duitstalige en Nederlandstalige proefpersonen zonder afasie zijn soms inhoudelijk verwant, maar verschillen wel van elkaar. Ook al zijn het dezelfde scenario's die zijn afgenomen, toch zijn de reacties van Nederlandstalige proefpersonen en Duitstalige proefpersonen niet identiek. Het samenvoegen van een Nederlandstalige steekproef en een Duitstalige steekproef zal nu niet mogelijk zijn, omdat er in de toekomst voor de Nederlandstalige en de Duitstalige ANTAT twee verschillende CU-schema's gebruikt zullen worden. De CU-schema's zijn afhankelijk van de responsies van de Nederlandstalige of Duitstalige proefpersonen zonder afasie. Het samenvoegen van twee steekproeven uit twee verschillende landen tot een grote steekproef is dan voor normering niet meer mogelijk.

5.2. Hypothesen toetsing

Hypothese 1: De scores op de ANTAT-CU4 (met CU-schema) op meetmoment 1 en meetmoment 2 met een tussentijdse periode van 8 weken bij een Nederlandstalige groep proefpersonen zonder afasie ($n=6$) laten geen significante verschillen zien.

Een belangrijke assumptie voor de test-hertest procedure was dat de variabelen van de proefpersonen niet veranderen tijdens het testinterval (Field, 2009). Hieraan kon voldaan worden, omdat alle proefpersonen zonder afasie binnen de tussentijdse periode van 8 weken niet zijn veranderd. Een andere assumptie van de test-hertest procedure was dat de items van de ANTAT-CU4 voor iedere proefpersoon random aangeboden moesten worden (Brinkmann, 2011). Aan deze assumptie kon bij de ANTAT-CU4 niet helemaal voldaan worden, omdat men de items van de ANTAT-CU4 in twee versies wilde houden. Niet bij iedere proefpersoon zonder afasie werd een compleet nieuwe volgorde gekozen, maar de versies zijn in een andere volgorde afgenomen (zie hiervoor hoofdstuk 3.2.1., tabel 3). Concluderend kan worden gezegd, dat aan deze assumptie bij de ANTAT-CU4 niet helemaal voldaan kon worden.

Uit de statistische analyse kwam naar voren, dat de mean scores van meetmoment 1 en 2 maar licht van elkaar afweken (T1: $M=108,83$; T2: $M=110$). Dat was al een eerste indicatie voor een goede test-hertest stabiliteit. Dit werd door de goede correlatie bevestigd. De correlatie tussen de ANTAT-CU4 score op meetmoment 1 en op meetmoment 2 was sterk, namelijk $.939$ bij een significantie van $p=.005$. Volgens Weiß (2010) is de correlatie tussen test en hertest goed, als de correlatiecoëfficiënt zo dicht mogelijk aan 1 ligt en een slechtere correlatie is te zien als de correlatiecoëfficiënt dichterbij 0 ligt. De scores op meetmoment 1 en op meetmoment 2 waren (bijna) gelijk aan elkaar, waardoor een sterke correlatie te zien was, die heel dicht aan 1 lag. Omdat er een hoge correlatie tussen de twee meetmomenten bestaat, kan geconcludeerd worden, dat er ook een hoge maat aan betrouwbaarheid bestaat (Field, 2009).

De scores op de ANTAT-CU4 lieten geen significant verschil zien tussen meetmoment 1 en meetmoment 2. De test-hertest resultaten van de proefpersonen zonder chronische afasie ($n=6$) geven een eerste indicatie, dat de ANTAT-CU4 een goede test-hertest stabiliteit en betrouwbaarheid heeft.

Hypothese 2: De Duitstalige ANTAT met nieuwe scenario's en oude kwalitatieve beoordelingsmethode, laat veranderingen in de functionele communicatie bij Duitstalige mensen met chronische afasie ($n=8$) zien.

In de praktijk wordt op individueel niveau gemeten. De logopedist wil vooral weten of een afasiepatiënt door bijvoorbeeld de logopedische therapie op T2 vooruit is gegaan. Uit de kwalitatief verkregen resultaten van de proefpersonen met afasie ($n=8$) kwam naar voren dat maar bij 3 proefpersonen met afasie een toename op de A-schaal voor begrijpelijkheid gezien kon worden. Alle anderen proefpersonen hadden op meetmoment 2 een slechtere score op de A-schaal behaald dan tijdens meetmoment 1. Dit laat zien dat de scenario's van de Duitstalige ANTAT, veranderingen (positieve en negatieve) in de functionele communicatie zichtbaar kunnen maken.

In overleg met de logopedistes van de proefpersonen met afasie ($n=8$) kwam naar voren, dat de proefpersonen met afasie die tijdens meetmoment 2 een lagere score behaalden ook binnen de therapie meer stimulatie en begeleiding nodig hadden. Een aantal proefpersonen met afasie was binnen de tussentijdse periode van 8 weken, opgenomen in het ziekenhuis of was vanwege ziekte thuisgebleven. Hierdoor hebben de proefpersonen met afasie niet de gemiddelde therapiefrequentie (2-3 keer per week á 45 min.) ontvangen. Dit zou een mogelijke verklaring voor de lagere testcores kunnen zijn.

Een andere belangrijke factor is, is dat het niet mogelijk was de inhoud van de logopedische behandelingen binnen de logopedie praktijk te controleren. Deze factor kon binnen deze masterscriptie niet gecontroleerd worden en heeft mogelijk invloed op de scores op de Duitstalige ANTAT met nieuwe scenario's gehad. In een vervolgonderzoek zou deze factor mogelijk beter te controleren zijn door vaste richtlijnen voor de logopedische behandeling op te stellen.

Er moet vooral benadrukt worden, dat de resultaten van de proefpersonen met afasie ($n=8$) alleen op een kwalitatieve beoordeling berusten met behulp van de oude beoordelingsmethode van de ANTAT (Blomert et al., 2008). De resultaten geven nog geen weerspiegeling van de kwaliteit van de Duitstalige ANTAT met nieuwe scenario's. Hiervoor zou in een vervolgonderzoek gewerkt kunnen worden met een Duitstalig CU-schema. Het zal een meerwaarde hebben om de responsies van de proefpersonen met afasie ($n=8$), die binnen deze masterscriptie zijn verzameld, ook in een vervolgonderzoek met een Duitstalig CU-schema te beoordelen. Op deze manier zouden de oude beoordelingsmethode en de nieuwe beoordelingsmethode met elkaar vergeleken kunnen worden. Bovendien zou het interessant zijn, om meerdere beoordelaars beide beoordelingsmethodes te laten gebruiken. Op deze manier zou ook de intra- en interbeoordelaars betrouwbaarheid geanalyseerd kunnen worden.

5.3. De ANTAT-CU4

In hoofdstuk 3.1.1. werd beschreven dat de scenario's 11 (bril), 16 (gips) en 20 (gemeente) met twee versies zijn opgenomen. De responsen van de proefpersonen zijn nog niet nader onderzocht, maar tijdens het afnemen van de ANTAT-CU4 viel op dat scenario 11.2 (afbeelding 9 in bijlage B) beter begrepen werd door de proefpersonen zowel met als zonder afasie. Afbeelding 9 laat een bril zien, die duidelijk kapot is en gerepareerd zou moeten worden. Afbeelding 4 laat een bril zien met een verbogen pootje. Het verbogen pootje was voor sommige proefpersonen niet goed te herkennen of geen duidelijke reden om naar een opticien te gaan. Ik zal daarom adviseren om afbeelding 9 (kapotte bril) voor de afname van de ANTAT-CU4 te gebruiken.

Scenario 16.1. en 16.2. werden door de gezonde proefpersonen altijd even goed begrepen. De zin in scenario 16.2. "Hij kijkt u verbaast aan." zou eventueel bij mensen met afasie tot begripsproblemen kunnen leiden, omdat dit extra informatie is die door de persoon met afasie moet worden onthouden. Dit zou in een vervolgonderzoek met Nederlandstalige proefpersonen met afasie nog nader onderzocht kunnen worden.

Scenario 20 blijft een moeilijk te begrijpen scenario. Tijdens de ANTAT-CU4 afnames hebben meerdere proefpersonen de instructie zowel voor scenario 20.1. als 20.2. niet goed begrepen. De proefpersonen dachten bij beide versies van het scenario dat ze al met de gemeente verbonden waren en binnen de gemeente doorverbonden moesten worden. Voor dit scenario zou ik adviseren om een compleet nieuwe formulering te ontwikkelen. Een voorstel voor een nieuwe formulering zou kunnen zijn: "U wilt iemand van de gemeente spreken. U belt, maar u krijgt de kapper aan de lijn. Wat zegt u?". Als deze aangepaste versie gebruikt zal worden, zal duidelijk zijn dat ze met iemand verbonden zijn die niet bij de gemeente werkt en die ook niet door kan verbinden. Dit kan binnen een vervolgonderzoek nader onderzocht worden. Bovendien moeten voor de scenario's 11 (bril), 16 (gips) en 20 (gemeente) nog de CU's middels CU-analyse vastgesteld worden. Het huidige CU-schema bevat nog de scenario versies van Giessen (2015) en Aan de Stegge (2015). Voor het vaststellen van de CU's zou de data van de proefpersonen zonder afasie ($n=30$) gebruikt kunnen worden, omdat bij deze proefpersonen beide scenario versies zijn afgenomen.

In hoofdstuk 4 werd in tabel 7 het advies gepresenteerd voor twee parallelle testversies voor de ANTAT-CU4. Bij de verdeling van de scenario's over de twee testversies werd ervoor gezorgd dat beide versies evenveel CU's bevatten en dat beide versies inhoudelijk vergelijkbare scenario's bevatten. Dit houdt in dat bijvoorbeeld scenario 3 (uit versie 1) en scenario 10 (uit versie 2) scenario's zijn waarbij de proefpersoon in een winkel moet beschrijven wat hij/zij wil hebben. Bovendien zijn er nu in beide versies evenveel scenario's met een afbeelding (3 scenario's) en zonder een afbeelding (7 scenario's).

De verdeling van de scenario's over twee testversies is nog niet nader onderzocht op de testparalleliteit. In een vervolgonderzoek zal het noodzakelijk zijn, de testparalleliteit nader te onderzoeken. Hiervoor moeten beide versies binnen een test-hertest procedure bij een groep proefpersonen met afasie afgenomen worden. Eerst zal versie 1 afgenomen moeten worden en met een tussentijdse periode van 8 weken zou dan versie 2 afgenomen moeten worden. Vervolgens zou gekeken moeten worden, of beide versies dezelfde scores en resultaten per proefpersoon laten zien.

5.4. Het CU-schema

Met behulp van de Content Unit analyse (CU-analyse) is het CU-schema voor de ANTAT-CU4 ontwikkeld. Er moest een grens bepaald worden, om de te kunnen bekijken welke CU's, welke verzoeken en welke aanleidingen essentieel waren en welke niet.

Voor het vaststellen van essentiële en niet essentiële CU's werd een drempelwaarde van 30% gehanteerd. Voor het wel of niet opnemen van een aanleiding en/of een verzoek per scenario, werd een drempelwaarde van 50% gehanteerd en voor het scheiden van essentiële en niet essentiële aanleidingen en verzoeken binnen een scenario werd een drempelwaarde van 10% gehanteerd. Deze keuzes werden gemaakt op basis van de subjectiviteit van het onderzoek. In de literatuur waren geen richtlijnen te vinden die een advies gaven over het hanteren van drempelwaardes. Met behulp van de tabellen in de bijlage D, E en F werd gekeken welke invloed het verlagen of verhogen van de drempelwaarde op het aantal verzoeken en aanleidingen, het aantal proposities en CU's heeft. Aan de hand van de drempelwaarde werd gekeken welke drempelwaarde wel variatie in de antwoorden toelaat, maar tegelijkertijd niet essentiële informatie-eenheden uitfiltert.

Als gewerkt wordt met drempelwaarden bestaat altijd een dilemma, namelijk dat alles dat niet de drempelwaarde behaald als niet essentieel beschouwd wordt. Terwijl een uiting die maar door een persoon gebruikt wordt wel adequaat kan zijn, deze alleen niet vaak genoeg voorkomt binnen een groep sprekers zonder afasie (Lotgering, 2009). Deze uitingen worden nu door de gekozen drempelwaardes niet meegenomen in het CU-schema. Voor toekomstig onderzoek is het wenselijk het CU-schema te baseren op een nog grotere normgroep (bijvoorbeeld $n > 50$), omdat de kans dan groter is dat meer proefpersonen dezelfde proposities met onderliggende CU's produceren en daardoor de drempelwaarden eerder behaald kunnen worden van al vertegenwoordigde proposities in het CU-schema van de ANTAT-CU4. Tegelijk biedt een grotere steekproef de mogelijkheid om nieuwe proposities te kunnen destilleren en te karakteriseren (Filipinski, 2014).

Dat het aantal CU's tussen de verschillende CU-schema's (Ruiter et al, 2011; Lotgering 2009; Filipinski, 2014) veranderd is zal tabel 8 duidelijk maken. Eerder onderzoek (Lotgering, 2009; Ruiter et al., 2011; Filipinski, 2014) heeft voor de ANTAT met de originele scenario's ook CU-schema's ontwikkeld. In tabel 8 is een overzicht gegeven van het maximaal te behalende aantal CU's voor ANTAT-CU1, 2, 3 en 4. Er is te zien, dat het aantal CU's tussen 107 en 129 CU's varieert. Het maximaal te behalen aantal CU's van het CU-schema van de ANTAT-CU4 is 118-120 CU's. De maximale score varieert tussen 118 en 120, omdat twee scenario's (scenario 15 en scenario 18) opgenomen zijn waar men of 5 punten of 6 punten kan halen. Dit komt omdat twee mogelijke verzoeken geproduceerd kunnen worden, die of 5 CU's bevatten of 6 CU's bevatten.

Tabel 8: *Maximaal aantal CU's in de ANTAT-CU1, 2, 3 en 4.*

ANTAT	Maximaal aantal CU's
ANTAT-CU1	107
ANTAT-CU2	129
ANTAT-CU3	126
ANTAT-CU4	118-120

In de verschillende CU-schema's was het bovendien interessant te bekijken, welke verschillen in het aantal CU's per scenario te vinden waren. Hier konden alleen de scenario's 2, 8, 15, 16 en 18 bekeken worden, omdat alle andere scenario's in de ANTAT-CU4 vervangen werden door nieuwe scenario's. Tabel 9 laat zien, dat het aantal CU's tussen de verschillende CU-schema's verschilde. Afhankelijk op welke groep proefpersonen een CU-schema gebaseerd is, afhankelijk daarvan varieert ook het aantal CU's per scenario en per CU-schema.

Tabel 9: Aantal CU's van de oude scenario's in vergelijking tussen ANTAT-CU1, 2, 3 en 4.

Oude scenario's	Aantal CU's			
	ANTAT-CU1	ANTAT-CU2	ANTAT-CU3	ANTAT-CU4
Scenario 2: Voetballen	9	8	8	7
Scenario 8: Hond	6	10	8	6
Scenario 15: Dokter	5	6	6	5/6
Scenario 16: Gips	4	6	6	2
Scenario 18: Kat	8	8	8	5/6

In bijlage M, tabel 27 is een overzicht van het aantal CU's voor de nieuwe scenario's weergegeven. In een vervolgonderzoek zou gekeken kunnen worden of het aantal CU's per scenario verandert als de ANTAT-CU4 afgenomen wordt bij een nieuwe groep proefpersonen zonder afasie of dat er eventueel nieuwe proposities in het CU-schema opgenomen kunnen worden.

Het is wenselijk duidelijke instructies voor het gebruiken van het CU-schema op te stellen en deze instructies toe te voegen aan een handleiding voor de ANTAT-CU. Het CU-schema zou kunnen verleiden tot een scoren van "woorden" in plaats van CU's. Het construct "propositie" en "CU" moet voor de gebruiker van de ANTAT-CU4 duidelijk zijn. In de handleiding van Blomert et al. (2008) staat namelijk niet uitgelegd wat een propositie precies inhoudt, terwijl dit begrip regelmatig in de handleiding gebruikt wordt. Een voorstel voor een nieuwe handleiding zou kunnen zijn een lijst van voorbeelden van mogelijke uitingen op te nemen, zodat het voor de beoordelaar duidelijk is wat de bedoeling van het CU-schema is. Een uitgebreide handleiding zou eventueel ook een positief effect kunnen hebben op de overeenstemming tussen en binnen beoordelaars.

Over de kwaliteit van het huidige CU-schema van de ANTAT-CU4 kan nog geen oordeel gegeven worden, omdat het CU-schema maar bij 6 proefpersonen voor de beoordeling gebruikt werd. De test-hertest resultaten lieten zien, dat de test scores stabiel waren. Dit geeft een eerste indruk over de kwaliteit van het huidige CU-schema.

Binnen dit scriptieonderzoek is alle data geanalyseerd door één onderzoeker, de auteur van deze masterscriptie. Vanuit die reden kan er nog geen uitspraak gedaan worden over de intra- en interbeoordelaars overeenstemming en betrouwbaarheid van het CU-schema. Om dit te controleren zou het CU-schema voor de ANTAT-CU4 door meerdere beoordelaars gebruikt moeten worden en de intra- en interbeoordelaars betrouwbaarheid en overeenstemming nader geanalyseerd moeten worden. Hiervoor zou bijvoorbeeld binnen een vervolgonderzoek de data van de groep proefpersonen met afasie ($n=8$) gebruikt kunnen worden. Middels een Duitstalig CU-schema zouden de responsies door meerdere beoordelaars gescoord en vergeleken kunnen worden om de interbeoordelaars betrouwbaarheid en overeenstemming te analyseren. De scores van de proefpersonen met afasie ($n=8$) op het Duitstalige CU-schema zouden bovendien ook vergeleken kunnen worden met de scores die binnen deze

masterscriptie middels de A-schaal verkregen zijn. Op deze manier zou ook nog een vergelijking tussen de oude kwalitatieve beoordelingsmethode en de nieuwe kwantitatieve beoordelingsmethode mogelijk zijn.

Voor het Nederlandstalige CU-schema zou het ook raadzaam zijn om in toekomstig onderzoek responsies van Nederlandstalige proefpersonen met afasie te verzamelen om op die manier het Nederlandstalige CU-schema verder op de inter- beoordelaars betrouwbaarheid en overeenstemming te analyseren.

Een manco van het huidige CU-schema is het aantal proefpersonen waarop het CU-schema gebaseerd is. Het huidige CU-schema werd aan de hand van responsies van 30 proefpersonen zonder afasie opgesteld, terwijl het derde CU-schema (Filipinski, 2014) opgesteld werd met behulp van de responsies van 50 proefpersonen zonder afasie. Ten opzichte van het derde CU-schema (Filipinski, 2014) is het huidige CU-schema beter, omdat het gebaseerd is op de nieuwe scenario's. Het huidige CU-schema moet nog aangevuld worden met een manier die niet alleen de effectiviteit maar ook de efficiëntie kan meten. Waarom dit belangrijk is wordt in de hierop volgende sectie beschreven.

5.5. Responsietijden, effectiviteit en efficiëntie

De Duitstalige proefpersonen met afasie ($n=8$) die binnen deze masterscriptie onderzocht zijn, zijn allemaal gediagnosticeerd met een afasiesyndroom (zie bijlage G). De verschillende afasie syndromen zijn te herkennen aan de verschillende responsietijden van de proefpersonen met afasie. De responsietijd is per proefpersoon en per scenario opgenomen in bijlage L, tabel 26 en grafisch samengevat in de resultaten in figuur 8.

Het viel op dat 5 van de 8 proefpersonen met afasie kortere reactietijden op meetmoment 2 hebben gehad dan op meetmoment 1. Het zou kunnen, dat de responsies op meetmoment 2 efficiënter zijn overgedragen dan op meetmoment 1, omdat het minder tijd heeft gekost de communicatieve boodschap over te dragen. Om een kleine inschatting van het belang van de responsietijd te geven, zijn de responsietijden van proefpersoon 3, 8 en 10 vergeleken met de scores op de A-schaal.

Proefpersoon 3 had op T1 een responsietijd van 543 seconden en op T2 van 590 seconden. Op de A-schaal behaalde deze proefpersoon op T1 een score van 79 en op T2 van 74. Deze proefpersoon had tijdens de testafname veel woordvindingsproblemen, semantische en fonematische parafasieën. De proefpersoon merkte de fouten tijdens het spreken op en verbeterde de fouten door conduite d'approche. Hierdoor had de proefpersoon meer tijd nodig om een goede responsie te geven.

Een ander beeld was bij proefpersoon 8 te zien. Proefpersoon 8 had op T1 een responsietijd van 161 seconden en op T2 178 seconden. Op de A-schaal behaalde deze proefpersoon op T1 een score van 112 en op T2 een score van 113. Deze proefpersoon kwam heel dicht bij de maximale score die op de A-schaal mogelijk te behalen was (max. 115). In vergelijking met de anderen proefpersonen had proefpersoon 8 de beste responsietijden in verhouding met de scores op de A-schaal. Deze proefpersoon was zowel effectief als ook efficiënt in vergelijking met de andere proefpersonen.

Proefpersoon 10 had op T1 een responsietijd van 656 seconden en op T2 van 477 seconden. Op de A-schaal had deze proefpersoon op T1 een score van 73 en op T2 een score van 62. Deze proefpersoon had op meetmoment 2 een lagere responsietijd dan op meetmoment 1. Tijdens

meetmoment 1 en 2 liet deze proefpersoon vaak conduite d 'écart zien, waardoor de proefpersoon de doelwoorden niet bereikte en tijdens de testafname vaak vastliep. Bovendien waren in de audio opname van proefpersoon 10 korte stukken van logorroé terug te horen. Hierdoor werden de responsies voor proefpersoon 10 langer dan van de anderen proefpersonen.

De beschrijvingen van proefpersoon 3, 8 en 10 zijn alleen interpretaties en moeten nog nader onderzocht worden. In de toekomst zou het interessant zijn de CU's van de proefpersonen met afasie ($n=8$) nader te onderzoeken op de communicatieve effectiviteit en efficiëntie. Binnen deze masterscriptie werd niet op het meten van de communicatieve efficiëntie ingegaan. Om in de toekomst een precieze inschatting van de communicatieve efficiëntie en de effectiviteit te kunnen geven is het noodzakelijk de beoordeling met het CU-schema aan te vullen met een methode die in staat is de communicatieve efficiëntie bij mensen met afasie te meten. Zoals in hoofdstuk 2.1. besproken werd, kan alleen dan een betrouwbare inschatting van de communicatieve efficiëntie gegeven worden, als de informatie-eenheden per minuut (CU's/min) geteld worden.

Voor toekomstig onderzoek is het wellicht nog interessant de factor leeftijd nader te onderzoeken en mee te nemen in de interpretatie van de scores op het CU-schema en het meten van de communicatieve efficiëntie. Binnen dit scriptieonderzoek werd erna gestreefd om uit iedere leeftijdscategorie evenveel proefpersonen deel te laten nemen. Het viel op, dat naarmate de proefpersonen ouder waren ook de testafname met de ANTAT-CU4 langer duurde en minder relevante CU's geproduceerd werden. De responsietijden en het aantal geproduceerde CU's van de proefpersonen uit deze masterscriptie zijn nog niet nader onderzocht en blijven tot dat ze nader onderzocht zijn een subjectieve inschatting van de auteur van deze masterscriptie. Voor toekomstig onderzoek zou het interessant zijn om bij de Duitse en bij de Nederlandse proefpersonen per leeftijdscategorie te analyseren of een verschil tussen de verschillende leeftijdscategorieën bestaat wat betreft de responsietijd en het aantal geproduceerde CU's.

Voor de normering van de ANTAT-CU is het mogelijk de normgroep te baseren op de oudere leeftijdscategorieën, omdat hersenletsel, en daarmee afasie, vaker voorkomen op een oudere leeftijd (Schneider, Wehmeyer, & Grötzbach, 2012). Een normgroep die gemiddeld jonger is dan de gemiddelde groep mensen met afasie zou daarom ook niet representatief zijn voor de prestatie die deze laatstgenoemde groep zonder afasie zou leveren (Filipinski, 2014). Er moet voorkomen worden dat het CU-schema gebaseerd is op responsies van gemiddeld jongere proefpersonen, omdat men dan eventueel niet alleen het onderscheid wel/geen afasie onderzoekt, maar eerder het verschil in leeftijd toetst. Voor de ANTAT-CU is het eventueel een idee de scores op de ANTAT te corrigeren op leeftijd. Dit wordt bijvoorbeeld bij de Duitstalige Token Test van de AAT (Graetz et al., 1992) gedaan. Afhankelijk in welke leeftijdscategorie de testpersoon valt, worden de scores op de Token Test gecorrigeerd door punten af te trekken. Dit zou een mogelijke optie voor de ANTAT-CU kunnen zijn, om de factor leeftijd bij de testcores erbij te betrekken.

5.6. Doelgroep van de ANTAT-CU4

In hoofdstuk 2.3.1. werd beschreven, dat het voor de normering belangrijk is alle hoofdsyndromen van afasie bij de normering erbij te betrekken. Misschien is bij de normering van belang, de ANTAT-CU niet voor alle afasie hoofdsyndromen te normeren. In de handleiding van Blomert et al. (2008) staat beschreven, dat proefpersonen met een globale afasie gemiddeld slechter op de A-schaal scoorden dan proefpersonen met andere afasie syndromen. Proefpersonen met een amnestische afasie en rest-afasie scoorden beter op de A-schaal dan proefpersonen met andere afasie syndromen. Amnestische en rest-afasie patiënten verschilden significant van elkaar op de A-schaal. Proefpersonen met een Wernicke afasie bleken gemiddeld slechter te scoren op de A-schaal dan proefpersonen met een Broca afasie (Blomert et al., 2008).

In het huidige onderzoek kon de ANTAT-CU4 bij proefpersoon 9 met afasie niet afgenomen worden. Proefpersoon 9 is met behulp van de AAT gediagnosticeerd met een globale afasie (zie bijlage G, tabel 18). Het taalbegrip van proefpersoon 9 was te slecht om de instructies en de inhoud van de scenario's te kunnen begrijpen. Dit maakte dat de testafname afgebroken moest worden.

Een ander probleem tijdens de afname van de ANTAT-CU4 was perseveratie. Proefpersoon 6 is gediagnosticeerd met een amnestische afasie (zie bijlage G, tabel 18). Deze proefpersoon werd door de perseveratie gehinderd om op de scenario's te reageren. De woorden uit voorafgaande scenario's werden steeds herhaald, wat ertoe leidde dat de proefpersoon niet op nieuwe scenario's kon ingaan.

Een manco van de ANTAT-CU4 is, dat deze test wellicht meer geschikt voor patiënten met een Broca-afasie of een anomische afasie dan voor patiënten met een Wernicke afasie of een globale afasie. Om dit van tevoren te controleren, zou het afasiesyndroom met behulp van de AAT bepaald kunnen worden en zouden richtlijnen in de handleiding neergezet kunnen worden.

Het is voor de toekomst belangrijk in de handleiding van de ANTAT-CU te vermelden, dat de ANTAT-CU4 alleen bedoeld is voor mensen bij die het auditieve en leesinhoudelijk taalbegrip redelijk intact is. Om dit van tevoren te kunnen beoordelen zou het een idee kunnen zijn de stanine waardes van de AAT (Graetz et al., 1992) als indicatie te gebruiken. Een voorstel zou kunnen zijn, dat het onderdeel "taalbegrip" niet slechter gescoord mag worden, dan een stanine waarde (SN) van ≥ 4 (licht/matig gestoord). Dit is eventueel ook een exclusie criterium voor de normering van de ANTAT-CU.

5.7. Conclusie en aanbevelingen

De huidige masterscriptie heeft een aantal producten opgeleverd die in toekomstig onderzoek nog verder ontwikkeld moeten worden om de ANTAT-CU4 verder te verbeteren. Voor toekomstig onderzoek zijn binnen deze masterscriptie data van zowel Nederlandstalige proefpersonen zonder afasie ($n=30$), Duitstalige proefpersonen met afasie ($n=10$) en Duitstalige proefpersonen zonder afasie ($n=30$) verzameld. Middels deze databases is het mogelijk de test-hertest stabiliteit van de ANTAT-CU4 nader te onderzoeken, een Duitstalig CU-schema voor de Duitstalige ANTAT met nieuwe scenario's op te stellen en de intra- en interbeoordelaars betrouwbaarheid en overeenkomst nader te onderzoeken. Een eerste indruk van de test-hertest betrouwbaarheid van de ANTAT-CU4 werd binnen deze masterscriptie al gegeven. Deze resultaten laten zien, dat de ANTAT-CU4 voor 6 proefpersonen zonder afasie tussen de test en de hertest stabiel is gebleven. Het is bovendien mogelijk in toekomstig onderzoek met behulp van een paralleliteitonderzoek, twee parallelle testversies van de ANTAT-CU4 verder uit te werken en op hun paralleliteit te onderzoeken. Het is wenselijk de ANTAT-CU4 verder te ontwikkelen, zodat het mogelijk is deze test in de klinische praktijk te kunnen gebruiken.

Referenties

- Allen, M.J., & Yen, W.M. (1979). *Introduction to measurement theory*. California: Brooks/Col Publishing Company.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Blomert, L., & Buslach, D. C. (1994). *Amsterdam-Nijmegen Everyday Language Test (ANELT) – Deutsche Fassung*. Lisse, Nederlande: Swets & Zeitlinger.
- Blomert, M., Koster, C., van Mier, H., & Kean, M-L. (1987). Verbal communication abilities of aphasic patients: The everyday language test. *Aphasiology*, 1, (6), 463-474.
- Blomert, L. (1988). *Eindrapport*. Praeventiefonds project 28-1530, Amsterdam.
- Blomert, L. (1990). What functional assessment can contribute to setting goals for aphasia therapy. *Aphasiology*, 4, 307-320.
- Blomert, L. (1993). *Afasie Partner Vragenlijst (APV)*. Eindrapport project Stichting Afasie Nederland.
- Blomert, L., & Kean, M. (2008). *ANTAT, Amsterdam-Nijmegen Test voor Alledaagse Taalvaardigheden: Handleiding*. Hogrefe Uitgevers B.V.
- Blomert, L., Kean, M-L., Koster, Ch., & Schokker, J. (1994). Amsterdam-Nijmegen Everyday Language Test: construction, reliability and validity. *Aphasiology*, 8, 381-407.
- Boxum, E., van der Scheer, F., & Zwaga, M. (2013). *ASTA – Analyse voor Spontane Taal bij Afasie*. Vereniging voor Klinische Linguïstiek.
- Brinkmann, J. (2011). *Cijfers spreken, overtuiging met onderzoek en statistiek*. Noordhoff Uitgevers Groningen/Houten.
- Bühner, M. (2004). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson.
- Centraal Bureau voor de Statiek (2015). *Standaardonderwijsindeling (SOI)*. Opgeroepen op 21.08.2015 via: <http://www.cbs.nl/nl-NL/menu/methoden/classificaties/overzicht/soi/default.htm>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, Vol. XX, 1.
- College voor Toetsen en Examens, (2016). *Staatsexamens Nederlands als tweede taal (NT2)*. Opgeroepen op 21.02.2016 via https://www.hetcvte.nl/item/staatsexamens_nederlands_als.
- Corpus gesproken Nederlands, (2000). *Protocol voor Orthografische Transcriptie*. Opgeroepen op 12.7.2015 via http://lands.let.ru.nl/cgn/doc_Dutch/topics/version_1.0/annot/orthography/ort_prot.pdf.
- Craig, H.K., Hinckley, J.J., Winkelseth, M., Carry, L., Walley, J., Bardach, L., Higman, B., Hilfinger, P., Schall, C., & Sheimo, D. (2007). Quantifying connected speech samples of adults with chronic aphasia. *Aphasiology*, 7, 155-163.
- De Vet, H.C.W., Terwee, C.B., Knol, D.L., & Bouter, L.M. (2006). When to use agreement versus reliability measures. *Journal of Clinical Epidemiology*, 59, 1033-1039.
- Dharmaperwira-Prins, R., & Maas, W. (2006). *Afasie: Beschrijving, onderzoek en behandeling*. Amsterdam: Wimpel-uitgeverij.

- Dijkstra, N. (2006). *Begrijpelijkheid of propositiemodel? Het beoordelen van verbale zelfredzaamheid met de Amsterdam-Nijmegen Test voor Alledaagse Taalvaardigheid*. Afstudeerproject Taal- en Spraakpathologie, Radboud Universiteit Nijmegen.
- Egberink, I.J.L., Janssen, N.A.M., & Vermeulen, C.S.M (2010). *Documentatie van Tests en Testresearch in Nederland, Researchbeschrijving Amsterdam – Nijmegen Test voor Alledaagse Taalvaardigheden (ANTAT)*. Nederlands Instituut van Psychologen (NIP), Amsterdam: Boom test uitgevers.
- Elias, J., & Saucier D. (2006). *Neuropsychology: Clinical and experimental foundations*. Boston: Pearson Education, Inc.
- Evers, A., Caminada, H., Koning, R., Laak, J. ter, Maesen de Sombreff, P. van der, & Starren, J. (1988). *Richtlijnen voor ontwikkeling en gebruik van psychologische tests en studie toetsen*. Amsterdam: NIP.
- Evers, A., Lucassen, W. Meijer, R., & Sijtsma, K. (2010). *COTAN Beoordelingssysteem voor de kwaliteit van tests*. Opgeroepen op 18.06.2015 via http://www.psynip.nl/tests_cotan/cotan-beoordeling.html
- Field, A. (2009). *Discovering Statistics Using SPSS*. London: Sage Publications.
- Filipinski, F. (2014). *Het scoringsmodel ANTATcu voor de kwantitatieve meting van de functionele communicatievaardigheid bij afasie Laveren tussen een model met standaardreacties en het behoud van de variatie in spontaan taalgebruik*. Afstudeerproject Taal- en Spraakpathologie, Radboud Universiteit Nijmegen.
- Frattali, C. M. (1992). Functional assessment of communication: Merging public policy with clinical views. *Aphasiology*, 6, 63–83.
- Giessen, M. (2015). *Aanpassing van scenario's van de Amsterdam-Nijmegen Test voor Alledaagse Taalvaardigheden, een verbetering van de scenario's met als doel een duidelijkere rolverdeling en verbeterde actualiteit en inleefbaarheid*. Afstudeerproject Taalwetenschap, Radboud Universiteit Nijmegen.
- Goodglass, H., & Kaplan, E. (1972). *Boston Diagnostic Aphasia Examination*. Philadelphia: Lea & Febiger.
- Graetz, R., De Bleser, R., & Willmes, K. (1992). *De Akense Afasie Test, Nederlandse Versie*. Lisse: Swets & Zeitlinger. Amsterdam: Hogrefe Uitgevers.
- Huber, W., Poeck, K., Weniger, D., & Willems, K. (1983). *AAT - Aachener Aphasia Test*. Göttingen, Toronto, Zürich: Hogrefe Verlag.
- Huber, W., Poeck, K., Weniger, D., & Willems, K. (1983). *AAT - Aachener Aphasia Test. Handanweisung (134f)*. Göttingen, Toronto, Zürich: Hogrefe Verlag.
- IBM (2015). *IBM SPSS Statistics Version 23*. IBM Corporation.
- International Test Commission (2000). *ITC Test Adaptation Guidelines*. Opgeroepen op 19.08.2015 via www.intestcom.org.
- Jöreskog, K.G., & Sörbom, D.G. (1988). *Analysis of linear structural relationships by maximum likelihood, Guide to the program and applications*. Chicago: SPSSInc.

- Kersting, M. (2006). *"DIN SCREEN". Leitfaden zur Kontrolle und Optimierung der Qualität vor Verfahren und deren Einsatz bei beruflichen Eignungsbeurteilungen*. Lengerich: Pabst Science Publishers.
- Krippendorff, K. (1970). Estimating the reliability, systematic error and random error of interval data. *Educational and Psychological Measurement*, 30, 61-70.
- Liepold, M., Ziegler, W., & Brendel, B. (2003). *Hierarchische Wortlisten: Ein Nachsprechtest für die Sprechapraxiediagnostik*. Opgeroepen op 06.10.2015 via http://ekndata.userweb.mwn.de/hpdf/MAT_13_HWL.pdf.
- Lomas, J., Pickard, L. Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (1989). *CETI- Communicative Effectivness Index*. Schlenck, C. & Schlenck, K.J. (1994, vertaald). Opgeroepen op 19.11.2015 via <http://www.aphasie.org/de/fachpersonen/frageboegen>
- Lotgering, E. (2009). *Functionele communicatie bij afasiepatiënten, Het meten van de communicatieve efficiëntie met de ANTATcu als operationalisatie van de eenheden van de ANTATprop*. Afstudeerproject Taal- en Spraakpathologie, Radboud Universiteit Nijmegen.
- Microsoft Corporation (2016). *Microsoft Office Home and Student 2016*. Microsoft-Unternehmensgruppe.
- Nicholas, L.E., & Brookshire, R.H. (1993). A System for Quantifying the Informativeness and Efficiency of the Connected Speech of Adults With Aphasia. *Journal of speech, language, and hearing research*, 36, 338-350.
- Nicholas, L., & Brookshire, R.H. (1995). Presence, completeness and accuracy of main concepts in the connected speech of nonbrain injured adults. *Journal of Speech Hearing Research*, 38, 145-156.
- Rentsch, H.-P. (2005). *Grundlagen der International Classification of Functioning, Disability and Health (ICF)*. ICF in der Rehabilitation, 17-40. Idstein: Schulz-Kirchner.
- Rietvel, T. (2013). *Chapter 3: Categorical data obtained by subjective measuring*. Cursus: Experimentele technieken, Radboud Universiteit Nijmegen.
- Ritzen, K., Spreuwenberg, F., & Senses, K. (2005). *Uitgebreide toelichting van het meetinstrument, Amsterdam – Nijmegen Test voor Alledaagse Taalvaardigheden (ANTAT)*. Opgeroepen op 26.05.2015 via http://meetinstrumentenzorg.nl-/Portals/0/bestanden/133_1_N.pdf
- Ruiter, M.B., Kolk, H.H.J., Rietveld, T., Dijkstra, N., & Lotgering, E. (2011). Towards a quantitative measure of verbal effectiveness and efficiency in the Amsterdam-Nijmegen Everyday Language Test (ANELT). *Aphasiology* 25 (8), 961-975.
- Sage (2013). *Sage researchmethods*. Opgeroepen op 28.06.2015 via <http://srmo.sagepub.com/view/encyc-of-research-design/n206.xml>.
- Schneider, B., Wehmeyer, M., & Grötzbach, H. (2012). *Aphasie: Wege aus dem Sprachdschangel*. Berlin, Heidelberg: Springer.
- Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing reliability. *Psychological Bulletin*, 86, 420-428.
- Spearman, C. (1910). Correlation calculated with faulty data. *British Journal of Psychology*, 3, 271-295.

- Aan de Stegge, S. (2015). *Aanpassing van ANTAT-scenarios, een onderzoek naar de mogelijkheid tot aanpassing van scenario's uit de ANTAT ter verbetering van actualiteit, inleefbaarheid en verbale eenduidigheid*. Afstudeerproject Taalwetenschap, Radboud Universiteit Nijmegen.
- Technische Universität Dresden (2014). *Übersicht über die Begrifflichkeiten der methodischen Grundlagen und der Versuchsplanung*. Opgeroepen op 16.06.2015 via <http://elearning.tu-dresden.de/versuchsplanung/e35/e234/e255>.
- Van der Meulen, A.C., & van de Sandt-Koenderman, W.M.E. (2013). *Scenario Test*. Houten: Bohn Stafleu van Loghum.
- Van der Meulen, I., van de Sandt-Koenderman, W.M.E., Duivenvoorden, H.J., & Ribbers, G.M. (2010). Measuring verbal and non-verbal communication in aphasia: reliability, validity, and sensitivity to change of the Scenario Test. *International Journal of Language & Communication Disorders* 45 (4), 424–435.
- Verhagen, A., & Alessie, J. (2014). *Evidence based diagnostiek van het bewegingsapparaat*. Bohn Stafleu van Loghum, onderdeel van Springer Media BV.
- Weiß, C. (2010). *Basiswissen Medizinische Statistik*. Heidelberg: Springer Medizin Verlag.
- WHO (2005). *ICF – Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit*. Genf: Deutsches Institut für Medizinische Dokumentation und Information (DIMDI).
- Wijk, C. van. (2000). *Toetsende Statistiek: Basistechnieken*. Bussum: Coutinho.
- Yorkston, K. M., & Beukelman, D. R. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of speech and hearing disorders*, 45, 27–36.