Radboud Universiteit

DONDERS INSTITUTE

Zinsbegrip in primair progressieve afasie classificatie

Bruikbaarheid van de zinsbegripstest "SynTest" in differentiaaldiagnostiek van logopenische en niet-vloeiende primair progressieve afasie

Pauline G. M. van Barschot

Masterscriptie Taal- en Spraakpathologie
Faculteit der Letteren

Zinsbegrip in primair progressieve afasie classificatie

Pauline G. M. van Barschot

Inhoud

Inleiding 5
Methode 9
Methode doel 1 9
Participanten 10
Testen en Materiaal 10
Procedure 12
Statistische analyses 12
Methode doel 2 13
Participanten 13
Materiaal 14
Procedure 14
Statistische analyses 14
Resultaten 15
Resultaten doel 1 15
Resultaten doel 2 17
Discussie 20
Doel 1 Maximalisatie van het onderscheid tussen nfv-PPA en lv-PPA 20
Doel 2 Optimalisatie van de SynTest 23
Tekortkomingen 25
Implicaties voor vervolgonderzoek 26
Referenties 28
Bijlagen 32
Bijlage 1 Opleiding scoringssysteem volgens Verhage 32

Abstract

Primary progressive aphasia (PPA) is a neurodegenerative disorder which is primarily characterized by a gradual decline of language abilities. Three variants or subtypes can be distinguished: semantic PPA (sv-PPA), nonfluent PPA (nfv-PPA) and logopenic PPA (lv-PPA). Differentiation between the nfv-PPA and the lv-PPA variant is highly complex in particular, partly due to overlapping linguistic features. Both subtypes also present sentence comprehension deficits, but these are due to other causal underlying mechanisms. In lv-PPA patients this would result in sentence comprehension problems primarily when processing long sentences whereas in nfv-PPA patients when processing syntactic complex sentences. The main goal of this study was primarily to distinguish nfv-PPA and lv-PPA patients on the basis a sentence comprehension test (SynTest) that targets these different underlying mechanisms by testing grammatical complexity and sentence length in separate ways. A second goal of this study was to improve the SynTest, such that this test may also be used as a general sentence comprehension test in other patient groups.

For the purpose of the first goal, already collected SynTest data of 15 nfv-PPA and 12 lv-PPA patients were used, which consisted of total scores (of five items) for each of seven different sentence types. For explorative purposes, scores of the abbreviated Boston Nomination Task (BBTK) have been added. A linear discriminant function analysis (LDA) was conducted to investigate whether the two PPA variants can be distinguished based on a combination of sentence type scores. For the purpose of the second goal, SynTest data of 43 cognitively healthy controls were used. The frequency of the scores 2, 1 and 0 for each item and sentence type were analyzed and visualized with barplots.

Results concerning goal 1 showed that with a classification accuracy of 48.1% less than half (13/27) of the included PPA patients could be correctly classified by LDA based on SynTest (and BBTK) scores. The SynTest thus turned out to have an insufficient discriminative power to distinguish the nfv-PPA and lv-PPA variants from each other. Regarding goal 2, the results showed that on item 3 (simplest sentence type) and on all items of sentence type 7 (most complex sentence type) healthy controls obtained considerably more suboptimal scores (0 and 1) than all other sentence types and associated items.

Clearly, more sensitive tests specifically designed for PPA diagnostics is needed to enable and improve differentiation between PPA variants. Current tests have been mostly developed for aphasia diagnostics in the CVA-field, which have been adapted for PPA diagnostics The low classification accuracy may be due to the complexity involved in distinguishing nfv-PPA and lv-PPA, which may simply not be possible by a single test for which its sensitivity and specificity may also be time-dependent. An integrated approach combining the SynTest with multiple other discriminatory linguistic and non-linguistic features has likely more potential to differentiate between nfv-PPA and lv-PPA. Furthermore, item 3 and items of category 7 turned out to be problematic even for cognitively healthy individuals, presumably due to the items' inadequacy and overly complex sentence type, respectively. Suggestions for improvement of item 3 and exclusion of sentence type 7 have been made towards optimizing the SynTest.

Samenvatting

Primair progressieve afasie (PPA) is een neurodegeneratieve aandoening die primair wordt gekenmerkt door een geleidelijke achteruitgang in de taalfuncties. Er worden drie varianten/subtypen onderscheiden: semantische PPA (sv-PPA), niet-vloeiende PPA (nfv-PPA) en logopenische PPA (lv-PPA). Met name het onderscheid tussen de nfv-PPA en de lv-PPA varianten is zeer complex, onder meer door de overlap in talige kenmerken. Zo gaan beide varianten ook gepaard met zinsbegripsproblemen, waar echter een ander onderliggend mechanisme aan ten grondslag ligt. Bij lv-PPA patiënten zou dit resulteren in zinsbegripsproblemen bij de verwerking van lange zinnen en nfv-PPA patiënten bij de verwerking van syntactisch complexe zinnen. Het primaire doel van deze studie was om nfv-PPA en lv-PPA patiënten op basis van deze verschillende onderliggende mechanismen van elkaar te onderscheiden, middels een zinsbegripstest (SynTest) waarin grammaticale complexiteit en zinslengte zo goed mogelijk gescheiden zijn gehouden. Het tweede doel van deze studie was erop gericht de SynTest te verbeteren, zodat deze mogelijk tevens als zinsbegripstest bij andere patiëntgroepen ingezet kan worden.

Ten behoeve van het eerste doel zijn reeds verzamelde SynTest gegevens van 15 nfv-PPA en 12 lv-PPA patiënten gebruikt, bestaande uit totaalscores (per vijf items) van elk van de zeven verschillende zinstypen. Voor exploratieve doeleinden zijn scores op de Verkorte Boston Benoemtaak (BBTK) zijn daar aan toegevoegd. Een ‘linear discriminant function’ analyse (LDA) is uitgevoerd om te onderzoeken of de twee PPA varianten op basis van een combinatie van zinstypen van elkaar te onderscheiden zijn. In het kader van het tweede doel zijn reeds verzamelde SynTest gegevens van 43 cognitief gezonde controles gebruikt. De frequentie van de scores 2, 1 en 0 zijn per item en per zinstype geanalyseerd en weergegeven in barplots.

Uit de resultaten van het eerste doel bleek dat met een classificatie accuratesse van 48.1%, minder dan de helft (13/27) van de geïncludeerde PPA patiënten correct is geclassificeerd door een LDA op basis van SynTest (en BBTK) scores. De SynTest bleek daarmee over onvoldoende discriminatiekracht te beschikken om de twee varianten van elkaar te kunnen onderscheiden. Met betrekking tot doel 2 lieten de resultaten zien dat gezonde controles op item 3 (het meest eenvoudige zinstype) en op alle items van categorie 7 (het meest complexe zinstype) aanzienlijk meer suboptimale scores (0 en 1) behaalden vergeleken met de andere items en zinstypes.

Omdat testmateriaal dat veelal afkomstig is uit het CVA-veld op de PPA diagnostiek wordt toegepast en waarschijnlijk onvoldoende sensitief is om kleine verschillen tussen PPA varianten te detecteren, is er meer sensitief testmateriaal specifiek ontwikkeld voor PPA nodig om differentiatie tussen de PPA varianten te verbeteren. De lage classificatie accuratesse waarde zou verklaard kunnen worden vanuit het idee dat het complexe onderscheid niet gemaakt kan worden op basis van 1 enkele test, waarvan de sensitiviteit en specificiteit mogelijk ook nog tijdsafhankelijk is. Een integrale aanpak waarbij meerdere onderscheidende talige en niet-talige kenmerken met de SynTest worden gecombineerd, heeft wellicht meer potentie om de nfv-PPA en lv-PPA varianten van elkaar te kunnen onderscheiden. Item 3 en alle items van zinstype 7 bleken zelfs voor cognitief gezonde personen problematisch. Suggesties voor verbetering van item 3 en exclusie van zinstype 7 zijn gegeven ten behoeve van het optimaliseren van de SynTest.

Inleiding

Primair Progressieve Afasie (PPA) is een neurodegeneratieve aandoening waar een geleidelijke achteruitgang van de talige functies het meest prominente klinische kenmerk is. Deze achteruitgang vormt tevens de hoofdoorzaak van beperkingen in het dagelijks leven en functioneren (Matías-Guiu & García-Ramos, 2013; Mesulam & Weintraub, 1992). PPA is met een geschatte prevalentie van 3 op de 100.000 personen een zeldzame aandoening (Marshall et al., 2018), die vaak op relatief jonge leeftijd tussen de 50 en 70 jaar ontstaat (Tippett et al., 2015). PPA wordt geclassificeerd onder fronto-temporale dementie (FTD), een overkoepelende term die verwijst naar een familie van syndromen waar specifieke groepen zenuwcellen in de frontale en/of temporale hersenkwabben voortijdig en progressief afsterven. Deze familie syndromen is onderverdeeld in twee sub-syndromen: de gedragsvariant (ook wel 'behavioural variant'; bvFTD) waar gedragsveranderingen met betrekking tot persoonlijkheid en gedrag op de voorgrond staan, en PPA waar voornamelijk veranderingen in zowel expressieve als receptieve taalfuncties tot uiting komen (Kirshner, 2014). Volgens de diagnostische richtlijnen die zijn gedefinieerd door Gorno-Tempini en collega's (2011) wordt een patiënt initieel gediagnosticeerd met PPA, waarna deze daarop volgend wordt gecategoriseerd in een van de drie PPA varianten op basis van specifieke cognitieve en 'neuro-imaging' kenmerken: de semantische (sv-PPA), niet-vloeiende (nfv-PPA) en logopenische variant (lv-PPA). De sv-PPA variant wordt gekenmerkt door semantische tekorten en stoornissen in woordvinding en woordbegrip. De logopenische variant (lv-PPA) wordt gekarakteriseerd door een stoornis in het ophalen en benoemen van woorden in de spontane taalproductie en een stoornis in het herhalen van zinnen. De niet-vloeiende variant (nfv-PPA) kenmerkt zich door agrammatisme in de taalproductie en moeizame, vertraagde spraak ('effortfull speech'), mede als gevolg van de aanwezigheid van spraakapraxie (Gorno-Tempini et al., 2011). Vanwege overlap van deze karakteristieken blijkt het classificeren van de patiënten in de praktijk moeilijk. Classificatie is verder bemoeilijkt door een gebrek aan meetinstrumenten die specifiek zijn voor ontwikkeld voor PPA diagnostiek, waardoor grotendeels uit het CVA-veld afkomstige testen worden toegepast in PPA diagnostiek (Jiskoot et al., 2020).

Vroegtijdig en nauwkeurig diagnosticeren van PPA is van groot belang, omdat therapie op de gebieden van taal en spraak sterk symptoomafhankelijk zijn en iedere patiënt daarmee een andere benadering vereist. Daarnaast is de kans dat therapie effect heeft het grootst is wanneer er zo snel mogelijk na openbaring van de ziektesymptomen wordt gestart (Jiskoot et al., 2020). Vroegtijdig diagnosticeren is tevens cruciaal vanwege de tijdsgevoeligheid van de diagnostiek van (Mesulam et al., 2014). Wanneer dit immers te laat in het ziekteproces plaatsvindt, is het vaak niet meer mogelijk om een onderscheid te maken tussen de drie PPA subtypen, omdat de ziekte dan al te ver is gevorderd naar een meer globale dementie en de degeneratie verder is uitgebreid naar andere niet-talige cognitieve functies, zoals het geheugen. De diagnostiek dient daarentegen ook niet te vroeg verricht te worden, omdat kenmerkende symptomen van PPA zich nog niet voldoende kunnen manifesteren waardoor classificatie tot subtype nog niet mogelijk is met de beschikbare testen (Jiskoot et al., 2020; Mesulam et al., 2014). De diagnostiek van PPA geschiedt multidisciplinair waarbij een arts, neuropsycholoog en klinisch of neurolinguïst betrokken zijn. Voor de diagnostiek worden drie verschillende niveaus gehanteerd: 1) klinisch niveau, 2) ondersteunende beeldvorming en 3) *post-mortem* onderzoek (Jiskoot et al., 2020). Voor ieder niveau zijn er volgens de criteria van Gorno-Tempini et al. (2011) richtlijnen over welke kenmerken verplicht en welke eventueel aanwezig dienen te zijn. Ondanks uitgebreide diagnostiek blijkt echter dat ongeveer 15-40% van de PPA patiënten niet geclassificeerd kan worden (Marshall et al., 2018; Vandenberghe, 2016), met name het onderscheid tussen nfv-PPA en lv-PPA is gecompliceerd (Croot et al., 2012; Gorno-Tempini et al., 2011; Leyton & Hodges, 2014).

Discriminatie tussen niet-vloeiende PPA (nfv-PPA) en logopenische PPA (lv-PPA) wordt sterk bemoeilijkt vanwege de grote mate van overlap in de talige symptomen. Zo bleek uit het onderzoek van Janssen (2020), waarin de diagnostische waarde van de SYDBAT-NL (Eikelboom et al., 2017) voor het onderscheiden van de PPA varianten werd onderzocht, dat de meeste classificatiefouten werden gemaakt tussen de nfv-PPA en lv-PPA varianten. Uit de prestaties op de herhaaltaak op woordniveau van de SYDBAT-NL bleek dat zowel nfv-PPA als lv-PPA patiënten fouten in het herhalen maakten, maar zonder prestatieverschillen tussen deze patiëntgroepen. Woordvindingsproblemen komen ook bij allebei de PPA varianten voor (Gorno-Tempini et al., 2011; Grossman, 2018), waardoor woordvindingsproblemen lastig als criterium kan bijdragen aan het onderscheid tussen de twee varianten. Daarnaast is er bij zowel nfv-PPA als lv-PPA patiënten sprake van een niet-vloeiende taalproductie die in meeste gevallen het gevolg is van respectievelijk spraakapraxie en fonologische fouten (Croot et al., 2012). Het onderscheid tussen een spraakapractische en fonologische stoornis is echter niet eenvoudig. Daarnaast zijn er aanwijzingen dat apractische of fonologische storingen (‘disruption’) mogelijk niet eenduidig verdeeld zijn over de verschillende PPA subtypen (Croot et al., 2012). Dat maakt het onderscheid nog complexer, te meer omdat spraakapraxie een van de twee hoofdkenmerken van nfv-PPA is, en de aanwezigheid van fonologische fouten en afwezigheid van spraakmotorische fouten kenmerken van lv-PPA zijn (Gorno-Tempini et al., 2011). Door meer in te zoomen op het type fouten die werden gemaakt bij spraakapraxie en fonologische tekorten, vonden Croot en collega’s (2012) echter dat de ratio van het type fouten de twee PPA varianten voor het grootste gedeelte van elkaar kon onderscheiden. Tot slot gaat zowel nfv-PPA als lv-PPA vaak gepaard met zinsbegripsproblemen, waarmee wederom sprake is van een overeenkomstig talig symptoom (Gorno-Tempini et al., 2011; Tee & Gorno-Tempini, 2019). Deze studies presenteren echter sterke aanwijzingen dat de aard van deze zinsbegripsproblemen verschillend is voor de twee PPA varianten, vanwege een ander onderliggend pathologisch mechanisme. Het huidige onderzoek richt zich primair op het onderscheid tussen nfv-PPA en lv-PPA op het gebied van tekorten in de syntactische verwerking.

Bij patiënten met nfv-PPA en lv-PPA worden tekorten in de syntactische verwerking gemeten, die echter het gevolg zijn van een ander onderliggend mechanisme. De stoornis in zinsbegrip bij patiënten met nfv-PPA wordt duidelijk beïnvloed door de grammaticale complexiteit van de zin, wat duidt op een stoornis in de verwerking van grammaticaal complexe zinnen (Croot et al., 2012). Daarentegen worden de zinsbegripsproblemen bij lv-PPA patiënten beïnvloed door de lengte en de voorspelbaarheid van de zin (Gorno-Tempini et al., 2011; Wilson et al., 2012). De zinsbegripsproblemen van lv-PPA zijn gerelateerd aan een stoornis in de fonologische lus (Gorno-Tempini et al., 2008; Meyer et al., 2015). De fonologische lus is volgens Baddeley’s model van het werkgeheugen (Baddeley, 1992) een onderdeel van het korte termijngeheugen die auditieve informatie tijdelijk opslaat en herhaalt, waardoor fonologische informatie voor een periode van enkele seconden wordt vastgehouden en wordt ververst. Over de exacte capaciteit van de fonologische lus is geen consensus (Baddeley & Hitch, 2019; Cowan, 2001; Jonides et al., 2008), maar er is brede overeenstemming dat het vasthouden en herhalen nodig is om informatie actief en direct beschikbaar te houden (Peach & Shapiro, 2012). De fonologische lus ondersteunt het begrip door binnenkomende auditieve informatie actief te houden, waardoor syntactische verwerking en interpretatie van woordreeksen mogelijk is (Gorno-Tempini et al., 2008). Bij lv-PPA leidt de stoornis in de fonologische lus tot afbraak van de processen voor het vasthouden en herhalen, met begripsproblemen van lange zinnen als gevolg. De verschillende mechanismen die onderliggend zijn aan de tekorten in syntactische verwerking zijn dus gebruikt om een zinsbegripstest te ontwikkelen die nfv-PPA en lv-PPA patiënten van elkaar onderscheidt (SynTest; Janssen et al., 2018).

De SynTest is een zinsbegripstest die specifiek is ontwikkeld om de nfv-PPA en lv-PPA varianten beter van elkaar te kunnen onderscheiden. In tegenstelling tot de meeste zinsbegripstesten

waarin grammaticale complexiteit en zinslengte in grote mate met elkaar correleren (Janssen, 2020), zijn deze twee aspecten in de SynTest zo goed mogelijk gescheiden gehouden. Dit is bereikt door de zinstypes gedurende de test afzonderlijk van elkaar toe te laten nemen in grammaticale complexiteit en zinslengte. Daarnaast is door de rijkdom aan lexicale informatie in de SynTest zo veel mogelijk te beperken geprobeerd een meetinstrument te ontwikkelen waarmee de syntactische verwerking zo zuiver mogelijk kan worden gemeten. In de gehele test zijn slechts vier karakters gebruikt die door alle items heen dezelfde uiterlijke kenmerken behouden en is het aantal verschillende werkwoorden zo veel mogelijk beperkt. Dit is gedaan omdat nfv-PPA en lv-PPA patiënten verhoogd gevoelig zijn voor interferentie van semantisch gerelateerde informatie. Verschillende studies vonden een (abnormaal) groot semantisch interferentie effect voor de twee PPA varianten in een benoemtaak met semantisch gerelateerde afleiders en 'picture-word interference' taken (Janssen et al., 2020; Thompson et al., 2012; Vandenberghe et al., 2005). Janssen en collega's (2020) stelden vervolgens dat het grote interferentie effect in nfv-PPA en lv-PPA patiënten een algehele kwetsbaarheid voor interferentie van afleiders suggereert. Door de lexicale eenheden door de test heen zo veel mogelijk te minimaliseren, hoeft er per nieuw item slechts een minimale hoeveelheid aan nieuwe lexicale informatie verwerkt te worden, waardoor het geven van een correct antwoord zo min mogelijk afhankelijk is van het vermogen om nieuwe woorden te begrijpen. De mogelijke invloed van afleiders op het gebied van semantiek is hiermee zo veel mogelijk beperkt.

De SynTest heeft ook de potentie om bruikbaar te zijn als meetinstrument voor zinsbegripsproblemen bij aandoeningen ten gevolge van niet-aangeboren hersenletsel, anders dan PPA. Er zijn in de klinische praktijk immers slechts een klein aantal testen waarmee zinsbegripsproblemen kunnen worden gekwantificeerd. Zo is de Verkorte Token Test (De Renzi & Faglioni, 1978) een centrale test in de afasiediagnostiek waarin de patiënt auditieve instructies met figuren in verschillende vormen, kleuren en groottes uit moet voeren. De (auditief aangeboden) zinnen nemen gedurende de test echter tegelijkertijd toe in syntactische complexiteit én zinslengte (Leyton & Hodges, 2014), waardoor zinsbegrip niet zo zuiver voor deze twee aspecten afzonderlijk gemeten kan worden zoals met de SynTest. Ook bevat de Nederlandstalige bewerking van de linguïstische testbatterij Comprehensive Aphasia Test (CAT-NL; Visch-Brink et al., 2014) een subtest die zinsbegripsproblemen detecteert. Door de items heen zijn er echter steeds andere karakters gebruikt (man, hond, postbode, soldaat etc.) en zijn er op de afbeeldingen vaak meer semantische componenten afgebeeld dan de (doel)zin behelst. Net zoals voor PPA patiënten zou het voor afasiepatiënten een meerwaarde zijn wanneer de invloed van afleiders op het gebied van semantiek zo veel mogelijk wordt beperkt in onder meer het onderzoek naar syntactische verwerking. Bij deze patiënten worden immers ook regelmatig grotere semantische interferentie effecten gevonden (Dyson et al., 2021; Hashimoto & Thompson, 2010). De subtest voor zinsbegrip is daarnaast slechts één van de twintig talige subtesten, terwijl de CAT-NL enkel als complete testbatterij wordt verkocht met bijbehorende kosten. De Werkwoorden- en zinnentest (WEZT; Bastiaanse et al., 2000) bevat ook een subtest gefocust op zinsbegrip. De WEZT is echter gedateerd en zal binnenkort vervangen worden door een applicatie (app). Omdat met de SynTest de verwerking van zowel grammaticaal complexe als lange zinnen wordt onderzocht en de test op semantisch niveau zo simplistisch en neutraal mogelijk is gehouden, kent de SynTest voordelen ten opzichte van de reeds bestaande zinsbegripstesten. Terwijl de SynTest dus ook potentie heeft om als zinsbegripstest te fungeren in de klinische praktijk voor meer patiëntgroepen, is er nog niet onderzocht hoe gezonde volwassenen op de SynTest scoren. Het is noodzakelijk om te bepalen of een item en/of zinstype categorie bij controles niet overmatig incorrecte antwoorden oplevert. Indien dit wel het geval is, dienen items en/of zinstype categorieën van de SynTest aangepast of zelfs geëxcludeerd te worden.

Het eerste van de twee gestelde doelen in het huidige onderzoek had betrekking op complexe differentiaaldiagnostiek van PPA. Primair is onderzocht of de niet-vloeiende PPA (nfv-PPA) en logopenische PPA (lv-PPA) varianten beter van elkaar te onderscheiden zijn op basis van tekorten in de syntactische verwerking. De nadruk lag op het onderscheiden van de twee PPA varianten op basis van SynTest gegevens. Omdat van een groot aantal patiënten ook gegevens van de Verkorte Boston Benoemtaak (BBTK) beschikbaar waren, zijn deze tevens uit exploratief oogpunt toegevoegd tot het onderzoek. De hypothese was dat de SynTest zou kunnen discrimineren tussen nfv-PPA en lv-PPA patiënten als gevolg van de verschillende mechanismen die ten grondslag liggen aan de tekorten in de syntactische verwerking. De verwachting was dat nfv-PPA patiënten structureel lager zouden scoren op syntactisch complexe zinstypes en structureel hogere scores zouden behalen op lange zinnen ten opzichte lv-PPA patiënten. Het tweede doel had betrekking op het optimaliseren van de SynTest zodat deze tevens als diagnostisch instrument voor zinsproblemen in andere patiëntgroepen ingezet zou kunnen worden. Daarvoor is onderzocht of er specifieke zinstypen of items waren waar controles geen maximale score op behaalden. Op basis van patronen in voorgaande data was de hypothese dat het zinstype van de laatste zinstype categorie van de SynTest voor gezonde volwassenen mogelijk te moeilijk is vanwege de zeer complexe zinsconstructie.

Methode

SynTest

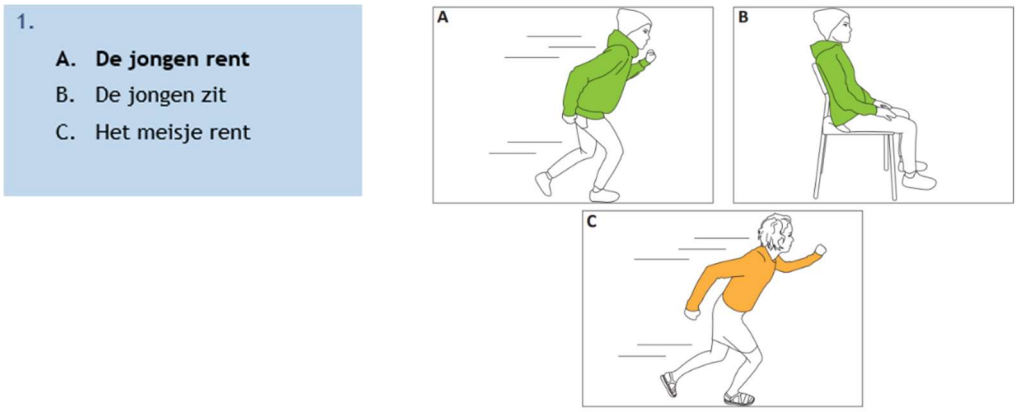
De SynTest (Janssen et al., 2018) is een test waarmee zinsbegrip met toenemende complexiteit wordt onderzocht door middel van matching van een auditief aangeboden zin met een afbeelding. De test is specifiek ontwikkeld om bij te dragen aan het onderscheid tussen lv-PPA en nfv-PPA patiënten. De test bestaat uit 35 items die in sets van 5 gelijkmatig zijn verdeeld over 7 categorieën, waarbij iedere categorie een specifiek zinstype representeert. De categorieën en de daartoe behorende zinstypen zijn weergegeven in tabel 1. Grammaticale complexiteit en zinslengte en de toename in complexiteit en lengte van de zinnen zijn in de verschillende zinstypen zoveel mogelijk van elkaar gescheiden gehouden. Zo is een item in een bepaalde categorie grammaticaal complexer of is de zinslengte toegenomen ten opzichte van een item uit de daaraan voorafgaande categorie. Naast de in totaal 35 zinnen bevatten alle items 4 afbeeldingen, met uitzondering van de items (1-5) die tot de eerste categorie met zinstype 'intransitief' behoren die 3 afbeeldingen bevatten. In de gehele test komen slechts 4 karakters voor (man, vrouw, jongen en meisje) en deze worden consequent met dezelfde uiterlijke kenmerken afgebeeld. Voor ieder item geldt dat er telkens één afbeelding volledig overeenkomt met de betekenis van de zin. Het is de bedoeling dat de patiënt naar aanleiding van een auditief aangeboden zin deze 'correcte' afbeelding selecteert. In figuur 1 is ter illustratie het eerste item van de SynTest weergegeven, waar zin A de doelzin en afbeelding A de doelafbeelding is.

De reactie van de patiënt wordt op een scoreformulier genoteerd en daaraan wordt vervolgens een score van 2, 1, of 0 toegekend. De score 2 wordt toegekend wanneer de doelafbeelding wordt aangewezen, de score 0 wanneer er een andere afbeelding dan de doelafbeelding wordt aangewezen. Score 1 wordt toegekend wanneer de doelafbeelding wordt aangewezen als gevolg van een zelfcorrectie, na herhaling van de zin door de onderzoeker of indien de reactietijd van de patiënt langer is dan 5 seconden. De afname van de SynTest resulteert middels de som van de scores van vijf items in een totaalscore per zinstype met een maximale waarde van 10. Opgeteld leidt dit tevens tot een algehele totaalscore voor de SynTest van maximaal 70.

Tabel 1

Categorieën en zinstypen van de SynTest

Categorie	*Items*	*Zinstype*	*Voorbeeld (eerste item per categorie)*
1	1 - 5	Intransitief	De jongen rent
2	6 - 10	Regelmatig actief	Het meisje draagt de jongen
3	11 - 15	Twee uitvoerders actief	De jongen eet een koek en het meisje drinkt de melk
4	16 - 20	Passief	De man wordt gevolgd door de vrouw
5	21 - 25	Een uitvoerder twee acties	Het meisje speelt een spel en drinkt de melk
6	26 - 30	Subject relatieve bijzin	De jongen die rent volgt het meisje
7	31 - 35	Relatieve bijzin met eind van zin oplossing	Het meisje kust de vrouw die de mannen volgen



Figuur 1. Voorbeeld van SynTest item. Een zin (doelzin is dikgedrukt) wordt auditief aangeboden en moet correct worden geassocieerd met de bijbehorende afbeelding

Methode doel 1

Participanten

Voor het onderzoek in het kader van doel 1 zijn data gebruikt van 31 patiënten, van wie 15 gediagnosticeerd waren met logopenische primair progressieve afasie (lv-PPA) en 16 met niet-vloeiende primair progressieve afasie (nfv-PPA). Alle patiënten hebben Nederlands als moedertaal. De data van deze patiënten waren reeds verzameld en afkomstig van twee verschillende bronnen. Gegevens van 11 patiënten waren afkomstig van data verzameld in het kader van het proefschrift van Janssen (2020), waarin onder meer onderzoek is gedaan naar primair progressieve afasie (PPA). De data van 20 andere patiënten waren afkomstig van het Erasmus Medisch Centrum te Rotterdam.

Er zijn twee lv-PPA patiënten van het onderzoek geëxcludeerd, omdat bleek dat deze twee patiënten tweemaal in de steekproef vertegenwoordigd waren. De oorzaak was dat in het onderzoek van Janssen (2020) patiëntdata zijn aangeleverd door het Erasmus MC waarin deze twee patiënten tevens voorkwamen. Doordat de data afkomstig van de twee bronnen in het huidige onderzoek werden samengevoegd tot één databestand, kwamen de twee desbetreffende patiënten tweemaal voor in de data. Daarom moesten de gegevens van deze patiënten er eenmaal uitgehaald worden, waarbij ervoor is gekozen om de gegevens die het eerst zijn verzameld te behouden. Tot slot is er nog één lv-PPA patiënt en één nfv-PPA patiënt van het onderzoek geëxcludeerd vanwege het ontbreken van respectievelijk een waarde in de SynTest en de score van de Verkorte Boston Benoemtaak (BBTK). Dit resulteerde uiteindelijk in een steekproefgrootte van 27 patiënten die bestond uit 12 lv-PPA patiënten en 15 nfv-PPA patiënten. Voorafgaand aan de analyse van deze data is eerst onderzocht of er sprake was van significante groepsverschillen op het gebied van leeftijd, opleidingsniveau, ziekteduur en prestatie op een cognitieve screening.

Materiaal

Ten behoeve van doel 1 zijn als materiaal primair scores van de SynTest en daarnaast de scores van de Verkorte Boston Benoemtaak (BBTK; van Loon - Vervoorn, 2005) gebruikt. De patiëntdata uit het onderzoek van Janssen (2020) waren reeds voorbewerkt. Echter, de patiëntdata afkomstig van het Erasmus MC bestonden echter nog uit ruwe scoreformulieren van de SynTest en de Boston Benoemtaak

(BBT; van Loon - Vervoorn et al., 1994). Deze data zijn eerst gedigitaliseerd, voorbewerkt en vervolgens toegevoegd met de data van Janssen (2020).

Voor de gegevens van de SynTest geldt dat de scores op de ruwe scoreformulieren per item allereerst zijn omgezet naar een accuratesse code. Het voornaamste verschil tussen deze scores was dat er met de *accuratesse codes* wordt gedifferentieerd tussen een correct antwoord na zelfcorrectie, herhaling of langere reactietijd, in tegenstelling tot de *scores* waarbij er in elk van deze drie gevallen de score 1 wordt toegekend. In tabel 2 zijn de accuratesse codes weergegeven. Voor ieder scoreformulier zijn afzonderlijk de scores omgezet naar accuratesse codes. Daarnaast is van elk incorrect antwoord, al dan niet gevolgd door een zelfcorrectie (antwoorden met accuratesse codes 0, 2 en 4), genoteerd in welk opzicht de incorrect gekozen afbeelding verschilde van de doelafbeelding. Hierin waren er drie mogelijke opties: 1) 'verb' is genoteerd wanneer er een andere actie werd uitgevoerd dan in de doelafbeelding, 2) 'role' is genoteerd wanneer de actie door een ander karakter werd uitgevoerd dan in de doelafbeelding en 3) 'role-verb' is genoteerd wanneer er sprake van zowel optie 1 als 2. Indien de resultaten op basis van de totaalscores onderscheidend zijn voor de twee PPA subtypen, zou onderscheid tussen type fouten aanvullend inzicht kunnen geven. Een PPA variant zou dan bijvoorbeeld vaker een specifiek type fout kunnen hebben gemaakt vergeleken met de andere PPA variant.

De aangeleverde BBT scores moesten omgezet worden naar BBTK scores, om deze samen te kunnen voegen met bestaande BBTK data. In het onderzoek van Janssen (2020) zijn ook patiëntgegevens aangeleverd door het Erasmus MC, inclusief BBT scores die vervolgens door Janssen op een specifieke manier omgezet naar BBTK scores. In het huidige onderzoek zijn de BBT scores van ruwe BBT scoreformulieren afkomstig van het Erasmus MC op eenzelfde wijze omgezet. Elk scoreformulier bestond uit 60 items waarvan de geproduceerde uitingen dichotoom waren gescoord met 'correct' of 'incorrect'. Alle uitingen die afweken van de doeluiting waren reeds op het scoreformulier genoteerd. Ook was genoteerd of een semantische en/of fonologische cue was verschaft door de onderzoeker en of dit wel of niet had geleid tot productie van het doelwoord. Een productie van het doelwoord naar aanleiding van een semantische of fonologische cue was op het scoreformulier altijd gescoord als 'incorrect'.

Uit elk scoreformulier zijn de 29 items die in de BBTK voorkomen geëxtraheerd. Vervolgens zijn de op het scoreformulier genoteerde antwoorden op deze 29 items gescoord volgens het scoringssysteem van de BBT (Roomer et al., 2011). Een correct antwoord als gevolg van een semantische of fonologische cue is consequent de score 0 toegekend. Bij enkele ruwe scoreformulieren kwam het voor dat er bij een bepaald item geen reactie of een foutieve reactie is genoteerd, maar daarnaast was ook het doelwoord onderstreept en het item als correct gescoord. In deze gevallen is er respectievelijk een score 3 (correct antwoord) en een score 2 (correct antwoord als gevolg van adequate zelfcorrectie) toegekend. Volgens het scoringssysteem van de BBT dient een correcte productie van het doelwoord na een periode van 10 seconden de score 2 toegekend te worden, maar omdat deze informatie niet beschikbaar was, kon reactietijd niet worden meegenomen in de scoring. Echter, omdat langere reactietijden niet waren gespecificeerd met een tijdsaanduiding is reactietijd niet meegenomen in de scoring. De toegekende scores van de 29 items zijn vervolgens bij elkaar opgeteld, waardoor er een BBTK totaalscore ontstond met een maximale waarde van 87. De omgezette BBTK scores zijn samengevoegd met de BBTK scores afkomstig van het onderzoek van Janssen (2020).

Tot slot zijn er gegevens van leeftijd, opleidingsniveau en ziekteduur gebruikt om te testen of patiëntgroepen verschilden. Gegevens van ziekteduur zijn gebaseerd op basis van hoelang de ziektesymptomen met betrekking tot PPA bij de patiënt (opgemerkt) aanwezig zijn. Gegevens met betrekking tot opleidingsniveau zijn uitgedrukt in opleidingscodes volgens het scoringssysteem van Verhage (Verhage, 1964). De zeven verschillende opleidingscodes die in dit scoringssysteem worden onderscheiden zijn weergegeven in bijlage 1. Daarnaast zijn de behaalde scores op de Nederlandse versies van de cognitieve screeningsinstrumenten Montreal Cognitive Assessment (MoCA; Dautzenberg & de Jonghe, 2004) en Mini-Mental State Examination (MMSE; Kok & Verhey, 2002)

gebruikt als globale indicatoren voor de algehele cognitieve achteruitgang. Doordat de data van het onderzoek van Janssen (2020) uitsluitend MoCA scores bevatte en de data van het Erasmus MC uitsluitend scores van de MMSE, zijn niet de scores gebruikt voor de analyse, maar is voor iedere patiënt bepaald of er een score lager dan de cutoff waarde (MoCA = 26; MMSE = 23) is behaald, of gelijk aan of hoger dan de cutoff waarde. Zo kon de data toch met elkaar verenigd en vergeleken worden, ondanks dat de scores afkomstig waren van twee verschillende screeningsinstrumenten.

Tabel 2

Accuratesse codering van de SynTest

Accuratesse code	*Antwoord*
1	correct
0	incorrect
2	adequate zelfcorrectie
3	correct na herhaling van de zin
4	adequate zelfcorrectie en herhaling van de zin
5	correct maar vertraagd

Procedure

Ten behoeve van het testen van groepsverschillen zijn data met betrekking tot leeftijd, opleidingsniveau, ziekteduur en prestatie op cognitieve screenings van alle patiënten ingelezen in R (R Core Team, 2021), het statistisch programma waarmee in het huidige onderzoek de analyses zijn uitgevoerd. Voor elk van de vier variabelen (leeftijd, opleidingsniveau, ziekteduur en prestatie op cognitieve screenings) visualiseert een dotplot het aantal en spreiding van de data. De packages tidyverse, ggplot2, RColorBrewer, ggpubr, rstatix en MASS zijn gebruikt voor de bewerkingen en analyse van de data (Kassambara, 2020, 2021; Neuwirth, 2014; Venables & Ripley, 2002; Wickham, 2016; Wickham et al., 2019).

Allereerst zijn de SynTest data en gegevens met betrekking tot diagnoses van alle patiënten ingelezen in R. De accuratesse codes waarin de SynTest data waren uitgedrukt, zijn vervolgens weer omgezet naar scores door aan accuratesse code 0 een score 0 toe te kennen, code 1 een score 2 en code 2, 3, 4 en 5 een score 1. Op basis van deze scores zijn totaalscores per categorie berekend, zodat de SynTest data voor iedere patiënt bestond uit zeven totaalscores (categoriescores) die maximaal een waarde van 10 konden aannemen. Daar zijn vervolgens BBTK scores van diezelfde patiënten aan toegevoegd, waardoor de inputdata van iedere patiënt uiteindelijk bestond uit zeven SynTest scores en een BBTK score. De packages tidyverse, caret (Kuhn, 2020) en MASS zijn gebruikt om de bewerkingen aan en de analyse over de data uit te kunnen voeren.

Statistische analyses

Om te onderzoeken of de twee patiëntgroepen in leeftijd en ziekteduur van elkaar verschillen, is er tweemaal een Welch's t-test uitgevoerd omdat het meetniveau van deze (afhankelijke) variabelen van ratio meetniveau zijn. In de resultatensectie zijn, waar van toepassing, afgeronde vrijheidsgraden gerapporteerd. Om groepsverschillen te testen voor de (afhankelijke) variabele opleidingsniveau is er een ordinale logistische regressie analyse uitgevoerd vanwege het ordinale meetniveau. Vanwege de binomiale aard van de afhankelijke variabele is op basis van Chi kwadraat scores getest of de twee

patiëntgroepen significant van elkaar verschilden in de kans dat patiënten onder of gelijk aan en boven de cutoff waarde van een cognitieve screening scoorden. De variabele diagnose was in alle vier de testen de onafhankelijke variabele.

Er is een 'linear discriminant function' analyse (LDA) uitgevoerd om de discriminatiekracht van de SynTest voor de twee patiëntgroepen te onderzoeken. LDA is gericht op het detecteren welke combinatie van variabelen het onderscheidend vermogen maximaliseert tussen categorieën die al bekend zijn. In het huidige onderzoek is getracht het onderscheidend vermogen tussen lv-PPA patiënten en nfv-PPA patiënten te maximaliseren, primair op basis van de SynTest en aan diezelfde SynTest LDA zijn gegevens van de BBTK complementair toegevoegd. De inputdata van de LDA waren 1) de totaalscores per categorie per patiënt van de SynTest en 2) de totaalscores van de BBTK, waarmee de LDA in één keer is uitgevoerd. De output van de LDA (inclusief 'leave-one-out' cross validatie) was de accuratesse van de voorspelling voor de diagnose op basis van de combinatie van SynTest met en zonder en BBTK gegevens.

Methode doel 2

Participanten

Voor het onderzoek in het kader van doel 2 zijn gegevens van 43 cognitief gezonde volwassenen gebruikt. Deze controle data zijn reeds geworven in het kader van het onderzoek van Janssen (2020) en van deze controle participanten zijn gegevens waaronder demografische gegevens en gegevens van de SynTest verzameld. Van deze controle participanten is bekend dat allen moedertaalsprekers van het Nederlands zijn. De demografische gegevens zijn weergegeven in tabel 3.

Tabel 3

Demografische gegevens controles

	Controles		
N	43		
Geslacht (aantal mannen/vrouwen)	22/21		
Gemiddelde leeftijd in jaren (SD), range	66.65 (8.58), 48-85		
Opleidingsniveau (aantal per opleidingscode)[a]	1 = 0 2 = 2 3 = 2	4 = 9 5 = 2 6 = 18	7 = 10
Gemiddelde score MoCA (SD)[b]	27 (1.76)		

Noot.

[a] Opleidingsniveau volgens Verhage (1964) met "1" = Basisonderwijs, "2" = Lager of voorbereidend beroepsonderwijs, VMBO-B, "3" = Middelbaar Algemeen Voortgezet Onderwijs (MAVO), VMBO-T, "4" = Middelbaar Beroepsonderwijs (MBO), Beroepsopleidende Leerweg (BOL), Beroepsbegeleidende leerweg (BBL), "5" = Hoger algemeen voorbereidend onderwijs (HAVO), VWO, "6" = Hoger Beroepsonderwijs (HBO), "7" = Wetenschappelijk Onderwijs (WO) (vollediger format van het scoringssysteem volgens Verhage is in bijlage 1 gepresenteerd)

[b] MoCA = Montreal Cognitive Assessment (Nederlandse versie)

Materiaal

Het materiaal bestond uit SynTest gegevens van de controles, die reeds waren verzameld in het kader van het al langer lopende onderzoek naar PPA aan het Donders Instituut te Nijmegen. Voor alle controles afzonderlijk waren de SynTest gegevens uitgedrukt in accuratesse codes (tabel 2) en de incorrecte antwoorden waren daarnaast gekenmerkt met het type fout op de manier zoals in het onderdeel 'materiaal' van de methodesectie van doel 1 beschreven.

Procedure

Om te onderzoeken of er bepaalde items en/of categorieën van de SynTest verbeterd of zelfs van de test geëxcludeerd zouden moeten worden, omdat dit item en/of categorie zelfs bij controles veel incorrecte antwoorden opleverde, zijn het aantal behaalde scores 0, 1 en 2 op een bepaald item per categorie in grafieken weergegeven. Hiertoe zijn de SynTest data ingelezen in R en zijn de accuratesse codes omgezet naar de scores 0, 1 of 2, op de manier zoals in het onderdeel 'procedure' van de methodesectie van doel 1 is beschreven. Voor iedere categorie is een staafdiagram gemaakt en de packages tidyverse, ggplot2 (Wickham, 2016), forcats (Wickham, 2021) en rstatix zijn gebruikt om de bewerkingen aan de data te kunnen doen en staafdiagrammen te maken.

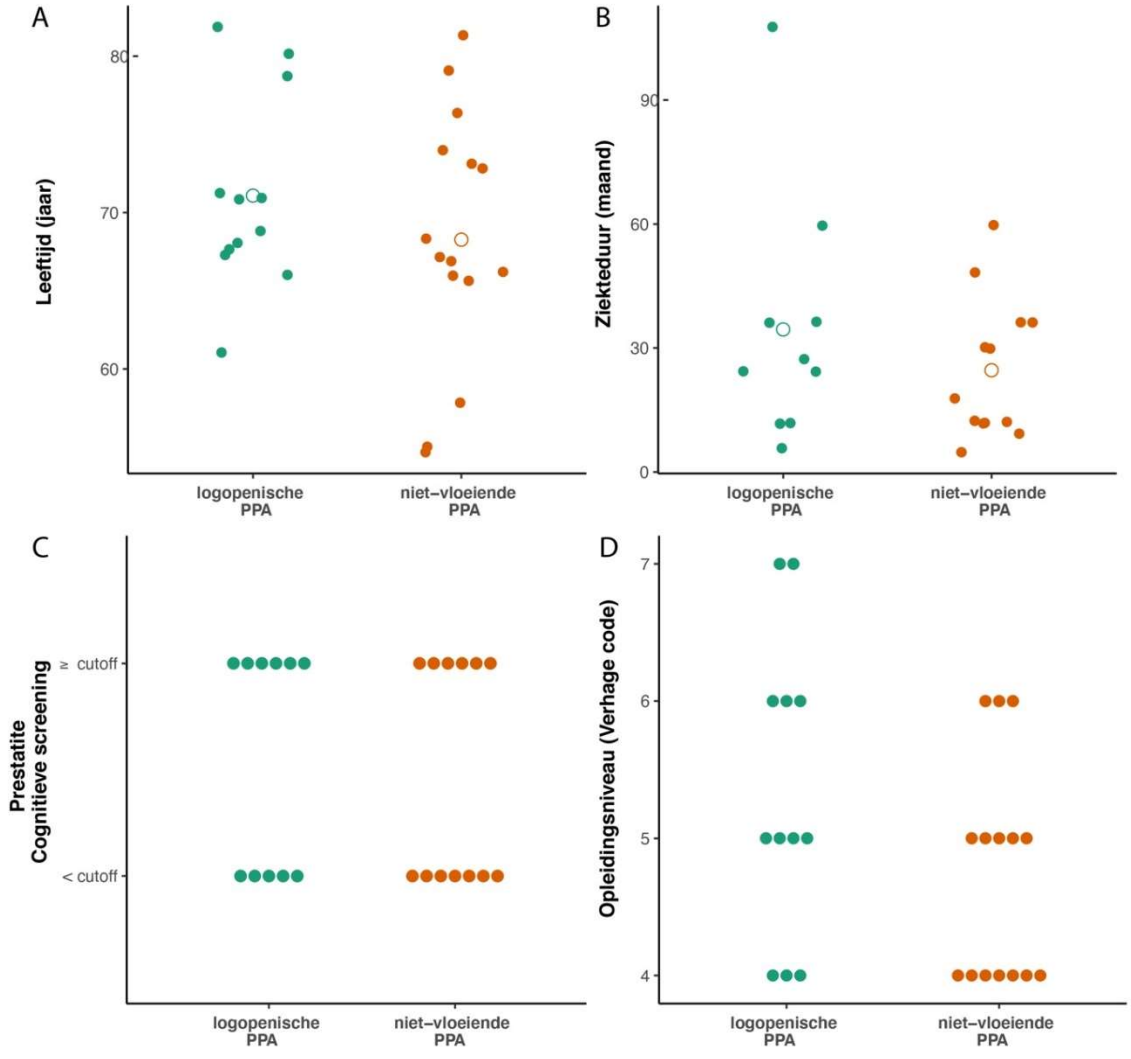
Statistische analyses

Er zijn twee strategieën toegepast om te onderzoeken of bepaalde categorieën ook voor controles problematisch zijn. De eerste strategie was gericht op het visualiseren van de data in staafdiagrammen, waarin de data is uitgedrukt in het aantal behaalde scores 2, 1 en 0 per item. Met deze staafdiagrammen is er per item inzichtelijk gemaakt hoe vaak er een suboptimale score 0 of 1 is behaald, waardoor het mogelijk was om de voor controles problematische of te complexe items en/of categorieën te detecteren. Met de tweede strategie is getracht middels statistische analyses per categorie van de SynTest te onderzoeken of de door de controles behaalde categoriescores significant afweken van de maximaal haalbare categoriescore (10). Hiertoe is de non-parametrische Wilcoxon rang-teken-toets uitgevoerd, waarbij de mediaan van de behaalde categoriescores per categorie steeds is vergeleken met de maximaal haalbare categoriescore.

Resultaten

Resultaten doel 1

Voordat een LDA is uitgevoerd op de Syntest en BBTK scores is onderzocht of er significante groepsverschillen waren op het gebied van leeftijd, opleidingsniveau, ziekteduur en prestatie op een cognitieve screening. Om allereerst een idee te krijgen van de spreiding in de data, zijn de datapunten per patiëntgroep voor de vier variabelen afzonderlijk (leeftijd, opleidingsniveau, ziekteduur en prestatie op een cognitieve screening) middels een dotplot gevisualiseerd (figuur 2). In figuur 2 D missen data voor één lv-PPA en twee nfv-PPA patiënten voor de variabele prestatie op een cognitieve screening. Deze drie patiënten zijn niet meegenomen in de Chi kwadraat test voor groepsverschillen. Ook voor de variabele ziekteduur gold dat er van twee lv-PPA en twee nfv-PPA patiënten een waarde ontbrak, waardoor deze patiënten niet zijn meegenomen in de ongepaarde t-test voor groepsverschillen.



Figuur 2. Verdeling van demografische gegevens van lv-PPA en nfv-PPA patiënten. Variatie in A) leeftijd, B) ziekteduur, C) opleidingsniveau en D) prestatie cognitieve screening per patiëntgroep. Gevulde cirkels representeren individuele patiënten. De open cirkels in plot A en B representeren gemiddelden per groep.

Tabel 4

Demografische gegevens

	lv-PPA	**nfv-PPA**	**N**
N	12	15	27
Sex (aantal mannen/vrouwen)	7/5	5/10	27
Gemiddelde leeftijd in jaren (SD)	71.08 (6.24)	68.27 (8.00)	27
Opleidingsniveau (aantal per opleidingscode)[a]	1 = 0 5 = 4 2 = 0 6 = 3 3 = 0 7 = 2 4 = 3	1 = 0 5 = 5 2 = 0 6 = 3 3 = 0 7 = 0 4 = 7	27
Gemiddelde ziekteduur in maanden (SD)	34.50 (30.11)	24.62 (16.89)	23
Prestatie cognitieve screening (aantal scores < cutoff / ≥ cutoff)[b]	5/6	7/6	24

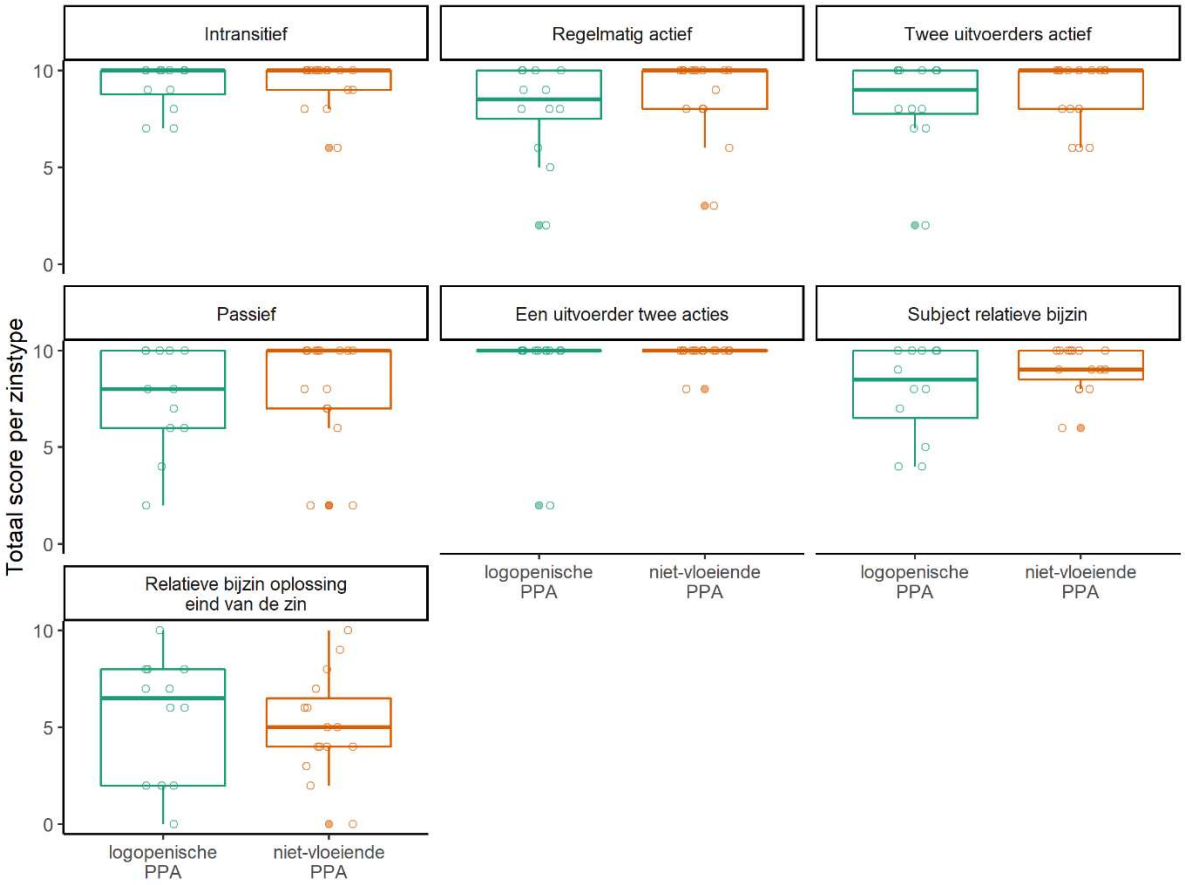
Noot.

[a] Opleidingsniveau volgens Verhage (1964) met "1" = Basisonderwijs, "2" = Lager of voorbereidend beroepsonderwijs, VMBO-B, "3" = Middelbaar Algemeen Voortgezet Onderwijs (MAVO), VMBO-T, "4" = Middelbaar Beroepsonderwijs (MBO), Beroepsopleidende Leerweg (BOL), Beroepsbegeleidende leerweg (BBL), "5" = Hoger algemeen voorbereidend onderwijs (HAVO), VWO, "6" = Hoger Beroepsonderwijs (HBO), "7" = Wetenschappelijk Onderwijs (WO) (vollediger format van het scoringssysteem volgens Verhage is in bijlage 1 gepresenteerd)

[b] Prestatie cognitieve screening is gebaseerd op of de score op de MMSE of MoCA lager (<) of gelijk aan/hoger dan de cutoff waarde (MMSE = 23; MoCA = 26)

De lv-PPA en nfv-PPA patiëntgroepen verschilden niet in de vier demografische variabelen. Statistische analyses bevestigden dat er daadwerkelijk geen sprake is van groepsverschillen. Allereerst voor de variabelen leeftijd en ziekteduur een t-test uitgevoerd. Het bleek dat er geen significant verschil was in leeftijd tussen de lv-PPA patiënten (M = 71.08, SD = 6.24) en de nfv-PPA patiënten (M = 68.27, SD = 8.00), ($t(25) = 1.03$, $p = .313$) (Fig. 2A). Ook bleek er geen significant verschil in ziekteduur tussen de lv-PPA patiënten (M = 34.50, SD = 30.11) en de nfv-PPA patiënten (M = 24.62, SD = 16.89), ($t(13) = .93$, $p = .368$) (Fig. 2B). Echter, de lv-PPA patiëntgroep bevatte een zeer hoge waarde (108), die als uitbijter gedefinieerd zou kunnen worden. Deze hoge waarde heeft mogelijk voor een groot gedeelte bijgedragen aan het hoge gemiddelde (34.50) en tevens aan de hoge standaarddeviatie (30.11). Er is echter geen correctie uitgevoerd voor deze waarde, enerzijds vanwege het feit dat de ziekte daadwerkelijk 108 maanden bij de desbetreffende patiënt aanwezig was en het excluderen van de uitbijter anderzijds niet zou leiden tot een significant groter groepsverschil, maar juist tot een kleiner groepsverschil. Om te onderzoeken of er sprake was van een groepsverschil op het gebied van prestatie op een cognitieve screening, is er een Chi kwadraat test uitgevoerd. Daaruit kwam naar voren dat de kans dat iemand onder of gelijk aan / boven de cutoff waarde scoort voor de twee patiëntgroepen niet (significant) verschillend was (X^2 (1) = .17, $p = .682$) (Fig. 2C). Tot slot is er middels een ordinale logistische regressie analyse onderzocht of er sprake was van een groepsverschil op het gebied van opleidingsniveau. Uit de resultaten kwam naar voren dat de kans dat iemand uit de nfv-PPA patiëntgroep hoger opgeleid was dan iemand uit de lv-PPA patiëntgroep niet significant was (log odds (*SE)* = 1.113 (.74), $t(1) = -1.51$, $p = 0.130$) (Fig. 2D). Deze resultaten indiceren dat de twee patiëntgroepen op geen van de onderzochte variabelen significant van elkaar verschilden.

Er is daarna onderzocht of de nfv-PPA en lv-PPA varianten van elkaar te onderscheiden zijn op basis van prestaties op de SynTest. Beschouwd per zinstype, verschillen de verdelingen van SynTest totaalscores van de nfv-PPA en lv-PPA patiëntgroepen niet (Fig. 3).



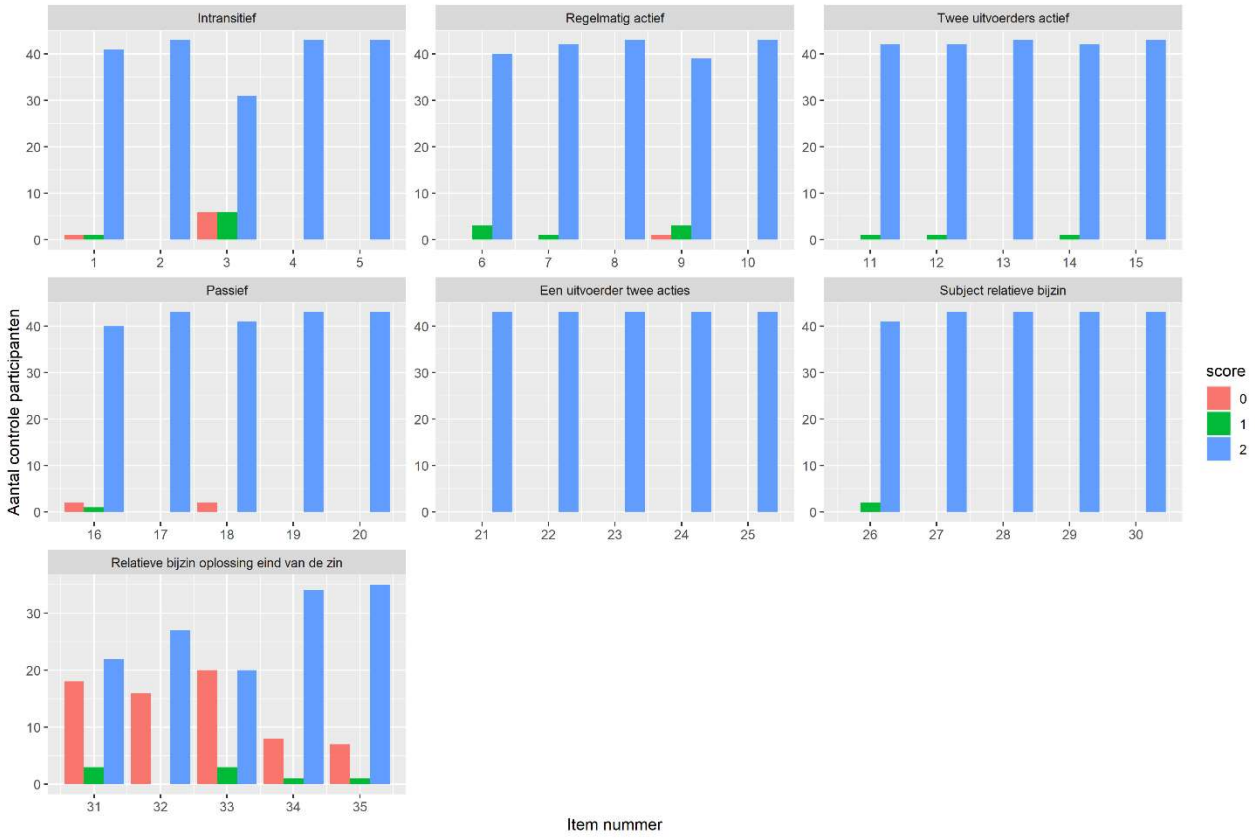
Figuur 3. Totaalscores per zinstype van de SynTest van de lv-PPA en nfv-PPA patiënten. De totaalscores per categorie met het daar toebehorend zinstype zijn van de logopenische PPA en de niet-vloeiende PPA patiënten weergegeven. De x-as representeert de twee patiëntgroepen en de y-as de behaalde totaalscore per zinstype. Elke open cirkel representeert een behaalde totaalscore op het desbetreffende zinstype van één patiënt. De gevulde cirkels duiden aan dat er uitbijters zijn.

De LDA met cross-validatie resulteerde in een classificatie accuratesse van slechts 48.1%. Dat houdt in dat op basis van de voorspellingen van het LDA model minder dan de helft van de patiënten (13/27) correct geclassificeerd werd. De accuratesse verbeterde niet wanneer de BBTK scores werden meegenomen in de analyse. Daarvoor is de LDA met cross-validatie nogmaals uitgevoerd met zowel de gegevens van de SynTest als de BBTK als input. Dit leverde echter een zelf nog lagere accuratesse op van 46.4%. Uit deze analyse is gebleken dat de SynTest en de BBTK geen discriminerend vermogen hebben om lv-PPA patiënten en de nfv-PPA patiënten van elkaar te onderscheiden.

Resultaten doel 2

Om te onderzoeken of er specifieke zinstypen of items waren waar controles geen maximale score op behaalden, is er allereerst gekeken naar de behaalde scores op ieder item van de SynTest (strategie 1). In figuur 4 zijn zeven diagrammen met de scores van controle participanten gepresenteerd waarbij elke grafiek de behaalde SynTest scores van één categorie representeert. In de diagrammen is voor ieder item (1-35) af te lezen hoeveel participanten de scores 0, 1 en 2 behaalden. Dit overzicht geeft per item én

per categorie inzicht in of en hoe vaak er een andere score dan respectievelijk de maximale score 2 en de maximale categoriescore 10 is behaald.



Figuur 4. Frequentie van behaalde scores door controle participanten. Per SynTest categorie zijn de scores voor alle 35 items weergegeven, waarbij de x-as de drie mogelijke scores (0, 1 en 2) per item representeert en de y-as het aantal controle participanten dat de scores 0, 1 en 2 heeft behaald.

In het staafdiagram dat de categorie 5 ‘een uitvoerder twee acties’ weergeeft, is te zien dat er op items 21-25 uitsluitend de score van 2 is behaald en dat de controles allemaal daarmee een maximale score van 10 voor de categorie hebben behaald. De items die behoren tot deze categorie hebben voor de controles geen incorrecte antwoorden opgeleverd. In de diagrammen van categorieën 3 ‘twee uitvoerders actief’ en 6 ‘subject relatieve bijzin’ is te zien dat er naast het groot aantal behaalde scores 2, bij enkele items de score 1 is behaald. Op item 11, 12 en 14 (categorie 3) is telkens slechts eenmaal de score 1 behaald en op item 26 (categorie 6) is tweemaal. Voor categorie 2 ‘regelmatig actief’ en 4 ‘passief’ geldt dat er naast de score 2 zowel enkele keren de scores 1 en 0 zijn behaald. In het diagram van categorie 2 is te zien dat er op de items 6, 7 en 9 respectievelijk 3, 1 en 3 keer de score 1 is behaald en op item 9 daarnaast ook eenmaal de score 0. Wanneer het diagram van categorie 1 ‘intransitief’ in beschouwing wordt genomen, is te zien dat item 1 slechts eenmaal de score 1 en 0 opleverde, maar item 3 leverde maar liefst 6 keer de score 1 en 6 keer de score 0 op. Omdat er op item 3 met de daartoe behorende zin ‘Het meisje ligt’ in totaal 12 keer een andere score dan de maximale score 2 is behaald, is er in het geval van een incorrect antwoord of zelfcorrectie gekeken naar het type fout dat werd gemaakt. Bij 4 van de 12 antwoorden was het type fout niet bekend, maar van de andere 8 antwoorden was het type fout gekenmerkt met ‘role’ wat inhoudt dat er 8 keer voor de afbeelding is gekozen waar een ander karakter (de jongen in plaats van het meisje) de handeling uitvoert. Tot slot is in het diagram van de laatste categorie, categorie 7 ‘relatieve bijzin met eind van de zin oplossing’, af te lezen dat er

op ieder item minimaal 8 (item 35) tot maximaal 23 keer een score anders dan 2 is behaald. In totaal is er binnen deze categorie 69 keer de score 0 en 8 keer de score 1 behaald. Vanwege het grote aantal antwoorden die een score 1 of 0 opleverde, is er voor al deze items naar het type fout gekeken in het geval van een incorrect antwoord of zelfcorrectie. Van 5 antwoorden was het type fout gekenmerkt met 'verb', van 67 antwoorden met 'role' en van 4 antwoorden met 'role-verb'. Daarnaast was van 3 antwoorden het type fout niet bekend.

Naast het visualiseren van de behaalde scores, is volgens strategie 2 de Wilcoxon toets uitgevoerd. Uit deze toets bleek dat de mediaan van de behaalde categoriescores van de categorieën 1, 2, 4 en 7 significant afweek van de maximaal haalbare categoriescore 10 ($p < .05$). Echter, wanneer van alle categorieën de mediaan werd bepaald, bleken de medianen van alle categorieën – met uitzondering van categorie 7 (Mdn = 7) – een waarde van 10 te hebben. Deze bevindingen leken niet met elkaar in overeenstemming. De inspectie van de staafdiagrammen met betrekking tot strategie 1 wees uit dat enkel categorie 7 problematisch leek voor controles, en vanwege de beperkte variatie in sommige categorieën bleek de statistische strategie 2 geen behulpzame aanpak. De inspectie van de plots in figuur 4 is voldoende om te kunnen concluderen dat categorie 7 voor controles problematisch was.

Discussie

In deze studie is primair onderzocht of de niet-vloeiende PPA (nfv-PPA) en logopenische PPA (lv-PPA) varianten beter van elkaar kunnen worden onderscheiden op basis van tekorten in de syntactische verwerking. De resultaten lieten met een classificatie accuratesse van 0.48 zien dat nog geen 50% van de patiënten op basis van de SynTest (en BBTK) correct geclassificeerd kon worden. Daarmee bleek tegen de verwachting in dat de SynTest onvoldoende discriminatiekracht heeft om de twee PPA varianten van elkaar te onderscheiden. Ten behoeve van het optimaliseren van de SynTest als zinsbegripstest is daarnaast onderzocht hoe cognitief gezonde volwassenen op de SynTest presteren, waarbij per item en categorie specifiek is gekeken naar de frequentie waarmee de suboptimale scores 0 en 1 zijn behaald. Controle participanten scoorden op item 3 en categorie 7 met het zinstype 'relatieve bijzin met eind van de zin oplossing' aanzienlijk minder goed dan op de andere items en categorieën, met scores die zelfs voor cognitief gezonde volwassenen problematisch te noemen zijn.

Een belangrijk aspect dat betrekking heeft op PPA in het algemeen en de subtypen binnen PPA, is dat het testmateriaal dat in de diagnostiek van PPA wordt gebruikt, vaak niet specifiek is ontwikkeld voor PPA maar veelal afkomstig is uit het CVA-veld (Henry & Grasso, 2018; Jiskoot et al., 2020). Dit testmateriaal blijkt niet voldoende toereikend om het onderscheid tussen de PPA varianten te maken, mede omdat het niet met dat specifieke doel en voor die patiëntgroep is ontwikkeld. Bij patiënten uit het CVA-veld is er sprake van een relatief plotseling ontstaan van stoornissen vanwege de acute aard van een ischemische of hemorragisch CVA. In de diagnostiek is het voornamelijk van belang om alle aangedane functies zo nauwkeurig mogelijk in kaart te brengen, om een gericht behandelplan op te maken ten behoeve van optimaal herstel. Bij PPA is daarentegen sprake van een geleidelijk ontstaan van dysfunctie. Vanwege de progressieve aard van PPA nemen deze functionele tekorten gedurende het ziekteproces toe in ernst en frequentie. De focus van therapie ligt, anders dan bij CVA, niet op herstel maar juist op het zo lang mogelijk behouden van functies. Vanwege de significante verschillen tussen deze twee patiëntgroepen op het gebied van het ontstaan en verloop van de ziekte waarbij nog niet eens is ingegaan op de neuropathologie en aanwezigheid van tekorten in de niet-talige cognitieve functies, kunnen er vraagtekens worden gezet bij de toepasbaarheid van het CVA testmateriaal in diagnostiek van PPA. Vermoedelijk is het testmateriaal niet voldoende sensitief om kleine verschillen in de vroegste fase van PPA te detecteren (Henry & Grasso, 2018) en is (een gedeelte van) het testmateriaal niet geschikt voor de diagnostiek van PPA. Men zou zich zelfs kunnen afvragen of de niet-vloeiende en logopenische PPA variant moeilijk van elkaar te onderscheiden zijn, *omdat* CVA testmateriaal toegepast op de PPA diagnostiek. Sensitiever testmateriaal dat specifiek voor de PPA doelgroep is ontwikkeld zou mogelijk kunnen bijdragen aan de oplossing voor het huidige classificatieprobleem van PPA. De Syntest neemt hier een vooraanstaande plaats in en is daarmee een belangrijke ontwikkeling ten opzichte van een groot gedeelte van de testen die nu in Nederland voor PPA diagnostiek worden gebruikt.

Doel 1 Maximalisatie van het onderscheid tussen nfv-PPA en lv-PPA

Ondanks dat de verwachting was dat met de SynTest het verschil in onderliggende mechanismen van de tekorten in de syntactische verwerking tussen nfv-PPA en lv-PPA kwantitatief aan te tonen, bleek de SynTest in het huidige onderzoek niet in staat te discrimineren tussen nfv-PPA en lv-PPA patiënten. De tijdsgevoeligheid van de diagnostiek van PPA die door Jiskoot en collega's (2020) wordt benadrukt zou hiervoor mogelijk een verklaring kunnen bieden (Jiskoot et al., 2020; Mesulam et al., 2014). Het zou betekenen dat indien onderzoek naar tekorten in de syntactische verwerking te vroeg wordt verricht, de stoornis nog niet voldoende is ontwikkeld om verschillen tussen de twee subtypen te kunnen detecteren. Maar wanneer het onderzoek echter te laat plaatsvindt, is de PPA al te ver gevorderd naar een globale dementie. Daardoor kunnen er ook stoornissen in de niet-talige cognitieve functies aanwezig zijn die het onmogelijk maken om op basis van tekorten in de syntactische verwerking nog een onderscheid te

maken tussen de PPA subtypen. Dat zou betekenen dat de sensitiviteit en specificiteit van de SynTest tijdsafhankelijk is, en deze mogelijk wel sensitief zou kunnen zijn om verschillen in tekorten in syntactische verwerking te kwantificeren indien afgenomen in een specifieke periode in het ziekteproces. Het huidige onderzoek is echter ontoereikend om hierover een uitspraak te kunnen doen, omdat er voorafgaand aan het onderzoek geen rekening kon worden gehouden met hoe lang ziektesymptomen bij iedere patiënt al aanwezig waren. Gegevens met betrekking tot ziekteduur zijn enkel gebruikt om te testen voor groepsverschillen. Vervolgonderzoek is wenselijk om te leren of de SynTest in een kritieke periode in het ziekteproces mogelijk wél discriminatiekracht heeft.

Een andere mogelijke verklaring voor het gebrek aan onderscheidend vermogen tussen de PPA subtypen is dat de differentiële diagnose tussen nfv-PPA en lv-PPA niet gemaakt kan worden op basis van één enkele test. Een combinatie van meerdere onderscheidende kenmerken zou discriminatie tussen de twee PPA varianten mogelijk kunnen maken. Met deze aanpak liet Janssen (2020) bijvoorbeeld zien dat semantische PPA (sv-PPA) op basis van de prestaties op de semantische associatie- en benoemtaak van de SYDBAT-NL met een hoge sensitiviteit en specificiteit van de andere twee PPA varianten onderscheiden kon worden. De subtesten van de SYDBAT-NL die op woordniveau testen bleken echter niet voldoende toereikend om accuraat te differentiëren tussen nfv-PPA en lv-PPA patiënten. In het huidige onderzoek bleek dat de SynTest als afzonderlijke test ook onvoldoende discriminatiekracht heeft. Eenzelfde aanpak als Janssen (2020), waarbij middels integratie van data van meerdere specifieke kenmerken een mogelijke route wordt gecreëerd voor toekomstig onderzoek naar de differentiatie tussen nfv-PPA en lv-PPA. Hiervoor zijn er verschillende mogelijkheden voor kenmerken met een potentiële toegevoegde waarde.

Ten eerste, problemen in zinnen herhalen hebben vanwege het verschil in onderliggend mechanisme en uitingsvormen de potentie om het onderscheid tussen de twee PPA varianten te maximaliseren. Bij lv-PPA worden problemen met zinnen herhalen geassocieerd met een stoornis in het fonologisch kortetermijngeheugen (Gorno-Tempini et al., 2008; Grossman, 2018) en deze uiten zich voornamelijk in omissies en fonologische parafasieën (Hohlbaum et al., 2018). Een stoornis in het herhalen van zinnen is geen hoofd- of ondersteunend kenmerk van nfv-PPA (Gorno-Tempini et al., 2011), maar de aanwezigheid van spraakapraxie, dat wél een hoofdkenmerk van nfv-PPA is, heeft ook effect op de vloeiendheid van de taalproductie en kan daarmee tevens de vloeiendheid van het herhalen van zinnen beïnvloeden. Guiu et al. (2020) lieten zien dat het toevoegen van prestaties op een zinsherhaaltaak aan prosodische parameters resulteerde in een hoger percentage correcte classificaties (96.2% versus 71.3%) in de categorieën nfv-PPA, lv-PPA en gezonde controles. Wanneer de diagnoses echter enkel op basis van de prestaties op de zinsherhaaltaak werden voorspeld, resulteerde dat slechts in een percentage van 66.2% correcte classificaties. Dit onderstreept opnieuw het belang van het combineren van onderscheidende kenmerken in de PPA diagnostiek.

Ten tweede, semi-spontane taalproductie maten als gevolg van een kwantitatieve analyse van de (semi-)spontane taalproductie zouden ook kunnen bijdragen aan het onderscheid tussen nfv-PPA en lv-PPA. In de masterscriptie van Wets (2020) is voor Nederlandstalige personen met PPA onderzocht wat op lexicaal en grammaticaal niveau typerend is voor de semi-spontane taalproductie met betrekking tot de drie PPA subtypen. Vanwege de kleine steekproefgrootte zijn voor de lexicale en morfo-syntactische variabelen geen significante verschillen tussen de subtypen gevonden, maar kwamen er wel bepaalde trends aan het licht. Het aantal lexicale werkwoorden per 100 woorden, het aantal koppelwerkwoorden per 100 woorden, de gemiddelde uitingslengte en de proportie finiete bijzinnen lijken veelbelovende variabelen om de nfv-PPA en lv-PPA te typeren. Herhaalonderzoek is wenselijk om te onderzoeken of deze trends op grotere schaal daadwerkelijk significante verschillen tussen de twee varianten opleveren. Indien er significante verschillen worden gevonden, zouden de meest typerende semi-spontane taalproductie maten als onderscheidende kenmerken kunnen fungeren in de differentiaaldiagnostiek van nfv-PPA en lv-PPA. Omdat analyse van de (semi-)spontane taalproductie

relatief arbeidsintensief is, dient de praktische haalbaarheid voor succesvolle implementatie in de klinische praktijk echter in acht genomen te worden.

Ten derde, naast de talige kenmerken hebben stoornissen in het geheugen een potentieel toegevoegde waarde in de differentiaaldiagnostiek van nfv-PPA en lv-PPA. Specifiek voor de nfv-PPA en lv-PPA varianten geldt namelijk dat de mate waarin functies met betrekking tot het geheugen zijn aangedaan verschillend is. Zo lieten Piguet et al. (2015) zien dat lv-PPA patiënten significant slechter presteerden op taken met betrekking tot het episodisch geheugen ten opzichte van nfv-PPA patiënten. Ramanam et al. (2016) vonden verschillen tussen de nfv-PPA en lv-PPA varianten op het gebied van het non-verbale geheugen. Op basis van de prestaties op een non-verbale geheugentaak (Rey Complex Figure delayed recall) kon 81% van de PPA patiënten correct geclassificeerd worden en dat bleek daarmee een goede maat om de twee PPA varianten te discrimineren. Eikelboom en collega's (2018) toonden aan dat lv-PPA patiënten significant slechter presteerden op zowel verbale als non-verbale geheugentaken ten opzichte van nfv-PPA patiënten en concludeerden dat geheugenstoornissen meer aanwezig waren bij lv-PPA vergeleken met nfv-PPA. De verschillen met betrekking tot geheugenstoornissen tussen de twee PPA varianten kunnen verklaard worden vanuit verschillen in neuropathologie die aan de PPA subtypen ten grondslag liggen. De nfv-PPA variant is gerelateerd aan het frontotemporale lobaire degeneratie (FTLD) spectrum (FTLD-tau pathologie; Grossman, 2012), terwijl de neuropathologie van lv-PPA patiënten in de meeste gevallen gerelateerde is aan de ziekte van Alzheimer (Gorno-Tempini et al., 2004; Spinelli et al., 2017). De verschillen in geheugenprofiel als gevolg van een ander onderliggende neuropathologie hebben tevens potentie om als onderscheidend kenmerk in de differentiaaldiagnostiek van PPA te fungeren. Een keerzijde is dat testprestaties binnen niet-talige cognitieve domeinen negatief beïnvloed kunnen worden door de aanwezige taalproblemen bij PPA en deze resultaten kunnen vertekenen (Jiskoot et al., 2020; Machulda et al., 2013).

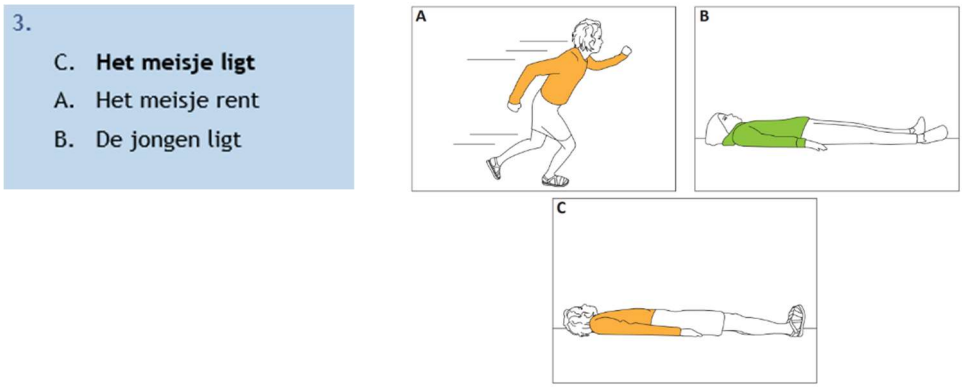
Er bestaat discussie over de huidige diagnostische criteria van PPA, onder meer over de kwestie dat de verschillende uitingsvormen van PPA niet zo discreet te onderscheiden zijn in de drie subtypen (semantische PPA, niet-vloeiende PPA, logopenische PPA). Vandenberghe (2016) stelt als gevolg van het relatief hoge percentage PPA patiënten voor wie classificatie naar subtype niet mogelijk is, een vierde PPA subtype voor: het gemixte subtype ('mixed variant'). Het gemixte subtype laat volgens Vandenberghe (2016) zowel woordbegripsproblemen als spraakapraxie of agrammatisme zien, een combinatie van kenmerken die niet voorkomt in de huidige classificatie die slechts drie PPA varianten omvat. Het gemixte subtype zou hierdoor ruimte moeten overlaten voor die patiënten die aan de criteria voor de basisdiagnose PPA voldoen, maar vervolgens niet naar een van de drie subtypen geclassificeerd kunnen worden. Ook wordt door Mesulam en collega's (2014) in hun benadering voor classificatie van PPA een gemixte PPA variant ('mixed PPA') voorgesteld. Deze gemixte PPA variant wordt volgens hen gekenmerkt door een combinatie van stoornissen op het gebied van begrip en grammatica die in de vroege stadia van de ziekte van ongeveer gelijke ernst zijn. Classificatie van PPA is volgens de benadering van Mesulam en collega's (2014) echter tijdsgevoelig. Dat zou inhouden dat indien deze te vroeg toegepast zou worden, patiënten die zich in de prodromale fase van niet-vloeiende PPA (in deze benadering 'agrammatic PPA') of semantische PPA bevinden geclassificeerd worden als logopenische PPA. Indien de classificatie volgens deze benadering te laat in het ziekteverloop wordt toegepast, zullen patiënten als de gemixte PPA variant geclassificeerd worden als gevolg van de toegenomen ernst van de afasie. Andere onderzoeksgroepen suggereren dat PPA zelfs beter als een continuümziekte beschouwd kan worden, waarbij een patiënt met een bepaalde uitingsvorm van PPA ook niet zo discreet onder het gemixte subtype geplaatst kan worden, maar zich ergens op het spectrum van vloeiend naar niet-vloeiende taalproductie bevindt (Jiskoot et al., 2020). Vanuit theoretisch oogpunt zou deze continuümbenadering meer recht doen aan de variatie binnen het ziektebeeld van PPA. Wanneer deze benadering in overweging wordt genomen, zou men zich ook kunnen afvragen of een dergelijk continuüm ook betrekking heeft op de syntactische verwerking. Indien dit het geval zou zijn, zou de SynTest mogelijk helpen om PPA patiënten die niet naar een van de drie subtypen kunnen worden geclassificeerd, op basis van syntactische tekorten ergens op het continuüm te plaatsen. Anderzijds is

classificatie voor de klinische praktijk wellicht meer wenselijk (Henry & Grasso, 2018) en verdient verder onderzoek naar mogelijkheden tot en optimaliseren van classificatie de voorkeur. Het introduceren van een 'mixed type' lijkt ook niet de volledige oplossing, omdat dit gemixte PPA subtype wellicht meestal 'mixed' is omdat er nog niet voldoende sensitief kan worden getest. Mogelijk zouden er op basis van een test of een combinatie van testen zoals de SynTest die specifiek voor PPA is ontwikkeld, alsnog twee groepen binnen dit gemixte type onderscheiden kunnen worden. In afwezigheid van 'mixed subtype' PPA patiënten biedt het huidige onderzoek hier geen inzicht in.

Doel 2 Optimalisatie van de SynTest

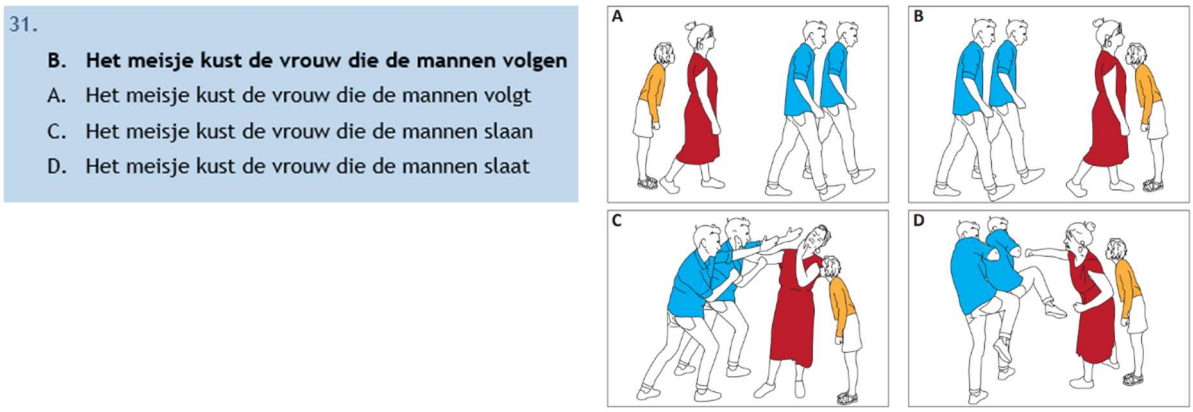
Controles bleken opvallend vaak een incorrect antwoord te geven op item 3 en op alle items van categorie 7 vergeleken met de andere items van de SynTest. Daarmee zijn deze items zelfs voor cognitief gezonde volwassenen problematisch. Omdat tekorten in de syntactische verwerking bij cognitief gezonde volwassenen uitgesloten zijn, mag van deze groep verwacht worden op alle items in de regel plafondscores te behalen. Dit laatste was echter niet van toepassing op item 3 en de items van categorie 7, waaruit geconcludeerd kan worden dat de incorrecte antwoorden worden geprovoceerd door de items zelf.

Item 3 wordt als problematisch beschouwd, omdat er op dat item in totaal zesmaal een score 1 en zesmaal een score 0 is behaald. Omdat dit het derde testitem van de SynTest is en zich daarmee aan het begin van de SynTest bevindt, bestaat de mogelijkheid dat de participanten nog niet volledig begrepen wat er van hen werd verwacht en nog moesten 'wennen' aan de test. Dit lijkt echter onwaarschijnlijk, allereerst omdat je zou verwachten dat testitems 1 en 2 een ongeveer gelijk aantal scores 0 en 1 opgeleverd zouden moeten hebben omdat deze items zich nog vóór item 3 in de SynTest bevinden. Daarnaast bevat de SynTest twee oefenitems die aan de testitems voorafgaan en die indien nodig toegelicht mogen worden. Het lijkt ook niet aannemelijk dat de categorie te moeilijk is, vanwege het feit dat item 3 zich in de eerste en daarmee de minst complexe categorie ten opzichte van de andere 6 categorieën bevindt. Verder is er met uitzondering van item 1 op geen van de andere items uit diezelfde categorie een score 1 of 0 werd behaald. Dit doet vermoeden dat specifiek item 3 in zijn huidige vorm geen goed item is. In de resultatensectie is beschreven dat bij alle incorrecte antwoorden waarvan informatie bekend was, was gekozen voor de afbeelding waarin de actie werd uitgevoerd door een ander karakter (gekenmerkt met 'role'). Voor de zin 'Het meisje ligt', werd als antwoord de (incorrecte) afbeelding waarop de jongen liggend werd afgebeeld gekozen. Dit suggereert dat het onderscheid tussen de karakters 'jongen' en 'meisje' op de afbeeldingen van item 3 onvoldoende duidelijk was voor de controle participanten, mogelijk omdat de uiterlijke kenmerken van de karakters minder duidelijk worden wanneer de karakters liggend zijn afgebeeld (Fig. 5). Vermoedelijk werd de jongen hierdoor (te) vaak voor het meisje aangezien met incorrecte antwoorden als gevolg. Mijn suggestie zou dan ook zijn om het onderscheid jongen versus meisje in ieder geval in de afbeeldingen van item 3 te verduidelijken en het item daarmee te verbeteren.



Figuur 5. Item 3 van SynTest. Doelzin (dikgedrukt) wordt auditief aangeboden en moet correct worden geassocieerd met de bijbehorende afbeelding C

Ook bleken de items die tot categorie 7 behoren problematisch voor de controles. Ter illustratie is een item van categorie 7 weergegeven (Fig. 6). Voor de vijf items (31 t/m. 35) gold dat er aanzienlijk meer scores 0 en 1 zijn behaald vergeleken met de andere items van de SynTest. Omdat dit niet op slechts één of twee maar op alle vijf de items van de categorie betrekking had, waarvan de doelzinnen allen eenzelfde zinsstructuur hadden, doet dit sterk vermoeden dat het grote aantal incorrecte antwoorden veroorzaakt wordt door het zinstype van categorie 7. Het zinstype van categorie 7 bevat een relatieve bijzin waarvan de oplossing van de zin pas aan het eind van de zin duidelijk wordt. Dit is tevens het meest complexe zinstype van de SynTest. Uit deze studie bleek dat het zinstype van categorie 7 zelfs voor cognitief gezonde volwassenen te complex was. Om die reden is mijn suggestie om categorie 7 geheel te excluderen van de SynTest. Immers, wanneer de SynTest voor een bredere doelgroep dan enkel voor PPA zal worden ingezet, is discrimineren tussen PPA varianten op basis van tekorten in de syntactische verwerking niet meer het doel. Het doel is dan om te onderzoeken of er bij een patiënt sprake is van zinsbegripsproblemen. Daarmee kan dan worden gediscrimineerd tussen personen met en zonder tekorten in de syntactische verwerking. Categorie 7 lijkt hierin nauwelijks tot niet informatief, omdat personen zonder tekorten tevens een grote kans hebben op incorrecte antwoorden binnen deze categorie, en de items in deze categorie daarom niet bijdragen aan de discriminatie tussen personen met en zonder tekorten in syntactische verwerking. Om deze reden verdient exclusie van deze categorie vooralsnog de voorkeur. Daarentegen, een leereffect binnen deze categorie bij een gezonde populatie zou een argument zijn voor het behouden van deze categorie in de SynTest, indien een dergelijk leereffect niet zou worden gevonden bij de patiëntpopulatie mét zinsbegripsproblemen. De twee groepen zouden dan op basis van dit criterium in categorie 7 alsnog van elkaar onderscheiden kunnen worden. Nader onderzoek is echter benodigd om deze hypothese te onderbouwen. Omdat categorie 7 zelfs voor cognitief gezonde personen dus problematisch bleek, bestaat de mogelijkheid dat de inclusie van deze categorie de resultaten met betrekking tot doel 1 heeft beïnvloed. Om afsluitend te verifiëren of de SynTest zonder categorie 7 de nfv-PPA en lv-PPA varianten al dan niet beter van elkaar kon onderscheiden, is de LDA eenmaal herhaald zonder categorie 7. De classificatie accuratesse bleef met 0.48 gelijk. Exclusie van categorie 7 resulteert dus niet in een beter onderscheid tussen de twee PPA varianten.



Figuur 6. Voorbeeld van SynTest item uit categorie 7. Doelzin (dikgedrukt) wordt auditief aangeboden en moet correct geassocieerd worden met de bijbehorende afbeelding B.

Tekortkomingen

Het huidige onderzoek kent een aantal tekortkomingen. Allereerst zijn de resultaten met betrekking tot de discriminatie tussen de nfv-PPA en lv-PPA varianten gebaseerd op een relatief kleine steekproef (n=27). Dit had gevolgen voor de manier waarop de LDA is uitgevoerd. Wanneer een steekproefgrootte voldoende groot is, heeft het de voorkeur bij een LDA de data te verdelen in een training set en een test set. Eerst wordt er met de data uit de training set een model gebouwd op basis van de training set. Vervolgens wordt de data uit de test set gebruikt om de nauwkeurigheid van het model te testen. Idealiter zou het model op nieuwe data toegepast worden, om een model te bouwen dat verder generaliseert dan de eigen data. Vanwege de relatief kleine steekproef is er in het huidige onderzoek echter voor gekozen om de data niet in twee data sets te verdelen, maar om op basis van de beschikbare data middels cross-validatie een zo accuraat mogelijk model te bouwen. Een keerzijde hiervan is dat er niet is getest hoe het model de (differentiaal)diagnoses van nieuwe data voorspelt.

Voor de data van de SynTest geldt dat de accuratesse codes zijn omgezet tot scores om een LDA uit te kunnen voeren. Zes accuratesse codes zijn hiertoe gereduceerd tot 3 scores; correct, incorrect, correct na herhaling, zelfcorrectie of tijd. Hierdoor is er in deze omzetting potentieel waardevolle informatie verloren gegaan. Dit heeft betrekking op de differentiatie tussen zelfcorrectie, herhaling door de onderzoeker, een combinatie van deze twee en een reactietijd van langer dan 5 seconden bij in het geval van een score 1.

Daarnaast bestonden er verschillen in de aangeleverde data van het Erasmus MC en de data die vanuit het onderzoek van Janssen (2020) was verkregen. Op van het Erasmus MC verkregen ruwe scoreformulieren van de Boston Benoemtaak (BBT) bestond de manier van scoren uit correct versus incorrect en was er bij de meeste patiënten de Mini-Mental State Examination (MMSE) afgenomen. De data afkomstig uit het onderzoek van Janssen (2020) bestonden uit scores van de Montreal Cognitive Assessment (MoCA) en uit de verkorte versie van de BBT (BBTK) waarbij het scoringssysteem van de BBT was gehanteerd. Om deze data met elkaar te kunnen verenigen, was het noodzakelijk om bewerkingen van de data te doen. Deze omzettingen hebben echter niet de voorkeur, omdat door deze bewerkingen de gegevens minder nauwkeurige benaderingen zijn van de daadwerkelijke gegevens. Dit pleit voor meer uniformiteit in keuze voor testmateriaal en manier van scoren in de diagnostiek van PPA maar dat geldt voor diagnostiek in het algemeen.

Patiënten bij wie de MoCA was afgenomen scoorden allen onder de cutoff waarde (26/30), terwijl het overgrote deel van de patiënten bij wie de MMSE is afgenomen een score gelijk aan of boven de cutoff waarde (23/30) behaalden. Dit suggereert dat de MoCA sensitiever is in het detecteren van zeer milde beperkingen in het cognitief functioneren ten opzichte van de MMSE. Verschillende studies ondersteunen dat de MoCA een sensitiever meetinstrument is om milde cognitieve beperkingen (mild cognitive impairment; MCI) te detecteren vergeleken met de MMSE (Ciesielska et al., 2016; Markwick et al., 2012; Nasreddine et al., 2005). Voor het huidige onderzoek houdt dat in dat het samenvoegen van MMSE en MoCA data op basis van de cutoff suboptimaal was. Met het oog op diagnostiek van PPA, en daarnaast diagnostiek van aandoeningen die samengaan met milde cognitieve beperkingen (MCI), zou de MoCA verkozen kunnen worden boven de MMSE.

Implicaties voor vervolgonderzoek

Als vervolg op het huidige onderzoek zou het interessant zijn om te onderzoeken of de tijdsgevoeligheid van de diagnostiek van PPA tevens betrekking heeft op het detecteren van verschillen in tekorten in syntactische verwerking tussen de nfv-PPA en lv-PPA varianten. Specifiek zou onderzocht kunnen worden of de gevoeligheid en de specificiteit van de SynTest tijdsafhankelijk zijn en de nvf-PPA en lv-PPA varianten in een specifieke periode in het ziekteproces wel te onderscheiden zijn op basis van tekorten in syntactische verwerking. De SynTest zou hiertoe bij een zo groot mogelijke groep nfv-PPA en lv-PPA patiënten op meerdere meetmomenten afgenomen moeten worden, waarbij de patiënten allemaal een ongeveer gelijke ziekteduur dienen te hebben. Een andere mogelijkheid zou zijn om een grote groep nfv-PPA en lv-PPA patiënten op basis van ziekteduur in meerdere groepen te verdelen en te onderzoeken of de SynTest voor een of meer van de ‘gemaakte’ groepen wel voldoende discriminatiekracht heeft om de nfv-PPA en lv-PPA varianten met een relatief hoge gevoeligheid en specificiteit van elkaar te onderscheiden. Vanuit theoretisch oogpunt zou dit potentieel waardevolle informatie kunnen opleveren. De praktische haalbaarheid van een dergelijk vervolgonderzoek is echter discutabel, omdat het accuraat bepalen van de ziekteduur bij een neurodegeneratieve aandoening als PPA – wat echter cruciaal is voor de voorgestelde onderzoeksdesigns – niet gemakkelijk is. De praktische haalbaarheid én de lastige klinische implementatie van eventuele uitkomsten van dergelijk vervolgonderzoek zullen een uitdaging zijn.

Daarnaast zou het in vervolgonderzoek de moeite waard zijn om een cutoff waarde voor de SynTest te bepalen die bepaalt wanneer een SynTest prestatie als pathologisch kan worden beschouwd. Een dergelijke cutoff kan tevens mogelijk helpen personen met tekorten in syntactische verwerking te discrimineren van personen zonder dergelijke tekorten. Om een cutoff score van de SynTest te bepalen waar het onderscheid tussen deze groepen maximaal is, zijn SynTest prestaties van een grote groep cognitief gezonde volwassenen en van een grote groep patiënten met niet-aangeboren hersenletsel en zinsbegripsproblemen cruciaal. Ook zou op basis van de SynTest prestaties van deze patiënt- en controlegroep kunnen worden onderzocht of er sprake is van een leereffect binnen categorie 7 ten behoeve van een nauwkeurigere besluitvorming om categorie 7 al dan niet van de SynTest te excluderen.

Verder wordt om een reactietijd in de SynTest als afwijkend te scoren een drempelwaarde van 5 seconden gehanteerd. Deze drempelwaarde is echter overgenomen van de CAT-NL waarin over de gehele testbatterij (indien reactietijden zijn meegenomen in de scoring) een drempelwaarde van 5 wordt gehanteerd. Wellicht is de optimale drempelwaarde van reactietijd voor de SynTest wel anders dan voor de CAT-NL en zou het de moeite waard zijn om dat voor de SynTest te onderzoeken.

Een integrale benadering waarbij de SynTest wordt gecombineerd met data van andere karakteristieken geeft misschien nog het meeste perspectief. Hoewel tijdsintensief, zou er in de toekomst nader kunnen worden onderzocht of ten eerste een kwantitatieve analyse van de (semi-)spontane taalproductie bijdraagt aan het maximaliseren van het onderscheid tussen de niet-vloeiende en

logopenische PPA varianten. Ook het herhalen van zinnen waarbij specifiek wordt gekeken naar de typen fouten die gemaakt worden hier mogelijk aan bijdragen. Ten slotte zouden stoornissen in de niet-talige cognitieve functies als geheugen en ook het executief functioneren onderwerpen voor verdere studie zijn. Integraal gebruik van data van deze verschillende kenmerken en testen verdient daarom in overweging genomen te worden.

Diagnostiek van PPA in het algemeen, en met name de differentiatie tussen de niet-vloeiende en logopenische variant, is zeer complex. Het onderscheid tussen de twee PPA varianten wordt onder meer bemoeilijkt door de overlap in linguïstische kenmerken tussen beiden en door het gebrek aan voldoende sensitief diagnostisch materiaal dat specifiek voor de PPA patiëntpopulatie is ontwikkeld en genormeerd. Met het huidige onderzoek is getracht het onderscheid tussen nfv-PPA en lv-PPA te maximaliseren op basis van syntactische tekorten gemeten middels SynTest, een test die specifiek voor PPA is ontwikkeld. De resultaten wezen echter uit dat de SynTest als op zichzelf staande test onvoldoende discriminatiekracht heeft om de twee PPA varianten van elkaar te onderscheiden. Een integrale karakterisering waarbij talige en niet-talige onderscheidende kenmerken in de differentiaaldiagnostiek worden gecombineerd, heeft de potentie om het onderscheid tussen nfv-PPA en lv-PPA te verbeteren. Indien de SynTest niet als afzonderlijke test wordt ingezet, maar deel zou uit gaan maken van een dergelijke integrale aanpak, zou het mogelijk bijdragen aan de differentiatie tussen nfv-PPA en lv-PPA. Desondanks zou er idealiter sensitiever testmateriaal specifiek voor de PPA patiëntgroep worden ontwikkeld en genormeerd, en zouden ook dat deel uit gaan maken van de integrale karakterisering van PPA.

De SynTest is daarnaast een potentieel waardevol meetinstrument om breder ingezet worden voor diagnostiek van zinsbegripsproblemen. Voor een optimale toepasbaarheid en implementatie in de klinische praktijk, zou het aan te bevelen zijn suggesties voor verbetering en exclusie van testitems en suggesties voor vervolgonderzoek die naar aanleiding van het huidige onderzoek zijn gedaan in acht te nemen.

Referenties

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. D., & Hitch, G. J. (2019). The phonological loop as a buffer store: An update. *Cortex*, *112*, 91–106. https://doi.org/10.1016/j.cortex.2018.05.015

Bastiaanse, R., Maas, E., & Rispens, J. E. (2000). *WEZT: Werkwoorden- en Zinnen Test*. Swets Test Publishers.

Ciesielska, N., Sokołowski, R., Mazur, E., Podhorecka, M., Polak-Szabela, A., & Kędziora-Kornatowska, K. (2016). Czy test Montreal Cognitive Assessment (MoCA) moze być skuteczniejszy od powszechnie stosowanego Mini-Mental State Examination (MMSE) w wykrywaniu łagodnych zaburzeń funkcji poznawczych u osób po 60. roku zycia? Metaanaliza. *Psychiatria Polska*, *50*(5), 1039–1052. https://doi.org/10.12740/PP/45368

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. https://doi.org/10.1017/S0140525X01003922

Croot, K., Ballard, K., Leyton, C. E., & Hodges, J. R. (2012). Apraxia of speech and phonological errors in the diagnosis of nonfluent/agrammatic and logopenic variants of primary progressive aphasia. *Journal of Speech, Language, and Hearing Research*, *55*(5). https://doi.org/10.1044/1092-4388(2012/11-0323)

Dautzenberg, P. L. J., & de Jonghe, J. F. M. (2004). *Montreal Cognitive Assesment Intructions (nederlands)*. 1–4.

De Renzi, E., & Faglioni, P. (1978). Normative Data and Screening Power of a Shortened Version of the Token Test. *Cortex*, *14*(1), 41–49.

Dyson, L., Morgan, J., & Herbert, R. (2021). Semantic processing in aphasia: evidence from semantic priming and semantic interference. *Language, Cognition and Neuroscience*, *36*(4), 491–516. https://doi.org/10.1080/23273798.2020.1849749

Eikelboom, W. S., Janssen, N., Berg, E., Roelofs, A., & Kessels, R. (2017). Differentiatie van primair-progressieve afasie varianten: De Nederlandse bewerking van de Sydney Language Battery (SYDBAT-NL). *Tijdschrift Voor Neuropsychologie*, *12*, 189–202. https://repository.ubn.ru.nl/handle/2066/178961

Eikelboom, W. S., Janssen, N., Jiskoot, L. C., van den Berg, E., Roelofs, A., & Kessels, R. P. C. (2018). Episodic and working memory function in Primary Progressive Aphasia: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, *92*(May), 243–254. https://doi.org/10.1016/j.neubiorev.2018.06.015

Gorno-Tempini, M. L., Brambati, S. M., Ginex, V., Ogar, J., Dronkers, N. F., Marcone, A., Perani, D., Garibotto, V., Cappa, S. F., & Miller, B. L. (2008). The logopenic/phonological variant of primary progressive aphasia. *Neurology*, *71*(16), 1227–1234.

Gorno-Tempini, M. L., Dronkers, N. F., Rankin, K. P., Ogar, J. M., Phengrasamy, L., Rosen, H. J., Johnson, J. K., Weiner, M. W., & Miller, B. L. (2004). Cognition and Amatomy in Three Variants of Primany Progressive Aphasia. *Annals of Neurology*, *55*(3), 335–346. https://doi.org/10.1002/ana.10825

Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F., Manes, F., Dronkers, N. F., Vandenberghe, R., Rascovsky, K., Patterson, K., Miller, B. L., Knopman, D. S., Hodges, J. R., Mesulam, M. M., &

Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, *76*(11), 1006–1014. https://doi.org/10.1212/WNL.0b013e31821103e6

Grossman, M. (2012). The non-fluent/agrammatic variant of primary progressive aphasia. In *The Lancet Neurology* (Vol. 11, Issue 6, pp. 545–555). Elsevier. https://doi.org/10.1016/S1474-4422(12)70099-6

Grossman, M. (2018). Linguistic Aspects of Primary Progressive Aphasia. *Annual Review of Linguistics*, *4*, 377–403. https://doi.org/10.1146/annurev-linguistics-011516-034253

Hashimoto, N., & Thompson, C. K. (2010). The use of the picture-word interference paradigm to examine naming abilities in aphasic individuals. *Aphasiology*, *24*(5), 580–611. https://doi.org/10.1080/02687030902777567

Henry, M. L., & Grasso, S. M. (2018). Assessment of Individuals with Primary Progressive Aphasia. *Seminars in Speech and Language*, *39*(3), 231–241. https://doi.org/10.1055/s-0038-1660782

Hohlbaum, K., Dressel, K., Lange, I., Wellner, B., Sáez, L. E., Huber, W., Grande, M., Amunts, K., Grodzinsky, Y., & Heim, S. (2018). Sentence repetition deficits in the logopenic variant of PPA: Linguistic analysis of longitudinal and cross-sectional data. *Aphasiology*, *32*(12), 1445–1467. https://doi.org/10.1080/02687038.2017.1423271

Janssen, N. (2020). *Staying connected as we speak* (PhD thesis). Radboud University Nijmegen.

Janssen, N., Piai, V., Beckmann, C. F., Kessels, R. P. C., & Roelofs, A. (2018). *SynTest*.

Janssen, N., Roelofs, A., Mangnus, M., Sierpowska, J., Kessels, R. P. C., & Piai, V. (2020). How the speed of word finding depends on ventral tract integrity in primary progressive aphasia. *NeuroImage: Clinical*, *28*(July). https://doi.org/10.1016/j.nicl.2020.102450

Jiskoot, L., Satoer, D., Seelaar, H., & Dopper, E. G. P. (2020). Primary Progressive Aphasia : Clinical practise , bottlenecks in diagnosis and treatment , suggestions for improvement , and future research [in Dutch]. *Tijdschrift Voor Neuropsychologie*, *15*(1).

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, *59*, 193–224. https://doi.org/10.1146/annurev.psych.59.103006.093615

Kassambara, A. (2020). *ggpubr: "ggplot2" Based Publication Ready Plots*. https://cran.r-project.org/package=ggpubr

Kassambara, A. (2021). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*. https://cran.r-project.org/package=rstatix

Kirshner, H. S. (2014). Frontotemporal dementia and primary progressive aphasia, a review. *Neuropsychiatric Disease and Treatment*, *10*, 1045–1055. https://doi.org/10.2147/NDT.S38821

Kok, R., & Verhey, F. (2002). *Dutch translation of the mini mental state examination (Folstein et al., 1975)*. GGZ.

Kuhn, M. (2020). *caret: Classification and Regression Training*. https://cran.r-project.org/package=caret

Leyton, C. E., & Hodges, J. R. (2014). Differential diagnosis of primary progressive aphasia variants using the international criteria. *Aphasiology*, *28*(8–9), 909–921. https://doi.org/10.1080/02687038.2013.869306

Machulda, M. M., Whitwell, J. L., Duffy, J. R., Strand, E. A., Dean, P. M., Senjem, M. L., Jack, C. R., & Josephs, K. A. (2013). Identification of an atypical variant of logopenic progressive aphasia. *Brain and Language*, *127*(2), 139–144. https://doi.org/10.1016/j.bandl.2013.02.007

Markwick, A., Zamboni, G., & De Jager, C. A. (2012). Profiles of cognitive subtest impairment in the Montreal Cognitive Assessment (MoCA) in a research cohort with normal Mini-Mental State Examination (MMSE) scores. *Journal of Clinical and Experimental Neuropsychology*, *34*(7), 750–757. https://doi.org/10.1080/13803395.2012.672966

Marshall, C. R., Hardy, C. J. D., Volkmer, A., Russell, L. L., Bond, R. L., Fletcher, P. D., Clark, C. N., Mummery, C. J., Schott, J. M., Rossor, M. N., Fox, N. C., Crutch, S. J., Rohrer, J. D., & Warren, J. D. (2018). Primary progressive aphasia: a clinical approach. *Journal of Neurology*, *265*(6), 1474–1490. https://doi.org/10.1007/s00415-018-8762-6

Matías-Guiu, J. A., & García-Ramos, R. (2013). Afasia progresiva primaria: Del síndrome a la enfermedad. *Neurologia*, *28*(6), 366–374. https://doi.org/10.1016/j.nrl.2012.04.003

Matias-Guiu, J. A., Suárez-Coalla, P., Pytel, V., Cabrera-Martín, M. N., Moreno-Ramos, T., Delgado-Alonso, C., Delgado-Álvarez, A., Matías-Guiu, J., & Cuetos, F. (2020). Reading prosody in the non-fluent and logopenic variants of primary progressive aphasia. *Cortex*, *132*, 63–78. https://doi.org/10.1016/j.cortex.2020.08.013

Mesulam, M. M., Rogalski, E. J., Wieneke, C., Hurley, R. S., Geula, C., Bigio, E. H., Thompson, C. K., & Weintraub, S. (2014). Primary progressive aphasia and the evolving neurology of the language network. *Nature Reviews Neurology*, *10*(10), 554–569. https://doi.org/10.1038/nrneurol.2014.159

Mesulam, M. M., & Weintraub, S. (1992). Spectrum of primary progressive aphasia. *Baillière's Clinical Neurology*, *1*(3), 583–609.

Meyer, A. M., Snider, S. F., Campbell, R. E., & Friedman, R. B. (2015). Phonological short-term memory in logopenic variant primary progressive aphasia and mild Alzheimer's disease. *Cortex*, *71*, 183–189. https://doi.org/10.1016/j.cortex.2015.07.003

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Neuwirth, E. (2014). *RColorBrewer: ColorBrewer Palettes* (R package version 1.1-2). https://cran.r-project.org/package=RColorBrewer

Peach, R. K., & Shapiro, L. P. (2012). Cognition and Acquired Language Disorders. In *Cognition and Acquired Language Disorders*. https://doi.org/10.1016/C2009-0-42045-2

Piguet, O., Leyton, C. E., Gleeson, L. D., Hoon, C., & Hodges, J. R. (2015). Memory and emotion processing performance contributes to the diagnosis of non-semantic primary progressive aphasia syndromes. *Journal of Alzheimer's Disease*, *44*(2), 541–547. https://doi.org/10.3233/JAD-141854

R Core Team. (2021). *R statistical software*.

Ramanan, S., Flanagan, E., Leyton, C. E., Villemagne, V. L., Rowe, C. C., Hodges, J. R., & Hornberger, M. (2016). Non-verbal episodic memory deficits in primary progressive aphasias are highly predictive of underlying amyloid pathology. *Journal of Alzheimer's Disease*, *51*(2), 367–376. https://doi.org/10.3233/JAD-150752

Roomer, E. K., Brok, S., Hoogerwerf, A. C., & Linn, D. E. (2011). *Een test voor woordvinding Hogeschool Utrecht Logopedie*.

Spinelli, E. G., Mandelli, M. L., Miller, Z. A., Santos-Santos, M. A., Wilson, S. M., Agosta, F., Grinberg, L. T., Huang, E. J., Trojanowski, J. Q., Meyer, M., Henry, M. L., Comi, G., Rabinovici, G., Rosen, H. J., Filippi, M., Miller, B. L., Seeley, W. W., & Gorno-Tempini, M. L. (2017). Typical and atypical pathology in primary progressive aphasia variants. *Annals of*

Neurology, *81*(3), 430–443. https://doi.org/10.1002/ana.24885

Tee, B. L., & Gorno-Tempini, M. L. (2019). Primary progressive aphasia: A model for neurodegenerative disease. *Current Opinion in Neurology*, *32*(2), 255–265. https://doi.org/10.1097/WCO.0000000000000673

Thompson, C. K., Cho, S., Price, C., Wieneke, C., Bonakdarpour, B., Rogalski, E., Weintraub, S., & Mesulam, M. M. (2012). Semantic interference during object naming in agrammatic and logopenic primary progressive aphasia (PPA). *Brain and Language*, *120*(3), 237–250. https://doi.org/10.1016/j.bandl.2011.11.003

Tippett, D. C., Hillis, A. E., & Tsapkini, K. (2015). Treatment of Primary Progressive Aphasia. *Current Treatment Options in Neurology*, *17*(8). https://doi.org/10.1007/s11940-015-0362-5

van Loon - Vervoorn, W. A. (2005). *De Boston BenoemingsTaak - Een test voor woordvinding bij afasie, 4e herziene druk*.

van Loon - Vervoorn, W. A., Stumpel, H. J., & de Vries, L. A. (1994). *De Boston BenoemingsTaak - Een test voor woordvinding bij afasie*.

Vandenberghe, R. (2016). Classification of the primary progressive aphasias: Principles and review of progress since 2011. *Alzheimer's Research and Therapy*, *8*(1), 1–9. https://doi.org/10.1186/s13195-016-0185-y

Vandenberghe, R. R., Vandenbulcke, M., Weintraub, S., Johnson, N., Porke, K., Thompson, C. K., & Mesulam, M. M. (2005). Paradoxical features of word finding difficulty in primary progressive aphasia. *Annals of Neurology*, *57*(2), 204–209. https://doi.org/10.1002/ana.20362

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth). Springer. https://www.stats.ox.ac.uk/pub/MASS4/

Verhage, F. (1964). *Intelligentie en leeftijd bij volwassenen en bejaarden - Research database - University of Groningen*. 2–5. https://research.rug.nl/en/publications/intelligentie-en-leeftijd-bij-volwassenen-en-bejaarden

Visch-Brink, E., Vandenborre, D., de Smet, H. J., & Mariën, P. (2014). *CAT-NL Comprehensive Aphasia Test - Nederlandstalige bewerking*. Pearson.

Wets, I. (2020). *Primair progressieve afasie : een complex plaatje*. Radboud Universiteit.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H. (2021). *forcats: Tools for Working with Categorical Variables (Factors)*. https://cran.r-project.org/package=forcats

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., Francois, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the {tidyverse}. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wilson, S. M., Galantucci, S., Tartaglia, M. C., & Gorno-Tempini, M. L. (2012). The neural basis of syntactic deficits in primary progressive aphasia. *Brain and Language*, *122*(3), 190–198. https://doi.org/10.1016/j.bandl.2012.04.005

Bijlagen

Bijlage 1 Opleiding scoringssysteem volgens Verhage (1964)

Codering Verhage	*Opleidingsniveau*
1	Basisonderwijs: Lager onderwijs, lager speciaal basisonderwijs
2	Lager of voorbereidend beroepsonderwijs: LBO, LTS, ambachtsschool, huishoudschool en nijverheidsonderwijs (LHNO); VMBO-B: basisberoepsgericht/praktijk
3	Middelbaar algemeen voortgezet onderwijs; MAVO, MULO, ULO, MBO-kort, VMBO-T: theoretische of kaderberoepsgerichte leerweg of gemengde leerweg
4	Middelbaar beroepsonderwijs, MBO-lang, beroepsopleidende leerweg (BOL), beroepsbegeleidende leerweg (BBL), MTS, MEAO
5	Hoger algemeen voorbereidend onderwijs; HAVO, vooropleiding hoger beroepsonderwijs (VHBO), middelbare meisjesschool (MMS); VWO, atheneum, gymnasium, hogere burgerschool (HBS)
6	Hoger beroepsonderwijs; HBO, kweekschool (PAB), conservatorium, MO-akten, nieuwe stijl hogescholen, HTS, HEAO
7	Wetenschappelijk onderwijs; WO, universiteit, technische/economische hogeschool oude stijl