



APROCSA-NL

**Auditief-perceptuele beoordeling
van natuurlijke spraak bij afasie**

Mara Barberis

Marianne Casilio

Stephen M. Wilson

Maaïke Vandermosten

KU LEUVEN





Colofon

APROCSA-NL

© Sig vzw - 2025

Alle rechten voorbehouden

Sig vzw, Pachthofstraat 1, 9308 Gijzegem (Aalst)

tel. +32 (0)53 38 28 18 - fax +32 (0)53 38 28 19

info@sig-net.be - www.sig-net.be

Eerste druk: 2025

D/2025/7746/05 - ISBN 978-90-5873-131-9

Niets uit deze uitgave mag worden gekopieerd zonder voorafgaande toestemming van de uitgever.

Inhoud

Achtergrond	1
 DEEL 1: PRAKTISCHE HANDLEIDING	
Afname: het verzamelen van een taalstaal	4
Scoring: de auditief-perceptuele benadering	5
APROCSA-NL features	6
APROCSA-NL scoreformulier	13
APROCSA-NL cheat sheet	14
 DEEL 2: PSYCHOMETRISCH ONDERZOEK	
Deelnemers met afasie	15
Beoordelaars en beoordelingsprocedure	16
Interbeoordelaarsbetrouwbaarheid van APROCSA-NL	17
Concurrente validiteit van APROCSA-NL	23
Besluit van het psychometrisch onderzoek	28
 Referenties	 29

Achtergrond

Achtergrond

Personen met afasie ondervinden vaak moeilijkheden bij het produceren van natuurlijke spraak, wat de alledaagse communicatie ernstig kan belemmeren. Dit kan leiden tot frustratie, een verminderde levenskwaliteit, psychosociale stress en sociale isolatie (Stark, 2023). Toch zijn afasie assessments vaak sterk gericht op geïsoleerde taaltaken met losse fonemen, woorden of zinnen (Leaman & Edmonds, 2021), die een beperkte ecologische validiteit hebben. Dit kan leiden tot discrepanties tussen de klinische beoordeling en de eigenlijke communicatieve vaardigheden die de persoon met afasie in het dagelijks leven ervaart (Cavanaugh & Haley, 2020; Fromm et al., 2017; Holland et al., 2019; Kim et al., 2022; Lesser & Algar, 1995; Stark & Fukuyama, 2021; Wallace & Kimelman, 2013). Daarom pleiten klinici en onderzoekers ervoor om systematisch natuurlijke spraaktaken te includeren bij het beoordelen van afasie (Azios et al., 2021; Dietz & Boyle, 2018).

Er werden reeds verschillende natuurlijke spraakfeatures gedefinieerd binnen de afasiologie (Bryant et al., 2016), maar de psychometrische eigenschappen van deze features zijn slechts beperkt onderzocht (Boyle, 2014, 2015; Brookshire & Nicholas, 1994; Pritchard et al., 2018; Stark & Dalton, 2024). Bijkomend wordt natuurlijke spraakanalyse nog te beperkt toegepast in de praktijk door de hoge workload (Bryant et al., 2019; Hallowell, 2016; Stark et al., 2021), wat typisch het manueel transcriberen en analyseren van natuurlijke spraak inhoudt. Dit proces kan tot één uur duren per minuut aan natuurlijke spraak (Armstrong et al., 2007). Als reactie hierop werden ‘transcription-less’ instrumenten ontwikkeld (zie bijvoorbeeld Armstrong et al., 2007; Casilio et al., 2019; Kim & Wright, 2020; Ruiter et al., 2022; Stark & Dalton, 2024), die meer bruikbaar zijn in de klinische praktijk aangezien de evaluatie van natuurlijke spraak op basis van het beluisteren van de audio-opname gebeurt, zonder dat er een transcriptie gemaakt dient te worden.

Om tegemoet te komen aan de nood voor een klinisch inzetbaar, comprehensief instrument om natuurlijke spraak bij personen met afasie te beoordelen, ontwikkelden Casilio en collega's in 2019 het ‘Auditory-Perceptual Rating of Connected Speech in Aphasia’ instrument. APROCSEA bestaat uit 27 natuurlijke spraakfeatures die elk beoordeeld worden op een vijfpuntschaal, variërend van 0 (niet aanwezig of binnen het bereik van gezonde, oudere sprekers) tot 4 (ernstig, kenmerk is bijna altijd aanwezig). Wat de scoring van APROCSEA-NL onderscheidt

van andere afasie-instrumenten, is dat er aan een bepaalde score niet enkel een ernstgraad (bv. mild, matig, ernstig), maar ook een beschrijving van deze ernstgraad gekoppeld wordt, wat de logopedist meer houvast geeft bij het scoren van de natuurlijke spraakfeatures. De 27 natuurlijke spraakfeatures zijn onderverdeeld in zeven categorieën: (1) ophalen van woorden, (2) selectie van woorden en klanken, (3) grammaticale constructies, (4) tempo en timing, (5) zelfcorrecties, (6) duidelijkheid en (7) diagnostisch.

Voor het Nederlands zijn er slechts een beperkt aantal natuurlijke spraak-instrumenten voor afasie beschikbaar. Deze hebben bovendien vaak beperkingen zoals een zeer arbeidsintensieve scoringsprocedure of een beperkte representatie van de diverse aspecten van natuurlijke spraak. Daarnaast ontbreekt het vaak aan een uitgebreide psychometrische validatie. De situatieplaat-beschrijving van de Nederlandstalige bewerking van de Comprehensieve Afasie Test (CAT-NL) (Visch-Brink et al., 2015) omvat een auditief-perceptuele beoordeling van de nauwkeurigheid en adequaatheid van de informatie, het tempo en de vloeiendheid, de grammaticale variatie en correctheid en een algemene beoordeling. Er is echter weinig houvast om aan deze aspecten een score toe te kennen, en er zijn geen psychometrische gegevens beschikbaar voor deze subtest van de CAT-NL. Meer recent werd de conceptuele scoring van de situatieplaat toegevoegd toegevoegd (Vandenborre et al., 2021), waarbij een objectieve, kwantitatieve maat voor het produceren van sleutelconcepten, het leggen van causale verbanden en het totaal aantal woorden wordt bekomen. De interbeoordelaarsbetrouwbaarheid van deze maten is uitstekend (ICC sleutelconcepten = 0.94, ICC aantal woorden = 0.99, ICC causale relaties = 0.92) (Vandenborre et al., 2018). De Nederlandstalige bewerking van de Main Concept Analysis (MCA) (Criel et al., 2021) kwantificeert de effectiviteit en efficiëntie van de informatieoverdracht op basis van de aanwezigheid, accuraatheid en volledigheid van de Main Concepts die geproduceerd worden tijdens vier sequentiële situatieplaatbeschrijvingen. De MCA scores toonden een goede tot uitstekende interbeoordelaarsbetrouwbaarheid (Criel et al., 2025). Bij de ANTAT vervolgens wordt natuurlijke spraak gemeten via de respons van de persoon met afasie op een alledaagse situatie (bv. een afspraak verzetten). De verbale responsen worden beoordeeld op een vijfpuntschaal met betrekking tot de verstaanbaarheid en de begripelijkheid van de boodschap. Van deze rating werd reeds aangetoond dat ze betrouwbaar en valide is (Blomert et al., 1994). Meer recent werd de Content Unit

(CU) scoring toegevoegd aan de ANTAT. Van deze CU-scoring werd eveneens aangetoond dat deze betrouwbaar en valide is (Ruiter et al., 2022). Een belangrijke kanttekening bij de zonet besproken testen is dat ze slechts beperkte aspecten van natuurlijke spraak meten, met een sterke focus op informatieoverdracht (cf. sleutelconcepten in de CAT-NL situatieplaat, Main Concepts in de MCA, Begrijpelijkheid / Content Units in de ANTAT). De Analyse voor Spontane Taal bij Afasiepatiënten (ASTA) (Boxum et al., 2019; Van Der Scheer et al., 2011) daarentegen is een meer comprehensieve methode die negentien natuurlijke spraakmetingen overheen verschillende taaldomeinen in kaart brengt. De interbeoordelaars-betrouwbaarheid voor negen ASTA features varieerde van redelijk tot uitstekend (Van Der Scheer et al., 2011). De ASTA is echter zeer tijdsintensief door het transcriberen en uitgebreid linguïstisch analyseren van het taalstaal, en wordt bijgevolg weinig ingezet in de klinische praktijk.

Het **doel van APROCSA-NL** is om op een tijdsefficiënte en toch comprehensieve manier een sterkte-zwakte profiel van een persoon met afasie op te stellen. Hierbij wordt uitgebreid rekening gehouden met opvallende kenmerken van natuurlijke taal- en spraakstoornissen bij afasie die representatief zijn voor problemen binnen alle kerndomeinen van spraak en taal (lexico-semantiek, morfosyntaxis, fonologie, spraakprogrammering en -uitvoering). Op deze manier zorgt APROCSA-NL voor een klinisch haalbaar assessment van natuurlijke spraak bij afasie.

DEEL 1:

Praktische handleiding

Afname: het verzamelen van een taalstaal

Zorg dat je in een rustige ruimte zit en er geen afleiders (bv. tv, radio) zijn. De testleider stelt open vragen aan de persoon met afasie. Enkele richtvragen kunnen de volgende zijn:

- Kan u mij iets vertellen over de dag dat u de beroerte kreeg?
- Hoe is uw herstel nadien verlopen?
- Kan u vertellen over een belangrijke gebeurtenis in uw leven?
- Wat heeft u afgelopen weekend gedaan?
- Wat voor beroep heeft u (gehad)? Kan u daar wat over vertellen?
- Kan u iets vertellen over uw familie?
- Heeft u hobby's? Wat doet u graag in uw vrije tijd?

Voor de meeste personen met afasie zullen de eerste twee vragen een voldoende representatief taalstaal uitlokken. In het geval dat er toch weinig verteld wordt (en dit niet te wijten is aan de ernst van de taal- of communicatieve problemen), of dat dit onderwerp te emotioneel is om over te spreken, kan er teruggevallen worden op de andere vragen. Deze lijst is uiteraard niet exhaustief.

Er is geen minimum vereiste voor de lengte van dit taalstaal, wat maakt dat de APROCSA-NL zelfs bij personen met ernstige afasie gebruikt kan worden. Het is wel belangrijk om te weten dat niet alle APROCSA-NL features dezelfde betrouwbaarheid hadden voor taalstalen van verschillende lengtes en verschillende ernstgraden van afasie. Het is daarom belangrijk om rekening te houden met de informatie uit het psychometrisch onderzoek (cf. Deel 2, p. 17) om te weten welke APROCSA-NL features betrouwbaar zijn in welke condities.

Neem het audiofragment op zodat je hier later terug naar kan luisteren. In het validatieonderzoek werd hiervoor een BOYA Lavalier BY-M1 microfoon en een Philips DVT1120 VoiceTracer gebruikt. Indien de persoon met afasie ermee instemt om ook een video-opname te maken, kan dit zinvol zijn om bv. zoekgedrag in het geval van comorbide spraakapraxie in rekening te kunnen brengen.

Scoring: de auditief-perceptuele benadering

APROCSA-NL hanteert een auditief-perceptuele benadering, waarbij 27 kenmerken ('features') of gedragingen worden beoordeeld op basis van wat er wordt gehoord of waargenomen op een audio- of video-opname. Dit is gelijkaardig aan de manier waarop spraakstoornissen (bv. dysartrie) het vaakst worden beoordeeld. De 27 features worden beoordeeld op een vijfpuntsschaal, waarbij elk punt gedefinieerd is op basis van de ernst en frequentie van het voorkomen van het feature. Deze definitie werd ontleend aan de Apraxia of Speech Rating Scale (Strand et al., 2014).

0	Niet aanwezig	Niet aanwezig of binnen het bereik van gezonde, oudere sprekers*.
1	Mild	Kenmerk is detecteerbaar, maar treedt niet frequent op.
2	Matig	Kenmerk is vaak aanwezig, maar het is niet pervasief**.
3	Gemarkeerd	Matig ernstig, pervasief**.
4	Ernstig	Kenmerk is bijna altijd aanwezig.

*Veel personen met afasie zullen slechts een subset van de features vertonen. Bovendien zullen gezonde personen zonder afasie misschien ook enkele van deze features vertonen. Meer specifiek, zouden bij gezonde sprekers ook hernemingen, valse starts of pauzes wanneer er naar een woord gezocht wordt, kunnen optreden. Andere mensen praten dan weer langzaam, en ook is het niet ongebruikelijk voor neurologisch gezonde sprekers om occasioneel een paragrammatische uiting te produceren. Als een persoon met afasie bijgevolg een feature vertoont dat binnen het verwacht bereik van een gezonde, oudere spreker valt, beoordeel het kenmerk dan als niet aanwezig (0).

**Pervasief gaat over de impact of invloed die het feature heeft op het gehele taalsample. Bij score 3 (gemarkeerd) betekent pervasief dat er het voorkomen van dit feature een significante impact heeft op het taalsample.

Luister aandachtig naar het audiofragment en bepaal de juiste beoordeling voor elk feature. Luister één keer naar het fragment. Terwijl je luistert, beoordeel je de features en neem je (indien gewenst) notities over observaties die je maakt. Pauzeer het fragment niet. Overloop nadien je scores en notities. Raadpleeg indien nodig de beschrijvingen van de APROCSA-NL features. Luister nadien nog éénmaal naar het audiofragment. Controleer of breng indien nodig wijzigingen aan. Op deze manier zou het beoordelen van de natuurlijke spraak bij een persoon met afasie slechts 15 minuten in beslag mogen nemen. Uit het validatieonderzoek van APROCSA-NL bleek de gemiddelde scoringsduur ongeveer 10 minuten te zijn.

Probeer zo objectief mogelijk te zijn bij het beoordelen van een feature. Ongeacht of je denkt dat een feature een onderliggende stoornis weerspiegelt, of eerder secundair is aan een ander taalkundig, cognitief of motorisch proces, beoordeel steeds eenvoudigweg wat je waarneemt in het audiofragment.

Beschouw de features binnen de context van een uiting. Hoewel het bepalen van een uiting enigszins subjectief kan zijn, is het toch een belangrijke variabele die de lengte en complexiteit van de spraak van de persoon met afasie reflecteert. Houd rekening met volgende factoren indien je een uiting begrenst:

- Een zin is een uiting,
- Zinnen die samengevoegd worden met 'en' zijn aparte uitingen,
- Een dalend intonatiepatroon suggereert het einde van een uiting.

Pauzes zijn onbetrouwbaar om de uitingsgrenzen bij personen met afasie te bepalen.

APROCSA-NL features

Anomie

Algemene indruk van woordvindingsproblemen. Kan op verschillende manieren geconcretiseerd worden, die ook allemaal gemeten worden als aparte features: woordvindingspauzes voor zelfstandige naamwoorden of werkwoorden, onvolledige uitingen nadat men er niet in geslaagd is het woord op te roepen, commentaar geven op het onvermogen om een woord op te roepen of uit te spreken, lege spraak, circumlocutie (omschrijving) of parafasieën.

Onvolledige uitingen

Uitingen worden niet afgemaakt. De spreker verandert van onderwerp, stopt met praten, probeert een andere modaliteit te gebruiken (bv. gebaren), of beëindigt de uiting eerder vaag (bv. schouders ophalen of 'je weet wel' zeggen).

Lege spraak

De spraak draagt weinig tot geen betekenis over door een gebrek aan specificiteit. Inhoudswoorden worden vervangen door voornaamwoorden of niet-specifieke woorden zoals 'ding'.

Circumlocutie (omschrijving)

Om woorden of concepten heen praten, zonder het doelwoord zelf te gebruiken. Dit kan één of meerdere uitingen omvatten (bv. hij was een acteur in de jaren '50 die veel in films speelde), of gewoon een lijstje van woorden (bv. acteur, jaren '50, films).

Semantische parafasieën

Substitutie van inhoudswoorden door gerelateerde of ongerelateerde inhoudswoorden, bv. kat i.p.v. hond of kaars i.p.v. hond. Substitutie van cijfers mag hier niet gescoord worden.

Fonematische parafasieën

Substitutie, insertie, deletie of transpositie van een duidelijk gearticuleerd foneem of syllabe in een deels herkenbaar woord, bv. pappel i.p.v. appel of palapru i.p.v. paraplu. De spreker is zich vaak bewust van de fout en zal deze proberen te corrigeren, waardoor deze score ook bijdraagt aan 'valse starts', 'hernemingen' en 'conduite d'approche'. De woordvormen zijn voornamelijk non-woorden, al kan het ook zijn dat de respons een bestaand Nederlands woord is. Fouten zijn vaker substituties van medeklinkers, komen voor bij multisyllabische woorden, en komen voor op het einde van het woord.



Met distorsie bedoelen we een afwijkende of onjuiste productie van een spraakklank, waardoor deze niet correct of afwijkend t.o.v. de norm uitgesproken wordt.

Neologismen

Woordvormen die geen bestaande Nederlandse woorden zijn en verscheidene substituties, inserties, deleties of transposities van duidelijk gearticuleerde fonemen bevatten. De foneemselectie is zo ernstig verstoord dat het bedoelde doelwoord slechts vluchtig duidelijk wordt, en de spreker herkent de fout meestal niet.

Jargon

Meestal vloeiende en prosodische correcte, maar grotendeels betekenisloze spraak die parafasieën, neologismen of onverstaanbare uitingen bevat. Lijkt op de Nederlandse grammatica en prosodie.



Perseveratie

Herhaling van eerder gebruikte woorden of uitingen in contexten waarin ze niet langer passend zijn.

Stereotypen en automatismen

Veelgebruikte woorden, neologismen of uitdrukkingen die relatief gemakkelijk en vloeiend worden geproduceerd, bv. 'tan', 'ik weet het', 'verdorie'.

Korte en vereenvoudigde uitingen

Uitingen worden gereduceerd in lengte of complexiteit, zoals de argumentenstructuur of het aantal bijzinnen. Antwoorden die niet uit een volledige zin bestaan*, moeten hier niet beschouwd worden. Score 1 (mild) is voorbehouden voor uitingen met een eenvoudige zinsstructuur, of een gebrek aan bijzinnen. Score 4 (ernstig) is voorbehouden voor éénwoorduitingen.

*Voorbeelden: 'Ben je met je vrouw gekomen?' 'Ja' of 'Met wie ben je gekomen?' 'Mijn vrouw'.

Omissie van functiewoorden

Functiewoorden (bv. lidwoorden, voorzetsels, voornaamwoorden, voegwoorden, hulpwerkwoorden) worden niet gebruikt wanneer dat wel zou moeten. Het weglaten van deze elementen resulteert over het algemeen in ongrammaticale uitingen (bv. wij gaan naar winkel, hij woont huis, kind school gegaan regen). Een score 4 (ernstig) is voorbehouden voor spraak die enkel éénwoorduitingen bevat.

Paragrammatisme

Ongepaste aaneenschakeling van zinsdelen en verkeerd gebruik van woorden, inclusief fouten tegen woordsoorten en substituties van grammaticale woorden en morfemen (bv. het is zo veel geweldig, afasie maakt het moeilijk om te spraak).

Pauzes tussen uitingen

Pauzes tussen uitingen kunnen verband houden met de formulering van de uiting. Pauzes tussen vragen van de testleider en het antwoord van de patiënt worden ook in rekening gebracht. Indien de patiënt er niet in slaagt om meerdere uitingen aan elkaar te rijgen, kan je dit ook scoren onder dit feature.

Pauzes binnen uitingen

Ongepulde of gepulde (bv. 'uh') pauzes binnen een uiting. Zowel de frequentie als de lengte van de pauzes moet hier in rekening gebracht worden.

Aarzelend en moeizaam

Het spreken vraagt veel moeite en is daardoor ongelijkmatig. De intonatie, het ritme of de klemtoon kunnen verminderd, afwezig of verkeerd gebruikt zijn. De spreker kan problemen ondervinden met het efficiënt communiceren van de boodschap.

Verminderd spreektempo

Het aantal woorden per minuut binnen uitingen is verminderd. Traag spreken en pauzes binnen uitingen tellen ook mee bij de scoring van dit feature. Pauzes tussen uitingen, die mogelijk de formulering van de uiting reflecteren, tellen niet mee.

Hernemingen

Opeenvolgingen van één of meerdere volledige woorden worden overbodig gemaakt door daaropvolgende herhalingen, herzieningen, wijzigingen of uitbreidingen, bv. 'de vlieger is ... de jongen bestuurt de vlieger'.

Valse start

Gedeeltelijke woorden worden afgebroken na één of twee fonemen, bv. 'het is een ka-kat'. De spreker produceert (al dan niet) het uiteindelijke doelwoord.

Conduite d'approche

Opeenvolgende pogingen voor een duidelijk aanwezig doelwoord (bv. pon-por-poneren-poberen voor het doelwoord 'proberen'). Het doelwoord kan wel of niet worden bereikt. De spreker is zich bewust van hun fouten. Deze gevallen dragen ook bij aan de scoring van 'hernemingen', 'fonematische parafasieën' of 'neologismen' afhankelijk van hoe dicht de poging aanleunt bij het doelwoord. Conduite d'approche kan taal- of spraakmoeilijkheden reflecteren.

Doelwoord of –uiting onduidelijk

De kenmerken van fonemen (articulatieplaats, -wijze of stemgeving) zijn verstoord. Dit wordt vaak veroorzaakt door distorsies, spraakapraxie, mompelen of jargon. De doelfonemen kunnen wel of niet waarneembaar zijn. In ernstige gevallen kan de betekenis van het doelwoord of de zin/uiting ongekend zijn. Fonemen zouden dan moeilijk of onmogelijk te transcriberen zijn.

Belangrijk: als je niet begrijpt wat de spreker zegt, maar de doelklanken zijn duidelijk, draagt dit eerder bij aan het feature 'betekenis onduidelijk' (zie hieronder) dan aan dit feature.

Betekenis onduidelijk

Het is niet duidelijk waar de spreker het over heeft, of het onderwerp is misschien wel duidelijk, maar wat erover gezegd wordt niet. Hoewel de reactie van de testleider verduidelijkend kan zijn m.b.t. waarover de spreker praat, mag dit niet in rekening gebracht worden bij de scoring.

Off topic

Het is duidelijk waar de spreker het over heeft, maar het is niet duidelijk hoe dit zich verhoudt tot de context. De uitingen van de spreker kunnen enigszins gerelateerd zijn aan het topic, bijvoorbeeld praten over een reis naar Italië bij de vraag 'wat is je favoriete Italiaanse gerecht?'. Dit feature kan bijdragen aan de scores bij 'lege spraak', 'betekenis onduidelijk' of 'semantische parafasie'.

Expressieve afasie

De taalproductie is verstoord; de spreker ervaart moeilijkheden om zichzelf te uiten. Verstoringen kunnen optreden in één van de of alle taaldomeinen (d.i. fonologie, morfologie, syntaxis, semantiek). Taalbegrip mag bij deze scoring niet in rekening gebracht worden.

Spraakapraxie

De spraak is langzaam en bevat verstoringen of verstoorde substituties, die vaker voorkomen in consonantclusters, aan het begin van een woord en in langere woorden. Lettergrepen zijn gesegmenteerd binnen en tussen woorden, en fonemen kunnen verlengd zijn, wat resulteert in atypische prosodie.

Dysartrie

De spraak is moeilijk te begrijpen en wordt gekarakteriseerd als onduidelijk, hakkelig (cf. 'staccato') of gemompeld. Fouten zijn consistent en zijn het resultaat van verminderde kracht, tonus, bewegingsbereik of coördinatie. Spraakademhaling, fonatie, resonantie, articulatie en prosodie kunnen verstoord zijn.

Algemene communicatiebeperking

Algemene indruk van de mate waarin de spreker gehinderd wordt in het overbrengen van hun boodschap. Score 1 (mild) moet een duidelijke spraaktaalstoornis weerspiegelen, maar zonder beperking in het bediscussiëren van allerhande onderwerpen. Score 2 (matig) moet gebruikt worden als de spreker gemakkelijk kan communiceren over eenvoudige, alledaagse onderwerpen, maar beperkt is in het bespreken van complexere onderwerpen. Score 3 (gemarkeerd) moet gebruikt worden indien communicatie over alledaagse onderwerpen mogelijk is met hulp van de testleider, maar de patiënt in de last van de communicatie* deelt. Score 4 (ernstig) moet gebruikt worden indien alle communicatie gefragmenteerd is, en de testleider de last van de communicatie* draagt.

*de last van de communicatie kan gezien worden als het aandeel dat je hebt om adequate informatieoverdracht in een conversatie te verzekeren.

Kanttekeningen bij de APROCSA-NL features

Het is niet de bedoeling om de APROCSA-NL features te vergelijken tussen personen met afasie en gezonde controles. Het gaat om stoorniskenmerken, in die zin dat men denkt dat ze alleen optreden (althans in frequentie of ernst) wanneer de taalverwerking verstoord is.

APROCSA-NL houdt (voorlopig) nog geen rekening met functionele communicatie, sociaal-pragmatische of andere aspecten van discourse (bv. informativiteit of efficiëntie). De focus van APROCSA-NL ligt momenteel op het vastleggen van tekortkoming op het subcomponentniveau van het taalsysteem en het begrijpen hoe deze zich in samenspel manifesteren in natuurlijke spraak.

Bij sommige vormen van afasie zullen personen hun fouten (proberen) herstellen. Fouten moeten, zelfs in het geval van een succesvolle herstelling, toch nog in rekening gebracht worden voor de beoordeling van het betreffende feature. Dit soort herstellingen zullen vooral voorkomen bij volgende features 'herneming', 'valse start', 'circumlocutie' of 'conduite d'approche'.

Het laatste feature, 'algemene communicatiebeperking' is geen gemiddelde van de andere features. Met andere woorden, een persoon met afasie krijgt niet automatisch een score 2 indien het merendeel van de voorgaande features een score 2 kreeg. De ernst van sommige features (bv. agrammatismen) of het effectieve gebruik van communicatiestrategieën (bv. circumlocuties) kunnen de algehele presentatie beïnvloeden. Probeer, net als bij de andere features, objectief te beoordelen wat er in het audiofragment aanwezig is.

APROCSA-NL

Auditief-perceptuele beoordeling van natuurlijke spraak bij afasie

Naam:

Datum:

Opmerkingen:

Beoordeel de natuurlijke spraak van de persoon met afasie gebruikmakend van volgende schaal:

Score	Ernst	Beschrijving
0	Afwezig	Niet aanwezig of binnen het bereik van gezonde, oudere sprekers.
1	Mild	Kenmerk is detecteerbaar, maar treedt niet frequent op.
2	Matig	Kenmerk is vaak aanwezig, maar is niet pervasief.
3	Gemarkeerd	Matig ernstig, pervasief.
4	Ernstig	Kenmerk is bijna altijd aanwezig.

Natuurlijke spraak feature		0	1	2	3	4
Ophalen van woorden	Anomie	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Onvolledige uitingen	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Lege spraak	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Circumlocutie (omschrijving)	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
Selectie van woorden en klanken	Semantische parafasieën	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Fonematische parafasieën	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Neologismen	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Jargon	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Perseveratie	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Stereotypen en automatismen	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
Grammaticale constructies	Korte en vereenvoudigde uitingen	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Omissie van functiewoorden	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Paragrammatisme	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
Tempo en timing	Pauze tussen uitingen	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Pauze binnen uitingen	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Aarzelend en moeizaam	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Verminderd spreektempo	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
Zelf-correcties	Hernemingen	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Valse start	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Conduite d’approche	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
Duidelijkheid	Doelwoord of -uiting onduidelijk	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Betekenis onduidelijk	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Off topic	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
Diagnostisch	Expressieve afasie	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Spraakapraxie	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Dysartrie	Afwezig	Mild	Matig	Gemarkeerd	Ernstig
	Algemene communicatiebeperking	Afwezig	Mild	Matig	Gemarkeerd	Ernstig

APROCSA-NL cheat sheet

Anomie = algemene indruk van woordvindingsproblemen.

Onvolledige uitingen = uitingen worden niet afgemaakt (ander onderwerp, modaliteit, vaag einde...).

Lege spraak = gebrek aan specificiteit → weinig betekenis ('ding').

Circumlocutie (omschrijving) = om woorden heen praten, zonder het woord zelf te gebruiken.

Semantische parafasieën = substitutie van inhoudswoorden door (on)gerelateerde inhoudswoorden.

Fonematische parafasieën = substitutie, insertie, deletie of transpositie van een duidelijk gearticuleerd foneem of syllabe in een deels herkenbaar woord.

Neologismen = woordvormen die geen bestaande Nederlandse woorden zijn.

Jargon = betekenisloze spraak met o.a. parafasieën & neologismen, meestal vloeiend en prosodisch correct.

Perseveratie = herhaling eerder gebruikt woord/uiting, niet langer passend in de context

Stereotypen & automatismen = veelgebruikte woorden, neologismen of uitdrukkingen die relatief gemakkelijk en vloeiend worden geproduceerd.

Korte en vereenvoudigde uitingen = uitingen worden gereduceerd in lengte of complexiteit.

1 = eenvoudige zinsstructuur, gebrek aan bijzinnen, 4 = uitsluitend éénwoorduitingen

Omissie van functiewoorden = functiewoorden worden niet gebruikt wanneer dat wel zou moeten.

4 = uitsluitend éénwoorduitingen

Paragrammatisme = ongepaste aaneenschakeling van zinsdelen en verkeerd gebruik van woorden.

Pauze tussen uitingen = m.b.t. formulering uiting / pauze na vraag testleider / aaneenrijgen uitingen

Pauze binnen uitingen = gevuld ('uh') en ongevuuld, frequentie & lengte in rekening brengen.

Aarzelend en moeizaam = spreken vraagt moeite en is ongelijkmatig (intonatie, ritme, klemtoon).

Verminderd spreektempo = ↓woorden/minuut binnen uiting (traag spreken, pauzes binnen uiting).

Hernemingen = redundante herhaling van één of meerdere volledige woorden, variabel in lengte.

Valse start = gedeeltelijke woorden worden afgebroken na één of twee fonemen, bv. 'een ka-kat'.

Conduite d'approche = opeenvolgende pogingen voor een duidelijk aanwezig doelwoord, bv. 'pon-por-poneren-poberen' voor het doelwoord 'proberen'.

Doelwoord of -uiting onduidelijk = de kenmerken van fonemen (articulatieplaats, -wijze of stemgeving) zijn verstoord.

Betekenis onduidelijk = het is niet duidelijk waar de spreker het over heeft, of het onderwerp is misschien wel duidelijk, maar wat er over gezegd wordt niet.

Off topic = het is duidelijk waar de spreker het over heeft, maar het is niet duidelijk hoe dit zich verhoudt tot de context; de uitingen kunnen enigszins gerelateerd zijn aan het topic.

Expressieve afasie = de taalproductie is verstoord; de spreker ervaart moeilijkheden om zichzelf te uiten.

Spraakapraxie = de spraak is langzaam en bevat verstoringen of verstoorde substituties, die vaker voorkomen in consonantclusters, aan het begin van een woord en in langere woorden. Lettergrepen zijn gesegmenteerd binnen en tussen woorden, en fonemen kunnen verlengd zijn, wat resulteert in atypische prosodie.

Dysartrie = de spraak is moeilijk te begrijpen en wordt gekarakteriseerd als onduidelijk, hakkelig ('staccato') of gemompeld. Fouten zijn consistent en zijn het resultaat van verminderde kracht, tonus, bewegingsbereik of coördinatie. Spraakademhaling, fonatie, resonantie, articulatie en prosodie kunnen verstoord zijn.

Algemene communicatiebeperking = Algemene indruk van de mate waarin de spreker gehinderd wordt in het overbrengen van hun boodschap.

1 = duidelijke spraak-taalstoornis, zonder beperking in het bediscussiëren van allerhande onderwerpen. 2 = gemakkelijk communiceren over eenvoudige, alledaagse onderwerpen, maar beperkt in het bespreken van complexere onderwerpen. 3 = communicatie over alledaagse onderwerpen is mogelijk met hulp van de testleider, maar de spreker deelt in de last van de communicatie, 4 = alle communicatie is gefragmenteerd, en de testleider draagt de last van de communicatie.

DEEL 2:

Psychometrisch onderzoek

Deelnemers met afasie

Aan deze studie participeerden 58 personen met chronische afasie ten gevolge van een beroerte (≥ 6 maanden post onset). De deelnemers werden gediagnosticeerd met afasie in de acute fase door een logopedist, en scoorden op het moment van deelname aan de studie afwijkend op de ScreeLing (Visch-Brink et al., 2010) of Nederlandse Benoem Test (NBT) (Van Ewijk et al., 2018), of rapporteerden communicatieve problemen volgens de Communicatieve Effectiviteits Index (CETI) (Lomas et al., 1989). Personen met afasie moesten geen specifiek type afasie hebben om geïnccludeerd te worden, aangezien we een zo representatief mogelijk sample van de heterogene populatie van personen met afasie wilden verzamelen. Deelnemers werden geëxcludeerd in het geval van een comorbide psychiatrische of neurodegeneratieve aandoening. De tabel hieronder geeft een demografische en klinische karakterisering van de deelnemersgroep weer.

Leeftijd	67.1 $\pm$ 13.4 (32-95) jaar
Geslacht	27 M; 31 V
Handvoorkeur	41 R; 14 L; 3 Am
Opleidingsniveau	7 LO; 27 MO; 16 H; 8 U
Tijd post onset	38.7 $\pm$ 58.2 (6-360) maanden
Afasie ernst ¹	58.9 $\pm$ 12.1 (19-72)
Woordvinding ²	208.4 $\pm$ 76.6 (7-274)
Communicatieve Effectiviteit ³	6.6 $\pm$ 1.8 (2.7-10.0)
Dysartrie ⁴	41.38%
Spraakapraxie ⁴	36.21%

M = Man, V = Vrouw, R = Rechtshandig, L = Linkshandig, Am = Ambidexter, LO = Lager onderwijs, MO = Middelbaar onderwijs, H = Hogeschool, U = Universiteit. ¹Gebaseerd op de ScreeLing (Visch-Brink et al., 2010). Totaalscores lager dan 68 wijzen op afasie, maximumscore = 72. ²Gebaseerd op de Nederlandse Benoem Test (Van Ewijk et al., 2018). Totaalscores lager dan 255 wijzen op woordvindingsproblemen, maximumscore = 276. ³Gebaseerd op de Communicatieve Effectiviteits Index (CETI) (Lomas et al., 1989). Scores op de zestien items werden gemiddeld tot een score op tien, waarbij lagere scores een lagere communicatieve effectiviteit weerspiegelen. ⁴Binaire classificatie (aanwezig/afwezig) door een logopedist.

Beoordelaars en beoordelingsprocedure

Twee afzonderlijke groepen van Nederlandstalige beoordelaars scoorden de APROCSA-NL voor het sample van 58 personen met afasie. De eerste groep bestond uit drie onervaren beoordelaars, meer specifiek, eerstejaars masterstudenten logopedie die theoretische achtergrond hadden over neurogene communicatiestoornissen, maar nog geen klinische ervaring opgedaan hadden. De tweede groep bestond uit drie ervaren raters. De eerste rater was logopedist van opleiding en werkzaam als onderzoeker, met vier jaar ervaring in het transcriberen en analyseren van natuurlijke spraak bij personen met afasie. De tweede en derde ervaren raters waren logopedisten met zeven tot negen jaar klinische ervaring en bijzondere expertise in neurogene communicatiestoornissen, werkzaam in zowel acute als chronische settings.

Voor het beoordelen van de eigenlijke validatiesamples, scoorden de ervaren beoordelaars elk vijf kalibratiesamples. Deze kalibratiesamples bestonden uit audio-opnames van de mondelinge beschrijving van de situatieplaat van de CAT-NL. De samples weerspiegelden diverse afasieprofielen en waren specifiek geselecteerd zodat de verschillende APROCSA-NL features aan bod zouden komen. Nadat de ervaren beoordelaars de kalibratiesamples afzonderlijk gescoord hadden, werden deze beoordelingen besproken en werd er consensus bereikt over de vijf kalibratiesamples. De onervaren beoordelaars kregen elk een 1,5 uur durende training met uitleg over het doel van APROCSA-NL, uitgebreide toelichting over de APROCSA-NL features en richtlijnen voor het scoren. Na deze trainingssessie oefenden de studenten onafhankelijk met de vijf kalibratiesamples. Een week later was er een tweede trainingssessie om hun ratings te vergelijken met de consensus ratings. Deze procedure zorgde ervoor dat alle beoordelaars gekalibreerd waren volgens dezelfde consensus scores. De beoordelaars scoorden elk de 58 samples van personen met afasie over een tijdspanne van twee maanden. Er werd aan de beoordelaars gevraagd om niet langer dan één uur aaneensluitend te scoren om vermoeidheidseffecten te elimineren. Het toekennen van de APROCSA-NL scores gebeurde op dezelfde manier als in de originele APROCSA studie: beoordelaars luisterden één keer naar het sample en scoorden vervolgens het merendeel van de features. Nadien was er de mogelijkheid om nogmaals te luisteren indien er over bepaalde features nog onduidelijkheden waren. Beoordelaars werden geadviseerd om notities te maken tijdens het beluisteren van het taalstaal.

Interbeoordelaarsbetrouwbaarheid van APROCSA-NL

Intraclass correlatie coëfficiënten (ICC's) werden berekend als maat voor interbeoordelaarsbetrouwbaarheid (IBB). ICC (A,k), die de absolute overeenstemming schat van metingen die gemiddelden zijn van k onafhankelijke metingen, werden berekend voor tweerichtingsmodellen. We onderzochten kwalitatief of het ervaringsniveau een invloed had op de ICC's van de APROCSA-NL features. ICC's werden als volgt geïnterpreteerd: slecht ($r < 0,40$), redelijk ($0,40 \leq r < 0,60$), goed ($0,60 \leq r < 0,75$) en uitstekend ($r \geq 0,75$) (Cicchetti, 1994). Vervolgens werd de invloed van de duur van de natuurlijke spraaksamples en de ernst van de afasie op de betrouwbaarheid van APROCSA-NL scores onderzocht aan de hand van de ICC's van de ervaren groep, ervan uitgaande dat de ervaren groep accuratere APROCSA-NL scores zou opleveren in vergelijking met de groep onervaren studenten.

Casilio en collega's analyseerden in de originele APROCSA studie natuurlijke spraaksamples van minimaal vijf minuten, aangezien dit de aanbevolen sample duur is door Boles & Bombard (1998) om communicatieve effectiviteit bij afasie te kunnen beoordelen. Brookshire & Nicholas (1994) suggereren daarentegen dat samples van 300 à 400 woorden voldoende zouden moeten zijn om betrouwbare informatie over de taalvaardigheden bij afasie te verkrijgen. Aangezien personen met afasie in onze studie een gemiddeld spreektempo van 110 woorden per minuut hadden, zou hun sample ongeveer drie minuten lang zijn om een sample tussen 300 en 400 woorden te bekomen. Daarom evalueerden we de invloed van sample duur op de IBB van de APROCSA-NL features gebruikmakend van zowel de drie als vijf minuten cutoff. We verdeelden de participanten bijgevolg in drie groepen: (1) sample duur van minder dan drie minuten ($n = 29$), (2) sample duur tussen drie en vijf minuten ($n = 14$) en (3) sample duur van meer dan vijf minuten ($n = 15$).

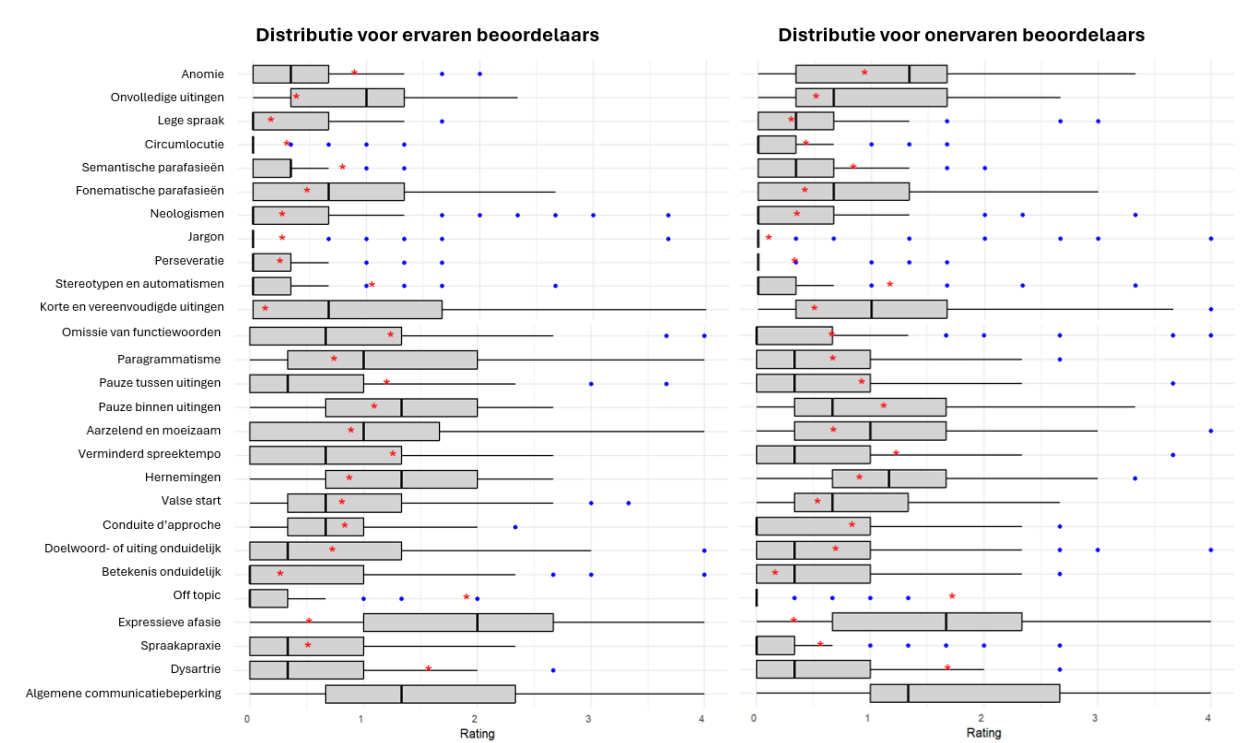
Om de invloed van de afasie-ernst op de IBB na te gaan, werden drie subgroepen gedefinieerd op basis van de ScreeLing en Nederlandse Benoem Test scores (Tabel 3). De eerste groep ($n = 5$) bestond uit personen met ernstige afasie met ScreeLing scores < 41.5 . De tweede groep ($n = 9$) bestond uit personen met matige afasie met ScreeLing scores tussen 42 en 53.5.

De derde groep (n = 35) bestond uit personen met milde afasie met ScreeLing scores tussen 54 en 67.5 of Nederlandse Benoem Test scores lager dan 255 in het geval dat de ScreeLing score 68 of hoger (en dus niet afwijkend) was. Dit laatste criterium werd toegevoegd om te vermijden dat personen met een zuiver anomische afasie over het hoofd gezien zouden worden, gezien de ScreeLing veelal receptieve taaltaken bevat.

Gemiddelde, standaarddeviatie en bereik van de ScreeLing en Nederlandse Benoem Test scores voor de ernstige, matige en milde afasie groepen.

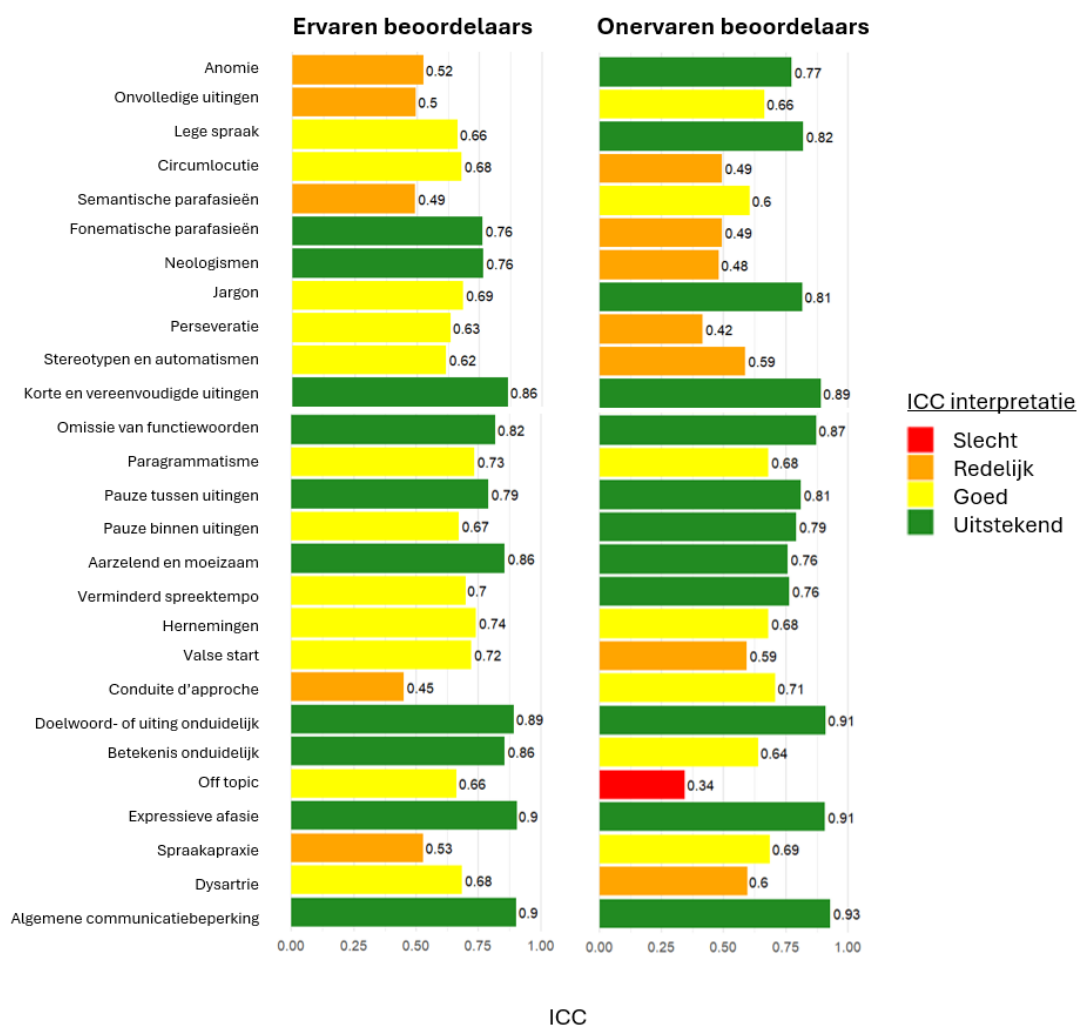
	Ernstig (n = 5)	Matig (n = 9)	Mild (n = 35)
ScreeLing	30.2 ± 9.67 (19-41)	47.2 ± 3.67 (42-53.5)	63.1 ± 4.9 (54-71)
NBT	64.6 ± 79.9 (7-179)	128.6 ± 93.6 (17-233)	234.7 ± 26.7 (141-269)

Onderstaande figuur toont de distributie van de APROCSA-NL scores voor zowel de ervaren als onervaren beoordelaars.



Box plots van de APROCSA-scores voor ervaren (links) en onervaren (rechts) beoordelaars. De rode asterisk duidt het gemiddelde aan, de blauwe punten zijn uitschieters.

Gebaseerd op ICC (A,k), behalen de ervaren beoordelaars een uitstekende interbeoordelaarsbetrouwbaarheid (IBB) voor tien features, een goede IBB voor twaalf features en een redelijke IBB voor vijf features. De onervaren beoordelaars hadden een uitstekende IBB voor twaalf features, een goede IBB voor zeven features, een redelijke IBB voor zeven features en een slechte IBB voor één feature (off topic).



Intraclass correlatie coëfficiënt [ICC (A,k)] en hun interpretatie (kleur-gecodeerd) voor de APROCSA-NL features.

Zelfs bij relatief korte natuurlijke spraaksamples van minder dan drie minuten, vertoonden de meerderheid van de features een goede tot uitstekende IBB gebaseerd op ICC (A,k) van de ervaren beoordelaars (Figuur 2). Anderzijds waren er ook enkele features (bv. stereotypen en automatismen, pauzes binnen en tussen uitingen, conduite d’approche en doelwoord of -uiting onduidelijk) die een relatief lagere ICC hadden voor langere (> 5 min) samples in vergelijking met kortere ($\leq$ 5 min) samples.



Intraclass correlatie coëfficiënt [ICC (A,k)] en hun interpretatie (kleur-gecodeerd) voor de APROCSA-NL features, onderverdeeld op basis van sample duur.

Onderstaande figuur toont dat niet alle APROCSEA-NL features even betrouwbaar zijn voor de verschillende ernstgraden van afasie. Desalniettemin is het een veelbelovende bevinding dat, zelfs voor ernstige afasie, de features jargon, perseveratie, omissie van functiewoorden, pauzes tussen uitingen, aarzelend en moeizaam, herneming, doelwoord of -uiting onduidelijk, betekenis onduidelijk, expressieve afasie en algemene communicatiebeperking een uitstekende IBB vertoonden.



Intraclass correlatie coëfficiënt [ICC (A,k)] en hun interpretatie (kleur-gecodeerd) voor de APROCSEA-NL features, onderverdeeld op basis van afasie-ernst. Er konden geen ICC's berekend worden voor circumlocutie, semantische parafasieën en betekenis onduidelijk in de groep 'Ernstig', aangezien er geen variabiliteit in de data was (alle raters waren het eens over score 0 (= afwezig) voor de vijf personen met ernstige afasie).

Bespreking van de interbeoordelaarsbetrouwbaarheid van APROCSA-NL

Een recente studie vond dat ervaren logopedisten hun betrouwbaarheid in het scoren van de APROCSA kunnen verbeteren met training, wat een veelbelovende bevinding is om de beoordeling door één logopedist te rechtvaardigen (Casilio et al., in press). Daarnaast suggereert de huidige studie dat - wanneer adequate training voorzien wordt - onervaren beoordelaars APROCSA-NL met een gelijkaardige IBB scoren als ervaren beoordelaars.

Zelfs bij relatief korte natuurlijke spraaksamples van minder dan drie minuten, bereikten de meerderheid van de APROCSA-NL features een goede tot uitstekende IBB, wat het gebruik van APROCSA-NL ondersteunt in het geval dat er geen sample van vijf minuten verzameld kon worden. Deze resultaten zijn in lijn met een latere studie van Casilio en collega's (2025), waaruit bleek dat natuurlijke spraaksamples van drie minuten nog steeds resulteerden in een goede IBB. Anderzijds toonden enkele features (bv. pauzes) een relatief lagere ICC voor langere (> vijf minuten) in vergelijking met kortere ($\leq$ vijf minuten) samples, wat zou kunnen suggereren dat deze features suboptimaal zijn voor auditief-perceptuele beoordeling, en best in kaart gebracht worden via een objectieve, automatische analyse. Wat de afasie-ernst betreft, is het veelbelovend dat vijftien APROCSA-NL features nog steeds een goede tot uitstekende IBB behaalden bij ernstige afasie. Samengevat, tonen de APROCSA-NL features een verschillend patroon van interbeoordelaarsbetrouwbaarheid afhankelijk van de duur van het natuurlijke spraaksample en de ernst van de afasie, wat belangrijk is om in het achterhoofd te houden wanneer de APROCSA-NL afgenomen en gescoord wordt.

Concurrente validiteit van APROCSEA-NL

Om de concurrente validiteit van APROCSEA-NL te onderzoeken, werden Benjamini-Hochberg gecorrigeerde Spearman correlaties berekend tussen enerzijds de APROCSEA-NL features (d.i., het gemiddelde van de scores van de drie ervaren beoordelaars) en anderzijds zestien kwantitatieve natuurlijke spraakmaten die berekend werden vanuit transcripties (zie tabel hieronder) en vijf scores op gestandaardiseerde afasie assessments (ScreeLing subtesten semantiek, fonologie en syntaxis, Nederlandse Benoem Test score en CETI score). De APROCSEA-NL features werden op voorhand gelinkt aan de kwantitatieve natuurlijke spraakmaten. Aangezien deze geen grammaticale maten bevatten, werden de grammaticale constructie features van APROCSEA-NL (korte en vereenvoudigde uitingen, omissie van functiewoorden en paragrammatisme) gelinkt aan de subtest syntaxis van de ScreeLing. Op deze manier konden 23 van de 27 APROCSEA-NL features gelinkt worden aan een kwantitatieve tegenhanger om de concurrente validiteit na te gaan.

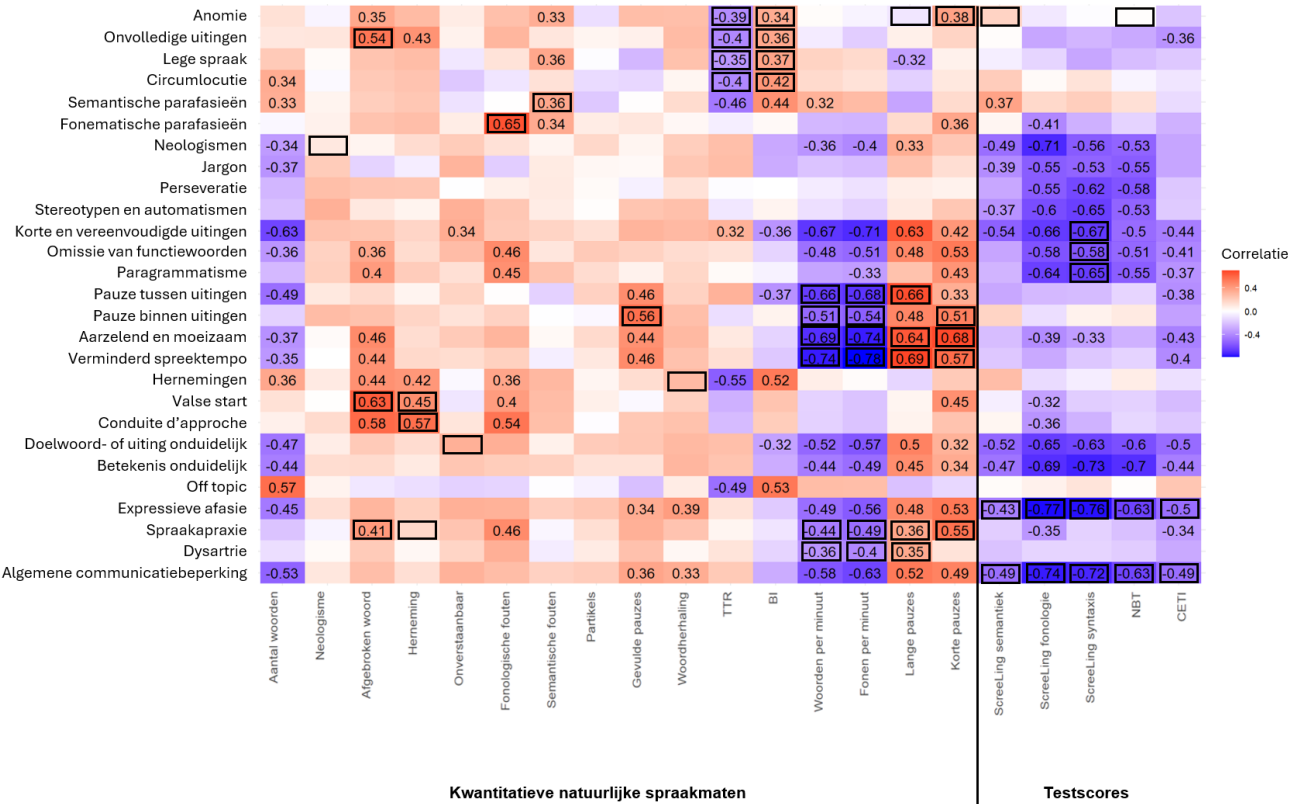
Kwantitatieve natuurlijke spraakmaten	Operationalisatie	APROCSEA-NL tegenhanger(s)
Aantal woorden	Totaal aantal geproduceerde woorden, inclusief non-woorden, fonematische fouten, herhalingen, minimale responsen, commentaren en stereotypen, in overeenstemming met Boxum et al. (2013) en Vandenborre et al. (2018).	/
Neologisme	Non-woorden, meer bepaald woordvormen die geen echte woorden zijn (Casilio et al., 2019), berekend als een proportie van het totaal aantal woorden.	Neologisme

Kwantitatieve natuurlijke spraakmaten	Operationalisatie	APROCSA-NL tegenhanger(s)
Afgebroken woord	Woorden worden afgebroken na één of twee fonemen; het doelwoord wordt niet geproduceerd. Berekend als een proportie van het totaal aantal woorden.	Onvolledige uitingen, valse start, spraakapraxie
Herneming	Een woord wordt initieel afgebroken, maar het doelwoord wordt uiteindelijk wel geproduceerd (bv. ca-camera). Berekend als een proportie van het totaal aantal woorden.	Valse start, conduite d'approche, spraakapraxie
Onverstaanbaar	Overstaanbare woorden, berekend als een proportie van het totaal aantal woorden.	Doelwoord- of uiting onduidelijk
Fonologische fouten	Foneem deletie, insertie, substitutie of transpositie (Casilio et al., 2019; Vermeulen et al., 1989), berekend als een proportie van het totaal aantal woorden.	Fonematische parafasieën
Semantische fouten	Substitutie van inhoudswoorden voor (on)gerelateerde inhoudswoorden (Casilio et al., 2019), berekend als een proportie van het totaal aantal woorden.	Semantische parafasieën
Partikels	Totaal aantal discourse partikels, berekend als een proportie van het totaal aantal woorden.	/

Kwantitatieve natuurlijke spraakmaten	Operationalisatie	APROCSA-NL tegenhanger(s)
Gevulde pauzes	Gevulde pauzes (bv. 'uh'), berekend als een proportie van het totaal aantal woorden.	Pauzes binnen uitingen
Woordherhaling	Onmiddellijke herhaling van een eerder gebruikt woord, berekend als een proportie van het totaal aantal woorden.	Herneming
Type Token Ratio (TTR)	Aantal verschillende woorden (types) gedeeld door het aantal woorden (tokens). Hogere waarden betekenen dat de spreker een meer gevarieerde woordenschat gebruikt (Sanborn et al., 2020).	'Ophalen van woorden' features
Brunet Index (BI)	Meting van lexical diversiteit die onafhankelijk is van de tekstlengte. Dit wordt berekend als $w^{(u^{-0.165})}$ waarbij w = totaal aantal tokens and u = totaal aantal unieke woord types (Parsa et al., 2021). Lagere waarden duiden op een rijkere woordenschat (Sanborn et al., 2020).	'Ophalen van woorden' features
Woorden per minuut	Totaal aantal woorden gedeeld door de totale duur in minuten.	'Tempo en timing' features
Fonen per minuut	Totaal aantal fonen gedeeld door de totale duur in minuten.	'Tempo en timing' features

Kwantitatieve natuurlijke spraakmaten	Operationalisatie	APROCSA-NL tegenhanger(s)
Lange pauzes	Aantal stille segmenten tussen woorden van > 400 ms (Le et al., 2018; Pakhomov et al., 2010), gedeeld door het totaal aantal woorden.	Anomie, pauzes tussen uitingen, aarzelend en moeizaam, verminderd spreektempo, spraakapraxie, dysartrie
Korte pauzes	Aantal stille segmenten tussen woorden tussen 150 en 400 ms (Le et al., 2018; Pakhomov et al., 2010), gedeeld door het totaal aantal woorden.	Anomie, pauzes binnen uitingen, aarzelend en moeizaam, verminderd spreektempo, spraakapraxie

De correlatiematrix met Benjamini-Hochberg gecorrigeerde Spearman correlaties tussen de APROCSEA-NL features en de kwantitatieve natuurlijke spraakfeatures enerzijds en testscores anderzijds zijn weergegeven in de figuur hieronder. Op voorhand werden verschillende APROCSEA-NL features gelinkt aan deze kwantitatieve natuurlijke spraakfeatures en testscores. Deze 59 paren zijn aangeduid met een zwarte omkadering in de correlatiematrix. Van deze 59 paren, waren 51 significant met elkaar gecorreleerd, waarvan er 30 sterk ($|r| \geq 0.50$) met elkaar gecorreleerd waren. Daarnaast konden we 23 van de 27 APROCSEA-NL features linken aan minstens één kwantitatieve natuurlijke spraakmaat of testscore om de concurrente validiteit na te gaan. Van deze 23 APROCSEA-NL features, waren er 19 significant gecorreleerd met hun kwantitatieve tegenhanger, waarvan er 14 sterk ($|r| \geq 0.50$) met elkaar gecorreleerd waren. We vonden geen significante verbanden met de kwantitatieve tegenhangers voor de volgende APROCSEA-NL features: neologismen, herneming en doelwoord of -uiting onduidelijk.



Spearman correlaties tussen de APROCSEA-NL features en de kwantitatieve natuurlijke spraakmaten en testscores. Enkel significante correlaties (Benjamini-Hochberg gecorrigeerde $p < 0.05$) worden weergegeven. Negatieve correlaties zijn aangeduid in het blauw, positieve correlaties in het rood. In het zwart omkaderd werden de a priori gedefinieerde relaties tussen APROCSEA-NL features en hun kwantitatieve tegenhanger aangeduid.

Bespreking van de concurrente validiteit van APROCSEA-NL

De validiteitsanalyse includeerde 23 APROCSEA-NL features die we a priori gelinkt hadden aan een transcriptie-gebaseerde natuurlijke spraakmaat of een testscore. Voor de meerderheid van deze APROCSEA-NL features vonden we een significante correlatie met hun kwantitatieve tegenhanger, wat suggereert dat APROCSEA-NL een valide instrument is om natuurlijke spraak bij afasie in kaart te brengen. We vonden geen significante correlaties voor de volgende APROCSEA-NL features: neologisme, herneming en doelwoord of -uiting onduidelijk. Herneming en doelwoord of -uiting onduidelijk toonden initieel significante correlaties met hun kwantitatieve tegenhanger, maar deze overleefden de correctie voor meerdere vergelijkingen niet. Bijkomend waren er slechts tien deelnemers die zich presenteerden met neologismen, wat hun detectie in de validiteitsanalyse mogelijks verhinderd heeft.

Besluit van het psychometrisch onderzoek

De meeste APROCSEA-NL features vertonen een goede tot uitstekende interbeoordelaarsbetrouwbaarheid voor zowel ervaren als onervaren beoordelaars, en zijn valide metingen van natuurlijke spraak. APROCSEA-NL kan eveneens gebruikt worden bij korte natuurlijke spraaksamples van minder dan drie minuten of bij zeer ernstige afasie, met als belangrijke kanttekening dat niet alle APROCSEA-NL features even betrouwbaar zijn voor alle duren van natuurlijke spraakstalen en ernstgraden van afasie.

Referenties

Armstrong, L., Brady, M., Mackenzie, C., & Norrie, J. (2007). Transcription-less analysis of aphasic discourse: A clinician's dream or a possibility? *Aphasiology*, 21(3-4), 355-374. <https://doi.org/10.1080/02687030600911310>

Azios, J. H., Archer, B., & Lee, J. B. (2021). Detecting behavioural change in conversation: procedures and preliminary data. *Aphasiology*, 35(7), 961-983. <https://doi.org/10.1080/02687038.2020.1812031>

Blomert, L., Kean, M. L., Koster, Ch., & Schokker, J. (1994). Amsterdam—Nijmegen everyday language test: construction, reliability and validity. *Aphasiology*, 8(4), 381-407. <https://doi.org/10.1080/02687039408248666>

Boles, L., & Bombard, T. (1998). Conversational discourse analysis: Appropriate and useful sample sizes. *Aphasiology*, 12(7/8), 547-560.

Boxum, E., Van Der Scheer, F., Van Werven, N., Dirksen, W., & Jonkers, R. (2019). Toevoeging van nieuwe linguïstische en communicatieve maten binnen de huidige ASTA. In *Stem-, Spraak- en Taalpathologie* (Vol. 24, pp. 23-40). Secretariaat cluster Nederlands - Rijksuniversiteit Groningen. <https://doi.org/10.21827/5ce3b9f148e00>

Boxum, E., van der Scheer, F., & Zwaga, M. (2013). Analyse voor Spontane Taal bij Afasie. Vereniging voor Klinische Linguïstiek.

Boyle, M. (2014). Test-retest stability of word retrieval in aphasic discourse. *Journal of Speech, Language, and Hearing Research*, 57(3), 966-978. https://doi.org/10.1044/2014_JSLHR-L-13-0171

Boyle, M. (2015). Stability of word-retrieval errors with the Aphasiabank stimuli. *American Journal of Speech-Language Pathology*, 24(4), S953-S960. https://doi.org/10.1044/2015_AJSLP-14-0152

Brookshire, R. H., & Nicholas, L. E. (1994). Speech Sample Size and Test-Retest Stability of Connected Speech Measures for Adults With Aphasia. *Journal of Speech, Language, and Hearing Research*, 37(2), 399-407. <https://doi.org/10.1044/jshr.3702.399>

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. In *Clinical Linguistics and Phonetics* (Vol. 30, Issue 7, pp. 489-518). Taylor and Francis Ltd. <https://doi.org/10.3109/02699206.2016.1145740>

Bryant, L., Ferguson, A., Valentine, M., & Spencer, E. (2019). Implementation of discourse analysis in aphasia: investigating the feasibility of a Knowledge-to-Action intervention. *Aphasiology*, 33(1), 31-57. <https://doi.org/10.1080/02687038.2018.1454886>

Casilio, M., Dutta, M., Bryan, K., Crouch, K., Dewall, Z., Feld, I., Kile, J., McMahan, K., Samora, J., Shibata, K., Steel, S., Kasdan, A.V., Walljasper, L., de Riesthal, M., & Wilson, S.M. (in press). A clinical training program for auditory-perceptual rating of connected speech in aphasia.

Casilio, M., Kasdan, A. V., Bryan, K., Shibata, K., Schneck, S. M., Levy, D. F., Entrup, J. L., Onuscheck, C., de Riesthal, M., & Wilson, S. M. (2025). Four dimensions of naturalistic language production in aphasia after stroke. *Brain*, 148(1), 291–312. <https://doi.org/10.1093/brain/awae195>

Casilio, M., Rising, K., Beeson, P. M., Bunton, K., & Wilson, S. M. (2019). Auditory-perceptual rating of connected speech in Aphasia. *American Journal of Speech-Language Pathology*, 28(2), 550–568. https://doi.org/10.1044/2018_AJSLP-18-0192

Cavanaugh, R., & Haley, K. L. (2020). Subjective Communication Difficulties in Very Mild Aphasia. *American Journal of Speech-Language Pathology*, 29(1S), 437–448. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0222

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>

Criel, Y., De Groote, E., Benbassat, E., Beirinckx, B., Van Damme, M., Kong, A. P.-H., & De Letter, M. (2025). The main concept analysis: Sensitivity, specificity and inter-rater reliability in Dutch-speaking people with aphasia. *Journal of Communication Disorders*, 116, 106543. <https://doi.org/10.1016/j.jcomdis.2025.106543>

Criel, Y., Deleu, M., De Groote, E., Bockstael, A., Kong, A. P. H., & De Letter, M. (2021). The dutch main concept analysis: Translation and establishment of normative data. *American Journal of Speech-Language Pathology*, 30(4), 1750–1766. https://doi.org/10.1044/2021_AJSLP-20-00285

Dietz, A., & Boyle, M. (2018). Discourse measurement in aphasia research: have we reached the tipping point? In *Aphasiology* (Vol. 32, Issue 4, pp. 459–464). Routledge. <https://doi.org/10.1080/02687038.2017.1398803>

Fromm, D., Forbes, M., Holland, A., Dalton, S. G., Richardson, J., & MacWhinney, B. (2017). Discourse Characteristics in Aphasia Beyond the Western Aphasia Battery Cutoff. *American Journal of Speech-Language Pathology*, 26(3), 762–768. https://doi.org/10.1044/2016_AJSLP-16-0071

Hallowell, B. (2016). Aphasia and Other Acquired Neurogenic Language Disorders: A Guide for Clinical Excellence. In *ProtoView* (Vol. 3, Issue 11). Ringgold, Inc.

Holland, A., Forbes, M., Fromm, D., & MacWhinney, B. (2019). Communicative Strengths in Severe Aphasia: The Famous People Protocol and Its Value in Planning Treatment. *American Journal of Speech-Language Pathology*, 28(3), 1010–1018. https://doi.org/10.1044/2019_AJSLP-18-0283

Kim, H., Berube, S., & Hillis, A. E. (2022). Core lexicon in aphasia: A longitudinal study. *Aphasiology*. <https://doi.org/10.1080/02687038.2022.2121598>

Kim, H., & Wright, H. H. (2020). Concurrent validity and reliability of the core lexicon measure as a measure of word retrieval ability in aphasia narratives. *American Journal of Speech-Language Pathology*, 29(1), 101–110. https://doi.org/10.1044/2019_AJSLP-19-0063

Le, D., Licata, K., & Mower Provost, E. (2018). Automatic quantitative analysis of spontaneous aphasic speech. *Speech Communication*, 100, 1–12. <https://doi.org/10.1016/j.specom.2018.04.001>

Leaman, M. C., & Edmonds, L. A. (2021). Assessing Language in Unstructured Conversation in People With Aphasia: Methods, Psychometric Integrity, Normative Data, and Comparison to a Structured Narrative Task. <https://doi.org/10.23641/asha>

Lesser, R., & Algar, L. (1995). Towards combining the cognitive neuropsychological and the pragmatic in aphasia therapy. *Neuropsychological Rehabilitation*, 5(1–2), 67–92. <https://doi.org/10.1080/09602019508520176>

Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (1989). The communicative effectiveness index: Developmental and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of Speech and Hearing Disorders*, 54(1), 113–124. <https://doi.org/10.1044/jshd.5401.113>

Pakhomov, S. V. S., Smith, G. E., Chacon, D., Feliciano, Y., Graff-Radford, N., Caselli, R., & Knopman, D. S. (2010). Computerized analysis of speech and language to identify psycholinguistic correlates of frontotemporal lobar degeneration. *Cognitive and Behavioral Neurology*, 23(3), 165–177. <https://doi.org/10.1097/WNN.0b013e3181c5dde3>

Parsa, M., Alam, M. R., & Mihailidis, A. (2021). Towards AI-powered Language Assessment Tools. Pritchard, M., Hilari, K., Cocks, N., & Dipper, L. (2018). Psychometric properties of discourse measures in aphasia: acceptability, reliability, and validity. *International Journal of Language and Communication Disorders*, 53(6), 1078–1093. <https://doi.org/10.1111/1460-6984.12420>

Ruiter, M. B., Otters, M. C., Piai, V., Lotgering, E. A. M., Theunissen, J. E. M. C., & Rietveld, T. C. M. (2022). A transcription-less quantitative analysis of aphasic discourse elicited with an adapted version of the Amsterdam-Nijmegen Everyday Language Test (ANELT). *Aphasiology*, 1–20. <https://doi.org/10.1080/02687038.2022.2109124>

Sanborn, V., Ostrand, R., Ciesla, J., & Gunstad, J. (2020). Automated assessment of speech production and prediction of MCI in older adults. *Applied Neuropsychology:Adult*. <https://doi.org/10.1080/23279095.2020.1864733>

Stark, B. C. (2023). Historical Review of Research in Discourse Deficits and Its Recent Advancement. In *Spoken Discourse Impairments in the Neurogenic Populations* (pp. 3–22). Springer International Publishing. https://doi.org/10.1007/978-3-031-45190-4_1

- Stark, B. C., & Dalton, S. G. (2024). A scoping review of transcription-less practices for analysis of aphasic discourse and implications for future research. *International Journal of Language & Communication Disorders*, 59(5), 1734–1762. <https://doi.org/10.1111/1460-6984.13028>
- Stark, B. C., Dutta, M., Murray, L. L., Fromm, D., Bryant, L., Harmon, T. G., Ramage, A. E., & Roberts, A. C. (2021). Spoken discourse assessment and analysis in aphasia: an international survey of current practices. *Journal of Speech, Language and Hearing Research (JSLHR)*, 64(11), 1–24. https://doi.org/10.1044/2021_JSLHR-20-00708
- Stark, B. C., & Fukuyama, J. (2021). Leveraging big data to understand the interaction of task and language during monologic spoken discourse in speakers with and without aphasia. *Language, Cognition and Neuroscience*, 36(5), 562–585. <https://doi.org/10.1080/23273798.2020.1862258>
- Van Der Scheer, F., Zwaga, M., & Jonkers, R. (2011). Normering van de ASTA, analyse voor spontane taal bij afasie. *Stem-, Spraak- En Taalpathologie*, 17(2), 19–30.
- Van Ewijk, L., Dijkhuis, L., Hofs-van Kats, M., Hendrickx, M., Wijngaarden, M., & De Hilster, C. (2018). Nederlandse benoemtest (NBT). Bohn Stafleu van Loghum.
- Vandenborre, D., Cornette, C., & Visch-Brink, E. (2021). CAT-NL White paper: Conceptuele Analyse Situatieplaat.
- Vandenborre, D., Visch-Brink, E., van Dun, K., Verhoeven, J., & Mariën, P. (2018). Oral and written picture description in individuals with aphasia. *International Journal of Language and Communication Disorders*, 53(2), 294–307. <https://doi.org/10.1111/1460-6984.12348>
- Vermeulen, J., Bastiaanse, R., & Van Wageningen, B. (1989). Spontaneous Speech in Aphasia: A Correlational Study. *Brain and Language*, 36(2), 252–274.
- Visch-Brink, E. G., van de Sandt-Koenderman, M., & El Hachoui, H. (2010). ScreeLing. Bohn Stafleu Van Loghum.
- Visch-Brink, E., Vandenborre, D., de Smet, H. J., Mariën, P., Swinburn, K., Porter, G., & Howard, D. (2015). Comprehensive Aphasia Test - Nederlandstalige bewerking (CAT-NL). Pearson.
- Wallace, S. E., & Kimelman, M. D. Z. (2013). Generalization of word retrieval following semantic feature treatment. *NeuroRehabilitation*, 32, 899–913. <https://doi.org/10.3233/NRE-130914>