



Proef ingediend met het oog op het behalen
van de graad van Educatieve Master Talen

CHATGPT ZET DE TAALLEERKRACHT (NOG NIET) BUITENSPEL

**Studie naar het gebruik van generatieve AI bij
taalleerkrachten in het volwassenenonderwijs**

JULIE SIMOENS
2023-2024

Promotor: Prof. Dr. Jo Tondeur
Multidisciplinair Instituut Lerarenopleiding



Ik verklaar plechtig dat ik de masterproef, "**ChatGPT zet de taalleerkracht (nog niet) buitenspel**", zelf heb geschreven.
Ik ben op de hoogte van de regels i.v.m. plagiaat en heb erop toegezien om deze toe te passen in deze masterproef.

03/06/2024, Naam en handtekening
Julie Simoens

Julie Simoens

Dankwoord

Tijdens dit intense jaar hebben we een werkplekleertaak en een praktijkonderzoek uitgevoerd. Het onderzoeksthema betrof het gebruik van generatieve AI bij de taalleerkrachten van het Centrum voor Volwassenenonderwijs (CVO). Dit was een collectieve realisatie en we vinden het belangrijk om alle mensen die er substantieel toe bijgedragen hebben daarvoor te bedanken.

In eerste instantie werden wij, een groepje van drie VUB-studenten, vakkundig en inspirerend begeleid in het hoogtechnologische onderwerp door onze promotor, Prof. Dr. Jo Tondeur. Hij heeft ons niet enkel aan zeer actuele bronnen geholpen, maar daagde ons ook steeds uit om de technologie op een praktijkgerichte manier te benaderen, wat de resultaten van het onderzoek zeker ten goede gekomen is. Zijn expertise in de gedragswetenschappen zorgde ervoor dat we methodologisch correct konden werken.

We willen in tweede instantie een warm dankwoord richten aan de CVO-collega's: Hans, Muriel, Sylvia en Vera. We werden hartelijk onthaald in jullie team, en jullie hebben ook een heel actieve rol opgenomen in het onderzoek: jullie verstuurd bevestigingen, spraken collega's aan, begeleidten focusgroepen, maakten transcripties en controleerden coderingen. We vonden het heel aangenaam samenwerken en we konden zo de deadlines en de kwaliteitscriteria respecteren.

Tenslotte willen we ook een speciaal dankwoordje richten aan elkaar. Teamwerk is intens, niet enkel door het produceren van een werkstuk, maar ook door het bouwen van een relatie. We zijn naar elkaar toegegroeid en echt op kruissnelheid gekomen een paar maanden terug. Bedankt aan elk van jullie om de laatste sprint met zoveel goede wil en toewijding gedaan te hebben!

Julie, Katrien en Nele.

Abstract

Het artikel onderzoekt het gebruik van generatieve artificiële intelligentie (GenAI) in het volwassenenonderwijs, in het bijzonder door de taalleerkrachten. De introductie van GenAI, waaronder bekende chatbots zoals ChatGPT, roept belangrijke vragen op over de rol en relevantie van taalleerkrachten in een sterk technologisch gedreven leeromgeving.

Het onderzoek was praktijkgericht, met een gemengd onderzoeksteam waar zowel academici als praktijkexperten aan deelnamen.

GenAI kan worden onderverdeeld in twee grote categorieën: *machinelearning*-systemen die vaste procedures volgen, en *deeplearning*-systemen die zich aanpassen aan een gegeven context en feedback. Hoewel GenAI-systemen drijven op krachtige taalmodellen, blijft menselijke intelligentie momenteel nog essentieel voor de interpretatie en toepassing van deze technologieën.

De onderzoeksmethode hanteerde focusgesprekken met twaalf taalleerkrachten om hun persoonlijke visies en praktijkervaringen met GenAI te verzamelen. De resultaten toonden een overwegend positieve houding en een groeiend gebruik, waarbij GenAI vooral werd ingezet ter voorbereiding van lessen, zoals het genereren van dictees en het aanpassen van teksten aan het beginniveau van de cursisten. De gegenereerde oefeningen viseerden de ontwikkeling van de vier taalvaardigheden (lezen, schrijven, luisteren, spreken), alsook van de ondersteunende taalsystemen (grammatica, woordenschat).

Er kwam kritiek op de betrouwbaarheid en de onnatuurlijke taal van sommige AI-gegenereerde antwoorden. Dit aspect was een breekpunt voor bepaalde talige toepassingen, en er was een grote consensus in de groep dat menselijke controle van de GenAI producten onmisbaar blijft.

De conclusie is dat hoewel GenAI nuttig is voor het creëren van lesmateriaal en het inspireren van nieuwe werkvormen, het de rol van de taalleerkracht niet overbodig maakt. De technologie biedt ondersteuning, maar de menselijke *touch* blijft cruciaal voor effectief en betekenisvol onderwijs. Toekomstige onderzoeksrichtingen zouden kunnen focussen op de ervaringen van cursisten en de verdere ontwikkeling van GenAI om tot betrouwbaardere en nauwkeurigere resultaten te komen binnen de praktijk van het taalonderwijs.

Gebruik van GenAI

We hebben voor dit werkstuk in meer of mindere mate gebruik gemaakt van generatieve AI. We maakten gebruik van Open AI (chat.openai.com) als hulp bij de oriëntatie over het onderwerp, om associaties te vinden met praktijkgerichte toepassingen, om experimenten uit te voeren alsook bij de verklaring van complexe technologische concepten. We hebben de output enkel gebruikt als inspiratiebron, het schrijfstuk is door onszelf geproduceerd. De taalkundige checks gebeurden door de Word-geïntegreerde software en door onszelf.

ChatGPT zet de taalleerkracht (nog niet) buitenspel

Generatieve AI vindt zijn intrede in het volwassenenonderwijs, maar het vraagt wel om een uitgekiende balans tussen technologie en de menselijke *touch* van de taalleerkrachten.

Beeld je in: volgend jaar wil je die promotie naar een leukere job bekomen, maar daarvoor moet je beter Frans leren spreken. Wat als je toekomstige taalleerkracht geen koffie nodig heeft, nooit nors is en altijd klaarstaat om je verder te helpen? Laat je niet bedotten. Dit is de beschrijving van Odile, de leerkracht Frans. En ChatGPT is slechts haar slimme knecht.

Een virtuele aftrap

Onze eerste kennismaking met de taalleerkrachten van het Centrum voor Volwassenenonderwijs (CVO) verliep online. We zouden uitwisselen over het centrale thema van ons gezamenlijke onderzoeksproject: het gebruik van generatieve artificiële intelligentie in het talenonderwijs. Er hing een sfeer van warme, menselijke openheid in de Zoom-ruimte, het voelde aan alsof we van bij meet af aan de technologie wouden tarten.

We deelden samen het groeiende gevoel van urgentie over een nieuwe technologie die onze woonkamers en werkplekken niet binnensloop maar -stormde. Deze innovatie wierp bijna existentiële vragen op, die als een mist tussen onze Zoom-lijnen hingen: wat gaat de rol van taalleerkrachten nog behelzen? Zullen cursisten beter vreemde talen kunnen leren met de nieuwe generatie van artificieel gestuurde taalapplicaties? Heeft het eigenlijk nog wel zin om voor taalleerkracht te studeren?

De hoge snelheid waarmee de nieuwe ontwikkelingen publiek gemaakt werden, bepaalde onze agenda. We zouden ons buigen over het gebruik van generatieve artificiële intelligentie (gemakshalve GenAI genoemd) en hoe die de onderwijspraktijk van de meer dan 250 taalleerkrachten van het CVO zou beïnvloeden. De voorbije maanden kregen ze soms berichten van ongeruste leerkrachten die reclamespots op sociale media gezien hadden, waar merken van taalapplicaties beweerden dat hun GenAI-oplossing de taalleerkracht overbodig zou maken...

GenAI voor dummies

Het is verleidelijk om de notie GenAI te gebruiken bij elk type van technologie dat mysterieus en slim lijkt. Als taalkundigen wouden we Babylonische toestanden vermijden in

het onderzoek en hebben we duidelijkheid geschept rond het centrale begrip. Of beter: relatieve duidelijkheid, want zelfs binnen het kennisdomein bestaat er geen eenduidige definitie over GenAI en andere intelligente technologieën (Holmes & Tuomi, 2022, p. 546).

We zochten naar definities van GenAI in de wetenschappelijke literatuur en weerhielden een paar verschillende perspectieven van wetenschappers uit de V.S., Europa en China. Wat daarin naar boven kwam, is het feit dat AI-systemen door de mens ontworpen zijn maar niet als doel hebben om de menselijke intelligentie na te bootsen. Ze beschikken over een eigen vorm van intelligentie, waarbij de systemen grote hoeveelheden aan gegevens verwerken op manieren die hen eigen zijn. Hier onderscheiden we twee grote categorieën.

Sommige systemen volgen vaste digitale procedures (algoritmen) op duidelijk afgebakende gegevens. Die brengen we thuis onder *machine learning* systemen. Denk hierbij aan Netflix, dat je aanbevelingen doet in je favoriete filmgenre, of aan Alexa, aan wie je een recept voor lekkere pannenkoeken vraagt. Ook taalapplicaties waarmee je je woordenschat in het Duits traint of vertalingen maakt, kunnen we onderbrengen in de categorie van machinaal leren. Deze systemen beschikken over vormen van intelligentie, m.a.w. ze verwerken en interpreteren gegevens, maar genereren daarom nog geen nieuwe inzichten.

Andere systemen doen dat wel. Ze beschikken over de mogelijkheid om hun analyse van gegevens, om hun methode van verwerking én hun uitkomsten aan te passen aan een context. Dat is het lerende aspect van de generatieve systemen, ze creëren nieuwe manieren om een welbepaald doel te bereiken. Elke wijze die ze gebruiken, slaan ze op. Elk resultaat dat ze behalen wordt geëvalueerd om eruit te leren. Elke feedback die ze krijgen, draagt bij tot dit proces. Je kan het vergelijken met neurale netwerken zoals die in onze hersenen, maar dan met een eigen logica. In tegenstelling tot het machinale leren, spreekt men voor deze systemen van *deep learning*, ofwel diep leren. Deze technologieën worden gebruikt voor processen die complexer zijn dan het aanbevelen van gepaste liedjes voor een romantisch diner. Indien je een systeem wil bouwen dat in staat is om te leren, heb je een gigantische hoeveelheid aan gegevens en dus ook een super krachtige verwerkingseenheid nodig.

De meest gekende van deze toepassingen bevinden zich op het gebied van de spraak- en tekstherkenning, waardoor het systeem in interactie kan gaan met mensen. Ze worden bestuurd door modellen die ervoor zorgen dat het systeem een bijna natuurlijke taal kan nabootsen. Die worden de “grote taalmodellen” genoemd in het vakjargon. Ze gebruiken

precies en creatief de vier taalvaardigheden: lezen, schrijven, luisteren en spreken, en dit in zoveel talen als er formeel bestudeerd worden.

Iedereen heeft intussen al gehoord van de befaamde chatbot ChatGPT, of diens soortgenoten Copilot, Google Bard en Gemini. Velen hebben ook al geëxperimenteerd met deze chatbots. Herinner je je nog de eerste keer toen je met open mond zat te staren naar het scherm, waar een uitgebreid antwoord op je vraag aan een snel tempo verscheen? Er bestaan ook systemen die het diepe leren toepassen op het verwerken en creëren van beeldgegevens, zoals DALL-E en Midjourney. Het principe blijft hetzelfde, enkel voedt de intelligentie van het systeem zich dit keer niet met talige gegevens, maar wel met beeld- en videomateriaal om er dan nieuwe vormen aan te kunnen geven.

GenAI in het taalonderwijs

Volgens ChatGPT hebben we zeven maal het woord “taal”, of een afgeleide ervan gebruikt bij de uitleg over GenAI. Dit is een voor de hand liggend bruggetje naar onze taalleerkrachten van het CVO. Indien er instrumenten bestaan die op zo’n krachtige manier met talen kunnen omspringen, waarom zouden we die dan niet massaal inzetten bij onze onderwijspraktijken? Het klinkt veelbelovend: een systeem dat in geen tijd een volledige taal cursus uit de virtuele mouw kan schudden, verrijkt met woordenschat- en grammatica oefeningen, met voorstellen om in dialoog te gaan, met persoonlijke feedback over je spreek- en schrijfvorderingen... Dit is momenteel de marketing-praat bij uitstek van menige taalapplicatie (Duolingo, Babbel, Busuu...), en die voor wat onrust zorgde bij het doelpubliek.

De recente en snelle opkomst van GenAI gestuurde chatbots was bepalend bij het aflijnen van de onderzoeksvraag. Het thema wakkert vurige debatten aan, maar we merkten snel dat we heel weinig concrete informatie hadden over het huidige gebruik van GenAI binnen de lespraktijken. De onderzoeksvraag diende zich op een natuurlijke wijze aan: hoe kunnen taalleerkrachten in het volwassenenonderwijs GenAI inzetten in hun onderwijspraktijk?

De wetenschap kon zich pas recent buigen over de verschillende toepassingen van GenAI, waardoor er momenteel vooral beschrijvende studies en experimenten te vinden zijn. In een recente metastudie van Thùy et al. (2024), maken de onderzoekers een analyse van hoe ChatGPT gebruikt wordt in onderwijs- en leerprocessen. Ze screenen in totaal 819 studies die gepubliceerd werden tussen 2022-2023, waarvan ze er na een selectie 51 inhoudelijk doorlichtten a.d.h.v. een SWOT-analyse. Geen van deze studies bevatte gegevens over het

volwassenenonderwijs, ze gingen in hoofdzaak over het secundair en hoger onderwijs. Een gelijkaardige literatuurstudie van de British Council bevestigt dit gegeven. Ze selecteerden 369 artikels gepubliceerd tussen 2014 en 2023 over het gebruik van AI in Engels taalonderwijs. Van de 43 weerhouden artikels was er amper één dat betrekking had op het volwassenenonderwijs (Crompton et al., 2023, p 8).

Intussen probeerden we met een schuchtere poging om een eerste beeld te krijgen van hoeveel taalleerkrachten er bij CVO reeds met GenAI geëxperimenteerd hadden. We nodigden hen uit voor een korte bevraging, maar daar kwam slechts 10% respons op, te weinig om algemene conclusies uit te trekken.

Terwijl we nog aan het redetwisten waren over de gepaste onderzoeksmethode, werden we schaamteloos ingehaald door de evoluerende realiteit. Een paar goed gedocumenteerde taalleerkrachten van het CVO gaf in een lesvrije week enkele workshops aan hun collega's over het gebruik van GenAI. Onze onderzoeksvraag bleef echter nog steeds relevant.

Een keuze voor diep leren

We opteerden uiteindelijk voor het verzamelen van kwalitatieve gegevens om een antwoord te vinden op onze onderzoeksvraag. Het nieuwbakken onderzoeksthema deed ons vermoeden dat er maar weinig leerkrachten echt beslagen zouden zijn in het gebruik van GenAI voor hun onderwijspraktijken. Na de beperkte respons op de online bevraging besloten we om de drempel wat te verlagen. We kozen voor de onderzoeksmethode van het focusgesprek om deelnemers met verschillende achtergronden en ervaring met elkaar in interactie te laten gaan. We waren geïnteresseerd in de “authentieke stemmen uit de leraarskamer”, er mocht met uiteenlopende blikken en overtuigingen gekeken worden naar de nieuwe technologie. Deze methode liet ons toe om informatie te verkrijgen over zowel feiten als persoonlijke overtuigingen.

We hielden twee focusgesprekken met in totaal twaalf deelnemers, allen taalleerkracht bij het CVO. Sommigen onder hen hadden nog geen GenAI gebruikt in hun lespraktijk, anderen experimenteerden er lustig op los. We werkten eerst deductief en hanteerden een vooraf bepaalde vraagstructuur. Er waren thema's waarover we gegevens wilden verzamelen: welke is je persoonlijke visie over GenAI? Hoe vaak heb je het al gebruikt in je lespraktijk? Voor welke onderwijspraktijken? Voor welke taalverwervingsprocessen? Welke voor- en nadelen

ervaar je? ... Anderzijds waren we ook geïnteresseerd in een meer spontane uitwisseling tussen de leerkrachten om te horen wat er bij hen leefde.

Onze dataverwerking werd daarop geënt. In de eerste ronde codeerden we de gegevens die betrekking hadden op de voorgelegde structuur. In een tweede fase bogen we ons over de data die we er niet onmiddellijk in konden onderbrengen. We vulden de structuur aan met nieuwe thema's en kregen zo een globaal en gedetailleerd overzicht op de inhoud van de focusgesprekken.

“De blauwe of de rode pil?”

Vooreerst wouden we kennismaken met de persoonlijke visie van de taalleerkrachten op GenAI. De meesten van hen hadden een praktische, positieve visie die belichaamd werd in de figuur van Iron Man. Tony Stark, de alledaagse alias van deze superheld, is een ingenieur en uitvinder die GenAI gebruikt om nieuwe technologieën te creëren die hem helpen om wereldse problemen op te lossen. Ze kozen ook voor de figuur van Meester Yoda, de groene ruimtetrol die uitblinkt in kennis en wijsheid, en het vermogen van GenAI symboliseert om ons te helpen navigeren in deze complexe wereld. Een enkele deelnemer koos voor Rachael, de robot met een menselijk uiterlijk uit de film Blade Runner. Ze werkte met mensen met een autismespectrumstoornis en geloofde erin dat GenAI hen zou helpen om beter menselijke emoties te begrijpen. Daarmee zat ze niet zo ver verwijderd van de visie van Sam Altman, het brein achter OpenAI, die op 13 mei 2024 een enigmatische tweet (X.com) de wereld in stuurde die maar één woord bevatte: “Her”. Hij verwees naar de titel van een visionaire film uit 2013 waar het hoofdpersonage verliefd wordt op een virtuele – lees: artificieel gegenereerde - assistente. In die toekomstvisie communiceert GenAI niet enkel via menselijke taal, maar gaat die ook in interactie met onze menselijke emoties. Al deze visies vertaalden de potentiële krachten van GenAI.

Slechts enkelen kozen de kant van Neo, de held van de Matrix-trilogie. Hij is een kei in het besturen van complexe technologieën, maar gelooft bovenal in het menselijke kunnen en in de vrijheid van handelen, los van enige sturing door krachtige algoritmen. Neo kijkt kritisch naar de technologie die hij gebruikt en is zich bewust van de mogelijke gevaren. Geen van de deelnemers opteerde voor het doemdenken van HAL9000, de supercomputer uit Kubricks “2001, A Space Odyssey”, die uiteindelijk helemaal doorslaat en ronduit gevaarlijk wordt.

Dankzij deze typetjes uit de populaire filmwereld startten we op een laagdrempelige manier het focusgesprek. Naar het einde van het gesprek begonnen de deelnemers spontaner uit te wisselen en kregen we nieuwe elementen te horen die meer beducht waren voor GenAI. We brachten die onder in de categorie “kritisch bezorgde visies”. Er werd gesproken over de digitale kloof in onze maatschappij, die met de komst van GenAI mogelijks nog dieper zou worden. Ze verwezen ook naar de volgende generaties en het belang van de ontwikkeling van het kritische denken om te vechten tegen intellectuele luiheid. We hoorden ook dat de steeds meer sturende algoritmen ons massaal naar geijkte paden zouden leiden, wat aanvoelde als een beperking van onze vrijheid van denken en handelen.

Wij concludeerden dat de deelnemers een overwegend positieve visie hadden van GenAI, maar dat ze ook kritisch waren voor de huidige en toekomstige toepassingen van de technologie.

Ontwikkeling van taalvaardigheden

Die genuanceerde visie trok zich door naar de discussie over de ervaringen van de taalleerkrachten met de beloftevolle technologie.

Het huidige gebruik van GenAI bij de onderwijspraktijken van de taalleerkrachten toonde aan dat ze in hoofdzaak grijpen naar de grote taalmodellen, en meer specifiek naar ChatGPT. De leerkrachten geven aan de technologie te gebruiken voor de ontwikkeling van de verschillende taalvaardigheden (spreken, schrijven, lezen, luisteren), zoals voorzien in de leerplannen, alsook voor de ondersteunende taalsystemen (grammatica, woordenschat).

Bij de inhoudelijke bepaling van deze vaardigheden, merkten we op dat GenAI vooral gebruikt wordt ter voorbereiding van situaties waar de vaardigheden ingezet worden tijdens de les. Een paar citaten van de leerkrachten zullen dit duidelijker stellen. “Voor schrijven heb ik al zinnnetjes laten genereren voor een dictee”, zegt Kris. “Ik gebruik ChatGPT voor het creëren van interessante spreekopdrachten” stelt Emma. Voor de leesvaardigheid wordt er verwezen naar de mogelijkheid van GenAI om een tekst aan te passen aan het beginniveau van de cursist. Over schrijfvaardigheid is er al wat meer onenigheid in de groep. Sommige leerkrachten moedigen de cursisten aan om via GenAI inspiratie te vinden voor schrijfoopdrachten, anderen storen zich aan het feit dat cursisten hun volledige schrijftaak door ChatGPT laten maken: “Wat is dan nog het nut van de taak? Als ik tien teksten krijg die perfect zijn, wat heb ik daaraan? Niks. Dan hoef ik de taak ook niet te geven” oppert Melissa

stellig. Sommigen onder hen pleitten voor een terugkeer van de klassikale schrijfoopdrachten, een beweging die niet echt mogelijk bleek te zijn voor de online leertrajecten.

Deze belichting contrasteert wat met de definitie die we vonden in de studie van British Council (Edmett et al., 2023, p. 15-17), een bevraging naar het gebruik van AI in taalonderwijs. Ook zij verwijzen naar de taalvaardigheden, maar focussen op de manier waarop AI die kan versterken tijdens de leermomenten zelf. Voor spreekvaardigheid was uitspraak een belangrijk criterium, en werd er o.a. AI-technologie ingezet om die te corrigeren. Voor de ontwikkeling van de schrijfvaardigheid registreren ze een overwicht van grammatica- en woordenschatoefeningen, iets wat de taalleerkrachten van het CVO ook uitdrukkelijk aangaven. Voor de ontwikkeling van de leesvaardigheid werd er dan weer enkel over woordenschatonderwijs gerept. Een interessant gegeven is dat ze veel minder voorbeelden van het gebruik van AI vonden voor de zogenaamde receptieve taalvaardigheden (lezen en luisteren) dan voor de productieve (spreken en schrijven). Voor de ontwikkeling van de luistervaardigheid werd er nog geen AI-toepassing gebruikt.

We voegen daarmee een kritische kanttekening bij de conclusie over het gebruik van GenAI voor de taalvaardigheden. In onze bespreking met de focusgroepen hebben de taalleerkrachten vooral voorbeelden aangehaald van situaties ter voorbereiding van een vaardigheid die behandeld zou worden tijdens de les. Ze focusten op de ondersteunende taalsystemen zoals grammatica- en woordenschatoefeningen. Het leermoment tijdens de les zelf was nog in de regie van de taalleerkracht, er werd geen GenAI gebruikt als live ondersteuning van bv. een spreekopdracht. Op vandaag helpt GenAI de taalleerkrachten bij het uitwerken van een gepaste leercontext, net zoals het gebruik van handboeken en audiovisueel materiaal.

Didactische vitaminekuur

Er was grote eensgezindheid binnen de focusgroepen over het nut van GenAI bij het vinden van inspiratie voor nieuwe en alternatieve werkvormen, de combinaties zijn eindeloos. “Ik gebruik ChatGPT om zinnetje te maken. Ik vraag dan voor het perfectum en imperfectum, NT2, ik geef een specifiek niveau en een aantal werkwoorden, en hij maakt een invuloefening voor me” poneert Emma. Andere collega’s vermelden de creatie van dialogen om een spreekoefening te begeleiden, nog anderen vermelden mogelijkheden om het creatief schrijven te stimuleren.

Het gebruik van GenAI bij specifieke didactische processen, zoals het evalueren, zorgt voor een meer kritische houding. Of het nu gaat over het voorbereiden van evaluaties of als hulp bij het evalueren zelf, telkens wordt er vermeld dat de huidige systemen nog niet betrouwbaar genoeg zijn om deze processen zonder menselijke controle te laten uitvoeren. Maarten zegt hier: “Als we gaan denken dat de technologie automatisch kan corrigeren, dan zal het nog even duren eer dat correct zal kunnen verlopen”. Zijn buikgevoel vertaalt wat de onderzoekers van de British Council ook meegaven: “Assesment in English Language Teaching is an area in need of further research” (Edmett et al., 2023, p. 57). Sommige leerkrachten hadden al een paar kleine tests uitgevoerd op de lagere taalbeheersingsniveaus, maar die zijn niet concluderend. Hun bevinding wordt bevestigd in het rapport: “AI can create assessment tasks, but for lower-level cognition, and results so far are not overly impressive” (Edmett et al., 2023, p.58).

De leerkracht blijft de regisseur van de behandelde leerdoelen en werkvormen in de les. Hij/zij kan beroep doen op een indrukwekkend arsenaal aan didactische variaties dankzij GenAI, maar moet nog steeds de coherentie van het leerproces bewaken.

Voorzichtig enthousiasme

Wanneer we de taalleerkrachten vragen om een balans op te maken van de voor- en nadelen van het gebruik van GenAI, zijn de antwoorden in evenwicht.

Bij de voordelen horen we twee aspecten veelvuldig terugkomen. Het eerste hoeft niet te verwonderen, GenAI wordt geprezen voor de tijdswinst die leerkrachten kunnen maken bij het opmaken van oefeningen en het zoeken naar inspiratie. Dit aspect is een doorslaggevende factor in een beroep dat onder druk staat door personeelstekorten en toenemende administratieve druk. Het tweede voordeel is gelinkt aan de natuur van het systeem zelf: het creatieve vermogen van de taalmodellen doet menig taalkundig hart sneller slaan. Lena verwoordt het als volgt: “Je kan die vragen om een essay te schrijven op Shakespeareaanse wijze en die doet dat subliem!”.

Er zit een schaduwkant aan het geweldige potentieel van GenAI: indien je het doelgericht en efficiënt wil inzetten, kan je maar best een kenner zijn in het schrijven van *prompts*, instructies die je geeft aan de chatbot. Wanneer die niet voldoen aan bepaalde criteria (volledigheid, context, duidelijkheid...), is de kwaliteit van het antwoord van je chatbot dan ook navenant.

Wanneer we de nadelen aankaarten, verschijnt de olifant in de kamer: de betrouwbaarheid van ChatGPT (of andere chatbots) is ver beneden het peil dat onze taalleerkrachten verwachtten. Victoria geeft uiting aan haar frustratie: “het antwoord komt heel snel, maar er zitten altijd fouten in, en heel flagrante fouten waar je zou overkijken omdat het met zo’n stelligheid geponeerd wordt!”. Er was geen enkele situatie waarin de leerkrachten zonder menselijke controle een puur GenAI product op de cursisten wouden loslaten. Maarten merkt zelfs een neerwaartse trend in kwaliteit: “In het begin vond ik dat ChatGPT me echt verraste met vaak goede vragen. Ik heb die oefeningen een paar maanden later opnieuw gedaan, en dan kreeg ik antwoorden met fouten in”. De variabele kwaliteit per taal (in het Engels zijn de resultaten beter), is ook een storende factor. Het onnatuurlijke taalgebruik van de favoriete chatbot wordt een paar keer vermeld als een beperking. De zeer gestandaardiseerde taal doet vragen opwerpen rond het niveau van authenticiteit van teksten en de normen die erin gehanteerd worden (wat is correct en wat niet?).

Het gebruik van GenAI door de cursisten wordt dubbelzijdig belicht. Enerzijds is het een nadeel omdat ze meer en beter kunnen valsspelen. Aan de andere kant is het herkenbare taalgebruik van de chatbots een voordeel om knoeiers te identificeren, maar is het ook een gelegenheid om cursisten op te leiden in een efficiënter gebruik van de technologie, één die het verwerven van hun denk- en taalvaardigheden ten gunste komt.

Kortom, hoewel GenAI handig is voor het uitwerken van lesmateriaal, benadrukken de leerkrachten het belang van menselijke controle en kritische denkvaardigheden.

Voulez-vous ChatGPT avec moi ?

ChatGPT is op vandaag geen directe concurrentie voor de goede taalleerkracht, die er echter wel alle baat bij heeft om zich het instrument eigen te maken en er didactisch voordeel uit te halen. Om een effectief gebruik van GenAI in de lespraktijk te promoten, geven onze deelnemers nog een paar aanbevelingen.

Om de meer terughoudende collega’s te overtuigen, is het nodig om met een bepaalde regelmaat momenten te organiseren waar leerkrachten samen kunnen experimenteren. Meerderen onder hen gaven aan dat ze niet goed wisten waar te beginnen, en dat de nuttige tips gegeven door een meer ervaren collega hen op weg geholpen hebben. Ook voor de leerkrachten die reeds aan de slag zijn met GenAI is het een meerwaarde verder uit te wisselen en te checken of er meer efficiënte manieren bestaan. Het samen experimenteren

leidt tot nieuwe inzichten en toepassingen en maakt de leerkrachten meer zelfzeker, wetende dat cursisten sowieso ook experimenteren met GenAI. Collectief leren geeft zin in leren. Daarmee hebben we al een paar ingrediënten voor de opzet van een professionele leergemeenschap bij CVO.

Om die leergemeenschap te voeden met nieuwe oefeningen, inzichten en aanbevelingen, is het nodig om een continue vinger aan de AI-pols te houden. Het kan lonen om een paar experts aan te duiden die een goed overzicht hebben van de markt, de technologieën en de toepassingen. Gezien GenAI nog sterk zal evolueren in de komende jaren, zal de lerende organisatie voordeel halen uit de onderbouwde adviezen van enkele interne connaisseurs. Zij kunnen ook onderzoeken of samenwerkingen met AI gestuurde taalapplicaties interessant zijn.

Het CVO zou kunnen overwegen om te investeren in een op maat gemaakte chatbot, die gevoed wordt met de reeds uitgebreide kennisdatabank van de school. ChatGPT kent immers niet het verschil tussen een taalbeheersingsniveau A2 en B2, het zou krachtig zijn om een beroep te kunnen doen op de lerende algoritmen van een softwarebedrijf en die dan te voeden met het specifieke lesmateriaal van het CVO. Op die manier zou de betrouwbaarheid van de resultaten aanzienlijk stijgen.

Onderzoekskeuzes

In dit onderzoek werd weinig aandacht besteed aan de beleving van de cursisten bij het gebruik van GenAI. Gezien de technologie nog zo nieuw is, vonden we het prioritair om de ervaringen van de leerkrachten te kunnen voeden met ideeën en tips, maar het zou een toegevoegde waarde zijn om ook de stem van de cursisten te horen, en hoe GenAI hun leerproces beïnvloedt.

We hadden een vermoeden en stelden ook vast tijdens het focusgesprek dat vele leerkrachten de stap naar het gebruik van GenAI nog niet hadden gezet. We wouden met het onderzoek niet te veel afwijken met vragen over percepties, overtuigingen en digitale vaardigheden, maar die zijn wel bepalend indien er een breed gebruik door een heel team van leerkrachten beoogd wordt. We hebben een tipje van de sluier opgelicht door te peilen naar de visie op GenAI, maar dit thema verdient ook een grondiger onderzoek.

De mogelijkheden van de GenAI zijn nog vrijwel onontgonnen. De praktijken die we op vandaag beschrijven, zijn overmorgen achterhaald. Het voelt aan als de beschrijving van een

landschap vanuit een hogesnelheidstrein op kruissnelheid. We hebben voornamelijk toepassingen van de grote taalmodellen belicht, en exclusief die met geschreven taal. Voor het taalonderwijs kan het ook interessant zijn om de mogelijkheden van GenAI met spraaktechnologie te bestuderen.

Odile of ChatGPT?

Indien je gemotiveerd bent om je Frans op te frissen, lijkt Odile je beste optie om snel tot resultaat te komen. Ze is de architect van je leertraject, en samen met haar virtuele assistent ChatGPT zal ze ervoor zorgen dat je de juiste dingen leert voor het bereiken van je doelen, dat de werkvormen gevarieerd zijn en afgestemd op je noden, en dat je steeds een zicht op je vooruitgang hebt. Ze zal je zelf een duwtje in de rug geven als je het even niet meer ziet zitten.

En of ze veel koffie drinkt en nooit nors is? Laat ons nog maar even onze typisch menselijke grillen hebben, dat houdt de spanning erin. Et maintenant ? Au travail !

Referenties

Crompton, H., Edmett, A., & Neenaz Ichaporia. (2023). Artificial intelligence and English language teaching: A systematic literature review. *www.britishcouncil.org*,
2. https://www.britishcouncil.org/sites/default/files/ai_in_english_language_teaching_systematic_review.pdf

Edmett, A., Ichaporia, N., Crompton, H., & Crichton, R. (2023). Artificial intelligence and English language teaching: Preparing for the future. British Council.
<https://doi.org/10.57884/78EA-3C69>

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal Of Education*, 57(4), 542–570. <https://doi.org/10.1111/ejed.12533>

Thùy, D. T., Da, C., & Van Hanh, N. (2024). The use of ChatGPT in teaching and learning: a systematic review through SWOT analysis approach. *Frontiers in Education*, 9.
<https://doi.org/10.3389/feduc.2024.1328769>

X.com (2024). X (Formerly Twitter). Geraadpleegd op 3 juni 2024, van
<https://x.com/sama/status/1790075827666796666?lang=fr>



Onderzoeksrapport

MASTERPROEF
EDUCATIEVE MASTER



MULTIDISCIPLINAIR
INSTITUUT
LERARENOPLEIDING

Inhoudstafel

Gegevens masterproefonderwerp.....	3
Inleiding	4
Artificiële intelligentie (AI)	5
Types van AI en AI-modellen	6
Taalkundige modellen.....	6
AI in educatie.....	7
AI-applicaties in taalonderwijs.....	8
Spanningsvelden	8
Onderzoeksopzet en methode.....	13
Resultaten	16
Discussie – conclusie	21
Referentielijst.....	24

Gegevens masterproefonderwerp

- Titel:
 - AI in Education (titel MaThis)
 - ChatGPT zet de taalleerkracht (nog niet) buitenspel (titel populair-wetenschappelijk artikel)
- Promotor: Professor Tondeur
- Studenten + EduMa: Katrien Sterckx (Educatieve Master Talen), Nele Libbrecht (Educatieve Master Cultuurwetenschappen + NT2), Julie Simoens (Educatieve Master Talen)
- Geschatte indienmoment van de masterproef: juni
- Wordenaantal (exclusief referenties):
 - 7981 woorden (onderzoeksrapport)
 - 3955 woorden (populair-wetenschappelijk artikel)

Inleiding

- [Wat is de maatschappelijk relevante probleemstelling?](#)

AI is niet meer weg te denken uit onze maatschappij. Na de lancering van ChatGPT eind 2022 volgde al snel de vraag of we chatbots moeten toestaan of verbieden in het onderwijs. In het begin was er veel tegenstand maar tegenwoordig beginnen steeds meer mensen de voordelen van het gebruik in te zien. Steeds meer leerkrachten gebruiken chatbots zoals ChatGPT en CoPilot voor hun lesvoorbereiding en in de les: “Mijn lessen voorbereiden gaat nu dubbel zo snel als vroeger’: veel leerkrachten omarmen AI in de klas” (Van Droogenbroeck, 2024).

In het begin ging academisch onderzoek vooral over de percepties van leerkrachten over AI in het onderwijs maar tegenwoordig spitst onderzoek zich steeds meer toe op hoe leerkrachten AI kunnen inzetten in het onderwijs. Ons onderzoek onderzoekt hoe taalleerkrachten in het volwassenenonderwijs GenAI kunnen inzetten in hun onderwijspraktijk. Tot op vandaag is er nog maar weinig onderzoek dat zich specifiek focust op het volwassenenonderwijs.

- [Wat is het algemene onderzoeksdoel?](#)

Onze initiële onderzoeksvraag was de volgende: “Hoe kan GenAI de leerkrachten helpen met hun onderwijspraktijk in het volwassenentaalonderwijs?”

We herformuleerden onze onderzoeksvraag als volgt: “Hoe kunnen taalleerkrachten in het volwassenenonderwijs GenAI inzetten in hun onderwijspraktijk?”

Literatuurstudie

- Wat zijn de centrale concepten/begrippen binnen je onderzoek?

Artificiële intelligentie (AI)

Bij het zoeken naar de definitie van AI hebben we meerdere (academische) bronnen geconsulteerd. AI snel merkten we dat we in de literatuur veel verschillende definities tegenkwamen. In een van de academische artikels (Holmes & Tuomi, 2022) werd er aangehaald dat er niet één dominante definitie bestaat die door de meeste AI-experts wordt aanvaard:

“There have been many attempts to define and clearly describe what we are talking about when we talk about Artificial Intelligence (AI), a name that we capitalise to highlight that it is a specific field of inquiry and development, and not simply a type of intelligence that is artificial. The literature provides many alternative definitions and it is often claimed that there exists no single dominant definition accepted by most AI experts.” (Holmes & Tuomi, 2022, p. 546)

Voor onze literatuurstudie lijkten we 3 verschillende definities op uit verschillende delen van de wereld:

EU: *“We propose to use the following updated definition of AI: “Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions. As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).”* (High Level Expert Group on Artificial Intelligence, 2019, p. 6)

VS: *“It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.”*
“Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines.” (McCarthy, J., 2007, p. 2)

China: *“AI machines do not necessarily have to obtain intelligence by thinking like a human and that it is important to make AI solve problems that can be solved by a human brain. Brain science and brainlike intelligence research and machine-learning represented by deep neural networks represent the two main development directions of core AI technologies, with the latter referring to the use of specific algorithms to direct computer systems to arrive at an appropriate model based on existing data and use the model to make judgment on new situations, thus completing a behavior mechanism...In general, the artificial intelligence we*

know today is based on modern algorithms, supported by historical data, and forms artificial programs or systems cap." (China Institute for Science and Technology Policy at Tsinghua University, 2018, p. 9)

Types van AI en AI-modellen

Binnen de artificiële intelligentie (AI) zijn er twee grote subgroepen, namelijk *Machine Learning* (ML) en *Deep Learning* (Janiesch et al., 2021)

Machine Learning is (semi) geautomatiseerde technologie die ontwikkeld is voor de oplossing van specifieke problemen. De algoritmes "leren" uit de beschikbare vormen van data in een bepaald vakgebied. De *Machine Learning*-algoritmes kunnen grofweg worden onderverdeeld in drie categorieën namelijk *supervised learning*, *unsupervised learning* en *reinforced learning*.

Deep Learning is een subgroep van *Machine Learning* modellen die over complexere neurale netwerken en ook over verschillende niveaus van gegevensbehandeling beschikken, waardoor de "lerende" capaciteit van de technologie aanzienlijk stijgt.

AI-systemen worden getraind in het verwerken van verschillende soorten gegevens, waar het dan meerdere taken op kan uitvoeren. De grote gegevensmodellen zijn: taal, beeld en cijfermatig materiaal.

Taalkundige modellen

In het licht van dit onderzoek naar het gebruik van artificiële intelligentie in taalonderwijs focussen we ons op de taalkundige modellen binnen AI. Deze taalmodellen zijn getraind op grote tekstdatasets en bieden verschillende mogelijkheden zoals het genereren van tekst, het samenvatten van documenten, het vertalen en het beantwoorden van vragen. We onderscheiden *small language models* (SLMs) en *large language models* (LLMs). Deze twee modellen verschillen van elkaar door het aantal gebruikte parameters bij het uitvoeren van een linguïstische taak.

Onder de *small language models* rekenen we de taalmodellen binnen AI die minder dan 100 miljoen parameters gebruiken (Ghosh 2023). De efficiëntie wordt ervaren als hoger aangezien de modellen sneller werken doordat ze minder parameters verwerken, wat ook een impact heeft op de vereiste opslagruimte. De algemene capaciteiten van deze kleine taalmodellen is echter ook lager. Dit betekent dat de kleine modellen zich vaker toespitsen op specifieke functies.

De grote technologiebedrijven, zoals Google, Microsoft, Amazon, IBM en Facebook, focussen zich steeds meer op de ontwikkeling van *large language models*. De bekendste grote taalmodellen zijn op dit moment GPT3 en ChatGPT, Bard en LLaMA. Deze modellen zijn *deep learning* algoritmes die verschillende taalverwervingstaken kunnen uitvoeren (Floridi 2023). Tegenwoordig kan AI de elektromagnetische eigenschappen beheren om teksten te verwerken op een manier die niet te onderscheiden is van menselijke resultaten (Floridi 2023). Dit noemen we *Natural Language Processing* (NLP).

NLP is ontstaan op het snijvlak van artificiële intelligentie en taalkunde in de jaren 50 (Nadkarni 2011). Natuurlijke taalverwerving binnen AI maakt gebruik van een statistische methode waarbij computeralgoritmen gecombineerd worden met *machine learning* en *deep-learning*-modellen om de bovengenoemde elementen uit tekst- en spraakgegevens automatisch te onttrekken, classificeren en labelen. Vervolgens wordt er een statistische waarschijnlijkheid toegewezen aan elke mogelijke

betekenis van de woorden binnen de context. De *deep learning*-modellen maken gebruik van leertechnieken gebaseerd op convolutionele neurale netwerken (CNNs) en recurrente neurale netwerken (RNNs) die ervoor zorgen dat de NLP-systemen al doende leren en zo steeds nauwkeuriger worden bij het bepalen van betekenissen uit enorme hoeveelheden ruwe tekst- en spraakdata.

Het is belangrijk te begrijpen dat AI-modellen niet voor zichzelf denken, zelf redeneren of verstaan (Bishop 2021). De modellen werken met datasets en algoritmes. De hoeveelheid data wordt steeds groter en de snelheid en nauwkeurigheid van de algoritmes steeds beter. Zo slagen AI-systemen erin op een statistische manier, dat wil zeggen aan de hand van formele structuren, te bereiken wat wij op een semantische manier, aan de hand van betekenis, verwerken (Floridi 2023).

AI in educatie

De mogelijkheden van artificiële intelligentie sijpelen door in verschillende werkomgevingen. Velen zien AI als het drijf wiel voor een vierde industriële revolutie (Zhai et al. 2021). De economische productiemodi in samenlevingen, aldus ook de industriële revoluties, hebben steeds vormgegeven aan de sociale organisatie van het leren. AI zal dus ook haar weg vinden in het onderwijslandschap (Yang et al. 2011). AI-applicaties kunnen de huidige digitalisatie van onderwijsmiddelen, gamificatie en gepersonaliseerde leerervaringen bijstaan.

De invoer van AI in het onderwijs zal echter gepaard gaan met aanpassingen voor zowel leerkracht als leerling. De rol van de leerkracht kan door het verhoogd gebruik van bijvoorbeeld *intelligent tutoring systems* (ITSs) veranderen en zal geherconceptualiseerd moeten worden (Albin-Clark et al. 2011). Daarnaast moet er aandacht zijn voor de attitudes van leerlingen ten opzichte van AI. De leerling moet AI effectief leren inzetten voor bepaalde leercontexten, anders kan het foutief gebruik van AI leiden tot negatieve attitudes ten opzichte van leren (Seni 2012).

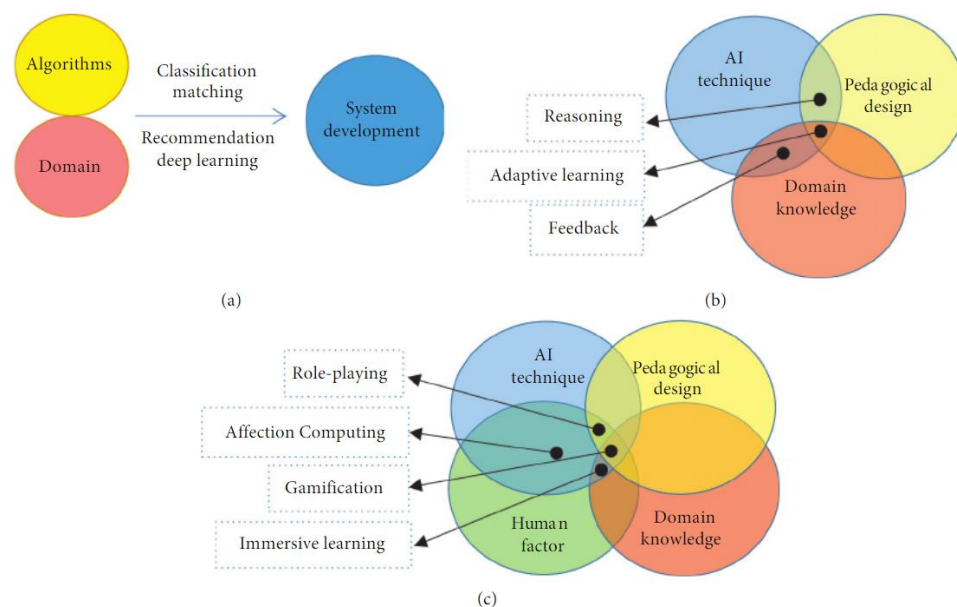


FIGURE 1: The hierarchy of artificial intelligence in educational implementation. (a) The dimension of system development, (b) the dimension of extraction, and (c) the dimension of application.

(Nguyen en Yang 2012)

In de metastudie van Zhai et al. (2021), worden drie types van pedagogische toepassing erkend: feedback, redeneren en adaptief leren. Feedback is een belangrijke strategie om de leerpatronen van

de leerling te optimaliseren (Melo et al. 2014). AI-systemen hebben de mogelijkheid leerlingen van onmiddellijke feedback te voorzien op basis van de input van deze leerlingen. AI kan tevens op verschillende manieren helpen om het redeneervermogen van leerlingen te versterken. Redeneerprocessen zijn grotendeels onzichtbaar en het is dus moeilijk voor de leerkracht om deze processen op gang te brengen door het observeren van gedrag. AI-technieken zoals de visualisatietechniek kunnen worden toegepast om het redeneren te stimuleren bij leerlingen. Een voorbeeld zijn de *argument-mapping* tools. Een derde pedagogische toepassing is het adaptief leren, waarbij AI zich aanpast aan de kenmerken, behoeften en voorkeuren van individuele leerlingen. Het verschil met de eerder vermelde feedback is dat het feedbacksysteem standaardreacties geeft. Het adaptieve leersysteem daarentegen is een formatief en correctief geautomatiseerd systeem dat zichzelf kan aanpassen naargelang het doel van de interventie om de gepersonaliseerde ontwikkeling van leerlingen te ondersteunen. (Jones 2011).

AI-applicaties in taalonderwijs

Er zijn enkele veelgebruikte AI-applicaties die ook ingezet worden in taalonderwijs zoals chatGPT en Bard. Deze conversatiemodellen helpen leerlingen om te interageren in een nieuwe taal, helpen bij creatieve opdrachten en kunnen ingezet worden om te vertalen of om te controleren op grammatica.

Er zijn ook specifieke taalapps die ingezet kunnen worden bij het leren van een taal voor alle taalvaardigheden. Enkele voorbeelden zijn Duolingo, Rosetta Stone, Babbel, Bussuu, MEMRISE, etc. Tot voor kort gebruikten deze taalapps vooral *Machine Learning* maar deze taalapps zijn ook aan het evolueren en gebruiken nu meer *Deep Learning* om de taalverwerving steeds meer te personaliseren.

Spanningsvelden

Artificiële intelligentie kent, zoals eerder besproken, verschillende toepassingen in het onderwijs en creëert zo vele mogelijkheden, zoals slimme studiebegeleidingssystemen of hulpmiddelen voor leerkrachten. Deze toepassing kent een steeds grotere rol toe aan databanken en algoritmes (Maedche 2019). Hier schuilen tevens enkele gevaren of valkuilen binnen AI.

Toepassingen in het onderwijs vereisen dus, net zoals elke datagedreven toepassing, persoonlijke gegevens. De Vlaamse toetsen, die in het schooljaar 2023-2024 geïntroduceerd werden, zullen onder andere bijdragen tot de gegevensverzameling omtrent kwaliteit van het onderwijs en het leerproces van leerlingen. Deze gegevens, in combinatie met andere data, zouden gebruikt kunnen worden om nieuwe algoritmes te creëren en toe te passen. Het is echter cruciaal rekening te houden met de mening van het doelpubliek. Er zijn namelijk verschillende zorgen betreffende privacy. Onderzoek door het Kenniscentrum Data & Maatschappij (Martens et al., z.d.) onderzocht de mening van ouders en leerlingen. Hieruit kan geconcludeerd worden dat zowel ouders als leerlingen een sterke nood hebben aan controle. Ze willen geïnformeerd worden en de mogelijkheid hebben te bepalen waarvoor welke gegevens gebruikt worden, alsook het gebruik van hun gegevens weigeren.

Het werken met data gegenereerd door mensen betekent uiteraard dat AI menselijke fouten of bias kan overnemen. Bias duidt op een vooringenomenheid en vormt ook in AI een struikelblok (Mitchell 1997). De algoritmes die AI gebruikt zijn namelijk een deel van de reeds bestaande (biased) instituties en structuren. Daarnaast bestaat het gevaar dat AI nieuwe vooroordelen introduceert aangezien AI-systemen de voorkeur geven aan aspecten van het menselijk leven die gemakkelijk telbaar zijn, en dus niet die aspecten die moeilijker of zelfs onmogelijk telbaar zijn. Ook bestaat het gevaar dat AI foute correlaties of reducties maakt. In het geval dat datasets steeds zouden bestaan uit foto's van honden

met een bal, zal het AI-systeem denken dat honden en ballen steevast met elkaar voorkomen. Wanneer de datasets enkel uit bruine honden zou bestaan, zal het AI-systeem denken dat alle honden bruin zijn.

Daarnaast is er nog de uitdaging betreffende veiligheid en ethiek (EU 2022). AI-systemen zijn namelijk gevoelig aan polarisatie. Ze kunnen online echokamers vormen en zo een gevaar zijn voor de democratie. Ook wat betreft misbruik, plagiaat, fraude, cybercriminaliteit en inbraak op privacy bestaan er verschillende gevaren. Het is belangrijk om de leerlingen bewust te maken van deze valkuilen zodat ze de mogelijke bias en echokamers niet onwetend overnemen.

Bij de mogelijkheden van AI vermeldden we het nut van technologische assistenten die routinetaken kunnen overnemen en automatiseren. De keerzijde hiervan is uiteraard dat de werkcontext ontmenselijkt wordt, wat een impact heeft op jobs.

- [Wat is het relevante theoretisch/conceptueel kader?](#)

In ons onderzoek zijn er linken naar twee grote onderzoeks- en toepassingsgebieden: enerzijds is er het concept van de generatieve artificiële intelligentie als technologisch artefact en anderzijds hebben we de onderwijskundige modellen en concepten. Die laatste kunnen dan nog verder afgelijnd worden in het specifieke sub-domein van het taalonderwijs.

Het model dat die multidisciplinaire aspecten naadloos integreert is het TPACK-model van Mishra en Koehler (2006). Dit raamwerk linkt technologische inhoud aan pedagogisch didactische modellen en aan inhoudelijke thema's. Het initiële opzet ervan was om aanschouwelijk te maken hoe technologie de leerprocessen kan ondersteunen en leiden tot betere leerresultaten. De drie belichte kennisgebieden zijn tevens wederzijds afhankelijk in dit model, dit zet ons aan om in het onderzoek ook aandacht te besteden aan de "tussengebieden", niet vanuit een statische beschrijving, maar inzoomend op hun specifieke interacties. Het T-pack model is fundamenteel contextgebonden en dynamisch geconcipeerd, wat bijdraagt tot zijn kracht en universaliteit, we kunnen er in dit onderzoek ook weer een beroep op doen.

In de technologische kennisgebieden kunnen we beroep doen op theoretische kaders m.b.t. "*Natural Language Processing*" (voor quasiauthentieke interactie), "*deep learning*" en neurale netwerken, of ook sommige informatieverwerkingsmodellen.

In de pedagogisch-didactische kennisgebieden zijn er meerdere theoretische kaders die kunnen ingezet worden afhankelijk van de context en de na te streven leerdoelen. Er lijkt een vrij goede matching te zijn met de constructivistische insteek, zeker in combinatie met collaboratieve werkvormen en de ontwikkeling van de 21st Century Skills, maar we sluiten ook geen behavioristische opbouw van leerpaden uit, zeker voor het verwerven van taalkundige basiskennis. Ook cognitivistische inzichten kunnen helpen om leeractiviteiten met een goede cognitieve belasting van het werkgeheugen op te zetten, of specifieke hulp te bieden bij leerstoornissen.

De inhoudelijke kennisgebieden situeren zich in deze studie op het vlak van de didactiek van vreemde talen. We besteden aandacht aan de volgende aspecten:

- Grammaticaonderwijs
- Woordenschatonderwijs
- Spreekvaardigheid

- Luistervaardigheid
- Schrijfvaardigheid
- Leesvaardigheid
- Welke antwoorden zijn er al op je onderzoeksvraag vanuit vorig onderzoek?

Cambridge University

Cambridge University doet onderzoek en experimentatie rond het gebruik en de voordelen van *Generative AI* in taalleerprocessen. De huidige experimenten hebben tot nog toe gefocust op het potentieel van de technologie voor de creatie van leerinhouden en evaluaties.

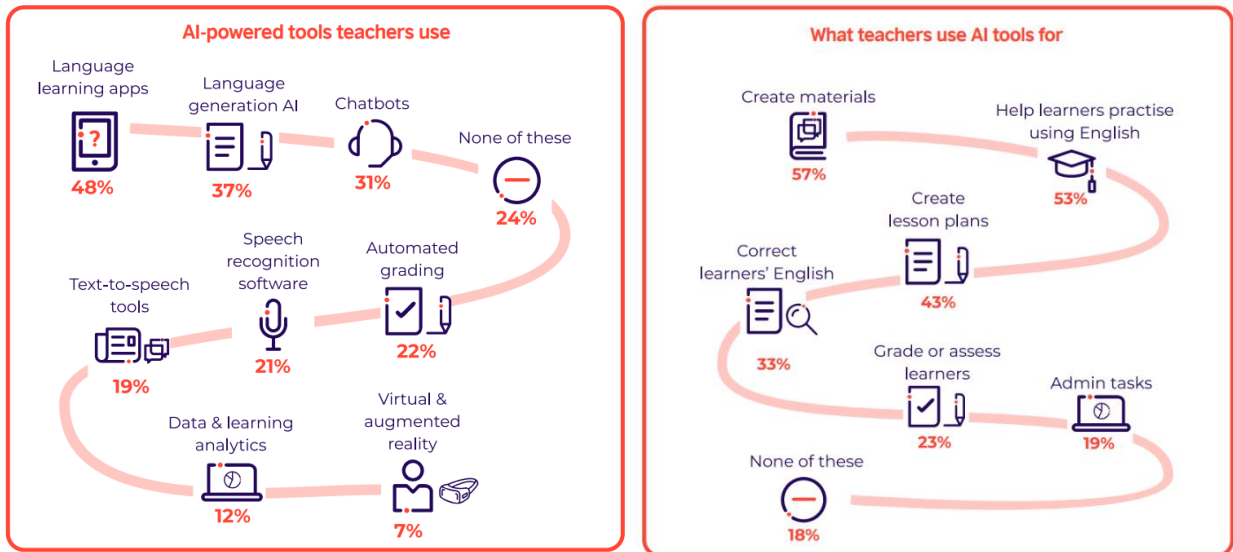
“In experiments to date, we have used third party commercial LLMs and have shown that they can be applied to creating a range of content types, e.g. multiple-choice reading comprehension questions at different levels of the CEFR or the design of prompts. These GenAI models could also enable the calibration of assessment content.” (Cambridge University, 2023, p. 5)

In hetzelfde artikel (Cambridge University, 2023) wordt er aangehaald dat Cambridge University werkt rond meerdere onderzoeksthema's in verband met AI in educatie. Een van de thema's is evaluatie en we vonden volgend interessant experimenteel onderzoek van Xu et al. (2021). In dit onderzoek werd de betrouwbaarheid van autocorrectie onderzocht aan de hand van antwoorden van kandidaten in een online mondelinge toets Engels. Ze lieten deze toets verbeteren door drie leerkrachten Engels en door de autocorrectie. De gemiddelde score van de drie leerkrachten Engels werd vergeleken met de score van de autocorrectie. Het onderzoek wees uit dat de autocorrector iets toegeeflijker verbeterde dan het gemiddelde van de leerkrachten, voornamelijk voor de kandidaten met een lage spreekvaardigheid.

Onderzoeksrapport in opdracht van de British Council

In het onderzoeksrapport van Edmett et al. (2023) in opdracht van de British Council werd er door middel van een kwantitatieve bevraging getracht om een antwoord te geven op de volgende vraag: Hoe wordt AI gebruikt voor het onderwijzen en leren van de Engelse taal?

De survey werd ingevuld door 1348 leerkrachten Engels uit 118 landen en regio's. In de survey werd er gepeild naar welke AI-tools de leerkrachten gebruiken en waarvoor leerkrachten deze tools gebruiken. De resultaten werden voorgesteld in deze twee kaders (Edmett et al., 2023, p. 24-25):



In de volgende sectie van de survey moesten de respondenten hun meningen geven over een aantal stellingen. Dit waren de meest opmerkelijke stellingen:

Statement 8:
I have received enough training to incorporate AI into my teaching.

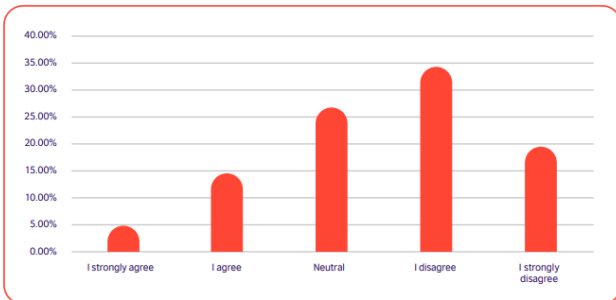


Figure 7 Responses from 1,112 teachers to survey statement 8

(Edmett et al., 2023, p. 30)

Statement 11:
By 2035, AI will be able to teach English without a teacher.

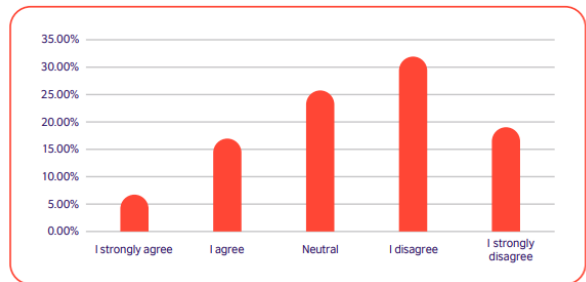


Figure 10 Responses from 1,112 teachers to survey statement 11

(Edmett et al., 2023, p. 33)

Metastudie van Thùý et al. (2024)

In de metastudie van Thùý et al. (2024) werd een SWOT-analyse gemaakt van het gebruik van generatieve AI, in het bijzonder de chatbot ChatGPT, in de verschillende aspecten en fasen van het leerproces. Het onderzoek belicht volgende sterktes en zwaktes van het gebruik van AI in het onderwijs:

Sterktes:

- Het verstrekken van basiskennis en reduceren van de *workload* van de leerkracht.
- Het aanbieden van gepersonaliseerde leerervaringen.
- Het verbeteren van leeruitkomsten.

Zwaktes:

- Bias in resultaten.

- Gegeneerde informatie door ChatGPT is niet altijd correct.
- [Moet je onderzoeksvraag geherformuleerd worden?](#)

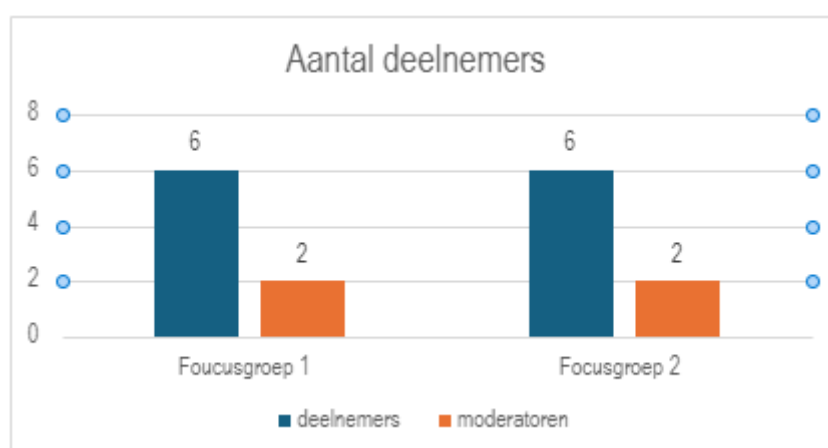
Onze initiële onderzoeksvraag was de volgende: “Hoe kan GenAI de leerkrachten helpen met hun onderwijspraktijk in het volwassenentaalonderwijs?” Bij het opstellen van ons instrument voor de focusgroepen in maart vonden we dat we onze initiële onderzoeksvraag neutraler moesten formuleren omdat we ook ingaan op de nadelen van het gebruik van GenAI. Het werkwoord “helpen” in de initiële onderzoeksvraag heeft immers een positieve connotatie. We herformuleerden onze onderzoeksvraag als volgt: “Hoe kunnen taalleerkrachten in het volwassenenonderwijs GenAI inzetten in hun onderwijspraktijk?”

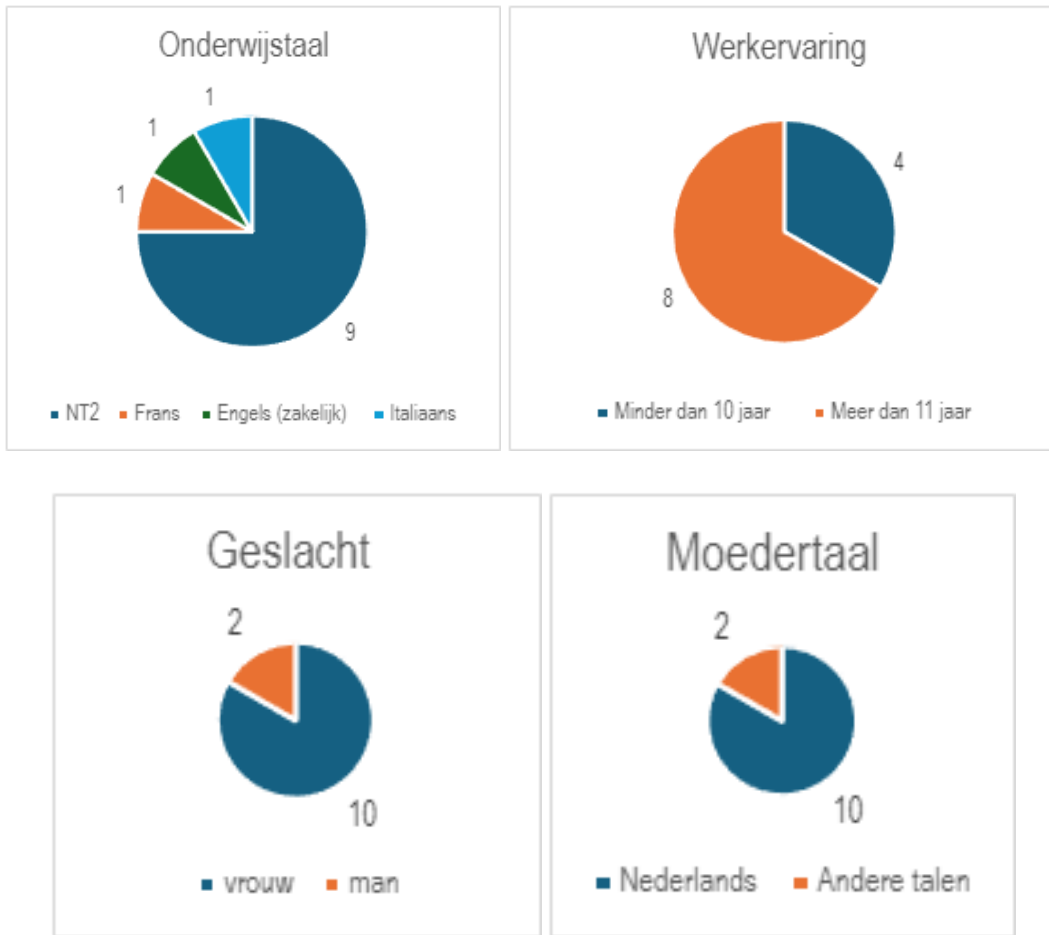
Onderzoeksopzet en methode

In dit onderzoek wordt gepeild naar de huidige AI-praktijken van taalleerkrachten, alsook hun percepties. Het onderzoek bestaat uit twee blokken: enerzijds de dataverzameling, en anderzijds een literatuurstudie om het onderzoek te onderbouwen en de resultaten te staven. Wij hebben geopteerd voor het verzamelen van overwegend kwalitatieve gegevens. De kwalitatieve methode genoot onze voorkeur omwille van verschillende voordelen. Kwalitatief onderzoek brengt persoonlijke belevingen en gevoelens in kaart. De dataverzameling gebeurde aan de hand van focusgesprekken met een aantal taaldocenten van CVO Semper. Voorafgaand werd ook een vragenlijst uitgestuurd om wat kwantitatieve gegevens over de doelgroep te verzamelen. De kwalitatieve focusgesprekken lieten ons toe in gesprek te gaan met enkele participanten. Deze methode stelt de stakeholders in staat om met elkaar in discussie te gaan over het onderzoeksonderwerp. Dit is interessant voor het onderzoek aangezien de bijdragen van een deelnemer op zijn beurt andere deelnemers kan inspireren en aanzetten tot het delen van ervaringen en ideeën. Zo bouwt men steeds verder op elkaars opvattingen en komt men tot diepgaandere inzichten. Zo kregen we inzicht in de percepties, emoties en ervaringen van de participanten, alsook consensus en/of verdeeldheid in de groep.

Dit onderzoek maakt deel uit van een bredere cyclus die over meerdere jaren uitgedragen kan worden met het uiteindelijke doel een professionele leergemeenschap te creëren waarin men competent is in het gebruik van artificiële intelligentie en waarbij men elkaar hierin ondersteunt. Dit oriënterend onderzoek vormt de vroege onderzoeksfase waarin wordt beoogd een zicht te krijgen op huidige praktijken en percepties rond artificiële intelligentie, alsook te toetsen naar de verwachtingen, uitdagingen, noden en aanbevelingen van de taalleerkrachten.

Voor dit onderzoek werden twee focusgroepen opgericht. Elke groep nam deel aan één focusgesprek. De focusgroepen bestonden uit taalleerkrachten uit CVO Semper. Aanvankelijk werd in december een oproep gedaan in de nieuwsbrief van CVO Semper om deel te nemen aan een experimentatie. Hier kwam echter slechts weinig reactie op. In februari werden in CVO Semper twee workshops rond AI verzorgd door twee leerkrachten die reeds competent zijn in het gebruik van AI. Hierdoor wijzigden we onze aanpak van experimentatie naar focusgesprekken. De participanten van de workshops werden uitgenodigd voor de focusgroepen. Uiteindelijk namen twaalf respondenten deel aan de focusgroepen. Onderstaande grafieken geven de deelnemerskenmerken weer:





De deelnemers kregen in dit werk pseudoniemen toebedeeld.

Focusgesprek 1 en 2 vonden respectievelijk plaats op 27 maart en 29 maart en werden in Microsoft Teams uitgevoerd. Voor elk gesprek werd één uur voorzien. Vóór de gesprekken werd reeds een vragenlijst uitgestuurd die de achtergrondkenmerken van de participanten bevroeg, alsook een formulier van *informed consent*. De gesprekken zelf verliepen aan de hand van een interviewleidraad die op voorhand opgesteld werd. Deze leidraad volgt de algemene opstelling voor een focusgesprek, bestaande uit een inleiding, een ijsbreker, de kernvragen en groepsdiscussie, en de nabespreking. Dit algemene kader werd vervolgens aangepast aan de specifieke gegevens van de deelnemers van de gesprekken, alsook om een antwoord te bieden op de onderzoeksvraag. Tijdens het focusgesprek werd deze leidraad gevolgd met ruimte voor uitweiding, input en vragen van de deelnemers. De vragen uit de interviewleidraad werden na de gesprekken in de vorm van een bundel per e-mail verspreid naar alle participanten. Zo kregen zij de mogelijkheid om eventuele inzichten naderhand alsnog met ons te delen. De participanten kregen hier de vrijheid de bundel in de eigen moedertaal in te vullen om het delen van ideeën of gevoelens te faciliteren.

De onlinefocusgesprekken werden met de toestemming van alle deelnemers opgenomen en automatisch getranscribeerd met de transcriptiefunctie van Microsoft Teams. Deze automatische transcriptie was echter gebrekkig en moest nadien manueel bijgewerkt worden. Aan de hand van deze transcriptie kon een codeboom opgesteld worden. Hiervoor werkten we in het programma MAXQDA. De codeboom bestaat uit hoofdcodes en subcodes. Hiervoor werd een mixed method gebruikt waarbij een deductieve en inductieve aanpak gehanteerd werden. De hoofdcodes werden op deductieve manier bepaald. De hoofdcategorieën werden afgeleid uit de interviewleidraad. Aan elke vraag werd

één hoofdcode gelinkt. De subcodes werden vervolgens inductief bepaald aan de hand van de antwoorden van de participanten. De transcripties werden doorgelezen. Aan elke betekenisvolle quote werd een subcode gelinkt. Op deze manier kwam een codeboom tot stand en werden quotes gegroepeerd onder subcodes. De codeboom ging door enkele rondes van iteratie om zo de codes uit de twee focusgesprekken op elkaar af te stemmen en de categorisering scherp te krijgen. De codeboom werd in een Excel-bestand geëxporteerd. Dit bestand werd aangevuld met de meest betekenisvolle quotes, alsook een indicatie voor perceptie of ervaring, en consensus of conflict (zie bijlage). Ook de transcriptie met alle gelinkte subcodes werd geëxporteerd in een PDF-bestand (zie bijlage).

In dit onderzoek werd de kwalitatieve data geanalyseerd vanuit een generieke benadering. Dit houdt in dat percepties, meningen en ervaringen van mensen onderzocht worden. De data onderging een thematische analyse. We hanteren hier het zesstappenmodel van Braun en Clarke (2006). Dit model begint met de familiarisatie van de data en het genereren van codes. Tevens werden de thema's gegenereerd, beoordeeld, gedefinieerd en benoemd. Voor deze thema's vertrokken we reeds vanuit een conceptueel kader dat tevens houvast geeft bij het analyseren van de data. De interviewleidraad met vragen werd immers reeds per thema opgedeeld. Deze overkoepelende thema's vloeien voort uit de literatuurstudie. De laatste stap van de thematische analyse is het opstellen van het onderzoeksrapport.

Resultaten

Zoals eerder vermeld, zal de analyse van de resultaten gebeuren aan de hand van een codeboek dat opgesteld werd. De hoofdcodes werden deductief bepaald aan de hand van de vragen uit de interviewleidraad van de focusgesprekken. De subcodes werden vervolgens inductief bepaald aan de hand van de input van de deelnemers tijdens de gesprekken.

We onderscheiden acht hoofdcodes: visie AI, meest gebruikte AI-tools, taalverwervingsaspecten, onderwijspraktijken, invloed op leerervaring cursisten, invloed op werkervaring leerkrachten, aanbevelingen en uitdagingen.

De visie op AI werd grotendeels afgeleid uit de ijsbreker. De deelnemers moesten kiezen tussen enkele sciencefiction personages die elk gelinkt zijn aan een visie op AI. We onderscheiden een pragmatisch positieve visie (Iron Man), een inspirationale visie (Yoda uit de Star Wars-saga), een kritisch positieve visie (Neo uit The Matrix), een negatieve visie (HAL9000 uit A Space Odyssey) en een antropomorfische visie (Rachael uit Blade Runner). Enkel de doorgedreven negatieve visie werd niet gekozen. Dit wil zeggen dat geen enkele participant volledig tegen AI is. De deelnemers zagen verschillende positieve kanten van AI, weliswaar op voorzichtige wijze. Zo werd vermeld hoe AI informatie op een snelle manier kan verwerken en problemen kan oplossen (pragmatisch positief) en hoe het ons kan helpen om de wereld beter te begrijpen en hierin te navigeren (inspiracioneel). Het belang van het bewaren van het menselijke aspect werd echter ook benadrukt. Tot slot erkennen we ook een kritisch bezorgde visie. Deze komt niet voort uit de ijsbreker maar komt wel aan bod tegen het einde van focusgesprek twee. Bij het bespreken van de nadelen en uitdagingen rond het gebruik van AI kwamen steeds meer bezorgdheden aan de oppervlakte die leven bij de leerkrachten wat betreft de integratie van AI in ons dagelijkse leven. Dit zijn angsten zoals de overmatige afhankelijkheid en verspreiding van AI waardoor mensen kritisch denkvermogen (kunnen) verliezen, lui worden of ons denken gestuurd wordt door AI-systemen, en de sociale kloof tussen mensen met en zonder kennis.

De tweede categorie omvat de meest gebruikte AI-tools. De deelnemers van de focusgesprekken gaven aan welke tools ze reeds gebruiken. Al deze tools worden in deze categorie opgenomen en geclassificeerd onder 'grote taalmodellen', 'beeldmodellen' of 'machinaal leren'. Onder de grote taalmodellen vinden we ChatGPT en MS CoPilot. Van de beeldmodellen werden DALL.E, Midjourney en Stable Diffusion Web reeds gebruikt. Voor het machinaal leren is het gebruik beperkt tot Wisdolia en Cockatoo transcription. ChatGPT is bij uitstek de meest gebruikte en gekende AI-tool onder de taalleerkrachten. Opvallend is dat alle andere AI-tools die hierboven vermeld werden elk slechts door één persoon gebruikt werden.

Categorie drie en vier geven een beeld van hoe de taalleerkrachten AI momenteel inzetten in hun onderwijspraktijk. Categorie drie focust hierbij op de taalverwervingsaspecten: spreekvaardigheid, schrijfvaardigheid, leesvaardigheid, luistervaardigheid, woordenschat- en grammaticaonderwijs. Al deze kennisgebieden werden inductief, vanuit de focusgesprekken, bepaald als subcodes. Dit wil zeggen dat de deelnemers van de focusgesprekken gezamenlijk reeds voor elk taalverwervingsaspect AI als hulp ingeschakeld hebben. Voor de luistervaardigheid is het relevant te vermelden dat er geen audiofragmenten gemaakt werden door AI. De leerkrachten gebruiken tools, zoals Wisdolia, waarbij vanuit de transcriptie van bestaande fragmenten meerkeuzevragen opgesteld kunnen worden. Het uiteindelijke doel is dus een luisteroefening, maar de AI-tools verwerken enkel tekst. Daarnaast werd inductief een laatste subcode bepaald die zeer specifiek is voor het gebruik van taalleerkrachten, namelijk het gebruik van AI per taalniveau. In de focusgesprekken wordt dit specifiek vermeld voor de NT2-niveaus 'Threshold' en 'Vantage'. Dit zijn de hogere niveaus in het leertraject van een NT2-cursist. Maarten geeft hier aan dat de hogere niveaus gelijk staan aan een grotere tijdsbelasting. Het

lesmateriaal voor de lagere niveaus is simpel en kort. Dit vergt minder energie en tijd om op te stellen. Pas wanneer AI effectief tijd kan besparen is het de moeite waard om in te zetten.

Naast de taalverwervingsaspecten in categorie drie, focust categorie vier zich op de algemene onderwijspraktijken. Onder onderwijspraktijken worden alle taken van de leerkracht verstaan. Uit de focusgesprekken leiden we vier subcategorieën af: inspiratie bij werkvormen, oefeningen opstellen, evaluaties opstellen en cursisten laten evalueren door AI. Didactische werkvormen hebben pedagogische consequenties. Het is dus belangrijk steeds je werkvormen af te stemmen op het leerdoel, alsook je doelpubliek. Daarnaast is er best een regelmatige afwisseling van werkvormen tijdens een les zodat de les niet monotoon wordt. Hier raadplegen de leerkrachten AI als hulp door inspiratie te voorzien, aangepast aan hun doelen en klasgroep. Het opstellen van oefeningen wordt zeer ruim geïnterpreteerd. Het gaat hier om al het lesmateriaal dat door de cursisten tijdens de les gebruikt zal/kan worden, bijvoorbeeld invuloefeningen of een tekst voor een lees- of luisteropdracht. We kunnen concluderen dat dit type gebruik van AI het populairst is onder de taalleerkrachten aangezien het 13 keer werd vermeld in deze focusgesprekken door verschillende deelnemers. Ook het opstellen van evaluaties omvat verschillende types toetsingen. Het verschil met de vorige categorie is dat de evaluaties dienen om het eindproduct van een leerproces te toetsen. Dit gebruik is echter veel minder populair dan het maken van generiek lesmateriaal. De input van de cursisten op deze toetsing kan ook door AI beoordeeld worden. Dit valt onder de laatste subcategorie. Hier bepaalt artificiële intelligentie wat er goed, niet goed, of fout is, en waarom. Slechts één participant paste dit reeds toe in de praktijk.

De volgende twee codes omvatten de invloed van artificiële intelligentie op de ervaring van zowel cursisten als leerkrachten. Wat de cursisten betreft, omvat code vijf de invloed op hun leerervaring. Belangrijk te herhalen, is dat de focusgroepen uitsluitend uit leerkrachten bestonden. Het gaat hier om hun percepties of ervaringen die zij gehad hebben met het gebruik van AI door hun cursisten, en dus niet de stem van de cursisten zelf. Uit de focusgesprekken leiden we drie subcodes af: meer klassikale schrijfopdrachten, als inspiratie voor schrijfopdrachten en meer gemakzucht. De eerste subcode duidt op de reactie van leerkrachten op het gebruik van AI door cursisten. Om misbruik door middel van AI tegen te gaan, geven verschillende leerkrachten meer klassikale oefeningen om zo het gevaar van misbruik uit de weg te gaan. Hierop kwam echter de reactie van Melissa dat zij dat niet kan toepassen aangezien ze enkel online lesgeeft. De volgende twee subcodes duiden op het gedrag van de cursisten zelf. Net zoals de leerkrachten AI als inspiratiebron inzetten voor werkvormen, doen cursisten dit voor hun schrijfopdrachten. Het gebruik van AI door cursisten ter inspiratie, wordt door de leerkrachten aanvaard. Ze benadrukken echter dat het gebruik duidelijk beperkt moet blijven tot het bieden van ideeën of als hulp om te starten met een werk. Tot slot geven de leerkrachten aan dat er een gemakzuchtige houding aangenomen kan worden door cursisten wanneer ze artificiële intelligentie overmatig inzetten. Wanneer het wangebruik van AI door cursisten vermeld wordt, wordt dit steeds gekoppeld aan de relevantie van het maken van de taak. Dit lijkt de cursisten te ontglippen. De leerkrachten zijn echter gefocust op de vaardigheden en kennis die de cursisten moeten bezitten om aan de leerplandoelstellingen te voldoen.

De invloed van AI op de werkervaring van de leerkrachten (code zes) wordt opgedeeld in de voor- en nadelen. Deze worden op hun beurt onderverdeeld in inductief bepaalde subcategorieën.

Onder de nadelen vinden we terug: misbruik door cursisten, verschil in output per taal, doelgericht gebruik is moeilijk, niet volledig betrouwbaar / nood aan menselijke check, onnatuurlijk taalgebruik van AI, beperkingen huidige AI-taalmodellen, en sturende algoritmes en gestandaardiseerde output. Zoals hierboven aangegeven kan AI zorgen voor gemakzucht bij de cursisten wanneer zij de tools overmatig gebruiken. Dit kan echter ook tot gevolg hebben dat de cursist de leerdoelen zelf niet haalt

doordat AI al het werk levert. Het is echter moeilijk voor leerkrachten om dit te controleren of beperken. Dit resulteert in gevoelens van frustratie en onmacht. Daarnaast geven de leerkrachten een verschil in output per taal aan. De output in het Engels is beter dan in het Nederlands. Doelgericht gebruik is een volgende moeilijkheid. Zoals eerder vermeld is ChatGPT de meest gebruikte AI-tool. Het is echter belangrijk te beseffen wat de aard en het doel van ChatGPT zijn om de tool op een correcte manier te gebruiken. ChatGPT is een groot taalmodel (LLM) dat natuurlijke taal kan begrijpen (NLP), informatie kan genereren en in interactie kan gaan. Hierdoor vormt ChatGPT bijvoorbeeld een handig hulpmiddel ter ondersteuning van een schrijfproces. Dit taalmodel is echter niet geschikt om de taak op te nemen van andere types van AI-modellen, zoals het AI-beeldmodel DALL-E. Daarnaast kun je ook nog steeds best een zoekmachine gebruiken voor simpele kennisverwerving aangezien ChatGPT niet volledig betrouwbaar is. Voor doelgericht gebruik moet men dus begrijpen onder welk type AI-model de gebruikte technologie valt, wat het doel is en wat de beperkingen zijn. Zelfs wanneer je dit begrijpt, moet je nog leren accurate *prompts* gebruiken. Deze *prompts* zijn aanwijzingen die je het AI-systeem geeft om een zo nauwkeurig mogelijke output te ontvangen. Deze elementen zorgen ervoor dat het correct gebruik van AI niet voor iedereen zo vanzelfsprekend is, wat een nadeel vormt. Zoals hierboven aangegeven, is ChatGPT niet volledig betrouwbaar. Zo kan er foutieve of verouderde informatie gedeeld worden. Het is voor de gebruiker cruciaal om hier alert voor te blijven. Een menselijke check op de AI-output is steeds nodig. Voor een ervaren gebruiker kan dit zorgen voor frustratie. Een ongetraind gebruiker is misschien niet op de hoogte zijn van deze valkuil en kan hierdoor ongemerkt foutieve informatie overnemen. Ondanks dat ChatGPT een taalmodel is dat werkt aan de hand van menselijke input en pre-training, gaven verschillende leerkrachten aan dat het taalgebruik van ChatGPT onnatuurlijk is. Dit omvat alle output die inhoudelijk misschien wel correct is, maar waarvan de vorm gekunsteld en “artificieel” klinkt. We bespraken reeds dat doelgericht gebruik moeilijk kan zijn voor gebruikers van AI-technologieën. Je moet weten met welk doel een model ontworpen is. Daarnaast zijn er echter ook beperkingen bij de huidige AI-taalmodellen. Hieronder verstaan we de limitaties op vaardigheden die wel onder het takenpakket van de technologie vallen. Zo kunnen de tools beperkt zijn in het aantal afbeeldingen dat ze kunnen genereren of de lengte van een tekst of heb je geen toegang tot alle mogelijkheden in de gratis versie. Er kunnen ook beperkingen zijn door bijvoorbeeld de lange duur voor het genereren van een tekst of afbeelding. De laatste subcategorie onder de nadelen is de gestandaardiseerde output en de sturende algoritmes. AI-technologieën worden getraind op basis van algoritmes en patronen. Dit zorgt ervoor dat ook de output een zeker patroon volgt. Hierdoor lijken bijvoorbeeld AI-gegenereerde teksten over eenzelfde onderwerp vaak op elkaar. Daarnaast kunnen de algoritmes de gebruiker ook een bepaalde richting uitsturen door bijvoorbeeld niet alle opties te tonen. Ook hiervoor moet de gebruiker alert zijn wat een zeker bewustzijn vereist.

Naast de nadelen, ervaren of zien de deelnemers uit de focusgroepen ook verschillende voordelen: geïndividualiseerde feedback, herkenbaarheid output van de taalmodellen, het associatieve/creatieve vermogen van taalmodellen, schrijfstijl –of niveau aanpassen aan beginsituatie, tijdsbesparing en inspiratie voor werkvormen. De taalleerkrachten uit de focusgroepen geven allen les aan klasgroepen met veel cursisten. Dit kan het moeilijk maken om iedereen van geïndividualiseerde feedback te voorzien. AI kan hier een hulpmiddel bieden om toch iedereen van feedback op individueel niveau te voorzien. Dit voordeel werd echter slechts door één leerkracht aangehaald. Daarbij gaat het hier om een perceptie. De leerkracht in kwestie erkent de mogelijkheid van AI om hier een voordeel te bieden, maar paste het zelf ook nog niet toe. Daarnaast bespraken we onder de nadelen reeds de gestandaardiseerde output van AI-technologieën. De gestandaardiseerde output zorgt echter ook voor een herkenbaarheid van de output van de taalmodellen. Dit werd door de taalleerkrachten ook als voordeel ervaren als controlemechanisme voor misbruik van AI door cursisten. Onder het

associatieve en creatieve vermogen van AI verstaan we de mogelijkheid om connecties en linken te laten leggen door AI-systemen. Dit kan dienen ter inspiratie, voor het verruimen van de eigen blik of als basis voor lesmateriaal. We weten ondertussen dat AI-systemen niet feilloos zijn. Lena geeft in focusgesprek 1 echter aan dat bij het maken van connecties zelfs de foutieve connecties interessant kunnen zijn aangezien deze ons aan het denken kunnen zetten. Het creatieve aspect duidt op het vermogen van AI om inventieve werken te maken. Uiteraard werkt AI op basis van reeds bestaande data aangevuld met de input van de gebruiker. Het leggen van nieuwe connecties tussen verschillende types data, kan echter nieuwe, recyclede, creatieve werken opleveren die opnieuw kunnen dienen als bijvoorbeeld lesmateriaal. ChatGPT heeft tevens de capaciteit om de schrijfstijl of het leesniveau van een tekst aan te passen. Als taalleerkracht kan je een tekst invoeren in ChatGPT en vervolgens vragen om de schrijfstijl aan te passen door bijvoorbeeld de tekst te converteren in een ander genre of verteltrant. Daarnaast kan het artificieel aanpassen van woordkeuze of taalgebruik de taalleerkracht ook helpen in het aanpassen van een tekst naar het niveau van de cursisten. Een tekst kan zo moeilijker of gemakkelijker gemaakt worden. Ook dit is een zeer populair gebruik onder de participanten van de focusgesprekken. De voorlaatste subcategorie van de voordelen is de tijdsbesparing. Hieronder wordt verstaan dat het uitvoeren van bepaalde onderwijspraktijken sneller verloopt bij het inschakelen van AI. Tot slot is er de inspiratie voor werkvormen. Deze categorie kwam reeds aan bod onder de hoofdcategorie 'onderwijspraktijken'.

Hoofdcategorie zeven omvat alle aanbevelingen die de deelnemers van de focusgroepen formuleerden. Deze aanbevelingen zijn adviezen die, indien toegepast, de taalleerkrachten zouden helpen bij het implementeren van AI in hun onderwijspraktijk op een nuttige manier. Er werden zeven aanbevelingen gedaan: doelgericht gebruik door leerkrachten en cursisten stimuleren, betere samenwerking met gespecialiseerde taalapps, een continue *watch* van AI door experts, experimentatie met AI aanmoedigen, kennis/ervaringen over AI uitwisselen, prompts logboek/bibliotheek bijhouden en een eigen CVO Semper AI bot. Eerder werd bij de nadelen reeds aangegeven dat het doelgerichte gebruik van AI moeilijk is aangezien dit een zekere kennis vereist. Daarnaast is een ander nadeel dat cursisten AI kunnen misbruiken. Een aanbeveling die uit deze twee nadelen voortvloeit is het stimuleren van het doelgerichte gebruik, zowel door leerkrachten als door cursisten. Het doelgerichte gebruik is het gevolg van een kennisbasis over de AI-systemen. Zo leren zowel leerkrachten als cursisten hoe ze AI kunnen inzetten en waar het nuttig is. Voor leerkrachten kan dit efficiënt werk stimuleren waardoor zij zoveel mogelijk voordelen kunnen halen uit het werken met AI. Voor de cursisten betekent dit doelgerichte gebruik het genieten van de voordelen van AI als hulp terwijl ze zelf de leerdoelen halen. Er bestaan ook reeds taalapps die ondersteund worden door AI-technologie. Deze apps kunnen door taalleerkrachten als bedreiging gepercipieerd worden. Lena deed echter de aanbeveling om de meerwaarde van die apps te incorporeren in de taallessen van CVO Semper om een zo krachtig mogelijk leerklimaat te creëren. Om op de hoogte te blijven van alle ontwikkelingen en mogelijkheden op het vlak van artificiële intelligentie kan een AI-*watch* opgericht worden. Dit houdt in dat enkele experts binnen het CVO-team de veranderingen in het oog houden om zo de rest van het team te informeren bij interessante vondsten. De laatste drie subcategorieën zetten in op de kracht van het collectieve team om elkaar aan te moedigen en te inspireren om AI op een bruikbare manier te gebruiken. Experimentatie kan zowel individueel als in team gebeuren. Dit houdt in dat je, om AI correct te gebruiken, er gewoon mee aan de slag moet gaan en aan de hand van *trial-and-error* goed gebruik kunt leren. Het uitwisselen van kennis en ervaringen over AI bouwt hierop verder. Het gaat hier echt om het communiceren met elkaar. Men kan elkaar nieuwe dingen bijleren, alsook delen wat niet werkt om zo het collectieve succes te verhogen. Er werd ook voorgesteld om een AI-logboek –of bibliotheek bij te houden. In dit logboek zou je terecht kunnen om alle nuttige prompts te raadplegen. Dit maakt het gebruik van AI ook voor nieuwe gebruikers

toegankelijker. Tot slot sprak Maarten over een op maat gemaakte CVO Semper AI-bot, die gevoed zou worden met de reeds uitgebreide kennisdatabank van de school. Een op maat gemaakte CVO Semper bot zou een grote meerwaarde bieden voor de kwaliteit van de output omdat er enkel correcte data in de eigen databank zit en de data ook gerangschikt is volgens niveau. Aangezien de output veel pertinenter en correcter is, zouden leerkrachten nog veel meer tijd besparen door deze bot te gebruiken. Een op maat gemaakte bot is ook veel geschikter voor het evalueren van cursisten. Ook hier vermeldt Maarten dat het menselijke vermogen niet overbodig zal zijn. Wanneer de 'AI auto-marker' zou twijfelen, wordt de menselijke expertise weer ingeschakeld. Aan de hand van deze output leert het AI-systeem bij.

Tot slot zijn er naast de aanbevelingen ook enkele uitdagingen die hindernissen kunnen vormen voor het implementeren van AI door de taalleerkrachten van CVO Semper. Er zijn vijf uitdagingen: remmingen bij leerkrachten, *collective teacher efficacy* rond gebruik AI, cursisten experimenteren meer en beter met AI, toekomstige rol van leerkrachten, beperkte vaardigheden en beperkt gebruik bij leerkrachten. Er worden twee types remmingen erkend. Eerst en vooral weten leerkrachten soms niet waar en hoe te beginnen met het gebruik van AI. Die onwetendheid kan ervoor zorgen dat leerkrachten volledig wegblijven bij de technologieën. Het is duidelijk dat ChatGPT inmiddels gekend is door alle deelnemers en door velen op zekere hoogte gebruikt wordt. Verschillende leerkrachten gaven aan buiten ChatGPT geen andere AI-tools te kennen. Daarnaast is ook het gebruik met ChatGPT nog oppervlakkig. De leerkrachten zien de expliciete toepassing voor hun lespraktijk niet altijd. Ze zouden graag genieten van de voordelen van AI maar weten niet steeds hoe. Daarnaast kost het tijd om te starten met het werken met AI. Zoals eerder vermeld moet er geëxperimenteerd worden om een goede werkwijze te vinden en/of moet er een zekere kennis zijn van de systemen. Dit kan voor leerkrachten die het sowieso reeds druk hebben een belemmering vormen. Zoals vermeld bij de aanbevelingen willen de leerkrachten die deelnamen aan de focusgesprekken een collectief leerklimaat zien. *Collective teacher efficacy* houdt in dat leerkrachten zich verbonden voelen met het team om zo samen successen te bereiken, erkennen en vieren. Er is echter nog geen basis aanwezig op het vlak van AI-experimentatie en communicatie. Het zou een uitdaging kunnen zijn om alle leerkrachten van het team mee te krijgen. Daarnaast experimenteren cursisten wel reeds meer en beter met AI. Dit zorgt bij de leerkrachten ook voor een gevoel dat ze steeds achterophinken. Het ongekende zorgt vaak voor angsten. Een volgende uitdaging kan de angst zijn van leerkrachten ten opzichte van AI-systemen. Deze systemen worden vaak geadverteerd als de "nieuwe en persoonlijke tutor". Zo kunnen leerkrachten het gevoel krijgen dat hun positie in het gedrang komt. De laatste uitdaging zijn de beperkte vaardigheden en het beperkte gebruik van AI door de leerkrachten vandaag. Dit kwam in de focusgesprekken meermaals door verschillende personen aan bod. Zoals eerder vermeld vormt een beperkte kennis van AI-systemen en -doelen alsook een tekort aan experimentatie moeilijkheden op verschillende vlakken, zoals doelgericht gebruik. Ook dit kan de leerkrachten belemmeren om zich (verder) te verdiepen in de technologieën waardoor ze zich beperken tot hun gekende materialen.

Discussie – conclusie

- [Wat is samenvattend een antwoord op je initiële onderzoeksvraag?](#)

Hoe kunnen taalleerkrachten AI inzetten in hun onderwijspraktijk? Om deze vraag te kunnen beantwoorden, werden verschillende zaken onderzocht: de visie van de leerkrachten, het huidige gebruik, de voordelen en kansen, en de nadelen en uitdagingen. Wat de visie betreft, zien de leerkrachten verschillende positieve aspecten van AI, zoals snelle informatieverwerking en probleemoplossend vermogen, maar ze benadrukken ook het belang van het menselijke aspect. Daarnaast kwamen zorgen naar voren over overmatige afhankelijkheid van AI, mogelijk verlies van kritisch denkvermogen en een groeiende sociale kloof door technologische ongelijkheid. Wat het huidige gebruik betreft, is ChatGPT bij uitstek de meest gebruikte tool. Andere AI-tools die vallen onder de grote taalmodellen, beeldmodellen en machinaal leren werden vaak slechts door één persoon gebruikt. AI wordt ingezet voor alle taalverwervingsaspecten: spreekvaardigheid, luistervaardigheid, leesvaardigheid en schrijfvaardigheid. Vooral voor de hogere taalniveaus blijkt AI nuttig te zijn doordat de tijdsbesparing significant is. AI wordt tevens gebruikt voor het bieden van inspiratie voor werkvormen, het opstellen van oefeningen en evaluaties en het evalueren van cursisten. De evaluatieve functie is nog niet zeer populair. De voordelen van AI omvatten inspiratie voor schrijfoopdrachten of werkvormen, de herkenbaarheid van de output van taalmodellen, het aanpassen van schrijfstijl –of niveau aan de beginsituatie, het associatieve en creatieve vermogen, tijdsbesparing en geïndividualiseerde feedback. Experimenteren en kennisdeling zijn cruciaal om collectief van deze voordelen te kunnen genieten. De leerkrachten zijn zich echter ook bewust van de nadelen van AI, zoals de onbetrouwbaarheid van AI-output, het onnatuurlijke taalgebruik van AI, de beperkingen van huidige modellen en sturende algoritmes. Bovendien vereisen effectieve AI-toepassingen specifieke kennis en vaardigheden, wat tijd en training vraagt. Daarnaast moeten de leerkrachten nu ook rekening houden met de groeiende gemakzucht bij cursisten waardoor leerdoelen niet gehaald worden. Momenteel voelen de leerkrachten een zekere onmacht bij het controleren van het gebruik van AI waardoor ze kiezen voor het reageren op misbruik met behulp van AI of preventief oefeningen in de les laten maken. Ook hangt er een onzekerheid over de toekomst van de rol van leerkrachten in een AI-gedomineerd onderwijslandschap, wat angst en weerstand kan oproepen. Er moet rekening gehouden worden met al deze nadelen en uitdagingen om AI zo efficiënt mogelijk in te zetten in de onderwijspraktijk.

- [Geef een interpretatie van je onderzoeksresultaten en koppel terug naar je theoretisch/conceptueel kader.](#)

We kunnen het gebruik van AI voor taalleerkrachten koppelen aan het TPACK-model. Dit model geeft weer welke types kennis een leerkracht nodig heeft om technologie, pedagogie en inhoud te kunnen combineren. Onder de *technological knowledge* (TK) verstaan we de kennis van de leerkrachten van AI-technologieën, en training en experimentatie. Zoals vermeld in de resultaten is er een kennisbasis en experimentatie nodig om AI-tools doelgericht te kunnen gebruiken. Wat de *pedagogical knowledge* (PK) betreft, verstaan we de onderwijspraktijken in het algemeen, zoals het genereren van oefenmateriaal en het bieden van gepersonaliseerde feedback. Daarnaast moeten de leerkrachten hun pedagogie aanpassen op het gebruik van AI door cursisten door bijvoorbeeld meer klassikale opdrachten te geven. De *content knowledge* (CK) omvat alle taalinhoud en taalverwervingsaspecten waarvoor de leerkrachten AI-hulp inschakelen. We vermeldden schrijfvaardigheid, spreekvaardigheid, luistervaardigheid, leesvaardigheid, woordenschat- en grammaticaonderwijs.

De leerkrachten zijn reeds vertrouwd met de *pedagogical content knowledge* (PCK) waarbij ze gepaste didactische methoden ontwikkelen voor taalonderwijs. We concluderen dat leerkrachten tevens AI kunnen integreren in de pedagogiek (*technological pedagogical knowledge* (TPK)) en dit reeds op zekere hoogte doen. Hiervoor is het cruciaal dat de leerkrachten weten hoe ze AI-technologieën kunnen gebruiken om pedagogische doelen te ondersteunen. Tot slot kunnen AI-tools ingezet worden voor taalonderwijs specifiek (*technological content knowledge* (TCK)). Kennis over gespecialiseerde taalapps –en technologieën zijn hier relevant.

Kennis en experimentatie door uitwisseling, bijscholing en andere hulpmiddelen kunnen de leerkracht het vermogen bieden om technologische, pedagogische en inhoudelijke kennis op een holistische wijze te integreren om zo AI effectief te kunnen inzetten in hun onderwijspraktijk.

- [Wat zijn de sterktes en/of beperkingen van je studie?](#)

We identificeren volgende sterktes:

- In ons onderzoek hebben we ons specifiek gefocust op de toepassing van GenAI in het volwassenenonderwijs aangezien er daar nog niet veel studies over gepubliceerd zijn.
- We werkten met een kwalitatieve bevraging, in plaats van met een kwantitatieve bevraging, waarbij getoetst werd naar de percepties, ervaringen en noden van de taalleerkrachten. In de focusgroepen konden we ingaan op wat de deelnemers zeiden om zo meer informatie te weten te komen, iets wat je bij een survey niet kan doen. Daarnaast gingen de deelnemers ook in interactie met elkaar, wat ook voor nieuwe inzichten zorgde.

We identificeren volgende zwaktes:

- De resultaten van ons onderzoek zijn niet per se generaliseerbaar naar andere mensen of settings.
- De deelnemers die meededen aan onze focusgroepen gaven zich vrijwillig op, wat impliceert dat ze reeds enige interesse en motivatie hebben om bij te leren over artificiële intelligentie en hier eventueel zelf mee aan de slag te gaan. Er was geen enkele participant met een uitgesproken mening tegen AI. Dit wil niet zeggen dat er geen taaldocenten in CVO Semper zijn met een pessimistische visie op AI. De focusgroepen belicht dus slechts een deel van de ervaringen en percepties van de taalleerkrachten in CVO Semper.
- In een focusgroep kunnen de deelnemers elkaar beïnvloeden waardoor iemand met een andere mening zijn mening misschien niet gaat durven uiten.

- [Welke aanbevelingen voor de onderwijspraktijk en beleid kan je maken op basis van je studie?](#)

Op basis van de aanbevelingen van de deelnemers van de focusgroepen, doen wij de volgende aanbevelingen:

- Elke leerkracht zou zoveel mogelijk moeten experimenteren met AI om AI doelgericht te leren gebruiken. Bovendien is het ook aangeraden om een prompts logboek bij te houden (micro-niveau).
- Binnen CVO Semper moet er op meso-niveau meer samengewerkt worden om het volgende te realiseren:
 - Een prompts bibliotheek aanleggen
 - Een continue AI *watch* door de experts binnen CVO Semper
 - Kennis/ervaringen over AI uitwisselen door op regelmatige basis samen te komen.

Volgend jaar zal er bij Onderzoekende School onderzocht worden hoe CVO Semper professionele leergemeenschappen kan opstarten zodat er tegemoet wordt gekomen aan de vraag van leerkrachten voor meer ondersteuning bij de implementatie van AI in hun onderwijspraktijk.

Referentielijst

Noteer hier de geraadpleegde bronnen. Hanteer APA 7 als referentiestijl.

- Albin-Clark, A., Howard, T.L.J., Anderson, B. (2011). Real-time computer graphics simulation of blockplay in early childhood. *Computers & Education*, 57 (4), p. 2496–2504.
- Bishop, J.M. (2021). Artificial Intelligence is stupid and causal reasoning will not fix it. In *Frontiers in Psychology*, 11, 2603.
- Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77-101. 10.1191/1478088706qp063oa.
- Cambridge University. (2023). *English language education in the era of generative AI: our perspective*. Geraadpleegd op 25 mei 2024, van <https://www.cambridgeenglish.org/Images/685411-english-language-education-in-the-era-of-generative-ai-our-perspective.pdf>
- China Institute for Science and Technology Policy at Tsinghua University (2018), *AI Development Report*. China's State Council (2017), Next Generation Artificial Intelligence Development. Geraadpleegd op 25 mei 2024, van [China AI development report 2018.pdf \(indianstrategicknowledgeonline.com\)](#)
- Content Studio. (2021, 15 april). *Wat is Machine Learning?* Geraadpleegd op 22 februari 2024, van <https://www.purestorage.com/nl/knowledge/what-is-deep-learning.html>
- Edmett, A., Ichaporia, N., Crompton, H., & Crichton, R. (2023). *Artificial intelligence and English language teaching: Preparing for the future*. British Council. <https://doi.org/10.57884/78EA-3C69>
- Floridi, L. (2023). AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models. In *Psychology & Technology*, 36(1). <https://doi.org/10.1007/s13347-023-00621-y>
- Ghosh, B. (2023). The Rise of Small Language Models— Efficient & Customizable. *Medium*. Geraadpleegd op 21 februari 2024, van <https://medium.com/@bijit211987/the-rise-of-small-language-models-efficient-customizable-cb48ddee2aad>
- High Level Expert Group on Artificial Intelligence (2019), *A definition of AI: Main capabilities and disciplines*. 88 HM Government: Department for Business, Energy & Industrial Strategy (2017) Industrial Strategy. Building a Britain fit for the future. Geraadpleegd op 25 mei 2024, van [A definition of Artificial Intelligence: main capabilities and scientific disciplines | Shaping Europe's digital future \(europa.eu\)](#)
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal Of Education*, 57(4), 542–570. <https://doi.org/10.1111/ejed.12533>
- IBM (z.d.) *What is Natural Language Processing?* | IBM. Geraadpleegd op 21 februari 2024, van <https://www.ibm.com/topics/natural-language-processing>
- Janiesch, C., Zschech, P. & Heinrich, K. (2021). Machine learning and deep learning. *Electron Markets* 31, 685–695 (2021). <https://doi.org/10.1007/s12525-021-00475-2>

- Jones, A. (2011). Philosophical and socio-cognitive foundations for teaching in higher education through collaborative approaches to student learning. *Educational Philosophy and Theory*, 43 (9), p. 997–1011.
- Maedche, A., et al. (2019). AI-Based Digital Assistants: Opportunities, Threats, and Research Perspectives. *Bus Inf Syst Eng*, 61(4), 535-544. <https://doi.org/10.1007/s12599-019-00600-8>
- Martens, M., De Wolf, R., Evens, T. (z.d.). *Algoritmes in de onderwijscontext: Een studie naar de perceptie, mening en houding van leerlingen en ouders in Vlaanderen*. Kenniscentrum Data & Maatschappij. Geraadpleegd op 21 februari 2024, van [https://data-en-maatschappij.ai/uploads/Rapport Algoritmes in het onderwijs.pdf](https://data-en-maatschappij.ai/uploads/Rapport_Algoritmes_in_het_onderwijs.pdf)
- McCarthy, J. (2007) *What is Artificial Intelligence*. Geraadpleegd op 25 mei 2024, van [whatisai.pdf \(stanford.edu\)](https://whatisai.stanford.edu)
- McLaren, B.M., DeLeeuw, K.E., Mayer, R.E. (2011). Polite web-based intelligent tutors: can they improve learning in classrooms?. *Computers & Education*, 56 (3), p. 574–584.
- Melo, F.R. Flores, E.L. Carvalho, S.D. Teixeira, R.A.G., Loja, L.F.B., de Sousa Gomide, R. (2014). Computational organization of didactic contents for personalized virtual learning environments. *Computers & Education*, 79, p. 126–137.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017-1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Mitchell, T.M. (1997). *Machine learning*. 1. NY: McGraw-Hill.
- Nadkarni, P.M., Ohno-Machado, L., Chapman, W.W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association*, 18 (5), p. 544–551, <https://doi.org/10.1136/amiajnl-2011-000464>
- Nguyen, B.A., Yang, D.L. (2012) A semi-automatic approach to construct Vietnamese ontology from online text. *International Review of Research in Open and Distributed Learning*, 13 (5), p. 148–172.
- Seni, D.A. (2012). Do the modern neurosciences call for a new model of organizational cognition?. *Science & Education*, 21 (10), p. 1485–1506.
- Thuy, D. T., Da, C., & Van Hanh, N. (2024). The use of ChatGPT in teaching and learning: a systematic review through SWOT analysis approach. *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1328769>
- Van Droogenbroeck., K. (6/03/2024). ‘Mijn lessen voorbereiden gaat nu dubbel zo snel als vroeger’: veel leerkrachten omarmen AI in de klas. *De Morgen*. Geraadpleegd op 30 mei 2024, van <https://www.demorgen.be/nieuws/mijn-lessen-voorbereiden-gaat-nu-dubbel-zo-snel-als-vroeger-veel-leerkrachten-omarmen-ai-in-de-klas~b95c2380/>
- Xu., J., Jones., E., Laxton., V., & Galaczi., E. (2021) Assessing L2 English speaking using automated scoring technology: examining automarker reliability, *Assessment in Education: Principles, Policy & Practice*, 28:4, 411-436, DOI:10.1080/0969594X.2021.1979467

Yang, C.W., Kuo, B.C., Liao, C.H. (2011). A HO-IRT based diagnostic assessment system with constructed response items. *Turkish Online Journal of Educational Technology-TOJET*, 10 (4), p. 46–51.

Zhai, X., Chu, X., Chai, C. S., Jong, M. S., Starčič, A. I., Spector, M., Liu, J., Jing, Y., & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, 2021, 1–18. <https://doi.org/10.1155/2021/8812542>