

# Het effect van de ervaring van basisscholen met tweetalig onderwijs op de leerprestaties

*Worden de Engelse taalvaardigheden van leerlingen beïnvloed door de mate van ervaring?*

Florine van den Bulk

6168981

MA Meertaligheid en Taalverwerving

Masterscriptie

Universiteit Utrecht

Begeleider: Dr. Anne-France Pinget

Tweede lezer: Dr. Jacomine Nortier

Datum: Juli 2022

Aantal woorden: 13.207



**Universiteit  
Utrecht**

## Samenvatting

Het tweetalig primair onderwijs (tpo) in Nederland staat nog in de kinderschoenen. In het schooljaar 2014/2015 is de pilot begonnen met het invoeren van tweetalig onderwijs in groep 1, waarna er elk jaar trapsgewijs nieuwe tweetalige groepen 1 zijn bijgekomen. Dit betekent dat in het schooljaar 2020/2021 de eerste lichting van tweetalige groepen 7 is ontstaan, waarna in het schooljaar daarop de tweede lichting tot stand is gekomen. De tpo-scholen en hun leerkrachten bevinden dus zich nog in nieuw terrein wat betreft het onderwijzen van een tweetalige groep 7. De scholen hebben aangegeven dat zij naar aanleiding van ervaringen met de eerste lichting van tweetalige groepen 7 veranderingen in het curriculum door hebben gevoerd. Dit geeft aanleiding tot de hypothese dat het tweetalig onderwijs voor de tweede lichting is verbeterd. De verwachting is daarom dat de tweede lichting betere Engelse taalvaardigheden heeft dan de eerste lichting. Dit is in het huidige onderzoek getest door de scores op de productieve vaardigheden van de *English Benchmark* toets, namelijk spreken en schrijven, te vergelijken tussen de twee lichtingen. Om zeker te weten dat enig verschil tussen de lichtingen werd veroorzaakt door de ervaring van de school met het tpo-systeem, en niet door ongerelateerde factoren, was het echter noodzakelijk om ook een aantal kind- en schoolgebonden factoren mee te nemen in het onderzoek (e.g. de leeftijd van de leerling en het percentage onderwijstijd in het Engels). Op deze manier kon ook de invloed van deze factoren op de productieve taalvaardigheden worden bepaald. Uit het onderzoek bleek dat de twee lichtingen niet significant van elkaar verschilden wat betreft hun productieve Engelse taalvaardigheden. Wel werd er een effect gemeten van de meertaligheid van een leerling, de gehanteerde onderwijsmethode (OTOL, OSOL of Sandwich) en de ervaring van een leerkracht met het onderwijzen van het Engels als een vreemde taal. Uit het onderzoek kan geconcludeerd worden dat één jaar aan schoolervaring met het onderwijzen van een tweetalige groep 7 het onderwijs niet dusdanig verbetert dat er een effect op de Engelse taalvaardigheden van leerlingen zichtbaar is.

# Inhoudsopgave

Hoofdstuk 1: Inleiding .....	4
Hoofdstuk 2: Theoretisch kader .....	6
2.1 Productieve taalvaardigheden .....	6
2.2 Content and Language Integrated Learning (CLIL) en zijn effecten .....	8
2.3 Variabelen die van invloed zijn op de Engelse taalvaardigheden van T2-leerders .....	9
2.3.1 Moedertaal van de leerling .....	10
2.3.2 Meertaligheid van de leerling .....	10
2.3.3 Geslacht van de leerling .....	12
2.3.4 Leeftijd van de leerling .....	13
2.3.5 Ervaring van de leerkracht .....	14
2.3.6 Kwantiteit en kwaliteit van de taalinput .....	15
Hoofdstuk 3: Probleemstelling en hypotheses .....	17
Hoofdstuk 4: Methode .....	18
4.1 Participanten .....	18
4.2 English Benchmark .....	19
4.3 Analyse .....	20
Hoofdstuk 5: Resultaten .....	22
5.1 Correlatie tussen de twee taalvaardigheden .....	22
5.2 Voorwaardes van een meervoudige lineaire regressie .....	22
5.3 De spreekvaardigheid .....	23
5.4 De schrijfvaardigheid .....	27
Hoofdstuk 6: Discussie .....	31
Hoofdstuk 7: Conclusie .....	36
Bijlage A .....	46

## Hoofdstuk 1: Inleiding

In Nederland is het wettelijk verboden om meer dan 15% van de onderwijstijd op de basisschool in een andere taal dan het Nederlands aan te bieden (Expertisecentrum Nederlands, z.d.). Globalisering heeft echter geresulteerd in een verhoogde populariteit van Engelstalig onderwijs. In het schooljaar 2020/2021 waren er 1192 basisscholen die vroeg vreemdetalenonderwijs (vvto) aanboden; dat komt neer op zo'n 17,9% van alle basisscholen in Nederland (CBS, 2021; Nuffic, 2020). Vanwege deze grote belangstelling om kinderen zo vroeg mogelijk bloot te stellen aan het Engels, heeft het ministerie van Onderwijs, Cultuur en Wetenschap in 2014 een pilot tweetalig primair onderwijs (tpo) opgezet. Deze pilot houdt in dat er op zeventien basisscholen vanaf groep 1 tussen de 30 en 50% van de onderwijstijd in het Engels wordt lesgegeven (Nuffic, z.d.). Dit proces is stapsgewijs gegaan, wat betekent dat de pilot in groep 1 begon en er elk schooljaar een nieuwe groep 1 bijkwam. Zodoende is in het schooljaar 2020/2021 de eerste lichting van groepen 7 ontstaan die tweetalig onderwijs hebben gevolgd.

De Engelse taalvaardigheden van deze eerste lichting zijn gevolgd via het Flankerend onderzoek Tweetalig primair onderwijs (FoTo) (Driessen et al., 2016; Jenniskens et al., 2018; Jenniskens et al., 2020). Alle daaropvolgende lichtingen worden daarentegen buiten beschouwing gelaten in dit onderzoek. Het is echter aannemelijk dat er een verschil zit in de taalvaardigheden tussen de eerste lichting en latere lichtingen. Omdat tpo zich in Nederland nog in zo'n vroeg stadium bevindt, is er namelijk geen duidelijk kader van hoe het onderwijs moet worden ingericht. De tpo-scholen zijn dus afhankelijk van hun eigen ervaringen om erachter te komen wat goed werkt en wat niet. Tijdens mijn stage kwam ik in contact met deze scholen en kreeg ik de kans om kennis te maken met het tweetalig onderwijs. Uit deze gesprekken is naar voren gekomen dat de tpo-scholen op basis van hun ervaringen met de eerste lichting aanpassingen aan hun tweetalige onderwijs hebben gedaan. Deze aanpassingen werden vaak het jaar daarna meteen al ingevoerd, waardoor de tweede lichting al heeft kunnen profiteren van de ervaring die de scholen hebben opgedaan met het tpo-systeem. Aanpassingen betroffen onder andere de leeftijd waarop kinderen mochten beginnen met lezen in het Engels, de lesmethodes die werden gebruikt en de hoeveelheid samenwerking tussen verschillende jaarlagen. Onderzoek naar het invoeren van tweetalig onderwijs in andere Europese landen heeft aangetoond dat dit soort veranderingen bijdragen aan de kwaliteit van het onderwijs (Bovellan, 2014; Fernández & Halbach, 2011; Infante, Benvenuto & Lastrucci, 2009). In dit onderzoek wordt daarom de vraag onderzocht of er een verschil in Engelse taalvaardigheden zit tussen de eerste en tweede lichting van groepen 7 uit het Nederlandse tweetalig primair onderwijs. De focus ligt hierbij op de productieve vaardigheden: spreken en schrijven.

Om de twee lichtingen op een valide manier met elkaar te kunnen vergelijken, is het echter noodzakelijk om ook allerlei kind- en schoolgebonden factoren mee te nemen, zoals het percentage onderwijstijd dat in het Engels wordt lesgegeven. Aangezien het effect van schoolervaring op de taalvaardigheden niet direct kan worden gemeten, zal een eventueel verschil in resultaten tussen de twee lichtingen in de eerste instantie geprobeerd te worden verklaard door de invloed van deze factoren. Wanneer er alsnog een verschil overeind blijft nadat er wordt gecontroleerd voor deze factoren, kan er gesteld worden dat de ervaring van de school met het tpo-systeem bijdraagt aan de productieve Engelse taalvaardigheden van leerlingen.

Het maatschappelijk belang van dit onderzoek is het leveren van een bijdrage aan de inzichten over tweetalig primair onderwijs in Nederland. Aangezien dit onderwijssysteem zich nog in een vroeg stadium bevindt, is het belangrijk om de factoren die van invloed zijn op de Engelse taalvaardigheden van leerlingen te inventariseren. Daarnaast is er nog maar weinig onderzoek gedaan naar de directe invloed van schoolervaring met een onderwijssysteem op de schoolprestaties van leerlingen. Op deze manier kan de kwaliteit van het tweetalig onderwijs hopelijk worden verhoogd.

Deze scriptie is als volgt opgebouwd. In hoofdstuk 2 wordt een theoretisch kader geschetst, waarbij er wordt ingegaan op de factoren die een effect kunnen hebben op de tweedetaalvaardigheden van leerlingen. Daaropvolgend worden de onderzoeksvragen en hypothesen geformuleerd in hoofdstuk 3. In hoofdstuk 4 wordt de onderzoeksmethode besproken, waarna er wordt ingegaan op de resultaten in hoofdstuk 5. In hoofdstuk 6 worden de gevonden resultaten vervolgens ter discussie gesteld. Tot slot wordt er in hoofdstuk 7 een conclusie getrokken.

## Hoofdstuk 2: Theoretisch kader

In dit hoofdstuk wordt eerst achtergrondtheorie gegeven over de twee productieve taalvaardigheden en hoe deze met elkaar samenhangen. Vervolgens wordt er een korte omschrijving gegeven van het immersieonderwijs dat de tpo-scholen aanbieden, ook wel *CLIL* genoemd, en wat voor voordelen dit onderwijssysteem biedt. Tot slot wordt er ingegaan op de kind- en schoolgebonden factoren waarvan uit de literatuur bekend is dat deze van invloed zijn op de tweedetaalvaardigheden van leerlingen.

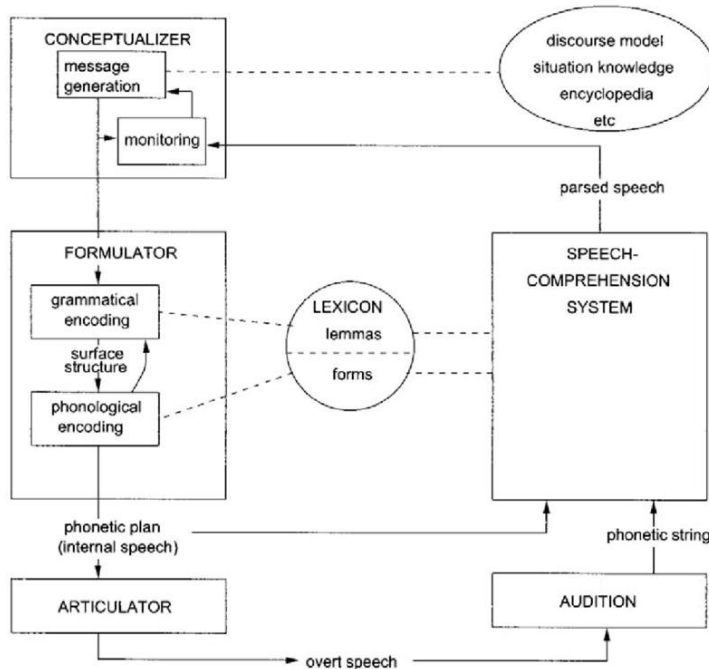
### 2.1 Productieve taalvaardigheden

Taalvaardigheid is een complex begrip dat uit veel verschillende onderdelen bestaat. Eén veelgebruikte manier om deze term onder te verdelen is in de receptieve vaardigheden en de productieve vaardigheden (Golkova & Hubackova, 2014). Spreken en schrijven staan bekend als de productieve vaardigheden omdat ze beide een vorm van taaloutput vereisen. Daartegenover staan de receptieve vaardigheden lezen en luisteren, die juist de verwerking van taalinput vergen (Masduqi, 2016).

In dit onderzoek ligt de focus op de productieve vaardigheden. Als onderdelen hiervan kunnen de volgende deelvaardigheden onderscheiden worden: uitspraak of spelling, woordenschat, grammatica, inhoud en structuur (Kwakernaak, 2009). Hoe deze onderdelen samenwerken tijdens het spreken kan deels gevisualiseerd worden door het klassieke taalproductiemodel van Levelt (1989), te vinden in Figuur 1. Dit model geeft een schematisch beeld van de processen die in het brein plaatsvinden wanneer een spreker taal produceert. In het Levelt-model bestaat de productie van spraak (aan de linkerkant van het schema) uit drie onafhankelijke systemen: de *conceptualizer*, de *formulator* en de *articulator*. In de *conceptualizer* wordt de boodschap die een spreker over wil brengen bedacht en vervolgens doorgestuurd naar de *formulator*. De boodschap van de *conceptualizer* is preverbaal, wat betekent dat deze nog niet in woorden wordt omschreven. Wanneer de boodschap aankomt in de *formulator* gebeuren er drie dingen. Lexicale inhoudswoorden worden geselecteerd uit het lexicon (lexicalisatie), morfemen worden aan de inhoudswoorden gehecht en functiewoorden worden toegevoegd (grammaticale codering) en de boodschap wordt omgezet tot een fonetisch plan (fonologische codering). Vervolgens wordt het fonetisch plan in de *articulator* omgezet tot spierbewegingen voor de articulatoren (tong, lippen, strottenhoofd etc.). Dit resulteert ten slotte in overte spraak. Voordat een spreker het fonetisch plan realiseert, heeft hij echter nog een mogelijkheid tot interne feedback doordat het plan naar zijn spraakbegripssysteem wordt gestuurd. Op deze manier kan de spreker fouten corrigeren voordat hij zijn uiting produceert. Soms worden fouten echter niet opgemerkt tijdens de interne feedback. In dit geval kan een spreker via het externe controlesysteem controleren wat voor spraak hij heeft geproduceerd, en zichzelf verbeteren indien nodig (Guerrero, 2005; Kwakernaak, 2009).

Alhoewel het Levelt-model op de productie van spraak is gebaseerd, komen de stappen die tijdens het schrijven nodig zijn overeen. Tijdens het schrijven wordt de articulatiestap echter vervangen door een transcriptiestap (Schoonen et al., 2009). In dat geval is er kennis van de spelling nodig in plaats van de uitspraak. Ondanks dat deze twee processen vergelijkbaar lopen, is spreken in een vreemde taal volgens Kwakernaak (2009) moeilijker dan schrijven. Dit komt doordat tijdens het spreken de taken veel sneller uitgevoerd moeten worden, waardoor de cognitieve belasting hoger ligt. Tijdens het schrijven

is er meer tijd om zowel de inhoud van de uiting te bedenken (*conceptualizer*), de inhoud om te zetten naar een correcte zin (*formulator*) en de juiste schrifttekens te vinden en neer te zetten op papier (*articulator*). Dit wordt ondersteund door bewijs dat T2-leerders een meer gevarieerde woordenschat gebruiken tijdens het schrijven dan tijdens het spreken (Baba, Takemoto & Yokochi, 2014; Kormos & Trebits, 2012).



Figuur 1. Taalproductiemodel van Levelt (1989, p. 9)

Er is nog maar weinig onderzoek uitgevoerd naar de relatie tussen de spreek- en schrijfvaardigheid bij T2-sprekers. De algemene bevindingen kunnen echter in twee punten worden samengevat volgens Kang (2020). De eerste bevinding is dat de ontwikkeling van de spreekvaardigheid niet altijd voorafgaat aan die van de schrijfvaardigheid, in tegenstelling tot bij moedertaalsprekers. Weissberg (2000) onderzocht bijvoorbeeld hoe volwassen Spaanse leerders van het Engels syntactische vaardigheden ontwikkelden in hun gesproken en geschreven taal. Hij ontdekte dat meer dan de helft van de nieuwe syntactische vormen voor het eerst in hun geschreven taal voorkwam, terwijl maar 20% voor het eerst in hun gesproken taal voorkwam. Daarnaast werden de syntactische vormen vaker correct geproduceerd in hun geschreven taal.

De tweede bevinding is dat er een grote individuele variabiliteit is in de ontwikkeling van de twee taalvaardigheden, en dat sommige leerders een asymmetrische ontwikkeling doormaken. Weissberg (2006) voerde bijvoorbeeld casestudies uit bij drie volwassen leerders met verschillende leeftijden, moedertalen en onderwijsachtergronden. Hij beschreef hoe zij hun spreek- en schrijfvaardigheid ontwikkelden tijdens een cursus Engels. Eén leerder, die een ervaren taalleerder was, toonde een gebalanceerde ontwikkeling tussen het spreken en schrijven. De andere twee leerders vertoonden echter een asymmetrische ontwikkeling. Bij de ene leerder verliep de ontwikkeling van het schrijven een stuk sneller, terwijl dit bij de andere leerder juist voor het spreken gold. Dit toont aan dat er geen vaste volgorde is waarop de productieve taalvaardigheden verworven worden. Daarnaast hebben Dickinson, Wolf en Stotsky (1993) laten zien dat de schrijfvaardigheid van T2-kinderen sterk verschilt per individu, en dat hun ontwikkeling niet op hetzelfde tempo verloopt.

Kortom gezien lijkt de schrijfvaardigheid van een tweede taal zich vaak sneller te ontwikkelen dan de spreekvaardigheid. Dit wordt aangetoond door de meer gevarieerde woordenschat en hogere syntactische correctheid in de geschreven teksten van tweedetaalleerders (Baba, Takemoto & Yokochi, 2014; Kormos & Trebits, 2012; Weissberg, 2000). Aangezien sommige leerders echter een asymmetrische ontwikkeling doormaken, hoeft dit niet voor iedereen te gelden (Weissberg, 2006). Hoe dan ook kan er geconcludeerd worden dat er veel individuele variabiliteit bestaat in de ontwikkeling van de productieve taalvaardigheden.

## **2.2 Content and Language Integrated Learning (CLIL) en zijn effecten**

Op tweetalige basisscholen wordt er gebruik gemaakt van immersieonderwijs. Dit wordt in de internationale literatuur ook wel *content and language integrated learning* (CLIL) genoemd. CLIL is een overkoepelende term die een breed spectrum aan educatieve benaderingen omvat. Onder deze term vallen alle lesprogramma's waarin een vreemde taal wordt gebruikt als instructietaal voor het onderwijzen van inhoudelijke vakken in het reguliere onderwijs (Gabillon, 2020). CLIL stimuleert een duale vorm van instructie, waarbij de focus wordt gelegd op zowel taal als inhoud (Coyle, Hood & Marsh, 2010). Dit houdt in dat het gericht is op het ontwikkelen van zowel vaardigheden in een inhoudelijk vak als in de taal waarin het vak wordt onderwezen (i.e. de doeltaal). CLIL-benaderingen kunnen echter verschillen in hun mate van intensiteit en blootstelling aan de vreemde taal; van programma's met een hoge intensiteit, zoals tweetalig onderwijs, tot programma's met een lage intensiteit, zoals losse CLIL-vakken (Mehisto, Marsh & Frigols, 2008).

De effectiviteit van CLIL is snel een aandachtspunt geworden van wetenschappelijke onderzoeken naar de tweedetaalverwerving. Uit deze onderzoeken blijkt over het algemeen dat CLIL-leerlingen de leerlingen uit reguliere klassen overtreffen op de T2-leerresultaten (Lorenzo et al., 2010; Pérez-Cañado, 2012). Van de Craen et al. (2007) benadrukken echter dat de positieve impact van CLIL op de T2-ontwikkeling consistent wordt gevonden in het primair onderwijs dan in het secundair onderwijs. De laatstgenoemde uitkomsten zijn minder uniform, met resultaten die veel meer lijken af te hangen van individuele variatie en leraar kenmerken, zoals attitude, taalvaardigheid en CLIL-training. Ook uit een recentere meta-analyse van Goris, Denessen en Verhoeven (2019) blijkt dat er veel variatie bestaat in de grootte van de gevonden effecten in het secundair onderwijs. Desondanks concluderen zij dat CLIL voor betere leerprestaties zorgt dan de reguliere vorm van talenonderwijs, waarbij leerlingen een tweede taal leren door middel van formele instructie tijdens één schoolvak.

Met betrekking tot de specifieke taalvaardigheden worden de receptieve vaardigheden het meest positief beïnvloed door CLIL (Mattheoudakis, Alexiou & Laskaridou, 2014). Interessanter voor het huidige onderzoek is echter het effect van CLIL op de productieve vaardigheden. Onderzoeken in dit gebied hebben inconsistente resultaten opgeleverd. In de eerste instantie is het niet duidelijk wat het effect van CLIL op de schrijfvaardigheid is. Dalton-Puffer (2008) heeft uit een systematische evaluatie geconcludeerd dat de schrijfvaardigheid niet wordt beïnvloed door CLIL-benaderingen. Ook recentere onderzoeken hebben geen significant verschil gevonden tussen CLIL-klassen en reguliere klassen op het gebied van de T2-schrijfvaardigheid (De Diezmas, 2016; Lahuerta, 2020; Roquet & Pérez-Vidal, 2017). Dit geldt voor zowel het primaire als secundaire onderwijs. Een aantal onderzoeken hebben echter wel positieve resultaten gevonden (Lasagabaster, 2008; Lorenzo, Casal & Moore, 2010; Van Mensel et al., 2020). In de tweede instantie lijkt de



sprekvaardigheid duidelijk te profiteren van een CLIL-aanpak, zoals blijkt uit onderzoeken van onder andere Delliou en Zafiri (2016), De Diezmas (2016) en Dalton-Puffer (2011). CLIL-leerders zijn in vergelijking met hun leeftijdsgenoten over het algemeen vloeiender in de T2. Hun mondelinge gebruik van de T2 is creatiever, ze zijn bereid risico's te nemen en ze experimenteren op uitdagende manieren met de taal (Mattheoudakis, Alexiou & Laskaridou, 2014).

Er zijn verschillende onderzoeken uitgevoerd om de oorzaken van de verbeterde T2-prestaties bij CLIL-benaderingen vast te stellen. Wolff (2007) noemt als belangrijkste factoren de grotere blootstelling, de frequentere aanwezigheid van moedertaalsprekers, de betere leeromstandigheden door authentiek lesmateriaal en de over het algemeen rijkere taalkundige inhoud van CLIL-lessen. Dalton-Puffer (2007) oppert dat het CLIL-klaslokaal een realistischere situatie biedt, waarin de vreemde taal echt wordt gebruikt. Dit levert volgens haar mogelijkheden op om informatie dieper te verwerken. Andere onderzoekers hebben aangetoond dat leerlingen in CLIL-programma's een grotere motivatie hebben, wat kan resulteren in betere leerprestaties (Navarro Pablo & García Jiménez, 2018; Pfenninger, 2016; Pladevall-Ballester, 2019).

In conclusie zien we dat CLIL binnen het primaire onderwijs een positief effect heeft op de algemene T2-vaardigheden van leerlingen. Het effect van CLIL op de schrijfvaardigheid is echter minder duidelijk, met sommige onderzoeken die een positief effect vinden (Lasagabaster, 2008; Lorenzo, Casal & Moore, 2010; Van Mensel et al., 2020) en sommige onderzoeken die geen effect vinden (De Diezmas, 2016; Lahuerta, 2020; Roquet & Pérez-Vidal, 2017). Voor de spreekvaardigheid worden consistentere resultaten gevonden, waarbij CLIL-leerlingen significant hoger scoren dan leeftijdsgenoten uit reguliere taalklassen (Dalton-Puffer, 2011; De Diezmas, 2016; Delliou & Zafiri, 2016).

### **2.3 Variabelen die van invloed zijn op de Engelse taalvaardigheden van T2-leerders**

De tweedetaalverwerving is een uiterst complex proces waarbij een verscheidenheid aan interne en externe variabelen met elkaar interageren en zo de leerprestaties beïnvloeden (Martinez Agudo, 2022). De interne variabelen hebben te maken met de kenmerken die elk individu met zich meebrengt naar de leercontext. Deze worden beïnvloed door andere variabelen die aanwezig zijn in de omgeving waarin een leerling zich bevindt, ook wel de externe variabelen genoemd. De interne variabelen omvatten volgens Dörnyei (2005) een breed scala aan kenmerken, zoals leeftijd, geslacht, taalaanleg, motivatie en eerdere taalervaring. Van dergelijke individuele kenmerken is aangetoond dat ze de snelheid waarmee een T2 wordt geleerd, beïnvloeden. Daarnaast spelen ook externe variabelen, zoals de leerkrachtkenmerken en de hoeveelheid blootstelling aan een taal, een rol tijdens de tweedetaalverwerving. In deze paragraaf wordt eerst ingegaan op de interne variabelen, waarna de invloed van de externe variabelen wordt besproken. De paragraaf begint met de rol die de moedertaal van een leerder speelt, en gaat daarbij in op de verschillen tussen T1- en T2-sprekers. Vervolgens worden de individuele kenmerken geslacht, leeftijd en meertaligheid besproken. Daarna wordt er ingegaan op het effect van de ervaring van de leerkracht met het onderwijzen van een vreemde taal op de leerprestaties. Tot slot wordt er gekeken naar de kwantiteit en kwaliteit van de taalinput die door de school wordt geboden. Bij de bespreking van elke variabele wordt de focus gelegd op de productieve vaardigheden waar mogelijk. Daarnaast worden er zoveel mogelijk onderzoeken uit een CLIL-context besproken, zodat de resultaten vergelijkbaar zijn met het huidige onderzoek.

### 2.3.1 Moedertaal van de leerling

Een belangrijke voorspeller van de leerprestaties op het Engels is de moedertaal van een leerling. Wanneer een leerling het Engels van huis uit spreekt, kan er immers verwacht worden dat zijn Engelse taalvaardigheden verder ontwikkeld zijn dan die van zijn Nederlandstalige klasgenoten. Meerdere onderzoeken hebben aangetoond dat tweedetaalsprekers lager scoren op gestandaardiseerde toetsen dan moedertaalsprekers. Dit geldt voor zowel hun syntactische en morfologische vaardigheden als hun woordenschat (Chondrogianni & Marinis, 2011; Scheele, Leseman & Mayo, 2009).

Een verschil tussen tweedetaalsprekers en moedertaalsprekers is ook aangetoond voor de productieve vaardigheden. Babayigit (2014) heeft onderzoek gedaan naar de Engelse schrijfvaardigheid van kinderen die van huis uit een andere taal dan het Engels spraken. Deze groep werd vergeleken met Engelstalige moedertaalsprekers die geen ervaring hadden met een andere taal. De participanten werden getoetst op vijf aspecten: woordenschat, spelling, structuur, inhoud en vloeiendheid. Zoals verwacht werd er een voordeel voor de moedertaalsprekers gevonden op alle aspecten, met uitzondering van de spellingvaardigheid. Dit komt overeen met de resultaten van Verheyden et al. (2010), waarin Turks-Nederlandse kinderen werd gevraagd een stripverhaal na te vertellen door deze in het Nederlands op te schrijven. Ook hier scoorde de eentalige controlegroep significant beter op elk gemeten aspect van de schrijfvaardigheid, behalve op de spellingvaardigheid.

Naar de invloed van de moedertaal op de spreekvaardigheid is minder onderzoek gedaan. Wel hebben Gagarina et al. (2015) en Blom en Boerma (2020) aangetoond dat de narratieve vaardigheden van successief tweetalige kinderen van vier tot negen jaar oud niet verschillen met die van eentalige leeftijdsgenoten. Daarnaast lijkt de uitspraak van successieve tweetaligen ook niet significant te verschillen met die van moedertaalsprekers (Abrahamsson & Hyltenstam, 2009; Alario et al., 2010).

De gevonden verschillen tussen moedertaalsprekers en tweedetaalsprekers kunnen door twee variabelen verklaard worden. Ten eerste speelt de *Age of Onset* (AoO) een rol, oftewel de leeftijd van eerste blootstelling aan een taal. Onderzoek heeft aangetoond dat de eerdere verwerving van een taal resulteert in significant betere taalvaardigheden (Bylund, Hyltenstam & Abrahamsson, 2020). Ten tweede is de cumulatieve Length of Exposure (CLoE) van invloed op de verwerving van een taal. De CLoE wordt berekend door informatie te verzamelen over het aantal jaren blootstelling aan een taal en het aantal uren dat per week wordt besteed aan het horen of gebruiken van de taal. Uit onderzoek is gebleken dat de CLoE één van de belangrijkste verklarende factoren is voor variabiliteit in de prestaties van tweetalige kinderen (Mitrofanova et al., 2018; Smolander et al., 2020; Unsworth, 2012).

Er kan kortom gesteld worden dat kinderen die het Engels als thuistaal spreken, betere Engelse taalvaardigheden zullen hebben dan kinderen die pas op school voor het eerst aan het Engels zijn blootgesteld. Als verklaring kunnen we hun hogere cumulatieve Length of Exposure en eerdere Age of Onset aanwijzen.

### 2.3.2 Meertaligheid van de leerling

Onderzoek in de afgelopen drie decennia heeft aangetoond dat meertaligheid gepaard gaat met een verscheidenheid aan voordelen (Montanari, 2019). In het cognitieve domein gaat het kennen van meer dan één taal samen met betere executieve functies, zoals een verhoogd vermogen tot selectieve aandacht (Bialystok & Barac, 2013). In het linguïstische domein vertonen meertalige sprekers een groter metalinguïstisch bewustzijn dan eentalige sprekers,

en ontwikkelen ze dit bewustzijn op een jongere leeftijd (Allgäuer-Hackl & Jessner, 2019). Het is gesuggereerd dat dergelijke cognitieve en linguïstische voordelen resulteren in het gemakkelijker leren van nieuwe talen bij sprekers die al minstens twee talen kennen (Paradis, 2008; Witney & Dewaele, 2018). Bij het leren van een nieuwe taal zijn meertaligen inderdaad doorgaans beter dan sprekers die slechts ervaring hebben met één taal (Hirosh & Degani, 2018). Festman (2020) onderscheidt drie effecten die dit voordeel bij meertaligen kunnen verklaren: een faciliterend effect, katalyserend effect en modulerend effect.

Bij het leren van een nieuwe taal kunnen leeders profiteren van hun beschikbare taalrepertoire. Dat wil zeggen dat zij gebruik kunnen maken van de reeds bestaande taalkennis die is opgeslagen in hun geheugen. Hierdoor kunnen meertalige sprekers een *faciliterend effect* ervaren bij het leren van nieuwe talen. Dit effect wordt ook wel positieve transfer genoemd (Festman, 2020). Grammaticale structuren, klanken en cognaten kunnen bijvoorbeeld overgedragen worden naar de nieuw geleerde taal (Jarvis & Pavlenko, 2008). Bij de verwerving van derde talen is linguïstische transfer vanuit beide andere talen waargenomen (Krenca, Hipfner-Boucher & Chen, 2020; Leung, 2005).

Het tweede effect dat volgens Festman (2020) het leren van talen gemakkelijker maakt bij meertaligen is een *katalyserend effect*. In plaats van de directe transfer van talige kennis, heeft dit effect te maken met de transfer van vaardigheden die zijn ontwikkeld op basis van eerdere taalleerervaringen. Trembley en Sabourin (2012) hebben bijvoorbeeld aangetoond dat het spreken van meerdere talen bijdraagt aan de ontwikkeling van verbeterde spraakperceptiecapaciteiten. Dit resultaat werd geïnterpreteerd als een gevolg van de ervaring die meertaligen hebben met het leren van nieuwe klanken.

Tot slot omschrijft Festman (2020) het *modulerende effect* dat van invloed is op meertalige sprekers. Hoe meer talen iemand kent, hoe gemakkelijker het is om hun bijzonderheden op te merken en hun overeenkomsten te identificeren. Dit wordt ook wel een metalinguïstisch bewustzijn genoemd. Onderzoek van Rauch, Naumann en Jude (2012) heeft laten zien dat meertaligen inderdaad een hoger metalinguïstisch bewustzijn hebben dan eentaligen.

Samengevoegd kunnen deze effecten ervoor zorgen dat meertaligen een nieuwe taal sneller zullen oppikken dan eentaligen. Deze bewering wordt ondersteund door meerdere onderzoeken die de verwerving van een vreemde taal op school tussen eentaligen en tweetaligen hebben vergeleken. Zo onderzochten Hopp et al. (2019) de productieve woordenschat in het Engels bij kinderen van acht tot tien jaar oud. Ze rapporteerden dat tweetalige studenten beter presteerden dan hun eentalige leeftijdsgenoten wanneer er rekening werd gehouden met de individuele kenmerken van leerlingen. Ook andere onderzoeken hebben een voordeel voor meertaligen gevonden op hun algemene taalvaardigheid van een derde taal (Abu-Rabia & Sanitsky, 2010; Fleckenstein, Möller & Baumert, 2018; Maluch & Kempert, 2017). Weinig onderzoeken hebben echter specifiek gekeken naar verschillen in de spreek- en schrijfvaardigheid tussen eentaligen en meertaligen. Een dergelijk onderzoek is uitgevoerd door Mady (2014), die eentalige kinderen van elf jaar oud vergeleek met tweetalige leeftijdsgenoten op hun taalvaardigheden van het Frans als een vreemde taal. Haar onderzoek omvatte twee tweetalige groepen: immigranten die tijdens de basisschool in Canada waren komen wonen (successieve tweetaligen) en kinderen die in Canada waren geboren (simultane tweetaligen). Beide tweetalige groepen scoorden hoger dan de eentaligen op het schrijfonderdeel van de toets. Op het spreekonderdeel werd echter alleen een voordeel gevonden voor de immigranten. Met behulp van een regressieanalyse toonden

de auteurs aan dat dit voordeel te wijten was aan sociale factoren, waaronder de bereidheid om te communiceren.

In conclusie lijkt er een duidelijk voordeel voor meertalige leerlingen te bestaan bij het leren van een nieuwe taal. Op basis van het onderzoek van Mady (2014) kan er verwacht worden dat meertalige leerlingen hoger zullen scoren op de schrijfvaardigheid. Daarnaast kan er verwacht worden dat meertalige leerlingen ook beter zullen presteren op de spreekvaardigheid, vanwege hun grotere productieve woordenschat en betere algemene taalvaardigheden (Abu-Rabia & Sanitsky, 2010; Fleckenstein, Möller & Baumert, 2018; Hopp et al., 2019; Maluch & Kempert, 2017).

### 2.3.3 *Geslacht van de leerling*

Een ander individueel kenmerk dat de tweedetaalverwerving kan beïnvloeden is het geslacht van de leerling. In het onderzoeksveld van de eerstetaalverwerving bestaat er een enorme hoeveelheid literatuur over geslachtseffecten, en wordt geslacht over het algemeen geaccepteerd als een factor die verschillen tussen mannelijke en vrouwelijke leerlingen kan verklaren (Etchell et al., 2018). Daarentegen heeft de invloed van geslacht op de tweedetaalverwerving relatief weinig aandacht gekregen (Martínez Agudo, 2022). Het schaarse bewijs dat tot nu toe beschikbaar is over geslachtsverschillen in T2-onderzoek komt echter overeen met de bevindingen die zijn verkregen in T1-onderzoek, waarbij er een licht voordeel voor vrouwen bestaat. Zo hebben Van der Slik, Van Hout en Schepens (2015) aangetoond dat volwassen vrouwelijke leerders consequent beter presteren dan mannelijke leerders op de spreek- en schrijfvaardigheden van een tweede taal. Ook voor jongere leerders heeft onderzoek laten zien dat meisjes beter schrijven in de tweede taal dan jongens (Babayigit, 2014; Sabra, 2018).

De T2-spreekvaardigheid lijkt eveneens beïnvloed te worden door het geslacht, zij het in mindere mate dan de schrijfvaardigheid. Erdiana, Bahri Ys en Akhmal (2019) en Walczak & Geranpayeh (2015) hebben beide aangetoond dat vrouwelijke middelbare scholieren significant betere spreekvaardigheden hebben voor het Engels als een tweede taal dan mannelijke scholieren. In beide onderzoeken was het verschil echter klein genoeg dat de auteurs het verwaarloosbaar noemden.

Verklaringen voor de geslachtsverschillen in tweedetaalverwerving kunnen worden gevonden in sociaal-psychologische dimensies zoals leerstrategieën, motivaties en attitudes (Iwaniec, 2019; Salahshour et al., 2013; Sroka et al., 2015). Vooral de rol van motivatie heeft veel aandacht gekregen door onderzoekers. Uit onderzoek uit verschillende landen werd geconcludeerd dat vrouwelijke leerlingen meer gemotiveerd zijn om een tweede taal te leren dan hun mannelijke leeftijdsgenoten (Iwaniec, 2019; Öztürk & Gürbüz, 2013; Thompson & Kerstin Sylvén, 2015). Aangezien motivatie een belangrijke factor is voor de tweedetaalverwerving (Hernández, 2008), kan dit de gevonden geslachtsverschillen mogelijk verklaren.

Ook in de context van CLIL is er onderzoek gedaan naar de rol van geslacht. Onderzoekers zoals Lahuerta (2020) hebben gesuggereerd dat een CLIL-benadering ertoe kan leiden dat geslachtsverschillen bij het leren van een tweede taal worden geëlimineerd, aangezien CLIL de motivatie van leerlingen kan vergroten. Onderzoeken naar motivatie bij reguliere leerlingen en CLIL-leerlingen hebben inderdaad laten zien dat verschillen tussen jongens en meisjes verdwijnen in een CLIL-context (Doiz et al., 2014; Gallardo-del-Puerto & Blanco-Suárez, 2021; Heras & Lasagabaster, 2015; Merisuo-Storm, 2007). Of dit ook van

invloed is op de leerprestaties, is echter nog niet duidelijk. Onderzoek van Lahuerta (2020) wijst erop van wel, aangezien vrouwelijke deelnemers niet significant beter presteerden dan hun mannelijke leeftijdsgenoten op schriftelijke vaardigheden in de CLIL-groep, terwijl dit wel het geval was in de reguliere groep. Deze resultaten worden echter tegengesproken door die van Roquet, Llopis en Pérez-Vidal (2016) en Lasagabaster (2008), die allebei een hogere schrijfvaardigheid bij vrouwelijke CLIL-leerlingen waarnamen. Lasagabaster (2008) vond daarentegen geen significant effect van geslacht op de spreekvaardigheid. Op de vraag of CLIL-programma's helpen om geslachtsverschillen in leerprestaties te compenseren, kan dus nog geen definitief antwoord worden gegeven.

Concluderend lijkt het zo te zijn dat de T2-spreekvaardigheid niet of nauwelijks wordt beïnvloed door het geslacht van een leerder. In de context van formele instructie blijkt het effect zo klein te zijn dat deze als verwaarloosbaar kan worden beschouwd, terwijl er in CLIL-context helemaal geen effect is gevonden (Erdiana, Bahri Ys & Akhmal, 2019; Lasagabaster, 2008; Walczak & Geranpayeh, 2015). Op de T2-schrijfvaardigheid lijkt geslacht een grotere rol te spelen, alhoewel het onduidelijk is in hoeverre dit effect overeen blijft in CLIL-programma's (Babayigit, 2014; Lahuerta, 2020). Op basis van de onderzoeken van Roquet, Llopis en Pérez-Vidal (2016) en Lasagabaster (2008) kan er een klein effect van geslacht worden verwacht op de T2-schrijfvaardigheid, waarbij meisjes beter zullen scoren dan jongens.

#### *2.3.4 Leeftijd van de leerling*

Een ander individueel kenmerk dat van invloed is op de T2-taalvaardigheden, is de leeftijd van een leerling. In het Nederlandse onderwijssysteem wordt over het algemeen 1 oktober gebruikt als de grensdatum voor de klassengroepering van leerlingen (Onderwijsinspectie, z.d.). Dit betekent dat een kind dat in september is geboren, in hetzelfde schoolcohort kan zitten als een ander kind dat 12 maanden eerder is geboren. Dergelijke leeftijdsverschillen in een schoolcohort worden omschreven met de term relatieve leeftijd (Cobley et al., 2009). Er is sterk bewijs dat de oudste leerlingen uit een klas significant beter presteren in vergelijking met hun relatief jongere tegenhangers. Zo hebben zowel Dhuey et al. (2019) als Mavilidi et al. (2021) relatieve leeftijdseffecten gevonden op de lees-, schrijf- en rekenprestaties van leerlingen tussen de zes en achttien jaar oud. De effecten waren groot in het primaire onderwijs, maar namen af in het secundaire onderwijs.

Er zijn in de literatuur verschillende hypothesen voorgesteld om prestatieverschillen veroorzaakt door de relatieve leeftijd te verklaren. De meest waarschijnlijke hiervan is volgens Cobley et al. (2009) de maturatiehypothese. Deze stelt dat de minst volwassen leerlingen uit een schoolcohort waarneembare cognitieve verschillen zullen vertonen. Deze verschillen resulteren in uiteenlopende schoolprestaties bij leerlingen. Specifieke gegevens over variaties in de cognitieve ontwikkeling binnen één schoolcohort zijn echter schaars, waardoor de hypothese niet bevestigd kan worden.

Ook binnen het vreemdetalenonderwijs is er onderzoek gedaan naar het effect van leeftijd. Zo heeft Bret (2011) het effect van leeftijd op de spreekvaardigheid binnen de context van formele instructie en CLIL onderzocht. Voor het onderzoek werden leerlingen uit groep 7 en 8 getest. In groep 7 werd er geen significant verschil gevonden tussen de controlegroep en CLIL-groep, terwijl in groep 8 de CLIL-groep significant beter presteerde op de spreekvaardigheid. De auteur concludeerde dat CLIL een veel duidelijker effect lijkt te hebben op de oudere leerlingen dan op de jongere, ondanks dat deze maar één jaar van elkaar

verschilden. Artieda, Roquet en Nicolás-Conesa (2020) vonden een vergelijkbaar resultaat voor de schrijfvaardigheid, in dit geval bij leerlingen die twee jaar van elkaar verschilden.

Tot slot kan er ook worden gekeken naar de rol van leeftijd in het FoTo-onderzoek (Jenniskens et al., 2020). De meest recente meting, die in groep 5 werd afgenomen, heeft laten zien dat de relatieve leeftijd binnen één schoolcohort al een significante invloed kan hebben op de prestaties. Voor de productieve vaardigheden, die werden getoetst door middel van de onderdelen woordenschat en grammatica, bleek namelijk dat een hogere leeftijd een voorspeller was van een hogere score.

In conclusie kan worden gesteld dat de relatieve leeftijd van leerlingen een rol speelt bij hun schoolprestaties. Dit is aangetoond voor zowel de algemene schoolprestaties (Dhuey et al., 2019; Mavilidi et al., 2021; Sakic, Burusic & Babarovic, 2013) als de prestaties binnen CLIL-contexten (Artieda, Roquet & Nicolás-Conesa, 2020; Bret, 2011; Jenniskens et al., 2020). Tot slot hebben de onderzoeken van Bret (2011) en Artieda, Roquet en Nicolás-Conesa (2020) laten zien dat dit leeftijdseffect voor zowel de spreek- als schrijfvaardigheid bestaat. In het huidige onderzoek kan dus voor beide taalvaardigheden worden verwacht dat oudere leerlingen beter scoren.

### *2.3.5 Ervaring van de leerkracht*

Een externe factor die invloed kan hebben op de leerprestaties van leerlingen, is de ervaring van de leerkracht. De belangrijkste taak van een leerkracht is om leerlingen betrokken te krijgen bij het bestuderen van de lesinhoud. Soms doen zich tijdens dit proces echter problemen voor, en leerkrachten ontwikkelen in de loop van de tijd routines om terugkerende problemen aan te pakken (Lampert, 2003). Het is dus aannemelijk dat er een relatie bestaat tussen het aantal jaren aan ervaring van een leerkracht en het leerproces van zijn leerlingen. Een grote hoeveelheid literatuur ondersteunt dit idee. De afgelopen decennia zijn er veel onderzoeken uitgevoerd die beginnende leraren hebben vergeleken met deskundige leraren (Stigler & Miller, 2018). Deze onderzoeken vinden over het algemeen een zwak maar significant effect. Zo hebben Nye, Konstantopoulos en Hedges (2004) onderzoek gedaan naar de relatie tussen leerkrachtkenmerken en de lees- en rekenprestaties van kinderen uit groep 2 tot en met 5. Hieruit bleek dat het aantal jaren ervaring van de leerkracht een lichte invloed heeft op de leerprestaties van leerlingen. Rockoff (2004) heeft in een vergelijkbaar onderzoek ook aangetoond dat er een zwak effect van de leerkrachtervaring op de lees- en rekenprestaties bestaat. Het effect van leerkrachtervaring bleek echter voor sommige onderdelen af te zwakken na zes jaar, wat aantoont dat na een bepaald punt extra onderwijservaring geen invloed meer heeft.

Ook de ervaring die leerkrachten hebben met CLIL-programma's kan een effect hebben op de prestaties van hun leerlingen. Zo hebben Fernández en Halbach (2011) vier jaar na de start van een tpo-project in Madrid onderzoek gedaan naar de ervaringen van leerkrachten. Een groot deel van de leerkrachten gaf aan dat hun manier van lesgeven was veranderd tijdens het project. Deze methodologische veranderingen schreven zij deels toe aan hun ervaringen met tweetalig lesgeven, zoals de noodzaak om meer visuele ondersteuning te gebruiken om ervoor te zorgen dat leerlingen de uitleg begrijpen, en meer hands-on onderwijs om de cognitieve belasting bij leerlingen te verlagen. Onderzoek heeft aangetoond dat er een significant verband bestaat tussen lesmateriaal en leerprestaties, en dat de verschillende manieren waarop materialen in de klas worden gebruikt een grote impact op het leerproces van leerlingen kunnen hebben (Hildén, 2011, geciteerd in Bovellan, 2014). Zodoende kunnen

de methodologische veranderingen die CLIL-leerkrachten door middel van hun ervaring doorvoeren, de taalvaardigheden van hun leerlingen positief beïnvloeden.

Een vergelijking tussen de manier waarop beginnende en ervaren CLIL-leerkrachten lesgeven, is uitgevoerd door Custodio-Espinar (2020). In haar onderzoek evalueerden leerkrachten hun eigen lespraktijken op basis van hoe vaak ze rekening hielden met CLIL-principes tijdens het plannen van lessen. De vragenlijst was ontworpen op basis van een reeks methodologische principes die volgens de literatuur aanwezig horen te zijn in een CLIL-lesplanmodel. Uit de resultaten bleek dat de leerkrachten met meer dan zeven jaar aan CLIL-ervaring de CLIL-principes significant vaker integreerden in hun lessen dan de leerkrachten met één tot drie jaar aan ervaring. Dit wijst wederom op een positief effect van leerkrachterservaring op het leerproces van CLIL-leerlingen.

Tot slot heeft ook Bovellan (2014) beginnende en ervaren CLIL-leerkrachten tegenover elkaar gezet door hun overtuigingen over het leren van talen te vergelijken. Ervaren en beginnende leerkrachten bleken vooral verschillende opvattingen te hebben over het soort materiaal dat moet worden gebruikt om CLIL-leerlingen te onderwijzen in de tweede taal. De ervaren leerkrachten gebruikten meestal lesmateriaal dat afkomstig was uit de doeltaal. Beginnende leerkrachten waren echter vaker van mening dat lesmateriaal uit het reguliere lesboek afkomstig moest zijn, dat ze vervolgens ofwel gedeeltelijk naar het Engels vertaalden ofwel behandelden door Engelstalige klassikale gesprekken. Daarnaast waren de beginnende leerkrachten terughoudender om hun materiaal door collega's te laten proeflezen.

In conclusie kan er gesteld worden dat de ervaring van een leerkracht een kleine invloed heeft op de leerprestaties van zijn leerlingen. Daarnaast tonen de onderzoeken van Fernández en Halbach (2011), Custodio-Espinar (2020) en Bovellan (2014) aan dat er duidelijke methodologische verschillen zichtbaar zijn tussen de lespraktijken van ervaren en minder ervaren CLIL-leerkrachten. Er kan dus in het huidige onderzoek verwacht worden dat de ervaring van leerkrachten met lesgeven in het algemeen en lesgeven binnen het tweetalig onderwijs een effect heeft op de leerprestaties van leerlingen.

### *2.3.6 Kwantiteit en kwaliteit van de taalinput*

Een externe variabele die door de school wordt veroorzaakt, is het percentage onderwijstijd in het Engels dat wordt gehanteerd binnen een CLIL-programma. Er is niet veel onderzoek uitgevoerd dat CLIL-klassen met elkaar heeft vergeleken op basis van de kwantiteit van de Engelse taalinput die wordt aangeboden. Desalniettemin kunnen een paar onderzoeken een licht werpen op de rol van deze variabele.

Ten eerste heeft Ruiz de Zarobe (2008) drie groepen middelbare scholieren vergeleken die het Engels aan het leren waren; een groep die een regulier vreemdetalenprogramma volgde, een groep die daarnaast ook nog een CLIL-vak van 4 uur per week had gevolgd vanaf hun veertiende (CLIL1) en een groep die twee CLIL-vakken van in totaal 6 uur per week volgde (CLIL2). Hun mondelinge taalvaardigheid in het Engels werd beoordeeld op vijf categorieën: uitspraak, woordenschat, grammatica, vloeiendheid en inhoud. Op het eerste meetmoment, toen de leerlingen vijftien jaar oud waren, werd er geen verschil gevonden tussen de CLIL-groepen. Op het tweede meetmoment van een jaar later scoorde de CLIL2-groep daarentegen hoger op vier van de vijf categorieën – alleen op het onderdeel woordenschat was geen effect van de hoeveelheid blootstelling zichtbaar. In een vergelijkbaar vervolgonderzoek heeft Ruiz de Zarobe (2010) dezelfde resultaten gevonden, maar dan voor de schriftelijke vaardigheden. De CLIL2-groep scoorde op het tweede

meetmoment wederom significant hoger dan de CLIL1-groep, dit keer op drie van de vijf vaardigheden: woordenschat, grammatica en spelling. Deze onderzoeken laten zien dat er na twee jaar al een verschil zichtbaar is tussen CLIL-groepen op basis van hun hoeveelheid blootstelling aan het Engels.

Ten tweede hebben Lecocq et al. (2007a, 2007b, 2009, geciteerd in Hiligsmann et al., 2017) ook een longitudinaal onderzoek uitgevoerd naar het effect van het percentage onderwijstijd in een vreemde taal. Twee groepen van Franstalige basisschoolleerlingen werden gevolgd terwijl zij een Nederlandstalig CLIL-programma volgden. De ene groep kreeg 75% van het curriculum in de T2 aangeboden, terwijl dit percentage voor de andere groep op 50% lag. In de loop der jaren werden de kinderen getoetst op verschillende onderdelen. De CLIL-groep met hogere intensiteit behaalde voor de meeste taken betere resultaten dan de andere groep, waaronder de schrijftaken. Deze uitkomst pleit voor een CLIL-programma waarbij het aantal uren dat in de tweede taal wordt onderwezen zo hoog mogelijk ligt.

Ten derde hebben Unsworth et al. (2014) aangetoond dat een hoog percentage onderwijstijd in de vreemde taal tijdens groep 1 al invloed heeft op de leerprestaties. De resultaten van het onderzoek suggereerden echter dat de Engelse taalvaardigheid van de leerkracht een betere voorspeller is van de leerprestaties. Leerlingen die een leerkracht hadden met een B2-niveau Engels scoorden significant lager dan de groep die onderwezen werd door een moedertaalspreker. Dit toont aan dat niet alleen de kwantiteit, maar ook de kwaliteit van de T2-input een rol speelt bij de tweedetaalontwikkeling van leerlingen.

Tot slot is uit het FoTo-onderzoek gebleken dat ook de manier waarop de T2-input wordt aangeboden een effect heeft op de T2-vaardigheden van leerlingen. Tpo-scholen kunnen ingedeeld worden in drie groepen op basis van de onderwijsmethode die zij hanteren: de *One Teacher, One Language* benadering (OTOL: één leerkracht spreekt Nederlands en één leerkracht spreekt Engels), de *One Situation, One Language* benadering (OSOL: beide talen worden gescheiden per situatie aangeboden door één leerkracht) en de *Sandwich* benadering (beide talen worden gedurende de week aangeboden als een mix). Leerlingen bij wie de OTOL-benadering werd gehanteerd, behaalden gemiddeld betere resultaten dan leerlingen uit de andere twee groepen (Jenniskens et al., 2020).

Er kan geconcludeerd worden dat leerlingen die een CLIL-programma volgen met een hoger percentage onderwijstijd in het Engels ook hogere leerprestaties zullen vertonen. Daarnaast heeft het T2-niveau van de leerkracht ook een effect op de leerprestaties; leerkrachten die een moedertaalspreker zijn, zullen voor betere taalvaardigheden bij hun leerlingen zorgen dan niet-moedertaalsprekers. Ten slotte kan er verwacht worden dat leerlingen die worden onderwezen via de OTOL-benadering, beter zullen presteren dan leerlingen die worden onderwezen via de Sandwich- of OSOL-benadering.



## Hoofdstuk 3: Probleemstelling en hypothesen

De onderzoeksvraag die centraal staat in dit onderzoek luidt als volgt:

*Onderzoeksvraag 1: In hoeverre is er een effect van de ervaring van de school met het tpo-systeem op de Engelse productieve taalvaardigheden van leerlingen?*

De hypothese is dat lichting 2 betere productieve taalvaardigheden zal hebben dan lichting 1, vanwege de extra ervaring met het tpo-systeem die de school heeft kunnen opdoen.

Wel moet er rekening worden gehouden met het effect van andere variabelen om op een valide manier het verschil tussen de twee lichtingen vast te kunnen stellen. De tweede onderzoeksvraag luidt daarom als volgt:

*Onderzoeksvraag 2: In hoeverre spelen andere factoren (de moedertaal, de meertaligheid, het geslacht en de leeftijd van de leerling, de ervaring van de leerkracht, het percentage onderwijstijd in het Engels en de gehanteerde onderwijsmethode) ook een rol bij de Engelse productieve taalvaardigheden van de leerlingen?*

Uit het theoretisch kader kunnen de volgende hypothesen worden opgesteld:

- Leerlingen met het Engels als moedertaal zullen beter scoren op de productieve taalvaardigheden dan leerlingen zonder het Engels als moedertaal.
- Meertalige leerlingen zullen beter scoren op de productieve taalvaardigheden dan eentalige leerlingen.
- Vrouwelijke leerlingen zullen beter scoren op de schrijfvaardigheid dan mannelijke leerlingen.
- Oudere leerlingen zullen beter scoren op de productieve taalvaardigheden dan jongere leerlingen.
- Leerlingen die worden onderwezen door een leerkracht met meer onderriservaring zullen beter scoren op de productieve taalvaardigheden.
- Leerlingen die een CLIL-programma met een hoger percentage onderwijstijd in het Engels volgen zullen beter scoren op de productieve taalvaardigheden.
- Leerlingen die een moedertaalspreker als leerkracht hebben zullen beter scoren op de productieve taalvaardigheden dan leerlingen zonder.
- Leerlingen die worden onderwezen via de OTOL-benadering zullen beter scoren op de productieve taalvaardigheden dan leerlingen die worden onderwezen via de Sandwich- of OSOL-benadering.

Tot slot worden er ook nog interactie-effecten verwacht tussen de lichting enerzijds en de schoolgebonden factoren anderzijds. Op basis hiervan zijn de volgende hypothesen opgesteld:

- Naar mate een leerkracht meer ervaring heeft, zal het verschil in resultaten tussen de twee lichtingen groter zijn.
- Naar mate een hoger percentage onderwijstijd in het Engels wordt gehanteerd, zal het verschil in resultaten tussen de twee lichtingen groter zijn.

## Hoofdstuk 4: Methode

Om de onderzoeksvraag te beantwoorden zijn er via de *English Benchmark Young Learners* toets resultaten verzameld over de productieve taalvaardigheden van de tpo-leerlingen. Deze zijn vervolgens vergeleken tussen de eerste en tweede lichting. In het huidige hoofdstuk wordt de onderzoeksmethode verder omschreven.

### 4.1 Participanten

151 leerlingen uit de groep 7 van het schooljaar 2020/2021 (lichting 1) en 208 leerlingen uit de groep 7 van het schooljaar 2021/2022 (lichting 2) hebben deelgenomen aan het onderzoek. De leerlingen kwamen van zes verschillende tweetalige basisscholen, en hadden allemaal genoten van minimaal 20% aan Engelstalig onderwijs sinds groep 1. Lichting 1 had een gemiddelde leeftijd van 11;0 (bereik: 10;2–12;5), en lichting 2 had een gemiddelde leeftijd van 11;0 (bereik: 9;8–12;5). Het aantal leerlingen per school, verdeeld per geslacht, kan gevonden worden in Tabel 1.

Tabel 1

*Aantal participanten per school, verdeeld per geslacht en lichting*

School	Lichting 1			Lichting 2		
	Jongens	Meisjes	Totaal	Jongens	Meisjes	Totaal
School 1	8	11	19	17	11	28
School 2	11	9	20	9	10	19
School 3	11	7	18	21	30	51
School 4	16	11	27	9	18	27
School 5	29	26	55	32	28	60
School 6	6	6	12	10	13	23
Totaal	81	70	<b>151</b>	98	110	<b>208</b>

De deelnemende scholen komen allemaal uit de eerste lichting van het FoTo-onderzoek (Driessen et al., 2016), wat betekent dat deze scholen sinds 2014 zijn begonnen met het geven van tweetalig onderwijs aan groep 1. Trapsgewijs zijn er elk jaar groepen bijgekomen, waardoor in het schooljaar 2020/2021 de eerste tweetalige groepen 7 zijn ontstaan. Het schooljaar daarop is de tweede lichting tweetalige groepen 7 tot stand gekomen.

Via een vragenlijst voor de tpo-coördinatoren is informatie verzameld over het percentage onderwijstijd in het Engels en de gehanteerde onderwijsmethode in groep 7 (i.e. *One Teacher One Language, One Situation One Language* of *Sandwich*). Deze informatie is te vinden in Tabel 2. Zoals te zien is, liggen de percentages Engelse onderwijstijd tussen de 20 en 50%. Daarmee is er een grote spreiding tussen de scholen wat betreft de blootstelling die leerlingen aan het Engels krijgen. Daarnaast is er door deze vragenlijst ook informatie verkregen over de leerkrachten die voor de klas stonden in groep 7. De leeftijden van de leerkrachten liepen uiteen van 21 tot 51 jaar oud. Geen van de leerkrachten bleek een moedertaalspreker van het Engels te zijn, waardoor hypothese 7 niet getest kon worden. Hun ervaring met lesgeven binnen het tweetalig primair onderwijs en de totale ervaring die zij hebben met het onderwijzen van het Engels als een vreemde taal (*Teaching English as a Foreign Language; TEFL*) is te vinden in Bijlage A. In deze tabel is te zien dat de tpo-ervaring uiteenliep van 0 tot 7 jaar, terwijl de TEFL-ervaring varieerde van 0 tot 24 jaar.

Tot slot hebben de tpo-coördinatoren ook informatie verstrekt over de leerlingen zelf. Naast de toetsscores voor de productieve onderdelen van de English Benchmark zijn ook het geslacht, de geboortedatum en de gesproken talen van leerlingen op anonieme wijze verzameld.

Tabel 2

*Informatie over de onderwijsmethode en het percentage onderwijstijd in het Engels per school*

School	Onderwijsmethode	Onderwijstijd Engels
School 1	Sandwich	20%
School 2	One Teacher, One Language	40%
School 3	One Situation, One Language	50%
School 4	One Teacher, One Language	45%
School 5	One Situation, One Language	30%
School 6	One Situation, One Language	30% (in cohort 1) 20% (in cohort 2)

## 4.2 English Benchmark

De productieve vaardigheden van de leerlingen voor het Engels zijn gemeten met de English Benchmark Young Learners. Dit is een toets voor leerders van zes tot veertien jaar oud, die kan worden gebruikt om de taalvaardigheden van leerlingen te beoordelen op het gebied van spreken, luisteren, lezen en schrijven. De English Benchmark is beschikbaar op zes verschillende niveaus, en biedt vijf toetsen per niveau aan (Pearson, 2021). De toets wordt afgenomen op een tablet of computer, waarna de score automatisch wordt beoordeeld door middel van kunstmatige intelligentie, geavanceerde spraakherkenning en examenbeoordelingstechnologieën. Dit houdt in dat de resultaten binnen een uur beschikbaar zijn, en er automatisch gedetailleerde rapporten worden gecreëerd die aangeven op welke onderdelen leerlingen kunnen verbeteren (Manning, 2019).

De verschillende niveaus van de English Benchmark zijn gebaseerd op de Global Scale of English (GSE). Dit is een internationale Engelse taalstandaard, ontwikkeld door Pearson in samenwerking met onderzoekers en leerkrachten uit verschillende landen (Pearson, 2022b). Deze schaal bouwt voort op het Europees Referentiekader voor de Talen (ERK) door het aantal leerdoelen uit te breiden. De GSE bestaat uit een schaal die loopt van 10 tot 90 (Pearson, 2022b). De GSE is hierdoor meer geschikt om de taalvaardigheden van jonge leerders te meten dan het ERK, aangezien de scores 10-22 nog onder het eerste niveau van het ERK liggen. Daarnaast creëert het grotere bereik van de GSE de mogelijkheid om de vooruitgang van leerlingen nauwkeuriger in kaart te brengen (Moonen, 2021). In Tabel 3 is te zien hoe de scores op de GSE zich verhouden tot de ERK-niveaus en de English Benchmark-niveaus.

Het type vragen dat in de toets wordt gesteld, varieert afhankelijk van het niveau waarop een leerling wordt getoetst. Daarnaast kan een enkele vraag ook meer dan één

vaardigheid testen, zoals luisteren én spreken (Manning, 2019). Vragen bestaan bijvoorbeeld uit het samenvatten van een uitgebeeld verhaal, het hardop herhalen van een zin of het selecteren van een gesproken woord (Pearson, 2021). Op welk niveau leerlingen in het huidige experiment zijn getoetst, hangt af van het niveau waarop hun leerkracht ze heeft ingeschat. Wanneer een leerling lager scoort dan het bereik van het getoetste niveau, krijgt hij echter een Below Level als score, wat inhoudt dat zijn niveau niet vastgesteld kan worden. Ook kan een leerling niet hoger scoren dan de hoogste score van het getoetste niveau; wanneer een leerling op een te laag niveau wordt getoetst, kan er dus een plafondeffect optreden. De resultaten uit het huidige onderzoek zijn dus sterk afhankelijk van het inschatten van het juiste niveau door de leerkracht. Aangezien leerlingen maar op één niveau getoetst kunnen worden voor alle vier de taalvaardigheden, en zij soms beduidend slechter zijn in één taalvaardigheid, is het echter niet altijd mogelijk om een Below Level-score te voorkomen.

De English Benchmark verschilt op een aantal punten van reguliere taaltoetsen. Ten eerste is de English Benchmark zo ontworpen dat de toets meer als een spel voelt voor leerlingen, door middel van aansprekende animaties en interactieve opdrachten (Pearson, 2021). Het toepassen van spelelementen bij toetsing kan volgens onderzoek stress verminderen en afleidingen minimaliseren (Kayımbaşıoğlu, Oktekin & Hacı, 2016; Lieberoth, 2015; Pitoyo, Sumardi & Asib, 2019). Op deze manier worden leerlingen gemotiveerd om elke taak in de toets te voltooien.

Ten tweede krijgen leerlingen in de toets te maken met ‘internationaal Engels’. Dit betekent dat er niet alleen Brits-Engelse accenten te horen zijn, maar een mengeling van verschillende soorten accenten. Op deze manier komt de toets meer overeen met het dagelijks leven van leerlingen, waar ze ook andere accenten zullen tegenkomen (Pearson, 2022a).

Tabel 3

*Omschrijving van de verschillende niveaus van de English Benchmark Young Learners (Pearson, 2021)*

	Niveau 1	Niveau 2	Niveau 3	Niveau 4	Niveau 5	Niveau 6
<b>Aantal toetsen</b>	5	5	5	5	5	5
<b>Toetsduur</b>	16 min.	22 min.	30 min.	30 min.	36 min.	36 min.
<b>GSE</b>	10-27	16-34	22-40	27-45	33-52	42-58
<b>ERK</b>	<A1-A1	<A1-A2	A1-A2+	A1-B1	A2- B1+	A2+-B1+

### 4.3 Analyse

Dit experiment omvatte een between-subjects design. Dit houdt in dat er twee groepen participanten met elkaar zijn vergeleken op basis van de lichting waar ze zich in bevonden. De afhankelijke variabelen waren de Engelse spreek- en schrijfvaardigheid van de leerlingen, zoals aangegeven door de scores op de English Benchmark. De onafhankelijke variabele was de lichting waar de leerling in zat. Daarnaast werd de invloed van een aantal interne variabelen (geslacht, leeftijd, meertaligheid en moedertaal) en externe variabelen (tpo-ervaring en TEFL-ervaring van de leerkracht, onderwijstijd Engels en onderwijsmethode) op de afhankelijke variabelen meegenomen in het experiment.

De analyse is uitgevoerd met het programma IBM SPSS Statistics 25. Aangezien er in een regressieanalyse alleen gebruik wordt gemaakt van variabelen op schaalniveau, zijn de nominale variabelen omgezet tot dummy-variabelen. De nominale variabelen die uit drie

niveaus bestonden (i.e. de onderwijsmethode en de meertaligheid) zijn, zoals gebruikelijk is, omgezet tot twee dummy-variabelen.

Niet alle verzamelde data kon worden gebruikt in de analyse, aangezien sommige leerlingen een Below Level als score hadden gekregen. Voor het schrijfonderdeel waren dit 52 leerlingen (14,5%) en voor het spreekonderdeel waren dit 10 leerlingen (2,8%). Deze leerlingen waren ongeveer evenredig verdeeld over beide lichten. De Below Levels zijn ingevuld als *missing data*.

## Hoofdstuk 5: Resultaten

### 5.1 Correlatie tussen de twee taalvaardigheden

Om te beginnen is er gekeken in hoeverre de scores op de twee taalvaardigheden met elkaar samenhangen. Er bleek een sterke positieve correlatie te bestaan tussen de spreek- en schrijfvaardigheid ( $r = .75$ ,  $p < .001$ ). Leerlingen die beter scoorden op de spreekvaardigheid, deden dat ook op de schrijfvaardigheid, en andersom. In Tabel 4 zijn de beschrijvende statistieken voor beide taalvaardigheden weergegeven. De gemiddelde score voor de spreekvaardigheid was 43,91 punten, en de gemiddelde score voor de schrijfvaardigheid was 43,22 punten. Voor de 298 leerlingen van wie data beschikbaar was voor beide taalvaardigheden, is gekeken of er een significant verschil bestond tussen de twee scores. Dit bleek het geval te zijn; volgens een paired-samples t-test scoorden leerlingen significant hoger op de spreekvaardigheid dan op de schrijfvaardigheid ( $t(297) = 2.54$ ;  $p = .011$ ).

Tabel 4

*Beschrijvende statistieken van de scores op de English Benchmark*

	M	SD	Min	Max	N
Score voor de spreekvaardigheid	43.91	8.23	23	58	349
Score voor de schrijfvaardigheid	43.22	9.32	22	58	307

### 5.2 Voorwaardes van een meervoudige lineaire regressie

Voordat de meervoudige lineaire regressies zijn uitgevoerd, is er gekeken of de data voldeed aan de volgende voorwaardes: outliers, geschikt aantal participanten, multicollineariteit en normale verdeling. Uit de boxplots bleek dat er zich geen outliers in de data bevonden. Ook bleken er voldoende participanten te zijn; Van Voorhis en Morgan (2007) stellen voor om minstens tien deelnemers per variabele te gebruiken, wat zou resulteren in een N van 90. Hier steekt de groep van 359 participanten ver bovenuit. Aan de voorwaarde van multicollineariteit werd in de eerste instantie niet voldaan; de factor ‘Engels als thuistaal’ correleerde met de factor ‘meertaligheid’ ( $V = .93$ ,  $p < .001$ ), de factor ‘school’ correleerde met de factoren ‘onderwijsmethode’ ( $V = 1.00$ ,  $p < .001$ ) en ‘percentage Engelse onderwijstijd’ ( $V = .98$ ,  $p < .001$ ), en de factor ‘ervaring tpo’ correleerde met de factor ‘ervaring TEFL’ ( $r = .35$ ,  $p < .001$ ). Aangezien de factor ‘ervaring TEFL’ een groter bereik heeft (0 – 24 jaar) dan de factor ‘ervaring tpo’ (0 – 7 jaar), is ervoor gekozen om de eerstgenoemde mee te nemen in de analyse. De factor ‘Engels als thuistaal’ is weggelaten aangezien het overgrote deel van de Engelstalige kinderen ook meertalig was, en deze factor dus al werd meegenomen in de optie ‘meertalig met Engels’ bij de factor ‘meertaligheid’. Tot slot is ervoor gekozen om de factor ‘school’ weg te laten, aangezien deze een een-op-eenrelatie had met de factoren ‘onderwijsmethode’ en ‘percentage Engelse onderwijstijd’; iedere school hanteerde immers maar één onderwijsmethode en maar één percentage onderwijstijd in het Engels voor al zijn tpo-klassen (met uitzondering van school 6, die van percentage was veranderd). Tot slot is er gekeken of de data normaal verdeeld was. Dit bleek voor beide taalvaardigheden niet het geval te zijn. Volgens Hayes (2018) is deze voorwaarde echter een van de minst belangrijke bij het uitvoeren van een lineaire regressie. Onderzoek suggereert namelijk dat alleen de meest ernstige schendingen van de normaliteitsaannname een

significant effect hebben op de validiteit van een regressieanalyse, tenzij de participantengroep relatief klein is. Er is daarom alsnog ervoor gekozen om gebruik te maken van een regressieanalyse.

### 5.3 De spreekvaardigheid

Om te kijken in hoeverre de variantie van de scores op de spreekvaardigheid verklaard kon worden door de variabelen, is een meervoudige lineaire regressie uitgevoerd. De resultaten toonden aan dat het model 18,6% van de variantie verklaarde en dat er een significante lineaire relatie was tussen de spreekvaardigheid en tenminste één van de onafhankelijke variabelen ( $F(9,339) = 9.82, p < .001$ ). Tabel 5 geeft de significanties van de variabelen in het model weer.

Tabel 5

*Meervoudige lineaire regressie voor de spreekvaardigheid*

Variabele	B	SD B	$\beta$	t	p
(Constant)	55.61	11.41		4.87	.000**
Leeftijd	-0.01	<0.01	-.09	-1.72	.087
Onderwijsmethode OTOL	2.77	1.19	.15	2.34	.020*
Onderwijsmethode Sandwich	0.69	1.63	.03	0.42	.672
Meertalig zonder Engels	2.91	0.99	.15	2.96	.003**
Meertalig met Engels	7.62	1.24	.31	6.13	.000**
Onderwijstijd Engels	0.09	0.05	.12	1.78	.076
Ervaring TEFL	0.20	0.07	.19	2.97	.003**
Geslacht	-0.38	0.80	-.02	-0.48	.634
Lichting	-1.07	0.82	-.06	-1.30	.194

Note. \*\*  $p < .01$  tweezijdig. \*  $p < .05$  tweezijdig.

Zoals in de tabel te zien is, bleken er drie variabelen significante voorspellers van de scores op de spreekvaardigheid te zijn. Dit waren de TEFL-ervaring van de leerkracht, de gehanteerde onderwijsmethode en de meertaligheid van een leerling. Zowel de groep meertalige leerlingen met het Engels als één van hun moedertalen ( $B = 7.62, p < .001$ ) als de groep meertalige leerlingen zonder het Engels als moedertaal ( $B = 2.91, p = .003$ ) hebben een positieve B-coëfficiënt, wat inhoudt dat deze groepen beide hoger scores op de English Benchmark dan eentalige leerlingen. Meertalige leerlingen met het Engels als één van hun moedertalen scoren gemiddeld het hoogst; hun gemiddelde score ligt 6,7 punten hoger dan die van eentalige leerlingen. In Tabel 6 zijn de gemiddelde scores per meertalige conditie weergegeven.

Tabel 6

*Scores op de spreekvaardigheid, uitgesplitst per meertalige conditie*

Meertaligheid	M	N	SD
Meertalig met Engels	49.09	45	7.33
Meertalig zonder Engels	44.98	86	9.36
Eentalig	42.41	218	7.42

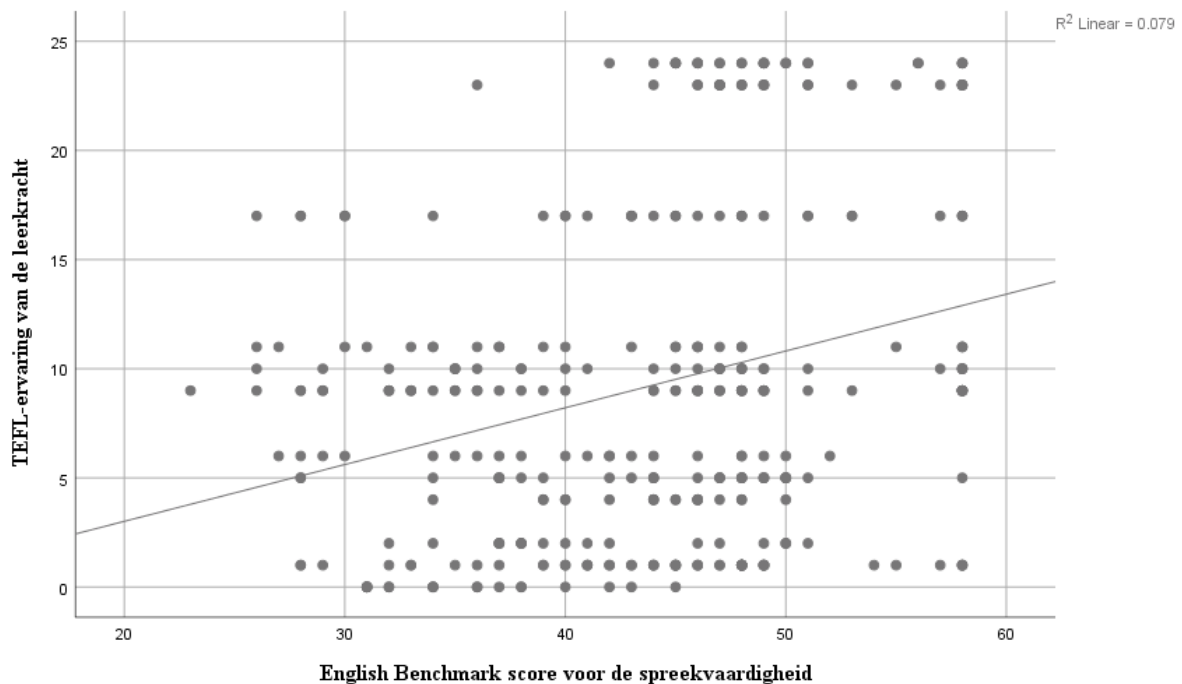
Ook de gehanteerde onderwijsmethode had een significante invloed op de scores van de leerlingen. De *One Teacher, One Language* benadering bleek voor significant hogere resultaten te zorgen dan de *One Situation, One Language* benadering ( $B = 2.77, p = .020$ ). Er was echter geen significant verschil tussen de *One Situation, One Language* benadering en de *Sandwich* benadering ( $B = .69, p = .672$ ). De gemiddelde scores per gehanteerde onderwijsmethode zijn in Tabel 7 te vinden. Er zat gemiddeld een verschil van 6,0 punten tussen leerlingen die werden onderwezen met de *One Teacher, One Language* benadering en de *Sandwich* benadering.

Tabel 7

*Scores op de spreekvaardigheid, uitgesplitst per gehanteerde onderwijsmethode*

Onderwijsmethode	M	N	SD
One Teacher, One Language	47.12	90	6.39
One Situation, One Language	43.17	212	8.95
Sandwich	41.09	47	5.88

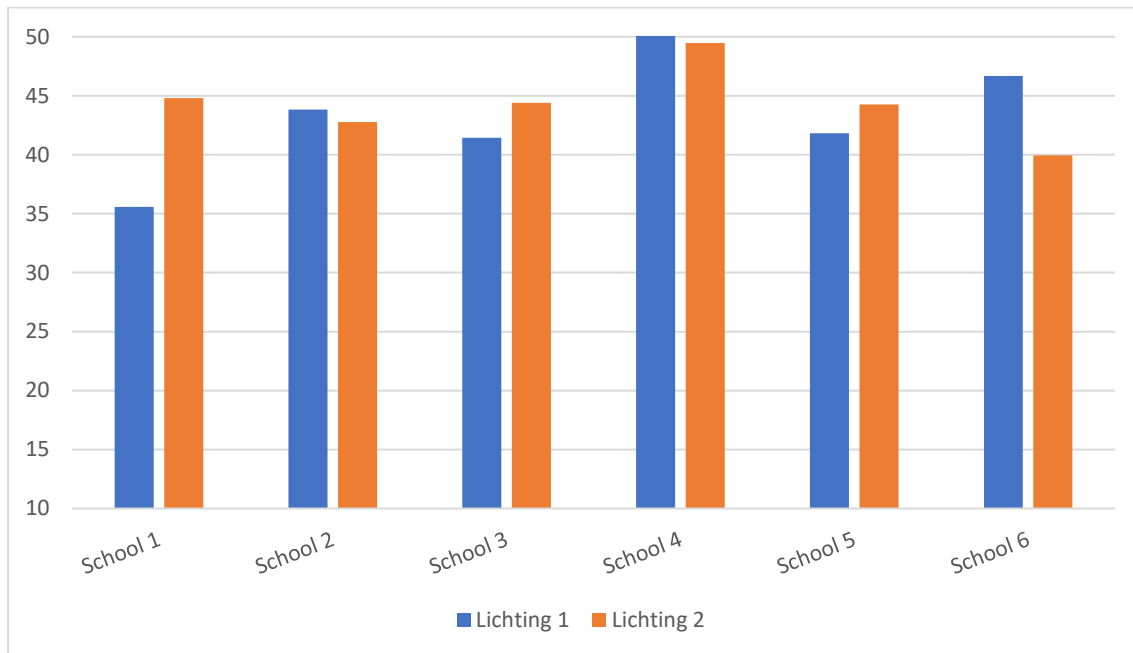
Tot slot beïnvloedde de ervaring van een leerkracht met het onderwijzen in het Engels de scores op de spreekvaardigheid significant ( $B = .20, p = .003$ ). De positieve B-coëfficiënt geeft aan dat elk jaar aan ervaring van de leerkracht resulteerde in een verhoging van de score met 0,2 punten. In Figuur 2 wordt het verband tussen de TEFL-ervaring en de scores visueel weergegeven.



*Figuur 2.* Het verband tussen de TEFL-ervaring van de leerkracht (in jaren) en de score voor de spreekvaardigheid.



Vervolgens is er voor elke school individueel gekeken naar de resultaten van de twee lichten. In Figuur 3 wordt dit visueel weergegeven. Alleen voor school 1, 3 en 5 waren de resultaten hoger voor lichting 2. Met een independent samples t-test is daaropvolgend per school berekend of het verschil in resultaten tussen lichting 1 en 2 significant was. Dit bleek alleen het geval te zijn voor school 1 ( $t(45) = 8.36; p < .001$ ) en school 6 ( $t(30.91) = -3.33; p = .002$ ). Waar voor school 1 lichting 2 significant beter presteerde, bleek dit voor school 6 juist lichting 1 te zijn.

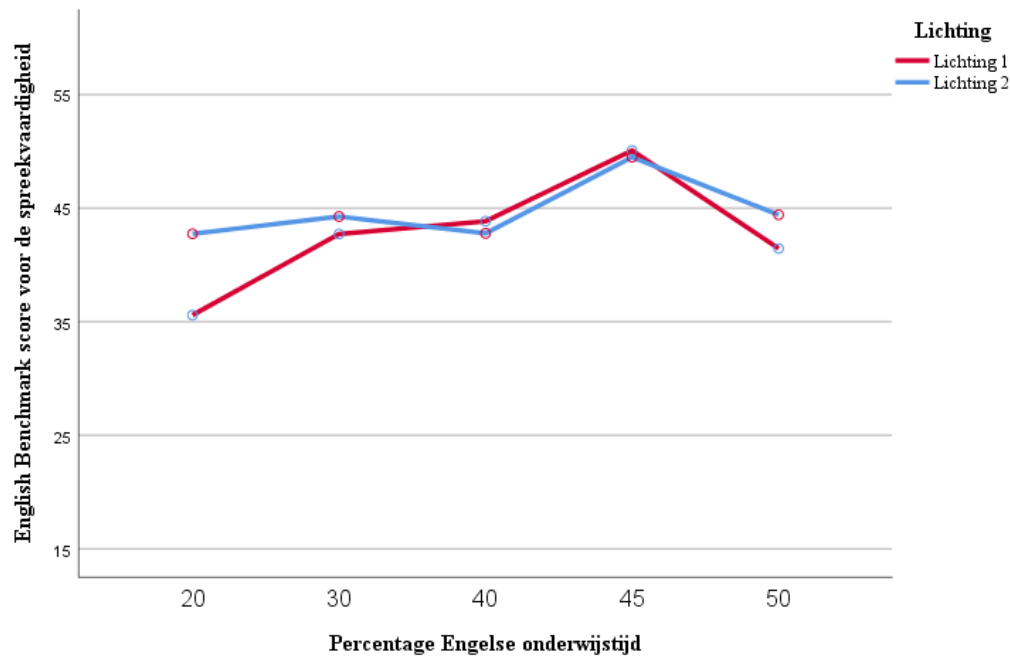


*Figuur 3.* Visuele weergave van de scores op de spreekvaardigheid per lichting, uitgesplitst per tpo-school.

Om een mogelijke verklaring te vinden voor dit verschil is opnieuw een meervoudige lineaire regressie uitgevoerd, dit keer voor school 1 en school 6 apart. Niet elke variabele kon echter worden meegenomen in de analyses. Aangezien de gehanteerde onderwijsmethode en het percentage onderwijstijd in het Engels gelijk waren tussen beide lichteningen voor school 1, zijn deze variabelen niet meegenomen. Het percentage onderwijstijd in het Engels verschilde wel tussen de lichteningen voor school 6 (zie Tabel 2). Deze variabele kon echter alsnog niet worden meegenomen. Het verschil in de TEFL-ervaring van de leerkracht tussen de lichteningen was namelijk één jaar op beide scholen, waardoor deze variabele een een-op-eenrelatie had met de variabele ‘onderwijstijd Engels’ in het geval van school 6. Dit gold ook voor de variabele ‘lichting’ op beide scholen. De regressieanalyse is dus alleen uitgevoerd met de variabelen ‘TEFL-ervaring van de leerkracht’, ‘meertaligheid’, ‘geslacht’ en ‘leeftijd’. Uit de resultaten bleek dat de ervaring van de leerkracht de enige significante voorspeller van de scores was voor zowel school 1 ( $B = 8.58, p < .001$ ) als school 6 ( $B = -7.06, p = .010$ ). Dit resultaat moet echter voorzichtig geïnterpreteerd worden, aangezien er zoals net gezegd geen onderscheid gemaakt kan worden tussen deze variabele en de variabele ‘lichting’ en, in het geval van school 6, de variabele ‘onderwijstijd Engels’.

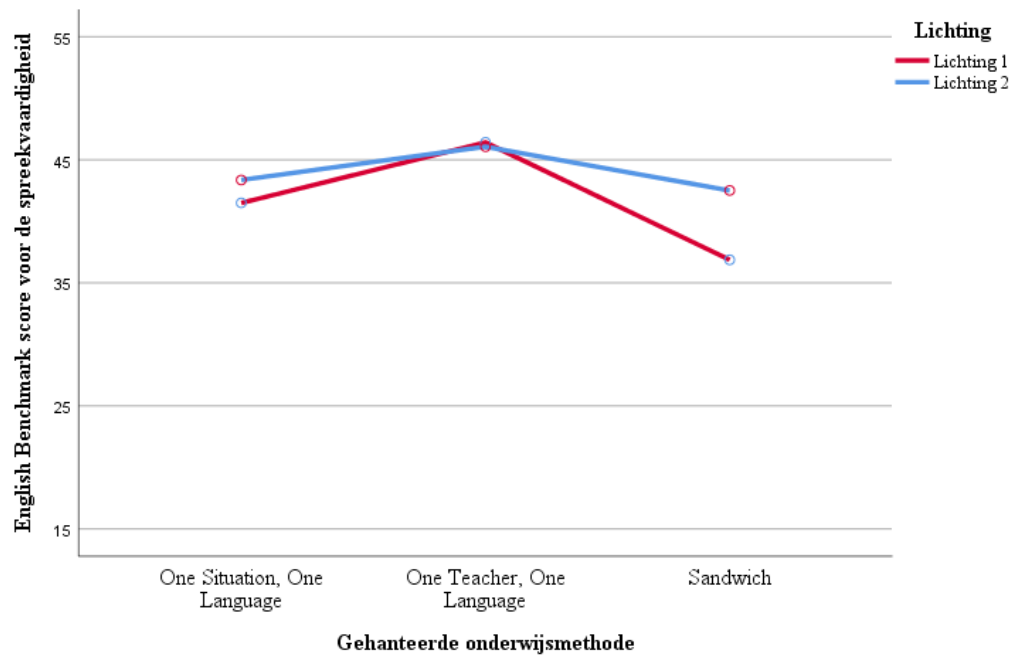
Tot slot zijn de hypothesen getoetst die stellen dat er een interactie-effect bestaat tussen de lichting enerzijds en de TEFL-ervaring van de leerkracht en het percentage onderwijstijd in het Engels anderzijds. Daarnaast is er ook nog gekeken of er een interactie-effect bestond tussen de lichting en de gehanteerde onderwijsmethode. Uit een more-way

ANOVA bleek ten eerste dat er een interactie-effect is tussen de lichteing en het percentage onderwijstijd in het Engels ( $F(4, 339) = 2.35; p = .050$ ). In Figuur 4 is dit effect zichtbaar. Tegen verwachtingen in werd het verschil tussen de twee lichteingen niet groter naarmate het percentage onderwijstijd toenam; deze was juist het grootst bij een onderwijspercentage van 20%.



*Figuur 4.* Het interactie-effect tussen de lichteing en het percentage Engelse onderwijstijd op de spreekvaardigheid

Ten tweede bleek uit een more-way ANOVA dat er een interactie-effect bestaat tussen de lichteing en de gehanteerde onderwijsmethode ( $F(2, 343) = 6.35; p = .002$ ). Figuur 5 laat dit effect zien. Uit de grafiek wordt duidelijk dat het voor de *One Teacher, One Language* en *One Situation, One Language* methodes niet uitmaakt in welke lichteing een leerling zit. Voor de Sandwich methode is de lichteing echter wel van belang; leerlingen uit lichteing 2 scoorden gemiddeld 9 punten hoger. Ten derde bleek er geen interactie-effect te zijn tussen de lichteing en de TEFL-ervaring van de leerkracht ( $F(1, 335) = 3.56; p = .060$ ).



Figuur 5. Het interactie-effect tussen de lichting en de gehanteerde onderwijsmethode op de spreekvaardigheid

Deze resultaten suggereren dat er een interactie-effect bestaat tussen de lichting enerzijds en de gehanteerde onderwijsmethode en het percentage Engelse onderwijstijd anderzijds. Als we dieper ingaan op de resultaten, lijkt dit echter niet het geval te zijn. In beide gevallen wordt het verschil namelijk alleen veroorzaakt door school 1, die als enige gebruikmaakt van de *Sandwich* benadering en een percentage van 20%. Een verklaring voor het verschil tussen de twee lichtingen dat wordt gevonden bij school 1 wordt besproken in de discussie.

#### 5.4 De schrijfvaardigheid

Vervolgens zijn dezelfde analyses uitgevoerd voor de schrijfvaardigheid. Ten eerste is er een meervoudige lineaire regressie gebruikt om te zien welke variabelen significante voorspellers waren van de scores op de schrijfvaardigheid. Het model bleek significant te zijn en 14,7% van de variantie in de scores te verklaren ( $F(9,297) = 6.88, p < .001$ ). In Tabel 8 zijn de significanties van de variabelen in het model weergegeven.

Tabel 8

*Meervoudige lineaire regressie voor de schrijfvaardigheid*

Variabele	B	SD B	$\beta$	t	p
(Constant)	66.40	14.38		4.62	.000**
Leeftijd	-0.01	<0.01	-.11	-1.99	.047*
Onderwijsmethode OTOL	2.89	1.59	.13	1.82	.070
Onderwijsmethode Sandwich	-1.35	2.13	-.04	-0.63	.527
Meertalig zonder Engels	2.89	1.24	.13	2.34	.020*
Meertalig met Engels	8.41	1.47	.32	5.72	.000**
Onderwijstijd Engels	0.01	0.06	.01	0.12	.902

Ervaring TEFL	0.18	0.09	.14	2.00	.046*
Geslacht	0.34	1.00	.02	0.34	.732
Lichting	-1.11	1.03	-.06	-1.08	.280

Note. \*\*  $p < .01$  tweezijdig. \*  $p < .05$  tweezijdig.

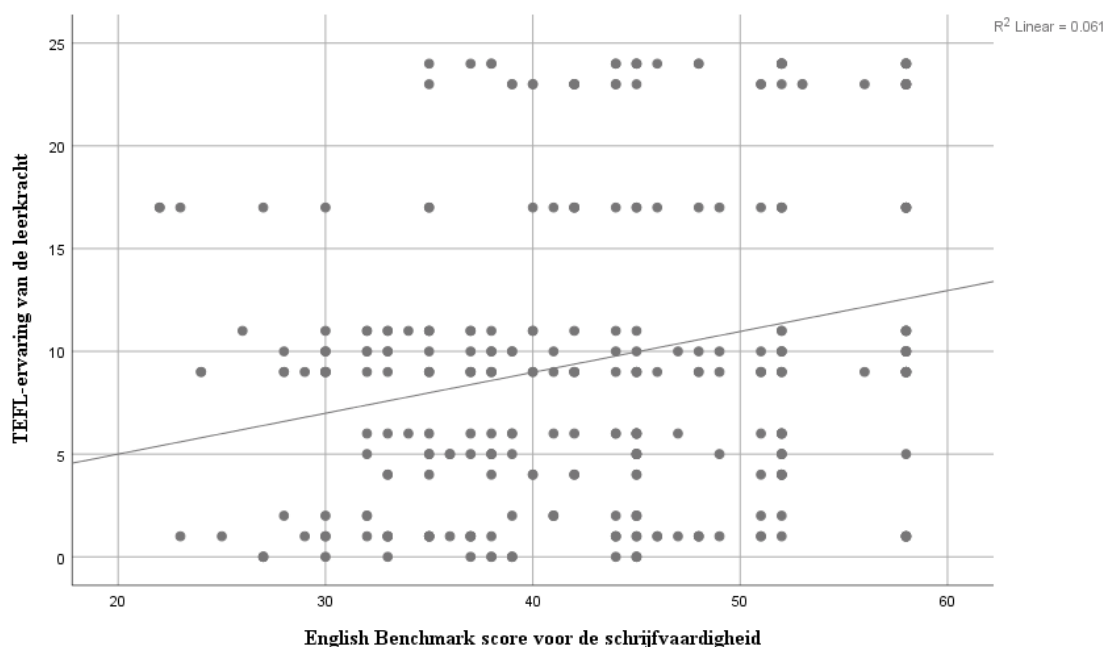
Wederom bleken niet alle variabelen de leerlingprestaties significant te voorspellen. Alleen de TEFL-ervaring van de leerkracht en de meertaligheid en leeftijd van de leerling beïnvloedden de resultaten significant. Voor de variabele ‘meertaligheid’ is wederom te zien dat zowel meertalige leerlingen met ( $B = 8.41$ ,  $p < .001$ ) als zonder ( $B = 2.89$ ,  $p = .020$ ) het Engels als één van hun moedertalen beter scoorden dan de eentalige leerlingen. Het verschil tussen de gemiddelde scores is te vinden in Tabel 9. De groep meertalige leerlingen met het Engels als moedertaal scoorde gemiddeld 7,7 punten hoger dan de groep eentalige leerlingen.

Tabel 9

*Scores op de schrijfvaardigheid, uitgesplitst per meertalige conditie*

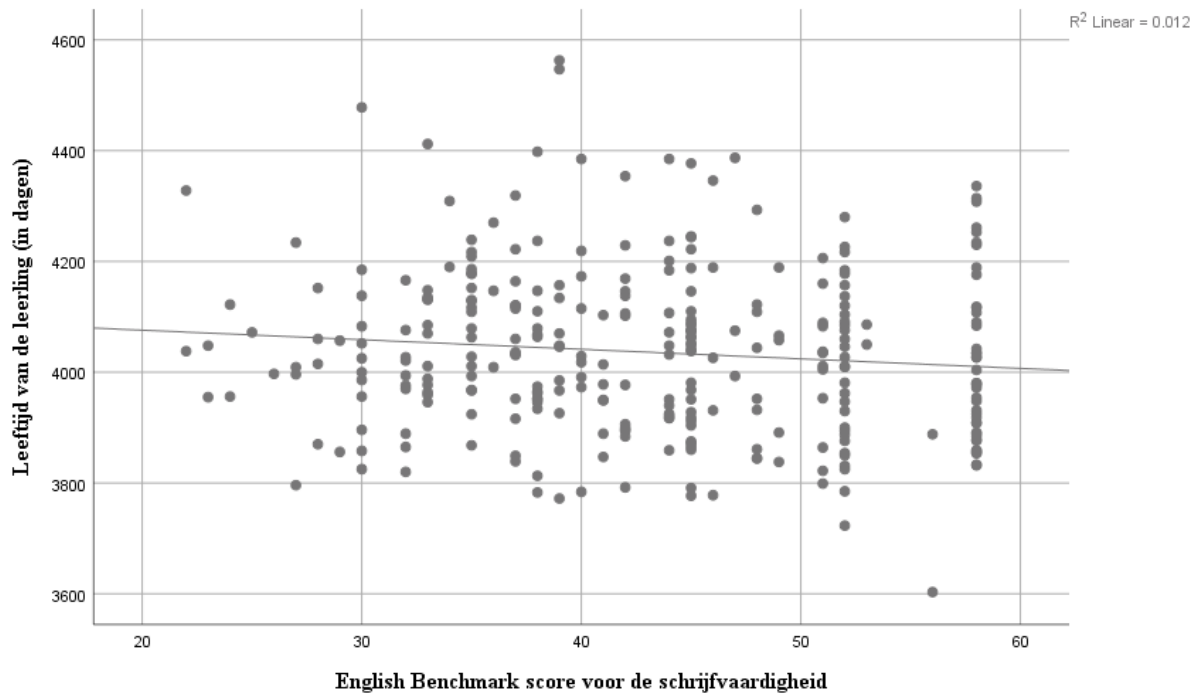
Meertaligheid	M	N	SD
Meertalig met Engels	49.18	44	8.68
Meertalig zonder Engels	44.11	74	10.51
Eentalig	41.48	189	8.33

Ook de variabele ‘TEFL-ervaring van de leerkracht’ bleek opnieuw een significante voorspeller van de scores te zijn ( $B = .18$ ,  $p = .046$ ). Hoe meer jaren aan ervaring een leerkracht had met het lesgeven in het Engels, hoe hoger de scores van zijn leerlingen op de schrijfvaardigheid waren. Dit wordt op een visuele manier weergegeven in Figuur 6.



Figuur 6. Het verband tussen de TEFL-ervaring van de leerkracht (in jaren) en de score voor de schrijfvaardigheid.

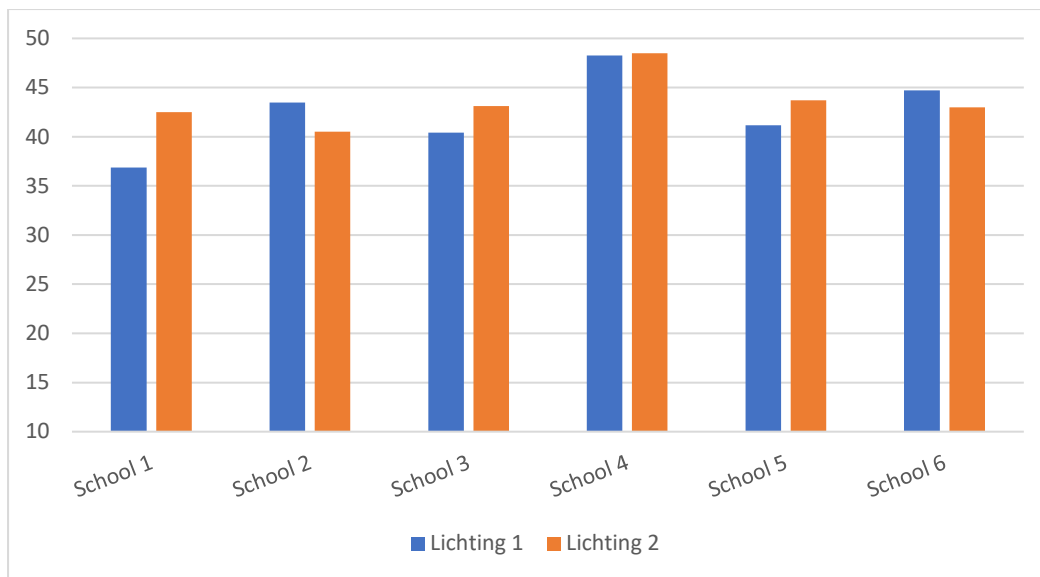
Tot slot bleek de leeftijd van de leerling de scores op de schrijfvaardigheid significant te beïnvloeden ( $B = -.01$ ,  $p = .047$ ). In Figuur 7 wordt het verband tussen de leeftijd en de scores visueel weergegeven. Tegen de hypothese in is dit verband negatief; hoe jonger een leerling is, hoe beter hij scoort op de English Benchmark. Aan de B-coëfficiënt is echter te zien dat dit effect maar zeer klein is.



*Figuur 7.* Het verband tussen de leeftijd van de leerling (in dagen) en de score voor de schrijfvaardigheid.

Vervolgens is opnieuw per school het verschil in resultaten tussen de twee lichtingen geanalyseerd. Figuur 8 toont een visuele weergave hiervan. Alleen de resultaten van school 1, 3, 4 en 5 waren hoger voor lichting 2. Een independent samples t-test toonde aan dat het verschil alleen voor school 1 significant was ( $t(29) = 2.29$ ;  $p = .030$ ); lichting 2 presteerde significant beter dan lichting 1. Daaropvolgend is een meervoudige lineaire regressie uitgevoerd met alleen de data van school 1 om te zien in hoeverre dit verschil verklaard kon worden door de onafhankelijke variabelen. Dit model bleek echter niet significant te zijn ( $F(5,25) = 2.52$ ,  $p = .057$ ). Deze resultaten kunnen dus niet geïnterpreteerd worden.

Tot slot is er nog een keer gekeken naar mogelijke interactie-effecten tussen de lichting waar de leerling zich in bevond en de variabelen ‘gehanteerde onderwijsmethode’, ‘TEFL-ervaring van de leerkracht’ en ‘percentage Engelse onderwijstijd’. Een more-way ANOVA toonde aan dat er geen interactie-effect was voor zowel de onderwijsmethode ( $F(2, 301) = 1.18$ ;  $p = .309$ ), TEFL-ervaring van de leerkracht ( $F(1, 293) = 3.24$ ;  $p = .073$ ) en het percentage onderwijstijd in het Engels ( $F(4, 297) = 1.09$ ;  $p = .362$ ).



*Figuur 8.* Visuele weergave van de scores op de schrijfvaardigheid per lichting, uitgesplitst per tpo-school.

## Hoofdstuk 6: Discussie

Dit experiment is opgezet om twee onderzoeksvragen te beantwoorden. Ten eerste is er onderzocht of er een effect is van de ervaring van basisscholen met het tweetalige onderwijssysteem op de productieve Engelse taalvaardigheden van leerlingen. Ten tweede is er gekeken in hoeverre andere school- en kindgebonden variabelen een rol spelen bij de productieve taalvaardigheden. Beide onderzoeksvragen konden worden beantwoord met de bevindingen van het experiment.

Allereerst bleek er geen effect te zijn van de schoolervaring op de leerprestaties van de leerlingen; er zat geen significant verschil tussen lichting 1 en lichting 2. De eerste onderzoeksvraag moet dus negatief beantwoord worden. Wanneer er dieper op de resultaten werd ingegaan door per school de twee lichtingen te vergelijken, werd er voor school 1 (voor spreken en schrijven) en school 6 (voor spreken) wél een significant verschil gevonden. Deze verschillen lijken echter verklaard te kunnen worden door veranderingen in het percentage onderwijstijd en de leerkrachtervaring. School 6, waar lichting 1 beter scoorde dan lichting 2, was van 30% aan onderwijstijd in het Engels afgezaakt naar 20%. Alhoewel het percentage onderwijstijd volgens de meervoudige lineaire regressie geen significante voorspeller was van de variantie in de scores, lijkt het aannemelijk dat de daling van de scores hierdoor verklaard kan worden. Er zaten verder namelijk geen grote verschillen tussen de twee lichtingen wat betreft de verhouding van jongens en meisjes, de gemiddelde leeftijd en de leerkrachtervaring. Wel zat er een verschil tussen de hoeveelheid leerlingen met het Engels als moedertaal, maar deze lag juist in het voordeel van lichting 2. Aangezien het onlogisch is dat een jaar extra aan schoolervaring met het tpo-systeem voor een verlaging van de scores zou zorgen, kan er alleen naar de verandering in het percentage Engelse onderwijstijd worden gewezen als verklaring. Voor school 1, waar lichting 2 beter scoorde dan lichting 1, is het wel een mogelijkheid dat het verschil werd veroorzaakt door het jaar extra schoolervaring. Het is echter opvallend dat op deze school werd lesgegeven door een leraar in opleiding die begon met 0 jaar aan ervaring. Onderzoek heeft aangetoond dat het eerste jaar van lesgeven een belangrijk onderdeel is van de loopbaan van een leraar (Hebert & Worthy, 2001). De overgang van de lerarenopleiding naar een echt klaslokaal wordt in de literatuur gekarakteriseerd als een realiteitsschok waarbij beginnende leraren beseffen dat de kennis die ze tijdens hun opleiding hebben opgedaan niet altijd overeenkomt met de realiteit waarmee ze worden geconfronteerd tijdens het lesgeven (Farrell, 2006). Het lijkt dus aannemelijk dat de hogere scores in lichting 2 worden veroorzaakt door de ervaring die de leerkracht heeft opgedaan tijdens zijn eerste jaar lesgeven, en niet door de algemene ervaring van de school. Ook wanneer er naar de individuele scholen wordt gekeken, lijkt er in conclusie geen effect van schoolervaring zichtbaar te zijn.

Vervolgens kan de tweede onderzoeksvraag beantwoord worden. Uit de resultaten bleek dat er drie variabelen invloed hadden op de spreekvaardigheid van leerlingen: de TEFL-ervaring van de leerkracht, de meertaligheid van een leerling en de gehanteerde onderwijsmethode. Deze variabelen gedroegen zich geheel zoals de verwachtingen. Ten eerste zorgde een hogere ervaring van de leerkracht met lesgeven in het Engels voor lichtelijk betere scores. Dit komt overeen met de literatuur, die stelt dat de ervaring van een leerkracht een kleine invloed heeft op de leerprestaties van zijn leerlingen (Rivkin, Hanushek & Kain, 2005; Rockoff, 2004; Stigler & Miller, 2018). Ten tweede scoorden meertalige leerlingen beter dan eentalige leerlingen, wat de hypothese bevestigt dat meertaligen een nieuwe taal

sneller oppikken dan eentaligen (Abu-Rabia & Sanitsky, 2010; Fleckenstein, Möller & Baumert, 2018; Hopp et al., 2019; Maluch, Kempert, Neumann & Stanat, 2015). Wel spreekt dit de resultaten van het onderzoek van Mady (2014) tegen, waarin werd gevonden dat meertalige leerlingen geen hogere spreekvaardigheid hadden dan eentalige leerlingen. Ook scoorden leerlingen die meertalig waren én het Engels als moedertaal hadden zoals verwacht nog hoger (Chondrogianni & Marinis, 2011; Scheele, Leseman & Mayo, 2009). Dit laat zien dat niet alleen de uitspraak belangrijk is voor de spreekvaardigheid, die volgens onderzoek van Abrahamsson en Hyltenstam (2009) en Alario et al. (2010) niet veel verschilt tussen moedertaalsprekers en succesief tweetaligen. Ten derde bleek er net zoals in het FoTo-onderzoek een verschil te zitten tussen de effectiviteit van de gehanteerde onderwijsmethodes. De *One Teacher, One Language* benadering zorgde voor significant betere resultaten dan de andere twee benaderingen, wat overeenkomt met bevindingen uit de derde meting van het FoTo-onderzoek (Jenniskens et al., 2020). In het FoTo-onderzoek kon dit verschil worden verklaard doordat de Engelse taalvaardigheid van de OTOL-leerkrachten hoger bleek te liggen dan die van de Sandwich- en OSOL-leerkrachten. Aangezien in het huidige onderzoek de taalvaardigheden van de leerkrachten niet zijn vastgesteld, is het echter niet mogelijk om te zeggen of dit hier eveneens het geval is.

Voor de schrijfvaardigheid bleken de TEFL-ervaring en de meertaligheid ook significante voorspellers te zijn, maar was de onderwijsmethode dat niet. Een mogelijke verklaring hiervoor is dat de gehanteerde onderwijsmethode meer invloed heeft op de spreekvaardigheid van een leerling dan de schrijfvaardigheid, omdat er tijdens de lessen meer wordt gesproken dan geschreven in het Engels. Een verschil in de manier waarop het Engels wordt aangeboden, is in dat geval meer van invloed op het spreken. Tot slot bleek de leeftijd van de leerling ook een significant effect te hebben op de schrijfvaardigheid. Tegen verwachtingen in scoorden jongere leerlingen iets beter dan oudere leerlingen. De effectgrootte was echter zeer klein ( $B = -0.01$ ). Een mogelijke verklaring die niet aannemelijk bleek te zijn is dat dit effect zou worden veroorzaakt door een vertekening van de data door het grote aantal Below Levels. Het zou namelijk mogelijk zijn dat vooral jongere leerlingen een Below Level hebben gehaald, en dit de slechtst scorende leerlingen uit de jongere groep heeft geëlimineerd. Wanneer het aantal Below Levels per leeftijdsgroep wordt geteld, lijkt het echter niet zo te zijn dat deze zich vooral bij de jongere leerlingen voordoen. Er kan dus geen verklaring voor dit onverwachte resultaat worden gegeven.

De hypotheses over de andere getoetste variabelen zijn niet bevestigd. Om te beginnen bleken vrouwen niet beter te scoren op de schrijfvaardigheid dan mannen. Dit gaat in tegen de resultaten van Roquet, Llopis en Pérez-Vidal (2016) en Lasagabaster (2008), die beide een effect van geslacht op de schrijfvaardigheid van CLIL-leerlingen vonden. Het bevestigt daarentegen de hypothese van Lahuerta (2020), die stelt dat CLIL-instructie de 'genderkloof' dicht die normaal gezien in de tweedetaalverwerving wordt gevonden. Of dit resultaat ook echt wordt veroorzaakt door een verhoogde motivatie bij jongens, zoals de hypothese stelt, kan echter niet gezegd worden. Om deze vraag te beantwoorden, zou vervolgonderzoek de motivatie van leerlingen mee moeten nemen als variabele. Ten tweede bleek een hoger percentage onderwijstijd in het Engels niet voor significant betere resultaten te zorgen. Dit is een frappant resultaat, aangezien onderzoek keer op keer heeft aangetoond dat de hoeveelheid blootstelling aan het Engels een directe invloed heeft op de taalvaardigheden (Lecocq et al., 2009; Ruiz de Zarobe, 2010; Unsworth et al., 2014). Er kunnen twee mogelijke verklaringen worden gegeven. Ten eerste is niet alleen de kwantiteit



maar ook de kwaliteit van de taalinput belangrijk. Het niveau Engels van de leerkracht speelt dus ook een rol bij de tweedetaalontwikkeling van leerlingen. Aangezien er in het huidige onderzoek niet gevraagd is naar het ERK-niveau van de leerkrachten, zou het zo kunnen zijn dat deze lager ligt op de scholen met een hoger percentage Engelse onderwijstijd. In dat geval zou het effect van de lagere kwaliteit taalinput het effect van de grotere hoeveelheid taalinput kunnen tegengaan. Een tweede verklaring is dat in het huidige onderzoek er te veel variabelen met elkaar samenhangen om het effect van het percentage onderwijstijd te kunnen meten. Elk percentage kwam namelijk maar op één school voor, waardoor deze variabele correleerde met de gehanteerde onderwijsmethode en de ervaring van de leerkracht. Voor de *Sandwich* benadering was het zelfs het geval dat deze maar op één school werd gehanteerd, waardoor er geen onderscheid gemaakt kan worden tussen het effect van deze benadering en het effect van het percentage onderwijstijd op deze school, wat 20% was. Voor een vervolgonderzoek zou het dus interessant zijn om naar meerdere scholen met hetzelfde percentage Engelse onderwijstijd te kijken, zodat alle schoolgebonden variabelen van elkaar gescheiden kunnen worden.

Tot slot is er nog gekeken naar mogelijke interactie-effecten tussen de lichte en andere schoolgebonden factoren. Deze werden niet gevonden voor de schrijfvaardigheid, maar wel voor de spreekvaardigheid. Ten eerste bleek er een interactie-effect te zijn tussen de onderwijsmethode en de lichte. In Figuur 3 is duidelijk zichtbaar dat dit door de *Sandwich* benadering wordt veroorzaakt; tussen lichte 1 en 2 zat gemiddeld een verschil van 9 punten. In werkelijkheid is dit echter precies hetzelfde effect als dat voor school 1 werd gevonden, aangezien deze school de enige is die de *Sandwich* benadering hanteerde. Er kan dus wederom gesteld worden dat dit verschil in lichten hoogstwaarschijnlijk te verklaren valt door de ervaring die de leerkracht in opleiding heeft opgedaan tijdens zijn eerste jaar lesgeven. Ten tweede bleek er tussen het percentage onderwijstijd en de lichte een interactie-effect te zijn. Ook dit effect valt echter te verklaren door de samenhangende variabelen; het verschil is vooral zichtbaar voor de lichten met 20% Engelse onderwijstijd, wat wederom alleen school 1 is. Allebei de gevonden interactie-effecten lijken dus in werkelijkheid door een andere factor te worden veroorzaakt.

Dit onderzoek heeft een aantal limitaties. Veel van de validiteit van het huidige onderzoek hangt ten eerste af van de English Benchmark toets waarmee de productieve taalvaardigheden van de leerlingen zijn gemeten. Deze toets heeft echter een aantal beperkingen. Zo waren er voor de spreekvaardigheid 32 leerlingen en voor de schrijfvaardigheid 42 leerlingen die de hoogst mogelijke score hebben gehaald. Door dit plafondefect is het niet mogelijk om van deze leerlingen de werkelijke taalvaardigheid vast te stellen. Andersom kan ook van de te laag scorende leerlingen de echte taalvaardigheid niet bepaald worden. Wanneer een leraar niet het juiste niveau voor iedere leerling uitzoekt, of een leerling een stuk slechter is in één van de vier taalvaardigheden, krijgt de leerling immers een Below Level. Uit deze score kan verder weinig geconcludeerd worden, waardoor de English Benchmark op dit punt inferieur is ten opzichte van reguliere toetsen die uit één niveau bestaan. Tot slot zorgt de automatische beoordeling van de spreekvaardigheid voor obstakels waar een beoordelende leraar geen last van zou hebben. Zo kan een leerling een No Score krijgen wanneer het de spraakherkenning niet lukt om de spraak vast te stellen. Het is voor een leraar echter niet duidelijk of de No Score is veroorzaakt doordat de leerling te zacht praatte of helemaal niks heeft gezegd, er te veel achtergrondruis was of de microfoon niet goed werkte. Aan deze nadelen hangen echter ook voordelen vast. Het feit dat de

sprekvaardigheid door automatische spraakherkenning wordt vastgesteld en beoordeeld, zorgt ervoor dat leerlingen sneller en objectiever kunnen worden getoetst. Ook zorgt het gebruik van niveaus binnen de toets ervoor dat, wanneer leerlingen wél op het juiste niveau worden getoetst, hun taalvaardigheid nauwkeuriger kan worden vastgesteld. Voor het tweetalig primair onderwijs, waar het niveau Engels tussen leerlingen enorm uiteenloopt, is dit een groot voordeel. Het gebruik van de English Benchmark om de taalvaardigheden van leerlingen vast te stellen kent dus zowel voor- als nadelen.

Een andere limitatie van het huidige onderzoek was dat er niet naar elke relevante variabele kon worden gekeken. Het lag buiten de omvang van dit onderzoek om de taalaanleg, intelligentie, leerstijl, motivatie en sociaal-economische status van de leerlingen te meten. Dit zijn echter allemaal belangrijke factoren die de tweedetaalontwikkeling kunnen beïnvloeden (Arabski & Wojtaszek, 2011; Cook & Singleton, 2014; Trebits et al., 2022; Wen, Biedroń & Skehan, 2017). De verwachting is dus dat de verklaarde variantie in de scores hoger dan de huidige 18,6% voor de spreekvaardigheid en 14,7% voor de schrijfvaardigheid zou liggen wanneer deze factoren waren meegenomen. Ook buitenschoolse omgang met de taal, zoals gamen en het gebruik van social media, zouden volgens de literatuur een deel van de variantie kunnen verklaren (De Wilde, Brysbaert & Eyckmans, 2020; Jensen, 2017; Sundqvist, 2009). Tot slot is het ERK-niveau Engels van de leerkrachten niet meegenomen in het onderzoek. Dit komt doordat uit het FoTo-onderzoek is gebleken dat de zelfinschatting en het werkelijke niveau niet altijd overeenkomen (Jenniskens et al., 2020). Het leek dus maar weinig zinvol om op basis van zelfinschattingen deze factor mee te nemen in de analyse. Voor een vervolgonderzoek zou het echter interessant zijn om het ERK-niveau van de leerkrachten vast te stellen door middel van interviews, net zoals in het FoTo-onderzoek is gedaan.

De resultaten van het huidige onderzoek hebben een aantal implicaties. Ten eerste is er gebleken dat één jaar aan ervaring met het tweetalige onderwijssysteem niet genoeg is om significant hogere leerprestaties teweeg te brengen bij leerlingen. Dit betekent echter niet dat de schoolervaring helemaal geen rol speelt bij de leerprestaties. Het is goed mogelijk dat er wel een verschil zal worden gevonden wanneer de eerste lichting zal worden vergeleken met bijvoorbeeld de vijfde lichting. Het spreekt immers voor zich dat er door de jaren heen meer aanpassingen aan het tweetalig onderwijs zullen worden gedaan. Aangezien er op het moment van het onderzoek echter nog maar twee lichtingen van groepen 7 bestaan, is het niet mogelijk om hier een conclusie over te trekken. Ten tweede blijken de ervaring van de leerkracht, de meertaligheid van de leerling en de gehanteerde onderwijsmethode belangrijke voorspellers van de scores te zijn. Over de gehanteerde onderwijsmethode kunnen geen harde conclusies worden getrokken, aangezien deze factor erg samenhangt met de andere schoolgebonden factoren. Het feit dat de ervaring van de leerkracht als significante voorspeller naar voren kwam, laat echter zien dat het voordelig is om leerkrachten die al langer lesgeven in het Engels in te zetten in het tpo-onderwijs. Het feit dat meertalige leerlingen beter presteren dan eentalige leerlingen, zelfs wanneer deze niet het Engels als moedertaal hebben, laat daarnaast zien dat er in het tpo-onderwijs rekening gehouden moet worden met de hoeveelheid talen die een leerling spreekt. Ook bevestigt dit een voordeel van meertaligheid die in de literatuur vaak wordt aangehaald, namelijk dat meertaligen een nieuwe taal sneller oppikken dan eentaligen (Abu-Rabia & Sanitsky, 2010; Fleckenstein, Möller & Baumert, 2018; Hopp et al., 2019; Maluch, Kempert, Neumann & Stanat, 2015). Bovendien laat het zien dat ouders niet bang hoeven zijn om hun meertalige kind naar een

tweetalige basisschool te sturen; hun talige voorkennis is geen hindernis, maar juist een voordeel.

## Hoofdstuk 7: Conclusie

Het doel van het huidige onderzoek was om te onderzoeken of de Engelse productieve taalvaardigheden van leerlingen uit de tweede lichting van tweetalige groepen 7 waren verbeterd ten opzichte van de eerste lichting. De hypothese was dat er een stijging in de leerprestaties zichtbaar zou zijn, aangezien de scholen naar aanleiding van ervaringen binnen het eerste jaar verbeteringen in hun tweetalig onderwijs hebben toegepast. Daarnaast was de verwachting dat een aantal andere kind- en schoolgebonden factoren de leerprestaties zouden beïnvloeden, namelijk het geslacht, de leeftijd, de meertaligheid en de moedertaal van een leerling, het percentage onderwijstijd in het Engels, de ervaring van de leerkracht en de gehanteerde onderwijsmethode. Uit de resultaten bleek dat de eerste hypothese moest worden verworpen; er was geen verschil zichtbaar tussen de productieve taalvaardigheden van de eerste en tweede lichting. Ook bleken niet alle kind- en schoolgebonden factoren van invloed te zijn op de scores. Alleen de ervaring van een leerkracht met lesgeven in het Engels en de meertaligheid van een leerling bleken significante voorspellers van de variantie in de scores te zijn, waarbij meertaligheid en meer ervaring beide resulteerden in hogere scores. Daarnaast bleek er nog een andere factor van belang te zijn, namelijk de gehanteerde onderwijsmethode. Leerlingen die werden onderwezen met de *One Teacher, One Language* benadering presteerden beter dan de leerlingen die werden onderwezen met de *Sandwich* en *One Situation, One Language* benaderingen. Uit het onderzoek kan geconcludeerd worden dat één jaar aan schoolervaring met het tweetalig onderwijzen van groep 7 niet genoeg is om een significante verbetering te zien in de Engelse productieve taalvaardigheden van leerlingen. Vervolgonderzoek kan zich richten op latere lichtingen, om na te gaan of meer ervaring met het tweetalig onderwijs wel in een verschil zal resulteren.

### Literatuurlijst

- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language learning*, 59(2), 249-306.
- Abu-Rabia, S., & Sanitsky, E. (2010). Advantages of bilinguals over monolinguals in learning a third language. *Bilingual Research Journal*, 33(2), 173-199.
- Alario, F. X., Goslin, J., Michel, V., & Laganaro, M. (2010). The functional origin of the foreign accent: evidence from the syllable-frequency effect in bilingual speakers. *Psychological science*, 21(1), 15-20.
- Allgäuer-Hackl, E., & Jessner, U. (2019). Cross-linguistic Interaction and Multilingual Awareness. *Multidisciplinary Perspectives on Multilingualism: The Fundamentals*, 19, 325.
- Arabski, J., & Wojtaszek, A. (Eds.). (2011). *Individual learner differences in SLA*. Bristol: Multilingual Matters.
- Artieda, G., Roquet, H., & Nicolás-Conesa, F. (2020). The impact of age and exposure on EFL achievement in two learning contexts: formal instruction and formal instruction+ content and language integrated learning (CLIL). *International Journal of Bilingual Education and Bilingualism*, 23(4), 449-472.

- Baba, K., Takemoto, Y., & Yokochi, M. (2014). Relationship between second language speaking and writing skills and modality preference of university EFL students. *Treatises and Studies by the Faculty of Kinjo Gakuin University. Studies in Social Sciences*, 10(1), 56-68.
- Babayigit, S. (2015). The dimensions of written expression: Language group and gender differences. *Learning and Instruction*, 35, 33-41.
- Bialystok, E., & Barac, R. (2013). Cognitive effects. *The Psycholinguistics of Bilingualism*, 83(2), 192-213.
- Blom, E., & Boerma, T. (2020). Bilingual children's lexical and narrative comprehension in Dutch as the majority language. *Developing Narrative Comprehension: Multilingual Assessment Instrument for Narratives*, 61, 197.
- Bovellan, E. (2014). *Teachers' beliefs about learning and language as reflected in their views of teaching materials for Content and Language Integrated Learning (CLIL)* [Doctoraal proefschrift, University of Jyväskylä]. Geraadpleegd op 26 mei 2022, van [https://jyx.jyu.fi/bitstream/handle/123456789/44277/1/978-951-39-5809-1\\_vaitos20092014.pdf](https://jyx.jyu.fi/bitstream/handle/123456789/44277/1/978-951-39-5809-1_vaitos20092014.pdf)
- Bret, A. (2011). *Implementing CLIL in a primary school in Spain: The effects of CLIL on L2 English learners' oral production skills* [Masterscriptie, Universitat Autònoma de Barcelona]. Geraadpleegd op 12 mei 2022, van <https://ddd.uab.cat/record/77027>
- Bylund, E., Hyltenstam, K., & Abrahamsson, N. (2021). Age of acquisition—not bilingualism—is the primary determinant of less than nativelike L2 ultimate attainment. *Bilingualism: Language and Cognition*, 24(1), 18-30.
- CBS (2021). *Onderwijsinstellingen; grootte, soort, levensbeschouwelijke grondslag*. Geraadpleegd op 16 juni 2022, van <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/03753/table?fromstatweb>
- Chondrogianni, V., & Marinis, T. (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children. *Linguistic Approaches to Bilingualism*, 1(3), 318-345.
- Cobley, S., McKenna, J., Baker, J., & Wattie, N. (2009). How pervasive are relative age effects in secondary school education?. *Journal of Educational Psychology*, 101(2), 520.
- Cook, V., & Singleton, D. (2014). *Key Topics in Second Language Acquisition* (Vol. 10). Multilingual Matters.
- Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL*. Cambridge University Press.
- Custodio-Espinar, M. (2020). Influencing Factors on In-Service Teachers' Competence in Planning CLIL. *Latin American Journal of Content & Language Integrated Learning*, 12(2), 207-241.
- Dalton-Puffer, C. (2007). *Discourse in Content and Language Integrated Learning (CLIL) Classrooms*. Amsterdam: John Benjamins Publishing Company.

- Dalton-Puffer, C. (2008). Outcomes and processes in content and language integrated learning (CLIL): Current research from Europe. *International Journal of Bilingual Education and Bilingualism*, 10, 543–562.
- Dalton-Puffer, C. (2011). Content-and-language integrated learning: From practice to principles? *Annual Review of applied linguistics*, 31, 182-204.
- De Diezmas, E. N. M. (2016). The impact of CLIL on the acquisition of L2 competences and skills in primary education. *International Journal of English Studies*, 16(2), 81-101.
- De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020). Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important?. *Bilingualism: Language and Cognition*, 23(1), 171-185.
- Delliou, A., & Zafiri, M. (2016). Developing the speaking skills of students through CLIL: A case of sixth grade Primary School students in Greece. *EIIC – Electronic International Interdisciplinary Conference*, 5(1).
- Dhuey, E., Figlio, D., Karbownik, K., & Roth, J. (2019). School starting age and cognitive development. *Journal of Policy Analysis and Management*, 38(3), 538-578.
- Doiz, A., Lasagabaster, D., & Sierra, J. M. (2014). CLIL and motivation: The effect of individual and contextual variables. *The language learning journal*, 42(2), 209-224.
- Driessen, G., Krikhaar, E., De Graaff, R., Unsworth, S., Leest, B., Coppens, K., & Wierenga, J. (2016). *Evaluatie pilot Tweetalig Primair Onderwijs: Startmeting schooljaar 2014/15*. FoTo.
- Erdiana, N., Bahri Ys, S., & Akhmal, C. N. (2019). Male vs. female EFL students: Who is better in speaking skills? *Studies in English Language and Education*, 6(1), 131-140.
- Etchell, A., Adhikari, A., Weinberg, L. S., Choo, A. L., Garnett, E. O., Chow, H. M., & Chang, S. E. (2018). A systematic literature review of sex differences in childhood language and brain development. *Neuropsychologia*, 114, 19-31.
- Expertisecentrum Nederlands (z.d.). *Zicht op vroeg vreemdetalenonderwijs*. Geraadpleegd op 16 juni 2022, van <https://expertisecentrumnederlands.nl/zicht-op-vroeg-vreemdetalenonderwijs>
- Farrell, T. S. (2006). The first year of language teaching: Imposing order. *System*, 34(2), 211-221.
- Fernández, R., & Halbach, A. (2011). Analysing the situation of teachers in the CAM bilingual project after four years of implementation. *Content and Foreign Language Integrated Learning. Contributions to multilingualism in European contexts*. New York: Peter Lang.
- Festman, J. (2020). Learning and processing multiple languages: The more the easier?. *Language Learning*, 71(S1), 121-162.
- Fleckenstein, J., Möller, J., & Baumert, J. (2018). Mehrsprachigkeit als Ressource. *Zeitschrift für Erziehungswissenschaft*, 21(1), 97-120.
- Gabillon, Z. (2020). Revisiting CLIL: Background, Pedagogy, and Theoretical Underpinnings. *Contextes et didactiques. Revue semestrielle en sciences de l'éducation*, (15).

- Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balèiūnienė, I., et al. (2015). Assessment of narrative abilities in bilingual children. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Assessing multilingual children: Disentangling bilingualism from language impairment* (pp. 241–274). Bristol: Multilingual Matters.
- Gallardo-del-Puerto, F., & Blanco-Suárez, Z. (2021). Foreign Language Motivation in Primary Education students: The effects of additional CLIL and gender. *Journal of Immersion and Content-based Language Education*, 9(1), 58-84.
- Golkova, D., & Hubackova, S. (2014). Productive skills in second language learning. *Procedia-Social and Behavioral Sciences*, 143, 477-481.
- Goris, J. A., Denessen, E. J. P. G., & Verhoeven, L. T. W. (2019). Effects of content and language integrated learning in Europe: A systematic review of longitudinal experimental studies. *European Educational Research Journal*, 18(6), 675-698.
- Guerrero, R. (2005). *Task complexity and L2 narrative oral production* [Doctoraal proefschrift, University of Barcelona]. Geraadpleegd op 26 mei 2022, van [https://www.tdx.cat/bitstream/handle/10803/1662/00.CHAPTER\\_0.pdf?sequence=1](https://www.tdx.cat/bitstream/handle/10803/1662/00.CHAPTER_0.pdf?sequence=1)
- Hayes, A. F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis* (2de editie). Guilford Publications.
- Hebert, E., & Worthy, T. (2001). Does the first year of teaching have to be a bad one? A case study of success. *Teaching and teacher education*, 17(8), 897-911.
- Heras, A., & Lasagabaster, D. (2015). The impact of CLIL on affective factors and vocabulary learning. *Language Teaching Research*, 19(1), 70-88.
- Hiligsmann, P., Van Mensel, L., Galand, B., Mettwie, L., Meunier, F., Szmalec, A., & Simonis, M. (2017). Assessing Content and Language Integrated Learning (CLIL) in French-speaking Belgium: Linguistic, cognitive, and educational perspectives. *Les Cahiers de Recherche du Girsef*, 109, 1-24.
- Hirosh, Z., & Degani, T. (2018). Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic bulletin & review*, 25(3), 892-916.
- Hopp, H., Vogelbacher, M., Kieseier, T., & Thoma, D. (2019). Bilingual advantages in early foreign language learning: Effects of the minority and the majority language. *Learning and Instruction*, 61, 99-110.
- Infante, D., Benvenuto, G., & Lastrucci, E. (2009). The effects of CLIL from the perspective of experienced teachers. *CLIL practice: Perspectives from the field*, 156-163.
- Iwaniec, J. (2019). Language learning motivation and gender: The case of Poland. *International Journal of Applied Linguistics*, 29(1), 130-143.
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. London, UK: Routledge.

- Jenniskens, T., Leest, B., Wolbers, M., Bruggink, M., Krikhaar, E., De Graaff, R., Spätgens, T., & Unsworth, S. (2020). *Evaluatie pilot Tweetalig Primair Onderwijs: Vervolgmeting schooljaar 2018/19*. FoTo.
- Jenniskens, T., Leest, B., Wolbers, M., Krikhaar, E., Teunissen, C., De Graaff, R., Unsworth, S., & Coppens, K. (2018). *Evaluatie pilot Tweetalig Primair Onderwijs: Vervolgmeting schooljaar 2016/17*. FoTo.
- Jensen, S. H. (2017). Gaming as an English language learning resource among young children in Denmark. *Calico Journal*, 34(1), 1-19.
- Kang, J. (2020). Speaking and Writing Connections in L2: The Roles of Multimodal Teaching and Learning. *어학연구 (Language Research)*, 56.
- Kayımbaşoğlu, D., Oktekin, B., & Hacı, H. (2016). Integration of gamification technology in education. *Procedia Computer Science*, 102, 668-676.
- Kormos, J., & Trebits, A. (2012). The role of task complexity, modality, and aptitude in narrative task performance. *Language Learning*, 62(2), 439-472.
- Krenca, K., Hipfner-Boucher, K., & Chen, X. (2020). Grammatical gender-marking ability of multilingual children in French immersion. *International Journal of Bilingualism*, 24(5-6), 968-983.
- Kwakernaak, E. (2014). *Didactiek van het vreemdetalenonderwijs*. Uitgeverij Coutinho.
- Lahuerta, A. (2020). Analysis of accuracy in the writing of EFL students enrolled on CLIL and non-CLIL programmes: the impact of grade and gender. *The Language Learning Journal*, 48(2), 121-132.
- Lampert, M. (2003). *Teaching problems and the problems of teaching*. New Haven, CT: Yale University Press.
- Lasagabaster, D. (2008). Foreign Language Competence in Content and Language Integrated Courses. *Open Applied Linguistics Journal*, 1, 30-41.
- Leung, Y. K. I. (2005). L2 vs. L3 initial state: A comparative study of the acquisition of French DPs by Vietnamese monolinguals and Cantonese-English bilinguals. *Bilingualism: Language and cognition*, 8(1), 39-61.
- Levelt, W. (1989). *Speaking. From intention to articulation*. Cambridge, MA-London: ACL-MIT Press.
- Lieberoth, A. (2015). Shallow gamification: Testing psychological effects of framing an activity as a game. *Games and Culture*, 10(3), 229-248.
- Lorenzo, F., Casal, S., & Moore, P. (2010). The effects of content and language integrated learning in European education: Key findings from the Andalusian bilingual sections evaluation project. *Applied linguistics*, 31(3), 418-442.



- Mady, C. (2014). Learning French as a second official language in Canada: Comparing monolingual and bilingual students at Grade 6. *International Journal of Bilingual Education and Bilingualism*, 17(3), 330-344.
- Maluch, J. T., & Kempert, S. (2017). Bilingual profiles and third language learning: The effects of the manner of learning, sequence of bilingual acquisition, and language use practices. *International Journal of Bilingual Education and Bilingualism*.
- Maluch, J. T., Kempert, S., Neumann, M., & Stanat, P. (2015). The effect of speaking a minority language at home on foreign language learning. *Learning and instruction*, 36, 76-85.
- Manning, J. (2019, 14 januari). *Introducing English Benchmark: An innovative test for young learners*. Pearson English. Geraadpleegd op 4 mei 2022, van <https://www.english.com/blog/english-benchmark-innovative-test-for-young-learners/>
- Martinez Agudo, J. D. D. (2022). To what extent does parental educational background affect CLIL learners' content subject learning? Evidence from research in Spain. *International Journal of Bilingual Education and Bilingualism*, 25(2), 598-611.
- Masduqi, H. (2016). Integrating receptive skills and productive skills into a reading lesson. In *Proceeding of the International Conference on Teacher Training and Education* (No. 1, pp. 507-511).
- Mattheoudakis, M., Alexiou, T., & Laskaridou, C. (2014). To CLIL or not to CLIL? The Case of the 3rd Experimental Primary School in Evosmos. *Major Trends in Theoretical and Applied Linguistics 3: Selected Papers from the 20th ISTAL*, 3, 215.
- Mavilidi, M. F., Marsh, H. W., Xu, K. M., Parker, P. D., Jansen, P. W., & Paas, F. (2021). Relative age effects on academic achievement in the first ten years of formal schooling: A nationally representative longitudinal prospective study. *Journal of Educational Psychology*.
- Mehisto, P., Marsh, D., & Frigols, M. J. (2008). *Uncovering CLIL: Content and language integrated learning in bilingual and multilingual education*. Oxford: Macmillan Education.
- Merisuo-Storm, T. (2007). Pupils' attitudes towards foreign-language learning and the development of literacy skills in bilingual education. *Teaching and teacher education*, 23(2), 226-235.
- Mitrofanova, N., Rodina, Y., Urek, O., & Westergaard, M. (2018). Bilinguals' sensitivity to grammatical gender cues in Russian: the role of cumulative input, proficiency, and dominance. *Frontiers in Psychology*, 1894.
- Montanari, S. (2019). Facilitated language learning in multilinguals. In S. Montanari & S. Quay (Eds.), *Multidisciplinary perspectives on multilingualism: The fundamentals* (pp. 302–324). Berlin, Germany: De Gruyter.
- Moonen, M. (2021, 2 december). *Global Scale of English*. SLO. Geraadpleegd op 4 mei 2022, van <https://www.slo.nl/thema/vakspecificieke-thema/mvt/global-scale-english/>
- Navarro Pablo, M., & García Jiménez, E. (2018). Are CLIL Students More Motivated? An Analysis of Affective Factors and their Relation to Language Attainment. *Porta Linguarum*, 71, 90.

- Nuffic (2020). *Dutch primary schools in Nuffic networks*. Geraadpleegd op 16 juni 2022, van <https://www.nuffic.nl/en/subjects/facts-and-figures/dutch-primary-schools-in-nuffic-networks>
- Nuffic (z.d.). *Tweetalig primair onderwijs (tpo)*. Geraadpleegd op 16 juni 2022, van <https://www.nuffic.nl/onderwerpen/tweetalig-primair-onderwijs/tweetalig-primair-onderwijs-tpo>
- Nye, B., Konstantopoulos, S., & Hedges, L. V. (2004). How large are teacher effects?. *Educational evaluation and policy analysis*, 26(3), 237-257.
- Onderwijsinspectie (z.d.). *Wie bepaalt of een kind overgaat naar groep 3?* Geraadpleegd op 13 mei 2022, van <https://www.onderwijsinspectie.nl/documenten/vragen-en-antwoorden/wie-bepaalt-of-een-kind-overgaat-naar-groep-3>
- Öztürk, G., & Gürbüz, N. (2013). The impact of gender on foreign language speaking anxiety and motivation. *Procedia-Social and Behavioral Sciences*, 70, 654-665.
- Paradis, M. (2008). Language and communication disorders in multilinguals. In B. Stemmer & H. A. Whitaker (Eds.), *Handbook of the neuroscience of language* (pp. 341–349). Amsterdam, Netherlands: Elsevier.
- Pearson (2019a). *English Benchmark Young Learners: Customer Readiness & Implementation Guide* [Brochure]. Geraadpleegd op 4 mei 2022, van [https://mypearsonhelp.com/helpconsole7/kb\\_pep\\_en\(old\)/docs/Pearson%20Benchmark%20-%20Customer%20Readiness%20&%20Implementation%20Pilot%20Users%20Guide.pdf](https://mypearsonhelp.com/helpconsole7/kb_pep_en(old)/docs/Pearson%20Benchmark%20-%20Customer%20Readiness%20&%20Implementation%20Pilot%20Users%20Guide.pdf)
- Pearson (2019b). *Global Scale of English Learning Objectives for Young Learners* [Brochure]. Geraadpleegd op 4 mei 2022, van <https://online.flippingbook.com/view/872842/>
- Pearson (2021). *English Benchmark Young Learners* [Brochure]. Geraadpleegd op 4 mei 2022, van <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/english/SupportingDocs/PEBYL/english-benchmark-young-learners-brochure.pdf>
- Pearson (2022a). *Waarom English Benchmark Young Learners?* Geraadpleegd op 4 mei 2022, van [https://www.pearson.com/nl/nl\\_NL/po/alle-groepen/english-benchmark.html](https://www.pearson.com/nl/nl_NL/po/alle-groepen/english-benchmark.html)
- Pearson (2022b). *Global Scale of English*. Geraadpleegd op 4 mei 2022, van <https://www.pearson.com/english/about-us/global-scale-of-english.html>
- Pérez-Cañado, M. L. (2012). CLIL research in Europe: Past, present, and future. *International Journal of Bilingual Education and Bilingualism*, 15(3), 315-341.
- Pfenninger, S. E. (2016). All good things come in threes: Early English learning, CLIL and motivation in Switzerland. *Cahiers du Centre de Linguistique et des Sciences du Langage*, (48), 119-147.
- Pitoyo, M. D., Sumardi, S., & Asib, A. (2019). Gamification based assessment: A Test Anxiety Reduction through Game Elements in Quizizz Platform. *International Online Journal of Education and Teaching (IOJET)*, 6(3), 456–471.

- Pladevall-Ballester, E. (2019). A longitudinal study of primary school EFL learning motivation in CLIL and non-CLIL settings. *Language Teaching Research*, 23(6), 765-786.
- Rauch, D. P., Naumann, J., & Jude, N. (2012). Metalinguistic awareness mediates effects of full biliteracy on third-language reading proficiency in Turkish–German bilinguals. *International Journal of Bilingualism*, 16(4), 402-418.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2), 417-458.
- Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American economic review*, 94(2), 247-252.
- Roquet, H., & Pérez-Vidal, C. (2017). Do productive skills improve in content and language integrated learning contexts? The case of writing. *Applied Linguistics*, 38(4), 489-511.
- Roquet, H., Llopis, J., & Pérez-Vidal, C. (2016). Does gender have an impact on the potential benefits learners may achieve in two contexts compared: Formal instruction and formal instruction + content and language integrated learning? *International Journal of Bilingual Education and Bilingualism*, 19(4), 370–386.
- Ruiz de Zarobe, Y. (2008). CLIL and foreign language learning: A longitudinal study in the Basque country. *International CLIL Research Journal*, 1(1), 60–73.
- Ruiz de Zarobe, Y. (2010). Written production and CLIL. In C. Dalton-Puffer, T. Nikula, & U. Smit (Eds.), *Language use and language learning in CLIL classrooms* (pp. 191–212). Berlin: John Benjamins.
- Sabra, S. (2018). *Gender differences in young learners' English skills in Swedish schools* [Masterscriptie, University of Gävle]. Geraadpleegd op 18 mei 2022, van <https://www.diva-portal.org/smash/get/diva2:1212666/FULLTEXT01.pdf>
- Sakic, M., Burusic, J., & Babarovic, T. (2013). The relation between school entrance age and school achievement during primary schooling: Evidence from Croatian primary schools. *British journal of educational psychology*, 83(4), 651-663.
- Salahshour, F., Sharifi, M., & Salahshour, N. (2013). The relationship between language learning strategy use, language proficiency level and learner gender. *Procedia-Social and Behavioral Sciences*, 70, 634-643.
- Scheele, A. F., Leseman, P. P., & Mayo, A. Y. (2009). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, 31(1), 117-140.
- Schoonen, R., Snellings, P., Stevenson, M., & Van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. *Writing in foreign language contexts: Learning, teaching, and research*, 77-101.
- Smolander, S., Laasonen, M., Arkkila, E., Lahti-Nuutila, P., & Kunnari, S. (2020). L2 vocabulary acquisition of early sequentially bilingual children with TD and DLD affected differently by

- exposure and age of onset. *International Journal of Language & Communication Disorders*, 56(1), 72-89.
- Sroka, M. C., Vannest, J., Maloney, T. C., Horowitz-Kraus, T., Byars, A. W., & Holland, S. K. (2015). Relationship between receptive vocabulary and the neural substrates for story processing in preschoolers. *Brain imaging and behavior*, 9(1), 43-55.
- Stigler, J., & Miller, K. (2018). Expertise and Expert Performance in Teaching. In A. Ericsson, R. Hoffman, A. Kozbelt, & M. Williams (Eds.), *The Cambridge Handbook*.
- Sundqvist, P. (2009). The impact of spare time activities on students' English language skills. *Vägar till språk och litteratur*, 63-76.
- Thompson, A. S., & Sylvé, L. K. (2015). "Does English make you nervous?" Anxiety profiles of CLIL and non-CLIL students in Sweden. *Apples-Journal of Applied Language Studies*.
- Trebits, A., Koch, M. J., Ponto, K., Bruhn, A. C., Adler, M., & Kersten, K. (2022). Cognitive gains and socioeconomic status in early second language acquisition in immersion and EFL learning settings. *International Journal of Bilingual Education and Bilingualism*, 1-14.
- Unsworth, S. (2012). Assessing the role of current and cumulative exposure in simultaneous bilingual acquisition: The case of Dutch gender. *Bilingualism: Language and cognition*, 16(1), 86-110.
- Unsworth, S., Persson, L., Prins, T., & De Bot, K. (2014). An investigation of factors affecting early foreign language learning in the Netherlands. *Applied Linguistics*, 36(5), 527-548.
- Van De Craen, P., Mondt, K., Allain, L., & Gao, Y. (2007). Why and How CLIL Works. An Outline for a CLIL Theory. *Vienna English Working papers*, 16(3), 70-78.
- Van Der Slik, F. W., Van Hout, R. W., & Schepens, J. J. (2015). The gender gap in second language acquisition: Gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *PloS one*, 10(11), e0142056.
- Van Mensel, L., Bulon, A., Hendriks, I., Meunier, F., & Van Goethem, K. (2020). Effects of input on L2 writing in English and Dutch: CLIL and non-CLIL learners in French-speaking Belgium. *Journal of Immersion and Content-Based Language Education*, 8(2), 173-199.
- Van Voorhis, C. W., & Morgan, B. L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in quantitative methods for psychology*, 3(2), 43-50.
- Verheyden, L., Van den Branden, K., Rijlaarsdam, G., Van den Bergh, H., & De Maeyer, S. (2010). Written narrations by 8-to 10-year-old Turkish pupils in Flemish primary education: A follow-up of seven text features. *Journal of Research in Reading*, 33(1), 20-38.
- Walczak, A., & Geranpayeh, A. (2015). *The Gender Gap in English Language Proficiency? Insights from a Test of Academic English* [PowerPoint-slides]. Cambridge English. Geraadpleegd op 18 mei 2022, van <https://www.cambridgeassessment.org.uk/Images/gender-differences-cambridge-english.pdf>
- Weissberg, B. (2000). Developmental relationships in the acquisition of English syntax: writing vs. speech. *Learning and Instruction*, 10, 37-53.

- Weissberg, B. (2006). *Connecting speaking and writing in second language writing instruction*. Ann Arbor, MI: University of Michigan Press.
- Wen, Z. E., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, 50(1), 1-31.
- Witney, J., & Dewaele, J. M. (2018). Learning two or more languages. *The Cambridge guide to learning English as a second language*, 43-52.
- Wolff, D. (2007). CLIL: Bridging the gap between school and working life. In D. Marsh & D. Wolff (Eds.), *Diverse contexts—converging goals. CLIL in Europe* (pp. 15–25). Frankfurt am Main, Germany: Peter Lang.

## Bijlage A

Tabel 1

*Informatie over de tpo-ervaring en TEFL-ervaring (in jaren) van de leerkracht van elke deelnemende groep 7, verdeeld per lichter*

School		Lichter 1	Lichter 2
School 1	Ervaring tpo	0	1
	Ervaring TEFL	0	1
School 2	Ervaring tpo	2	3
	Ervaring TEFL	4	5
School 3	Ervaring tpo	2	5, 1
	Ervaring TEFL	2	9, 1
School 4	Ervaring tpo	2	3
	Ervaring TEFL	23	24
School 5	Ervaring tpo	6, 7	7, 6
	Ervaring TEFL	9, 11	10, 17
School 6	Ervaring tpo	5	6
	Ervaring TEFL	5	6

*Note:* aangezien sommige scholen meer dan één groep 7 tegelijk hadden, worden er soms twee getallen weergegeven.