

G I O N

Gronings Instituut voor Onderzoek van Onderwijs
Grote Rozenstraat 3 9712 TG Groningen



rijksuniversiteit
 groningen / gion



Universiteit Utrecht
 Faculteit Geesteswetenschappen

Voorstudie Pilot Primair Tweektalig Onderwijs

Penvoerder:

GION, Rijksuniversiteit Groningen

Ellen van den Broek

Rick de Graaff

Sharon Unsworth

Vicky van der Zee

Universiteit Utrecht, Faculteit Geesteswetenschappen

Voorstudie Pilot Primair Tweektalig Onderwijs Ellen van den Broek Rick de Graaff Sharon Unsworth Vicky van der Zee

Voorstudie Pilot Tweektalig Primair Onderwijs

verplichtingnr. 50500-79653
NARCIS onderzoeksnummer: OND1355031

31 januari 2014

Penvoerder:
GION, Rijksuniversiteit Groningen

Auteurs:
Ellen van den Broek, Rick de Graaff, Sharon Unsworth, Vicky van der Zee
Universiteit Utrecht, Faculteit Geesteswetenschappen



Universiteit Utrecht
Faculteit Geesteswetenschappen

Klankbordgroep:
K. de Bot, Rijksuniversiteit Groningen
A. Corda, Hogeschool van Amsterdam
E. Krikhaar, Radboud Universiteit Nijmegen

Opdrachtgever:
Ministerie van OCW, directie Primair Onderwijs
Stuurgroep TPO

ISBN 978-90-6690-000-4

2014 GION, Gronings Instituut voor Onderzoek van het Onderwijs

Niets uit deze opgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook zonder voorafgaande schriftelijke toestemming van de directeur van het instituut.

No part of this book may be reproduced in any form, by print, photo print, microfilm or any other means without written permission of the director of the institute.

Inhoudsopgave

Inhoudsopgave.....	I
Samenvatting	1
Voorstudie Pilot Tweektalig Primair Onderwijs.....	5
0. Inleiding	5
0.1 Introductie	5
0.2 Terminologie.....	5
Talen.....	5
Tpo-onderwijs.....	5
0.3 Onderzoeksvragen.....	6
0.4 Methode.....	6
0.5 Experts	8
0.6 Opzet van de rapportage	8
0.7 Werkwijze	8
1. Wat voor voorbeelden zijn er in het buitenland die wat betreft opzet en inrichting van de programma's relevant en inzetbaar kunnen zijn voor de tpo-pilot?	11
1.1 Organisatie scholen en programma's	11
1.2 Startleeftijd	11
1.3 Intensiteit L2	12
1.4 Onderwijsaanbod L2	12
1.5 Externe factoren.....	13
Taalafstand	13
Buitenschools taalcontact	14
Overige factoren	14
1.6 Conclusie	14
2. Welke evaluatie-instrumenten bestaan er die ingezet kunnen worden voor kwalitatief en kwantitatief evaluatieonderzoek naar tpo?	17
2.1 L1	17
2.2 L2	17
2.3 Vakinhoudelijke kennis.....	18
2.4 Overige vaardigheden	18
2.5 Conclusie	19

3. Welke opbrengsten van tpo zijn gerapporteerd wat betreft de eerste schooltaal en de tweede doeltaal?	21
3.1 L1	21
3.2 L2	22
3.3 Vakinhoudelijke kennis	23
3.4 Externe factoren.....	24
Anderstalige achtergrond	24
Opleidingsniveau ouders	24
SES.....	25
Cognitieve vaardigheden	25
3.5 Conclusie	25
4. Wat zijn de relevante kenmerken en talige en didactische vaardigheden van tpo-leerkrachten?	27
4.1 Profiel leerkracht	27
4.2 Taalachtergrond	27
4.3 Scholing en didactische vaardigheden.....	28
4.4 Tpo-training.....	28
4.5 Ontwikkeling materialen	29
4.6 Conclusie	30
5. Welke attitudes zijn er t.a.v. het onderwijsprogramma, het Engels en de eerste schooltaal bij leerlingen, leerkrachten en ouders?	31
5.1 Evaluatie	31
Leerlingen	31
Leerkrachten	32
Ouders	33
5.2 Evaluatievormen	33
Leerlingen	33
Leerkrachten	34
Ouders	34
Hawthorne-effect	34
5.3 Conclusie	34
6. Wat zijn mogelijkheden voor internationale samenwerking op het terrein van curriculumontwikkeling en onderzoek?.....	37
7. Discussie en aanbevelingen	39
7.1 Inrichting tpo-programma's en deelnemende scholen	39

7.2 Opzet onderzoek – wat te meten en hoe?	40
Op leerlingniveau	40
Op leerkrachtniveau	41
Op organisatieniveau	41
Onderzoeksmatige uitdagingen	42
8. Referenties.....	45
Appendix I – Summaries of expert interviews	49
Kristin Kersten (Universit�t Hildesheim, Germany).....	49
Marianne Nikolov (University of P�cs, Hungary).....	52
Pilar Medrano (Ministry of Education, Madrid, Spain)	57
Fred Genesee (McGill University, Montr�al, Canada).....	63
Johanna Watzinger-Tharp (University of Utah, US).	70
Hanna �sterlund (Skolverket, Sweden)	75
Appendix II – Overview of evaluation instruments	81
Appendix III – Matrix Literature Review	85

Samenvatting

Dit rapport doet verslag van een internationale literatuurstudie en expertconsultatie ten behoeve van de opzet, uitvoering en evaluatie van een pilot tweetalig primair onderwijs (tpo) in Nederland. Hierbij krijgen maximaal twintig basisscholen de komende jaren de gelegenheid om tweetalig onderwijs in te richten, waarin 30% tot 50% van de onderwijstijd in het Engels, Duits of Frans wordt aangeboden (L2).

In dit rapport wordt antwoord gegeven op de volgende onderzoeksvragen:

1. Wat voor voorbeelden zijn er in het buitenland die wat betreft opzet en inrichting van de programma's relevant en inzetbaar kunnen zijn voor de tpo-pilot?
2. Welke evaluatie-instrumenten bestaan er die ingezet kunnen worden voor kwalitatief en kwantitatief evaluatieonderzoek naar tpo?
3. Welke opbrengsten van tpo zijn gerapporteerd wat betreft de eerste schooltaal en de tweede doeltaal?
4. Wat zijn de relevante kenmerken en talige en didactische vaardigheden van tpo-leerkrachten?
5. Welke attitudes zijn er t.a.v. het onderwijsprogramma, het Engels en de eerste schooltaal bij leerlingen, leerkrachten en ouders?
6. Wat zijn mogelijkheden voor internationale samenwerking op het terrein van curriculumontwikkeling en onderzoek?

In het rapport geven we aan de hand van bovenstaande onderzoeksvragen een overzicht van bestaande studies naar tpo, didactische kenmerken van de beschreven tpo-praktijken, de mogelijke evaluatie-instrumenten om het effect van tpo op de taalvaardigheid in de L1 en de L2 te meten en de opbrengsten in de L1 en L2 uit reeds uitgevoerde tpo-studies. In hoofdstuk 1 worden gepubliceerde studies en resultaten van tpo-studies uit een aantal landen binnen en buiten Europa besproken. In hoofdstuk 2 worden de mogelijkheden van evaluatieonderzoek naar tpo besproken, met voorbeelden van evaluatie-instrumenten zoals deze ingezet zijn in andere studies. In hoofdstuk 3 worden de opbrengsten van tpo met betrekking tot de L1 en L2 geëvalueerd. In hoofdstuk 4 worden relevante didactische vaardigheden van tpo-leerkrachten besproken, evenals achtergrondinformatie met betrekking tot scholing en onderwijservaring. In hoofdstuk 5 wordt de houding van diverse belanghebbenden ten opzichte van het tpo-programma en daarbij horende evaluatie-instrumenten besproken. In hoofdstuk 6 zullen vervolgens de mogelijkheden tot internationale samenwerking worden besproken. Ten slotte worden in hoofdstuk 7 aanbevelingen gegeven voor de inrichting van de tpo-pilot en de opzet en focus van het onderzoek.

Onderzoeksvraag 1: Wat voor voorbeelden zijn er in het buitenland die wat betreft opzet en inrichting van de programma's relevant en inzetbaar kunnen zijn voor de tpo-pilot?

Aan de hand van geïnventariseerde literatuur is een selectie gemaakt van 29 studies die qua opzet en inrichting relevant zijn voor de tpo-pilot. In nagenoeg alle studies wordt tussen de 30% en 50% instructie in de L2 aangeboden en ligt de startleeftijd van de leerlingen tussen de vier en zes jaar. Verschillende studies bevinden zich nog in de opstartfase van tpo en vertonen daardoor overeenkomsten met de beginsituatie van de tpo-pilot in Nederland. Er zijn ook diverse studies die doorlopend testen hebben uitgevoerd en de ontwikkeling over de gehele basisschoolperiode hebben gemeten. De invulling van het onderwijsaanbod in de L2 verschilt sterk in de geïnventariseerde studies en is vaak afhankelijk van het curriculum en de beschikbare docenten. Wel blijkt er een voorkeur uitgesproken te worden voor vakken die een relevante functie buiten de school kunnen vervullen, zoals wereldoriënterende vakken.

Onderzoeksvraag 2: Welke evaluatie-instrumenten bestaan er die ingezet kunnen worden voor kwalitatief en kwantitatief evaluatieonderzoek naar tpo?

Er bestaat een grote variëteit aan evaluatie-instrumenten die ingezet kunnen worden voor kwalitatief en kwantitatief onderzoek naar de opbrengsten van tpo. Hierbij wordt niet alleen een onderscheid gemaakt tussen evaluatie-instrumenten voor beoordeling van de L1 en L2, maar ook voor de evaluatie van vakinhoudelijke of cognitieve vaardigheden. Veel studies maken gebruik van een combinatie van gestandaardiseerde en zelfontwikkelde instrumenten om de diverse vaardigheden te testen. Het voordeel van het gebruik van gestandaardiseerde instrumenten is dat leerlingen en scholen op nationaal en internationaal niveau met elkaar vergeleken kunnen worden. Voorbeelden van zulke gestandaardiseerde instrumenten die gebruikt kunnen worden voor de beoordeling van de L1 en L2 vaardigheden zijn het PIRLS onderzoek voor leesvaardigheid, de PPVT voor woordenschat en de TROG voor de kennis van grammatica. Zelfontwikkelde instrumenten hebben als voordeel dat ze beter afgestemd kunnen worden op het curriculum waardoor specifiekere vaardigheden getest kunnen worden. Zelfontwikkelde testen kunnen daarom als goede aanvulling dienen op de genormeerde evaluatie-instrumenten. Vakinhoudelijke kennis, van vakken die in de L2 onderwezen zijn, wordt in de geïnventariseerde studies vaak beoordeeld aan de hand van nationale testen of leerlingvolgsystemen in de L1, zodat ook hier een goede vergelijking met leerlingen in reguliere onderwijsprogramma's mogelijk is. De niet-talige vakinhoudelijke opbrengsten als gevolg van tpo zijn nog weinig onderzocht.

Onderzoeksvraag 3: Welke opbrengsten van tpo zijn gerapporteerd wat betreft de eerste schooltaal en de tweede doeltaal?

De geïnventariseerde studies laten positieve effecten van tpo op de vaardigheden in de L1 en L2 en op de verworven vakinhoudelijke kennis zien. Tpo blijkt geen nadelig effect te hebben op de ontwikkeling van de eerste schooltaal van leerlingen (L1). Ook niet als dit niet de oorspronkelijke moedertaal of thuistaal van de leerlingen is. Vaak is de L1-ontwikkeling van tpo-leerlingen gelijk aan die van leerlingen in regulier onderwijs en soms levert tpo zelfs een voordeel op in de L1-ontwikkeling. Dit voordeel is echter vaak pas aan het eind van de basisschool waarneembaar; tot die tijd laten leerlingen in tpo een iets langzamere of meestal gelijke L1-ontwikkeling zien als leerlingen in reguliere onderwijsprogramma's. Wat de L2 betreft laten leerlingen in tpo hogere taalvaardigheidsscores zien in de deelvaardigheden lezen, schrijven, spreken en luisteren dan leerlingen in reguliere onderwijsprogramma's waarin de L2 alleen als vak wordt onderwezen. Deze verschillen zijn, afhankelijk van de testleeftijd, niet altijd direct met de start van tpo meetbaar; naarmate het tpo-programma vordert, nemen de verschillen in L2-vaardigheden toe tussen tpo-leerlingen en leerlingen in regulier onderwijs en worden deze verschillen richting het eind van de basisschoolperiode steeds significanter. De vakinhoudelijke kennisontwikkeling in tpo is in de geïnventariseerde studies alleen gemeten aan de hand van het vak rekenen, waarbij deze studies geen negatieve effecten van tpo op de rekenvaardigheden hebben gevonden en soms levert tpo zelfs een voordeel op in de ontwikkeling van rekenvaardigheden. Daarnaast is het belangrijk om de invloed van diverse externe factoren zoals bijvoorbeeld taalachtergrond, buitenschools taalcontact, motivatie, cognitieve vaardigheden en SES mee te nemen in de evaluatie van de opbrengsten van tpo, omdat deze aspecten van invloed kunnen zijn op de uiteindelijke resultaten. Studies die de mogelijke invloed van deze factoren hebben gecontroleerd, vonden nog steeds betere taalvaardigheidsscores voor leerlingen in tpo ten opzichte van leerlingen in reguliere onderwijsprogramma's. Enkele studies laten een cognitief voordeel zien voor tpo-leerlingen ten opzichte van leerlingen in regulier onderwijs; naast het meten van talige en vakinhoudelijke kennis is dit een ander belangrijk aspect dat een ontwikkeling laat zien als gevolg van tpo.

Onderzoeksvraag 4: Wat zijn de relevante kenmerken en talige en didactische vaardigheden van tpo-leerkrachten?

Uit een inventarisatie van de relevante kenmerken en didactische vaardigheden van tpo-leerkrachten blijkt dat er nog geen ideaalprofiel voor een tpo-leerkracht bestaat. Ook is er weinig bekend over de mogelijkheden tot het volgen van tpo-trainingen en –cursussen. Dit heeft tot gevolg dat er een grote variëteit aan leerkrachtprofielen binnen de diverse tpo-studies bestaat. Vaak bestaan er wel minimale vereisten voor het lesgeven in tpo, maar dit blijkt in de praktijk vaak af te wijken van wat gewenst is. Het aanstellen van moedertaalsprekers wordt in veel studies genoemd als een relevante toevoeging aan het tpo-programma, maar dit blijkt vaak niet haalbaar. Echter, goede algemene didactische vaardigheden, in combinatie met voldoende L2-kennis, blijken van groter belang dan het aanstellen van moedertaalsprekers. Dit benadrukt ook direct het belang van goede training voor tpo-leerkrachten, zowel voorafgaand als gedurende het programma. Op die manier kan een leerkracht zijn talige, didactische en organisatorische vaardigheden verbeteren, wat ten goede komt aan de ontwikkeling van het tpo-programma en daaropvolgend de ontwikkeling van de vaardigheden van leerlingen in tpo. Over het gebruik van lesmaterialen wordt in de literatuurstudies weinig gezegd, maar aan de hand van expertinterviews blijkt dat het gebruik van authentieke materialen in combinatie met zelfontwikkelde materialen het meest geschikt lijkt.

Onderzoeksvraag 5: Welke attitudes zijn er t.a.v. het onderwijsprogramma, het Engels en de eerste schooltaal bij leerlingen, leerkrachten en ouders?

Evaluatie van de houding van zowel leerlingen, leerkrachten en ouders ten opzichte van het nieuwe onderwijsprogramma is van groot belang voor het slagen en de verdere ontwikkeling van tpo. Er dient voldoende steun en draagkracht te zijn voor de invoering van tpo en een positieve houding van diverse belanghebbenden is daarbij belangrijk. Over het algemeen kan geconcludeerd worden dat leerlingen vaak erg positief zijn over deelname aan tpo en dat ze het leren van en in een nieuwe taal vooral leuk en leerzaam vinden. Hierbij moet wel worden opgemerkt dat dit aan het begin van de basisschool vaak sterker is dan aan het eind, omdat het volgen van donderwijs via een tpo-programma dan als meer vanzelfsprekend wordt beschouwd. Leerkrachten zijn vaak enthousiast over de invoering van tpo, maar kunnen aan het begin nog wel wat terughoudend of onzeker zijn wat betreft hun eigen didactische en talige vaardigheden in deze nieuwe vorm van onderwijs. Dit verdwijnt vaak snel wanneer de positieve opbrengsten zichtbaar worden en men meer ervaring krijgt door het lesgeven en het deelnemen aan trainingen. Ook ouders zijn vaak overwegend positief over de invoering van tpo, omdat ze het belang voor de toekomst van hun kind ervan inzien. Een deel van de ouders uit aan het begin soms nog enige bezorgdheid omdat ze vrezen voor de hogere moeilijkheidsgraad van het onderwijs die gepaard zou kunnen gaan met deelname aan tpo, maar door een juiste informatievoorziening kunnen deze zorgen verminderd worden. Diverse evaluatiemethoden kunnen bijdragen aan een verbeterde informatievoorziening naar de leerlingen, leerkrachten en ouders, een verbetering van de didactische vaardigheden van leerkrachten en de ontwikkeling van het tpo-programma. Bij leerlingen kan de attitude, afhankelijk van de leeftijd, geëvalueerd worden aan de hand van leerkrachtobservaties, interviews of vragenlijsten. Ook voor ouders kan het zinvol zijn om vragenlijsten te gebruiken of informatiebijeenkomsten te organiseren. Leerkrachten zijn vaak wat sceptisch over deelname aan evaluatie-onderzoek, omdat zij door het onderwijzen in een tpo-programma al extra werklast ervaren. Echter, voor de voortgang en ontwikkeling van tpo is het erg belangrijk om ook de ervaringen van leerkrachten te evalueren, bijvoorbeeld aan de hand van vragenlijsten of discussiebijeenkomsten.

Onderzoeksvraag 6: Wat zijn mogelijkheden voor internationale samenwerking op het terrein van curriculumontwikkeling en onderzoek?

In de gevonden literatuurstudies is nauwelijks aandacht besteed aan de mogelijkheden tot internationale samenwerking. Uit de expertinterviews blijkt echter dat hier wel degelijk interesse voor

is. Voorbeelden van internationale samenwerking kunnen bijvoorbeeld plaatsvinden op het gebied van curriculum-, materiaal- en toetsontwikkeling, door het uitwisselen van ervaringen in een school- of leerkrachtennetwerk, of door het organiseren van een symposium waar onderzoekers, beleidsmakers en leerkrachten hun visie op en ervaringen met tpo kunnen uitwisselen. Naast deze samenwerking zou het voor deze tpo-pilot erg interessant zijn om een intensievere samenwerking aan te gaan met een land waar al een start gemaakt is met tpo, en waar op sociocultureel, politiek en taalkundig vlak voldoende overeenkomst is met de situatie in Nederland. Op die manier kan worden bijgedragen aan een nog betere voorbereiding, opzet en uitvoering van het tpo-onderzoek in Nederland.

Aanbevelingen en conclusie

In dit rapport is op basis van een literatuuronderzoek en expertconsultatie een antwoord gegeven op zes onderzoeksvragen met betrekking tot de opzet en uitvoering van tpo in Nederland. De literatuurstudie bestond uit een inventarisatie en samenvatting van resultaten uit andere overzichtsstudies, in combinatie met relevante resultaten uit recente studies naar tpo zowel binnen als buiten Europa. Door middel van de expertinterviews is specifieke en aanvullende informatie en praktische ervaringen verzameld uit een aantal andere landen die al meer ervaring hebben met tpo. Op basis van de gevonden resultaten zijn diverse aanbevelingen voor de tpo-pilot en het flankerende onderzoek geformuleerd. Deze aanbevelingen hebben betrekking op de inrichting van de tpo-programma's, de deelnemende scholen, de opzet van het onderzoek op leerling-, leerkracht-, en organisatieniveau, en de onderzoeksmatige uitdagingen.

Voorstudie Pilot Tweektalig Primair Onderwijs

Ellen van den Broek, Rick de Graaff, Sharon Unsworth, Vicky van der Zee

0. Inleiding

Deze inventariserende studie is uitgevoerd ter voorbereiding op het pilot-project Tweektalig Primair Onderwijs. Het doel van deze literatuur- en expertstudie is om vanuit een internationaal perspectief een bijdrage te leveren aan de opzet, uitvoering en evaluatie van de pilot tpo.

Er is veel literatuur beschikbaar over tpo in verschillende vormen in landen met een onderwijssysteem dat in meer of mindere mate vergelijkbaar is met het Nederlandse. Wij hebben ons gericht op een inventarisatie en samenvatting van de resultaten die in eerdere overzichtsstudies verschenen zijn, in combinatie met andere recente en relevante studies. Daarnaast hebben wij een aantal internationale experts geïnterviewd uit contexten die informatief zijn voor de ontwikkeling van tpo in Nederland. Wij focussen daarbij op die publicaties en ervaringen die specifiek gaan over programma's waarin 30-50% van de onderwijstijd in de vreemde taal wordt verzorgd.

0.1 Introductie

In de Nederlandse onderwijswet is vastgelegd dat alle vakken in het primair onderwijs, met uitzondering van vakken in de vreemde taal, in het Nederlands gegeven moeten worden. Bij wijze van experiment krijgen maximaal twintig basisscholen de komende jaren gelegenheid om tweektalig onderwijs in te richten, waarin 30-50% van de onderwijstijd in het Engels, Duits of Frans wordt gegeven. Deze inventarisiestudie heeft als doel om vanuit een internationaal perspectief een bijdrage te leveren aan de opzet, uitvoering en evaluatie van de tpo-pilot.

Door de grote verschillen in taalkundige, onderwijskundige en culturele diversiteit, binnen en buiten Europa, lopen de realisaties van de verschillende tpo-programma's behoorlijk uiteen. Met het in acht nemen van deze specifieke verschillen inventariseert deze studie relevante praktijken en opbrengsten voor de Nederlandse tpo-pilot. De studie beoogt bij te dragen aan een gestructureerd en gecontextualiseerd raamwerk met duidelijke doelen en verwachtingen.

0.2 Terminologie

Talen

In dit rapport wordt gebruik gemaakt van de term L1 wanneer het over de eerste/schooltaal gaat, en L2 wanneer het over een tweede/vreemde/doeltaal gaat. Met L1 bedoelen we dus de hoofdtaal zoals die in het betreffende land en op school gebruikt wordt. Dit betekent niet noodzakelijk dat dit ook voor ieder kind zijn/haar L1 is, iets dat ook in de tpo-pilot terug zal komen. Sommige leerlingen hebben namelijk een anderstalige achtergrond waarbij de eerste schooltaal als tweede taal geleerd wordt. Hierbij moet wel opgemerkt worden dat dit niet per definitie ook om leerlingen met een lagere sociaaleconomische status (SES) gaat. Wanneer hier expliciet verschil in wordt gemaakt in een bepaalde studie, zal dit worden aangegeven. Bij L2 gaat het in veel gevallen om Engels, maar ook andere talen komen voor in de onderzochte onderwijspraktijken. In enkele gevallen, bijvoorbeeld in Friesland of Baskenland, is er sprake van een derde taal als vreemde taal. Die wordt aangeduid met L3.

Tpo-onderwijs

Er zijn vele termen die gekoppeld zijn aan intensief en geïntegreerd tweedetaalonderwijs. In de internationale literatuur is (early) partial immersion een veelgebruikte term. Voor meer duidelijkheid betreffende de diverse termen volgt hier een korte omschrijving.

- Immersion: over het algemeen wordt met *immersion* bedoeld dat ten minste 50% van de instructie in de tweede taal (L2) plaatsvindt, waarbij aan het begin van het basisonderwijs vaak begonnen wordt met volledige instructie in de L2 en naar het eind van de basisschoolperiode wordt afgebouwd naar 50% (Genesee, 1987). Het doel is vaak om tweetaligheid te bereiken aan het eind van de basisschool (Seikkula-Leino, 2007).
- Partial immersion: met *partial immersion* wordt vaak bedoeld dat tussen de 30% en 50% van de instructie in de L2 plaatsvindt (Johnson & Swain, 1997), minder dan 30% is volgens Genesee (p.c.) geen *immersion* meer maar verrijkt of versterkt taalonderwijs, zoals in de onlangs afgeronde 15% project voor vroeg vreemdetaalonderwijs (vvtto) in Nederland.
- Content and Language Integrated Learning/CLIL: wordt als algemene term gebruikt om diverse vormen van tweetalig onderwijs te beschrijven, waarin de hoeveelheid L2 instructie en de uiteindelijke einddoelen voor taalvaardigheid behoorlijk kunnen verschillen (Eurydice, 2006; Seikkula-Leino, 2007). Uitgangspunt bij CLIL is dat de vreemde taal wordt gebruikt als middel om vakinhoudelijke doelen te bereiken, en dat inhoudelijke vakken als middel en context worden gebruikt voor taalontwikkeling (de Graaff, 2013).
- Two-way immersion/TWI: in *two-way immersion* worden leerlingen met een achtergrond in de eerste taal (L1) en leerlingen met een achtergrond in de L2 samengebracht, waarbij beide groepen leerlingen gelijk vertegenwoordigd zijn en de instructie 50-50 over beide talen verdeeld is (Christian, 1996). Een voorbeeld hiervan zijn van huis uit Spaanstalige en Engelstalige kinderen die samen voor de helft van de tijd les in het Spaans krijgen en voor de andere helft in het Engels.

In deze studie zullen we één term hanteren, namelijk tpo (*tweetalig primair onderwijs*). De studies in dit onderzoek zijn allemaal geselecteerd op een vroeg begin en hun hoeveelheid instructie in de L2, waardoor ze in voldoende mate overeenkomen met de beoogde tpo-focus van deze studie. Ze richten zich dus grotendeels op *partial immersion* onderwijs.

0.3 Onderzoeksvragen

In dit rapport wordt antwoord gegeven op de volgende onderzoeksvragen:

1. Wat voor voorbeelden zijn er in het buitenland die wat betreft opzet en inrichting van de programma's relevant en inzetbaar kunnen zijn voor de tpo-pilot?
2. Welke evaluatie-instrumenten bestaan er die ingezet kunnen worden voor kwalitatief en kwantitatief evaluatieonderzoek naar tpo?
3. Welke opbrengsten van tpo zijn gerapporteerd wat betreft de eerste schooltaal en de tweede doeltaal?
4. Wat zijn de relevante kenmerken en talige en didactische vaardigheden van tpo-leerkrachten?
5. Welke attitudes zijn er t.a.v. het onderwijsprogramma, het Engels en de eerste schooltaal bij leerlingen, leerkrachten en ouders?
6. Wat zijn mogelijkheden voor internationale samenwerking op het terrein van curriculumontwikkeling en onderzoek?

0.4 Methode

Om antwoord te geven op de bovenstaande onderzoeksvragen is zowel een literatuuronderzoek als een expertconsultatie uitgevoerd, waarbij we ons hebben gefocust op gepubliceerd onderzoek en recente ontwikkelingen waar de experts ons op hebben gewezen. Voor het literatuuronderzoek is gebruikt gemaakt van een aantal overzichtsstudies (Corda et al., 2012; Genesee, 1985; Maas, 2013; Nikolov & Mihaljević Djigunović, 2011) en zoektermen zoals voorgesteld in de projectaanvraag (zie sectie hieronder). Deze zijn aangevuld met enkele extra zoektermen om specifiekere resultaten te

verkrijgen. De focus tijdens het zoeken heeft vooral gelegen op zoekopdrachten die o.a. de termen (early/partial) immersion en/of CLIL en primary education bevatten en minder op resultaten van two-way immersion programma's, omdat dit gezien de status van L1 en L2 in Nederland minder relevant bleken te zijn. Resultaten met deze zoekwoorden zijn echter niet bij voorbaat al uitgesloten, omdat ze toch nog relevante informatie over bijvoorbeeld didactiek of evaluatie-instrumenten kunnen geven.

De belangrijkste gebruikte zoektermen zijn:

(Early) partial immersion, Content and Language Integrated Learning/CLIL, Content-based Instruction, immersion program, immersion education, (early) language immersion, two-way bilingual program, two-way immersion/TWI, elementary school, primary education/school, achievements, outcomes, attitude, didactics.

Aan de hand van deze zoektermen zijn de verschillende (internationale) taal- en onderwijskundige tijdschriften geraadpleegd, waarbij voor de definitieve selectie van artikelen vooral is gekeken of ze antwoord gaven op één of meerder onderzoeksvragen. Ook is via de ERIC (Education Resources Information Center) database veel relevante literatuur gevonden. Verder hebben referenties in verschillende relevante studies weer geleid naar andere interessant tijdschriften of artikelen.

De belangrijkste geraadpleegde tijdschriften zijn:

International Journal of Bilingualism, International Journal of Bilingual Education and Bilingualism, Journal of Multilingual and Multicultural Development, Bilingual Research Journal, International Multilingual Research Journal, Journal of Language, Identity and Education, International Journal of Multilingualism, Evaluation & Research in Education, The Modern Language Journal, Annual Review of Applied Linguistics, Applied Linguistics, Foreign Language Annals, International Review of Applied Linguistics in Language Teaching, Language Acquisition, Language and Education, Language Learning, Language Teaching, Language Testing, Modern Language Quarterly, Modern Language Review, Multilingual Education, Second Language Research, Studies in Second Language Acquisition, Child Development, Developmental Science, Review of Educational Research, The Elementary School Journal, The Alberta Journal of Educational Research, Teaching and Teacher Education, Journal of Immersion and Content-Based Language Education, AILA Review, Bilingualism: Language and Cognition, European Journal of Developmental Psychology, Harvard Educational Review, International Education, International CLIL Research Journal, International Journal of the Sociology of Language, Language Awareness, Language, Culture, and Curriculum, Neuroscience Research, Zeitschrift für Pädagogische Psychologie, Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie.

Uiteindelijk zijn er 29 artikelen geselecteerd die een antwoord geven op één of meerdere van de hierboven genoemde onderzoeksvragen. Drie studies zijn uitgevoerd in Azië (China: 2, Japan: 1), zeven studies in Noord-Amerika (USA: 6, Canada: 1), 16 studies in Europa (Duitsland: 5, Spanje: 4, Finland: 3, Frankrijk/Zwitserland/Estland/Zweden: 1) en drie studies in Nederland, uitgevoerd in Friesland. De onderstreepte tijdschriften hierboven hebben één of meerdere artikelen gepubliceerd die terug te vinden zijn in de matrix. De overige tijdschriften zijn geraadpleegd om andere artikelen en achtergrondinformatie met betrekking tot tpo te bestuderen.

Bij het selecteren van de artikelen is rekening gehouden met het jaartal van publicatie, waarbij geprobeerd is zo recent mogelijke artikelen te selecteren. Omdat tpo buiten Noord-Amerika relatief nieuw is, zijn uiteindelijk 24 van de 29 geselecteerde artikelen gepubliceerd na 2000. Slechts vijf artikelen zijn gepubliceerd vóór die tijd en dit zijn met name artikelen uit Noord-Amerika, waar tpo al veel langer toegepast wordt en *total immersion* in plaats van *partial immersion* gebruikelijker is. Het was daarom niet mogelijk om recentere artikelen uit deze regio te selecteren die ook qua opzet

en inrichting overeenkwamen met deze tpo-pilot. Voor een overzichtelijke weergave van de artikelen, en hoe ze een bijdrage leveren aan de beantwoording van onze onderzoeksvragen, is een matrix gecreëerd. Deze matrix is te vinden in Appendix II.

0.5 Experts

Naast het literatuuronderzoek is ook via gestructureerde Skype-interviews informatie gevraagd bij verschillende tpo-experts binnen en buiten Europa. Op basis van een shortlist van experts zoals voorgesteld in de projectaanvraag hebben we contact gelegd vijf experts (Fred Genesee, Kristin Kersten, Gun Lundberg, Jelena Mihaljević Djigunović en Carmen Muñoz), waarbij we een goede spreiding over landen en regio's wilden hebben. Enkele experts vonden zichzelf niet het meest geschikt om onze vragen te beantwoorden en hebben ons doorverwezen naar een andere expert in het desbetreffende land. De definitieve lijst met geïnterviewde experts is als volgt:

1. Fred Genesee (McGill University, Montréal, Canada)
2. Kristin Kersten (Universität Hildesheim, Duitsland)
3. Pilar Medrano (Ministerie van Onderwijs, Madrid, Spanje)
4. Marianne Nikolov (University of Pécs, Hongarije)
5. Johanna Watzinger-Tharp (University of Utah, VS)
6. Hanna Österlund (Skolverket [Nationaal agentschap voor onderwijs], Zweden)

Relevante informatie die verkregen is uit de expertinterviews is verwerkt bij de betreffende onderzoeksvragen in het rapport. Een samenvatting van de expertinterviews, aan de hand van onze onderzoeksvragen, is te vinden in de Appendix I.

De samenvattingen en de matrix zijn opgenomen in het Engels, omdat er ook vanuit de experts grote interesse bestond voor het ontvangen van een literatuuroverzicht en een samenvatting van de bevindingen uit de verschillende interviews. De experts hebben de samenvatting van hun eigen interview ingezien en goedgekeurd.

0.6 Opzet van de rapportage

In dit rapport geven we aan de hand van bovenstaande onderzoeksvragen een overzicht van bestaande studies naar tpo, didactische kenmerken van de beschreven tpo-praktijken, de mogelijke evaluatie-instrumenten om het effect van tpo op de taalvaardigheid in de L1 en de L2 te meten en de opbrengsten in de L1 en L2 uit reeds uitgevoerde tpo-studies. In hoofdstuk 1 worden gepubliceerde studies en resultaten van tpo-studies uit een aantal landen binnen en buiten Europa besproken. In hoofdstuk 2 worden de mogelijkheden van evaluatieonderzoek naar tpo besproken, met voorbeelden van evaluatie-instrumenten zoals deze ingezet zijn in andere studies. In hoofdstuk 3 worden de opbrengsten van tpo met betrekking tot de L1 en L2 geëvalueerd. In hoofdstuk 4 worden relevante didactische vaardigheden van tpo-leerkrachten besproken, evenals achtergrondinformatie met betrekking tot scholing en onderwijservaring. In hoofdstuk 5 wordt de houding van diverse belanghebbenden ten opzichte van het tpo-programma en daarbij horende evaluatie-instrumenten besproken. In hoofdstuk 6 zullen vervolgens de mogelijkheden tot internationale samenwerking worden besproken. Ten slotte worden in hoofdstuk 7 aanbevelingen gegeven voor de inrichting van de tpo-pilot en de opzet en focus van het onderzoek.

0.7 Werkwijze

Het projectteam bestaat uit twee projectleiders (de Graaff en Unsworth), een uitvoerder (Ellen van den Broek, MA), een student-assistent (Vicky van der Zee) en een klankbordgroep met een drietal experts (de Bot, Corda en Krikhaar).

Het analyseformat, het interviewformat en de literatuurinventarisatie zijn in november 2013 voorgelegd aan de klankbordgroep. Op 20 december 2013 is de concept-rapportage verzonden aan de klankbordgroep en op 10 januari besproken. De concept-rapportage is op 17 januari naar de opdrachtgever/stuurgroep verzonden en op 24 januari met de stuurgroep tpo besproken. De feedback is verwerkt in de definitieve versie van 31 januari 2014.

1. Wat voor voorbeelden zijn er in het buitenland die wat betreft opzet en inrichting van de programma's relevant en inzetbaar kunnen zijn voor de tpo-pilot?

Internationaal is al behoorlijk wat onderzoek gedaan naar tpo. In dit hoofdstuk wordt een overzicht gegeven van wetenschappelijke studies die hierover gepubliceerd zijn. De 29 onderzochte studies komen zoveel mogelijk overeen met de beoogde tpo-opzet, zodat zo veel mogelijk relevante informatie uit deze studies verzameld kan worden voor het beantwoorden van de onderzoeksvragen. In dit hoofdstuk zullen de kenmerken van deze tpo-programma's besproken worden.

1.1 Organisatie scholen en programma's

Het doel van deze studie is om na te gaan wat de ervaringen, resultaten en randvoorwaarden zijn om tussen de 30% en 50% van de onderwijstijd les te geven in een L2. Voor een goede vergelijking tussen internationale studies en de huidige tpo-pilot is het van belang dat de programma's qua opzet en uitvoering dermate overeenkomen dat ze een relevante bijdrage kunnen leveren aan de uitvoering van deze studie. Dus studies waarbij de L1 ook de meerderheidstaal in het land is en de L2 een vreemde taal, hebben de voorkeur. Een overzicht van de gevonden studies, zowel Europese als studies uitgevoerd buiten Europa, is te vinden in de matrix.

Zoals in de introductie al gezegd werd, hebben landen en basisscholen verschillende beweegredenen om voor een bepaalde taal als L2 in het basisonderwijs te kiezen. Enkele studies in bijvoorbeeld Baskenland (Lasagabaster, 2001) en Zwitserland (Serra, 2007) focussen op een minderheidstaal als eerste L2 in het basisonderwijs. Jäppinen (2005) onderzoekt zowel de invoering van een minderheidstaal als de belangrijkste L2 in het basisonderwijs. De meeste voorbeelden, uit onder andere China, Japan, de Verenigde Staten, Duitsland, Spanje en Finland (e.g., Cheng et al., 2010; Bostwick, 2001; Gebauer, Zaunbauer & Möller, 2012; Turnbull, Hart & Lapkin, 2003; Lorenzo, Casal & Moore, 2009; Merisuo-Storm, 2007), beschrijven echter de implementatie van een vreemde taal als L2. Dit is in de meeste gevallen Engels, maar ook Frans of Spaans komen voor.

Voor een goede vergelijking tussen leerlingen in een tpo-programma en een regulier onderwijsprogramma is het van belang dat beide onderwijsprogramma's op een vergelijkbare manier zijn ingericht. Van belang is bijvoorbeeld dat in beide groepen hetzelfde curriculum onderwezen wordt en dat scholen en klassen een vergelijkbare didactische aanpak en SES hebben. Daarnaast is het van belang dat leerlingen uit het tpo-programma en reguliere onderwijsprogramma gekoppeld zijn op basis van bijvoorbeeld IQ of algemene vaardigheden. De meeste studies die in dit rapport geanalyseerd worden hebben een geschikte controlegroep gevonden en ook zijn de meeste studies uitgevoerd op openbare basisscholen, wat een eventuele vergelijking met Nederlandse scholen makkelijker maakt. Ook de invloed van zelfselectie speelt een belangrijke rol, waarbij in enkele gevallen het enthousiasme voor tpo zo groot was, dat niet direct aan alle aanvragen voldaan kon worden en er alsnog een selectie gemaakt moest worden, bijvoorbeeld op volgorde van inschrijving. De Rivière (2004) concludeert wel dat naar aanleiding van een teveel aan aanmeldingen op een Belgische school de leerlingen die toegelaten zijn tot het tpo-programma uit verschillende socioeconomische en culturele milieus afkomstig zijn (De Rivière, 2004), waardoor de diversiteit van leerlingen met verschillende achtergronden voldoende gegarandeerd is.

1.2 Startleeftijd

Verschillende studies hebben de effecten van zowel beginleeftijd, als het aantal uren blootstelling aan de tweede taal op de taalvaardigheid besproken (Muñoz, 2006; Nikolov & Mihaljević Djigunović, 2011; Ojima et al. 2011). De aanname 'hoe eerder, hoe beter' wordt nog steeds veel gehoord, maar lijkt alleen te gelden in een natuurlijke tweedetaalverwervingssituatie (Muñoz, 2006; Genesee &

Lindholm-Leary, 2013). De studies beschreven in dit rapport betreffen bijna allemaal vroege tpo-programma's, dat wil zeggen dat met het onderwijzen van de L2 is begonnen op de kleuterschool of in *grade 1*, overeenkomend met groep 3 in Nederland. Verschillen in startleeftijd en L2 aanbod worden vaak bepaald door didactische kenmerken, socioculturele en politieke factoren (Genesee, 2004). De keuze voor een bepaalde startleeftijd in de studies die in dit rapport beschreven worden, is voornamelijk bepaald door de onderwijsinrichting in het betreffende land en de leeftijd waarop kinderen verplicht naar school gaan. Zowel de studies die zijn begonnen in groep 1 als in groep 3 behalen positieve resultaten. Over het startmoment van het leren lezen en schrijven wordt in de gevonden literatuur weinig gezegd, al beschrijven enkele studies wel het testen van de lees- en/of schrijfvaardigheid in zowel de L1 en L2 na slechts één of twee jaar tpo (Cheng et al., 2010; Knell et al., 2007; Thomas, Collier & Abbott, 1993; Burmeister & Pasternak, 2004; Gebauer, Zaunbauer & Möller, 2012; Zaunbauer & Möller, 2010), wat impliceert dat hiermee (vrijwel) gelijktijdig begonnen wordt. Medrano (persoonlijke communicatie (p.c.), 12 december 2013) geeft aan dat in Spanje met zowel het lezen als schrijven in de L1 en L2 op een jonge leeftijd wordt begonnen en dat dit geen negatieve gevolgen heeft voor de vaardigheden in de L1 (getest aan de hand van nationale testen) en de L2 (gebaseerd op de beoordeling van leerkrachten).

1.3 Intensiteit L2

Rekening houdend met het doel van dit onderzoek zijn de meeste studies geselecteerd op een L2 aanbod van tussen de 30% en 50%. De meeste onderzochte contexten lijken echter een voorkeur te hebben voor een 50-50 model, waarbij de L2 voor 50% van de onderwijstijd wordt gebruikt. Enkele studies uit Duitsland (Burmeister & Pasternak, 2004; Gebauer, Zaunbauer & Möller, 2012; Zaunbauer & Möller, 2010; Kuska, Zaunbauer & Möller, 2010) zitten hier boven met 70% L2 instructie en Merisuo-Storm (2007) beschrijft een studie uit Finland waar 20% van de onderwijstijd in de L2 plaatsvindt. Een hoge intensiteit van L2 instructie met een minimum van 40% (Medrano, p.c.), 50% (Genesee, p.c., 17 december 2013) of zelfs 70% (Kersten, p.c., 12 december 2013) wordt door de geïnterviewde experts aangeraden. Genesee benadrukt dat alle overige communicatie buiten de lessen om hoogstwaarschijnlijk in de L1 plaatsvindt, waardoor zelfs bij een 50-50 model de L1 alsnog de overhand krijgt en het daarom niet raadzaam is om voor minder dan 50% L2 instructie te gaan.

1.4 Onderwijsaanbod L2

Wanneer tussen de 30% en 50% van de onderwijstijd lesgegeven wordt in de L2, betekent dat vanzelfsprekend dat er één of meerdere inhoudsvakken in de L2 onderwezen zullen worden. Volgens Cummins (2009) zou de beslissing over welke vakken in de L2 onderwezen worden, op lokaal niveau gemaakt moeten worden en rekening moeten houden met bijvoorbeeld de voorkeur van ouders, de beschikbaarheid van tekstboeken, de ervaring van de docent met de L2 en de vakinhoud en tenslotte het regionale/nationale toetscurriculum. Scholen kiezen er bijvoorbeeld eerder voor om vakken die in de L1 en/of het nationale curriculum getest worden (bijv. leesvaardigheid of rekenen) ook in de L1 te onderwijzen voor een bepaalde periode of zelfs de gehele basisschoolperiode, om er zeker van te zijn dat leerlingen optimaal scoren op deze toetsonderdelen. Deze uitleg van Cummins is ook terug te zien in de variëteit aan vakken die juist wel of niet in de L2 gegeven worden. Niet alle studies beschrijven specifiek welke vakken in de L2 gegeven worden, maar benoemen slechts dat 'een deel van het curriculum in de L2 wordt onderwezen' (Campbell, Gray, Rhodes & Snow, 1985). In veel studies in zowel Japan, de Verenigde Staten en Europa komt onder andere rekenen voor in het L2 curriculum, aangevuld met bijvoorbeeld sociale studies, gym, muziek en/of aardrijkskunde in de L2. In studies uitgevoerd in Duitsland (e.g., Gebauer, Zaunbauer & Möller, 2012; Burmeister & Pasternak, 2004) werd ervoor gekozen om alle vakken, behalve de vakken Duits en Duitse cultuur, te geven in de L2. Cheng et al. (2010) en Knell et al. (2007) geven voorbeelden van studies waar rekenen in de L1 en

bijvoorbeeld gym of sociale studies in de L2 worden onderwezen. Een verklaring waarom juist wel of niet voor een bepaald vak in de L1 of L2 is gekozen, wordt niet vaak gegeven.

De meeste studies in het literatuuroverzicht benoemen verschillende vakken die in de L2 onderwezen zouden kunnen worden. Het Skolverket (2010) rapport geeft aan dat met name sociale en wereldoriëntatievakken in de L2 onderwezen worden, waarbij scholen zelf kunnen beslissen of ze een vak geheel of gedeeltelijk in de L2 onderwijzen, zolang de totale instructie in de L2 niet meer is dan 50% (Österlund, p.c.). Rekenen is verder het enige vak dat er verder uitgelicht en ook daadwerkelijk getest wordt. Jäppinen (2005) verklaart dat juist voor rekenen in de L2 is gekozen, omdat dit vak in zowel het basisonderwijs als het voortgezet onderwijs gegeven wordt. Dit maakt het mogelijk om in zowel het basisonderwijs als het voortgezet onderwijs in alle jaren een vergelijkbare controlegroep te hebben. Ook in andere studies uit het literatuuroverzicht wordt rekenen structureel in de L2 onderwezen gedurende de hele basisschoolperiode (Cheng et al., 2010; Turnbull, Hart & Lapkin, 2003; Thomas, Collier & Abbott, 1993; Kuska, Zaunbauer & Möller, 2010; Zaunbauer & Möller, 2010).

Door het grote verschil in vakaanbod in de L2 is het moeilijk om een eenduidige conclusie te trekken over welke vakken meer en minder geschikt zijn om in de L2 te worden onderwezen. Er zijn echter wel studies (Skolverket, 2010; Jäppinen, 2005) die aangeven dat sommige (abstracte) natuurkundige onderwerpen misschien minder geschikt zijn voor jonge kinderen om in een L2 geleerd te worden, omdat deze cognitief te veeleisend zouden zijn. Genesee (p.c.) raadt aan om vakken in de L2 te onderwijzen die ook relevant en bruikbaar zijn buiten de school. Rekenen is bijvoorbeeld erg abstract met veel specifieke wiskundetermen die kinderen buiten schooltijd om weinig zullen gebruiken. Sociale of natuurkundige studies hebben een grotere relevantie met de belevingswereld van leerlingen buiten school om en dit zijn daarom geschiktere vakken om in de L2 te onderwijzen, omdat leerlingen dan ook makkelijker de koppeling naar de praktijk leggen. Österlund (p.c.) voegt hieraan toe dat zowel sociale als wereldoriëntatievakken geschikt zijn, maar dat sociale vakken zoals maatschappij en geschiedenis bij voorkeur ook gedeeltelijk in de L1 worden onderwezen, omdat deze vakken vaak nationale aspecten behandelen die beter in de L1 onderwezen kunnen worden.

1.5 Externe factoren

Factoren zoals taalafstand, blootstelling aan de L2, motivatie, cognitieve vaardigheden, SES, L1 achtergrond, kwaliteit van scholen en verwachtingen van docenten zijn ook van invloed op de uiteindelijke leerprestaties van leerlingen (Cummins, 2009). Het is van belang om rekening te houden met deze factoren wanneer de taalvaardigheid van leerlingen getest wordt.

Taalafstand

Een studie van Lindgren en Muñoz (2013) naar het effect van diverse externe factoren in vvtto, uitgevoerd in het kader van het ELLiE-onderzoek (Enever, 2011), laat zien dat taalafstand de belangrijkste voorspeller is van de lees- en luistervaardigheid van 10- en 11 jarige leerlingen in vvtto-programma's in zeven Europese landen. Leerlingen uit Zweden en Nederland, met een kleine taalkundige afstand tot het Engels (beide talen zijn Germaanse talen), scoorden significant beter dan leerlingen uit bijvoorbeeld Spanje (Romaanse taal) en Kroatië (Slavische taal), met een grotere taalafstand tot het Engels. Een studie van D'Ydewalle en Van de Poel (1999) laat ook een positief effect van een kleine taalafstand zien. Nederlandstalige leerlingen van 8 tot 12 jaar oud die een Deense (kleine taalafstand, Germaans) of Franse (grote taalafstand, Romaans) film te zien kregen en daarna een woordenschattest maakten, lieten een positief effect op de woordenschat zien wanneer ze een Deense film gezien hadden. Dit effect was niet waarneembaar bij de leerlingen die een Franse film te zien kregen.

Buitenschools taalcontact

Cummins (2009) stelt dat voor de meerderheid van de leerlingen in een tpo-programma het uiteindelijke taalvaardigheidsniveau voor het grootste deel afhankelijk is van de input op school en dat blootstelling aan de L2 buiten de schoolcontext een minder grote invloed heeft. Om die reden zou een minimum van 50% van de onderwijstijd gedurende de gehele basisschoolperiode aan de L2 moeten worden besteed. De studie van Lindgren en Muñoz (2013) onderzocht ook de invloed van buitenschoolse L2 blootstelling op de lees- en luistervaardigheid van de hierboven genoemde groep leerlingen in het ELLiE-onderzoek. Zij vonden een grote variatie aan out-of-school exposure in verschillende landen in Europa en concludeerden dat het aanbod van L2 buiten de schoolcontext, na taalafstand, de belangrijkste voorspeller was voor lees- en luistervaardigheid, waarbij het kijken van films (na muziek luisteren en computerspelletjes spelen) de grootste invloed had op de lees- en luistervaardigheid in de L2. In het FLiPP-onderzoek is onder vvto-leerlingen in groep 1 en 2 echter geen invloed van out-of-school exposure gevonden (Unsworth, Persson, Prins & De Bot, in druk); de auteurs merken op dat dit aan de jonge leeftijd van de deelnemende kinderen te wijten zou kunnen zijn.

Overige factoren

De studies die in dit rapport besproken worden, noemen vaak het belang van het controleren van de verschillende externe factoren, maar hebben niet altijd de middelen om dit te onderzoeken. Enkele voorbeelden van externe factoren die wel besproken worden zijn zelfvertrouwen en motivatie (Seikkula-Leino (2007), SES (Bostwick, 2001; Kuska, Zaunbauer & Möller, 2010), cognitieve vaardigheden (Kuska, Zaunbauer & Möller, 2010; Zaunbauer & Möller, 2007), opleidingsniveau en betrokkenheid ouders (Skolverket, 2010; Zaunbauer & Möller, 2007) en L1 achtergrond (Skolverket, 2010; Zaunbauer & Möller, 2007; Bostwick, 2001).

1.6 Conclusie

In dit hoofdstuk is aan de hand van de gevonden literatuur een overzicht gegeven van 29 studies naar tpo die overeenkomen met de tpo-pilot. Organisatorisch komen de meeste scholen en programma's uit de geïnventariseerde studies overeen met de tpo-pilot. Zo wordt in nagenoeg alle studies tussen de 30% en 50% instructie in de L2 aangeboden, waarbij uit zowel de literatuur als de expertinterviews eerder 50% dan 30% wordt aangeraden. In alle studies ligt de startleeftijd van tpo tussen de vier en de zes jaar, wat overeenkomt met uiterlijk groep 3 in Nederland. Gezien het nieuwe karakter van tpo-onderwijs hebben enkele studies veel overeenkomsten met de beginsituatie van deze tpo-pilot. Hierdoor wordt niet in iedere studie getest tot en met het einde van de basisschool en is daarom niet altijd een ontwikkeling over de gehele basisschoolperiode onderzocht. Wel hebben deze studies verschillende testmomenten ingelast om voor zover mogelijk naar een ontwikkeling in de verschillende basisschooljaren te kunnen kijken. Er zijn echter ook verschillende studies waarbij leerlingen wel de gehele basisschoolperiode hebben doorlopen en waar doorlopend in ieder jaar getest is op taalvaardigheid. Voorbeelden hiervan zijn onderzoeken uit Zweden en Spanje, die beide door het ministerie werden ondersteund (Skolverket, 2010; Dobson, Pérez-Murilo & Johnstone, 2010). Ook enkele studies uit China, Canada, Zwitserland en Finland hebben testen uitgevoerd tot en met het eind van de basisschool en daarnaast is ook de vvto-studie van Muñoz (2006), uitgevoerd in Catalonië, een goed voorbeeld van een studie die inhoudelijk ingaat op de ontwikkeling over een langere periode.

Het onderwijsaanbod van vakken in de L2 verschilt sterk en is in de meeste gevallen afhankelijk van het (toets)curriculum, de leeftijd van de leerlingen en de beschikbaarheid van docenten. Uit zowel de literatuur als de expertinterviews blijkt dat in principe alle vakken geschikt zijn om in de L2 te onderwijzen, waarbij een voorkeur wordt uitgesproken voor vakken met een relevante functie buiten school, zoals bijvoorbeeld sociale studies of wereldoriënterende vakken. Naast de

inrichting van de programma's zijn ook diverse externe factoren zoals taalachtergrond, buitenschools taalaanbod, motivatie, cognitieve vaardigheden en SES belangrijke aspecten om mee te nemen in de evaluatie van tpo, omdat ook deze factoren van invloed kunnen zijn op de uiteindelijke opbrengsten.

2. Welke evaluatie-instrumenten bestaan er die ingezet kunnen worden voor kwalitatief en kwantitatief evaluatieonderzoek naar tpo?

Met een groeiende interesse in tpo, stijgt ook de behoefte aan relevante en geschikte evaluatie-instrumenten. De grootste uitdaging hierbij is het nieuwe karakter van tpo-studies, zodat vaak nog geen algemene richtlijnen en/of einddoelen voor tpo zijn opgesteld (Massler & Stotz, 2013). De studies die in dit rapport geëvalueerd worden, maken gebruik van een grote variëteit aan evaluatie-instrumenten voor zowel de L1 als de L2, waarbij zowel gestandaardiseerde als zelfontwikkelde testen worden gebruikt. Dit wordt mede bepaald door de beschikbaarheid van diverse testen, de vakken die in de L2 onderwezen zijn en de testleeftijd van de leerlingen op bepaalde testmomenten. In dit hoofdstuk worden evaluatie-instrumenten voor de L1 en L2 vaardigheid besproken, als ook evaluatie van de overige vakken en niet-taalkundige aspecten. Bij het meten van de taalvaardigheid, in zowel de L1 als de L2, wordt een onderscheid gemaakt tussen het meten van receptieve (lezen en luisteren) en productieve (schrijven en spreken) taalvaardigheden. In appendix III wordt een overzicht gegeven van relevante evaluatie-instrumenten uit de geïnventariseerde studies.

2.1 L1

Evaluatie-instrumenten die de L1 evalueren, meten vaak een of meerdere receptieve of productieve vaardigheden. De meeste studies maken gebruik van (inter)nationale of regionale testen of door de school ontwikkelde testen om de taalvaardigheid in de L1 te meten (e.g., Cheng et al., 2010; Bostwick, 2001; Turnbull, Hart & Lapkin, 2003; Kuska, Zaunbauer & Möller, 2010). In een Finse studie uitgevoerd door Seikkula-Leino (2007) wordt de leesvaardigheid in de L1 gemeten en beoordeeld aan de hand van het PIRLS-onderzoek (*Progress in International Reading Literacy Study*). Een internationaal onderzoek dat iedere vijf jaar wordt uitgevoerd in ongeveer 50 landen om de leesvaardigheid van basisschoolleerlingen in groep 6 te testen. Knell et al. (2007) gebruikt de *Peabody Picture Vocabulary Test* (Dunn & Dunn, 2007) en de Wide Range Achievement Test (WRAT-3; Wilkinson, 1993) om de taalvaardigheid in de L1 te testen. In het FLiPP-project naar vvto in Nederland is gebruik gemaakt van de Nederlandse versie van de *Peabody Picture Vocabulary Test* (PPVT-III-NL; Dunn, Dunn & Schlichting 2005) en voor zover beschikbaar van gegevens uit de Cito leerlingvolgsystemen (FLiPP, 2012; Unsworth et al., in druk).

2.2 L2

In de L2 wordt vaak hetzelfde onderscheid gemaakt tussen receptieve en productieve taalvaardigheden. Enkele studies meten, vanwege een testmoment aan het begin van het leertraject, alleen receptieve taalvaardigheden in de L2 (e.g., Merisuo-Storm, 2007), andere studies beoordelen alleen de productieve vaardigheden in de L2 (e.g., Thomas, Collier & Abbott, 1993), maar de meeste studies meten de effecten van tpo op zowel de receptieve als de productieve taalvaardigheden in de L2 (e.g., Bostwick, 2001; Cheng et al., 2010; Dobson, Pérez-Murillo & Johnstone, 2010; Knell et al., 2007). Welke vaardigheden precies getest worden is ook hier weer afhankelijk van de testleeftijd.

In enkele studies wordt gebruik gemaakt van L2 evaluatie-instrumenten die voor een specifiek land of voor leerlingen met een specifieke L1 ontwikkeld zijn. Een voorbeeld hiervan is de *EIKEN Test in Practical English Proficiency* (Eiken, n.d.), gebruikt om zowel receptieve als productieve Engelse taalvaardigheid van Japanse leerlingen te meten. Deze test is echter oorspronkelijk ontwikkeld voor leerlingen in het voortgezet onderwijs, maar met een groeiend aantal tpo-programma's wordt dit evaluatie-instrument nu ook in het basisonderwijs gebruikt. In Spaans Baskenland wordt de *Galbahe Language Tests* (Olaziregi & Sierra, 1988) gebruikt om de Baskische taalvaardigheid van Spaanstalige kinderen te testen.

Naast deze evaluatie-instrumenten die voor een specifieke context zijn ontwikkeld en dus niet direct relevant zijn voor deze TPO-pilot, worden er ook verschillende L2 evaluatie-instrumenten onafhankelijk van een specifieke L1-achtergrond gebruikt in studies die de L2 vaardigheden testen. Voorbeelden hiervan zijn de *LAS Links Assessment for English Language Learners* (CTB/McGraw-Hill, n.d.), de *MLA Cooperative Foreign Language Tests* (MLA, 1963) en de *Cambridge Young Learners English Tests* (Cambridge ESOL, 2003), die zowel de lees-, schrijf-, luister-, en spreekvaardigheid in de L2 testen. De *Woodcock Language Proficiency Battery* (Woodcock, 1991) test de lees-, schrijf-, en spreekvaardigheid en de *Peabody Picture Vocabulary Test* (PPVT-4; Dunn & Dunn, 2007) test de woordenschat in de L2. Ook in het FLiPP-project is de PPVT-4 gebruikt om de L2-woordenschat te meten; om de receptieve grammaticavaardigheden in de L2 te meten is er gebruik gemaakt van de *Test of Reception of Grammar* (TROG-2; Bishop, 2003).

2.3 Vakinhoudelijke kennis

Naast taalvaardigheden in de L1 en L2 is het in tpo ook van belang om de opgedane kennis in specifieke schoolvakken te evalueren, deze vakinhoud wordt namelijk overgebracht via de L2 en dus is het van belang te evalueren wat voor effect dit heeft op de resultaten. Studies in dit overzicht beschrijven verschillende vakken die in de L2 aangeboden worden, maar vakinhoudelijke kennis wordt alleen getest voor het vak rekenen, omdat rekenen vaak in alle schooljaren onderwezen wordt. Testen worden afgenomen in de L1 aan de hand van nationale testen (e.g., Kuska, Zaunbauer & Möller, 2010; Skolverket, 2010; Thomas, Collier & Abbott, 1993) of een door de school ontwikkelde rekentest (Cheng et al., 2010), zodat een goede vergelijking mogelijk is tussen leerlingen uit een tpo-programma en regulier onderwijs.

2.4 Overige vaardigheden

Slechts enkele studies uit het literatuuroverzicht hebben het effect van tpo op de ontwikkeling van metalinguïstische en cognitieve vaardigheden getest. De studie van Reder, Daigle en Demont (2013) heeft aan de hand van mondelinge testen kinderen in groep 3 en 4 getest op hun vaardigheden om structurele eigenschappen van taal (zoals morfologische, fonologische en syntactische kenmerken) te kunnen herkennen en verwerken. Zij vonden dat leerlingen in tpo-onderwijs significant beter scoorden dan leerlingen in regulier onderwijs op taken die syntactische of morfologische bewustzijn betroffen. Fonologische taken lieten geen verschil zien tussen de twee groepen, iets dat verklaart wordt door de fonologische overlap tussen L1 en L2 Jäppinen (2005) onderzocht met behulp van zelfontworpen maatstaven voor analoog redeneren het vermogen van leerlingen om te gaan met relaties, vergelijkingen, het overbrengen van concepten van de ene naar de andere taal en het uitvoeren van probleemoplossende taken. Deze vaardigheden werden getest aan de hand van verworven vakinhoudelijke kennis van rekenen en wereldoriëntatie en hoe dit hun denkproces beïnvloedt. In de eerste testgroep (7-9 jaar oud) werden geen leerprestatieverschillen tussen tpo-leerlingen en leerlingen in regulier onderwijs gemeten in rekenvakken. In de wereldoriëntatievakken was een initiëel voordeel te zien op het eerste en tweede testmoment voor leerlingen in regulier onderwijs, maar lieten de tpo-leerlingen een significant hoger niveau zien op het vierde testmoment. In de tweede testgroep (10-12 jaar oud) lieten de reken- en wereldoriëntatievakken nog geen significant verschillend niveau zien voor leerlingen in tpo en regulier onderwijs. Op de volgende drie testmomenten lieten leerlingen in tpo-onderwijs echter een significant hoger prestatieniveau zien dan leerlingen in regulier onderwijs. Cognitieve ontwikkeling van zowel tpo-leerlingen als leerlingen in regulier onderwijs verloopt de eerste jaren vrijwel gelijkmatig, maar oudere leerlingen in tpo laten uiteindelijk een snellere cognitieve ontwikkeling zien dan leerlingen in regulier onderwijs in zowel reken- als wereldoriëntatievakken.

Enkele recente onderzoeken laten zien hoe metalinguïstische voordelen zijn verbonden aan meertaligheid (Hilchey & Klein, 2011; Morales, Calvo & Bialystok, 2013; Poarch & Van Hell, 2012).

Deze studies zijn vanwege hun inrichting weliswaar niet direct vergelijkbaar met de tweetalige situatie in de tpo-pilot, maar zijn toch relevant omdat de mogelijkheid dat dergelijke voordelen ook te halen zouden kunnen zijn in een gestuurde context nog niet volledig onderzocht is. De studie van Morales, Calvo en Bialystok (2013) met 5- tot 7-jarige één- en tweetalige kinderen laat zien dat tweetalige kinderen significant beter scoren op taken die voornamelijk in het werkgeheugen worden uitgevoerd, waarbij het voordeel voor tweetalige kinderen het grootst is in cognitief veeleisendere taken. De testresultaten laten zien dat het niet de individuele componenten van uitvoerende functies in het brein of het werkgeheugen alleen zijn die worden beïnvloed door tweetaligheid, maar dat de gehele tweetalige ervaring een geïntegreerde set van vaardigheden beïnvloed, waarbij de efficiëntie het hoogst is bij cognitief veeleisende taken. Hilchey en Klein (2011) tonen met behulp van diverse cognitieve evaluatie-instrumenten aan dat tweetalige kinderen een uitgebreider cognitief verwerkingsvoordeel hebben ten opzichte van eentalige kinderen, een voordeel dat met name zichtbaar is in de reactietijden bij cognitief moeilijkere taken en niet-talige interferentietaken. Poarch en Van Hell (2012) onderzochten in drie groepen meertalige kinderen (L2-leerders, tweetaligen en drietaligen) en eentalige kinderen van 5 tot 8 jaar met behulp van de Simon-taak (Simon & Rudell, 1967) en de Attentional Networks Task (ANT) (Rueda et al., 2004) het vermogen om conflicten op te lossen, taken te controleren en aandacht te verschuiven, vermogens verbonden aan het executieve controlenetwerk. De Simon-taak liet een voordeel zien voor twee- en drietalige kinderen boven L2-leerders, maar ook een voordeel van L2-leerders ten opzichte van eentalige kinderen. Dit verschil was echter niet statistisch significant. De ANT liet wel een significant voordeel zien voor twee- en drietalige kinderen ten opzichte van L2-leerders. De resultaten suggereren dat L2-leerders al wel een verbeterde aandachtscontrole laten zien ten opzichte van eentalige kinderen, maar nog niet het prestatieniveau van volledige twee- of drietalige kinderen hebben bereikt, iets dat mogelijk nog kan veranderen doordat de L2-leerders zich nog in het beginstadium van de L2-verwerving bevonden (Poarch & Van Hell, 2012).

2.5 Conclusie

Na inventarisatie van de literatuurstudies blijkt dat er een grote variëteit aan evaluatie-instrumenten bestaat die allemaal ingezet zouden kunnen worden voor kwalitatief en kwantitatief evaluatieonderzoek naar tpo. Veel studies die beschreven worden in dit rapport maken gebruik van zowel gestandaardiseerde als zelfontwikkelde testen om de L1 en L2 vaardigheden en ook vakinhoudelijke of cognitieve vaardigheden te testen. Aan beide soorten testen zitten voor- en nadelen. Aan de ene kant is een continue en systematische evaluatie met gestandaardiseerde instrumenten, waarbij de L1, L2 en de vakinhoud worden beoordeeld, essentieel wanneer leerlingen en scholen op nationaal niveau met elkaar worden vergeleken (Massler & Stotz, 2013). Aan de andere kant kan met zelfontwikkelde testen de onderwezen vakinhoud en L2-kennis beter afgestemd worden op het curriculum, iets dat erg belangrijk kan zijn in verdere ontwikkeling van het curriculum en de einddoelen van het L2-onderwijs (Genesee, p.c.). Het gebruik van zelfontwikkelde testen maakt het vergelijk van resultaten op nationaal niveau echter wel moeilijk, maar het kan als goede aanvulling dienen naast de genormeerde evaluatie-instrumenten.

In dit hoofdstuk zijn diverse evaluatie-instrumenten voor de L1 en L2 besproken, evenals evaluatiemogelijkheden voor de overige vakken en niet-taalkundige vaardigheden. Bij het meten van de L1 en L2 vaardigheid wordt een onderscheid gemaakt tussen het meten van receptieve (lezen en luisteren) en productieve (schrijven en spreken) taalvaardigheden. Het gebruik van bepaalde testen wordt mede bepaald door de toepasbaarheid in een specifieke context, bijvoorbeeld de testleeftijd van leerlingen op een bepaald testmoment. Voorbeelden van deze evaluatie-instrumenten zijn het PIRLS onderzoek voor leesvaardigheid, de PPVT voor woordenschat en de TROG voor grammatica. Naast gestandaardiseerde testen worden zelfontwikkelde testen gebruikt om meer specifiekere aspecten van de taalvaardigheid in de L1 en L2 te meten.

Naast de taalvaardigheid is het ook van belang om de vakinhoudelijke kennis van schoolvakken die in de L2 worden onderwezen te evalueren om na te gaan wat voor effect tpo heeft op de verworven vakkennis. De meeste studies in dit rapport hebben een evaluatie van het vak rekenen gedaan, waarbij in de meeste gevallen nationale of door de school ontwikkelde testen in de L1 zijn gebruikt, zodat een goede vergelijking met leerlingen in reguliere onderwijsprogramma's kon worden gemaakt. Ook hebben enkele studies het effect van tpo op de ontwikkeling van metalinguïstische en cognitieve vaardigheden getest, waarbij leerlingen in tpo uiteindelijk een snellere cognitieve ontwikkeling laten zien dan leerlingen in regulier onderwijs en waarbij meertalige leerlingen een cognitief voordeel ontwikkelen ten opzichte van eentalige leerlingen.

3. Welke opbrengsten van tpo zijn gerapporteerd wat betreft de eerste schooltaal en de tweede doeltaal?

In het vorige hoofdstuk werden mogelijke evaluatie-instrumenten van tpo besproken. Met behulp van de overzichtsstudies (Corda et al., 2012; Maas, 2013; Nikolov & Mihaljević Djigunović, 2011) hebben we 29 studies gevonden die qua opzet en inrichting overeenkomen met de tpo-pilot. In dit hoofdstuk zullen we de effecten van tpo op de L1 en L2 vaardigheid bespreken. Ook het effect van tpo op andere vakken, zoals rekenen, wordt geïnventariseerd, omdat tpo niet alleen een effect heeft op de uiteindelijke taalvaardigheden, maar ook op de vakinhoudelijke kennis van vakken die in de L2 onderwezen worden.

3.1 L1

Onderzoek van Baker (2001) en Genesee (1987, 2004) laat zien dat leerlingen in tpo-onderwijs tijdelijk een vertraagde ontwikkeling kunnen laten zien in lezen, schrijven en spelling in vergelijking met leerlingen in regulier onderwijs, maar dit onderzoek toont ook aan dat leerlingen in tpo-onderwijs na enkele jaren weer hetzelfde taalvaardigheidsniveau laten zien als leerlingen in een regulier onderwijsprogramma. Genesee (1987, 2004) beschrijft dit effect alleen bij leerlingen in volledige immersieprogramma's, waarbij instructie in de eerste jaren volledig in de L2 is. Dit verschil in L1 ontwikkeling is na enkele jaren verdwenen wanneer ook de instructie in de L1 wordt ingevoerd. Baker (2001) vond dit effect ook bij leerlingen in tpo-programma's waarbij de helft van het onderwijs in de L2 gegeven wordt. Ook bij deze groep is het verschil in L1 ontwikkeling tussen leerlingen in tpo-onderwijs en regulier onderwijs na enkele jaren volledig verdwenen en laten beiden groepen dezelfde scores in lees-, schrijf-, en spellingsvaardigheid zien aan het eind van de basisschool.

Ook enkele studies die in dit rapport worden besproken vonden bewijs voor deze tijdelijke vertraging. Wode (2004) concludeerde dat de vaardigheden van tpo-leerlingen in groep 3 zich niet zo snel ontwikkelden als van leerlingen in regulier onderwijs, maar dat dit verschil aan het eind van groep 4 al verdwenen was. Turnbull, Hart en Lapkin (2003) leggen het grote taalvaardigheidsverschil tussen beide groepen leerlingen in groep 8 ten opzichte van groep 5 ook uit aan de hand van deze tijdelijke vertraging in de L1 ontwikkeling. De studie van Turnbull, Hart en Lapkin (2003) toont aan dat de verschillen tussen tpo-leerlingen en leerlingen in regulier onderwijs groter wordt naar mate het onderwijs in een tpo-programma vordert. In een vergelijking van lees- en schrijfvaardigheidsscores van beide groepen vonden zij op het eerste testmoment in groep 5 nog geen grote verschillen, maar op het tweede testmoment in groep 8 (na 6 jaar tpo) wel aanzienlijk betere lees- en schrijfvaardigheidsscores voor leerlingen in tpo-programma's. Het verschil wordt door het cross-sectionele karakter van de studie echter uitgedrukt in de behaalde niveaus en niet statistisch onderbouwd.

Cheng et al. (2010) rapporteerden ook hogere taalvaardigheidsscores (Chinese karakterherkenning, schrijfvaardigheid, leesvaardigheid) voor tpo-leerlingen ten opzichte van leerlingen in regulier onderwijs, waarbij de scores niet-significant verschillen op het eerste en tweede testmoment in groep 4 en 6, maar wel significant op het derde testmoment in groep, $t(180.49) = 6.61$, $p < 0.001$). Ook Dobson, Pérez-Murillo en Johnstone (2010) vonden duidelijk hogere schrijfvaardigheidsscores aan het eind van de basisschool voor leerlingen in tpo-onderwijs vergeleken met leerlingen in reguliere onderwijsprogramma's, maar statistische gegevens ontbreken omdat de studie niet was ingericht om te rapporteren over de verschillen tussen leerlingen maar slechts ter promotie van tpo. Merisuo-Storm (2007) vond hogere schrijfvaardigheidsscores voor leerlingen in tpo, maar deze resultaten verschilden niet significant met die van leerlingen in regulier onderwijs.

De studie van Thomas, Collier en Abbott (1993) rapporteerden al na twee jaar statistisch significante verschillen in leesvaardigheidsscores in groep 4 in het voordeel van leerlingen in tpo, $F(1, 725) = 9.98$, $p < .0016$, $\eta^2 = .24$). Merisuo-Storm (2007) vond al na één jaar significant hogere

leesvaardigheidsscores voor leerlingen in tpo-programma's, maar heeft deze effectgroottes niet gerapporteerd. Het Skolverket (2010) rapport uit Zweden heeft betere L1 vaardigheidsscores op nationale testen voor leerlingen in tpo-programma's gevonden in vergelijking met het landelijk gemiddelde. Kuska, Zaunbauer en Möller (2010) concludeerden dat spellingsscores een significant effect hebben op de leerprestaties: $F(1, 393) = 47.53, p < .001$) en vonden een significant verschil in groep 5 en 6 tussen tpo-leerlingen en leerlingen in regulier onderwijs in het voordeel van tpo-leerlingen: $F(1, 400) = 6.81, p < .05$). Reder, Daigle en Demont (2013) concludeerden dat het vermogen van leerlingen om morfologische en syntactische kenmerken uit hun L1 te herkennen en verwerken al aanwezig is vanaf het eerste testmoment in groep 4, maar dat hierin pas een significant in het voordeel van tpo-leerlingen optreedt naar mate langer tpo-onderwijs gevolgd wordt.

Studies van De Rivière (2004) en Cenoz (2005) vonden geen negatieve effecten van tpo op de L1 ontwikkeling en ook Knell et al. (2007) en Seikkula-Leino (2007) vonden geen significante verschillen in leesvaardigheid tussen leerlingen in tpo en regulier onderwijs. Gebauer, Zaunbauer en Möller (2012) vonden gelijke leesvaardigheidsscores voor leerlingen in groep 3 tot en met 6 in regulier en tpo-onderwijs, maar zij concludeerden wel dat leerlingen in tpo-onderwijs een snellere stijging in leesvaardigheid lieten zien over deze vier schooljaren. Zaunbauer en Möller (2010), Serra (2007) en Lasagabaster (2000; 2001) vonden gelijke of (niet-significant) hogere lees-, spreek-, en/of schrijfvaardigheidsscores voor leerlingen in tpo in vergelijking met leerlingen in regulier onderwijs.

Seikkula-Leino (2007) toont aan dat een onderverdeling naar leerlingen met sterke en zwakke leerprestaties geen verschil laat zien in L1 vaardigheden tussen leerlingen die tpo of regulier onderwijs hebben gevolgd. Dit geeft aan dat zowel zwakkere als sterkere leerlingen, gebaseerd op hun verwachte maximumcapaciteiten, dezelfde resultaten en vooruitgang in hun L1 kunnen verwezenlijken, onafhankelijk van het onderwijsprogramma. tpo heeft dus geen nadelig effect op leerlingen met zwakke leerprestaties (Seikkula-Leino, 2007) en het kan zelfs een gunstig effect op hun leerprestaties hebben (Genesee, 2007). Ook leerlingen in tpo-programma's waarvan de nationale L1 niet hun L1 is, laten gemiddeld betere resultaten zien dan dezelfde groep leerlingen op controlescholen met een regulier onderwijsprogramma (Skolverket, 2010).

Een ander factor die van invloed kan zijn op de uiteindelijke taalvaardigheden is de taalafstand. In de meeste studies die hier beschreven worden, maken zowel de L1 als de L2 gebruik van hetzelfde schriftsysteem en komen ze soms ook nog uit dezelfde taalfamilie. Er zijn echter ook een aantal studies waarin zowel het schriftsysteem als de taalfamilie van de L1 en L2 verschillend zijn (Knell et al., 2007; Bostwick, 2001; Mehisto & Asser, 2007; Thomas, Collier & Abbott, 1993). Hier wordt geen negatief effect gevonden van het tpo-programma op de L1 vaardigheden. Bij leerlingen met Engels als L1 die Spaans, Frans of Japans als L2 leerden, lieten de leerlingen die Japans leerden, en dus te maken hadden met een ander schriftsysteem en een andere taalgroep, zelfs de grootste vooruitgang zien in de gemeten L1 vaardigheden (Thomas, Collier & Abbott, 1993).

3.2 L2

Ook voor de gemeten L2 vaardigheid geldt dat afhankelijk van de studie, testleeftijd en het aantal jaren instructie via een tpo-programma de receptieve, productieve of beide taalvaardigheden getest kunnen worden. De studies in dit rapport laten veelal positieve effecten zien van tpo op de vier basisvaardigheden lezen, schrijven, spreken en luisteren zien (e.g., Cheng et al., 2010; Knell et al., 2007; Lorenzo, Casal & Moore, 2009; Lasagabaster, 2000; Lorenzo, Casal & Moore, 2009; Skolverket, 2010). Significant hogere taalvaardigheidsscores voor leerlingen in tpo ten opzichte van leerlingen in regulier onderwijs zijn soms al na één jaar tpo zichtbaar in receptieve vaardigheden zoals woordherkenning of woordenschat (Cheng et al., 2010; Knell et al., 2007). Daarnaast worden ook andere taalvaardigheden genoemd waarop tpo-leerlingen significant beter scoren dan leerlingen in reguliere onderwijsprogramma's. Voorbeelden hiervan zijn leesvaardigheid (Serra, 2007; Lasagabaster, 2000), schrijfvaardigheid (Dobson, Pérez-Murillo & Johnstone, 2010), luistervaardigheid (Campbell, Gray, Rhodes & Snow, 1985; Serra, 2007) en spreekvaardigheid (Cheng et al., 2010; Knell

et al., 2007; Burmeister & Pasternak, 2004; Campbell, Gray, Rhodes & Snow, 1985; Serra, 2007), deze studies hebben echter niet altijd een effectgrootte gerapporteerd.

Cheng et al. (2010) vonden significant hogere taalvaardigheidsscores voor tpo-leerlingen in vergelijking met leerlingen in reguliere onderwijsprogramma's in de deelgebieden lees-, schrijf-, luister-, en spreekvaardigheid in groep 4, $t(375.60) = 4.97, p < 0.001$, groep 6, $t(382.05) = 4.99, p < 0.001$, en groep 8, $t(86.44) = 13.69, p < 0.001$. Knell et al. (2007) rapporteerden significant hogere vaardigheidsscores in groep 3, 4, en 5 voor tpo-leerlingen op mondelinge taalvaardigheid, $F(1,181) = 16.847, p < .001$, woordidentificatie, $F(1,181) = 8.210, p < .001$, en woordenschat, $F(1,181) = 37.295, p < .001$, maar zij vonden geen significante verschillen in fonologisch bewustzijn tussen tpo-leerlingen en leerlingen in regulier onderwijs. Campbell, Gray, Rhodes en Snow (1985) rapporteerden significant hogere luister- en spreekvaardigheidsscores voor tpo-leerlingen in vergelijking met leerlingen in regulier onderwijs in groep 6, maar geen significant hogere lees- en schrijfvaardigheidsscores en Lasagabaster (2000) rapporteert significant hogere leesvaardigheidsscores voor tpo-leerlingen ten opzichte van leerlingen in regulier onderwijs in groep 7, $t = 2.91, p < 0.05$.

Leerlingen laten vaak een geleidelijke stijging zien gedurende de basisschoolperiode op de geteste L2-vaardigheden (Bostwick, 2001; Thomas, Collier & Abbott, 1993), waarbij de stijging aan het eind van de basisschoolperiode sterker doorzet en leidt tot toenemende verschillen in taalvaardigheid tussen leerlingen in tpo-programma's en leerlingen in reguliere onderwijsprogramma's (Lasagabaster, 2000; Riemersma & De Jong, 2007; Van Ruijven & Ytsma, 2008). Ook is er een nauwe samenhang tussen de behaalde L1 en L2 resultaten, waarbij leerlingen met de hoogste L1 scores ook de hoogste scores in (receptieve) L2 vaardigheden behalen (Merisuo-Storm, 2007).

3.3 Vakinhoudelijke kennis

Naast de opbrengsten in de L1 en L2 rapporteren vele studies ook over de resultaten in andere vakken die in de L2 onderwezen werden. Vaak is in de onderzoeken gekozen voor rekenen als één van de vakken in de L2, omdat dit vak altijd in het basisschoolcurriculum voorkomt en daarom een vergelijking in rekenresultaten tussen leerlingen in tpo-programma's en reguliere onderwijsprogramma's goed mogelijk is. De resultaten laten zien dat tpo geen negatief effect heeft op de rekenresultaten (Turnbull, Hart & Lapkin, 2003) of laten zelfs significant betere rekenresultaten zien voor leerlingen in tpo-programma's (Cheng et al., 2010; Serra, 2007; Kuska, Zaunbauer, Möller, 2010; Skolverket, 2010; Thomas, Collier & Abbott, 1993). Cheng et al. (2010) vonden nog geen significante verschillen tussen tpo-leerlingen en leerlingen in regulier onderwijs in groep 4 en 6, maar wel significante verschillen in de rekenresultaten in groep 8 in het voordeel van tpo-leerlingen: $t(138.35) = 4.10, p < .001$. Thomas, Collier en Abbott (1993) vonden na twee jaar een significant verschil in rekenresultaten tussen tpo-leerlingen en leerlingen in regulier onderwijs in het voordeel van tpo-leerlingen: $F(1,613) = 24.45, p < .001, d = .42$. Rekenen is in deze studies volledig of gedeeltelijk in de L2 onderwezen en door middel van eigen schooltesten of nationale testen in de L1 getest. Vaak zijn de rekenresultaten aan het begin van de basisschoolperiode nog gelijk aan of niet-significant verschillend van leerlingen in reguliere onderwijsprogramma's, maar behalen tpo-leerlingen aan het eind van de basisschool significant betere rekenresultaten dan leerlingen in reguliere onderwijsprogramma's (Turnbull, Hart & Lapkin, 2003,) en laten tpo-leerlingen een snellere stijging zien in hun rekenresultaten dan leerlingen in reguliere onderwijsprogramma's (Zaunbauer & Möller, 2010).

In tegenstelling tot in de hierboven genoemde studies laat de studie van De Courcy en Burston (2000) als enige geen betere rekenresultaten zien voor leerlingen in tpo-programma's. In groep 5 wordt door leerlingen in tpo en regulier onderwijs nog gelijk gescoord, maar in groep 7 laten leerlingen in regulier onderwijs betere testresultaten rekenen zien dan leerlingen in tpo-programma's ($p < 0.05$). Hierbij moet echter worden opgemerkt dat leerlingen in reguliere programma's de rekentest in de L1 hebben gemaakt, terwijl tpo-leerlingen dezelfde test in de L2 maakten, waardoor dit testdesign niet zonder meer vergelijkbaar is met de andere studies die rekenresultaten hebben

gemeten. De cognitieve uitdaging van het maken van een rekentest in de L2 wordt als verklaring gegeven voor deze resultaten en in de discussie wordt aangegeven dat wanneer tpo-leerlingen de test ook in de L1 gemaakt hadden of meer tijd kregen, ze waarschijnlijk minimaal dezelfde resultaten hadden behaald als leerlingen uit reguliere onderwijsprogramma's. Het feit dat de tpo-leerlingen verschillende wiskundige concepten wel begrepen hebben, laat zien dat ze de kennis die via de L2 is verworven wel kunnen overbrengen naar hun L1.

Waar in veel landen op basis van het curriculum gekozen wordt voor rekenen in de L2, wordt in Zweden juist gedacht dat het onderwijzen van rekenen in de L1 zou moeten plaatsvinden, omdat het vanwege abstracte wiskundebegrippen te moeilijk zou kunnen zijn in de L2 en omdat het voor leerlingen makkelijker zou zijn om de nationale einddoelen te halen wanneer het vak ook in de L1 onderwezen wordt (Skolverket, 2010). Sociale vakken en natuuroriënterende vakken zoals aardrijkskunde en biologie krijgen daar de voorkeur als het gaat om instructie in de L2. Leerlingen in tpo-programma's behalen gemiddeld ook hogere scores in deze vakken in nationale testen dan leerlingen in reguliere onderwijsprogramma's (Skolverket, 2010).

3.4 Externe factoren

Verschillende studies die in dit rapport worden besproken, hebben de invloed van diverse externe factoren op de taalvaardigheid van de leerlingen geanalyseerd. Uit het ELLiE-onderzoek (Enever, 2011) en de daarop gebaseerde studie van Lindgren en Muñoz (2013) blijkt dat zowel taalafstand als out-of-school exposure van grote invloed kunnen zijn op de uiteindelijke taalvaardigheden van vto-leerlingen, waarbij een kleine taalafstand en groot L2-aanbod buiten school de L2-verwerving kan faciliteren. Aan de andere kant toont de studie van Bostwick (2001) aan dat een grote taalafstand (L1 Japans, L2 Engels) geen nadelige effecten laat zien op de ontwikkeling van de L1.

Anderstalige achtergrond

Naast het meten van de L1 en L2 vaardigheden dient er ook rekening gehouden te worden met leerlingen met een anderstalige achtergrond. Dit kan gaan om tweede-taalleerders van de eerste schooltaal (zoals NT2-leerders in Nederland), maar ook om leerlingen die naast de L1 nog een andere thuistaal hebben, zoals bijvoorbeeld in Friesland het geval kan zijn. Uit de expertinterviews kwam duidelijk naar voren dat er geen enkele reden is om deze kinderen uit te sluiten van tpo, en dat onderwijs in een L3 zelfs kan leiden tot betere prestaties in hun moedertaal (Genesee, p.c.; Nikolov, p.c.; Watzinger-Tharp, p.c.). Het Skolverket (2010) rapport gaat uitgebreid in op de invloed van L1 achtergrond en maakt een onderscheid tussen leerlingen met een anderstalige achtergrond geboren in Zweden of in het buitenland, afkomstig uit landen met een groot aandeel migranten of uit landen waar juist al Engels gesproken wordt, en het opleidingsniveau van de ouders van deze anderstalige leerlingen. Het aantal leerlingen in dat onderzoek met een anderstalige achtergrond is groter in tpo-onderwijs (36%) dan het landelijk gemiddelde (17%), waarbij het opleidingsniveau van de ouders een belangrijke rol speelde in de keuze voor het type onderwijs. Wanneer deze factoren statistisch gecontroleerd worden, blijkt dat anderstalige leerlingen in tpo betere resultaten te behalen op nationale testen in de L1, L2 en rekenen dan het landelijk gemiddelde van anderstalige leerlingen. Ook is het prestatieverschil tussen leerlingen met een anderstalige achtergrond en Zweedse achtergrond kleiner in tpo-scholen dan in reguliere scholen (Skolverket, 2010).

Opleidingsniveau ouders

Wanneer rekening wordt gehouden met het opleidingsniveau van de ouders, en de mogelijke invloed van deze factor op de uiteindelijke resultaten wordt uitgeschakeld, behalen leerlingen in tpo-programma's gemiddeld alsnog betere resultaten op nationale testen in L1, L2 en rekenen dan het landelijk gemiddelde (Skolverket, 2010). De studie van Kuska, Zaunbauer en Möller (2010) heeft deze factor niet als covariaat meegenomen, maar liet wel een significant positieve correlatie zien tussen

opleidingsniveau van de ouders en L1 vaardigheden bij zowel tpo-leerlingen als leerlingen in regulier onderwijs. L2-vaardigheden correleerden significant positief met het opleidingsniveau van de ouders bij tpo-leerlingen en (niet-significant) positief bij leerlingen in regulier onderwijs.

SES

De studie van Dobson, Pérez-Murillo en Johnstone (2010) concludeert dat tpo-onderwijs voordelen meebrengt voor leerlingen met lagere socioeconomische achtergrond. Ook Genesee (2004) beschrijft dat leerlingen met een onvoordelige thuissituatie en lage SES kunnen profiteren van tpo-onderwijs. Zijn studie laat zien dat socioeconomisch benadeelde leerlingen in tpo-onderwijs dezelfde ontwikkeling in L1 en rekenen laten zien als leerlingen in regulier onderwijs. Daarnaast presteren leerlingen uit een lagere socioeconomische groep over het algemeen beter op alle L2 vaardigheden in tpo-onderwijs dan in regulier onderwijs (Genesee, 2004). Kuska, Zaunbauer en Möller (2010) vonden bij zowel leerlingen in tpo als regulier onderwijs geen significante correlatie tussen SES en prestaties in de L2. Wel werd er een significante positieve relatie gevonden tussen L1 prestatie en SES bij tpo-leerlingen.

Cognitieve vaardigheden

Kuska, Zaunbauer en Möller (2010) tonen aan dat intelligentie positief correleert met L1 en L2 vaardigheden bij zowel tpo-leerlingen als leerlingen in regulier onderwijs. Een studie van Zaunbauer en Möller (2007) laat zien dat leerlingen in tpo-onderwijs significant hogere scores op testen van nonverbale intelligentie dan leerlingen in regulier onderwijs. Seikkula-Leino (2007) geeft aan dat gebaseerd op vooraf geteste cognitieve vaardigheden leerlingen in tpo-programma's minder vaak boven hun verwachte leerprestaties uit lijken te stijgen dan leerlingen in reguliere onderwijsprogramma's, omdat de gelijktijdige verwerving van de L2 extra cognitieve vereisten met zich meebrengt.

3.5 Conclusie

In dit hoofdstuk zijn de effecten van tpo op de L1 en L2 vaardigheden en de vakinhoudelijke kennis besproken. Uit bovenstaande onderzoeken, en ook uit de overzichtsstudies van Genesee, Lindholm-Leary, Saunders en Christian (2005) en Corda et al. (2012), blijkt dat tpo geen nadelig effect heeft op de L1 ontwikkeling van leerlingen. Vaak is de L1 ontwikkeling gelijk voor tpo-leerlingen en leerlingen in regulier onderwijs en soms levert het volgen van tpo-onderwijs zelfs een voordeel op in de L1 ontwikkeling. Dit voordeel is vaak echter pas aan het eind van de basisschool waarneembaar, tot die tijd vertonen tpo-leerlingen soms een langzamere of vaak een gelijke ontwikkeling in de L1 als leerlingen in regulier basisonderwijs.

Wat de L2 betreft laten tpo-leerlingen, afhankelijk van het testmoment, hogere taalvaardigheidsscores in de deelvaardigheden luisteren, spreken, lezen en/of schrijven zien dan leerlingen in regulier onderwijs. Naar mate het tpo-programma vordert, laten tpo-leerlingen ook een geleidelijke stijging zien in de ontwikkeling van hun L2-vaardigheden, waarbij de stijging aan het eind van de basisschool sterker doorzet en de verschillen in L2-vaardigheid tussen tpo-leerlingen en leerlingen in regulier onderwijs steeds signifikanter worden.

Vakinhoudelijke kennis is in bovenstaande studies vaak alleen gemeten aan de hand van het vak rekenen, waarbij de studies die de invloed van tpo op rekenen hebben onderzocht geen negatieve effecten van de L2 op de rekentaalvaardigheid hebben gevonden. Dit geldt zowel voor studies waarbij rekenen geheel in de L2 als gedeeltelijk in de L2 werd onderwezen en ondanks uiteenlopende meningen over het onderwijzen van rekenen in de L2 lijkt er wat behaalde resultaten betreft geen reden om dit niet te doen.

De invloed van diverse externe factoren op de verworven taalvaardigheden in tpo is in verschillende studies meegenomen in de analyses. Een anderstalige achtergrond, opleidingsniveau

van de ouders, SES en cognitieve vaardigheden zijn hierbij het vaakst onderzocht en in enkele studies zijn deze factoren ook als covariaat meegenomen om enig effect hiervan op de uiteindelijke taalvaardigheid uit te sluiten. Studies waarin dit gedaan is, laten onafhankelijk van diverse externe factoren betere prestaties zien voor leerlingen in tpo in vergelijking met leerlingen in reguliere onderwijsprogramma's.

4. Wat zijn de relevante kenmerken en talige en didactische vaardigheden van tpo-leerkrachten?

De rol van leerkrachten in het onderwijs is ontzettend belangrijk. In de geïnventariseerde studies in dit rapport is nagegaan hoe groot de invloed is van taalvaardigheden en didactische vaardigheden van leerkrachten op de uiteindelijke verworven (taal)vaardigheden bij leerlingen. Naast vakinhoudelijke kennis en L2 kennis zijn verschillende factoren geïnventariseerd, zoals het vermogen om leerlingen te motiveren en stimuleren in hun leerproces, problemen te herkennen en op te lossen, en deelname aan cursussen en trainingen, die van invloed kunnen zijn op de uiteindelijke leerprestaties. De implementatie van tpo heeft ongetwijfeld ook gevolgen voor de talige, didactische en organisatorische vaardigheden van de leerkracht voor het L1-gedeelte van het onderwijsprogramma. Dit is echter een onderwerp dat niet besproken wordt in de geïnventariseerde studies en de focus in dit hoofdstuk zal daarom op de vaardigheden van de L2-leerkracht liggen.

4.1 Profiel leerkracht

Gezien het nieuwe karakter van tpo in Nederland is er nog nauwelijks onderzoek gedaan naar de mogelijkheden tot tpo-trainingen en –cursussen en wat het ideale profiel van een tpo-leerkracht zou zijn. Het Europees Platform (2009) heeft daarentegen wel een ideaalprofiel voor een vvto-leerkracht opgesteld, waarbij een onderscheid gemaakt wordt tussen de taalvaardigheid, pedagogische en didactische kennis, en kennis van het vvto-onderwijs. Als minimaal beheersingsniveau in vvto moet een leerkracht kennis van het Engels hebben op minimaal het B2-niveau van het CEFR voor de vaardigheden spreken, luisteren en lezen, en minimaal B1 voor de vaardigheid schrijven (wanneer aan kinderen jonger dan 8 wordt lesgegeven). Ook in het ELLiE-onderzoek wordt het B2-niveau genoemd als minimaal taalvaardigheidsniveau, waarbij nog wel wordt opgemerkt dat C1 wenselijk zou zijn voor een vvto-leerkracht om ook volledig functioneel te zijn in de informele en incidentele taalsituaties die voorkomen tijdens de lessen (Enever, 2011). Daarnaast worden kennis van schoolspecifieke taaluitdrukkingen, het geven van juiste feedback, het toegankelijk aanbieden van de L2, het omvormen van taaldoelen in praktische lessen, het bijdragen aan nieuwe lesmaterialen en voldoende kennis van taalverwervings- en taalontwikkelingsprocessen genoemd in het ideaalprofiel van een vvto-leerkracht (Europees Platform, 2009). Het ELLiE-onderzoek laat echter zien dat er nog veel variatie bestaat in Nederland tussen leerkrachten wat betreft hun taalvaardigheidsniveau en didactische vaardigheden (Enever, 2011). Uit het ELLiE-onderzoek blijkt verder dat er in de zeven bestudeerde landen geen vast onderwijstraject bestaat om leerkrachten voor te bereiden op vvto-onderwijs en te garanderen dat zij zowel voldoende L2 als didactische vaardigheden bezitten. Daarentegen lag er in deze landen wel een grote nadruk op het aanbieden van extra cursussen ter bevordering van de taalvaardigheid en het geven van vreemdetaalonderwijs (Enever, 2011).

4.2 Taalachtergrond

Eén van de kwesties die steeds weer terugkomt bij het opstarten van tpo-programma's is de keuze tussen het aanstellen van moedertaalsprekers of docenten met een (ver)gevorderde L2 kennis. Veel studies spreken een voorkeur uit voor moedertaalsprekers, maar zeggen hier ook bij dat dit vaak niet haalbaar is. In de meeste gevallen wordt er gekozen voor L1 moedertaalsprekers met een goede kennis van de L2 (Cheng et al., 2010; Knell et al., 2007; Wode, 2004; Burmeister & Pasternak, 2004; Kuska, Zaunbauer & Möller, 2010; Zaunbauer & Möller, 2010; Gebauer, Zaunbauer & Möller, 2012; Thomas, Collier & Abbott, 1993; Reder, Daigle & Demont, 2013; Cenoz, 2005; Skolverket, 2010), waarbij de vereisten voor L2 vaardigheden vaak behoorlijk verschillen (zie paragraaf scholing en didactische vaardigheden). Een combinatie van moedertaalsprekers en docenten met (ver)gevorderde L2 vaardigheden (Cirino et al., 2007; Serra, 2007; Lasagabaster, 2000) of soms zelfs een combinatie

van moedertaalsprekers, moedertaalsprekers met goede L2 vaardigheden en taalassistenten met de L2 als moedertaal (Lorenzo, Casal & Moore, 2009; Dobson, Pérez-Murillo & Johnstone, 2010) behoren ook tot de mogelijkheden, waarbij de verschillende typen leerkrachten elkaar kunnen ondersteunen bij zowel vakinhoudelijke als L2 aspecten. Een volledige invulling door L2 moedertaalsprekers komt ook voor, maar lijkt eerder een uitzondering te zijn (Bostwick, 2001) of om een situatie te gaan waarbij de onderwezen L2 de nationale taal is en dus aan een minderheid onderwezen wordt (Mehisto & Asser, 2007).

In Zweden worden, afhankelijk van het schooltype, zowel Zweedse leerkrachten als moedertaalsprekers van het Engels ingezet. Uit evaluatie blijkt dat sommige Zweedse leerlingen klagen over het onvermogen of de onzekerheid van de leerkracht om les te geven in het Engels, terwijl leerlingen aan de andere kant aangaven het ook problematisch te vinden wanneer hun Engelstalige leerkracht over onvoldoende kennis van het Zweeds beschikt (Skolverket, 2010; Österlund, p.c.). In het kader van vvtu-onderwijs in Nederland hebben de resultaten van het FLiPP-project aangetoond dat er in dat onderzoek geen verschil was tussen de prestaties van leerlingen met een moedertaaldocent en leerlingen met een gevorderde L2-docent (CEFR niveau C), maar dat leerlingen met een gemiddelde L2-docent (CEFR niveau B) significant lager scoorden dan de andere groepen (Unsworth et al., in druk).

4.3 Scholing en didactische vaardigheden

De opleiding van leerkrachten loopt vaak sterk uiteen, met name ook door de relatief nieuwe opzet van tpo-programma's, waardoor reguliere leerkrachten zonder relevante tpo-ervaring lesgeven in een tpo-programma. Onderwijservaring wordt uitgelegd in termen van volledig gediplomeerd en/of gecertificeerd (Cirino et al., 2007; Thomas, Collier & Abbott, 1993), het bezitten van L2 onderwijsbevoegdheden (Wode, 2004; Knell et al., 2007; Skolverket, 2010; Cenoz, 2005; Lasagabaster, 2000) en het hebben van relevante ervaring in het basisonderwijs (Dobson, Pérez-Murillo & Johnstone, 2010; Serra, 2007; Bostwick, 2001). Onderwijservaring op de basisschool loopt uiteen van enkele jaren tot meer dan tien jaar, waarbij gemiddeldes van acht jaar (Bostwick, 2001), tien jaar (Cirino et al., 2007) of meer dan tien jaar basisschoolervaring (Dobson, Pérez-Murillo & Johnstone, 2010) genoemd worden voordat leerkrachten met een tpo-programma aan de slag gingen. Bij gebrek aan specifieke tpo-ervaring zijn algemene didactische vaardigheden en onderwijservaring namelijk van groot belang. Bostwick (2001) concludeert dat samenwerking tussen L1 en L2 leerkrachten wordt gezien als een belangrijke succesfactor bij het samenstellen van activiteiten voor de leerlingen en het afstemmen van de lessen die in de L1 en L2 onderwezen worden. Burmeister en Pasternak (2004) noemen het ontwikkelen van nieuwe materialen als belangrijke factor bij het succes van tpo en Cirino et al. (2007) concluderen dat er een sterke positieve relatie is tussen de kenmerken van leerkrachten (academische vaardigheden, onderwijservaring, vakinhoudelijke kennis, cursusdeelname) en de behaalde taalvaardigheden van leerlingen.

4.4 Tpo-training

Tpo-programma's zijn in Europa relatief nieuw en naast ervaring in het basisonderwijs en eventueel met het onderwijzen van Engels als vak, hebben leerkrachten vaak weinig tot geen ervaring met lesgeven via een tpo-aanpak. Het evaluatierapport van tpo in Zweden (Skolverket, 2010), geïnitieerd door het ministerie, geeft aan dat slechts 10% van de leerkrachten een achtergrond specifiek gericht op het onderwijzen van een vreemde taal of het onderwijzen van taal- en vakgeïntegreerd onderwijs bezit en dat ervaring en kennis over tpo bij veel leerkrachten, die betrokken waren bij de start van tpo, ontbreekt (Österlund, p.c.). Verder worden er op diverse scholen extra trainingen en cursusdagen voor de leerkrachten aangeboden. De studie van Dobson, Pérez-Murillo en Johnstone (2010), uitgevoerd in opdracht van het ministerie van onderwijs in Spanje in samenwerking met de

British Council, beschrijft wel hoe een goed georganiseerde, initiële training werd aangeboden aan alle scholen en leerkrachten. In deze training werden leerkrachten begeleid en ondersteund door experts van het ministerie of de British Council door bijvoorbeeld het bezoeken van de scholen, het organiseren van bijeenkomsten en trainingen, en hulp bij de ontwikkeling van materialen en testen. Kersten, Frey & Hähnert (2009) benadrukken het belang van tpo-training voorafgaand aan de start van het programma om leerkrachten voor te bereiden op het onderwijzen in een L2 en hun eventuele onzekerheid over het onderwijzen van vakinhoud in een L2 te verminderen. Enkele scholen bieden trainingsprogramma's en (zomer)cursussen aan om zo de didactische vaardigheden van het onderwijzen in een L2 en de kennis van de L2 zelf te verbeteren (Thomas, Collier & Abbott, 1993; Knell et al., 2007). Montecel en Cortez (2002) noemen het organiseren van tussentijdse bijeenkomsten om leerkrachten constant op de hoogte te houden van nieuwe ontwikkelingen op het gebied van tpo. Ook volgens Cirino et al. (2007) blijken goede algemene didactische vaardigheden, voortkomend uit bijvoorbeeld onderwijservaring en juiste training, van grotere doorslaggevendheid in het uiteindelijke (taal)vaardigheidsniveau van leerlingen dan bijvoorbeeld het aanstellen van moedertaalsprekers. Dit werd ook door meerdere experts geconcludeerd en zij benadrukten ook allen het belang van een goede voorbereiding en ondersteuning (Genesee, p.c.; Nikolov, p.c.; Medrano, p.c.; Kersten, p.c.; Watzinger-Tharp, p.c.; Österlund, p.c.). In hetzelfde onderzoek van Cirino et al. (2007) werd ook geconcludeerd dat goede algemene didactische vaardigheden een positieve invloed hebben op de betrokkenheid en motivatie van de leerlingen. Het kunnen toepassen van verschillende strategieën (Christian, 1996), een goede taalvaardigheid in zowel L1 als L2 (Cirino et al., 2007), de leerlingen een rijke leeromgeving met afwisselende materialen kunnen aanbieden (Lorenzo, Casal & Moore, 2009), enthousiasme en ondersteuning van de school en (andere) leerkrachten, de mogelijkheid tot het volgen van trainingen en cursussen (Lasagabaster, 2000) en (financiële) hulp bij het ontwikkelen van nieuwe studiematerialen (Cenoz & Lindsay, 1996) bleken uiteindelijk belangrijke factoren te zijn bij het wel of niet slagen van tpo.

4.5 Ontwikkeling materialen

In de geïnventariseerde literatuurstudies wordt nauwelijks aandacht besteed aan de gebruikte lesmaterialen. Uit de expertinterviews blijkt echter dat dit als een erg belangrijk aspect wordt gezien in de inrichting van tpo en dat leerkrachten vaak enthousiast zijn over het ontwikkelen van nieuwe lesmaterialen (Kersten, p.c.; Nikolov, p.c.; Medrano, p.c.; Watzinger-Tharp, p.c.). In het ELIAS-project in Duitsland (Kersten et al., 2009) lag de nadruk van tpo op milieu-educatie en is in samenwerking met de lokale dieren tuin nieuw Engelstalig lesmateriaal ontwikkeld om zo theorie en praktijk direct aan elkaar te koppelen (Kersten, p.c.). In het BEP-project in Spanje (Dobson, Pérez-Murille & Johnstone, 2010) zijn diverse strategieën toegepast voor de ontwikkeling van nieuwe, voor tpo geschikte, lesmaterialen. Leerkrachten ontwikkelden, in samenwerking en met ondersteuning van experts van de British Council, nieuwe materialen met behulp van internet en beschikbare lesmaterialen uit Engeland. Daarnaast werden lesmaterialen uit het Spaanse curriculum vertaald in de L2. Hierdoor blijft de Spaanse culturele context van de materialen behouden en blijft de inhoud hetzelfde, zodat ook de doelstellingen uit het nationale curriculum behaald kunnen worden. Deze vertaalstrategie wordt ook in Utah, VS toegepast. Hier wordt tpo aangeboden in vijf verschillende talen (Spaans, Portugees, Frans, Duits en Chinees) en wordt onder andere rekenen in de L2 onderwezen. Met behulp van staatssteun is een contract opgezet met een uitgeverij om zo het tekstboek dat normaal voor rekenen in de L1 gebruikt wordt te vertalen naar een van de vijf tpo-talen. Op die manier blijft de inhoud hetzelfde, wat belangrijk is om de doelstellingen van het curriculum te behalen. Dit is ook interessant omdat in Utah voor iedere taal één leerkracht is en ieder vak voor de helft in de L1 en de L2 onderwezen wordt. Door het vertalen van de materialen wordt dus voor ieder vak in beide talen dezelfde lesmethode gebruikt, wat ten goede komt aan de continuïteit. Voor de taal- en cultuurvakken in de L2 worden authentieke lesmaterialen en werkboeken uit de betreffende landen gebruikt (Watzinger-Tharp, p.c.). Ook scholen in Zweden

maken gebruik van authentieke lesmaterialen uit Engeland, met name voor de natuurkundige onderwerpen die internationaal veel overeenkomsten vertonen (Österlund, p.c.).

4.6 Conclusie

Uit inventarisatie van de geselecteerde studies naar relevante kenmerken en didactische vaardigheden voor tpo-leerkrachten blijkt dat er nog geen ideaalprofiel voor een tpo-leerkracht bestaat en er nog nauwelijks onderzoek is gedaan naar de mogelijkheden tot tpo-trainingen en – cursussen. Om die reden bestaat er erg veel variëteit in de profielen van tpo-leerkrachten. Het minimum vereiste taal- en opleidingsniveau, ervaring in het basisonderwijs en specifiek met het doceren in een vreemde taal en het aanbod van voorbereidende en/of tussentijdse training en cursussen verschilt sterk binnen de verschillende landen. Het vereiste taalniveau voor tpo-leerkrachten loopt uiteen van minimaal een B1-niveau tot een C1-niveau van het CEFR en ook in de praktijk blijken de minimale taalvereisten sterk te verschillen. Veel studies spreken een voorkeur uit voor moedertaalsprekers, maar geven ook aan dat dit vaak niet haalbaar is en dat goede algemene didactische vaardigheden in combinatie met voldoende L2-kennis uiteindelijk van doorslaggevendere aard zijn voor het slagen van tpo dan het aanstellen van moedertaalsprekers.

Diverse studies benadrukken het belang van tpo-training voorafgaand aan en gedurende het tpo-programma om zo de didactische vaardigheden met betrekking tot tpo te verbeteren. Kenmerkend voor veel van de geïnventariseerde studies is echter dat ze zich nog in de opstartfase van tpo bevinden. Dit heeft tot gevolg dat er vaak wel aandacht besteed wordt aan één of meerdere didactische vaardigheden, maar dat door een gebrek aan middelen of tijd vaak niet met ieder aspect rekening gehouden kan worden. Hierdoor ontstaat een grote diversiteit aan leerkrachtcompetenties, iets dat vaak gecompenseerd wordt door een goede samenwerking tussen diverse leerkrachten met ieder hun specifieke vaardigheden. Over het gebruik van lesmaterialen kan op basis van de gevonden literatuur weinig worden gezegd. Uit de expertinterviews blijkt dat er diverse strategieën worden toegepast voor de ontwikkeling van tpo-materialen en dat authentiek materiaal in combinatie met zelfontwikkelde materialen het meest geschikt lijkt. Verder blijkt uit de expertinterviews dat er veel interesse is vanuit diverse landen voor het opzetten van een netwerk, waarin leerkrachten ervaringen, lesmaterialen en evaluatiemethoden uit kunnen wisselen om zo hun eigen vaardigheden, en daarmee ook het tpo-programma, te verbeteren.

5. Welke attitudes zijn er t.a.v. het onderwijsprogramma, het Engels en de eerste schooltaal bij leerlingen, leerkrachten en ouders?

De attitude van diverse belanghebbenden ten opzichte van het onderwijsprogramma is van invloed op het wel of niet slagen van tpo. Ook kan evaluatie van de houding ten opzichte van tpo-onderwijs nieuwe inzichten opleveren in zaken die van belang zijn voor het succes van tpo. In dit hoofdstuk bespreken we de houding van leerlingen, leerkrachten en ouders, evenals evaluatie-instrumenten die hun houding kunnen meten.

5.1 Evaluatie

Leerlingen

Leerlingen zijn over het algemeen erg positief over hun ervaringen met het tpo-programma. Ze vinden het leuk en interessant om een nieuwe taal te leren (Bostwick, 2001; Dobson, Pérez-Murillo & Johnstone, 2010; Jessner & Cenoz, 2007; Cenoz & Lindsay, 1996; Skolverket, 2010; Merisuo-Storm, 2007). Een evaluatie van Bostwick (2001) toont aan dat wanneer leerlingen zelf de keuze konden maken, ze ook op een school wilden zitten waar in twee talen onderwezen werd. Veel leerlingen geloven ook dat het belangrijk en nuttig is voor hun toekomstige studie en carrière om een goede L2 vaardigheid te hebben (Skolverket, 2010, Merisuo-Storm, 2007) en zijn ze trots op hun verworven L2 kennis (Merisuo-Storm, 2007).

Leerlingen geven aan dat ze veel vertrouwen hebben in hun L2 vaardigheden en hun zelfvertrouwen in het algemeen (Bostwick, 2001; Christian, 1996; Dobson, Pérez-Murillo & Johnstone, 2010; Merisuo-Storm, 2007). De opgedane L2 kennis helpt ze ook om de kennis van andere vakken te verbreden en een groter vertrouwen in andere vakkennis te ontwikkelen (Dobson, Pérez-Murillo & Johnstone, 2010). Opmerkelijk is dat tpo-leerlingen in één studie door middel van zelfevaluatie hun L2 vaardigheden gemiddeld lager inschatten dan leerlingen in een regulier onderwijsprogramma (Seikkula-Leino, 2007). Leerlingen in een tpo-programma lijken een sterkere zelfkritiek te hebben, ondanks het feit dat ze een goede kennis van de L2 hebben verworven. Een mogelijke verklaring hiervoor is dat ze de moeilijkheden die horen bij het leren en spreken van een L2 zien als falen of onbekwaamheid en daarom meer kritisch zijn op hun taalvaardigheden in plaats van in te zien hoe goed hun vooruitgang eigenlijk is.

De studies van Cenoz (2005), Merisuo-Storm (2007) en De Rivière (2004) tonen aan dat tpo-leerlingen een grotere motivatie laten zien bij het leren van een taal dan leerlingen in een regulier onderwijsprogramma, waarbij Merisuo-Storm (2007) een significant verschil van $t = 5.71$, $p < .001$ rapporteert. Ook staan tpo-leerlingen positiever tegenover spreken, $t = 4.70$, $p < .001$, en schrijven, $t = 2.73$, $p = .007$, in de L2 dan leerlingen in reguliere onderwijsprogramma's.

Ytsma en Beetsma (2001) concluderen dat leerlingen na enkele jaren tpo-onderwijs gemotiveerd zijn om door te leren in de L2. De studie van Merisuo-Storm (2007) in reguliere onderwijsprogramma's laat zien dat meisjes een significant positievere leerhouding hebben dan jongens. Dit verschil is significant in regulier onderwijs ($t = -2.98$, $p = .004$), terwijl dit verschil in tpo-programma's kleiner en niet-significant verschillend is, wat duidt op een positieve stimulatie van tpo op de leerhouding (Merisuo-Storm, 2007). Cenoz (2005) benadrukt dat jonge leerders over het algemeen een positievere houding ten opzichte van het leren van de L2 laten zien dan oudere leerders, iets dat volgens Cenoz mogelijk verklaard kan worden door psychologische (leeftijd) en onderwijskundige (lesmethode) factoren. Uit de interviews met diverse experts blijkt dat kinderen over het algemeen ontzettend enthousiast zijn over het leren van een nieuwe taal, waarbij Watzinger-Tharp (p.c., 14 januari 2014) opmerkte dat bij een gestructureerd 50-50 model leerlingen na een aantal jaar niet eens meer doorhebben dat ze op een speciale manier onderwezen worden, omdat zowel de L1 als de L2 lessen als vanzelfsprekend worden beschouwd. Verder laten leerlingen een

positievere houding zien ten opzichte van andere talen en culturen, d.w.z. een verhoogde talensensibilisering, en hebben een groter sociocultureel bewustzijn (Christian, 1996; Montecel & Cortez, 2002; De Rivière, 2004; Dobson, Pérez-Murillo & Johnstone, 2010; Frijns, Sierens, van Gorp, Sercu, van Avermaet & Devlieger, 2011). Christian (1996) beschrijft een studie waarin leerlingen met behulp van vragenlijsten aangaven makkelijker vriendschappen te ontwikkelen onafhankelijk van de etniciteit van andere leerlingen in de klas of school. Watzinger-Tharp (p.c.) benadrukte het belang van onderzoek naar sociocultureel bewustzijn en identiteit. Het onderzoek naar 50-50 tpo-programma's dat wordt uitgevoerd in Utah besteedt veel aandacht aan de evaluatie van dit sociocultureel bewustzijn als gevolg van tpo-onderwijs. Watzinger-Tharp gaf aan dat dit aan de hand van interviews en vragenlijsten werd onderzocht, maar dat er nu voornamelijk grote interesse was in de keuze van leerlingen in hun overstap van het basisonderwijs naar het voortgezet onderwijs. Met een keuze uit diverse middelbare scholen, waarbij enkele een vervolg op tpo aanbieden, willen ze onderzoeken in hoeverre tpo heeft bijgedragen aan het sociocultureel en globaal bewustzijn van leerlingen en in hoeverre dit hun keuze voor vervolgonderwijs beïnvloedt.

Leerkrachten

Massler (2012) en Bostwick (2001) merken op dat er in het beginstadium vaak enige bezorgdheid is bij leerkrachten over de uitvoering van tpo. Met name wanneer het tpo-programma volledig nieuw is en er nog geen ervaring is met het onderwijzen in een L2 uiten leerkrachten bezorgdheid over hun didactische vaardigheden en over de grote werklast die gepaard gaat met het opstarten van het nieuwe onderwijsprogramma. Massler (2012) stelt dat leerkrachten voorafgaand aan de start van een tpo-programma soms wel de kans krijgen om trainingen en cursussen te volgen, maar dat dit vaak niet genoeg is om hun onzekerheid over grotere verantwoordelijkheden en het onderwijzen in een tpo-programma volledig te laten verdwijnen (Massler, 2012).

Massler (2012) noemt het verbeteren van de L2 vaardigheden, algemene didactische vaardigheden, de aanschaf en ontwikkeling van materialen en de extra tijd die gepaard gaat met het invoeren van een nieuwe onderwijsmethode zowel als positieve factoren die bijdragen aan persoonlijke en professionele ontwikkeling, maar ook als negatieve factoren die zorgen voor extra werkbelasting. Dobson, Pérez-Murillo en Johnstone (2010) concluderen dat naar mate het tpo-programma is opgestart en de eerste zorgen zijn weggenomen, leerkrachten het tpo als zeer waardevol en nuttig beschouwen voor zowel de leerlingen, de school en voor zichzelf. Bostwick (2001) en Cenoz en Lindsay (1996) noemen dat leerkrachten na verloop van tijd positiever worden over tpo en met name de mogelijke positieve intellectuele uitkomsten ervan voor de leerlingen als belangrijk effect van tpo beschouwen. Ook Dobson, Pérez-Murillo en Johnstone (2010) noemen de verworven taalvaardigheid in de L2, cognitieve ontwikkeling, sociale en culturele ontwikkeling, openheid en meer zelfvertrouwen van leerlingen als de belangrijkste opbrengsten van tpo.

Dobson, Pérez-Murillo en Johnstone (2010) benoemen het ontwikkelen van nieuwe vaardigheden, het verbeteren van lesmethoden, het vergaren van nieuwe kennis, voldoening van het onderwijzen in tpo en het opzetten van nieuwe contacten als de belangrijkste positieve aspecten voor leerkrachten die lesgeven in een tpo-programma. Montecel en Cortez (2002) benadrukken dat leerkrachten zich verantwoordelijk voelen voor het succes van leerlingen in een tpo-programma en dat succes en positieve resultaten van tpo bijdragen aan een positievere houding van leerkrachten ten opzichte van tpo. Iets dat leerkrachten ervan kan overtuigen door te gaan met het tpo-programma of anderen juist kan motiveren om dit te doen.

Massler (2012) en Mehisto en Asser (2007) noemen training als voorbereiding op het onderwijzen in een tpo-programma, de beschikbaarheid van relevante materialen en de mogelijkheid tot het verzamelen en ontwikkelen van nieuwe materialen als belangrijke factoren die bijdragen aan de kwaliteit van het tpo-programma. Een betere voorbereiding zorgt ook voor meer zekerheid bij leerkrachten, wat ten goede komt aan hun didactische vaardigheden. Mehisto en Asser (2007) noemen ook feedback en tussentijdse training en cursussen als een belangrijke factor in het succes

van tpo, evenals het evalueren van bepaalde strategieën die focussen op de minder gemotiveerde of academisch minder sterke leerlingen.

Ouders

Ouders laten over het algemeen een positieve houding zien ten opzichte van tpo en vinden het nuttig dat hun kind een extra taal leert (Dobson, Pérez-Murillo & Johnstone, 2010; De Rivière, 2004; Cenoz & Lindsay, 1996, Ytsma & Beetsma, 2001). Het belang van tpo, de verworven vaardigheden in de L2, de verbreding van toekomstperspectieven, de ontwikkeling van algemene vaardigheden en een groter inlevingsvermogen in andere landen en culturen worden genoemd als belangrijke positieve uitkomsten die ouders toeschrijven aan tpo-programma's (Bostwick, 2001; Thomas, Collier & Abbott, 1993; Dobson, Pérez-Murillo & Johnstone, 2010; Skolverket, 2010, Mehisto & Asser, 2007). De Rivière (2004) toont ook aan dat de motivatie onder ouders om de L2 te leren toeneemt wanneer hun kind deelneemt aan tpo.

Ondanks de overwegend positieve reacties op tpo uiten ouders bij de start van een tpo-programma nog vaak hun bezorgdheid en zijn ze wat sceptisch over de mogelijke positieve opbrengsten van tpo (Lasagabaster, 2001; Bostwick, 2001). Volgens Bostwick (2001) wordt deze bezorgdheid vaak toegeschreven aan onwetendheid over de mogelijke effecten van zo'n programma en ook het gebrek aan informatie bij de opzet van een tpo-programma. Lasagabaster (2000, 2001) noemt de mogelijke verwarring tussen de L1 en L2 (en L3 in de situatie in bijvoorbeeld Baskenland) als reden tot bezorgdheid (Lasagabaster 2000, 2001). Naar mate meer bekend is over het programma en leerlingen zelf ook positief zijn, zien ook ouders in dat de voordelen belangrijker wegen dan de eventuele moeilijkheden die het programma met zich meebrengt.

De SES van ouders speelt ook een rol in hun houding ten opzichte van het tpo-programma (Skolverket, 2010; Nikolov, p.c., 12 december 2013). Hoger opgeleide ouders, eventueel met internationale studie- en/of werkervaring, zijn vaak minder sceptisch over deelname van hun kind aan een tpo-programma en laten een positievere houding zien ten opzichte van onderwijs in een L2, iets wat vaak terug te zien is in een beter schoolondersteunend thuisclimaat. Ook een juiste informatievoorziening kan tot een beter schoolondersteunend thuisclimaat leiden. Mehisto en Asser (2007) en Bostwick (2001) benadrukken het belang van een juiste informatievoorziening. Ouders vinden het belangrijk om in dialoog te kunnen gaan met de leerkracht en regelmatig te worden geïnformeerd over de ontwikkeling van het tpo-programma, de kennis van leerkrachten wat betreft tpo, de lesmethoden en de behaalde resultaten van leerlingen (Genesee, p.c.).

5.2 Evaluatievormen

Leerlingen

Het is belangrijk om de houding van leerlingen ten opzichte van tpo, de L2 en het leren via die L2 te evalueren. Verschillende studies (e.g., Merisuo-Storm, 2007; Dobson, Pérez-Murillo & Johnstone, 2010) tonen aan dat een positieve houding kan zorgen voor een grotere motivatie, waardoor het leren van de L2 gestimuleerd wordt. Aan de hand van leerkrachtobservaties of mondelinge interviews kan bijvoorbeeld de houding van leerlingen in de onderbouw van de basisschool onderzocht worden (Massler, 2012). Wat oudere leerlingen kunnen vragenlijsten invullen, om hun houding ten opzichte van het onderwijsprogramma aan te geven en bijvoorbeeld ook wat ze geleerd hebben, wat ze belangrijk vinden, wat ze graag anders zouden willen zien en hoe vaak ze de L2 gebruiken buiten school om. Voorbeelden van zulke leerlingvragenlijsten zijn beschikbaar in de studies van Dobson, Pérez-Murillo en Johnstone (2010) en Bostwick (2001). In Zweden zijn enkele tweetalige basisscholen gestart met het aanbieden van proeflessen in het Engels, zodat kinderen, maar ook ouders, kunnen ervaren wat tpo inhoudt en gebaseerd op deze ervaring een keuze kunnen maken voor een basisschool (Österlund, p.c.).

Leerkrachten

De meest toegepaste methoden om de houding van leerkrachten ten opzichte van tpo-onderwijs te onderzoeken is door het gebruik van interviews, discussies of vragenlijsten (Massler, 2012; Dobson, Pérez-Murillo & Johnstone, 2010). Op die manier kan feedback op verschillende tpo-onderwerpen verkregen worden. Met behulp van vragenlijsten uit de studies van Massler (2012) en Dobson, Pérez-Murillo en Johnstone (2010) is bijvoorbeeld nagegaan hoe leerkrachten denken over de voortgang in taalvaardigheid en de balans tussen het leren van vakinhoud en taal, maar ook de professionele ontwikkeling van leerkrachten, hun L2 vaardigheid, de werkdruk, het aanbod van lerarenopleidingen en extra tpo-trainingen en -cursussen, de beschikbaarheid van lesmaterialen en voldoende ondersteuning vanuit de school zijn zo onderzocht.

Ouders

De houding van ouders ten opzichte van het tpo-programma kan gepeild worden door het organiseren van informatiebijeenkomsten of door het gebruik van vragenlijsten (Massler, 2012; Dobson, Pérez-Murillo & Johnstone, 2010; Bostwick, 2001; Lindgren & Muñoz, 2013). Ouders vinden het belangrijk om bij te dragen aan de verdere ontwikkeling van het onderwijsprogramma van hun kind en staan daarom vaak open voor het invullen van bijvoorbeeld vragenlijsten ter bevordering van het onderwijs (Massler, 2012). Door middel van opbrengsten van vragenlijsten of informatiebijeenkomsten kan bijvoorbeeld voorafgaand aan de start van het tpo-onderwijs ingespeeld worden op eventuele zorgen wat betreft het programma of bezorgdheid over de ontwikkeling van hun kind dat onderwezen wordt in een tpo-programma. Gebaseerd op deze uitkomsten kan de informatievoorziening vanuit de school naar de ouders worden aangepast of uitgebreid. Door meer kennis van het tpo-programma en een beter zicht op de vooruitgang en resultaten van hun kinderen leveren vragenlijsten die gedurende of aan het eind van het tpo-programma afgenomen worden vaak meer gedifferentieerde antwoorden op. Hierdoor is het erg nuttig om houding van ouders ten opzichte van tpo op meerdere momenten te evalueren en zo bij te dragen aan een verbetering van het tpo (Massler, 2012).

Hawthorne-effect

Het 'Hawthorne-effect' kan een factor zijn die van invloed is op de manier waarop leerlingen over hun attitude rapporteren. Het effect laat zien dat verschillen op de onafhankelijke variabele niet alleen veroorzaakt hoeven te worden door die variabele zelf, maar ook beïnvloed kunnen zijn door andere factoren, bijvoorbeeld doordat de respondenten weten dat ze onderzocht worden (Mellow, Reeder & Forster, 1996). Mellow et al. (1996) beargumenteren dat het risico op dit effect kan worden verkleind of uitgeschakeld wanneer proefpersonen volledig willekeurig deelnemen en er verschillende testmomenten, bijvoorbeeld door het doen van longitudinaal onderzoek, worden ingelast. Bijna alle geselecteerde studies in deze rapportage voldoen aan één of beide van deze voorwaarden en dus maakt dit het Hawthorne-effect onwaarschijnlijk.

5.3 Conclusie

Het meten en evalueren van attitudes bij leerlingen, leerkrachten en ouders ten aanzien van het nieuwe onderwijsprogramma is van belang voor het slagen en voor de verdere ontwikkeling van tpo. Over het algemeen kan geconcludeerd worden dat leerlingen vaak zeer positief zijn over het deelnemen aan een tpo-programma en het leren van een nieuwe taal, en dat ze ook het nut inzien van het leren van deze nieuwe taal voor hun toekomst. Verder lijkt tpo een positief effect te hebben op de algemene leerhouding van leerlingen, waarbij opgemerkt moet worden dat dit aan het begin van de basisschool vaak sterker is dan aan het eind, omdat het onderwijs via de tpo-methode na enkele jaren als meer vanzelfsprekend wordt gezien. Leerkrachten kunnen in het beginstadium van tpo soms nog wat terughoudend of bezorgd zijn over de gevolgen van tpo voor de leerlingen en voor

hun eigen verantwoordelijkheden, maar dit verdwijnt vaak snel wanneer de positieve effecten van tpo zichtbaar worden en wanneer ze door extra training meer zekerheid krijgen over hun eigen didactische vaardigheden in de vreemde taal. Ouders zijn overwegend positief over de implementatie van tpo en zijn vaak erg enthousiast over deelname van hun kind aan een tpo-programma. Ondanks dat sommige ouders aan het begin sceptisch zijn over de mogelijke extra last die tpo met zich meebrengt voor de leerlingen, is in veel landen en op veel scholen uit de geïnventariseerde studies de vraag naar tpo groter dan het aanbod. Een juiste informatievoorziening, voor zowel leerlingen, ouders als leerkrachten, is van groot belang om de eventuele zorgen met betrekking tot deze nieuwe onderwijsmethode weg te nemen.

Ook diverse evaluatiemethoden kunnen bijdragen aan een verbeterde informatievoorziening, betere didactische vaardigheden en een verbetering van het tpo-programma. De attitude van leerlingen kan, afhankelijk van de leeftijd, geëvalueerd worden door leerkrachtobservaties, mondelinge interviews of door het gebruik van vragenlijsten. Voor ouders kan het organiseren van informatiebijeenkomsten erg zinvol zijn om ze vooraf en tussentijds in te lichten over de voortgang van hun kinderen en het tpo-programma. Ook kan de attitude van ouders geëvalueerd worden aan de hand van vragenlijsten, waarin ze zowel de voortgang van hun eigen kind als hun eigen houding ten opzichte van tpo kan worden aangegeven. Massler (2012) concludeert dat leerkrachten niet altijd even positief zijn over deelname aan evaluatie-onderzoek, omdat zij vanwege de extra tijd die gepaard gaat met het onderwijzen in een tpo-programma bijvoorbeeld niet geïnteresseerd zijn in nog meer taken naast de extra uren die ze al besteden aan tpo. Voor de voortgang van tpo is het echter van belang om ook de ervaringen van leerkrachten te evalueren, bijvoorbeeld met behulp van vragenlijsten of discussiebijeenkomsten, om zo gerichte aanpassingen in het tpo-programma te kunnen doen en leerkrachten eventueel extra trainingen te kunnen aanbieden.

6. Wat zijn mogelijkheden voor internationale samenwerking op het terrein van curriculumontwikkeling en onderzoek?

Met het interviewen van verschillende experts is al een eerste stap gezet wat betreft mogelijkheden tot internationale samenwerking. Deze experts komen voort uit het brede en internationale netwerk dat de projectleiders en leden van de klankbordgroep bezitten en dat ook voor verder tpo-onderzoek van groot belang kan zijn. De geïnterviewde experts waren allemaal erg enthousiast over dit inventarisatieonderzoek en de voorgestelde mogelijkheden tot internationale samenwerking, zoals bijvoorbeeld op het gebied van curriculum-, materiaal- en toetsontwikkeling, om ervaringen uit te wisselen of om een internationale samenwerking met scholen en leerlingen aan te gaan.

Een voorbeeld van internationale samenwerking zou bijvoorbeeld het organiseren van een symposium of congres kunnen zijn, waarbij onderzoekers, leerkrachten en beleidsmakers ervaringen kunnen uitwisselen. Ook het opzetten van een schoolnetwerk, waarbij niet alleen leerkrachten en scholen kunnen samenwerken, maar waarbij ook tpo-leerlingen contacten met andere tpo-leerlingen in het buitenland kunnen aangaan via internet, behoort tot de mogelijkheden. Vanuit de expertinterviews blijkt dat hiervoor voldoende interesse is en er zijn zelfs al enkele suggesties gedaan voor mogelijkheden hiertoe.

Op schoolniveau zou onderzoek met betrekking tot internationale oriëntatie of interculturele benadering en de uitwerking daarvan in de praktijk erg interessant zijn. In de geïnventariseerde studies is dit geen onderwerp van onderzoek geweest, maar uit de expertinterviews blijkt dat scholen hier wel degelijk mee bezig zijn. Met name in Utah vindt een uitbreiding plaats naar het testen en evalueren van biculturele kennis en bewustzijn. Naast het testen van de taalvaardigheid wordt in Utah ook onderzoek gedaan naar de vorming van biculturele identiteit als gevolg van tpo, hoe leerlingen zich verbonden voelen met andere culturen als gevolg van tpo en wat voor invloed dit heeft op hun verdere onderwijskeuzes (Watzinger-Tharp, p.c.). Ook in Duitsland vond aan de hand van interviews op scholen een observatie en evaluatie van interculturele vaardigheden en cultureel bewustzijn plaats (Kersten, p.c.).

Daarnaast zou het voor de opzet van tpo in Nederland erg interessant zijn om contact te leggen met landen waar een vergelijkbare onderwijssituatie is, waar al een start gemaakt is met tpo en waar op sociocultureel, politiek en taalkundig vlak voldoende overeenkomst is met de situatie in Nederland. Een voorbeeld van zo'n land is Zweden, waar vanuit het ELLiE-project al diverse contacten gelegd zijn en dit dus gemakkelijk weer geïntensiveerd kan worden. Contact met onderzoekers, beleidsmakers, scholen en leerkrachten in een land dat in meerdere aspecten overeenkomt met de Nederlandse onderwijssituatie kan bijdragen aan een nog betere voorbereiding, opzet en uitvoering van het tpo-onderzoek in Nederland en daarnaast kan het ook bijdragen aan de materiaal-, toets- en curriculumontwikkeling.

7. Discussie en aanbevelingen

In dit rapport hebben we door middel van een literatuurstudie en een expertconsultatie antwoord gegeven op de zes onderzoeksvragen met betrekking tot de opzet en uitvoering van tpo in Nederland. We hebben ons in de literatuurstudie gericht op een inventarisatie en samenvatting van resultaten uit andere overzichtsstudies, in combinatie met relevante resultaten uit andere recente studies naar tpo zowel binnen als buiten Europa. Door middel van de expertinterviews hebben we specifieke en aanvullende informatie en praktische ervaringen verzameld uit een aantal andere landen die al meer ervaring hebben met tpo.

Uit de resultaten van deze voorstudie komen de volgende aanbevelingen voor de tpo-pilot en het flankerende onderzoek voort:

7.1 Inrichting tpo-programma's en deelnemende scholen

- Werk zo veel mogelijk toe naar 50% onderwijstijd in het Engels; minder tijd lijkt minder effectief (Genesee, p.c.; Kersten, p.c.).
- Beslissingen over welke vakken in de L2 onderwezen moeten worden, kunnen op lokaal niveau genomen worden; informatie van de experts en resultaten uit de literatuurstudie geven geen reden om bepaalde vakken niet in het Engels te geven (incl. rekenen). Er zijn weinig voorbeelden gevonden waarbij vakken deels in L1 en deels in L2 worden gegeven (of de ene periode in L1 en de andere in L2), maar waar dit wel gerapporteerd wordt, blijkt dit geen nadelig effect te hebben (Genesee, p.c.; Medrano, p.c.; Skolverket, 2010).
- Uit de literatuurstudie is weinig informatie over de herkomst van lesmaterialen naar voren gekomen. Naar aanleiding van de expertinterviews blijkt dat een combinatie van authentieke en zelfontwikkelde materialen het meest wenselijk is. Er dient daarom een goede inventarisatie plaats te vinden van de huidige lesmaterialen, beschikbare internationale lesmaterialen en eventueel zelf te ontwikkelen lesmaterialen (Genesee, p.c.; Medrano, p.c.; Watzinger-Tharp, p.c.; Österlund, p.c.).
- Er is over het algemeen weinig gekeken naar leerlingen met een andere L1 achtergrond (anders dan de landstaal, in ons geval het Nederlands dus). In de studie waar dit wel uitgebreid onderzocht is (Skolverket, 2010; Österlund, p.c.), bleek deelname aan tpo een positief effect te hebben op de ontwikkeling van de landstaal bij kinderen met een andere L1 achtergrond. Er is daarom geen enkele reden om NT2-leerlingen of leerlingen met een andere thuistaal (bijvoorbeeld in Friesland) uit te sluiten van deelname aan tpo. Belangrijk is om de resultaten hiervan goed te evalueren en ouders en leerkrachten vooraf goed te informeren over (positieve) verwachtingen m.b.t. het effect van tpo op de taalontwikkeling van deze leerlingen (Genesee, p.c.).
- Er is geen reden om leerlingen vooraf te selecteren op hun leerprestaties en vaardigheden. Alle leerlingen, met alle niveaus van schoolse vaardigheden, kunnen functioneren in tpo en leerlingen met zwakkere leerprestaties kunnen zelfs profiteren van tpo (Genesee, 2007, p.c.). Vanzelfsprekend vraagt dit om leerkrachten die goed kunnen differentiëren binnen de groep.
- Betrokkenheid van moedertaalsprekers Engels als docenten of klassenassistenten is zinvol, maar niet per se noodzakelijk. Leerlingen behalen dezelfde resultaten bij docenten met goede L2-vaardigheden als bij moedertaalsprekers. Goede didactische vaardigheden en competenties zijn minimaal net zo belangrijk (FLIPP, 2012; Genesee, p.c.; Medrano, p.c.).

- Er dient voldoende tijd vrijgemaakt te worden voor extra scholing en begeleiding van leerkrachten en scholen, zowel ter voorbereiding op het tpo-programma, als gedurende het tpo-programma (Genesee, p.c.; Massler, 2012; Österlund, p.c.).

7.2 Opzet onderzoek – wat te meten en hoe?

Op leerlingniveau

- Om een volledig overzicht te krijgen van de opbrengsten van tpo moet naast taalvaardigheid in het Engels (de L2, maar in sommige gevallen L3 of de L1) en het Nederlands (voor de meerderheid van de leerlingen is dit de L1, maar soms ook de L2), ook de schoolprestaties in verschillende vakken, attitude (m.b.t. het Engels en ook in bredere zin), motivatie en andere cognitieve vaardigheden (zoals executieve functies) gemeten worden. Deels kan dat op basis van in de school al beschikbare informatie (zoals via een leerlingvolgsysteem) (Genesee, p.c.). Zowel kwantitatieve gegevens over leerprestaties als kwalitatieve over leer- en interactieprocessen zijn relevant. Het flankerende onderzoek vraagt daarom om een brede en multidisciplinaire opzet.
- Het is belangrijk om ook externe factoren te meten die van invloed kunnen zijn op de uiteindelijke (taal)vaardigheden. Hierbij valt te denken aan buitenschools taalcontact, taalachtergrond, en opleiding en SES van de ouders (Dobson, Pérez-Murillo & Johnstone, 2010; Genesee, p.c.; Medrano, p.c.; Skolverket, 2010; Österlund, p.c.).
- In de evaluatie van diverse taalkundige en niet-taalkundige aspecten is het toepassen van meerdere methoden (*triangulation*) erg belangrijk om vanuit verschillende invalshoeken meerdere/betrouwbaardere resultaten te verkrijgen, bijvoorbeeld d.m.v. interviews, observaties, toetsen en attitude-onderzoek (Genesee, p.c.; Medrano, p.c.; feedback stuurgroep).
- Gezien de looptijd van de pilot (5 jaar) is een combinatie van longitudinaal en cross-sectioneel onderzoek en dataverzameling zinvol om het ontwikkelingstraject van leerlingen het beste in kaart te brengen (Genesee, p.c.; feedback stuurgroep).
- Maak voor het toetsen van de Nederlandse taalvaardigheid (waar mogelijk) gebruik van genormeerde instrumenten zodat er vergelijkingen gemaakt kunnen worden met leerlingen in het reguliere basisonderwijs (Genesee, p.c.; Medrano, p.c.). Het leerlingvolgsysteem (LVS) behoort tot de mogelijkheden, maar hierbij dient een goede inventarisatie plaats te vinden omdat niet alle scholen hetzelfde LVS gebruiken.
- Maak voor het toetsen van de Engelse taalvaardigheid (waar mogelijk) gebruik van instrumenten die al in internationaal onderzoek in vergelijkbare onderwijscontexten gebruikt worden en die onafhankelijk van specifieke L1-achtergrond zijn. Hierdoor wordt veel tijd bespaard in de pilotfase van het onderzoek en worden de resultaten zo vergelijkbaar mogelijk met al eerder verschenen onderzoek. Voorbeelden van dergelijke instrumenten en procedures zijn het PIRLS onderzoek voor leesvaardigheid, PPVT voor woordenschat, de TROG voor grammatica (Genesee, p.c.; Kersten, p.c.).
- Maak voor het toetsen van de overige Engelstalige vakken (waar mogelijk) ook gebruik van genormeerde instrumenten zodat er vergelijkingen gemaakt kunnen worden met leerlingen in het reguliere basisonderwijs. Op basis van het beschikbare onderzoek is onvoldoende duidelijk of deze toetsing het beste in de L1 of de L2 kan worden afgenomen. Zowel toetsing in het Nederlands als in het Engels kan daarom onderzoeksrelevant zijn en nader overwogen worden (Dobson, Pérez-Murillo & Johnstone, 2010; Medrano, p.c.).
- Maak voor het toetsen van de overige Engelstalige vakken (waar mogelijk) ook gebruik van genormeerde instrumenten zodat er vergelijkingen gemaakt kunnen worden met leerlingen in het reguliere basisonderwijs. Op basis van het beschikbare onderzoek is onvoldoende duidelijk of deze toetsing het beste in de L1 of de L2 kan worden afgenomen. Zowel toetsing in het Nederlands als in het Engels kan daarom onderzoeksrelevant zijn en nader overwogen worden (Genesee, p.c.; Dobson, Pérez-Murillo & Johnstone, 2010; Medrano, p.c.).

- Maak voor het toetsen van de overige Engelstalige vakken (waar mogelijk) ook gebruik van genormeerde instrumenten zodat er vergelijkingen gemaakt kunnen worden met leerlingen in het reguliere basisonderwijs. Op basis van het beschikbare onderzoek is onvoldoende duidelijk of deze toetsing het beste in de L1 of de L2 kan worden afgenomen. Zowel toetsing in het Nederlands als in het Engels kan daarom onderzoeksrelevant zijn en nader overwogen worden (Genesee, p.c.; Dobson, Pérez-Murillo & Johnstone, 2010; Medrano, p.c.).
- Voor een volledige beoordeling is het belangrijk om een goede combinatie te gebruiken van standaard toetsgegevens (LVS), gestandaardiseerde toetsen (PPVT, TROG) en zelfontwikkelde toetsen voor specifiekere aspecten van de taalvaardigheid (Genesee, p.c.; Watzinger-Tharp, p.c.; Österlund, p.c.).

Op leerkrachtniveau

- Er is tot nu toe relatief weinig bekend over het niveau van leerkrachten en de invloed van de L2 competenties van leerkrachten binnen tpo; om deze lacune in de huidige literatuur in te vullen, zou dit in kaart gebracht moeten worden in de tpo-pilot. L2 competenties van leerkrachten kunnen gemeten worden via internationaal gevalideerde toetsen (FLiPP, 2012).
- Er is ook relatief weinig bekend over de rol van specifieke didactische vaardigheden van leerkrachten binnen tpo; om deze lacune in de huidige literatuur in te vullen zou ook dit in kaart gebracht moeten worden in de tpo-pilot. Interactievaardigheden en didactische vaardigheden kunnen geëvalueerd worden aan de hand van klasobservaties. Hiervoor zijn enkele instrumenten beschikbaar uit onderzoek elders (Cirino et al., 2007; Dobson, Pérez-Murillo, Johnstone, 2010; Massler, 2012; Kersten, p.c.), maar in de meeste studies is dit niet grootschalig gevalideerd.
- Er is relatief weinig tot niets bekend over de invloed van tpo op het L1-deel van het curriculum en wat de invloed is op de talige, didactische en organisatorische vaardigheden van de L1-leerkracht. Dit kan geëvalueerd worden aan de hand van observaties en interviews of door het gebruik van vragenlijsten (feedback stuurgroep).
- Voor het bepalen van leerkrachtvaardigheden kan gebruik worden gemaakt van taaltoetsen om de L2 competentie te meten, van klassenobservaties om de didactische competenties in kaart te brengen, van vragenlijsten en interviews om de attitude t.o.v. de pilot tpo, de didactische activiteiten en de behoefte aan nascholing/ondersteuning in kaart te brengen (Cirino et al., 2007, Pérez-Murillo, Johnstone, 2010; Massler, 2012; Kersten, p.c.; Watzinger-Tharp, p.c.).

Op organisatieniveau

- Maak een inventarisatie en beschrijving van het innovatie-, implementatie- en kwaliteitszorgproces in de deelnemende scholen (feedback stuurgroep).
- Maak een inventarisatie van de opzet, uitwerking en waardering van de samenwerking met het VO. Hier is, mede door de nieuwe opzet van tpo in Europa, nog weinig onderzoek naar gedaan. Alleen het onderzoek in Spanje (Dobson, Pérez-Murillo & Johnstone, 2010) beschrijft hoe er vanaf het begin overeenstemming was met enkele VO-scholen, zodat leerlingen hun tpo konden voortzetten in het VO. Ook in Canada en de VS wordt hier zorg voor gedragen, maar daar bestaat al een langere historie met tpo (Genesee, p.c.; Watzinger-Tharp, p.c.).
- Maak een inventarisatie van de opzet, uitwerking en waardering van de samenwerking met Pabo's, nascholings- en begeleidingsorganisaties (feedback stuurgroep).
- Het organisatieniveau is in de gerapporteerde studies weinig onderzocht. We bevelen aan de opvattingen en activiteiten van schoolleiders te inventariseren in wat ze doen om het team te faciliteren tijdens de pilot. Dat kan o.a. via interviews en via focusgroepen (Genesee, p.c.; Watzinger-Tharp, p.c.; Medrano, p.c.; Österlund, p.c.; feedback stuurgroep).

Onderzoeksmatige uitdagingen

- Het is belangrijk om vooraf de voorwaarde te concretiseren voor scholen en leerkrachten in de tpo-pilot dat ze daarbij ook deelnemen aan evaluatie-onderzoek (Massler, 2012).
- Gezien bovenstaande aanbevelingen vraagt het flankerende onderzoek om een inter-/multidisciplinair onderzoeksteam, waarbij kennis vanuit taalwetenschappelijk (L1 en L2), taaldidactisch, onderwijskundig en beleidsmatig perspectief aanwezig is (feedback stuurgroep).
- Een eerste en zeer belangrijke stap van het onderzoek is het breed inventariseren van de stand van zaken m.b.t. de huidige en ingeplande invulling van de tpo-programma's, de samenstelling van de schoolpopulatie en de kenmerken van de docenten. Alleen op basis hiervan kan de preciezere opzet van de uit te voeren (deel)projecten binnen de pilot vastgesteld worden (Genesee, p.c.).
- Wanneer niet alleen gekeken wordt naar onderwijscontext met alle bijbehorende factoren wordt gekeken, maar ook naar sociale, culturele en politieke aspecten, komt de tpo-studie uit Zweden (Skolverket, 2010) in grote mate overeen met de situatie in Nederland. Deze studie zou als voorbeeld beschouwd kunnen worden voor het vervolg en de uitvoering van de tpo-pilot. Ook concrete samenwerking met het betreffende onderzoeksinstituut en de betrokken scholen kan relevant zijn.
- De meeste onderzoeken werken met zelf-geselecteerde scholen; enkele grote onderzoeken daargelaten is hier bijna niet aan te ontkomen. Er moet echter wel rekening gehouden worden met de generaliseerbaarheid van de resultaten en de rol van andere factoren buiten het tpo-programma om, zoals buitenschools taalaanbod, SES en opleidingsniveau van de ouders. Door een goede inventarisatie van de leerlingpopulatie op de deelnemende scholen en van de onderwijskwaliteit van de deelnemende scholen (inspectierapporten, 'Dronkers-lijst'), kunnen de opbrengsten in verband gebracht worden met die van vergelijkbare niet-tpo-scholen (Genesee, p.c.; Watzinger-Tharp, p.c.; Österlund, p.c.).
- Het meenemen van juiste controlegroepen is belangrijk, maar het is niet altijd evident wat de juiste controlegroep zou moeten zijn. Dit moet goed worden afgewogen aan de hand van specifieke onderzoeksvragen en de kenmerken van de deelnemende scholen. Dit maakt een goede inventarisatie aan het begin ook weer des te belangrijker. De ontwikkeling van de Engelse taalvaardigheid zal daarom vooral gemeten worden binnen het project. De ontwikkeling van de L1 kan bijvoorbeeld landelijk geëvalueerd worden (Genesee, p.c.; Nikolov, p.c.; feedback stuurgroep).
- Het dynamische en experimentele karakter van de tpo-pilot zal tot gevolg hebben dat de onderzoeksomgeving weinig stabiel zal zijn. In de loop van de vijf pilotjaren zullen de scholen kwalitatief en organisatorisch sterk in ontwikkeling zijn. Dit maakt het lastig om longitudinaal onderzoek goed te controleren. Dit komt in de literatuurstudies ook vaker voor, maar hier wordt niet specifiek verder op ingegaan (feedback stuurgroep).
- De leeftijdsrange over de gehele basisschoolperiode (groep 1 t/m 8) is groot. Dit heeft gevolgen voor (i) de vaardigheden die gemeten kunnen/moeten worden, (ii) het te gebruiken instrumentarium, (iii) de te verwachten rol van bepaalde factoren (buitenschoolse taalaanbod zal bijvoorbeeld een grotere rol spelen in de bovenbouw), en (iv) de te verwachten resultaten (dat motivatie misschien minder groot is bij oudere leerlingen dan bij jongere leerlingen. In de tpo-pilot wordt echter niet verwacht dat scholen integraal beginnen, maar dat ze allemaal opbouwen vanaf groep 1/2. De focus van het testen van verschillende vaardigheden zal hierdoor op groep 1 t/m 5 komen liggen, waardoor resultaten aan het eind van de basisschool slechts voorspeld kunnen worden op basis van de resultaten uit de onderbouw. Kwalitatieve aspecten zoals organisatie en leerkrachteigenschappen zouden wel schoolbreed geïnventariseerd kunnen worden (feedback stuurgroep).

- Eenduidigheid is wenselijk over de opbrengsten aan het eind van de pilotstudie (groep 4/5), zodat de resultaten van de deelnemende scholen onderling vergelijkbaar zijn (feedback stuurgroep).

8. Referenties

- Baker, C. (2001). *Foundations of bilingual education and bilingualism*. Clevedon: Multilingual Matters.
- Bishop, D. (2003). Test for reception of grammar second edition (TROG-2). New York, NY: Pearson.
- Bostwick, R. M. (2001). Bilingual education of children in Japan: Year four of a partial immersion programme. In M. G. Noguchi & S. Fotos (Eds.), *Studies in Japanese bilingualism* (pp. 272-311). Clevedon: Multilingual Matters.
- Burmeister, P., & Pasternak, R. (2004). Früh und intensiv: Englische Immersion in der Grundschule am Beispiel der Claus-Rixen-Grundschule in Altenholz. *Fachverband Moderne Fremdsprachen fmf Landesverband Schleswig-Holstein, Mitteilungsblatt August 2004*, 24-30. Retrieved from www.fmks-online.de
- Cambridge ESOL. (2003). *Cambridge young learners English tests: Handbook*. Cambridge: University of Cambridge ESOL Examinations.
- Campbell, R. N., Grey, T. C., Rhodes N. C., & Snow, M. A. (1985). Foreign language learning in elementary schools: A comparison of three language programs. *The Modern Language Journal*, 69(1), 44-54. doi:10.1111/j.1540-4781.1985.tb02526.x
- Cenoz, J. (2005). English in bilingual programs in the Basque Country. *International Journal of the Sociology of Language*, 171, 41-56. doi:10.1515/ijsl.2005.2005.171.41
- Cenoz, J., & Lindsay, D. (1996). English in primary school: Teaching a third language to eight year olds in the Basque Country. *Cuadernos de Filología Inglesa*, 5 (1), 81-102. Retrieved from <http://dialnet.unirioja.es/>
- Cheng, L., Miao, L., Kirby, J. R., Qiang, H., & Wade-Woolley, L. (2010). English language immersion and students' academic achievement in English, Chinese and mathematics. *Evaluation & Research in Education*, 23 (3), 151-169. doi:10.1080/09500790.2010.489150
- Christian, D. (1996). Two-way immersion education: Students learning through two languages. *The Modern Language Journal*, 80 (1), 66-76. Retrieved from <http://www.jstor.org/>
- Cirino, P. T., Pollard-Durodola, S. D., Foorman, B. R., Carlson C. D., & Francis, D. J. (2007). Teacher characteristics, classroom instruction, and student literacy and language outcomes in bilingual kindergartners. *The Elementary School Journal*, 107 (4), 341-364. doi:10.1086/516668
- Corda, A., Phielix, C., & Krijnen, E. (2012). *Wat weten we over vroeg Engels op de basisschool? Effecten van vroeg vreemde talenonderwijs op de moedertaal en de leerprestaties in de vreemde taal: Review van de onderzoeksliteratuur*. Leiden: Expertisecentrum mvt (National Centre on Modern Languages).
- CTB/McGraw-Hill (n.d.). *LAS Links assessment for English language learners*. Monterey, CA: CTB & McGraw-Hill.
- Cummins, J. (2009). Bilingual and immersion programs. In M. H. Long & C. J. Doughty (Eds.), *Handbook of second language teaching* (pp. 159-181). London: Blackwell.
- De Courcy, M. C., & Burston, M. (2000). Learning mathematics through French in Australia. *Language and Education*, 14 (2), 75-95. doi:10.1080/09500780008666780
- De Graaff, R. (2013). Taal om te leren: Iedere docent een CLIL-docent. *Levende Talen Magazine*, 100(7). 7-11.
- De Rivière, A. (2004). Immersie-onderwijs bij jonge kinderen in het basisonderwijs van de Franse gemeenschap. In Vlaamse Onderwijsraad (Ed.), *Meer talen in de basisschool? Een verkenning* (pp. 15-30). Brussel: Vlaamse Onderwijsraad & Garant Uitgevers.
- Dobson, A., Pérez-Murillo, M. D., & Johnstone, R. (2010). *Bilingual education in Spain. Evaluation report*. Madrid: Ministry of Education & British Council.

- D'Ydewalle, G., & van de Poel, M. (1999). Incidental foreign language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, 28 (3), 227-244. doi:10.1023/A:1023202130625
- Dunn, L. M., & Dunn, D. M. (2007). Peabody Picture Vocabulary Test (PPVT-4). Bloomington, MN: Pearson.
- Dunn, L. M., Dunn, L. M., & Schlichting, L. (2005). Peabody Picture Vocabulary Test (PPVT-III-NL). Amsterdam: Pearson.
- Eiken. (n.d.). *Test in practical English proficiency*. Retrieved from <http://stepeiken.org/>
- Enever, J. (Ed.). (2011). ELLiE: Early Language Learning in Europe. Manchester: British Council.
- Europees Platform (2009). *Modellenboekje vvto Engels. Streefmodellen voor vvto Engels in het basisonderwijs*. Retrieved from <http://europeesplatform.nl/wp-content/uploads/2013/09/Modellenboekje-vvto-Engels-februari-2009.pdf>
- Eurydice (2006). Content and Language Integrated Learning (CLIL) at School in Europe. Brussels: Eurydice. Retrieved from http://ec.europa.eu/languages/documents/studies/clil-at-school-in-europe_en.pdf
- Frijns, C., Sierens, S., van Gorp, K., Sercu, L., van Avermaet, P., & Devlieger, M. (2011). 't is goe, juf, die spreekt mijn taal! Literatuurstudie praktijkgericht onderwijsonderzoek in opdracht van de Vlaamse OnderwijsRaad (Vlor). Brussel: Vlaamse Onderwijsraad.
- Gebauer, S. K., Zaunbauer, A. C. M., & Möller, J. (2012). Erstsprachliche Leistungsentwicklung im Immersionsunterricht: Vorteile trotz Unterrichts in einer Fremdsprache? *Zeitschrift für Pädagogische Psychologie*, 26 (3), 183-196. doi:10.1024/1010-0652/a000071#
- Genesee, F. (1985). Second language learning through immersion: A review of U.S. programs. *Review of Educational Research*, 55 (4), 541-561. doi:10.3102/00346543055004541
- Genesee, F. (1987). *Learning through two languages: Studies of immersion and bilingual education*. Rowley, MA: Newbury.
- Genesee, F. (2004). What do we know about bilingual education for majority language students? In T. K. Bhatia & W. Ritchie (Eds.), *Handbook of Bilingualism and Multiculturalism* (pp. 547-576). Malden, MA: Blackwell.
- Genesee, F. (2007). French immersion and at-risk students: A review of research evidence. *Canadian Modern Language Review*, 63 (5), 655-726. doi:10.3138/cmlr.63.5.655
- Genesee, F., Lindholm-Leary, K., Saunders, W., & Christian, D. (2005). English language learners in U.S. schools: An overview of research findings. *Journal of Education for Students at Risk*, 10 (4), 363-385. Retrieved from eric.ed.gov/
- Genesee, F., & Lindholm-Leary, K. (2013). Two case studies of content-based language education. *Journal of Immersion and Content-Based Language Education*, 1 (1), 3-33. doi:10.1075/jicb.1.1.02gen
- Goorhuis-Brouwer, S., & de Bot, K. (2010). Impact of early English language teaching on L1 and L2 development in children in Dutch schools. *International Journal of Bilingualism*, 14 (3), 289-302. doi:10.1177/1367006910367846
- Hilchey, M. D., & Klein, R. M. (2011). Are there bilingual advantages on non-linguistic interference tasks? Implications for plasticity of executive control processes. *Psychonomic Bulletin & Review*, 18 (4), 625-658. doi:10.3758/s13423-011-0116-7
- Jäppinen, A.-K. (2005). Thinking and content learning of mathematics and science as cognitional development in Content and Language Integrated Learning (CLIL): Teaching through a foreign language in Finland. *Language and Education*, 19 (2), 148-169. doi:10.1080/09500780508668671
- Jessner, U. & Cenoz, J. (2007). Teaching English as a third language. In J. Cummins & C. Davison (Eds.), *The International Handbook on English Language Teaching* (pp. 155-167). Berlin/New York: Springer. doi:10.1007/978-0-387-46301-8_12
- Johnson, R. K. & Swain, M. (1997). *Immersion education: International perspectives*. Cambridge, UK: Cambridge University Press.

- Kersten, K., Frey, E., & Hähnert, A. (2009). *ELIAS: Early Language and Intercultural Acquisition Studies*. Magdeburg: EU Education, Audiovisual and Culture Executive Agency.
- Knell, E., Haivan, Q., Miao, P., Yanping, C., Siegel, L. S., Lin, Z., & Wei, Z. (2007). Early English Immersion and Literacy in Xi'an China. *The Modern Language Journal*, 91 (3), 395-417. doi:10.1111/j.1540-4781.2007.00586.x
- Kuska, S. K., Zaunbauer, A. C. M., & Möller, J. (2010). Sind Immersionsschüler wirklich leistungsstärker? Ein Lernexperiment. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42 (3), 143-153. doi:10.1026/0049-8637/000015
- Lasagabaster, D. (2000). Three Languages and Three Linguistic Models in the Basque Educational System. In J. Cenoz & U. Jessner (Eds.), *English in Europe: The Acquisition of a Third Language* (pp. 179-197). Clevedon: Multilingual Matters.
- Lasagabaster, D. (2001). Bilingualism, Immersion Programmes and Language Learning in the Basque Country. *Journal of Multilingual and Multicultural Development*, 22 (5), 401-425. doi:10.1080/01434630108666443
- Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, 10 (1), 105-129. doi:10.1080/14790718.2012.679275
- Lorenzo, F., Casal, S., & Moore, P. (2009). 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. doi:10.1093/applin/amp041
- Maas, E. (2013). *An international perspective of partial immersion in primary education. Effect studies and educational models* (Unpublished term paper). Utrecht: Utrecht University.
- Massler, U., & Stötz, D. (2013). Assessment in bilingual classrooms. In D. Elsner & J.-U. Keßler (Eds.), *Bilingual Education in Primary School* (pp. 72-92). Tübingen: Narr Verlag.
- Massler, U. (2012). Primary CLIL and its stakeholders: What children, parents and teachers think of the potential merits and pitfalls of CLIL modules in primary teaching. *International CLIL Research Journal*, 1 (4). Retrieved from <http://www.icrj.eu/>
- Mehisto, P., & Asser, H. (2007). Stakeholder perspectives: CLIL programme management in Estonia. *The International Journal of Bilingual Education and Bilingualism*, 10 (5), 683-701. doi:10.2167/beb466.0
- Mellow, J. D., Reeder, K., & Forster, E. (1996). Using time-series research designs to investigate the effects of instruction on SLA. *Studies in Second Language Acquisition*, 18 (3), 325-350. doi:10.1017/S0272263100015059
- Merisuo-Storm, T. (2007). Pupils' attitudes towards foreign-language learning and the development of literacy skills in bilingual education. *Teaching Teacher Education*, 23 (2), 226-235. doi:10.1016/j.tate.2006.04.024
- MLA (Modern Language Association). (1963). *MLA cooperative foreign language tests*. Princeton, NJ: Educational Testing Service.
- Montecel, M. R., & Cortez, J. D. (2002). Successful bilingual education programs: Development and the dissemination of criteria to identify promising and exemplary practices in bilingual education at the national level. *Bilingual Research Journal*, 26 (1), 1-22. doi:10.1080/15235882.2002.10668696
- Morales, J., Calvo, A., & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, 114 (2), 187-202. doi:10.1016/j.jecp.2012.09.002
- Muñoz, C. (2006). *Age and the rate of foreign language learning*. Clevedon: Multilingual Matters.
- Nikolov, M., & Mihaljević Djigunović, J. (2011). All shades of every color: An overview of early teaching and learning of foreign languages. *Annual Review of Applied Linguistics*, 31, 95-119. doi:10.1017/S0267190511000183
- Ojima, S., Matsuba-Kurita, H., Nakamura, N., Hoshino, T., & Hagiwara, H. (2011). Age and amount of exposure to a foreign language during childhood: Behavioral and ERP data on the semantic

- comprehension of spoken English by Japanese children. *Neuroscience Research*, 70 (2), 197-205. doi:10.1016/j.neures.2011.01.018
- Olaziregi, I., & Sierra, J. (1988). *Galbahe-E2 hizkuntza testa. Gida liburua*. Vitoria-Gasteiz: Central publications of the Basque government.
- Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, 113 (4), 535-551. doi:10.1016/j.jecp.2012.06.013
- Reder, F., Daigle, D., & Demont, E. (2013). Metalinguistic development in French learners enrolled in an immersion programme: A longitudinal study. *European Journal of Developmental Psychology*, 10 (4), 476-494. doi:10.1080/17405629.2012.734132
- Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Pappert Lercari, L., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, 42 (8), 1029-1040. doi:10.1016/j.neuropsychologia.2003.12.012
- Riemersma, A., & de Jong, S. (2007). *Frisian: The Frisian language in education in the Netherlands* (4th ed.). Leeuwarden/Ljouwert: Mercator-Education.
- Seikkula-Leino, J. (2007). CLIL learning: Achievement levels and affective factors. *Language and Education*, 21 (4), 328-341. doi:10.2167/le635.0
- Serra, C. (2007). Assessing CLIL at primary school: A longitudinal study. *The International Journal of Bilingual Education and Bilingualism*, 10 (5), 582-602. doi:10.2167/beb461.0
- Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology*, 51 (3), 300-304. doi:10.1037/h0020586
- Skolverket. (2010). *Undervisning på Engelska. Utvärdering av en försöksverksamhet i grundskolan*. [Teaching in English. Evaluation of a pilot project in primary school]. Stockholm: Skolverket.
- Thomas, W. P., Collier, V. P., & Abbott, M. (1993). Academic achievement through Japanese, Spanish, or French: The first two years of partial immersion. *The Modern Language Journal*, 77(2), 170-179. doi:10.1111/j.1540-4781.1993.tb01961.x
- Turnbull, M., Hart, D., & Lapkin, S. (2003). Grade 6 French immersion students' performance on large-scale reading, writing, and mathematics tests: Building explanations. *The Alberta Journal of Educational Research*, 49 (1), 6-23. Retrieved from <http://ajer.synergiesprairies.ca/>
- Unsworth, S., Persson, L., Prins, T., & de Bot, K. (in druk, 2014). An investigation of factors affecting early foreign language learning in the Netherlands. *Applied Linguistics*.
- Van Ruijven, B., & Ytsma, J. (2008). *Trijetalige skoalle yn Fryslân. Onderzoek naar de opbrengsten van het drietalig onderwijsmodel in Fryslân*. Ljouwert: Fryske Akademy.
- Wilkinson, G. S. (1993). *Wide Range Achievement Test - Revised (WRAT-3)*. Chicago, IL: Riverside Publishing Company.
- Wode, H. (2004). *Frühes Fremdsprachenlernen: Englisch ab Kita und Grundschule. Warum? Wie? Was bringt's?* Kiel: Verein für frühe Mehrsprachigkeit an Kindertageseinrichtungen und Schulen FMKS e.V.
- Woodcock, R. (1991) *Woodcock Language Proficiency Battery-Revise (WLPB-R)*. Chicago, IL: Riverside Publishing Company.
- Ytsma, J., & Beetsma, D. (2001). Three lingual primary school in Fryslân. *MOER*, 32 (1), 3-10. Retrieved from www.taalunieversum.org/moer/
- Zaubauer, A. C. M., & Möller, J. (2007). Schulleistungen monolingual und immersiv unterrichteter Kinder am Ende des ersten Schuljahres. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 39 (3), 141-153. doi:10.1026/0049-8637.39.3.141
- Zaubauer, A. C. M., & Möller, J. (2010). Schulleistungsentwicklung immersiv unterrichteter Grundschüler in den ersten zwei Schuljahren. *Psychologie in Erziehung und Unterricht*, 57(1), 30-45. doi:10.2378/peu2010.art03d

Appendix I – Summaries of expert interviews

Kristin Kersten (Universität Hildesheim, Germany)

Email: kristin.kersten@uni-hildesheim.de

Assistant Professor of Foreign Language Education and Second Language Acquisition; Department of English Language and Literature.

Main research areas: second language acquisition, early language teaching and learning, immersion education, bilingual subject teaching, CLIL.

Interview date: 12/12/2013

General introduction

In Germany, primary education consists of two hours of L2 teaching per week by ministerial decree. All federal states teach English at grade three, but some already start at grade one.

There is much enthusiasm in Germany concerning immersion programmes. The reason for many schools to start an (early partial) immersion programme is to promote the use of English as this language is gaining more and more influence worldwide. The parents want their children to take part in these kinds of programmes as they are often convinced that a greater skill set will be obtained during their child's education. Furthermore, 'the younger, the better' is a widespread idea amongst parents and the intensive early start is generally believed to be a good educational method. New teachers are also very enthusiastic, even though these new programmes are very challenging. However, Germany has also come across various difficulties when developing and teaching through immersion programmes; these concern teacher training, measuring proficiency and the pressure parents exert on their children.

TPO RESEARCH QUESTIONS

1. Which international examples are available that resemble our pilot study, both in organization and content of the programme?

Even though there are between 700-800 schools in Germany offering immersion programmes, there is not much information on this subject. KK indicated that she, therefore, could not tell us much about this. Much work is to be done in terms of centralization. The only thing she knew was what has been said by the Ministry, which is that every primary school should teach at least two hours of L2 each week. Some schools already start teaching the L2 during grade one.

Amount of L2

Most schools prefer 70% of the lessons to be taught in the L2. The other 30% is taught in the L1.

Best amount of L2

The viewpoint of the federal administration differs from the viewpoint the researchers and schools often have. The administration sets a maximum of 20% of the lesson time to be taught through the L2. According to KK, this will result in pupils not acquiring enough knowledge to fully comprehend the subject matter, and because of this, the results will, in the end, not be satisfactory. She recommends at least 70% of the total lesson time to be taught in the L2.

2. Which evaluation instruments are available for use in qualitative and quantitative evaluation research in bilingual primary education?

KK again indicates that nothing in Germany is centralized, so she can only tell us what she knows from her own research and experience.

L1/L2 assessment

To measure the proficiency in the L1, the SETK is often used. This test is used for all children, including children with migration backgrounds and the youngest children. For both groups, the results are debatable, but the test is used anyway as there is no better alternative at this point. L2 proficiency can be measured in different ways and by different methods. In the ELIAS project, KK used a new test which she and her colleagues designed: the ELIAS Grammar Test. Nowadays, many researchers use the CEFR in order to describe the proficiency levels. However, the criteria are open to personal interpretation which makes this less reliable as assessment tool.

Due to this lack of centralization, many schools do not consistently measure their students' proficiency and because of this, there is not much information on the subject. On the basis of her own experience, KK advised using the BPVS to measure vocabulary and the second version of the ELIAS Grammar Test to test grammar, both assess receptive proficiency. She did however add a note of caution as the norms for EAL (English as an Additional Language) students are no longer available in the third (most current) edition of the BPVS. This is of course important as EAL students need different criteria because of their migration background.

KK also emphasizes the importance of designing reliable tests in order to evaluate the immersion programmes.

3. What results have been reported concerning the development of children's first (L1) and second (L2) languages when involved in such programmes?

Selection procedure

In general, the schools themselves decide on whether to accept each student to their immersion programme or to have a selection procedure.

Development of L1/L2

As KK said before, there is no common framework along which the proficiency of the students can be reliably assessed. Because of this, no results have been reported yet.

4. What are the relevant characteristics and didactic skills of (early partial immersion) teachers?

The teachers themselves come across many difficulties as there are many educational programmes for teachers wanting to work in regular education, but not many programmes which focus on immersion programmes. Because of this deficit, many teachers wanting to work in the field of immersion education get this job without being fully qualified to teach in an immersion classroom. As a result, many teachers find themselves feeling insecure about their capabilities. Some schools do offer immersion teacher programmes, but again, nothing is centralized.

There are however many ongoing developments in this area. Many universities and schools are broadening their horizon and now offer dual-language teaching programmes. These programmes also oblige teachers-to-be to go abroad for half a year and hereby improve their proficiency. In this way, the problem of insecurity will probably be solved.

In CLIL, didactic principles are very important and should therefore be taught in immersion teaching education. These principles include task-based learning, focus on form and learner-based learning.

Teaching materials

Teachers should get enough time to do research and design their own materials. Standardized materials are barely available, one of the few materials is "Take Off". Teachers (especially those who have just started their jobs) have a hard time finding and compiling material. They should be given the opportunity to produce their own materials and evaluate these materials after using it for the first couple of times. Furthermore, they should be given time (especially when they are not fully credentialed to teach in immersion programmes) to research the programme and find the best ways to approach this educational method. There is not much communication between the schools offering bilingual programs in Germany, but sometimes teaching materials are shared among teachers.

School preferences

Many schools prefer to employ near-native speakers for two reasons. The first is that intercultural issues are known to be quite problematic in education and so a native German teacher will probably fit in well with the current staff. The second is that German schools rarely accept qualifications from abroad and international teachers therefore often have a hard time finding a job in education in Germany, even if they are fully qualified.

Late starters in an ongoing programme

There are no data about this. Some schools accept late starters, other schools do not.

Subjects

Schools are free to choose which subjects to teach in the L1 or L2. KK remarked that maths is often not taught in the L2 as it is sometimes considered 'high-risk' (as the children might have too much difficulty grasping the meaning of the words and the theory itself). However, she thinks of maths as a very suitable subject for being taught in the L2 indeed. General studies is often taught in the L2, as it is very suitable to use methods like hands-on experience. Furthermore, it is often self-explanatory. Teachers are free to design their own materials and can hereby shape their lessons themselves. KK advised against teaching Physical Education in the L2, as Germany's experience with this has not been very good.

When schools want to start their immersion programme, they sometimes start by teaching one or two subjects in the L2. This often fully depends on the availability of teachers, as not many are able to teach their subject through the L2.

5. How are the educational programme and the development of the children's different languages perceived by the students, teachers and parents?

KK advised to keep informing the parents at all times, as most are very interested in their child's bilingual education but do have their initial concerns (for instance about the consequences for the L1). Their involvement can be facilitated by inviting them for parental meetings, where they can hear about their child's progress and the programme itself. The most difficult are the parents who exert much pressure on their child, who are often academics themselves.

The students are said to have initial concerns as well, which turn into enthusiasm after having seen pictures and videos of the bilingual education. Initial concerns (like: 'Will I be able to do this?') are quickly proven to be unnecessary.

Marianne Nikolov (University of Pécs, Hungary)

Email: nikolov.marianne@pte.hu

Professor of English Applied Linguistics, Head of Department; Department of English Applied Linguistics.

Main research areas: English as a foreign language, early learning and teaching, individual differences, teacher education, language policy.

Interview date: 12/12/2013

General introduction

In Hungary the term used for (early partial) immersion education is dual language instruction. Dual language instruction has gained much influence over the last years, as the number of schools offering this type of education has increased. Typically, the schools in Hungary offer eight years of primary education starting from the age of six years old.

The reason why parents enrol their child in a dual language instruction class is the prestige involved. In addition, parents also see the advantages of second language learning. In the past, many German schools were crowded with children wanting to learn a second language and this tradition is still alive.

Schools in Hungary usually have two main reasons for introducing dual language instruction. The first is the great demand from parents and the second is the extra financial support they receive when they launch such a programme.

Selection procedure

In the past, children in Hungary used to go to school in the district they lived in. However, nowadays parents select schools of their own preference. This results in a very strict self-selection, related to the SES of the parents, with more high-ability students in immersion programmes.

Approach/differences with regular education

The efficiency of immersion programmes has always been a big question among parents. Hardly any scientific research has been conducted on this topic. One study was conducted by a doctoral student of MN last year. She published her dissertation in Hungarian and one paper in German (http://www.fachportal-paedagogik.de/fis_bildung/suche/fis_set.html?FId=1000427).

In her study, she interviewed parents, policy makers, as well as teachers in regular and immersion programmes, and she also observed classes. Her main finding was that teachers in both regular and immersion education use the same teaching techniques (with translation as the main strategy) and do the exact same things in both types of education.

Politics

The last important factor regarding immersion education MN wants to emphasize is the influence of the Hungarian government. Hungary recently elected a very strict, right-wing government. They introduced new requirements in terms of achievement targets concerning immersion education which are now laid out in law. However, these requirements have been specified without any scientific evidence. They describe proficiency requirements based on the CEFR and are, according to MN, overwhelmingly optimistic about the proficiency levels that should be achieved at the end of primary education (and also in secondary education). The first findings will come out next year when the first cohort of these children reaches the end of primary education. The law states that if schools fail to achieve the required levels by more than 60% of the students, they will lose the right to teach through immersion education.

In addition, decisions about the start of new programmes are now made by the ministry. Schools used to have the right to decide about new developments themselves, but this right has now been centralized by the current government.

TPO RESEARCH QUESTIONS

1. Which international examples are available that resemble our pilot study, both in organization and content of the programme?

As a starting point for the interview we sent all interviewees an overview of studies on immersion education that we analyzed to find an answer to our research questions. We asked the interviewees if they could suggest other studies that we had not included yet. MN emphasized that there has not been much scientific research in the field of immersion education in Hungary, let alone publications in English. There is one article written in English, describing a project on German dual language schools, that is currently in press that she wants to share with us:

Medgyes, P. & Nikolov, M. (2013). A Country in Focus. Research in foreign language education in Hungary (2006-2012). *Language Teaching*.

2. Which evaluation instruments are available for use in qualitative and quantitative evaluation research in bilingual primary education?

Since the government has just recently introduced the new specific immersion education requirements, there is no evidence yet about the results and about how realistic these requirements are. Official or standardized tests do not exist. The only information about student assessment so far comes from teachers who attend or have attended courses in immersion education teaching. They generally tend to rely on testing the student's reading and writing proficiency. Hardly any oral proficiency is tested, because teachers aren't familiar with this assessment or find it too time consuming.

Assessment seems to be related to the teaching materials, but there is not much data available on where the teaching and test materials originate from. This is a topic for further research. According to MN the best tests are in line with the curriculum, but this is not always feasible.

The Hungarian core curriculum states what level children have to achieve by which grades, but no systematic assessment exists throughout primary school, except for the school leaving exam. A framework like the CEFR is being looked at, but this is not a tool for assessment. The first language assessment will take place this year. MN currently runs a project on diagnostic assessment of young learners of English which will become available as a guideline for teachers to use online. She emphasizes that the key stakeholders are teachers and that the teachers' competence, qualifications, motivation, and positive feedback are decisive and can assure the success of immersion programmes.

3. What results have been reported concerning the development of children's first (L1) and second (L2) languages when involved in such programmes?

Streaming and abilities

In Hungary students are divided by ability groups, something that is not initiated by the government but by parents and schools. The immersion programmes are seen as more prestigious than monolingual programmes and therefore schools have to cope with an enormous number of applications.

MN is strictly against the streaming of students into those two programmes, because it causes a lot of harm to both groups of students. Students in an immersion programme often feel anxious because they keep comparing themselves to the more talented students in the programme. Students in the regular programme know from day one that they belong to the lower achieving group and will therefore achieve less than they possibly could. It also causes a bias amongst teachers, because they will not expect as much from students in the lower achieving groups than from students in the immersion groups.

Transition from primary to secondary

This has always been an enormous challenge. In a study from 2007 whereby every student in year nine (the first year of secondary education) was involved, two thirds of the students started their secondary language education in English at a lower level (or even as course beginners) than what they had achieved in primary school. Secondary schools often do not continue what the students have already achieved in primary school, which causes a loss of the previously done work, takes away the students' previous experiences, and moreover causes demotivation.

Native vs. minority languages

Besides the English language, which is nowadays seen as the most prestigious language, language teaching and assessment can be divided up into three categories:

- Heritage languages. German used to be one of those prestigious heritage languages, but it is now taught less and less because English has taken over its function.
- Ethnic languages. Languages spoken by minority groups, e.g. Croatian. Parents who think this language will be useful over time send their children to schools where they teach through ethnic languages. These languages are assessed in a traditional way, with no standardized tests, and mostly assessed on children's motivation and behavior.
- Romani or 'gypsy' languages. No 'gypsy' schools or language programmes exist. It would be interesting to assess the language development of this minority population, but as the law does not permit any data collection on ethnicity it remains a sort of non-existent issue, as MN explains.

4. What are the relevant characteristics and didactic skills of (early partial) immersion teachers?

Not much research has been published on this topic either. Most of the information that MN has is based on some studies in Hungarian and on her own experience.

Approach

As mentioned earlier, and based on the results of the study published in German, most teachers were trained as ethnic minority language teachers. Teachers in immersion education are the same teachers who used to teach German (or English) as a subject. Teachers tend to use the exact same teaching strategies as they did in regular teaching programmes, whereby often extremely traditional (grammar-translation and not a child-centered approach) and inappropriate classroom practices are applied in immersion classes.

MN mentions the PhD research on dual language education in Hungary conducted early 1990s by Patricia Duff. She was one of the few people who noticed that the main challenge in the education system was the socialization process and the one who pointed out that a change in teaching and learning from the traditional teaching model had to take place.

L2 proficiency/native speakers

Native speakers of English are very rare in Hungary, mainly due to the fact that Hungary is not perceived as a very attractive country to start teaching. Schools would be happy to receive more native speakers, but MN only mentions some examples of opportunistic (native English-speaking) teachers who came to Hungary for a short period with, for example, a Comenius programme exchange, but a national language programme cannot be built based on a few specific cases.

The acceptance of foreign teaching qualifications is also an issue, as teachers' certificates are not compatible or comparable to those in Hungary. Furthermore, foreign teachers often have difficulties in Hungary as they are not always familiar with the Hungarian culture and curriculum.

Teaching degrees in Hungary can roughly be divided into three groups:

1. Lower primary teaching degree (four-year teacher training) with an additional specialization in a language. These teachers are very enthusiastic about teaching to children, because they are trained to become a primary level teacher; however, their academic/language proficiency is modest.
2. College level language major degree. Trained as teacher of a language. These teachers have extensive pedagogical training, but often had aimed for university level, but did not get in.
3. University degree. These are the best possible English teachers, but not necessarily the best of the population; those would have aimed for other professions (e.g., medical school, IT specialist), because the profession of teacher does not have a high reputation. These teachers have a relatively solid background and have received strong TEFL training. Because MN focuses on young learners in her university programmes, many of her students show interest in this topic too now, but this is not necessarily the case for all university educated teachers.

Subject competence in L2

This is a bit of a controversy, because if a teacher is supposed to teach a subject through L2, s/he should also obtain a major degree in that subject. However, this is more of a topic relevant to secondary education, because the subjects taught in primary school differ from those taught to future teachers in college or university.

Subjects

Schools decide themselves which national curriculum courses to offer through the second language. No data are currently available about what subjects are taught in the L2. According to MN, the most common subjects for L2 instruction are music, PE, arts, environmental studies, and maybe mathematics. The new project MN is working on now also focuses on the subjects taught through English.

Best subjects

MN said any subject is suitable for immersion education. The decisive factor should be teacher competence and a supportive context, so that children can easily pick up the content.

Teaching materials

The development of new materials is definitely a problem that schools have to tackle. Initially, teachers have mainly used existing course books translated into the L2. Nowadays, teachers (especially in the university programmes MN runs) tend to be very enthusiastic about developing new course materials, but they also still rely on materials that are readily available from England or the US.

(Inter)national exchange of ideas

Secondary schools might cooperate with schools in Germany, but schools are traditionally never encouraged to exchange ideas about experiences and materials. In the past, courses were centralized and one book was used per course. In more recent years, Hungarian schools have been able to choose their own materials, but recently the government has centralized the education system again, so schools need to implement the method of one book per course again. According to MN, this shows that the government is not supportive of the immersion classrooms.

5. How are the educational programme and the development of the children's different languages perceived by the students, teachers and parents?

Some of the MA-students MN is supervising are examining parents' attitude or teacher motivation in their theses, but their studies only concern small-scale research.

Parents

Parents who care about the teaching programme tend to become very critical over the years and express their thoughts about what they think is missing in school.

MN describes three types of parents who deal with their child's education in a very different way.

- Well-educated parents who care about the education their children receives, but also tend to become very critical towards the programme and even embarrass teachers by trying to outsmart the teachers. These are usually professional or business people.
- 'Normal' parents. They see the benefits and the importance of education and also support the programme their child participates in, and they also believe in the schools' approach and trust its teachers.
- Lower-educated, low SES parents. They don't see the benefits of education and do not think education should be changed to immersion education or to any other model. They have no aspirations concerning their child's ambitions or future.

Teachers

Teachers MN works with in university degree programmes are more open towards immersion education, they are provided with better courses and new materials, and have better opportunities to work on their professional development. However, the teaching profession is still one of the worst paid professions in Hungary with very low prestige. The main reason for a student to become a primary school teacher is because they love working with children.

School system

The school system enhances the differences between children in their abilities and socioeconomic status when entering the school. The strong streaming selection makes it almost impossible for children with a lower socioeconomic status to enroll in immersion programmes.

Students

Students in general are very enthusiastic about learning new languages. However, the Hungarian school system streams them into ability groups (essentially according to their SES) and doesn't provide equal opportunities for children's education. MN totally disapproves of this streaming and emphasizes that all children should be given equal educational opportunities and should be equally challenged.

Conclusion

MN would be very interested in knowing how far this project develops, what can be set up and what results will be found. She would be very interested in being updated on the progress of the project.

Pilar Medrano (Ministry of Education, Madrid, Spain)

Email: pilar.medrano@mecd.es

Project Manager & Educational Adviser; Institute of Teacher Education, Educational Research and Innovation.

Main research areas: bilingual education, teacher advisory.

Interview date: 12/12/2013

General introduction

In Spain a clear distinction can be made between (early partial) immersion and regular (EFL) education. The most notable difference between the two types of education can be found in a) the number of hours taught in the L2, b) the starting age, and c) content taught through the L2.

Spanish schools with immersion programmes usually devote between 30% and 50% of their instruction to the L2, which is English in most cases. There are schools in autonomous regions with another minority language as their L1 besides the national language, Spanish. In these schools, therefore, English is taught as an L3 (but still the first *foreign* language). The number of schools starting immersion programmes is growing fast, but schools still differ a lot from one another in the number of hours taught in the L2, the subjects taught in the L2, the starting age (some start at the age of 3/preschool level, some at the age of 6/grade 1), and even in didactic approach.

In Spain an extensive research on immersion education was conducted by the Ministry of Education and the British Council: the Bilingual Education Project (BEP). Together, the Ministry and the British Council developed a new curriculum specifically for all schools participating in this study, something that differed from other schools that taught parts of the national curriculum in English. The BEP was the first national project to start in state schools and received great parental support. Because of this, the Ministry decided this was a good approach for other schools to follow in order to raise the national level of English. The results from the BEP encouraged local authorities to start similar programmes and the number of schools starting similar programmes in various regions in Spain is growing very fast. As an example, the Madrid region started in 2004 with 26 primary schools and now there are 318 primary schools and 91 secondary schools with bilingual education programmes.

TPO RESEARCH QUESTIONS

1. Which international examples are available that resemble our pilot study, both in organization and content of the programme?

As a starting point for the interview we sent all interviewees an overview of studies on immersion education that we have analyzed to find an answer to our research questions. We asked the interviewee if they could suggest other studies that we hadn't included yet. PM wondered why we did not include any examples from the UK. In fact, we did find some studies from Scotland, but they did not describe any results, which is why they were not included. PM suggested contacting Richard Johnstone for more information about the results which were obtained from immersion education in the UK.

Amount of L2

As explained in the introduction, most Spanish schools devote between 30% and 50% of their instructional time to the L2. In the BEP approximately 40% of the instruction was taught through English.

Best amount of L2

PM pointed out that the best amount of L2 instruction depends a lot on the context (e.g., language situation and language distance, out-of-school exposure), but she suggested starting with at least 40% if a reasonably high L2 proficiency level is one of the aims of immersion education.

Further research

- Low attainers
- Students with special needs
- Role of the school and effect of the school's approach

2. Which evaluation instruments are available for use in qualitative and quantitative evaluation research in bilingual primary education?

L1/L2 simultaneously

A shift has taken place in Spain in terms of the starting age of teaching reading and writing in the L1 and L2. Initially, students started the reading and writing process of the L2 only after they had acquired a basic knowledge of reading and writing in their L1. When schools/teachers realized that teaching of reading and writing in L1 and L2 simultaneously did not negatively affect the L1, schools started early, around the age of 5, with reading and writing in both languages.

The BEP was the first national study to start with both L1 and English at the same age and because of these initial concerns, they decided to watch the results closely. Slowly, schools and teachers started to realize that simultaneously teaching reading and writing in L1 and L2 did not negatively affect the L1. Moreover, research has shown that teaching through the L2 does not negatively affect the L1 development.

PM pointed out that there are many phonetic differences between English and Spanish. The earlier children assimilate this, the easier it will be for them to learn it, to produce it and to become independent learners. However, it might be different in other contexts with other languages pairs.

L1/L2 assessment

National tests in Spain assess the L1 development throughout primary school (after each cycle, at the age of 8/10/12). Mathematics is assessed by using national tests during primary education. The L2 development is not usually measured with an official test, but is based on the teachers' own assessment of the students' L2 proficiency and development. Therefore, no comparison with other schools is available.

Secondary schools were also included in BEP, so that the childrens' bilingual education could continue after their primary education. The students who had been educated through immersion primary education took the Cambridge IGCSE examinations in secondary school in L1, L2, and various other school subjects they were taught in English. It was used as an instrument to show that the aims of the project had been achieved.

The BEP report online supplement gives an overview of examples of internal tests which were designed by experts for the British Council project schools to get an idea of the student's level at the end of primary school. Examples of class observations as well as student interviews are also available. These are however not yet available as proper evaluation materials, due to staff and budget limitations.

Minority languages/areas

Education is a regional responsibility and the development of projects, programmes, materials, and assessment mainly depends on the context. Apart from schools in the BEP, other regions and schools use their own instruments for L2 assessment. In some schools the Cambridge ESOL exams (for young learners) are used for L2 assessment. This wide variety means that no valid comparison between different regions is possible. The approach to immersion education in bilingual and trilingual schools differs from one another. Especially regions with a minority language, such as the Basque Country, use their own materials designed for their trilingual language situation, not only focussing on English as the first foreign language, but also on the L1 Basque and L2 Spanish.

Subject assessment

In the BEP both primary and secondary schools were included. As noted above, students in the fourth year of secondary education, who had been taught through immersion education from primary school onwards, took the Cambridge IGCSE in L1, L2, and also various subjects. The successful performance on this test showed that students were able to tackle subject matter taught in English successfully.

The IGCSE is designed for 14-16 year old students and is thus not suitable for primary school students. A suggestion from PM was the Cambridge Checkpoint test for young learners. A test they haven't used themselves because of economic reasons, but it would be an appropriate test to assess primary school students' subject knowledge.

3. What results have been reported concerning the development of children's first (L1) and second (L2) languages when involved in such programmes?

Selection procedure

The BEP was based on a whole-school approach whereby every child in the project schools could participate in the immersion programme. Given that no data on this topic are available yet, it is difficult to provide an answer to the question of selection procedures. In general, it depends on the school, and some schools have a dual track programme and might not include every student when the project is initiated.

CEFR

The CEFR is being considered in Spain as objectives for teaching, but the level of proficiency for L2 pupils has not been established yet. At the moment more attention is paid to the teachers and their education.

Transition from primary to secondary

PM remarks much work is to be done in this field. In the BEP, an agreement was reached with associated secondary schools to continue the bilingual programmes, but in general this topic is not yet really a subject of discussion in Spain. Furthermore, most students have not reached secondary education yet.

4. What are the relevant characteristics and didactic skills of (early partial) immersion teachers?

Approach

With a wide variety of teachers available, the BEP encouraged a more hands-on approach as opposed to the more typical EFL teaching approach. Because of the fact that the number of schools offering an immersion programme is increasing, the demand for well-educated teachers also increases. However, at the moment, the language competence of teachers is not always sufficient to teach in immersion

programmes. They are used to and were trained to teach in a typical EFL setting and have to get used to the more hands-on approach (with for instance storytelling in the L2 and singing songs) and early teaching of reading and writing.

L2 proficiency/native speakers

The minimum proficiency level for teachers to teach in immersion programmes is B2. Many teachers are taking extra courses to improve their proficiency and reach a higher level.

Most primary school teachers are specialist teachers with an L2 degree. There is a preference for native speakers, especially among parents, but this is not feasible in every school. In many schools, specialist teachers work with the support of a native language assistant from an English-speaking country. In the BEP, there were two types of teachers available per school: qualified native speakers and qualified specialist teachers.

Most important though, PM pointed out, is the teachers' language competence and their didactic skills of knowing how to teach instead of having native language speakers in every school.

Subject competence in L2

Surprisingly, the content itself has not shown to be the greatest obstacle in implementing immersion education in Spain. The most important factor is the confidence in teaching through a foreign language. Most teachers are specialist in a certain topic, but the change from EFL to an immersion programme has proven to be quite difficult for them.

Late starters in an ongoing programme

There is not much data available on this subject. PM remarked that this should be a topic for more research.

Subjects

In the BEP, science, history, and geography were taught in English and together with arts, these are also the most common subjects to teach using the L2. In general, schools are free to decide which subjects they teach through the L2, but these choices basically depend on the availability of teachers and their subject and language competence.

Best subjects

In Spain there is a part of the curriculum called 'knowledge of the world', which includes science, geography and history. This is the most popular subject to teach in the L2 and PM also thinks these are indeed good subjects to teach in the L2, although she points out that didactic skills and teacher competence are once again decisive. Success also depends on the confidence of the school and the teachers about their didactic skills. Some are concerned that if the children learn a certain subject through English, they will not be able to talk about the subject in Spanish. This is something that has been proven to be untrue, but these initial worries need to be dealt with. For this reason, subjects are often taught 50-50 in the L1 and L2.

Teaching materials

When the BEP started six years ago, only EFL materials were available, so teachers had to work on new materials. They were usually enthusiastic about developing their own materials, which they did by:

- Using the internet
- Adopting materials from the UK.
- Combining different kind of sources to create materials for a certain topic

The fact that the BEP students follow a Spanish-English integrated curriculum means that there are no suitable textbooks, and therefore teachers are very clear that they must develop their materials to suit their specific needs.

Translation of the national curriculum is also widely applied in Spain. Many publishers translate the original Spanish textbooks for a certain subject into English, so the original content is maintained (which makes it easier to reach the goals of the curriculum) and the focus on the specific (Spanish) context is kept as well.

(Inter)national exchange of ideas

An official teacher network does not exist, but some schools, and particularly teachers, exchange ideas when they meet at courses or trainings. They usually exchange their own experiences with immersion teaching but are cautious about sharing their materials, because they might be unsure about the quality of their materials, although this is usually quite good.

Schools have some contacts with schools abroad to get involved in European projects, for instance through the Comenius projects. They have contacts with both British schools as well as with schools offering immersion programmes from other countries. They do not have any contacts through international school networks, but PM thought that this was a good idea and could be a very useful development.

5. How are the educational programme and the development of the children's different languages perceived by the students, teachers and parents?

Parents

In general parents show a very positive and supportive attitude. There is a huge demand in Spain from parents for immersion teaching and education to improve the English language skills of their child, as the general EFL method is not perceived as being very successful. At the same time, the parents do have their concerns and issues, for instance that children might lose their proficiency in their L1 and that the immersion programme will affect the academic level of certain subjects that are taught through the L2.

PM pointed out that parental involvement is an important topic for future research, especially for parents who do not speak English themselves and are not able to help their children with the L2 acquisition.

Assessment

The BEP used questionnaires to assess parents' perceptions of immersion education. Examples of questionnaires for parents with children in primary and secondary education are available in the BEP online supplement.

Teachers

Initially, some teachers reacted somehow negatively, because they were afraid of not being able to teach through the L2. Once they were involved in immersion education, had received training, and had been assured that what they did was correct, and when they started seeing the results, their attitude towards immersion education changed and they did not want to switch back to teaching EFL anymore. Teachers in immersion programmes also find it much more rewarding than regular programmes from a professional point of view. Not only do students' results improve, but the

increased motivation among students makes the programme much more rewarding and also increases the teachers' level of motivation.

In the beginning, teachers showed initial concerns about the start of the reading and writing process in the L2 at a very young age. Within the BEP project, the project administration tried not to push it and just see what would happen, which resulted in a more positive attitude when the teachers realized themselves it did not affect the L1 negatively. Another issue schools and teachers deal with is how to involve parents and how to stimulate them in supporting their children.

Assessment

As noted above, the BEP used questionnaires to assess teachers' perceptions of immersion education. Examples of questionnaires for teachers are available in the BEP online supplement. In addition, class observations were conducted.

Students

In general, students show a very positive attitude towards immersion education. They enjoy learning a new language and they like the new, more practical and hands-on approach of teaching. They are also positive about the benefits of English for the future and for their professional career. Young children even feel superior to their older siblings, because they know their English is better.

Assessment

In the BEP, only upper primary students (10 to 11 year olds) were asked about their perceptions by means of a questionnaire. Lower primary schools students were asked about their opinion via interviews.

Conclusion

PM would be delighted to hear more about the progress of this project and she said it would be great to receive some updates. She would also be interested in the possibility of setting up some closer collaboration in the future between immersion schools and teachers. Finally, she suggested that if we want to know more about evaluation instruments we should contact one of the authors of the BEP report (Richard Johnstone, Alan Dobson, María Dolores Pérez-Murillo).

Fred Genesee (McGill University, Montréal, Canada)

Email: fred.genesee@mcgill.ca

Professor Emeritus; Department of Psychology.

Main research areas: immersion education, bilingualism, second language acquisition, bilingual first language acquisition.

Interview date: 12/17/2013

General introduction

In order to be called 'immersion' in Canada, a school has to offer at least 50% of the curriculum in the L2. When a school devotes less than 50% of the instructional time to the L2, it is usually called 'enriched language instruction'. Most programmes in Canada are total immersion programmes, i.e., 100% of the curriculum in the L2, and this has become the standard model. (Early) partial immersion is not very common.

For L2 proficiency, total immersion is assumed to be better than partial immersion, because more instruction through the L2 leads to a higher language proficiency at the end of primary school. Research has also shown that teaching through an L2 does not negatively affect the students' L1 skills. In fact, as FG pointed out, students who receive more L2 instruction sometimes perform better in their native language.

The main reason why schools opt for partial immersion as opposed to total immersion is because parents are cautious and concerned about the effects of total immersion on their children's L1. The amount of L2 instruction therefore depends largely on what the community will tolerate.

FG remarked that when a school has a 50/50 programme, that usually means that overall there is less than 50% instruction in the L2 because many school activities are likely to take place in the L1 – for example, school announcements, school excursions (in Dutch in the case of the TPO project). A 30% programme is, therefore, going to be more like a 20% programme. FG said it is best to have as much time devoted to the target language as possible, and advises 70% of the instructional time in the L2 if you want 50%, especially because there is no disadvantage to offering more instruction in the L2. In his opinion, parents often want to start off with 50% and then expand to more as their children become more comfortable. According to FG, 30% L2 instruction would be more like enriched language instruction, instead of the partial immersion programmes aimed at in the TPO pilot study.

Decisions about immersion education programmes in Canada are generally made because the country is officially bilingual. Moreover, many people have a non-Canadian background, so they are generally very enthusiastic about and familiar with learning another language. Usually, there is more demand than there are places available. Therefore, the children who were not able to enrol, often end up in the regular programme. FG advised a whole-school approach, so that the language can be promoted in the entire school and at every level. It is also easier to administer for the school principals. In dual track schools it is harder to promote the L2, because there are more administration complexities and all other day-to-day activities will most likely be in the L1 (Dutch in the case of the TPO project). The downside of a whole-school approach is that these schools might be seen as elite. They will receive a lot of attention and will attract very good students and teachers.

No formal admission requirements exist in Canada, and therefore every student can attend immersion education, in principle. As noted above, demand is typically greater than the number of places. Sometimes some informal selection takes because of community and parental concerns. This usually happens in areas where parents are less familiar with the concept and are cautious about their child participating in an immersion programme. FG gave an example from the Quebec province, where

parents are much more comfortable with the concept of immersion education and where less self-selection takes place.

Research also shows there is no reason to exclude students from an immersion programme. Students who struggle might even benefit from such a programme. There is no empirical evidence that students who struggle in school are more likely to fail in an immersion programme. No research has been conducted on students with severe language problems yet, so this would be an interesting topic for further research.

However, English immersion schools do not like to take new students in (total) immersion programmes after grade 2, because the advantage of students who started from the beginning might be too large.

TPO RESEARCH QUESTIONS

1. Which international examples are available that resemble our pilot study, both in organization and content of the programme?

There are not many examples of partial immersion programmes in Canada, because most parents want total immersion. More research has been done on partial immersion programmes in the US with the 50/50 and 90/10 models of immersion education. However, most of those programmes are two-way immersion and they tend to focus on the minority language student instead of the majority language student (which will be the case in the Dutch context). FG suggested to either contact Kathryn Lindholm-Leary, who has a large data set on partial immersion/TWI, or someone at Utah University (which is what we have done). It might also be worth going to the Utah University conference on immersion education in Oct. 2014, where FG will also give a presentation.

In general, results that have been reported in Canada about partial and total immersion programmes show that partial immersion students performed better than regular programme students with the L2 as a subject, but not as well as full immersion programme students.

Decisions in Canada about new programmes/developments are made on a regional level, as education is a provincial jurisdiction.

Europe vs. Canada

FG pointed out that there is a clear difference between the research on immersion education in Europe and that carried out in Canada, mainly due to the fact that there is much more experience with immersion programmes in Canada. Outcomes in Canadian research are measured in a more traditional/conventional way with more consistent outcomes, whereas research in Europe is not very focused (yet) and there are no standard models for research on immersion education. In European studies one tends to look more at discourse and very specific aspects of language. This does not necessarily indicate the overall proficiency of students, which makes the results hard to generalize.

Further research

FG mentions a few topics in the field of immersion education that he thinks are very interesting to look at in future research and would also be interesting topics for our TPO study. The first topic he is currently working on himself as well.

1. Trying to identify students who are at risk for reading or language problems early in school so that they can be given additional support early (FG is working on an article on this subject).

2. (Immigrant) backgrounds and learner characteristics related to learning outcomes.
3. Pedagogy and teacher skills. What is the most effective way of teaching?

2. Which evaluation instruments are available for use in qualitative and quantitative evaluation research in bilingual primary education?

Education in Canada falls under provincial jurisdiction. Some Canadian provinces have their own language assessment tests, whereas others do not. Many instruments were however created when immersion education was just introduced. FG suggested developing new tests in English if the national curriculum specifies the learning objectives for English very clearly and precisely. In this way, one can also use the results to improve the curriculum.

He also suggested using the PPVT to measure receptive vocabulary in both L1 and L2 and to test productive proficiency by means of a storytelling task, such as The Frog Story.

Productive proficiency is not easy to measure with a standardized test. Many Canadian schools use storytelling exercises and analyse the speech using a Likert scale to assess the pronunciation, grammar etc. It's not very specific and depends a lot on the assessor, but it gives an overall idea of how the students' oral proficiency has improved.

Minority language students

Very little research has been conducted on this topic and FG thought it would be very interesting in our context to look at minority/immigrant students as well. In Canada these students are assessed the same way as native L1 speakers and this does not seem to cause any problems. In the Quebec province, students from immigrant families are obliged by law to go to French-speaking schools.

3. What results have been reported concerning the development of children's first (L1) and second (L2) languages when involved in such programmes?

L1/L2 literacy teaching

One of the challenges in immersion education is how and when schools are going to teach L1 and L2 literacy. In early partial immersion both languages are taught simultaneously, so that a programme can be developed in which both L1 and L2 literacy are taught at the same time.

At-risk students

In Canada, the St. Lambert evaluation tested students at the beginning of kindergarten and then after each grade. FG said that any pretesting would be ideal in order to assess whether there are at-risk students in the immersion programme and whether immersion students are academically stronger than typical students in the non-immersion programme. It might be that immersion students do better in comparison to non-immersion students because of the self-selection, they might be more academically gifted.

FG furthermore suggested that in the TPO pilot it would be useful taking socioeconomic status into account when comparing results, in order to eliminate variability and control for these factors.

CEFR

The CEFR is being adopted in Canada as a guideline for developing tests. It is not used as a test itself, because it only provides benchmarks assessing students in very general and widely interpretable terms. There is no alignment between a school's curriculum and the CEFR whatsoever,

so testing by making use of the CEFR would not necessarily indicate the effectiveness of the curriculum. Testing would be more useful if the curriculum was aligned with the CEFR, but this is rarely the case.

FG pointed out that using standardized tests and measures is very useful, because other schools/researchers can use this as well and this will make it easy to compare results, but on the other hand, there is then no alignment with the national curriculum.

Another framework that is being considered in Canada is the ACTFL Proficiency Guidelines (The American Council on the Teaching of Foreign Languages).

Transition from primary to secondary

In the Canadian province of Quebec, primary school students continue their immersion education programmes in secondary education. They can choose from a variety of courses and sometimes these have been designed according to the knowledge they have already gained in primary school. In some smaller communities, a collaboration between primary and secondary education exists, whereby children from one particular primary school continue their education at a particular secondary school. It is however difficult to account for a perfect transition between primary and secondary education, because one can never fully predict which secondary school primary school students will choose.

4. What are the relevant characteristics and didactic skills of (early partial) immersion teachers?

FG pointed out he is not an expert on didactics and suggests that if we want more information on this topic, we might want to contact Roy Lyster or Merrill Swain.

Language requirements

In Canada, teachers must hold a basic teaching certificate. In addition, supplementary training for teaching in immersion settings is offered. However, immersion training is often not as elaborated as would be desired and teachers usually only take some extra additional courses besides the general teaching degree they hold. For more information on teacher training, FG suggested consulting the Canadian Association of Immersion Teachers website.

L2 proficiency/native speakers

The feasibility of hiring native speakers as immersion teachers in Canada, and especially in the Québec province, is much higher than it would be in countries with only one official language. Regions in other provinces with a high population of French speakers often have more access to native speakers. However, FG emphasized that teacher qualifications and didactic abilities are more important than having native speakers. In the Dutch context, it would be more useful for (experienced) teachers to take language enrichment courses to improve their L2 knowledge.

Late starters in an ongoing immersion programme

In Canadian immersion programmes, schools do not usually accept new students beyond grade two, because it is thought that after this age, students may have difficulties catching up. However, the situation with early partial immersion could be slightly different, because less L2 is provided, so it mainly depends on the extra support from the school. It is important for parents to understand that the benefits of immersion education are long-term benefits and that in order to guarantee success, students should not jump in and out of these programmes.

Subjects in L2

There is not much research on whether some specific subjects are more suitable to teach through the L2 than other subjects. One important factor in teaching a subject through L2 is to make sure the teacher has a good proficiency in both the language and the subject matter. Teachers in primary school technically teach everything, but if a teacher is particularly good at a certain subject, they should definitely teach that in English. Another important factor is whether the language skills taught through L2 are useful outside the school context. In this regard, FG emphasized the need to make sure that subjects taught in the L2 are serious subjects and that if subjects like music, PE, or arts are taught in the L2, students will probably not take the immersion programme too seriously.

FG suggested that mathematics might as well be taught in L1/Dutch, because it contains a highly specialized variety of the language that is not very useful outside the math classroom. It is therefore not a particularly good subject to teach in English because the language skills are not all that useful in other domains. In contrast, areas in the social sciences and the natural science are probably particularly useful to teach in the L2, because the range of language skills is broad and these skills are generalizable to out-of-school contexts.

FG also added that it is possible in the Dutch context to teach a certain subject using the L2 throughout the entire primary school. However, some other researchers would suggest not to teach one subject in the L2 for more than two years.

(Inter)national exchange

There has not been much international exchange in Canada, but good opportunities will be available in the Dutch context for L2 immersion programmes. FG suggested looking into the UG website called e-pals. They arrange electronic exchanges between students in different countries. It also includes technology. Students like it and they enjoy the exposure to native speakers. The website is carefully constructed and monitored by teachers.

5. How are the educational programme and the development of the children's different languages perceived by the students, teachers and parents?

It is very important to assess the attitude of the community and school administration/principal of school, because they have a huge effect on the success of a programme. Communities in Canada are very favourable towards immersion, but especially at the start of the TPO pilot, enthusiasm on the part of the school/school administrators should be established.

About the start of early partial immersion in the Netherlands

FG suggested that if the selection of participating schools has not been made yet, this should be done on a competitive basis. Spread the message and see which schools are interested. Start with the most enthusiastic schools, which are most likely to succeed, and because of this success many other schools might want to follow the example that has been set at that point.

Furthermore, FG pointed out the importance of finding support in the community. Organize meetings for the most influential people in the community (administrators, business people) and keep informing the parents about the programme. In this way, advantages and positive impacts of the program can be shown and this will also improve the overall success as many people will now support the idea.

Teachers

Enthusiastic and well-qualified teachers are essential for the success of an immersion programme. It is very important to provide good training and have enthusiastic schools/teachers. Teachers must

believe that immersion is a good idea. Meetings with teachers are very important as well. New teachers should visit schools with existing programmes so they can see how it works and can lose any prejudices about the impact on the L1. In this regard, FG suggested creating a teacher network, for example an exchange between Dutch and Canadian/US schools. Talking with other teachers will be the best strategy, as is the case within the Canadian Association of Immersion Teachers.

Parents

In Canada, parents' attitude was never a cause for concern because the parents were the ones enrolling their child in an immersion programme. Therefore, no assessment of this aspect of immersion programmes has taken place. Of course there were also parents who were less willing, but those were mostly parents who thought their children would have problems in school. FG noted that in the TPO context, it would however be very useful to assess parents' attitude. He suggested self-designed questionnaires for this purpose.

In some communities in Canada, problems emerged when parents who did not have children in immersion were upset that there were very few good students in the regular program and this created conflict in the community. Parents who thought that immersion was going to be too difficult had kept their children in a regular programme and this caused a self-selection situation with stronger students in immersion programmes and parents with children in regular education felt their children got a second-class education. Providing information to parents is therefore crucial, and if there are parents who do not speak Dutch or English, make sure to have an interpreter or translation. Create a website with a clear and detailed description of the programme, research from other countries and FAQs.

One further remark concerning parents who do not speak the L2 was that parents who do not speak the L2 are often worried about what they can do and how they can help. While this is not such a big problem in Canada, FG mentioned that a solution may come from the teachers, as they can help with homework. Children can attend extra classes. Furthermore, parents can talk to other parents about the problems they are facing.

Students

With regard to assessment, the same holds as for parents, that is, it is best to use self-designed questionnaires. FG noted that examining the attitudes of minority students would be an interesting topic to look at, because not much research on this topic has been conducted so far. FG thought that immersion education could be particularly interesting for them because they would also feel encouraged to learn English. What is important is that teachers should understand that it is beneficial to minority language students to be in immersion programmes as well and that it will not negatively affect their acquisition of Dutch.

Conclusion

FG pointed out that there seems to be good control over the TPO project with lots of planning, which is critical to the success of the programme. He said that he would also like to be updated about the project and pointed out that it is very important to plan and foresee issues as addressed in this interview, as this will prevent many mistakes from being made during implementation of the program and furthermore improve its chances of success. He said that the TPO pilot should make sure new teachers get as much support as possible with respect to professional development as teachers' skills are critical. And it should provide teachers with some in- and/or pre-service training/summer courses. This is especially important during the first year. It is also important to create a community of teacher-learners. They are the ones who will know from experience what methods are beneficial to the children's proficiency.

Finally, with respect to the TPO pilot, FG suggested to start very small, with a few schools at the beginning and grow slowly. Starting with schools which are really enthusiastic and relatively well prepared leads to the best chance of success (this can be realised using a competitive selection procedure). Other schools will follow their example once they realize the success of the pilot schools.

Johanna Watzinger-Tharp (University of Utah, US).

Email: j.tharp@utah.edu

Associate Professor of Applied Linguistics & German; Department of Languages & Literature/Department of Linguistics

Main research areas: dual language immersion, foreign language studies, German linguistics, teacher education.

Interview date: 01/14/2014

General

Many schools have already started with an immersion program in the USA or are busy implementing one. The terms partial and full immersion are not used so much anymore; rather, the term 'dual language immersion' is used. In these programs, 50-100% of the lesson time is devoted to teaching in the target language. States often decide on the guidelines for the immersion education themselves. Utah, for instance, has adopted its own statewide 50/50 model, which started with legislative funding in 2008. From first grade on, pupils spend half of their school day in English, and the other half in the target language. Each immersion grade is split into two classrooms with two different teachers, each teaching in their own language. In addition to being very cost-effective, this model is also very efficient as the children listen and learn in their target language for 50% of the time. The present target languages include Spanish (because of the large Hispanic population in the US), French (because of the connections with Canada and Africa), and since 2012, because of business and trade connections, Chinese and Portuguese, and finally, as of Fall 2014, German, because of its connections to the overlapping STEM (science, technology, engineering, math) disciplines in education in Germany and Utah, and because it is one of the main languages in the EU.

Schools might decide on starting an immersion program because of the economical benefits. They always receive state funding and also private funds and funds from non-US governments, for example France and Spain. The immersion programs help students to become proficient in more than one language, which will improve their chances on the job market in the future. Moreover, the children become bilingual, biliterate and bicultural. As many schools only start teaching second languages in the 9th grade, the students do not get the opportunity to be fully immersed in the language and hereby enhancing their language skills. Because of this, a recent development is that Delaware and Georgia are following in Utah's footsteps of becoming a state which offers many immersion programs.

TPO RESEARCH QUESTIONS

- 1. Which international examples are available that resemble our pilot study, both in organization and content of the programme?**

Standardized tests

Utah uses standardized tests for measuring proficiency, including more tests for young learners. The ACTFL Oral Proficiency Interviews were created in the 1980s. The Foreign Language Standards in the late 1990s, with more recent adaptations for younger learners.

After 6th grade, when students are more likely to make their own choices, it is important to expand testing from language proficiency to bilingual and bicultural knowledge. Utah is starting to investigate pupils in the immersion program, asking the following questions: To what extent have students been bilingually educated? How has their identity been shaped during six years of immersion education? Upon leaving their current educational institution, will they freely choose to continue within an immersion programme or will they then prefer monolingual education? JWT is not interested in the intercultural competence per se, but more in how the children's bilingual identity has been shaped

through the years and whether they feel connected to both cultures. She also remarked that students' motivation about whether or not to continue is of much interest to Utah researchers. Utah is currently working on a curriculum for the 10th, 11th and 12th grade so the children wanting to continue their bilingual education can now do so.

2. Which evaluation instruments are available for use in qualitative and quantitative evaluation research in bilingual primary education?

L1/L2 assessment

The L2 proficiency and development can be measured by using the AAPPL-test, which measures both receptive and productive proficiency. The USA (and especially Utah) also uses can-do statements, a method which resembles the CEFR, for all four skills in the early grades. In order to compare the academic potential of the bilingual and monolingual students, the bilingual students also take part in the regular curriculum tests designed for the monolingual classes. The regular tests must be given to all students in elementary schools, so results can be compared, but more importantly, these can be used to document how well the immersion students are achieving academically in reading, writing, mathematics and, later on, in science as well.

The tests are designed according to the curriculum. A recent development in Utah is the shift towards the 'Common Core' method, a nationwide curriculum which measures skills that go beyond the regular subjects, like analytical and critical thinking. Common Core is now also integrated in some of the testing. Furthermore, there are standardized annual tests from 2nd grade on.

More information on the AAPPL-test can be found here: <http://www.languagetesting.com/aappl>
And a link to Utah standardized tests: <http://www.utaheducationfacts.com/index.php/get-the-facts/test-scores/117-statewide-tests>

Recommended evaluation instruments/tests

According to JWT, the best tests are designed according to the curriculum.

3. What results have been reported concerning the development of children's first (L1) and second (L2) languages when involved in such programs?

Selection procedure

The school decides which children to educate. Because of the great demand for immersion programs, many schools implement a selection procedure, most of the time by lottery. Parents are often very motivated and search outside their neighbourhood to find an immersion programme with the preferred target language. They try to enrol their students in their preferred programs, but many are rejected because of the great demand. Schools now have waiting lists, regardless of their target language, as the parents notice that immersion education will allow their children to obtain valuable assets.

Development of L1/L2

In Utah, five languages are currently taught as the L2: Spanish, Chinese, French, Portuguese, and German. All languages are said to be acquired in the same way and at approximately the same speed. However, JWT did remark that in Chinese literacy is extra challenging because of its different writing system.

CEFR

In the USA, the CEFR is a popular model for self-assessment, but it is seen as a supplement to some other assessment. The CEFR is not thought of as being complete, so it is often complemented by can-do statements. Moreover, the CEFR is mostly used in secondary schools.

4. What are the relevant characteristics and didactic skills of (early partial immersion) teachers?

Guest speakers/teachers often arrive from Spain, China or France in primary schools in Utah. Much of the time these teachers have also been educated in their country of origin and have experience as elementary teachers there. Nevertheless, they still go through training programmes in order to get their credentials to teach in the USA.

L1 teachers are native speakers of English and many L2 teachers are also native speakers, especially for Chinese. The requirements for L2 proficiency levels are very high: the ACTFA Advanced Mid level is the minimum language proficiency requirement.

Educational programmes

Utah universities have now established immersion programme teacher education. They offer a set of courses on dual language immersion and content-based instruction. The key to becoming a successful immersion teacher is to be able to stick to the L2. Optimal foreign languages strategies have been developed and are now spreading to secondary schools as well.

In order to maintain the high level of proficiency and teaching, teachers are continuously monitored by administrators at the schools; these administrators carry out observations and make sure the teachers are still able to teach their pupils successfully.

Subjects

As the primary curriculum is the same as the state curriculum, the following subjects are included: math, literacy, science, social studies, art, music and PE, of which the last three can be designed by the school itself.

Subject competence

As there is a two-teacher model, the content is not duplicated. The teachers reinforce each other and work together as every subject in school is taught in both languages (50/50).

Late starters in an ongoing program

Some schools allow the children to start during the 2nd grade, but most schools do not enrol these students in their bilingual program once the program has started, as these students will often come across many difficulties while trying to grasp the meaning of words and phrases other children have acquired already. Utah's goal is to spread immersion education widely and to make it mainstream (across the state), so that every child will be able to take part in a program if that is desired.

According to JWT, the combinations of the subjects and which language to use does not really matter in early immersion education. However, in later education, especially secondary education, it is very important for the courses in the target language to offer rich linguistic input. Various genres and many texts make sure there is much exposure to the L2.

Teaching materials

Utah has contracted a textbook publisher (Pearson) to translate textbooks for maths. This is the only subject for which translation has taken place, so that the subject matter can be taught in the L2 but the exact same content remains, which is important to meet the curriculum goals. JWT believes in using authentic materials from the target language and culture. Therefore, all the other subjects use authentic books and workbooks from the original country.

The curriculum is taken care of by various experts, who design materials and inspect the quality.

(Inter)national exchange

Utah organizes a large conference every August during which training and meetings are scheduled to promote exchange and collaboration between the schools. Utah also has bilateral exchanges and district-level partnerships with for instance France (Grenoble) and Brazil. Students are also involved in these exchanges of material and knowledge.

5. How are the educational programme and the development of the children's different languages perceived by the students, teachers and parents?

Teachers

The teachers are very motivated and always want more materials. They want a rich, cultural environment and want to use the target language as much as possible. Because of this, they are continuously looking for new authentic materials, sharing their ideas and exchanging materials and knowledge. Guest teachers are sometimes perceived to be stricter than American teachers due to cultural and pedagogical differences.

Students

The students are very motivated as well, especially in their first year. They find a new way of learning and this makes them very enthusiastic. However, after a couple years of taking part in the programme, they don't realize they are taking part in a special educational programme anymore and their immersion education turns into something normal.

Parents

Parents are in general very supportive, but some of them are cautious about using an L2 to educate their children. JWT described their doubts about their children not achieving the same level in maths and science as their monolingual peers. This has been proven to be untrue, however.

Before the programme starts, there are numerous parent meetings; parents also assist and are part of the programme, so they are for example often involved in the classroom. JWT remarks that parental involvement is very important, as many parents are interested in their children and want them to do well in school. Therefore, she advises organizing parental meetings at the start of the program and throughout in order to keep the parents informed at all times. As indicated above, Utah even allows the parents to be present in dual immersion classrooms and to assist teachers when appropriate.

Conclusion

JWT told us that she very much enjoyed speaking with us and she hopes to continue the conversation and the "cultural exchange". She is very interested in the language education developments in the Netherlands and Europe. She also looks forward to reading the report and our findings and was very interested in receiving recent articles about the announcement of immersion education in the Netherlands.

She also told us about the Immersion Conference in Utah (October 2014) and encouraged us to participate or attend this event, as a part of the conference will be dedicated to immersion education in Europe as well.

Hanna Österlund (Skolverket, Sweden)

Email: hanna.osterlund@skolverket.se

Project leader, The Swedish National Agency for Education

Main research areas: CLIL in Swedish primary education, educational evaluation of schools.

Interview date: 01/31/2014

General introduction

There are currently a growing number of primary schools where various subjects are, to a certain extent, taught in English. This has however not been researched very thoroughly and especially the evaluation of the programmes has been problematic. In 2003, researchers from Skolverket (The Swedish National Agency for Education) started tracking the number of primary schools which had implemented variants of immersion education and monitoring their progress. This change was largely instigated by schools, schools administrators, and parents, that is, it was a bottom-up rather than top-down change coming from the government.

The study on immersion education conducted by Skolverket included 42 primary schools. Those schools were located by the tracking process that was done before and these schools appeared to be the only schools that had started an immersion programme at that time. Most were situated in and around the bigger cities, but there was a certain spread throughout the country. The schools that were selected were willing to take part in the research and evaluation process to further improve their immersion programmes.

After the Skolverket report was published, many other schools became interested in starting their own immersion programme. However, the national school inspection (Skolinspektionen) has to approve new applications and not every school was considered good enough. The school inspection is supposed to make sure that only motivated, well-prepared and enthusiastic schools can implement this form of education and it is now policy that all new schools that want to implement an immersion programme need to apply and wait for approval of the school inspection.

The Skolverket report mentions a "social bias in immersion education", namely the fact that as a result of social selection in immersion programmes only academically motivated students follow this type of education. There are no results on whether this social bias has affected the proficiency of either the L1 or the L2. Note, however, that it is not the case that students are in fact selected, that is, a formal selection procedure is not allowed in schools, and – on paper at least – programmes are open to all students. HÖ remarks that it would be interesting to investigate this issue in more depth.

TPO RESEARCH QUESTIONS

1. Which international examples are available that resemble our pilot study, both in organization and content of the programme?

HÖ is not working on research concerning CLIL and immersion education anymore and is therefore not familiar with any new developments regarding this type of education in Sweden. To her knowledge, some researchers outside her agency are currently working on this topic and for more recent information about any developments or assessment she suggests contacting Maria Lim Falk.

Further research

The Skolverket study did not primarily focus on the language outcomes of participating students or the specific educational approach adopted. HÖ therefore mentions a number of topics that would require some more in-depth research:

- language proficiency (effects on the L1, immersion programmes vs. regular programmes)
- didactics (language teaching and training, teacher competences)
- assessment (specific assessment methods for the L2, assessment of difficult learners and students with foreign backgrounds).

2. Which evaluation instruments are available for use in qualitative and quantitative evaluation research in bilingual primary education?

Testing

National tests in Sweden are administered in grade three, six, and nine. Students are tested twice during primary education on English, and three times on Swedish and mathematics. There has been no pre-testing (testing before the start of immersion education), because language outcomes and comparison were not the main research area of this study. However, some schools advise parents whether or not to enrol their children in an immersion programme. This advice is based on interviews with parents, but schools can never refuse to educate a child in an immersion programme, as selection is strictly prohibited.

L1/L2 assessment

HÖ did not have any up-to-date information about evaluation or assessment methods. In Sweden, matrices are sometimes used to assess the language proficiency, but she is not familiar with any specific assessment tools. HÖ mentions that more specific language assessment tools would be required for further evaluation of immersion education in Sweden. As noted above, no specific language assessment tools were developed during the course of the Skolverket study.

3. What results have been reported concerning the development of children's first (L1) and second (L2) languages when involved in such programmes?

Selection procedure

As noted above, whenever a school gets approval from the school inspection to start the immersion programme, the teachers and principals are not allowed to have a selection procedure. The researchers in the Skolverket study found that, during the study, many schools had in fact conducted interviews with the children (and sometimes even their parents), in order to get a clear picture of the children and whether they were motivated to take part in an immersion programme. Carrying out such interviews is not allowed anymore.

CEFR

HÖ did not know whether the CEFR is used as a guideline or as an assessment tool, but Skolverket is working on the development of assessment tools for various subjects that could also be used in immersion schools.

Transition from primary to secondary

HÖ did not know much about the transition from primary to secondary education in general, but she knows that many of the children who have taken part in the Skolverket study at the time have continued their bilingual education in secondary education/gymnasium schools. Since the school system in Sweden is organized a bit differently than, for instance, in the Netherlands (for example, schools can receive funding from a certain municipality or business), various chains of schools exist that cooperate and offer appropriate and continuous secondary education.

4. What are the relevant characteristics and didactic skills of (early partial) immersion teachers?

HÖ did not know about any criteria for the teachers to teach in an immersion programme, nor does she have much detailed information about teacher education. She did know that teachers need to have the right qualifications, but she did not know what the current exact requirements are.

L2 proficiency/native speakers

At the start of immersion education, only a few schools were informed about required teacher competences and therefore many schools accepted teachers who would now be considered to be barely qualified. Their subject competence was sufficient, but their proficiency in the L2 did not suffice. Nowadays, teachers often have the opportunity to go abroad and study the L2 for a while, for instance in courses offered in England, and they can often take part in special training courses in college or universities.

Subject competence in L2

Depending on the goal of the immersion school (there are 'profile schools' with mostly Swedish native speakers and 'bilingual schools' with many native English speakers for the subjects taught in the L2), both Swedish and English native speakers can take part in training courses so they can either improve their English or Swedish language skills.

Whereas a preference for native speakers was expressed, a contradictory situation arose. Native English speakers had an impeccable knowledge of the L2, but often lacked sufficient L1/Swedish knowledge to explain topics in the children's first language. On the other hand, children sometimes complained about the inability of their native Swedish teacher to explain subjects in English. For this reason, as HÖ explains, more teacher training is required prior to the start of immersion education programmes.

Subjects

Mostly the social and nature-oriented subjects are taught in English, but HÖ remarks that the decision about what specific subject(s) to teach in the L2 differs among schools. The decisions were mostly made according to what competences the teachers had at that particular moment, but it also depends on the type of school.

Best subjects

Schools have different opinions about what topics are the most suitable to teach in the L2. The 'bilingual schools' for instance, are more internationally oriented and have hired more teachers coming from English-speaking countries. Those schools will more often choose for nature-oriented subjects, because that often includes international topics that are more suitable to teach in the L2. Social studies like society and history are suitable as well, but they often include national subjects that would ideally be taught in the L1.

For this reason, schools sometimes decide to teach a subject entirely (nature-oriented subjects) or partially (social studies) in the L2. Schools are free to decide about the amount of time devoted to the L2 per subject, as long as the total amount of L2 instruction does not exceed the maximum of 50%.

Teaching materials

HÖ did not know very much about the usage of teaching materials in general, as her study did not look into this topic and did not ask schools about what materials they used. From her own experience she knows some specific examples whereby schools bought authentic textbooks from England, especially for the natural science topics. Other schools teaching mathematics in the L2 still used the original L1 textbooks. HÖ added that when a subject is taught 50/50 in the L1 and L2, schools might

either use two different books just the one original L1 textbook, but this is also a topic for further investigation.

(Inter)national exchange of ideas

There has not been much research on this topic in Sweden yet. HÖ only knows about non-native English teachers who had the opportunity to go abroad and improve their L2 skills by taking part in language courses in England.

HÖ had heard something about students writing letters to students abroad, but this was not the main focus of her research and this is not a very common practice in Swedish schools yet.

(Inter)cultural approach

Some schools find it very important to include cultural subjects (related to the L1 and L2) in their curriculum. However, this is very different among the schools. It depends on the preferences of the individual schools.

5. How are the educational programme and the development of the children's different languages perceived by the students, teachers and parents?

Society

With the introduction of immersion programmes many people were concerned about the L1 proficiency development. Some of the L1-oriented people or researchers were very critical about the effects of this educational approach on children's L1 proficiency, which resulted in a little bit of resistance towards the programme. They thought this was not the best way to learn English or that children in primary school already learn enough English for their age. However, for most people English has a very high status and is seen as a prestigious language, so many parents are interested in these immersion programmes and are very supportive when it comes to their child taking part in such a programme.

Teachers

At the start, many teachers were very self-conscious about their L2 proficiency. They needed more training in order to be convinced of their own proficiency. Furthermore, they needed training on special methods to teach these students through an immersion programme.

Students

In general, students are very positive towards the immersion programmes. The students see the immersion classes as a great opportunity to learn more English.

Parents

Parents have been involved in the immersion programmes throughout. In fact, given that parents were often the ones who chose the school and programme, there were many information sessions and children could even take part in try-out lessons to see whether they liked it or not. Again, the status and prestige of the English language ensured that the parents were interested, even at the start of the programmes.

Assessment

Interviews were conducted to find out more about the opinions of both the teachers and the students. Questionnaires were designed to assess the opinion of the head teachers of the schools.

Conclusion

HÖ was very curious to know more about the reasons for the proposed educational change in the Netherlands and wondered if there was strong parental support, as in Sweden, or if it was mainly initiated by the government. She was interested in the subjects taught in the L2, the required teacher qualifications, and the time devoted to the L2. She also said it would be interesting to receive more information about the developments concerning immersion education in the Netherlands and especially information and developments regarding language assessment, as this was not the main goal of the study in Sweden and thus a topic they would like to research more as well. She said it could be helpful to exchange ideas about more specific language assessment tools. Furthermore, she added that if more information was required about developments in Sweden, it was always possible to contact her again and she could refer to colleagues in her agency who have more experience and information with specific topics regarding immersion education.

Appendix II – Overview of evaluation instruments

Author(s)	Evaluation instrument		Tested construct
	Standardized	Non-standardized	
Cheng et al. (2010)	Cambridge YLE (L2)	Reading and writing proficiency (L1)	L1 receptive and productive proficiency L2 receptive and productive proficiency Mathematics
Knell et al. (2007)	PPCLS, PPVT-R (L1) PPVT-R, WRAT-3, Fry Word List (L2)	Phoneme deletion, onset-rhyme detection (L1) Oral interview, onset-rhyme detection (L2)	L1 receptive and productive proficiency, vocabulary L2 receptive proficiency
Wode (2004) Burmeister & Pasternak (2004)	Hamburger Lesetest (HAMLET) (L1)	Oral production using picture-storytelling tasks (L2)	L1 receptive proficiency L2 productive proficiency
Kuska, Zaunbauer, Möller (2010)	Hamburger Schreib-Probe (HSP) (L1) Deutsche Mathematik-Test (Maths test for grade 2 and 3)	Thematic educational movie in L1 on a topic not discussed in class to assess the general learning skills	L1 productive proficiency L1 General learning and comprehension skills Mathematics
Zaunbauer & Möller (2010)	Coloured Progressive Matrices (intelligence) KT1 (Konzentrationstest für das erste Schuljahr) (concentration) Nonwort-Rekognitionsaufgabe (verbal intelligence)	Reading and writing test (6 items, e.g. recognizing and writing the letters of the alphabet) (L1) Mathematics test (4 items, e.g. recognizing and writing the numbers from 1-9)	L1: receptive and productive proficiency (Verbal) intelligence; concentration Mathematics
Gebauer, Zaunbauer & Möller (2012)	Würzburger Leise Leseprobe (WLLP) (L1) Hamburger Schreib-Probe (HSP) (L1) Raven's Coloured Progressive Matrices (intelligence)		L1 receptive and productive proficiency Intelligence
Bostwick (2001)	Eiken (STEP) test, Language Assessment Scale (LAS) Iowa Test of Basic Skills (ITBS) (L2)	Japanese language test (L1) Written test (L2)	L1 receptive and productive proficiency L2 receptive and productive proficiency

Turnbull, Hart & Lapkin (2003)	Education Quality and Accountability Office (EQAO) test (L1 and mathematics)		L1 receptive and productive proficiency Mathematics
Thomas, Collier & Abbott (1993)	Metropolitan Achievement Test (MAT) (L1) Student Oral Proficiency Rating (SOPR) (L2) POS Mathematics test		L1 receptive proficiency L2 productive proficiency Mathematics
Campbell, Gray, Rhodes & Snow (1985)	Modern Language Association Cooperative (MLA) test (L2)	Self-assessment forms for the students (L2)	L2 receptive and productive proficiency
Cirino et al. (2007)	TOPP-S, WLPB-R, Experimental Word List (L1) CTOPP, WLPB-R, LEXESP (L2)	School-issued test on character recognition (L1) School-issued test on character recognition (L2)	L1 receptive and productive proficiency, vocabulary L2 receptive and productive proficiency, vocabulary
Reder, Daigle & Demont (2013)	French standardized phonological test	Morphological test, syntactic test	L1 receptive and productive proficiency
Serra (2007)	Standardized tests in grade 3-6	Oral and written tests in grade 1-2 Mathematics test	L2 receptive and productive proficiency L1 receptive and productive proficiency Mathematics
Dobson, Pérez-Murillo & Johnstone (2010)	Cohort of students continuing in secondary education: Cambridge IGCSE (L1, L2)	Written test (L1) Interviews, analysis of lessons (especially the language skills during communication in the classroom), descriptive-narrative writing task, teacher assessment	L1 productive proficiency L2 productive and receptive proficiency
Lasagabaster (2000)	Standardized test (Galbahe) (L1) Standardized test (Galbahe) (L2)	Grammar and vocabulary test, reading and listening comprehension test, oral test (L3)	L1 receptive and productive proficiency L2 receptive and productive proficiency L3 receptive and productive proficiency
Lorenzo, Casal & Moore (2009)		Skill-based diagnostic tests	L2 receptive and productive proficiency
Skolverket (2010)	National tests in grade 3, 6, and 9		L1 proficiency L2 proficiency Mathematics

Seikkula-Leino (2007)	IEA reading skill assessment/PIRLS (L1) National Mathematics Test Raven's Non-verbal Progressive Matrices, Wechsler Adult Intelligence Scale	The relationship between levels of intelligence and school success (achievement level)	L1 receptive proficiency Mathematics Intelligence
Merisuo-Storm (2007)		Various tests, including auditory and visual perception, written task, mathematical skills (L1) Oral and written recognition (L2)	L1 receptive and productive proficiency L2 receptive proficiency
Jäppinen (2005)		Problem solving, picture tasks, generalisations, metaphors, proverbs, tables, graphics, relations, similarity.	L1 various linguistic and non-linguistic skills
van Ruijven & Ytsma (2008) Riemersma & de Jong (2007)		Vocabulary test, technical reading, spelling, reading comprehension Test bilingualism for toddlers. 'toets tweetaligheid voor kleuters' Vocabulary test, technical reading, spelling, reading comprehension	L1 receptive proficiency L2 receptive proficiency

Appendix III – Matrix Literature Review

The first column

Authors: the author(s) of the referred study, including year of publication.

In the first row information provided in the referred study is given.

Participants: the number of participating students with in between brackets (if mentioned in the study) the number of children taking part in the immersion programme. If mentioned in the study, the number of schools is also given here.

Age of onset/age of testing: age at which the students have started taking part in the immersion programme. Below the age of onset, the age(s) of testing is given.

Context & Aim of the study: the environment in which the testing has taken place and what is tested in the study.

Language allocation: the languages that are referred to in the study and the percentages of instructional time devoted to each language (L1, L2 and, if mentioned, the L3).

Subjects L1/L2: an overview of which subjects are taught in the L1 or L2 (and, if mentioned, the L3).

Tested construct: the constructs the study has investigated. Can be related to the L1 or L2, but also to different subjects.

Instruments: the tests which were used to assess the given constructs.

Teacher information: information about the educational background of the teachers.

Results: the overall results of the study.

In the second row an answer to the specific questions from the tpo pilot is given based on information from the referred study.

Examples: explanation of how the results of this study were comparable to the tpo pilot.

Evaluation instruments: relevance of the evaluation instruments mentioned in the study to the tpo pilot.

Results on L1/L2: results and specific outcomes for both languages.

Didactic characteristics: the didactic characteristics and educational backgrounds of the teachers included in this study.

Attitudes: student, teacher and/or parental attitudes are described.

“-” means info is not given in the article.

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Cheng et al. (2010)	<i>N</i> = 998 (618 imm.) 3 schools (1 private, 2 public)	6 yrs (grade 1) Tested in grade 2, 4, and 6.	Primary education in China. To test the effectiveness of immersion programs. To test the imm. students' performance in L1, L2, and maths compared to non-imm. students.	L1: Chinese (60-70%) L2: English (30-40%)	L1: Chinese, maths L2: English language and culture, science, social studies L1/L2: PE, art and music	L1: Receptive and productive proficiency L2: Receptive and productive proficiency Mathematics	School-issued test Cambridge YLE School-issued test	Non-native English speakers. Almost all (95%) have certificate diploma or degree. Experience: 21% 1-2 years, 75% 3-9 years, 4% over 11 years.
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
L1: No negative effects on L1 proficiency. L2: Immersion students perform signf. better on all linguistic measures. Maths: no negative impact on development of mathematics.			Tested the effectiveness of L2 programs in public and private primary schools. Time spent on L2 (30-40%) was comparable to what is aimed for in the present pilot study. Students had to learn a new writing system.	Apart from various school- issued tests in L1 and mathematics, the Cambridge YLE test was used to assess the L2 development. This test is specifically designed for children in primary education and tests their L2 ability at three different levels.	Immersion students performed better in English in all test grades compared to non-immersion students, with the advantage of English becoming more prominent at the end of primary school. Immersion and non-immersion students scored similarly in L1 in grade 2 and 4, while the immersion group outperformed the non- immersion students in grade 6.	Teachers in this study were L1 native speakers with relevant diplomas or degrees in English and L2 teaching. They were relatively young and had different levels of experience.	No information about student, teacher, or parental attitude was given in this study.	

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Knell et al. (2007)	<i>N</i> = 183 (92 imm.) 1 school	± 2/3 yrs (kindergarten/pr eschool) Tested in grade 1, 2, and 3.	Primary education in China. To examine possible differences in L2 literacy, vocabulary and oral proficiency between imm. and non-imm. To test for differences in L1 character recognition between imm. and non-imm. To find predictors of word identification.	L1: Chinese (50%) L2: English (50%)	L1: Chinese, maths L2: English language and culture, science, social studies	L1: receptive and productive proficiency, vocabulary L2: receptive proficiency	L1: Oral school- issued test, PPVT- R ¹ , WRAT-3 ² , Fry Word List, school- issued onset- rhyme detection and phoneme deletion test L2: Onset-rhyme detection, PPCLS ³ character identification, PPVT-R	L1 native speakers. Graduates of teaching colleges (most with university degree). Have passed an oral and written English exam and have participated in training seminars.
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
L1: no negative effects on L1 character recognition L2: imm. students performed signf. better on L2 word recognition, vocabulary and oral language proficiency			This study described a partial immersion situation in China which started in kindergarten, whereby half of the teaching time was allocated to the L2. Students had to learn a new writing system.	This study used an extensive list of measures to assess L1 and L2 development, some of which are applicable to the present study (PPVT, WRAT-3), while some were specifically designed for this Chinese context by immersion teachers or other schools or for Chinese primary education (PPCLS)	Despite the fact that this study involved languages with different alphabets/writing systems, no negative effect was found on L1 proficiency. Immersion students scored signf. better on various linguistic measures, especially on oral proficiency. Immersion students were better able to express and comprehend various concepts in L2.	Most L2 teachers had relevant teaching experience and had passed some training courses in L2. None of the teachers were (near) native-like in English and thus it is interesting to note that also with relevant experience and sufficient didactic skills, signf. results in L2 can be obtained.	No information about student, teacher, or parental attitude was given in this study.	

¹ Peabody Picture Vocabulary Test - Revised

² Wide Range Achievement Test 3

³ Preschool and Primary Chinese Literacy Scale

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Wode (2004) Burmeister & Pasternak (2004)	<i>N</i> = not given (half imm.) Four parallel classes in same school	3 yrs (nursery school) Tested in grade 1 (begin and end), grade 2, and 4.	Pre-school and primary education in Germany. To show the potential of early (partial) immersion programs.	L1: German (30%) L2: English (70%)	L1: German language and culture L2: All other subjects	L1: receptive proficiency L2: productive proficiency, vocabulary	Hamburger Lesetest (‘Hamburg reading test’) (HAMLET) School-issued test elicited oral production using picture-storytelling tasks	Native German speakers. English teaching qualification with a sufficient minimum English proficiency understood as they must be able to teach the curriculum content with varied speech. Teachers develop their own materials, teaching with illustrated books.
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
L1: Imm. students achieved better average scores than non-imm. students; no negative effects on L1. L2: German primary school children learned a foreign language up to a very high proficiency level, when early, intensive, and continuous L2 exposure was provided.			This study described the implementation of English in German preschool and primary education. A maximum of 70% of the instruction was given in L2, which was more than aimed for in the present study. Given the fact that the German school system and culture are rather similar to the Dutch system, the results are nevertheless valuable for comparison.	A school-issued storytelling task was designed to test the students’ vocabulary development. Their reading skills were assessed with a reading test specifically designed for the city-state of Hamburg.	L1: no negative effects on the L1 (reading comprehension) were found. L2: The results from immersion students were not compared to those from students in a control group, but given the fast increase in vocabulary knowledge and oral proficiency, this study revealed very positive outcomes in L2 development.	All teachers in this study were L1 native speakers. They had a good knowledge of English, qualifications in English teaching, and experience with developing new teaching materials. Teachers were instructed to only correct students’ mistakes indirectly and not to correct development-specific mistakes, including code-mixing.	No information about student, teacher, or parental attitude was given in this study.	

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Kuska, Zaunbauer, Möller (2010)	N = 402 (228 imm.)	6 yrs (grade 1) Tested in grade 3 and 4	Primary education in Germany. What differences in performance are found between monolingual and immersive taught students when other factors are statistically controlled for?	L1: German (30%) L2: English (70%)	L1: German language and culture L2: All other subjects	General learning and comprehen- sion skills L1: productive proficiency Mathematics	Thematic educational movie in L1 on a topic not discussed in class Hamburger Schreib-Probe (Spelling test) Deutsche Mathematik-Test (Maths test for grade 2 and 3)	Native German speakers.
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
Students in the immersion program showed signif. better general learning skills when taught through an L1 movie than non- immersion students. Although not the focus of this study, mathematics and spelling results were signif. higher for immersion than for non-immersion students.			This study addressed the implementation of an L2 in German primary education. 70% of the time the L2 was used as a medium of instruction, which included all subjects except German language and culture.	The general learning skills in a new topic were assessed based on an L1 multiple-choice listening task with independent variables like intelligence, mathematics and spelling skills controlled for. They included SES in their analysis.	L1: Immersion students scored signif. better on an L1 spelling test than non-immersion students.	Teachers were native L1 speakers.	No information about student, teacher, or parental attitude was given in this study.	

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Zaunbauer, Möller (2007)	N = 139 (75 imm.)	6 yrs (grade 1) Tested in grade 1.	Primary education in Germany. To test the development of immersion and monolingual students' achievements and cognitive skills in the first grade.	L1: German (30%) L2: English (70%)	L1: German language and culture L2: All other subjects	L1: receptive and productive proficiency Mathematics Intelligence Concentration Verbal intelligence	Reading and writing test (6 items, e.g. recognizing and writing the letters of the alphabet) Mathematics test (4 items, e.g. recognizing and writing the numbers from 1-9) Intelligence test – CPM ⁴ KT1 (Concentration test for the first year) Nonwort- Rekognitionsaufga be	Native German speakers.
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
Both immersion and non- immersion students showed a similar development in math performances. L1: No negative results on the L1 proficiency.			70% of the instruction time was devoted to English, all subjects except for German language and culture, were taught in the L2.	School-issued tests and standardized tests.	No negative results on the L1 acquisition.	Teachers were native L1 speakers.	No information about student, teacher, or parental attitude was given in this study.	

⁴ Coloured Progressive Matrices

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Gebauer, Zaunbauer & Möller (2012)	N = 657 (351 imm.) 5 schools	6 yrs (grade 1) Tested in grade 1, 2, 3, and 4	Primary education in Germany. To compare the development of L1 reading fluency and L1 orthography skills between immersion and non-immersion students.	L1: German (30%) L2: English (70%)	L1: German language and culture L2: All other subjects	L1: receptive proficiency L2: productive proficiency Intelligence	Würzburger leise Leseprobe (WLLP) (Reading test) Hamburger Schreib-Probe (HSP) (Writing test) Raven's Coloured Progressive Matrices	Native German speakers.
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
L1: reading fluency of immersion students increased more quickly than that of non-immersion students. Both immersion and non- immersion students showed a similar development in spelling performance. Immersion education had no negative effect on L1 compared to conventional primary education.			This study addressed the implementation of L2 education in German primary schools. The L2 was used as a medium of instruction 70% of the time and this included all subjects except German language and culture.	Except for intelligence, where a standardised test (Raven's Coloured Progressive Matrices) was used, German regional tests were used to assess students' receptive and productive proficiency in L1.	L1: Immersion education had no negative effect on the L1 development. Also immersion students showed a faster increase in reading fluency than non-immersion students. No differences in spelling performance were observed.	Teachers were native L1 speakers.	No information about student, teacher, or parental attitude was given in this study.	

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Mehisto & Asser (2007)	<i>N</i> = not given ⁵ 4 schools	7 yrs (grade 1) Tested in grade 3	Kindergarten and primary education in Estonia. To determine factors contributing to the success of the immersion program, conditions required for the programs' sustainability and measures for the programs' development. To provide age-appropriate functional language proficiency.	L1: Russian (50%) L2: Estonian (50%)	Half of the subjects in L1/L2	Lesson observations to ascertain coherence between program goals, parental and teacher opinions, and best immersion/CLIL methodologies	-	Almost all native Estonian speakers or near-native fluency in Estonian. All have teaching experience, but not all teachers were experienced in CLIL/immersion teaching. Training was offered to (inexperienced) teachers. Apply a variety of teaching strategies and actively engage students in learning.
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
<p>On average, Estonian language students met the curriculum expectations in all subjects tested (mathematics, science, L1 and L2 proficiency).</p> <p>L1: grade 3 immersion students achieved better results in written tests than non-immersion Russian students.</p> <p>It was important to include parents in training or activities to raise awareness about the program.</p> <p>It was important to improve and maintain home-school</p>			<p>The main focus in this study was on stakeholder perspectives and program management, but it also gave an example of the implementation of L2 education for minority group students in Estonia.</p> <p>The teaching time was divided 50-50% over L1/L2.</p> <p>This research will be useful to the TPO pilot because in Estonia, the country's L1 was taught as an L2 to minority students. The same is likely to happen in the Netherlands, as some children will learn English as their L3 and Dutch as their L2.</p>	-	L1: grade 3 immersion students achieved better results in written tests than students in a non-immersion control group.	In this study the target language was taught to a minority population (1/3 of the Estonian population is Russian). Therefore, almost all teachers were native speakers of the language that was taught as an L2 to the students.	<p>Teachers: Overall, teachers were very satisfied with the CLIL/immersion program.</p> <p>Teachers used a wide variety of teaching strategies, but indicated there were different challenges involved in teaching boys and girls: 83% stated that the boys caused more problems than the girls. The article reports the need to increase teachers' repertoire of strategies for dealing with those differences between the sexes.</p> <p>Teachers found it important to reconsider the participation of academically weak, undisciplined or unmotivated students or evaluate the strategies required for those groups.</p> <p>Teachers considered training opportunities for the immersion program and teaching feedback as the key success factors.</p>	

⁵ The main aim of this study was the assessment of stakeholders' attitude. Therefore they do not specify the number of participants.

	<p>communication and cooperation and increase parental involvement, for example via training components for parental participation.</p>					<p>Parents: parents were positive about placing their children in immersion programs, because via this program they would achieve bilingual proficiency in the country's L1 and an additional language, the L2, which is important for their academic and professional future. 'General academic achievement' was also an important consideration when placing their children in an immersion program.</p> <p>In general, parents found it important to develop a dialogue with the teachers and over 50% wished to be updated more on topics like program development, student achievement, methodology, and teacher qualifications.</p> <p>Parents were satisfied about the program and their children's L1 development, even only half the time was devoted to the L1.</p>
--	---	--	--	--	--	---

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Bostwick (2001)	<i>N</i> = 33 (with a control group of non-immersion 4th grade students from the previous school years, resp. 46 and 48 students)	6 yrs (grade 1) Tested in grade 4.	Primary education in Japan. The program attempts to achieve a form of additive bilingualism that is designed to maintain and develop students' L1 skills while adding an L2. To test the effect of four years of immersion education in L1, L2, and mathematics.	L1: Japanese (50%) L2: English (50%)	L1: Japanese language and culture, music, social studies, art. L2: maths, general studies, PE, music, art.	L1: receptive and productive proficiency L2: receptive and productive proficiency	Japanese language test Written test, Eiken (STEP) test, LAS ⁶ , ITBS ⁷	Native English speakers. Trained and certified in their home country. An average of 8 years elementary school teaching experience before becoming an immersion teacher.
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
<p>L1: Immersion students performed as well as non-immersion students and nearly one SD above the national average.</p> <p>L2: 57% passed an English test designed for junior high students and another 10% a test intended for high school students.</p> <p>Immersion students were able to maintain academic achievements in maths at approximately one SD above national average.</p>			<p>This study examined the L2 proficiency of students in private and public primary schools in Japan.</p> <p>50% of the instruction time was in L2.</p> <p>Students had to learn a new alphabetical system.</p>	<p>L1 proficiency was examined with a Japanese language test.</p> <p>L2 proficiency was measured with both a Japanese designed test (Eiken (STEP) test) as well as internationally developed tests (LAS, ITBS).</p>	<p>L1: no negative effect was found on the development of L1, even when different alphabetical systems are concerned. Immersion students performed as well as non-immersion students, and nearly one SD above the national average.</p> <p>L2: L2 proficiency improved steadily in all skill areas over the years.</p> <p>The Eiken test halfway the fourth year of the program revealed a level at or above junior high level in Japan. The ITBS scores revealed approximately the same level as second grade students in the US, but this test is designed for students following an American-based curriculum, so scores are not directly transferable.</p>	<p>Teachers in this Japanese program came from English-speaking countries, such as the US, Canada, and Australia, and were trained in their respective home countries.</p> <p>L2 teachers were paired with Japanese teachers to work together and develop activities for the students.</p>	<p>Students: Generally had a very positive attitude. They enjoyed studying the L2 and if they had the option, they would study in a school with both Japanese and English. They developed a great confidence in their English ability.</p> <p>Parents: Initially had concerns about the immersion program, possibly due to this being the first in Japan, and by lack of information about program.</p> <p>In the end they stated the benefits of the program outweighed the difficulties.</p> <p>They chose the program for the bilingual possibilities and to broaden their child's perspectives.</p>	

⁶ Language Assessment Scale

⁷ Iowa Test of Basic Skills

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Turnbull, Hart & Lapkin (2003)	N = 3108	6 yrs (grade 1) Tested in grade 3 and 6.	Primary education in Canada. To test whether reading, writing, and mathematical skills of immersion students were comparable to non-immersion students.	L1: English (30-50%) L2: French (50-70%)	-	L1: receptive and productive proficiency Mathematics	EQAO ⁸ test	-
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
L1: Grade 3 immersion students scored as well or only somewhat better than non-immersion students. Grade 6 students outperformed the regular program students on all skill areas (reading, writing, maths). No negative effect of immersion education on reading, writing, and mathematics skills.		This study addressed a TPO program in Canadian primary education. 50-70% of the instruction was given in L2, which means some schools have implemented more L2 instruction than we aim for in this study.	The EQAO test is a test based on the Ontario curriculum testing various key points in (primary) education, including language development.	L1: no difference for either reading, writing, or mathematical skills between immersion and non-immersion students. Intensive L2 programs did not negatively impact upon L1 skills. Grade 3 immersion students scored equally well or slightly better than non-immersion students. Grade 6 immersion students signif. outperformed non-immersion students on all three skill areas.	-	No information about student, teacher, or parental attitude was given in this study.		

⁸ Education Quality and Accountability Office

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Thomas, Collier & Abbott (1993)	N = 2039 (719 imm.)	6 yrs (grade 1) Tested in grade 2 and 3	Primary education in the USA. To compare mathematics skills and L1 achievement of immersion and non-immersion students. During the first two years the only goal for the target language was to begin development of oral proficiency at no cost to the students' academic achievement in L1.	L1: English (50%) L2: Spanish/Japanese/French (50%)	L1: English language and culture, social studies L2: mathematics, science, health	L1: receptive proficiency Productive proficiency Mathematics	MAT ⁹ SOPR ¹⁰ POS Mathematics test	One teacher per language. L2 teacher proficient in target language. Certified to teach the lower elementary grades. Courses provided in teaching in foreign language immersion and possibility to participate in summer schools and training programs.
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
<p>L1: immersion students signif. outperformed the non-immersion students.</p> <p>L2: Japanese immersion students improved to same extent or more than their counterparts in Spanish or French immersion on English and mathematics tests.</p> <p>Mathematics: immersion students did as well as or better than the non-immersion students. Statistically signif. differences are usually only seen after the third or fourth year.</p>		<p>This study described the first three years of TPO programs on primary schools in the USA.</p> <p>50% of instruction was taught in the L2.</p> <p>L2 immersion in languages with a similar writing system (Spanish, French) and a different writing system (Japanese).</p>	<p>The MAT test is a standardized test for children in kindergarten through twelfth grade that tests topics like maths, science, and language.</p> <p>The SOPR is a standardized performance measure and a modification of the Student Oral Language Observation Matrix (SOLOM), using teacher judgment on a rating scale, to rate oral proficiency in a foreign language.</p>	<p>L1: immersion students signif. outperformed non-immersion students on tests of receptive (reading) proficiency.</p> <p>L2: Japanese immersion students improved to same extent or more than their Spanish and French immersion counterparts on all measures in the L1 and in mathematics.</p> <p>L2 Japanese immersion students scored better on the grade 2 maths test than students in a local immersion comparison group, which is an important finding since not much research was previously conducted on acquisition of academic subjects through a non-Roman-alphabet language.</p>	<p>Specialist language teachers, but no native speakers. All teachers were certified and had the possibility to take part in training programs and summer schools to improve their foreign language teaching skills.</p>	<p>Parents: Student enrollment in the program was by parental choice and response was so overwhelming that not every student could be accepted to the immersion program, indicating that parents were very positive about teaching a foreign language through immersion education in school.</p>		

⁹ Metropolitan Achievement Test

¹⁰ Student Oral Proficiency Rating

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Cambell, Gray, Rhodes & Snow (1985)	N = 382 (98 imm.)	6 yrs (grade 1)	Primary education in the USA.	L1: English (70%)	L1:	L1:	MLA ¹¹ test Self-assessment	-
	15 schools	Tested in grade 4.	To examine the efficacy of three foreign language learning programs (FLES, part. imm, full imm.) at the elementary school level.	L2: Spanish (30%)	L2: "a portion of the school curriculum is taught in the foreign language"	L2: receptive and productive proficiency		-
	Results		Relevance to TPO pilot					
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
L2: partial immersion students score signif. better than regular L2 students on listening and speaking, and slightly better on reading and writing. Overall, full immersion students achieved the highest scores and outperformed partial immersion students and regular L2 students on all measures.		This study examines the effect of various foreign language learning programs in the USA, including partial immersion, and compares the efficacy of them. The time dedicated to teaching the L2 through a partial immersion program is approximately 30%.	The MLA test was originally developed to test language proficiency in secondary education, but was also used in this study because it was the only test that assessed all four language skills (reading, writing, listening, speaking). A previous pilot study with primary school immersion students was performed to verify if this test was appropriate to use with this group of students. However, two schools stated that the writing task was beyond the capabilities of their fourth grade students.	When comparing the L2 achievement level in grade 4, full immersion students obtained the highest scores, outperforming both partial immersion and regular L2 students. Concerning the present study, partial immersion students signif. outperformed regular L2 students on listening and speaking tests and also scored somewhat, but not signif., better on reading and writing tests.	-	No information about student, teacher, or parental attitude was given in this study.		

¹¹ Modern Language Association Cooperative

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Christian (1996)	<i>N</i> = ? 169 schools	Varies. Many US schools start their two-way immersion in kindergarten and some later (in primary school)	Two-way immersion in the USA. This study describes the current state of two-way immersion in the US.	L1: English (10-50%) L2: Spanish (and few others incl. Japanese/Cantonese/Russian) (50-90%)	L1: varies. Many subjects are taught 50-50 in L1/L2 L2: varies	-	-	Teachers employed a wide variety of instructional approaches and strategies.
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results	4. Didactic characteristics	5. Attitudes		
Many different forms of two-way immersion in the USA. 75-92% of the nonnative-English-speaking 5 th and 6 th graders were rated as fluent in English.		This study describes the various forms of two-way immersion in the USA.	-	Describes other US studies on two-way immersion.	Teachers should at all times be aware that they should make use of 'sheltered instruction' in order to clarify what they are saying because there are always students who are being taught in their L2.	Students develop friendships in the classroom quite independent of race or ethnicity and program. Two-way immersion might foster positive cross-cultural attitudes and self-esteem among students.		

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information						
Cirino et al. (2007)	N = 1451 35 schools	5/6 yrs (grade 1) Tested at the end of kindergarten	Primary education in the USA. To test the influence of teacher characteristics on student outcomes.	L1: Spanish (% differs among teachers/schools) L2: English (% differs among teachers/schools)	L1: - L2: -	L1: receptive and productive proficiency, vocabulary L2: receptive and productive proficiency, vocabulary	School-issued test TOPP-S ¹² WLPB-R ¹³ Experimental Word List School-issued test CTOPP ¹⁴ WLPB-R LEXESP ¹⁵	Experience: average of 9.7 years. Credentials: 2/3 bilingual or English as a Second Language (ESL), 1/3 other credentials or unspecified.						
									Results	Relevance to TPO pilot				
										1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes
	Positive correlations between teacher-quality ratings and student engagement. Initial (baseline) student and classroom performance and the oral language proficiency of the teacher in both Spanish and English predicted student outcomes.	This study describes the implementation of English in native-Spanish classrooms. Teacher characteristics were examined to see to what extent they affected the kindergarten students' language and literacy end- of-year outcomes.	School-issued tests were used. The CTOPP test was used partially; five out of the nine subtests were used in order to measure reliable outcomes. The TOPP-S test was a Spanish equivalent of the CTOPP test. The WLPB-R was used to measure oral proficiency. Children were also observed through teacher observation. They included SES in their analysis.	Positive correlations between teacher-quality ratings and student engagement. Teacher quality was related to word- reading skills, but teacher knowledge in either language was unrelated to these student literacy outcomes.	2/3 bilingual or English as a Second Language (ESL), 1/3 other credentials or unspecified. The teacher should have a high oral language proficiency in both languages. Teacher knowledge did not produce strong relations to outcomes.	No information about student, teacher, or parental attitude was given in this study.								

¹² Test of Phonological Processing - Spanish

¹³ Woodcock Language Proficiency Battery – Revisited (in both Spanish and English)

¹⁴ Comprehensive Test of Phonological Processing

¹⁵ Database of printed words frequencies in Spanish

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Montecel & Cortez (2002)	N = 219 – 1848 10 schools (8 elementary, 2 high schools, 1 middle school)	Differed among schools	Primary education in the USA. To identify common characteristics and criteria that are contributing to the high academic performance of students served by bilingual education programs.	L1: Spanish/ Navajo/Russian L2: English	L1: - L2: -	L1: receptive and productive proficiency, L1: content knowledge L2: content area mastery	- -	Fully qualified and English as a Second Language (ESL) teachers. Continuously acquiring new knowledge regarding best practices in bilingual education.
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results	4. Didactic characteristics	5. Attitudes		
By fifth grade, students were fully proficient in speaking, reading and writing in both the L1 and the L2.		This study identified 25 common characteristics and criteria contributing to the high academic performance of students in 10 bilingual education programs. These characteristics can serve to guide teachers and administrators.	-	L1: if students took part in the bilingual education since kindergarten, they were fully proficient in speaking, reading and writing by fifth grade. L2: if students took part in the bilingual education since kindergarten, they were fully proficient in speaking, reading and writing by fifth grade.	All teachers were fully qualified and were continuously informed about the recent findings in bilingual education.	All of the staff, parents and students felt responsible for maintaining a safe and orderly climate. Racial and cultural differences were valued. Parents should know the rationale and the critical components of bilingual and ESL programs and are strong advocates of the programs. Staff held themselves accountable for the success of all students.		

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Peregoy & Boyle (1999)	N = 61 2 schools	3-6 yrs (Kindergarten) Tested in kindergarten	Primary education in the USA. To discover how teachers structure opening activities in order to ensure comprehension and participation by children.	L1: English/Spanish (%) L2: English/Spanish (%)	English: English language and culture Spanish: all other activities	L1: - L2: -	-	-
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results		4. Didactic characteristics		5. Attitudes
Routines are scaffolded to help L2 learners participate appropriately.		Two bilingual Spanish immersion classrooms are studied in terms of how the teachers use structured opening activities to ensure comprehension and participation.	Routines are analyzed in terms of scaffolding provided.	-		Teachers model language and reading behavior, plus questions and commands.		No information about student, teacher, or parental attitude was given in this study.

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Reder, Daigle & Demont (2013)	<i>N</i> = 76 (33 imm.)	4/5 yrs (Kindergarten) Tested in grade 2.	Primary education in France. To determine whether early L2 exposure improves metalinguistic awareness. To examine the development of metalinguistic abilities between immersion and non-immersion students.	L1: French (50%) L2: German (50%)	L1: history, geography L2: mathematics, science	L1: receptive and productive proficiency	French standardized phonological test Morphological test Syntactic test	One teacher per language.
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
L1: early exposure to L2 has a positive impact on children's ability to reflect on morphological and syntactic features. The gap between immersion and non- immersion students increases over time.		This study focusses on the development of metalinguistic (morphological, syntactic, and phonological) awareness in L1 between immersion and non-immersion students in French primary education. Time spent on teaching L1 and L2 is divided 50-50%.	All tests were developed. Immersion and non- immersion students were matched based on the PPVT and Raven's Coloured Progressive Matrices. The tests conducted to measure their metalinguistic abilities were all French (standardized) tests.	Immersion students were better able to recognize and reflect on syntactic and morphological features than non-immersion students. No differences were observed in phonological awareness, most probably because the L1 and L2 in this study share similar sound patterns, but differ greatly in morphological compounds and syntactic features. The results indicate that metalinguistic awareness develops right from the beginning of primary education, but is not fully developed yet by the end of the second grade, so may continue to improve further during the third year.	Specialist language teachers, but not native speakers of the L2.	No information about student, teacher, or parental attitude was given in this study.		

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Serra (2007)	<i>N</i> = 44 (imm.) 3 schools	6 yrs (grade 1) Tested grade 1, 2, 3, 4, 5, and 6	Primary education in Switzerland. To evaluate the process of L2 acquisition and the development in L2 and mathematics.	L1: German (50%) L2: Italian/Romansh (50%)	L1: - L2: notably mathematics	L2: productive proficiency receptive proficiency	Verbal interactions between pupils and researcher, narrative task Oral and written comprehension tests	One teacher per language Bilingual teachers Teaching experience, but not necessarily bilingual teaching experience
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
L1: By the end of grade 6, all immersion classes reached equally satisfactory performances. Immersion students scored as well as or better than non-immersion students on oral and written comprehension. Mathematics: All three immersion groups scored signif. better than the control groups. No differences between the immersion groups was found.		This study addresses the process of L2 and mathematics development in Swiss primary schools. 50% of the time is spent on teaching the L2.	The narrative and comprehension tests were designed for the purpose of this study.	By the end of primary school, both immersion and non-immersion students reached equally satisfactory performances, indicating there is no negative effect of immersion education on the L1 performance. Immersion scored better in various comprehension tests, indicating that an L2 supports the thinking process.	There is one teacher per language, but both teachers are bilingual. They have valuable primary school teaching experience, but not all teachers are experienced with immersion education yet.	No information about student, teacher, or parental attitude was given in this study.		

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the program	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Dobson, Pérez-Murillo & Johnstone (2010)	N = 1200 (imm.) 44 schools	3 - 4 (preschool) Tested at the end of primary school and beginning of secondary school	Primary education in Spain Provide research-based evidence on students' L2 proficiency as developed and demonstrated through the study of subject matter in a bilingual context; and on their achievements in L1. To identify and disseminate good practice as occurring in the project schools To provide research-based evidence on awareness, attitudes and motivation	L1: Spanish (60%) L2: English (40%)	L1: - L2: science, history, geography	L1: productive proficiency L2: productive and receptive proficiency	L1: written test, Cambridge IGCSE ¹⁶ L2: interviews, analysis of lessons (especially the analysis of the teacher and language skills during communication in the classroom), descriptive-narrative writing task, IGCSE	Students were taught by <i>funcionarios</i> (teachers with civil service status and conditions) who were appointed after a number of competitive examinations. Native (or near-native) command of both spoken and written English European Qualified Teacher Status in infant/primary teaching Very experienced teachers: over 3/4 had over 10 years of classroom experience
Results		Present study						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
Immersion students were better at writing in both their L1 and L2 than their peers in a non-bilingual classroom.		This study describes an independent 3-year-investigation into the Ministry of Education/British Council Bilingual Schools project.	Oral interviews, written tests, IGCSE. The Cambridge IGCSE examinations were conducted in Spanish 1 (Spanish for mother tongue speakers), English 1 and 2 (a test for native speakers as well as a test for English language learners) and also in various subjects (e.g. geography, history) They included SES in their analysis.	L1: performance of the immersion students was significantly better than non-immersion students where writing was concerned L2: immersion students show better performances than non-immersion students. All students passed the IGCSE subject tests in geography, history and biology. Those tests were conducted in English, indicating that immersion students were able to tackle subject matter in a foreign language.	Native (or near-native) command of both spoken and written English European Qualified Teacher status in infant/primary teaching Classroom experience within a range of three to eleven years	Students: the great majority ... -- is (very) happy with the experience and found it (very) interesting; -- said it was (very) useful; -- is (very) confident while learning through English; -- believes the experience has helped to develop self-confidence and has helped them broaden their understanding of subjects learnt at school; -- says it has helped to broaden their understanding of Europe and Spain Parents ... -- 88% held a (very) favourable view of the BEP; -- 86% believed the BEP had contributed to their child's proficiency in English;		

¹⁶ International General Certificate of Secondary Education for 14-16 year old students

						<p>-- 55% believed studying other subjects in Spanish had helped their child's progress in these subjects, and 43% believed this about being taught in English;</p> <p>64% believed the BEP had helped their child to understand other countries better, and 44% believed it had enabled their child to make contacts with other countries.</p> <p>Teachers ...</p> <p>-- considered the program (highly) beneficial for students (96%), teachers (97%) and schools (99%).</p>
--	--	--	--	--	--	--

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information	
Lasagabaster (2000)	N = 252 (42 part. imm.)	L2: 6 yrs (grade 1) L3: 9 (grade 4) Tested grade 5 (and 8)	Primary education in the Basque Country To examine the effect of the degree of bilingualism on the acquisition of English (L3)	L1: Spanish (40%) L2: Basque (40%) L3: English (20%)	Varies depending on teacher	L1: receptive and productive proficiency L2: receptive and productive proficiency L3: receptive and productive proficiency	Standardized test (Galbahe) Standardized test (Galbahe) Grammar and vocabulary test, reading and listening comprehension test, oral test	Most teachers were qualified to teach in L2 Basque and had completed a degree in English studies (+ teacher training degree to teach in primary education). Possibility to take extra courses and attend summer schools.	
Results		Relevance to TPO pilot							
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes			
		L1: Partial immersion (and also full immersion) students obtained higher scores in L1 competence than regular programme students, although they received less teaching in L1. No neg. effect on L1. L2: Partial (and full) immersion students signif. outperformed regular programme students in L2 competence. L3: Partial immersion students signif. outperformed regular programme students on reading and also scored slightly higher on speaking and grammar.	Unlike most other studies, this study examines L2 <i>and</i> L3 acquisition in primary schools, in the Basque Country in Spain. Basque is taught as a second, but not a foreign, language. English is taught as a third, but first foreign, language. During the first years of primary school, L1 and L2 are both used for 50% of the time. When English is added the languages are divided 40/40/20%.	The tests to assess the L1 and L2 proficiency were standardized tests issued by the Basque government. No further details about the L3 proficiency tests was given.	This study reveals better results for immersion students in all three languages. L1: Immersion (full and partial) students obtained higher L1 scores, while they received much less instruction in this language than non-immersion students. No negative influence of immersion education on the development in L1. L2/L3: Immersion (full and partial) students signif. outperformed non-immersion students on various language skills.	Teachers are mostly bilingual in L1 and L2, but not necessarily native speaker in this L2. The level of L3 competence usually wasn't very proficient, but in this study teachers have completed a degree in English studies and are offered extra courses to improve their L3 proficiency.	Another study by Lasagabaster (2001) describing the same language situation states that some parents, especially from students in the full immersion program, were concerned that adding an L3 would only cause extra confusion and hinder the survival of the local minority language. Being bilingual supports learning an L3 and multilingual students also have other cognitive advantages over monolingual students. Jessner and Cenoz (2007) also found that children have a very positive attitude toward an early instruction in English in primary education.		

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Lorenzo, Casal & Moore (2009)	<i>N</i> = 525 (380 imm.) 403 schools (incl. German and French schools)	±7 yrs Tested in grade 4	Primary education in Spain. To present key findings from bilingual projects in the South of Spain including both the competence development in L1/L2 and organization of the program.	L1: Spanish (± 50%) L2: English (French/German) (± 50%)	L2: three subjects (depending on the teacher profiles)	L2: receptive and productive proficiency	Skill-based diagnostic tests	Content teachers teaching in L2, L2 specialist teachers, and native language assistants. L2 specialists and language assistants used more L2 in conversational exchange than content teachers.
	Results	Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
	L2: results incorporate both primary and secondary students; students in the immersion program signif. outperformed the non-immersion students on all language skills.	This study examines the L2 competence in Spanish primary schools. A maximum of 50% of the time is dedicated to the L2.	Various skill-based diagnostic tests were developed to assess the students' L2 competence.	L2: students in immersion programs signif. outscore non-immersion students on all tested language skills.	In addition to the content and specialist language teachers, there were also (native-speaker) language assistants. All teachers seem to deal with a specific area of language expertise, whereby (near)native speakers focused more on conversational praxis and the content teacher more on the textual level. This provided the students with a very rich learning environment.	No information about student, teacher, or parental attitude was given in this study.		

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Cenoz (2005)	<i>N</i> = 45,807	3 yrs (grade 1)	Bilingual and trilingual education in the Basque Autonomous Community in Spain. To examine the use of English, the teaching conditions, to summarize results of research on multilingual schooling in the Basque Country and to examine challenges of the multilingual system.	L1: Spanish (40%) L2: Basque (40%) L3: English (20%)	L1: - L2: -	L1: - L2: -	- -	Certificate of education from a teachers' college. Teachers who teach Basque or through the medium of Basque need a certificate of Basque.
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
L1: not adversely affected by the students' acquisition of English. L2: being bilingual was associated with higher scores in general writing proficiency, syntactic/lexical complexity, fluency and error production.		This study examined both bilingual and trilingual education in various schools in the Basque Country.	-	L1: not adversely affected by the students' acquisition of English. L2: being bilingual was associated with higher scores in general writing proficiency, syntactic/lexical complexity, fluency and error production.	Certificate of education from a teachers' college. Teachers who teach Basque or through the medium of Basque need a certificate of Basque.	Younger learners were more motivated toward learning English. The attitude toward learning three languages declined at the end of primary school. The article provided two possible reasons for this: psychological or educational, the latter referring to observation that children see the educational methods as traditional and less active.		

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Skolverket (2010)	N = 9400 42 schools	6 yrs (grade 1) Tested at the end of primary education	Primary education in Sweden. To examine the implementation of English in primary education.	L1: Swedish (50%) L2: English (50%)	L1: other subjects L2: mostly social studies and science, but varies per school	L1: - L2: -	National tests.	Two-third of the teachers teaching subjects in English had (part of their) education in an English-speaking country or Swedish education in combination with studies in English. Part of the immersion teachers had an English background or English educational background. About 10% had experience with immersion teaching.
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
Students in an immersion program obtained on average better results on national tests than students in non-immersion programs. Students in immersion programs obtained higher scores in the subjects taught in L2 in comparison to non-immersion students		This large scale study describes the situation concerning immersion education in primary schools in Sweden. In immersion schools 50% was taught in the L2.	All tests were part of the Swedish national test curriculum.	Immersion students achieved (signf.) higher scores on various skills in the national test than non- immersion students. Immersion students also obtained higher scores in the subjects they were taught in L2 than non- immersion students who were taught in their L1.	All teachers in this study had considerable teaching experience in primary education. About 10% of them also had experience with immersion teaching. Two-third of the English teachers had followed education in an English-speaking country or had a degree in English studies. A part of the immersion teachers had an English (educational) background.	Students, in the first place, found it fun to learn English and were eager to learn more English. Most students also believed it was useful for future studies and career. Students in regular primary education stated that they would wanted to receive more English education. Most parents considered English an important language in today's globalized world. Parents had different opinions about English education depending if their child was in an immersion program or not. Parents with children attending regular primary schools mentioned the value of a good learning environment, while parents with children in immersion programs often had an international background or decided on immersion education because of the need for English language skills in modern world., future needs, and work opportunities.		

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Seikkula- Leino (2007)	N = 217 (116 imm.) 10 classes	6 yrs (grade 1) Tested in grade 5 and 6.	Primary education in Finland. To investigate how successfully students can learn content through immersion programs. To examine how well students fared in different teaching methods in relation to their possible learning competence.	L1: Finnish (30-60%) L2: English (40-70%)	L1: - L2: -	L1: receptive proficiency Mathematics Intelligence Verbal intelligence	IEA reading skill assessment (PIRLS) National mathematics test Raven's non-verbal intelligence test Wechsler's vocabulary test	-
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
<p>L1: no sign. Differences between immersion and non-immersion students and no negative effects in L1.</p> <p>No signif. differences in general learning outcomes between the two groups, indicating that both weak and strong students could perform as well in both types of education.</p> <p>L2: Immersion students tend to overachieve less (according to their previously tested intelligence level) than non-immersion students. Indicating that learning an L2 may demand a lot of students' cognitive abilities.</p>			<p>This study examined immersion programs in primary schools in Finland.</p> <p>It assessed the achievement levels of immersion and non-immersion students and also looked at the relationship between the teaching method and their general achievement outcomes.</p> <p>Depending on the school 40-70% of the time the L2 was used as language of instruction.</p>	<p>The international IEA test was used to assess L1 reading proficiency.</p> <p>Independent factors such as intelligence and verbal intelligence were measured were tested with official tests as well.</p> <p>Mathematics achievement was tested with a national test.</p>	<p>L1: no differences were found on the L1 reading proficiency between immersion and non-immersion students.</p> <p>L2 proficiency was not measured with a test. On the other hand, they achieve average results in L1, comparable to non-immersion students, so all gains that are made in the L2 are extra benefits of the programme.</p>	-	<p>Students' self-evaluation in immersion programs was more clearly marked by self-criticism as they felt they were weaker in all language skills compared to non-immersion students. This is remarkable since they had have acquired a good knowledge in the L2. A possible explanation is that they feel less competent because of the communicative problems they are facing when learning the language and therefore might feel frustrated by this.</p>	

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Merisuo-Storm (2007)	N = 136 (78 imm.) 3 schools	6 yrs (grade 1) Tested in grade 2.	Primary education in Finland. To investigate the effects of bilingual teaching on the development of students' literacy skills and attitude.	L1: Finnish (80%) L2: English (20%)	-	L1: receptive and productive proficiency L2: receptive proficiency	Various tests, including auditory and visual perception, written task, mathematical skills Oral and written recognition	-
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2		4. Didactic characteristics	5. Attitudes	
<p>L1: immersion students scored signif. better than non-immersion students on reading comprehension and equally well or higher on writing performance; no negative effect on L1.</p> <p>Students with the highest L1 literacy obtained the highest scores in the English test, indicating a close connection between students' literacy level and L2 acquisition.</p>		<p>This study describes the effects of immersion education in primary schools in Finland.</p> <p>20% of the instructional time is in the L2, somewhat lower than we aim to assess in the present study, but since a specific attitude evaluation was carried out as well, it can hopefully provide some useful information on that topic.</p>	<p>Various Finnish L1 and L2 tests were administered, but no specific details were given in this study.</p>	<p>L1: Immersion students achieved signif. higher scores than non-immersion students on reading comprehension and equal or higher scores on writing proficiency scores than non-immersion students. No negative effects were found.</p> <p>Immersion students could read with greater speed than non-immersion students at the end of grade 2.</p> <p>In the group of students with the lowest literacy skills, those who took part in an immersion program achieved higher reading accuracy and comprehension scores in their L1 at the end of grade 2 than those in a non-immersion program.</p> <p>L2: However, a close connection between literacy development and L2 acquisition was observed, whereby students with a poor literacy level were less successful in L2 performance.</p>		-	<p>Students enjoyed using the L2 in test situations and they were proud of their L2 knowledge.</p> <p>An attitude study was carried out two years after the first study to find out if there was a difference between students in immersion and non-immersion programs. No signif. differences in attitude towards reading and writing were found between immersion and non-immersion students. However, a difference in language learning was found, whereby immersion students showed a more positive attitude than non-immersion students.</p> <p>Students in the immersion program found using the English language in everyday context natural and felt comfortable using it.</p> <p>All students considered a good knowledge of English important.</p> <p>Boys had a more negative attitude than girls, but this difference was smaller (and not signif.) in the immersion classes.</p>	

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Jäppinen (2005)	<i>N</i> = 669 (335 imm.) Incl. 1/3 sec. education <i>N</i> = 456 (229 imm.) 12 schools	7 yrs (grade 1) Tested in grade 3 and 5.	Primary education in Finland. To examine the thinking and learning processes of mathematics and science taught in L2.	L1: Finnish (0-80%) L2: English/French/ Swedish (20-100%)	L1: Other subjects L2: mathematics, science	L1: various skills	Problem solving, picture tasks, generalisations, metaphors, proverbs, tables, graphics, relations, similarity.	-
Results		Relevance to TPO pilot						
		1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes		
<p>Primary school tests were divided over two age groups: 7-9 and 10-12 years old.</p> <p>First group: no (signf.) differences in mathematics and science results between the immersion and non-immersion students, indicating that L1 and L2 teaching result in the same level of development.</p> <p>Second group: the starting level was the same, but in the end the immersion students outperformed the non-immersion students in both mathematics and science, indicating that L2 teaching seemed to support the required thinking and learning processes.</p>		<p>This study examined the thinking and learning processes of immersion and non-immersion primary school students in Finland.</p> <p>Depending on the age group, L2 was used as language of instruction for 20-100% of the time.</p>	<p>Finnish tests were used to assess the students' thinking and learning processes.</p>	<p>The influence of immersion and non-immersion teaching was evaluated by testing students mathematics and science achievements.</p> <p>In the first age group (the first years of primary education), no differences in mathematics and science results were measured between immersion and non-immersion students.</p> <p>In de second age group (the last years of primary education), immersion students signf. outperformed non-immersion students.</p>	-	<p>No information about student, teacher, or parental attitude was given in this study.</p>		

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
Ytsma & Beetsma (2001)	N = 281 7 schools	4 yrs (Preschool) Tested in grade 1, 2, 3, 4, 5, and 6.	Primary education in the Netherlands. To examine the implementation of a third language whereby the general educational goals will still be achieved.	L1: Dutch (50%) L2: Frisian (50%) Grade 5/6 L1: Dutch (40%) L2: Frisian (40%) L3: English (20%)	L1/L2: varies L3: geography, arts	L1: Dutch language proficiency L2: Frisian language proficiency	-	Sometimes one teacher per language. (Near) native speaker for L3/English.
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
<p>L1 Frisian-speaking children with Dutch as their L2 and L1 Dutch-speaking children with Frisian as their L2 participated in this study.</p> <p>By the end of preschool, L1 Frisian-speaking and L1 Dutch-speaking children show equal scores on Dutch language tests. Immersion education does not negatively affect L1 development.</p> <p>L1 Frisian-speaking children obtain higher proficiency scores in Frisian tests than L1 Dutch-speaking children, implying L1 Frisian-speaking children reach balanced bilingualism before the start of primary school, while L1 Dutch-speaking children still lag behind in their Frisian fluency.</p>			<p>This study shows the attitudes of the parents and children towards the trilingual programs. Furthermore, it discusses the model and the results of an interesting study during which six students from Belfast taught the children for over three months, completely immersing them in English.</p>	-	<p>L1: By the end of preschool, L1 Frisian-speaking and L1 Dutch-speaking children show equal scores on Dutch language tests. Immersion education does not negatively affect the L1 development.</p> <p>L2: L1 Frisian-speaking children obtain higher language proficiency scores in Frisian tests than L1 Dutch-speaking children, implying L1 Frisian-speaking children reach balanced bilingualism before the start of primary school, while L1 Dutch-speaking children still lag behind in their Frisian fluency.</p>	<p>Teacher must be (near) native speakers of the English language.</p>	<p>Students: The evaluation of the project has shown that 77% of the students wanted to continue their education in English after they had been taught in English only for three months.</p> <p>Parents: 95% of the parents found it very useful for their children to learn Frisian, and 70% said the same about English.</p>	

Authors	Participants	Age of onset/ Age of testing	Context & Aim of the study	Language allocation	Subjects L1/L2	Tested construct	Instruments	Teacher information
van Ruijven & Ytsma (2008) Riemersma & de Jong (2007)	<i>N</i> = not given	4 yrs (Preschool) Tested in grade 1, 2, 3, 4, 5, and 6.	Primary education in the Netherlands. To examine the results of three languages in Frisian primary education.	Grade 1/4 L1: Dutch (50%) L2: Frisian (50%) Grade 5/6 L1: Dutch (40%) L2: Frisian (40%) L3: English (20%)	L1/L2: Varies L3: geography, arts	L1: receptive proficiency L2: receptive proficiency	Vocabulary test, technical reading, spelling, reading comprehension Test bilingualism for toddlers. 'toets tweetaligheid voor kleuters' Vocabulary test, technical reading, spelling, reading comprehension	-
Results			Relevance to TPO pilot					
			1. Examples	2. Evaluation instruments	3. Results on L1/L2	4. Didactic characteristics	5. Attitudes	
<p>L1: the same results were obtained by students in immersion and non-immersion education throughout primary education.</p> <p>L2: during the first years of primary school there were no differences, but the second half of primary education immersion students showed better achievements than non-immersion students.</p> <p>Students in immersion education scored above the regional and national average.</p> <p>L3: No differences between project schools and control schools.</p>			<p>This study evaluates the results of a trilingual program. The students are taught using three different languages.</p>	<p>School-issued test, testing language proficiency in both Dutch and Frisian. In grades 7 and 8, the children were also tested on their proficiency in English.</p>	<p>L1: the same results are obtained by students in immersion and non-immersion education throughout primary education.</p> <p>L2: during the first years of primary school, there are no differences, but in the second half of primary education immersion students show better achievements than non-immersion students.</p> <p>Students in immersion education score above the regional and national average.</p> <p>L3: No differences between project schools and control schools.</p>	-	<p>No information about student, teacher, or parental attitude was given in this study.</p>	