

# 10. Bedöma språk eller innehåll? Språkets roll vid bedömning i språk- och ämnesintegrerad undervisning

Helena Reierstam

## Inledning

Är det möjligt att bedöma elevers ämneskunskap utan att samtidigt bedöma språklig förmåga? Eller omvänt, går det att bedöma språklig förmåga utan någon inblandning av ämnesmässigt innehåll? Relationen mellan språk och innehåll framträder extra tydligt, och blir samtidigt en utmaning, vid bedömning i en tvåspråkig undervisningskontext: Vad är det som bedöms? I det här bidraget, som sammanfattar min licentiatuppsats (Reierstam, 2015), diskuteras dessa frågor med utgångspunkt i en studie om lärares skriftliga bedömningspraktik i engelska, biologi och historia i tre gymnasieskolor där engelska används som undervisningsspråk istället för svenska. I internationella sammanhang betecknas detta sätt att arbeta ofta som *Content and Language Integrated Learning*, förkortat CLIL (Dalton-Puffer, 2007; Mehisto et al., 2008). I Sverige har det ibland gått under namnet Språk- och ämnesintegrerad undervisning, SPRINT (Nixon, 2000). Syftet med CLIL är att eleverna ska tillägna sig ett språk samtidigt som de lär sig ämnesinnehåll. Språk- och ämnesintegrerad undervisning kan innebära allt från enstaka ämnesövergripande projekt till sammanhang där i huvudsak all undervisning bedrivs på ett annat språk. I den aktuella studien har eleverna merparten av sin undervisning på engelska. Studien ingår i ett större projekt finansierat av vetenskapsrådet, CLISS-projektet, Content and Language

---

### Hur du refererar till det här kapitlet:

Reierstam, H. (2021). Bedöma språk eller innehåll? Språkets roll vid bedömning i språk- och ämnesintegrerad undervisning. I Bardel, C. et al. (Red.). *Forskarskolan FRAM — lärare forskar i de främmande språkens didaktik* (s. 219–244). Stockholm: Stockholm University Press. DOI: <https://doi.org/10.16993/bbg.k>. License: CC-BY 4.0.

Integration in Swedish Schools, med syfte att studera CLIL i svensk skola (för en närmare beskrivning, se Sylvén, 2019; Sylvén & Ohlander, 2014).

Förhållandet mellan språk och ämnesinnehåll vid bedömningstillfället har utpekats som ett viktigt område att studera i CLIL-sammanhang (Gablasova, 2014; Hönig, 2009; Massler et al., 2014; Quartapelle, 2012; Wewer, 2014), då detta är tämligen outforskat. Det blir relevant att fråga sig om hänsyn tas till eventuella språkliga begränsningar vid bedömning av ämneskunskap, så att inte språket utgör ett hinder då elever ska visa sitt ämnesrelaterade kunnande. Det har också visat sig att ämneslärare inte vet hur de ska undervisa språk i sina ämnen, eller vilka språkliga register eleverna behöver tillägna sig (Clegg, 2007; Reierstam, 2015).

Upprinnelsen till den aktuella studien och intresset för språk- och ämnesintegrering finns i författarens bakgrund som språklärare i olika skolformer. I undervisningen i engelska och franska har jag haft som mål att motivera eleverna genom autentisk språkanvändning i uppgifter med ämnesövergripande innehåll. I undersökningar som gjorts där elever har tillfrågats vad de anser vara motiverande för språkinläring, anger de uppgifter som känns autentiska, där de får möjlighet att använda språket på riktigt (Erickson, 2010). ”På riktigt” kan innebära att exempelvis beställa mat på restaurang, men också att skriva en laborationsrapport. Från språkklassrummet ledde intresset för autentisk språkanvändning vidare till en nyfikenhet på språkets roll i alla ämnen. En viktig inspirationskälla infann sig under en period då jag undervisade i yrkesengelska, men även i språk- och ämnesintegrerade mindre projekt med ämneskollegor. När det blev dags för bedömning uppstod frågor angående vad som skulle bedömas, hur det skulle göras och av vem, språkläraren eller ämnesläraren? De två första frågorna utgör huvudfrågor i den aktuella studien, rörande *vad* som bedöms och *hur*. Efter studiens avslut framträder stora beröringspunkter mellan CLIL och undervisning av elever i andra sammanhang med ett annat första språk än svenska. Frågor om språk- och kunskapsutvecklande arbetssätt har också fått alltmer utrymme i olika undervisningssammanhang och i forskning (Hajer & Meestringa, 2014; Skolverket, 2012b).

## Bakgrund

Förhållandet mellan språk och ämnesinnehåll har pekats ut som ett viktigt område att utforska vid bedömning i CLIL-sammanhang, vilket nämdes inledningsvis. Det råder brist på forskning kring lärares bedömningspraktik i tvåspråkig och ämnesintegrerad kontext, särskilt avseende språkets roll vid bedömning av ämneskunskap. I den här delen kommer jag redogöra mer för vad forskningen säger om CLIL men även beröra den aktuella diskursen rörande språk i alla ämnen och teorier som ligger till grund för forskningsmetoden. Avsnittet inleds med en diskussion om begreppen *språk* och *innehåll* samt vad de representerar i olika ämneskontext.

Språkets betydelse för lärande har, vilket redan nämnts, kommit att aktualiseras, delvis utifrån allt fler elevers flerspråkiga bakgrund, men också för att forskning påvisar språkets roll som redskap för lärande i alla ämnen (Cummins, 2007; Gibbons, 2013; Schleppegrell, 2004). I språkundervisning utgör språket både mål och medel. Fokus skiftar mellan språkets form och förmågan att använda språket för att interagera och göra sig förstådd. I den svenska kurs- och ämnesplanen för engelska (Skolverket, 2011) framgår exempelvis att undervisningen ska syfta till att eleverna får tillfälle att använda språket i funktionella och meningsfulla sammanhang för att utveckla en allsidig kommunikativ förmåga. I alla ämnen finns samtidigt ett ämnesinnehåll som eleverna ska tillägna sig, där språkförmågan anger elevers möjligheter att visa vad de kan. Språk och ämnesinnehåll kan därmed sägas utgöra två sidor av ett mynt, där den ena är intimt sammankopplad med den andra (Mohan, 1986).

Tidigare studier har diskuterat vilket språk som ska användas vid bedömning i tvåspråkiga undervisningssammanhang, elevens första språk eller undervisningsspråket, liksom vad som definierar andraspråket som man behöver ta hänsyn till vid bedömning (Gablasova, 2014; Hönig, 2009; Kiely, 2009; Morgan, 2006). Om elever tillåts använda sitt första språk vid bedömningstillfällen uttrycker de sig måhända mer ledigt, men viktiga ämnesbegrepp kan saknas eftersom de har lärt sig dessa på undervisningsspråket. En annan aspekt berör syftet med CLIL, nämligen att tillägna sig ett nytt språk. Om inte språket ska bedömas i samband med

ämnesundervisningen är frågan när och var eleven får ett kvitto på sin utvecklade språkförmåga, och vem som ger återkoppling för fortsatt språkutveckling (Morgan, 2006; Wewer, 2014).

Avseende språk i ämnesundervisningen är det viktigt att beakta att det finns dels ämnesspecifika begrepp som eleven behöver tillägna sig, dels ett ämnesöverskridande akademiskt skolspråk med ämnestypiska drag. I forsknings-sammanhang talas det om att ämnen har sin egen litteracitet vilket syftar på förmågan att använda läsande och skrivande för att skapa ett innehåll i skriftspråket så att det svarar mot behoven i en speciell kontext. För att elever ska behärska ämnesspråket räcker det inte med att de förstår och kan benämna olika begrepp. De behöver också kunna sätta in dessa i ett sammanhang för att i en provsituation kunna beskriva vad som händer i en process, särskilt om det handlar om öppna uppgiftsformat där eleven ska formulera sig med egna ord. I definitionen om vad som karakteriserar ett akademiskt språk, förutom specifika akademiskt frekventa ord (Olsson, 2016), brukar grammatiska drag som nominalfraser samt passiv form nämnas (Llinares et al., 2012; Schleppegrell, 2004; Skolverket, 2012). Clegg (2007) talar om att ämneslärare skulle behöva undervisa sina elever om dessa språkliga drag, liksom hur de är kopplade till olika ämnestypiska tankeaktiviteter, som exempelvis att *beskriva* eller *analysera*.

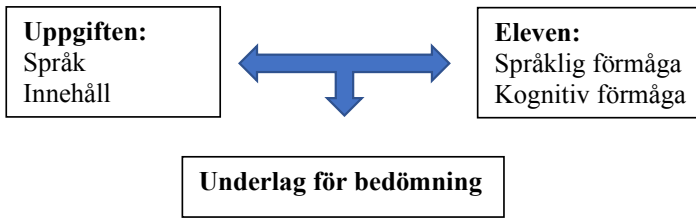
I en inlärnings-situation går eleverna från att först förstå innebörden i begrepp, till att börja använda dem, till att slutligen kunna applicera begrepp som en del av användandet av ett helt språkligt register, som exempelvis ett naturvetenskapligt språk. Under den här inlärningsprocessen kan ett så kallat *hybridspråk* uppstå som ett mellansteg, som innebär en blandning av vardagligt språk och det akademiska ämnesspråket. Vardagliga uttryck kan här ses som en resurs på vägen till det specifika målspråket (Olander & Ingerman, 2011). På samma sätt kan det vara relevant att utveckla elevers kompetens i olika språkliga register i språkklassrummet, att inte stanna vid vardagliga ämnen utan även använda ämnesövergripande akademiskt språk. Cummins (2000) skiljer på BICS, *Basic Interpersonal Communicative Skills*, och CALP, *Cognitive Academic Language Proficiency* där det första motsvarar vardagsspråk och det sista det akademiska

skolspråket. Internationella undersökningar visar att svenska elevers kompetens i engelska är mycket hög, särskilt inom de receptiva förmågorna. Samtidigt konstateras att svenska elevers akademiska skriftliga förmåga inte är lika väl utvecklad och att deras akademiska språk inte motsvarar samma höga nivå som det mer informella vardagsspråket (European Commission, 2012; Skolverket, 2012a). Språklärarnas undervisning av akademiskt skolspråk blir av ännu större betydelse då ämneslärare till viss del visar en omedvetenhet om behovet att undervisa akademiska språkliga register (Clegg, 2007; Yoxsimer Paulsrud, 2014).

Avseende innebörden av begreppet *innehåll* har det emellanåt visat sig vara svårare att definiera och avgränsa än språk. I relation till ett ämne finns i svenska kurs- och ämnesplaner det som beskrivs som centralt innehåll. Detta omfattar ofta ämnesmässigt stoff. I CLIL-sammanhang har innehåll definierats som *kunskap, förmågor* och *förståelse* (Coyle et al., 2010). I bedömningsanvisningar för uppsatser i nationella prov i engelska beskrivs innehåll som *sammanhang och struktur, begriplighet och tydlighet, fyllighet och variation*, liksom *anpassning till syfte och mottagare*. Innehåll skulle eventuellt kunna sammanfattas med ”hur man behandlar ett ämne med hjälp av språket”. I den aktuella studien (Reierstam, 2015) förhåller sig ämneslärarna till innehåll mestadels som ämnesmässigt stoff, men detta inkluderar även i viss mån ämnesspecifika begrepp, det vill säga språk. I och med CLIL-perspektivet blir det även naturligt för språklärarna att tänka på motsvarande sätt, att innehåll motsvarar ämnesstoff, vilket framgår i analysen av resultaten.

Frågan om språkets roll blir aktuell både vid själva bedömningstillfället, men även i den föregående utformningen av bedömningsuppgiften. Utgångspunkten vid konstruktion av en bedömningsuppgift är en definition av vad uppgiften avser att bedöma, även benämnt ett konstrukt (Messick, 1989). Det kan bestå i en förmåga eller en ämnesdomän som finns beskrivna i kunskapsmålen i kurs- och ämnesplaner i den aktuella kursen.

Läraren/provkonstruktören utgår från sina undervisningsmål och sitt kursinnehåll samtidigt som han/hon behöver väga in elevens förväntade språkliga och kognitiva förmåga, vad eleven förväntas kunna och göra.



**Figur 4.** Faktorer i uppgiften och hos eleven som bidrar till bedömningsunderlaget.

För att en uppgift ska ge ett rättvisande resultat, och för att eleven ska få visa sitt kunnande, behöver även möjliga svårigheter vägas in (Bachman & Palmer, 2012; Crooks et al., 1996; Messick, 1989). Elevers eventuella begränsade skrivförmåga ska exempelvis tas hänsyn till i val av uppgiftsformat, så att inte skrivförmåga hindrar en elev att visa sin ämnesförståelse. Figur 4 ovan illustrerar samspelet mellan uppgiften och eleven och de aspekter som därmed behöver vägas in vid uppgiftskonstruktionen för att ge ett så rättvisande underlag för bedömning som möjligt.

I provsituationen möts eleven av en uppgift som ställer krav på hans/hennes förmåga att svara både med korrekt innehåll och med ett språkligt uttryckssätt som motsvarar kraven som ställs i frågans formulering. Beroende på ordval och frågeformat följer olika höga krav på dessa förmågor.

Blooms (1956) reviderade taxonomi (Anderson & Krathwool, 2001) ger en hierarkisk beskrivning av de kognitiva förmågorna i sex steg, med de mindre kognitivt krävande i botten av pyramiden till successivt mer kognitivt krävande. Nivåerna definieras i tur och ordning med kognitiva processer som att *minnas*, *förstå*, *tillämpa*, *analysera*, *värdera* och till sist *skapa*. På motsvarande sätt innebär så kallade kognitiva funktionsord (se exempel nedan), av Dalton-Puffer (2016) benämnda som *cognitive discourse functions* förkortat CDFs, olika krav på kognitiv förmåga. På den lägre nivån återfinns ord som att *namnge*, *identifiera*, *återge* och *beskriva*. På en högre nivå finns ord som kräver mer av eleven såsom att *jämföra*, *förklara*, *analysera* och *generalisera* (jfr. Cummins, 2000). De kognitiva funktionsorden kan i sin tur sägas aktivera språkliga processer som visar på olika grad av språklig förmåga. Att

*identifiera* kan exempelvis innebära att ringa in rätt ord, vilket kräver förståelse men inte någon språklig produktion, att *förklara* medför däremot att producera eget språk. Frågeord kan även de kategoriseras utifrån vilka krav de medför för att leverera svaret. *Var, när, vem* och *vilken* innebär oftast att eleven ska minnas och återge fakta, medan *hur* och *varför* medför ett mer utvecklat resonerande svar. Frågeordet *vad* innehar en egen position då det kan innebära både och, antingen att namnge något, *vad heter* (exempelvis en person), eller att ge ett mer utförligt svar om *vad som hände* (Reierstam, 2015; Reierstam & Sylvén, 2019).

För att erbjuda ett verktyg för lärare vid utformningen av uppgifter, där kognitiv och språklig svårighetsgrad vägs samman med hur mycket stöd den omgivande kontexten erbjuder, vidareutvecklade Coyle (1999) en modell som är känd som Cummins kvadrant, eller matris (Cummins, 2000). Genom att erbjuda stöd från en rik kontext, menar Coyle att det är eftersträvansvärt att initialt utforma uppgifter som är kognitivt krävande utan att detta medför behov av att använda ett alltför krävande språk. Hon konstaterar att detta är en utmaning för lärare. Som ett exempel kan nämnas att detta enligt Coyle skulle motsvara uppgifter där eleven ska *generalisera, jämföra, sammanfatta, klassificera* och *recensera*. I sådana uppgifter erbjuder den omgivande kontexten ett visst språkligt stöd att luta sig mot där eleven kan hitta språkliga formuleringar att använda sig av. Uppgifter som på samma sätt är kognitivt krävande, men även medför högre språkliga krav, inkluderar att eleven exempelvis ska *argumentera, utveckla sina idéer, motivera sin åsikt, forma hypoteser, förutsäga resultat, analysera* och *föreslå lösningar*. I den senare kategorin finns paralleller till Blooms högsta nivå där eleven förväntas skapa egna slutsatser och argument. Därmed finns inte heller lika mycket stöd i en omgivande språklig kontext, utan eleven måste själv producera språk.

Resonemanget ovan lyfter fram olika aspekter som lärare i en flerspråkig kontext behöver tänka på vid utformandet av bedömningsuppgifter men även ha med i sin undervisning. Det kan sammanfattas med att beakta vilka språkliga krav olika frågeformuleringar och uppgiftsformat ställer på elever, samt att språk inte går att separera från innehåll.

## Forskningsfrågor

Innan vi går in på den aktuella studien presenteras här de tre övergripande forskningsfrågorna där den första jämför lärares bedömningspraktik i biologi och historia, den andra intresserar sig för undervisningsspråkets eventuella påverkan på engelsklärares bedömningspraktik på samma skolor, och den tredje frågan jämför bedömningsformer i alla tre ämnen. Den jämförande designen innebär att lärare som undervisar på engelska (CLIL), jämförs med kollegor som undervisar på svenska, här kallat icke-CLIL. Undervisningsspråket är det som skiljer de två åt medan ämnesmål och kunskapskrav är desamma. Likaså jämförs ämneslärare med engelsklärare då deras respektive kurser kan sägas integrera språk och ämne, fast med olika fokus och syfte. Den tredje och sista frågan jämför ämnesspecifika alternativt ämnesövergripande aspekter i uppgifternas utformning oberoende av CLIL.

- Finns det skillnader i bedömningspraktik beroende på om undervisningsspråket är engelska eller svenska, CLIL respektive icke-CLIL, i biologi och historia på gymnasiet? I så fall, *hur* skiljer den sig i åt och *på vilka grunder*?
- Påverkas innehåll och bedömningsformer i engelskkursen när andra ämnen undervisas på engelska? I så fall, *hur*?
- Hur ser bedömningsformerna ut i de olika ämnena med avseende på *språk, form* och *innehåll*, oavsett om det handlar om CLIL eller icke-CLIL? Finns det gemensamma drag?

Forskningsfrågorna härrör från ett intresse att belysa språkanvändning i alla ämnen, men likaså hur ämnesstoff kan integreras i språkkurser, och hur detta tar sig uttryck i bedömningsuppgifter. Kurs- och ämnesplaner sätter upp ramar som kursen ska hålla sig inom och stoff som ska hinnas med, vilket vissa av lärarna i studien hänvisar till. Utifrån elevperspektiv kan det vara viktigt att ställa sig frågan hur det är möjligt att lyfta språket i alla ämnen och att samtidigt identifiera meningsfulla autentiska uppgifter för språkanvändning i alla kurser som kan öppna för gemensamma språkutvecklande skriftliga bedömningsformer.

Nedan följer en presentation av studiens upplägg, de deltagande lärarnas bakgrund samt resultat av studien. Därefter



följer en diskussion om språkliga genrer och register som ingår i ämnena, samt behovet att beakta språkets roll för validiteten i bedömningsituationen.

## Studien

### Metod och informanter

Studien jämför tolv lärares syn på bedömning, CLIL och icke-CLIL, se förklaring av de två lärarkategorierna ovan. Tre av lärarna undervisar i biologi, tre i historia och sex i engelska vid tre gymnasieskolor i en stor samt två mellanstora svenska städer, se Tabell 9 nedan:

**Tabell 9.** Skolor och lärare som ingick i studien.

<i>Skola A Internationell skola med svensk läroplan</i>	<i>Skola B Två CLIL-program med en klass per år mestadels på engelska</i>	<i>Skola C Ett CLIL-program med två klasser per år delvis på engelska</i>
CLIL historia	CLIL biologi	CLIL historia
CLIL engelska	CLIL biologi	Icke-CLIL historia
CLIL engelska	Icke-CLIL biologi	CLIL engelska
	CLIL/Icke-CLIL engelska	Icke-CLIL engelska
	CLIL/Icke-CLIL engelska	

En av skolorna (A) är helt engelskspråkig, men med svensk läroplan, medan de andra två erbjuder två respektive ett gymnasieprogram (B och C) med internationell profil där undervisningsspråket i de flesta ämnen är engelska. Ingen av ämneslärarna som undervisar på engelska (CLIL) har engelska som sitt första språk, det har däremot två av engelsklärarna. Två av engelsklärarna undervisade både CLIL och icke-CLIL elever, medan de övriga endast undervisade på den ena eller den andra inriktningen. Eftersom studien ingår i det större CLISS-projektet, var samtliga informanter redan införstådda med den övergripande studiens syfte och omfattning.

Materialet som samlades in under 2013–14, består av halvstrukturerade intervjuer, en lärarenkät samt lärares skriftliga provexempel. Intervjuerna omfattar totalt 8 timmar, och varierar

från 16 till 58 minuter i längd. En intervjuguide användes med följande teman:

- Lärares erfarenheter av CLIL
- Lärares syn på det egna ämnet
- Lärares syn på bedömning och olika bedömningsverktyg
- Kursinnehåll och kursmaterial
- Förekomsten av kollegialt ämnesövergripande samarbete

De teman som användes framträdde efter läsning av tidigare studier om bedömning och CLIL, teori kring lärares egna uppfattningar (Borg, 2003) och så kallad *constructive alignment* (Biggs, 2003), det vill säga hur undervisning och bedömning kopplas samman med kursmål till en helhet.

Samtliga tolv lärare var med i den första inledande intervjun. Efter intervjun valde två av engelsklärarna att avsluta sitt deltagande och en av CLIL-biologilärarna gick i pension, vilket innebär att nio av de tolv lärarna svarade på enkäten. Samtliga lärare tillfrågades vid intervjun om att lämna in exempel på skriftliga prov från en av sina kurser, vilket ingen motsatte sig. Av de kvarvarande nio informanterna lämnade samtliga ämneslärare ( $n=5$ ) i bedömningsexempel, men endast en engelsklärare. Ytterligare en engelsklärare lämnade in ett par exempel på elevproducerade texter, men ingen beskrivning av uppgiften eller bedömningskriterierna.

Enkäten användes för att komplettera informationen från den första halvstrukturerade intervjun, detta efter att planen med en uppföljande intervju efter dokumentanalysen fick revideras. Enkäten, som var webbaserad, bestod av totalt 26 frågor uppdelade på tre teman enligt nedan (med några exempel under varje):

- Bakgrundsfrågor:
  - Lärares antal år i yrket och med undervisning i CLIL-kontext
- Bedömning och kursinnehåll:
  - Lärares vikt vid muntlig/skriftlig prestation, språkets roll vid bedömning, vilka faktorer som väger tyngst vid bedömning, vad läraren skulle vilja lära sig mer om avseende bedömning, kursplanernas roll

- Ämnesintegrering och kollegialt samarbete:
  - Om sambedömning förekommit med ämnes-/språklärrarkollega, eventuella hinder för sådant samarbete

Intervjuer och enkäter analyserades och jämfördes tematiskt med utgångspunkt i svenska kursmål med hjälp av intervjuguide och teman enligt ovan. I samband med standardiserade *high stakes* prov (dvs. där mycket står på spel) exempelvis nationella prov, finns ofta en så kallad *provspecifikation* som beskriver ett provs syfte, innehåll, uppgiftsformat och bedömningsgrunder. De skriftliga proven analyserades på motsvarande vis utifrån frågetyp, exempelvis flervalsfrågor, matchningsfrågor, kortsvarsfrågor eller uppsatsfrågor samt vilka krav proven ställer på elevers förväntade kognitiva och språkliga förmåga, bland annat beroende på vilket frågeord eller kognitivt funktionsord som användes (Bachman & Palmer, 2012; Wikström, 2013). För att bedöma det sistnämnda användes även Blooms reviderade taxonomi (Anderson & Krathwool, 2001) och en tillämpning av Cummins kvadrant som kombinerar kognitiv komplexitet med språklig komplexitet (Coyle, 1999), vilket beskrivits ovan. Beskrivningar som återfinns både i CLIL-litteratur och studier av bedömning i CLIL (Hönig, 2009; Wewer, 2014), samt i tidigare studier av bedömningsmaterial i andra ämnen (Lindmark, 2013; Odenstad, 2010; Rosenlund, 2011) utgjorde även viktiga referenser.

Eftersom det inte finns några tidigare gjorda studier av det här slaget inom CLIL, var det till att börja med svårt att hitta beskrivningsvariabler eller en lämplig metod för att kunna beskriva och jämföra provuppgifter från skilda ämnesdiscipliner på ett enhetligt sätt. I arbetet med Cummins kvadrant och Blooms taxonomi, liksom granskningen av ovan nämnda litteratur och tidigare gjorda studier, framkom flera gemensamma jämförbara kriterier och en egen modell utformades. Ett sätt att jämföra hur användandet av de kognitiva funktionsorden och frågeorden såg ut i jämförelsen mellan CLIL och icke-CLIL bestod i en kvantitativ analys där förekomsten av olika ord räknades i de prov som förekom i kurserna för att se om exempelvis ord med högre kognitiva och språkliga krav var mindre förekommande i CLIL jämfört med icke-CLIL.

Validitet är en viktig aspekt vid bedömning. En vanlig definition av validitet är om ett prov mäter vad det är avsett att mäta, vilket benämns som *inhållsvaliditet* eller *begreppsvaliditet*. I modern testteori används det sistnämnda, men då kallat *konstruktvaliditet*, som ett sammanhållet validitetsbegrepp som inbegriper de konsekvenser tolkningen av ett provs resultat kan få (Crohnbach & Meehl, 1955; Messick, 1989). Tidigare nämndes att utgångspunkten vid konstruktion av en bedömningsuppgift är en definition av vad som ska bedömas eller mätas. Messick har definierat två hot mot validiteten, nämligen *construct-underrepresentation* och *construct-irrelevant variance*. Det första innebär att för lite av det som ska bedömas finns med i provets utformning, medan den andra i princip betyder det motsatta, nämligen att störande faktorer som inte borde finnas med kan inverka negativt på resultatet. I den här studien användes en valideringsmodell för att identifiera sådana möjliga hot mot validiteten i bedömning i CLIL-kontext. Crooks, Kane och Cohens kedjemodell från 1996 ansågs tydlig och användbar för ändamålet. Den är utformad i åtta steg och styrkan i varje steg eller länk beror på om själva syftet med bedömningen är tydlig och lämpligheten i bedömningsuppgifterna står i relation till dessa syften. Den andra länken består exempelvis i poängsättning och ett hot mot denna länk kan utgöras av om för mycket vikt läggs vid någon parameter. I CLIL-kontext skulle ett sådant hot kunna utgöras av att för stor vikt fästs vid bristande stavning, då fokus i uppgiften är historisk analysförmåga. Detta innebär således att resultatet kan påverkas negativt av konstruktirrelevant varians.

## Resultat

Svaret på den första forskningsfrågan, om det finns skillnader i lärares bedömningspraktik beroende på om undervisningsspråket är svenska eller engelska, är i huvudsak nej, även om vissa skillnader återfinns. CLIL och användandet av engelska som undervisningsspråk har inte någon nämnvärd påverkan på lärares förhållningssätt till språket vid bedömning eller eventuella användning av andra bedömningsformer, enligt deras egen utsago i intervjuerna. De skillnader som kan spåras i användande

av olika bedömningsformer, framför allt i historia, verkar snarare bero på individuella preferenser eller lärarnas ämnessyn. De två CLIL-lärarna i historia föredrar uppsatser, men av helt olika karaktär. Den ena förordar en narrativ genre, med motiveringen att detta bäst avspeglar den historiska ämneskaraktären, medan den andra föredrar en mer akademisk form. Den sistnämnda läraren hänvisar till de texttyper som eleverna förväntas använda vid rapportskrivning eller uppsatser på universitet senare, en skriftlig genre som enligt henne är mer tydlig och strukturerad än narrativa uppsatser. Icke-CLIL-kollegan i historia föredrar frågeprov, men vissa av frågorna, framför allt mot slutet av kursen, har uppsatskaraktär. I biologi används både frågeprov och andra skrivuppgifter såsom laborationsrapporter och fältanteckningar. Två kollegor, en CLIL och en icke-CLIL, samarbetar kring utformningen, och flera provfrågor och uppgifter är närmast identiska, men översatta till engelska i det ena fallet. Läraren som undervisade på svenska använde ett något större antal uppsatsfrågor som kräver längre skriftliga svar.

Som svar på följdfrågan *hur* bedömningspraktiken skiljer sig åt och *på vilka grunder*, kan man konstatera att både intervjuer och enkätsvar visar att både CLIL-lärare och icke-CLIL-lärare föredrar skriftliga bedömningsformer framför muntliga. CLIL-lärarna hävdar likaså att språket inte ska spela någon roll, ändå visar det sig att det gör det. I samband med användandet av uppsatser som bedömningsform och längre provfrågor anger CLIL-lärarna att de ser en del problem med språkets negativa inverkan vid bedömning ibland, vilket de löser genom uppföljande förtydligande samtal med berörda elever. En av historielärarna beskriver att det ibland är svårt att veta om ett otydligt svar beror på att eleven inte förstått frågan, inte kan svaret, eller om språkliga begränsningar i användandet av engelska utgör ett hinder att formulera sig. CLIL-lärarna har inte någon bedömningsmatris eller rättningsmall att presentera som förtydligar vad som ska bedömas i de aktuella uppgifterna. En av dem värjer sig snarare mot detta och menar att bedömningskriterierna blir tydliga under processens gång i dialog med eleverna. Det som anses viktigt med avseende på validiteten vid bedömning, är att syftet med uppgiften går att redogöra för samt att bedömningskriterierna är transparenta. Alderson och

Banerjee (2002) hävdar exempelvis att ju mer strukturerad en uppgift är, desto mer reliabel/rättvisande är bedömningen.

Språket ska inte bedömas enligt ämneslärarna. En skillnad som framkommer i intervjuerna med biologilärarna, är att CLIL-lärarna medvetet låter eleverna arbeta med ämnesspecifika vokabulär, vilket delvis kan härledas till användandet av engelska. Även icke-CLIL läraren förser dock eleverna med ordlistor över ämnesspecifika begrepp. En av CLIL-lärarna i biologi medger att det generella akademiska ämnesspråket inte ägnats någon särskild fokus, men påpekar dock i intervjun att det skulle vara en intressant aspekt att ha i åtanke. Icke-CLIL kollegan däremot nämner spontant behovet att tillgodose utvecklandet av ett mer generellt akademiskt ämnesspråk.

CLIL-eleverna som haft engelska som undervisningsspråk har i vissa fall möjlighet att välja vilket språk de vill använda, ha tillgång till ordböcker eller få längre skrivtid vid prov. Dessutom finns möjligheten att göra förtydliganden i efterhand om språket tycks utgöra ett hinder. Gablasova (2014) påpekar i sin studie som nämndes ovan, att möjligheten att välja språk kan vara problematiskt, eftersom undervisningen inneburit att ämnesspråket är givet på ett språk och i provsammanhang kan viktiga begrepp saknas på modersmålet. Detta kan innebära ett hot mot validiteten i ett bedömningsunderlag. Alla CLIL-lärare i den aktuella studien ger uttryck för att det enda som bör skilja sig åt mellan CLIL och icke-CLIL är språket, inte innehållet eller vad som bedöms. De hänvisar till att de lyder under samma nationella ämnesplaner och kursmål. I enkäten anger alla lärare utom en, att nationella kursmål är mycket viktiga vid betygssättning. I ämnesplanerna i de aktuella ämnena, framgår det tydligt att språkliga förmågor lyfts fram: förmåga att använda kunskap i biologi för att kommunicera, förmåga att använda olika historiska teorier och koncept för att undersöka, förklara och dra slutsatser. Ämneslärarna betraktar inte detta som språk utan en del av innehållet.

Vissa modifieringar i bedömningspraktiken på grund av CLIL kan eventuellt spåras. Frekvensen av funktionsord på en lägre kognitiv nivå kan förefalla något högre i CLIL-prov. I biologi, oberoende av om det är CLIL eller ej, är *förklara* det mest frekventa. Ofta används det tillsammans med *beskriv*. I historia är

frågeordet *vad* det överlägset mest använda i CLIL-kontext, men inte lika mycket i icke-CLIL. Bruket av ordet varierar dock med frågor av olika karaktär. I samband med vissa frågor får eleverna hjälp i form av en föreslagen inledande bokstav (se exempel 1 nedan). I andra frågor förväntas eleven utveckla svaret något mer (exempel 2 nedan). I ytterligare andra exempel kan det handla om att beskriva ett helt händelseförlopp där en poängangivelse i anslutning till frågan anger en möjlig begränsning. Exempelen nedan kommer från två olika CLIL-lärare (se även Reierstam, 2015 och Reierstam & Sylvén, 2019).

- (1) *What was the popular name for the young Macedonian king who was in charge of the Macedonian troops and personalized this development? He also had a very famous teacher, almost as famous as himself. What was his name? (4)*

*Answer:*

*The king:*                    A \_\_\_\_\_

*His famous teacher:* Ar \_\_\_\_\_

- (2) *More explanatory answers needed. Points in brackets.*

*What does the concept appeasement mean? Also give an example of when one can say that the UK and France used that way of acting. (4)*

I den första frågan förekommer raka poäng, det vill säga poäng för rätt ifyllt ord, dock två poäng per svar trots att det rör sig om ett ord i vart och ett av svaren. I den andra frågan ska frågans kvalitet viktas. Svaret kräver en förklaring och det skulle kunna bli ganska långt.

Den andra forskningsfrågan i den aktuella studien handlar om och i så fall hur, innehållet och bedömningsformerna i engelskkursern påverkas i de fall där eleverna läser andra kurser på engelska. Svaret är, i likhet med den föregående frågan: inte nämnvärt. I CLIL-litteratur beskrivs CLIL som en undervisningsmetod vars framgång är beroende av graden av samarbete lärare emellan (Coyle, 1999). Interdisciplinärt samarbete tycks dock mindre

frekvent bland de deltagande lärarna i denna studie än i andra icke-CLIL kontexter. Endast en engelsklärare (icke-CLIL), som lämnade studien efter intervjuerna, angav att hon samarbetade med en historiekollega. Frågan som infinner sig är vad bristen på integrering av innehåll från ämneskurserna beror på? Enligt lärarna själva beror det på brist på tid att hinna med något annat än det kursinnehåll de är tvungna att klara av i relation till sina ämnesmål. Engelsklärarna fäster stor vikt vid de nationella proven. Dessa används som referens för vad som behöver ingå i kursen, och gamla prov används som bedömningsinstrument och som förberedelse för de ”riktiga” proven. En annan faktor som hämmar ämnesintegrering beskrivs som brist på tid för samarbete och samplanering. På den internationella skolan förekom det vid något enstaka tillfälle att elever eller ämneskollegor frågade engelsklärarna om hjälp med engelskspråkig kurslitteratur. Det ämnesövergripande arbetet var något skolledningen hade insett att man behövde stärka genom att exempelvis ge lärarna i uppdrag att studera varandras ämnesplaner. Slutsatsen är att CLIL verkar stanna inom ämneskurserna, och inte inbegripa språkkurserna, vilket bekräftas av andra studier (Clegg, 2007; Massler et al., 2014). Det som bör påpekas är dock att i avsaknaden av nationella riktlinjer för CLIL, saknas även beskrivningar för hur integrering och samarbete ska gå till.

Syftet med den tredje och sista forskningsfrågan är att jämföra bedömningsformerna i de tre ämnena avseende *språk, form* och *innehåll* och huruvida det finns gemensamma drag. En tydlig gemensam nämnare handlar om fokus på skriftlig genre. Även om formen delvis skiftar beroende på ämne finns ett behov av att eleverna lär sig bemästra akademiskt skrivande i form av resonering och argumenterande uppsatser i alla ämnen. I biologi förutsätts eleverna ge prov på analytisk och utvärderande förmåga i längre skrivuppgifter så som laborationsrapporter medan historieämnet innebär att elever behöver beskriva, jämföra och resonera kring händelseförlopp. I både historia och biologi ställer dessa skriftliga genrer krav på specifik språklig förmåga. Omvänt, i engelskkurserna, där ett av syftena är att utveckla både skriftlig narrativ och argumenterande uttrycksförmåga, används innehållsstoff som mycket väl kan vara av exempelvis



historisk art. Här finns med andra ord ett möjligt område för ämnesövergripande samarbete. En skillnad som dock framkommer berör dels lärarnas uppfattning om elevernas språkliga kompetens, dels hur lärarna ser på språkets plats i undervisningen. Det sistnämnda kan beskrivas med att språket inte utgör det primära fokuset i ämnena vilket lärarna ger uttryck för, men medan engelsklärarna ser CLIL-eleverna som avancerade språkinlärare, anser ämneslärarna att de måste stötta eleverna språkligt. Detta kan sägas återspegla att två olika språkliga register är i fokus i de två kontexterna, vilket i sin tur kan jämföras med Cummins (2000) BICS, *Basic Interpersonal Communicative Skills* och CALP, *Cognitive Academic Language Proficiency*, där det första är mer frekvent i engelskklassrummet, och det sistnämnda i ämneskurserna. Den första engelskkursen på gymnasiet, engelska 5, liksom årskurs 9 i grundskolan, motsvarar B1-nivå enligt Gemensam Europeisk Referensram för Språk (GERS, Skolverket, 2009). I årskurs två på gymnasiet läser eleverna kursen engelska 6 motsvarande B2-nivå. I engelska 5 fokuserar innehållet fortfarande den personliga domänen, medan engelska 6 rör sig vidare mot mer samhällrelaterade, yrkesmässiga och vetenskapliga domäner, vilket även påverkar språkets register. Elever som börjar på gymnasiet kan ha mycket gott språkligt flyt, men deras ordförråd utanför den personliga domänen är då mer beroende av vilka kontakter de har haft med engelska utanför skolan, så kallad extramural engelska. Sylvén och Sundqvist (2012) skriver om hur pojkar genom sitt datorspelade uppvisar ett bredare ordförråd än flickor, exempelvis.

På frågan om huruvida ämneslärarna tar hänsyn till språket i sin bedömning, svarade de lite olika. I intervjun svarade tre av fyra CLIL-ämneslärare att språket absolut inte spelar roll vid bedömning i deras ämnen, utan att det är upp till lärarna i engelska. Däremot problematiserade de språkets inverkan på elevernas svar, vilket beskrivits ovan. På frågan vilka faktorer lärarna anser väger tyngst vid bedömning, svarade samtliga ämneslärare i enkäten att de tyckte att användandet av ämnesspecifika begrepp är bland det viktigaste i bedömning av elevernas kunskap, samtidigt angav ingen av dem att användandet av ett generellt akademiskt språk skulle vara bland det viktigaste. Det tyckte däremot två

av engelsklärarna. De angav å andra sidan i intervjuerna att de inte var intresserade att integrera annat ämnesstoff i sina kurser då de inte ansåg sig kapabla att bedöma ett korrekt användande av ämnesbegrepp eller känner till ämnet väl nog för att göra en bedömning av språket.

I enkätsvaren framgick att fem av nio svarande gymnasielärare anser att bedömning och betygssättning i deras ämnen är svårt. Av dessa fem var fyra lärare i historia och biologi. Två av dessa undervisar sitt ämne på engelska, vilket kan bidra till deras uppfattning. De fyra lärare som ansåg att bedömning är förhållandevis lätt var alla engelsklärare, samma lärare var emellertid tveksamma till att visa sina prov och underlag för bedömning i kursen. I stället hänvisade de till nationella prov och elevproducerade texter, men angav i flera fall att de använder en bred repertoar av skriftliga vokabulär- och grammatikprov samt muntliga redovisningar och skrivuppgifter. I anslutning till den första forskningsfrågan beskrevs att samtliga lärare föredrar skriftliga bedömningsformer framför muntliga, det gällde inte bara ämneslärarna utan även engelsklärarna. Lärarna menade att skrivuppgifter ger eleverna mer tid att tänka, större möjlighet att använda språket och göra sig hörda samt att skriftliga prov är lättare och mer reliabla för lärare att bedöma.

En gemensam nämnare relaterar således till det ämnesöverskridande akademiska språket och användandet av skriftliga genrer. Det ämnesspecifika innehållet skulle kunna användas för att berika det språkliga uttrycket i engelskkurserna. Engelsklärarna skulle kunna integrera frågor från andra ämnen för diskussion på engelska, eller använda viss lektionstid till att hjälpa eleverna med skriftliga ämnesrelaterade genrer och ämnestypiskt språk. Begreppet innehåll, vilket tidigare konstaterats, är tydligt kopplat till ämnets karaktär. I biologi anses ämnesbegreppen vara en del av innehållet (Llinares et al., 2012), men detta är inte lika tydligt i historia. I den engelska ämnesplanen framstår begreppet innehåll som mest mångfacetterat, vilket kan sägas avspeglas i den breda och något spretiga bedömningspraktiken engelsklärare emellan. Ämneskollegorna i biologi och historia visade på mindre mångfald i bedömningsformer. Sammanfattningsvis kan sägas att den skriftliga produktionen framstår som den gemensamma

ämnesöverskridande faktorn där språk och innehåll möts, oberoende av hur dessa definieras.

## Didaktiska implikationer

Studien bekräftar det som redan konstaterats i tidigare forskning, nämligen en avsaknad av tydligt definierade riktlinjer för CLIL som undervisningsmetod (Reierstam & Sylvén, 2019; Sylvén, 2013; Yoxsimer Paulsrud, 2014), vilket för med sig frågor rörande validiteten i bedömning av ämnesinnehåll i CLIL-kontext. Ett syfte med CLIL kan tyckas vara ett parallellt utvecklande av både ämnesrelaterad och generell akademisk språklig kompetens. För validiteten och värdet i CLIL skulle ämneskurser och engelskkurser behöva samarbeta för att bidra till utvecklandet av CALP och ett generellt akademiskt språk. Om syftet med CLIL är att samtidigt kunna testa faktakunskap, analytisk förmåga och språklig repertoar kan användandet och undervisningen av skriftliga generer (Hyland, 2003, 2007) och portfolio för bedömning ses som en framkomlig väg. Vad avser frågeprov som är vanliga i ämneskurser, skulle en diskussion om hur olika frågekonstruktioner står i relation till förväntad språklig prestation, i synnerhet på ett annat språk än modersmålet, kunna vara till hjälp. Enligt Hyland (2007) kan genrepedagogik ge lärare möjlighet att i sin undervisning utgå från de texter eleverna förväntas skriva i den aktuella kontexten. Bachman och Palmer (2012) beskriver på samma sätt hur målspråket kan vara utgångspunkt i framtagandet av lämpligt innehåll för bedömningsuppgifter. De karakteristiska dragen i den specifika akademiska genren lyfts därmed fram och undervisas i den aktuella kursen. På så vis blir alla lärare på ett sätt språklärare.

I punkterna nedan sammanfattas några möjliga praktiska riktlinjer för bedömning i språk- och ämnesintegrerad undervisningskontext. Punkterna grundar sig på motsägelser i ämneslärares uttalanden i den aktuella studien om att de inte är språklärare men samtidigt menar att bedömningen påverkas av elevers språkliga begränsning. Om inte språket ska utgöra ett validitetshot mot rättvisa och rättvisande tolkningar av elevers ämneskunnande behöver ämneslärare bli mer av språklärare, såvida inte samarbetet med engelsklärarna kan bli mer utvecklat och kursinnehållet mellan ämnen mer integrerat. Slutsatserna i den aktuella studien

sammanfaller även med tidigare forskningsstudier om att språk, förutom begrepp, faller mellan stolarna.

Språket och skriftliga genrer:

- Synliggöra och undervisa såväl ämnesspecifika begrepp som ämnesspecifika uttryck
- Identifiera och undervisa relevanta ämnesrelaterade skriftliga genrer. Exemplifiera med måltexter och målspråk
- Lyfta fram och undervisa generellt akademiskt språk, t ex användandet av nominalfraser och passiv form

Bedömning:

- Utarbeta och använda transparenta/tydliga bedömningsmallar för olika skriftliga genrer (till exempel laborationsrapport, akademisk uppsats) som inkluderar både innehåll och språk, gärna i ämnesövergripande samarbete med kollegor
- Koppla funktionsord/frågeord i olika frågetyper till exempel på elevens förväntade språkliga produktion och förmåga, så kallade *benchmarks*

Slutsatsen blir att för att bedömning av ämnesinnehåll i språk- och ämnesintegrerad undervisning ska vara så valid och rättvisande som möjligt finns ett behov att undervisa det språk som uppgifterna kräver. Det inkluderar hela ämnesspråket med dess genrer och inte enbart begrepp, för att elever ska bli litterata och kunna utföra språkligt och kognitivt krävande uppgifter.

## **Vad kan man som forskande lärare bidra med till det språkdidaktiska fältet?**

Som forskande lärare finns möjligheter att knyta praxisnära forskning, vilket behövs, till sin egen undervisning och till sina elever. De frågor som ligger och pyr både hos mig själv som lärare och hos mina elever är oftast själva katalysatorn, som i mitt fall. Inom forskningsfältet tvåspråkig undervisning och bedömning, oavsett

om undervisningen kan kategoriseras som CLIL eller undervisning av nyanlända/flerspråkiga, behövs många goda exempel på hur bedömning kan gå till. Som noterat inledningsvis är det fortfarande ett relativt outforskat område. Kartläggningsstudier behövs för att veta mer om språklärares bedömningspraktik och dessa kan ta sin början med hjälp av kollegorna på den skola där läraren är. En enkät, eller en insamling av bedömningsexempel kan utgöra grund för en studie hur språkliga krav ser ut i alla ämnen. Vad gäller språkens status behövs det många goda forskningsexempel rörande hur vi kan stärka andra språk än engelskan i svensk skola. CLIL behövs kanske snarare i de andra språken än i engelska - här behövs forskning.

Efter avslutade forskarstudier har jag som lärare fått en mycket större förståelse för komplexiteten i forskningsprocessen: de etiska aspekterna, dilemman relaterade till att avgöra vilka faktorer som påverkar vad och insikten att som man frågar får man svar. Det finns alltid förbehåll när det gäller forskningsresultaten. Mina resultat måste jämföras med andras för att skapa en större helhet, men om jag bidrar med en del, och någon annan en till del, så blir det till slut en hel del.

## Referenser

- Alderson, J. C. & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35, 79–113. <https://doi.org/10.1017/S0261444802001751>
- Anderson, L. & Krathwool, D. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Bachman, L. F. & Palmer, A. S. (2012). *Language assessment in practice*. Oxford University Press.
- Biggs, J. (2003). *Teaching for quality learning at university*. Open University Press. <https://doi.org/10.1007/s10734-004-7092-z>.
- Bloom, B. S. (1956). *Taxonomy of educational objectives, the classification of educational Goals. Handbook I: Cognitive Domain*. David McKay Company.

- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe and do. *Language Teaching*, 36, 81–109. <https://doi.org/10.1017/S0261444803001903>
- Clegg, J. (2007). Analysing the language demands of lessons taught in a second language. I F. Lorenzo, S. Casal Madinabeitia, V. de Alba Quiñones, & P. Moore (Red.), *Models and Practice in CLIL*, Monografnummer av *Revista Española de Lingüística Aplicada* (s. 113–128). <https://dialnet.unirioja.es/servlet/articulo?codigo=2575499>
- Coyle, D. (1999). Supporting students in content and language integrated learning contexts: planning for effective classrooms. I Masih, J. (Red.), *Learning through a foreign language. Models, methods and outcomes* (s. 46–62). CiLT.
- Coyle, D., Hood, P. & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press.
- Crohnbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302. <https://doi.org/10.1037/h0040957>
- Crooks, T. J., Kane, M. T. & Cohen, A. S. (1996). Threats to the valid use of assessments. *Assessment in Education: Principles, Policy & Practice*, 3, 265–285. <https://doi.org/10.1080/0969594960030302>
- Cummins, J. (2000). *Language, power and pedagogy. Bilingual children in the crossfire*. Multilingual Matters.
- Cummins, J. (2007). *What works? Research into practice. Research monograph No 5*. The literacy and numeracy secretariat and the Ontario association of Deans of education. <http://www.edu.gov.on.ca/eng/literacynumeracy/inspire/research/cummins.pdf>
- Dalton-Puffer, C. (2007). *Discourse in content and language integrated learning (CLIL) classrooms*. John Benjamins.
- Dalton-Puffer, C. (2016). Cognitive discourse functions: Specifying an integrative interdisciplinary construct. I T. Nikula, E. Dafouz, P. Moore & U. Smit (Red.), *Conceptualizing integration in CLIL and multilingual education* (s. 29–54). Multilingual Matters.
- Erickson, G. (2010). Good practice in language testing and assessment – a matter of responsibility and respect. I T. Kao & Y. Lin

- (Red.), *A new look at teaching and testing: English as subject and vehicle* (s. 237–258). Bookman Books.
- European Commission (2012). *First European survey on language competences: final report*. Nedladdat från: [http://nellip.pixel-online.org/files/publications\\_PLL/European%20Survey%20on%20Language%20Competences.pdf](http://nellip.pixel-online.org/files/publications_PLL/European%20Survey%20on%20Language%20Competences.pdf)
- Gablasova, D. (2014). Issues in the assessment of bilingually educated students: expressing subject knowledge through L1 and L2. *The Language Learning Journal*, 42, 151–164. <https://doi.org/10.1080/09571736.2014.891396>
- Gibbons, P. (2013). *Lyft språket, lyft tänkandet – Språk och lärande*. Hallgren & Fallgren.
- Hajer, M. & Meestringa, T. (2014). *Språkinriktad undervisning. En handbok*. Studentlitteratur.
- Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of Second Language Writing*, 12, 17–29. [https://doi.org/10.1016/S1060-3743\(02\)00124-8](https://doi.org/10.1016/S1060-3743(02)00124-8)
- Hyland, K. (2007). Genre pedagogy: Language literacy and L2 writing instruction. *Journal of Second Language Writing*, 16, 148–164. <https://doi.org/10.1016/j.jslw.2007.07.005>
- Hönig, I. (2009). *Assessment in CLIL*. Diplomarbeit, University of Vienna. Philologisch-Kulturwissenschaftliche Fakultät. [http://othes.univie.ac.at/7976/1/2009-11-21\\_0107764.pdf](http://othes.univie.ac.at/7976/1/2009-11-21_0107764.pdf)
- Kiely, R. (2009). CLIL – The question of assessment. Developing teachers.com [https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/kiely\\_r\\_clil\\_assessment.pdf](https://www.unifg.it/sites/default/files/allegatiparagrafo/20-01-2014/kiely_r_clil_assessment.pdf)
- Lindmark, T. (2013). *Samhällskunskapslärares ämneskonceptioner* [Doktorsavhandling, Umeå universitet]. <http://umu.diva-portal.org/smash/get/diva2:663989/FULLTEXT01.pdf>
- Llinares, A., Morton, T. & Whittaker, R. (2012). *The roles of language in CLIL*. Cambridge University Press.
- Massler, U., Stotz, D. & Quessier, C. (2014). Assessment instruments for primary CLIL: the conceptualisation and evaluation of test tasks. *The Language Learning Journal*, 42, 137–150. <https://doi.org/10.1080/09571736.2014.891371>

- Mehisto, P., Marsh, D. & Frigols, J. M. (2008). *Uncovering CLIL: Content and language integrated learning in bilingual and multi-lingual education*. Macmillan Education.
- Messick, S. (1989). Validity. I R. L. Linn (Red.), *Educational measurement* (s. 13–103). American Council on Education, MacMillan.
- Mohan, B. (1986). *Language and Content*. Addison-Wesley Publishing Company.
- Morgan, C. (2006). Appropriate language assessment in content and language integrated learning. *Language Learning Journal*, 33, 59–67. <https://doi.org/10.1080/09571730685200121>
- Nixon, J. (2000). *SPRINT: Språk- och innehållsintegrerad inläring och undervisning. Rapport på uppdrag av Skolverket*. Skolverket. <https://www.skolverket.se/download/18.6bfaca41169863e6a6542do/1553957424040/Quality%20in%20Sprint.pdf>
- Odenstad, C. (2010). *Prov och bedömning i samhällskunskap. En analys av gymnasielärares skriftliga prov* [Licentiatuppsats, Karlstads universitet]. <https://www.diva-portal.org/smash/get/diva2:379242/FULLTEXT01.pdf>
- Olander, C. & Ingerman, Å. (2011). Towards an inter-language of talking science: exploring students' argumentation in relation to authentic language. *Journal of Biological Education*, 45, 158–164. <https://doi.org/10.1080/00219266.2011.591414>
- Olsson, E. (2016). *On the impact of extramural English and CLIL on productive vocabulary* [Doktorsavhandling, Göteborgs universitet]. [https://gupea.ub.gu.se/bitstream/2077/41359/1/gupea\\_2077\\_41359\\_1.pdf](https://gupea.ub.gu.se/bitstream/2077/41359/1/gupea_2077_41359_1.pdf)
- Quartapelle, F. (2012). *Assessment and evaluation in CLIL*. Ibis. <http://aeclil.altervista.org/Sito/wp-content/uploads/2013/02/AECLIL-Assessment-and-evaluation-in-CLIL.pdf>
- Reierstam, H. (2015). *Assessing language or content? A comparative study of the assessment practices in three Swedish upper secondary CLIL schools* [Licentiatuppsats, Göteborgs universitet]. [https://gupea.ub.gu.se/bitstream/2077/40701/4/gupea\\_2077\\_40701\\_4.pdf](https://gupea.ub.gu.se/bitstream/2077/40701/4/gupea_2077_40701_4.pdf)
- Reierstam, H. & Sylvén, L. K. (2019). Assessment in CLIL. I L.K. Sylvén (Red.), *Investigating content and language integrated*



- learning. Insights from Swedish high schools* (s. 59–75). Multilingual Matters.
- Rosenlund, D. (2011). *Att hantera historia med ett öga stängt: samstämmighet mellan historia A och lärares prov och uppgifter* [Licentiatuppsats, Forskarskolan i historia och historiedidaktik, Lunds universitet]. <https://portal.research.lu.se/ws/files/5890611/1971787>
- Schleppegrell, M. J. (2004). *The language of schooling. A functional linguistics perspective*. Lawrence Erlbaum Associates.
- Skolverket (2011). *Läroplan för gymnasieskolan*. Skolverket. <https://www.skolverket.se/undervisning/gymnasieskolan/laroplan-program-och-amnen-i-gymnasieskolan/laroplan-gy11-for-gymnasieskolan>
- Skolverket (2012a). *Bedömning av språklig kompetens. En studie av samstämmigheten mellan Internationella språkstudien 2011 och svenska styrdokument*. Skolverket. <https://www.skolverket.se/download/18.6bfaca41169863e6a659e88/1553964475581/pdf2831.pdf>
- Skolverket (2012b). *Greppa språket. Ämnesdidaktiska perspektiv på flerspråkighet*. Skolverket.
- Sylvén, L.K. (2013). CLIL in Sweden – why does it not work? A metaperspective on CLIL across contexts in Europe. *International Journal of Bilingual Education and Bilingualism*, 16, 301–320. <https://doi.org/10.1080/13670050.2013.777387>
- Sylvén, L. K. (2019). *Investigating content and language integrated learning. Insights from Swedish high schools*. Multilingual Matters.
- Sylvén, L.K. & Ohlander, S. (2014). The CLISS project: Receptive vocabulary in CLIL versus non-CLIL groups. *Moderna Språk*, 108, 80–114. <http://ojs.ub.gu.se/ojs/index.php/modernasprak/article/view/2789/2614>
- Sylvén, L.K. & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL*, 24, 302–321. <https://doi.org/10.1017/S095834401200016X>
- Wewer, T. (2014). *Assessment of young learners' English proficiency in bilingual content instruction CLIL* [Doktorsavhandling, Åbo

universitet]. <https://www.utupub.fi/bitstream/handle/10024/96838/AnnalesB385Wewer.pdf?sequence=2&isAllowed=y>

Wikström, C. (2013). *Konsten att göra bra prov*. Natur och Kultur.

Yoxsimer Paulsrud, B. (2014). *English-Medium instruction in Sweden. Perspectives and practices in two upper secondary schools* [Doktorsavhandling, Stockholms universitet]. <http://du.diva-portal.org/smash/get/diva2:727238/FULLTEXT01.pdf>