

Zeitschrift für Interkulturellen Fremdsprachenunterricht

Didaktik und Methodik im Bereich Deutsch als Fremdsprache

ISSN 1205-6545 Jahrgang 20, Nummer 2 (Oktober 2015)

Leistungsmessung

Lena Heine

Seminar für Sprachlehrforschung
Ruhr-Universität Bochum
Universitätsstraße 150
44780 Bochum
E-Mail: lena.heine@rub.de

Abstract: In dem vorliegenden Beitrag werden Instrumente zur Leistungsmessung im CLIL-Unterricht sowie die Frage nach der Gewichtung und Trennung sprachlicher und fachlicher Kompetenzen bei der Leistungsfeststellung und -beurteilung thematisiert.

The present contribution discusses possibilities and challenges of assessment in CLIL education associated with simultaneous language and content learning, in particular backwash effects and the question of separating language and subject competences.

Schlagwörter: CLIL, Messung von Sprachkompetenz, Messung interkultureller Kompetenz, Testkonstrukt; CLIL, assessing language proficiency, assessing intercultural competence, construct validation

1. Einleitung

Da der CLIL-Unterricht mit der Ausbildung von fremdsprachlichen, sachfachlichen und interkulturellen Kompetenzen mehrere Zielsetzungen gleichzeitig verfolgt (vgl. Heine 2015 in der vorliegenden Ausgabe), ergeben sich Fragen danach, mittels welcher Instrumente die jeweiligen Bereiche in Schülerleistungen erhoben werden können: In welchem Verhältnis sollen die Kompetenzdimension untereinander gewichtet werden? In welcher Sprache soll die Leistungen erhoben werden? Dabei ist zu diskutieren, inwiefern (fremd-)sprachliche und sachfachliche Kompetenzen getrennt voneinander messbar sind und welche Konsequenzen es hat, wenn die fremdsprachliche Leistung (vermeintlich zu Gunsten der Lernenden) nicht mitbewertet wird.

2. Instrumente zur Leistungsmessung im CLIL-Unterricht

Wie jede Form von Unterricht steht CLIL-Unterricht in Bezug auf eine adäquate Leistungsmessung vor der Herausforderung, die entwickelten Kompetenzen angemessen und valide sichtbar zu machen. In Bezug auf die sachfachliche Dimension dienen dem CLIL-Unterricht dabei als unterliegende Testkonstrukte die Kompetenzmodelle der einzelnen Fächer, die für Deutschland z.B. in den nationalen Bildungsstandards ausformuliert und durch exemplarische Testaufgaben illustriert werden. Damit hat die Messung der sachfachlichen Kompetenzen in CLIL dieselben Möglichkeiten sowie Konstruktions- und Legitimationsschwierigkeiten (vgl. z.B. Klieme, Avenarius, Blum, Döbrich, Gruber, Prenzel, Reiss, Riquarz, Rost, Tenorth & Vollmer 2007: 55-60) wie die Leistungsmessung in den Regelschulfächern. So läuft beispielsweise die empirische Untermauerung von (Ab-)Stufungen fachlicher Kompetenz für die Naturwissenschaften erst an, für viele Fächer ist sie noch gar nicht ausgearbeitet.

In Bezug auf die sprachliche Dimension ist vor allem der Gemeinsame Europäische Referenzrahmen für Sprachen (GeR) ein einflussreiches Werkzeug zur Einschätzung fremdsprachlicher Kompetenz, das u.a. in Beobachtungen mündlicher Leistungen, bei der Bewertung von schriftlichen Arbeiten, aber auch Portfolios zur Anwendung kommt

(vgl. KMK 2013: 15). Die Vor- und Nachteile des GeR (vgl. z.B. Harsch 2007; Hulstijn 2015) kommen dabei ebenfalls im CLIL-Unterricht zum Tragen. Als spezielle Schwierigkeit lässt sich im fachspezifisch orientierten CLIL-Kontext u.a. die Tatsache identifizieren, dass der GeR ein breites Spektrum allgemeiner sprachlicher Kompetenzen erfassbar zu machen sucht. Spezifika von wissenschaftspropädeutischem oder gar fachspezifischem Sprachgebrauch, wie sie für CLIL-Kontexte zentral sind, werden nicht gezielt in den Blick genommen (vgl. ausführlich Hulstijn 2015) und wesentliche Kompetenzdimensionen für den CLIL-Unterricht nicht abgebildet. Da in Deutschland im Gegensatz zum internationalen Feld der Bereich Prüfen und Testen von Fremdsprachenkompetenzen im schulischen Kontext insgesamt nur wenig entwickelt ist (vgl. z.B. Harsch 2007; Vogt 2010), muss man dies auch für CLIL-Unterricht bemängeln.

Vergleichbares lässt sich für den Bereich der interkulturellen Kompetenz konstatieren. Verschiedentlich ist auf die generelle Schwierigkeit der Messung interkultureller Kompetenz verwiesen worden (vgl. z.B. Byram 2009; Hu 2010). Von dominanten VertreterInnen wird eine Messbarkeit über Testformate überhaupt abgelehnt und Formen der Überprüfung favorisiert, die z.B. selbsteinschätzende und selbstreflektierende Momente beinhalten (vgl. z.B. Byram 2009). Diese aus der Perspektive des Fremdsprachenunterrichts stammende Problematik überträgt sich auch auf den CLIL-Unterricht. Im Gegensatz zu Fachkompetenz und (Fremd-)Sprach(en)kompetenz ist davon auszugehen, dass interkulturelle Kompetenz aufgrund fehlender Messinstrumente keine Rolle bei der Leistungsmessung im CLIL-Unterricht spielt.

3. Wie werden fachliche und sprachliche Kompetenz bei der Leistungsmessung gewichtet?

In den meisten CLIL-Formen im deutschsprachigen Kontext haben die Standards des Sachfachs Vorrang gegenüber der Dimension der fremdsprachlichen Kompetenz (vgl. Heine 2015 in der vorliegenden Ausgabe). Die Frage nach der Gewichtung beider Bereiche wird dabei unterschiedlich gehandhabt. Zumeist wird auf die Vorgaben aus dem regulären Sachfachunterricht zurückgegriffen. Sprachliche Kompetenz kann gemäß dieser Vorgaben dabei zu einem gewissen Teil bei der Bewertung der Gesamtleistung Berücksichtigung finden: Die „Darstellungsleistung“ soll mit bis zu 20 % der Gesamtleistung bewertet werden, „ein[...] Abzug von bis zu 2 Notenpunkten“ für gravierende sprachliche Normverstöße ist vorgesehen (KMK 2013: 55). Für eine Übertragung auf den CLIL-Unterricht ist hier zu fragen, woran sich die erwartete Norm orientieren soll; denn in den sich noch entwickelnden fremdsprachlichen Fähigkeiten der SchülerInnen sind Abweichungen von einer L1-Norm der erwartbare Regelfall. Auch ein Code-Switching zwischen der Erst- und der Fremdsprache, um etwa Lexiklücken zu überbrücken, entspricht weder der fremdsprachlichen noch der erstsprachlichen Norm des Fachdiskurses. Häufig wird in der Praxis daher davon abgesehen, sprachliche Leistungen überhaupt in eine Bewertung mit einfließen zu lassen; Sprache wird damit auf ihre Funktion, Gedankeninhalte zu kommunizieren, reduziert. Dies wiederum führt zu dem Problem, dass gute fremdsprachliche Leistungen nicht belegt und daher auch nicht positiv bemerkt werden können. Als eine Konsequenz davon zeigt Zydafiß (2007) für den Berliner Kontext auf, dass die höheren fremdsprachlichen Leistungen von CLIL-SchülerInnen im Vergleich zu den Leistungen von RegelschülerInnen durch die Benotungspraxis nicht honoriert werden, so dass eine erhebliche Schülerzahl den CLIL-Unterricht in der gymnasialen Oberstufe aus notenstrategischen Gründen nicht fortführt.

Eine weitere Schwierigkeit ergibt sich daraus, dass die Bedeutung einer normgerechten Sprachbeherrschung in Fachkontexten durch die fehlende Bewertungsdimension marginalisiert wird, wodurch ein Washback-Effekt entsteht: Wenn sie nicht als zentrale Leistung überprüft wird, rückt die Erarbeitung einer fachangemessenen Sprachkompetenz, z.B. in Bezug auf die Realisierung von fachspezifischen Textmustern, nicht gezielt als Lerngegenstand des CLIL-Unterrichts in den Blick und wird ergo auch nicht entwickelt. Coetzee-Lachmann (2007) etwa weist in ihrer Studie nach, dass es CLIL-SchülerInnen massiv an Kompetenzen in der Teilhabe an einem fachadäquaten Diskurs fehlt – was damit zusammenhängen kann, dass diese im Unterricht nicht expliziert und auch nicht eingefordert werden. Wichtig ist hier, dass dieses fachspezifische Sprach- und Diskurswissen auch in der Erstsprache sehr unterschiedlich entwickelt sein und daher nicht vorausgesetzt werden kann. Der Aufbau dieser sprachlichen Kompetenzdimensionen ist an die Aktivitäten des Faches gebunden und kann folglich nur in fachlichen Kontexten entstehen. Wird Lernenden beispielsweise in Prüfungssituationen angeboten, ihre Erstsprache zu verwenden – mit der Absicht, potentielle Schwierigkeiten bei der Abbildung von Fachkompetenz durch die Fremdsprache zu minimieren – so wird also implizit eine Kompetenz gemessen, deren Entwicklung nicht Gegenstand des Unterrichts gewesen ist.

Um in der Erstsprache für fachliche Kontexte adäquat gerüstet zu sein, müssen die relevanten Register- und Textsortenkompetenzen ebenfalls systematisch angebahnt worden sein. Hier verbindet sich CLIL direkt mit dem aktuellen Diskurs zu Sprachförderung in allen Fächern.

4. Zur Trennbarkeit der Konstrukte Fach und Sprache

In diesem Zusammenhang wird sehr deutlich, wie wenig eine klare Trennung von fachlicher und sprachlicher Kompetenz als Testkonstrukte möglich ist: In den traditionellen Prüfungsformen (schriftliche Arbeiten oder mündliche Prüfungen) der meisten Fächer dominiert die sprachliche Darstellung von Fachkompetenz weit über jede Art von nichtsprachlicher Kompetenzdemonstration. Da somit die Dimensionen der fachlichen Kompetenz in traditionellen Prüfungsformen in erster Linie über Sprache zugänglich werden, stehen Sachfachkompetenz und Kompetenz in der Arbeitssprache in einer zutiefst funktionalen Verbindung (die u.a. in den PISA-Studien immer wieder deutlich wird). Die enge Verwobenheit von Fach und Sprache wirft in diesem Zusammenhang die Frage auf, inwiefern fachliche Kompetenzen überhaupt ohne Rückgriff auf Sprache erhoben werden können. Zwar wäre denkbar, dass eine Bearbeitung von fachlichen Aufgabenstellungen je nach Gegenstandsbereich z.B. mittels physischer Aktivitäten wie dem tatsächlichen Konstruieren oder Manipulieren von Gegenständen oder dem Rückgriff auf visuelle Daten wie dem Erstellen von Grafiken, dem Malen von Bildern etc. möglich ist. Bereits bei der Vermittlung der Aufgabenstellung stößt man aber auf die Schwierigkeit, dass ohne Einbezug von Sprache zumindest komplexe oder spezifische Anforderungen kaum vermittelbar sind. Auch lässt sich ein nichtsprachliches Vorgehen sicher nicht für alle Fachzusammenhänge realisieren; so steht z.B. der Geschichtsunterricht aus seiner Selbstdefinition heraus in untrennbarer Verbindung mit der sprachlichen Vermitteltheit von historischen Ereignissen. Staschen-Dielmann (2012) vertritt in der Konsequenz die Position, dass sich sprachliches und fachliches Lernen im Geschichtsunterricht nicht unabhängig voneinander beurteilen lassen. Sie entwirft daher holistische Skalen und Niveaustufen zur Bewertung von narrativer Kompetenz im bilingualen Geschichtsunterricht. Dabei nimmt sie an, dass die von den SchülerInnen verwendete Sprache nicht nur Aufschluss über Sprachkompetenz, sondern auch den Kompetenzgrad in Hinblick auf fachspezifisches Denken gibt. Coetzee-Lachmann (2007) entwickelt getrennte Skalen zur detaillierten Einschätzung von fachlicher Vollständigkeit bzw. Angemessenheit einerseits und der Angemessenheit der sprachlichen Realisierung andererseits, die jeweils unterschiedliche Fokussierungen vornehmen und u.a. eine strukturierte Rückmeldung an Lernende möglich machen (vgl. Heine & Mainzer 2015 in der vorliegenden Ausgabe).

5. Fazit

Die Leistungsmessung im CLIL-Unterricht gestaltet sich komplex. Derzeit wird vor allem die fachliche Leistung bewertet, was nicht zuletzt auf dem Wunsch beruht, CLIL-SchülerInnen durch die Fremdsprache als Arbeitssprache nicht zu benachteiligen. Dies führt allerdings zu mehreren Schwierigkeiten: a) das Testkonstrukt „fachliche Kompetenz“ ist nicht sauber von Sprache trennbar, da der fachliche Kompetenzgrad fast ausschließlich über Sprache erhoben wird (bzw. werden kann) und damit sprachliche Kompetenz stets miterfasst wird; b) als Washback-Effekt wird ein Sprachfokus in CLIL vernachlässigt und relevante Registerkompetenzen nicht entwickelt. Auch ein Ausweichen auf die Erstsprache der Lernenden löst diese Probleme nicht, denn CLIL bedarf spezifischer sprachlicher Kompetenzen. Diese müssen – egal in welcher Sprache – gezielt entwickelt werden. Hieraus erwächst ein Plädoyer für eine Leistungsmessung, die sowohl fachliche als auch (fremd-)sprachliche Dimensionen abbildet, was in der Folge auch bedeutet, dass die interne Wechselwirkung beider Ebenen gezielt im Unterricht in den Blick genommen und der angemessene Umgang mit ihr ausgebildet werden muss.

Literatur

Byram, Michael (2009), Evaluation and/or assessment of intercultural competence. In: Hu, Adelheid & Byram, Michael (Hrsg.), *Interkulturelle Kompetenz und fremdsprachliches Lernen. Modelle, Empirie, Evaluation. Intercultural Competence and Foreign Language Learning. Models, Empiricism, Assessment*. Tübingen: Narr, 215-234.

- Coetzee-Lachmann, Debbie (2007), *Assessment of subject-specific task performance of bilingual learners. Analyzing aspects of subject-specific written discourse*. E-Dissertation Universität Osnabrück [Online unter https://repositorium.uni-osnabrueck.de/bitstream/urn:nbn:de:gbv:700-2009030617/2/E-Diss864_thesis.pdf 17.09.2015].
- Harsch, Claudia (2007), *Der gemeinsame europäische Referenzrahmen für Sprachen. Leistung und Grenzen*. Saarbrücken: VDM Verlag Dr. Müller.
- Harsch, Claudia (2015), Assessment Literacy – Trend oder Notwendigkeit? In: Böcker, Jessica & Stauch, Anette (Hrsg.), *Konzepte aus der Sprachlehrforschung – Impulse für die Praxis. Festschrift für Karin Kleppin*. Frankfurt a.M.: Peter Lang, 489-510.
- Heine, Lena (2015), Lernziele. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 20: 2, ??-?? [Online unter http://zif.spz.tu-darmstadt.de/jg-20-2/beitrag/Heine_Lernziele.pdf].
- Heine, Lena & Mainzer, Mirka (2015), Rückmeldung an Lernende. *Zeitschrift für Interkulturellen Fremdsprachenunterricht* 20: 2, ??-?? [Online unter http://zif.spz.tu-darmstadt.de/jg-20-2/beitrag/Heine_Mainzer.pdf].
- Hu, Adelheid (2010), Aspekte der Evaluation interkultureller Kompetenzen im Fremdsprachenunterricht. In: Berndt, Annette & Kleppin, Karin (Hrsg.), *Sprachlehrforschung: Theorie und Empirie. Festschrift für Rüdiger Grotjahn*. Frankfurt a.M.: Peter Lang, 61-72.
- Hulstijn, Jan (2015), *Language Proficiency in Native and Non-native Speakers. Theory and research*. Amsterdam: John Benjamins.
- Klieme, Eckhard; Avenarius, Hermann; Blum, Werner; Döbrich, Peter; Gruber, Hans; Prenzel, Manfred; Reiss, Kristina; Riquarz, Kurt; Rost, Jürgen; Tenorth, Hans-Elmar & Vollmer, Helmut J. (2007), *Zur Entwicklung nationaler Bildungsstandards. Expertise*. Berlin: BMBF.
- KMK (2013), *Bericht: Konzepte für den bilingualen Unterricht – Erfahrungsbericht und Vorschläge zur Weiterentwicklung. Beschluss der Kultusministerkonferenz vom 17.10.2013* [Online unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2013/2013_10_17-Konzepte-bilingualer-Unterricht.pdf 27.05.2015].
- Staschen-Dielmann, Susanne (2012), *Narrative Kompetenz im bilingualen Geschichtsunterricht. Didaktische Ansätze zur Förderung der schriftlichen Diskursfähigkeit*. Frankfurt a.M.: Peter Lang.
- Vogt, Karin (2010), Assessment Literacy bei Fremdsprachenlehrkräften – Bedarf und Bedürfnisse in Aus- und Weiterbildung. In: Altmayer, Claus; Mehlhorn, Grit; Neveling, Christiane; Schlüter, Norbert & Schramm, Karen (Hrsg.), *Grenzen überschreiten: sprachlich – fachlich – kulturell. Dokumentation zum 23. Kongress für Fremdsprachendidaktik der Deutschen Gesellschaft für Fremdsprachenforschung (DGFF) Leipzig, 30 September – 3. Oktober 2009*. Baltmannsweiler: Schneider, 267-277.
- Zydati, Wolfgang (2007), *Deutsch-Englische Züge in Berlin (DEZIBEL). Eine Evaluation des bilingualen Sachfachunterrichts an Gymnasien. Kontext, Kompetenzen, Konsequenzen*. Frankfurt a.M.: Peter Lang.