

## 5. Sprachdiagnose und Evaluation

### 5.1 Sprachdiagnose und Evaluation als Bausteine von Unterrichtsentwicklung

In den vorangegangenen Kapiteln wurden Prinzipien für die Entwicklung von Aufgaben vorgestellt, die den spezifischen Bedürfnissen von Schülern undH Rechnung tragen und das Lateinische als Brückensprache zum Deutschen als Zweitsprache nutzen. Der Entschluss, auf den hier vorgestellten Grundsätzen basierend neue Aufgaben und Arbeitsmaterialien zu entwickeln und im eigenen Unterricht zu erproben, kann als eine Maßnahme der Unterrichtsentwicklung betrachtet werden. Das Ziel von Unterrichtsentwicklung besteht in der Verbesserung der eigenen Unterrichtspraxis. Den Weg dorthin bilden kontinuierliche Diagnose und systematische Reflexion des Unterrichtshandelns. Unterrichtsentwicklung kann also beschrieben werden als ein „ständiger und kontinuierlicher Erfindungsprozess, als Experiment von Lehrerinnen und Lehrern [...], dessen systematische Anteile oft erst in der Rückschau erkennbar sind“ (BASTIAN 2007: 30). Ein zentraler Gedanke der Unterrichtsentwicklung ist, dass Unterrichtsqualität nicht durch einfache „Unterrichtsrezepte“ zu erreichen ist. Vielmehr entsteht guter Unterricht in einem Prozess der beständigen (kollegialen) Reflexion und Weiterentwicklung der unterrichtlichen Praxis unter den jeweiligen Bedingungen, die Lehrer an ihrer Schule und in ihren Klassen vorfinden. In diesem Zusammenhang kommen der diagnostischen Kompetenz der Lehrkräfte und der Evaluation des Unterrichts eine besondere Bedeutung zu. Ohne sie ist eine systematische Unterrichtsentwicklung kaum denkbar, da sie die Grundlage für die Reflexion und Verbesserung der eigenen Praxis bilden (vgl. PEEK 2007: 141). In diesem Kapitel werden daher einige zentrale Grundsätze der Sprachdiagnose und Evaluation vorgestellt und anhand einiger ausgewählter, im Lateinunterricht erprobter Messinstrumente illustriert.

### 5.2 Die Bedeutung diagnostischer Kompetenz

„Sprachdiagnoseinstrumente haben die Aufgabe, den erreichten Stand eines Lernenden in der Fremd- oder Zweitsprache zu ermitteln“ (RÖSCH 2011a: 37). Die Diagnose kann dabei bereits vor dem eigentlichen Unterricht erfolgen: Um im Unterricht die Bedürfnisse der Lernenden angemessen zu berücksichtigen, ist es sinnvoll, vor dem Beginn einer sprachfördernden Maßnahme den vorhandenen Lern-

stand und bereits erreichte Kompetenzen festzustellen (vgl. RÖSCH 2011a: 37). Anhand dieser ersten Einschätzung können gezielt differenzierende Lern- und Förderangebote in den Lateinunterricht integriert werden. In der Regel wird die Sprachdiagnose jedoch nicht auf das einmalige Erfassen des Status Quo reduziert, sondern muss vielmehr kontinuierlich im Unterricht verfolgt werden. Somit sollen nicht nur punktuell aktuelle Sprachstände, sondern linear längere Sprachenentwicklungen beobachtet werden, um daraus Erkenntnisse für eine langfristig wirksame Sprachförderung zu gewinnen, deren Erfolg dann durch Methoden der Evaluation festgestellt werden kann (vgl. RÖSCH 2011a: 37).

#### Methoden der Sprach- diagnose

Zur Sprachdiagnose haben sich verschiedene Methoden etabliert. Die wohl gängigste und von allen Lehrkräften praktizierte Form ist die Beobachtung. Diese orientiert sich am Sprachmaterial, das die Lernenden im Unterricht mündlich oder schriftlich produzieren. Obgleich es für Deutsch als Fremdsprache durchaus viele Tests zur Sprachdiagnose gibt, existieren bis heute „keine Verfahren zur Sprachstandsfeststellung, die für den Erwerb des Deutschen als Zweitsprache normiert sind“ (JEUK 2010: 83), ja sogar liegen „für Schülerinnen und Schüler der Sekundarstufe keine Verfahren vor“ (JEUK 2010: 91). Dennoch werden in der DaZ-Forschungsliteratur einige Methoden vorgestellt, die ursprünglich aus der Forschung zu Deutsch als Fremdsprache stammen.

#### Profilstufen- analyse

Eine systematische Sprachdiagnose kann z.B. mittels Profilstufenanalysen oder „C-Tests“ erfolgen, die mit wenig Aufwand ein recht valides Bild über die globalen Sprachfähigkeiten einer Lerngruppe liefern können (vgl. RÖSCH 2011a: 57–59). In Hinblick auf syntaktische Qualifikationen hat W. GRISSHABER<sup>1</sup> ein Konzept zur Erfassung des Sprachstandes entwickelt, das sich an den Erwerbssequenzen der Lernenden orientiert und die Verbstellung fokussiert (vgl. JEUK 2010: 62 ff.). Er arbeitet mit Schülertexten oder mündlichen Sprachprodukten und analysiert sie hinsichtlich ihrer sprachlichen Komplexität. Dabei wird erfasst, welches sprachliche Niveau der DaZ-Lernende im Deutschen bereits erreicht hat. In dieser Hinsicht unterscheidet sich GRISSHABERS Diagnoseform von den meisten anderen gängigen

<sup>1</sup> Eine Auswahl an Publikationen von W. GRISSHABER ist auf der Homepage der Universität Münster einzusehen: <http://spzwww.uni-muenster.de/~griessa/index.html>. (Stand: Juli 2013). Hier sind auch Informationen zur Profilstufenanalyse abrufbar. Vgl. Rösch 2011a: 51 ff.

Messmethoden, die eher auf die Auffindung von Defiziten ausgerichtet sind. GRIESSHABER entwickelte folgendes Profilstufensystem:

- Stufe 0: Bruchstückhafte Äußerungen: *Papa kommen.*
- Stufe 1: Finites Verb in einfachen Äußerungen: *Papa kommt.*
- Stufe 2: Verbalklammer: *Papa kommt zurück.*
- Stufe 3: Inversion: *Morgen kommt Papa.*
- Stufe 4: Nebensätze mit finitem Verb in Endstellung: *Papa kommt, nachdem er einkaufen gewesen ist.*
- Stufe 5: Insertion eines Nebensatzes: *Papa, der gerade einkaufen gewesen ist, kommt.*
- Stufe 6: Insertionen eines erweiterten Partizipialattributs in einer Nominalkonstruktion: *Papa hat die von Mama auf eine Liste geschriebenen Lebensmittel eingekauft.*

Nach GRIESSHABER entscheidet die Häufigkeit der jeweiligen Profilstufen, die in den Sprachprodukten festgestellt werden, über den Sprachstand der Testperson. Ist eine Profilstufe mindestens dreimal geäußert worden, gilt sie als erlangt<sup>2</sup>.

Ein weiteres (oben schon erwähntes) gängiges Instrument, das bereits im Fremdsprachenunterricht zum Einsatz kommt und mit dessen Hilfe Kompetenzen der Morphosyntax sowie des Leseverständnisses erfasst werden können, ist der sog. C-Test (vgl. JEUK 2010: 91, RÖSCH 2011a, 57–59). Gerade dieser Test gilt als probates Mittel und wird auch im schulischen Bereich als Instrument zur Bestimmung der allgemeinen sprachlichen Kompetenz eingesetzt (vgl. BAUR/GROTJAHN/SPETTMANN 2006). Er entwickelte sich aus den bis dahin üblichen Lückentexten („CLOZE-Tests“) und weist bis heute einige typische Merkmale dieser auf: Grundlage bildet ein inhaltlich zusammenhängender Basistext, der in vier Subtexte unterteilt wird und sich an der Erfahrungswelt und dem Sprachniveau der Lernenden orientieren sollte (vgl. RÖSCH 2011a: 57). Die Subtexte werden durch einen ersten vollständigen Satz eingeleitet, der keine Lücken hat und somit eine inhaltliche Vorerschließung ermöglicht. Nach diesem ersten Satz folgt ein Text mit insgesamt zwanzig Lücken (ca. bei jedem dritten Wort), wobei allerdings nie ganze Wörter, sondern immer nur die Wortenden getilgt werden und anschließend inhaltlich und gram-

C-Test

<sup>2</sup> Die Auswertung solcher Profilstufenanalysen ist mitunter recht aufwendig, da man für jeden Satz und Nebensatz die erreichte Profilstufe festlegen muss. Durch das Verfahren können aber auch typische Fehlerquellen deutlich werden, wenn die rein syntaktische Analyse durch eine Fehleranalyse ergänzt wird.

matikalisch korrekt von den Schülern zu ergänzen sind. Für die Bearbeitung der C-Tests sind insgesamt zwanzig Minuten nötig (pro Subtext fünf Minuten). Bei der Auswertung ist nicht die Lösung der einzelnen Subtexte oder gar einzelner Lücken von Belang, sondern der Summenwert des gesamten Tests. Somit gilt der C-Test als relativ ökonomisch und eindeutig hinsichtlich der Durchführung, Auswertung und Interpretation (vgl. BAUR/GROTJAHN/SPETTMANN 2006: 2). Als problematisch gilt, dass der C-Test nicht alle Aspekte der Sprachkompetenz misst. So bleiben produktive Bereiche, Lesefähigkeit sowie das Hörverstehen unbeachtet. Auch Fähigkeiten im Bereich der Morphologie, Syntax und Orthographie werden durch eine ausschließliche Richtig-falsch-Korrektur zumeist nicht erfasst. Dem kann allerdings durch eine komplexe Auswertung (z.B. durch die Vergabe von je einem Punkt für Worterkennung, morphologische Realisierung und Orthographie) entgegengewirkt werden (vgl. BAUR/GROTJAHN/SPETTMANN 2006: 6), sodass ein differenzierteres Bild der Sprachstände erfasst wird. Im Lateinunterricht kann ein solcher C-Test polyvalent eingesetzt werden: zur Überprüfung oder Vermittlung von Inhalten, aber auch zur Sprachförderung, da Lückentexte zu den sprachbildenden Aufgabenformaten gezählt werden (vgl. Kapitel 4.2.5).

#### Teilfertig- keitstest

Um gezielt grammatische Fähigkeiten zu überprüfen, sind außerdem Teilfertigkeits-tests ratsam, die durch systematische Tilgung bestimmter Wortgruppen die differenzierte Analyse des Sprachstandes ermöglichen (vgl. BAUR/GROTJAHN/SPETTMANN 2006: 11 ff.). Hier kann z.B. anhand einer Liste typischer Stolpersteine im Deutschen eine Orientierung zur Gestaltung der Lücken gegeben werden, die zum jeweils behandelten Thema des Lateinunterrichts passen (z.B. Dativendungen, Tempora, Präpositionen und Rektion etc.).

Die Welt der Römer	
<p>Reichtum und Armut, mutige Soldaten und finstere Machtkämpfe, Friede unter den Völkern, blutige Sklavenaufstände, grausame Zirkusspiele – all das ist Rom. Die Römer herr_____ über ein Weltr_____ voll von Gegens_____. Bis heute kö_____ wir die prächtigen Bau_____ und andere Hinterlassens_____ bewundern. Die lateinische Spr_____ prägte unser Wi_____ und Denken. Auch des_____ nennt man R_____ die „ewige Stadt“. Dab_____ gab es anfangs ni_____ weiter als e_____ paar kleine Bauernh_____ und Siedlungen am U_____ eines großen Flus_____. Angeblich haben d_____ Zwillinge Romulus und Remus diese spä_____ Weltstadt gegründet. Rom w_____ auf sieben Hügeln err_____.</p>	Korrekt gelöst: _____ / 20
<p>Die klassische „familia“ war keine Kleinfamilie aus Mutter, Vater und Kindern wie wir sie heute kennen. Man vers_____ darunter die gesa_____ Hausgemeinschaft, die unter ei_____ Dach lebte. Zu El_____ und Kindern kamen Verw_____ wie Großmutter, Schwiegertö_____ oder unverheiratete O_____ und Tanten hinzu. Auch Sk_____, die sich d_____ reichen Römer als Haushaltsgehilfen leist_____ konnten, gehörten zur Fa_____. Sie arbeiteten u_____ lebten im selben H_____. Alle Aufgaben waren n_____ strengen Regeln verteilt. D_____ Vater war der Hausvorst_____. Der Platz al_____ Frauen war der Hausha_____. Sie koc_____ und sorgten für Klei_____ und Möbel.</p>	Korrekt gelöst: _____ / 20

Abb. 1: C-Test zur globalen Sprachstandanalyse

Auch schriftliche Leistungs- und Wissenstests sind gängige Mittel zur Sprachdiagnose, die im Lateinunterricht regelmäßig zum Einsatz kommen und vor allem Hinweise auf den Stand der Sprachkompetenz im Lateinischen geben sollen. So zielen solche Tests zumeist auf die Erfassung einzelner Bereiche des Unterrichts ab (vgl. RÖSCH 2011a: 40), wie z.B. den Wortschatzerwerb in einem Vokabeltest. Durch leichte Änderungen im Testaufbau können aber auch Kompetenzen im Bereich der deutschen Sprachfertigkeiten erfasst werden, wie durch folgendes Beispiel illustriert werden soll.

Dieser Zwischentest wurde zum Thema Präpositionen und ihrer Rektion entwickelt und zur Erfassung der Fähigkeiten im Lateinischen sowie im Deutschen konzipiert. Ausgehend von vier lateinischen Sätzen werden zunächst dort die präpositionalen Verbindungen identifiziert, dann übersetzt und sodann in den halbfertigen deutschen Satz eingetragen (Abb. 2). Schließlich werden die präpositionalen Verbindungen tabellarisch gesammelt und einander im Deutschen und Lateinischen gegenübergestellt, um die Unterschiede in der Kasusreaktion deutlich zu machen.

Fragen zur lateinischen Grammatik
<p>1. a) Sieh dir bitte den folgenden lateinischen Text an. Unterstreiche in den lateinischen Sätzen jeweils die Präposition und ihr Bezugswort (pro Satz also zwei Wörter).</p> <p>1.1 Marcus et Valeria amicum ante basilicam exspectant.            1.2 Sed amicus ad tabernam currit.            1.3 Itaque liberi ad tabernam properant.            1.4 Cuncti in tabernam currunt.</p> <p>1. b) Nun ergänze in den folgenden Sätzen der Übersetzung die präpositionalen Ausdrücke und trage sie in die Lücken ein:</p> <p>1.1 Marcus und Valeria erwarten ihren Freund _____            1.2 Aber ihr Freund _____            1.3 Daher eilen die Kinder _____            1.4 Alle laufen _____</p>

Abb. 2: Auszug aus dem zweiten Teilfertigkeitstest der Studie „Pons Latinus“ zum Thema Präpositionen und Rektion

Die Diagnose von Sprachständen bildet zumeist die Grundlage für eine Interventionsmaßnahme, deren Erfolg oder Misserfolg anschließend gemessen werden soll, was anhand evaluativer Methoden geschehen kann.

### 5.3 Was versteht man unter Evaluation?

„Evaluation beinhaltet die systematische Anwendung empirischer Forschungsmethoden zur Bewertung des Konzeptes, des Untersuchungsplanes, der Implementierung und Wirksamkeit sozialer Interventionsprogramme“ (BORTZ/DÖRING <sup>2</sup>1995: 96), wobei es sich um einen systematischen und zielgerichteten Aus- und Bewertungsprozess handelt (vgl. PEEK 2007: 143).

#### Definition

Ausgangspunkt jeder Evaluation ist ein vorab festgelegtes Ziel bzw. eine Fragestellung. So ist die Problemformulierung oder Fragestellung Ausgangspunkt und erstes Kriterium für die Evaluationsforschung (vgl. BORTZ/DÖRING <sup>2</sup>1995: 103). Im schulischen Kontext könnte sich die Evaluation beispielsweise auf den Lernerfolg der Schüler richten. Es könnten jedoch auch andere Aspekte einbezogen werden oder möglicherweise sogar in den Vordergrund treten, beispielsweise die Motivation der Schüler, die Verständlichkeit der Aufgaben, der Bezug der neu entwickelten Materialien zum Lehrbuch usw. In jedem Fall sollte die Evaluation der eigenen Unterrichtspraxis einen speziellen Fokus bekommen, um sich nicht in der Fülle von Aspekten, die den Unterricht beeinflussen, zu verlieren. Ein Modell, das mögliche Fragestellungen und, darauf aufbauend, Vorgehensweisen bei der Evaluation systematisiert, ist das sog. *CIPP-Modell* von STUFFLEBEAM und SHINKFIELD (2007), das unten vorgestellt wird.

Eine Evaluation ist grundsätzlich datengestützt, d.h., zur Beantwortung der eingangs formulierten Fragen werden systematisch Daten erhoben und ausgewertet. Daten können durch Beobachtung oder Befragung gesammelt werden. Auch Arbeitsproben und Testergebnisse der Schüler können als Daten für die Analyse herangezogen werden. Oft ist es sinnvoll und aufschlussreich, verschiedene Datenquellen und Methoden zu kombinieren, um so eine multiperspektivische und facettenreiche Antwort auf die untersuchte Fragestellung zu erhalten. Die anschließende Bewertung sollte auf der Grundlage vorher (ggf. gemeinsam mit Kollegen) festgelegter Kriterien erfolgen. Bei der kollegialen Aus- und Bewertung im Zusammenhang mit der Unterrichtsentwicklung ist es besonders wichtig, dass die Ergebnisse der Evaluation transparent, also nachvollziehbar für Dritte, als Planungs- und Entscheidungshilfe für mögliche Verbesserungen des Unterrichts dienlich sind (vgl. PEEK 2007: 143).

Je nachdem, ob die Evaluation von Außenstehenden oder den Betroffenen selbst durchgeführt wird, unterscheidet man zwischen

### Externe und interne Evaluation

externer und interner Evaluation. Da dieses Kapitel für Personen geschrieben ist, die die eigene Unterrichtspraxis evaluieren möchten, wird im Weiteren der Fokus auf die interne Evaluation gelegt, wobei durchaus angeregt wird, Kollegen oder andere „kritische Freunde“ (vgl. COSTA/KALLICK 1993: 49f.) als externe Berater hinzuzuziehen, wobei in diesem Fall i.d.R. spezifische Datenschutzregularien zu beachten sind. Die Vorschläge für das methodische Vorgehen sind der Aktionsforschung entlehnt, einem Ansatz, der den Lehrer als forschenden Praktiker in den Mittelpunkt stellt (vgl. ALTRICHTER/POSCH 1998). Die Aktionsforschung sieht die Evaluation des Unterrichts klar im Dienste der Reflexion und Entwicklung des eigenen Handelns und betont die Praktikabilität der Evaluation sowie den kollegialen Austausch gegenüber Aspekten wie Objektivität und Repräsentativität.

### 5.4 Evaluation nach dem CIPP-Modell

Evaluation gehört in gewisser Hinsicht immer schon zur Unterrichtstätigkeit von Lehrern, da sie z.B. regelmäßig und systematisch die Lernleistungen der Schüler mit Blick auf bestimmte vorab festgelegte Lernziele überprüfen. Eine solche „lernzielbasierte Evaluation“ hat eine lange Tradition in der Evaluationsforschung (vgl. TYLER 1942) – sie greift jedoch im vorliegenden Fall zu kurz, da ihre Ergebnisse nur begrenzt Rückschlüsse auf die Unterrichtsqualität und vor allem Maßnahmen zu ihrer Verbesserung erlauben. Im Kontext der Unterrichtsentwicklung sollte eine Evaluation ganzheitlicher erfolgen. Ein Modell hierfür haben STUFFLEBEAM und SHINKFIELD (2007) vorgelegt: das *CIPP-Modell* (Context, Input, Process, Product). Dieses Modell entstammt zwar nicht primär der Unterrichtsentwicklung, lässt sich jedoch auf die Evaluation des schulischen Unterrichts beziehen (vgl. SCHMIDT/PERELS 2010: 14ff.). Es benennt vier wesentliche Ebenen, die bei der Beurteilung einer Unterrichtsmaßnahme betrachtet werden sollten, um zu aussagekräftigen Ergebnissen zu gelangen.

#### Context

*Context* meint die Rahmenbedingungen des Unterrichts, z.B. organisatorische Bedingungen wie Klassengröße, Unterrichtszeiten, räumliche Gegebenheiten. Weiterhin fallen darunter die allgemeinen Lernvoraussetzungen der Schüler, ihre Vorkenntnisse, ihr Interesse am Thema, ihr allgemeines Leistungsniveau und die Zusammensetzung der Lerngruppe. Schließlich können im Rahmen der Kontextevalua-



tion auch die Zielsetzungen für den Unterricht mit evaluiert werden (z.B. in welchem Ausmaß die Entwicklung fachlicher/überfachlicher Kompetenzen eine Rolle spielte).

Die Evaluation des *Input* untersucht die konkreten Aufgaben und Methoden, die im Unterricht eingesetzt werden, also das Material und die Ressourcen, mit denen die Schüler lernen. Aspekte, die hier betrachtet werden könnten, sind die aufgewendete Unterrichtszeit, die eingesetzten Sozialformen und Medien, Methodenwechsel und Motivierungsmaßnahmen, Unterstützungsmaßnahmen usw.

Input

Die *Process*-Evaluation richtet ihr Augenmerk auf die Lernprozesse der Schüler, die Art und Weise, wie sie das Lernmaterial nutzen bzw. im Unterricht mitarbeiten. Hier könnte zum Beispiel im Vordergrund stehen, wie die Schüler die Aufgaben einschätzen, welche Aufgaben ihnen Freude, welche ihnen Schwierigkeiten bereiten, an welchen Stellen sie Verständnisprobleme haben, ob das Schwierigkeitsniveau angemessen ist, inwieweit der Unterricht bzw. die verwendeten Aufgaben und Methoden für leistungsstarke und leistungsschwache Schüler geeignet sind usw.

Process

Im Rahmen der *Product*-Evaluation wird schließlich das Lernergebnis der Schüler betrachtet und analysiert, ob alle die angestrebten Lernziele erreichen und wo ggf. noch Lücken bestehen. Die Ebene der Evaluation kann weiterhin das subjektive Resümee der Beteiligten einschließen – sind die Lernenden, ist die Lehrkraft mit den erreichten Ergebnissen zufrieden?

Product

Eine umfassende Erhebung aller vier Ebenen ist außerordentlich komplex und im Rahmen der Unterrichtsevaluation durch Lehrer kaum zu leisten. Für die Unterrichtsentwicklung ist dies jedoch auch nicht unbedingt erforderlich. Das Modell bietet vielmehr eine Systematisierung und Strukturierung der Aspekte, die zur Beantwortung von Evaluationsfragestellungen relevant sind. Es sollte deshalb als Hilfestellung bei der Planung genutzt werden, wobei sich die Evaluation auf ausgewählte Teilfragen der vier Ebenen oder auch nur auf eine Evaluationsebene konzentrieren kann (vgl. SCHMIDT/PERELS 2010: 15).

## 5.5 Evaluationsmethoden

Die Datenerhebung im Rahmen der Evaluation kann mit unterschiedlichen Methoden erfolgen, wobei jede Methode grundsätzlich für mehrere der im Vorhergehenden dargestellten Evaluationsebenen

geeignet ist. Die wichtigsten bei der Evaluation verwendeten Erhebungsmethoden sind die Beobachtung, mündliche und schriftliche Befragungen sowie Tests und Arbeitsproben der Schüler.

### 5.5.1 Beobachtung

„Keine Datenerhebungsmethode kann auf Beobachtung verzichten.“ (BORTZ/DÖRING <sup>2</sup>1995: 240) Die Beobachtung ist für Lehrer selbstverständlicher Teil des Unterrichts, da sie kontinuierlich Verhalten und Lernprozesse der Schüler im Auge haben und ihren Unterricht hierauf abstimmen.

#### Kriterien der Methode

Als Evaluationsmethode sollte die Beobachtung systematisch und nachvollziehbar sein. Will man eine Selbstbeobachtung durchführen, kann man sich z.B. eines Tagebuchs bedienen, in dem man Gedächtnisprotokolle des eigenen Unterrichts anfertigt und reflektiert. ALTRICHTER und POSCH (1998: 26) sehen in solchen Tagebüchern einen wichtigen Ausgangspunkt für die Erforschung des eigenen Unterrichts. Für die Evaluation des Unterrichts Gewinn bringend sind darüber hinaus Fremdbeobachtungen, z.B. durch Kollegen (vgl. SCHMIDT/PERELS 2010: 55). Das Ziel der Beobachtung sollte vorher zwischen Beobachter und Beobachtetem geklärt werden, wobei der Beobachtete festlegt, zu welchen Aspekten er sich eine Rückmeldung wünscht. Auch sollte der Beobachter möglichst nicht ins Unterrichtsgeschehen involviert werden, sondern eine passive Rolle im Unterricht einnehmen.

Die Beobachtung kann, je nach Fragestellung, mehr oder weniger stark strukturiert erfolgen. In einer unstrukturierten Form protokolliert der Beobachter frei alle Ereignisse im Unterricht, die ihm im Rahmen des vorher vereinbarten Fokus auffallen. Für die strukturierte Beobachtung werden vorab Beobachtungsbögen entwickelt, auf denen der Beobachter die Häufigkeit bestimmter Verhaltensweisen festhält oder bestimmte Unterrichtsphasen oder Ereignisse einschätzt. Beobachtungsbögen können helfen, die Beobachtung zu fokussieren und für die Beteiligten nachvollziehbarer zu machen. Die Beobachtung von Unterricht sollte unbedingt zunächst auf der Ebene der sachlichen Beschreibung erfolgen, bevor interpretiert wird. Eine solche sachliche Beschreibung hilft, sich zunächst von den Geschehnissen zu distanzieren, um dann im zweiten Schritt verschiedene Bewertungen gegeneinander abzuwägen und sinnvolle Maßnahmen zur Optimierung zu entwickeln.

### 5.5.2 Befragung

Befragungen liefern gewissermaßen die „Innenperspektive“ des Unterrichts, insofern als Schüler in Befragungen zu ihrer subjektiven Wahrnehmung, die sich durch Beobachtungen nur schwer erschließen lässt, Auskunft geben. Befragungen bilden deshalb eine gute Ergänzung zu Beobachtungen.

Befragungen können schriftlich in Form von Fragebögen oder mündlich als Interview erfolgen. Ähnlich wie die Beobachtung kann auch die Befragung mehr oder weniger stark strukturiert werden. Bei der Befragung mithilfe von Fragebögen sind relativ stark strukturierte Formate üblich, bei denen den Befragten in der Regel Aussagen vorgelegt werden, zu denen sie den Grad ihrer Zustimmung angeben müssen.

#### Fragebögen

Fragebögen dieser Art haben den Vorteil, dass eine große Menge von Personen in relativ kurzer Zeit befragt werden kann und die Antworten sehr gut vergleichbar sind. Auch ist der Einfluss des Fragenden (im Vergleich zum Interview) relativ gering. Die Fragen sollten eindeutig sein, sie sollten nur Aspekte erkunden, zu denen der Befragte auch eine Antwort geben kann, komplizierte Formulierungen und doppelte Verneinungen sollten vermieden werden (vgl. PORST 2000). Die Fragen müssen dem sprachlichen Niveau der Sprach- und Lesekompetenz der Lernenden angepasst werden und können ggf. vorgelesen werden. Zur Strukturierung dieser persönlichen Einschätzungen dienen sog. *Rating-Skalen* (z.B. Zahlen, Beschreibungen, Symbole), zu denen der Proband den Grad des Zutreffens anhand der vorgegebenen Ausprägungsgrade einschätzen soll. Der Proband kreuzt diejenige Stufe der Rating-Skala an, der er hinsichtlich einer vorgegebenen Aussage subjektiv am meisten zustimmen kann (vgl. BORTZ/DÖRING <sup>2</sup>1995: 163f.). Bei diesem Frageformat ist darauf zu achten, die Kategorien möglichst allgemein zu fassen, wobei sich Formulierungen wie *trifft überhaupt nicht zu*, *trifft eher nicht zu*, *trifft teilweise zu*, *trifft eher zu* und *trifft genau zu* bewährt haben.

Fragen zu den Unterrichtsmaterialien					
Lies dir die folgenden Aussagen (linke Spalte) durch und kreuze an, in welchem Maß sie für dich zutreffen (bitte nur ein Kreuz machen).					
	trifft gar nicht zu	trifft eher nicht zu	trifft teilweise zu	trifft eher zu	trifft genau zu
2.1 Die Unterrichtsmaterialien fand ich allgemein gut.					
2.2 Die einzelnen Aufgaben waren für mich gut strukturiert.					
2.3 Die Hinweise zu Vokabeln und Grammatik waren für mich immer hilfreich und verständlich formuliert.					
2.4 Die Informationstexte waren für mich immer klar und verständlich formuliert.					

Abb. 3: Beispiel für eine verbale Rating-Skala zur Einschätzung der selbstentwickelten Unterrichtsmaterialien.

Um den Befragten die Gelegenheit zu geben, Aspekte einzubringen, die in den Fragen nicht angesprochen wurden, sollten Fragebögen möglichst auch einige offene Fragen enthalten, die die Möglichkeit eines frei formulierten Feedbacks einräumen. Hierzu zählen offene Fragen bzw. Aufgaben mit einem freien Antwortformat ohne vorgegebene Antwortalternativen (vgl. MOOSBRUGGER/KELAVA <sup>2</sup>2012: 40), also z.B. Kurzaufsätze, Ergänzungsaufgaben und Lückentexte. Ferner sind Aufgaben mit Fragen in einem gebundenen Antwortformat möglich. Hier wird eine gewisse Auswahl an Antwortalternativen vorgegeben (vgl. MOOSBRUGGER/KELAVA <sup>2</sup>2012: 43), wie z.B. bei Ordnungs-, Auswahl- (Multiple Choice) und Beurteilungsaufgaben (Rating-Aufgaben). Gerade bei der Befragung von Schülern empfiehlt sich ein „Probelauf“ mit einigen wenigen Personen der Zielgruppe, um die Fragen sowie die Antwortskala auf ihre Verständlichkeit zu testen.

## Interview

Mündliche Befragungen haben gegenüber schriftlichen eine größere Offenheit und geben den Schülern mehr Möglichkeiten, ihre persönlichen Wahrnehmungen, ihr Handeln und ihre Beweggründe zu erläutern. Insofern erlauben sie oft einen tieferen Einblick in Lernprozesse und Bewertungen des Unterrichts als Fragebögen. Der Erhebungs- und Auswertungsaufwand ist jedoch wesentlich größer, sodass mit dieser Methode deutlich weniger Personen befragt werden können. Eine im Rahmen von Evaluationen häufige Form des Interviews ist das sog. *Leitfaden-Interview*, für das vorab ein Katalog von Fragen entwickelt wird, die allen Befragten in möglichst identischer Form gestellt werden. Diese Form des Interviews lenkt das Gespräch auf die für die Evaluation relevanten Aspekte. Da alle Befragten dieselben Fragen erhalten, können die Antworten im Anschluss gut verglichen werden. Bei der Formulierung von Interviewfragen sollte man darauf achten, die Fragen so zu formulieren, dass sie nicht mit einem simplen „Ja“ oder „Nein“ zu beantworten sind. Sie sollten die Schüler dazu anregen, ihre Antworten auszuführen und zu begründen, z.B. durch W-Fragen (Was? Wie? Warum?). Weiterhin sollten die Fragen möglichst neutral formuliert werden, d.h. keine Antwort in eine bestimmte Richtung nahelegen. Wie im Fragebogen ist auch im Interview darauf zu achten, die Befragten nur zu solchen Aspekten zu befragen, zu denen sie auch eine reflektierte Auskunft geben können.

Zur Aufzeichnung des Interviews sollte ein Aufnahmegerät verwendet werden, da der Interviewer sich dann ganz auf das Gespräch konzentrieren kann. Das Gespräch kann dann im Nachhinein nochmals abgehört und transkribiert werden. In wissenschaftlichen Studien werden in der Regel vollständige Transkripte angefertigt. Ein weniger aufwendiges Verfahren besteht in der Anfertigung zusammenfassender Notizen aufgrund der Tonaufnahme, bei denen nur besonders wichtige Aussagen vollständig transkribiert werden. Wenn eine Aufnahme nicht möglich ist, sollten während des Gesprächs Notizen gemacht werden, die direkt im Anschluss an das Interview aus dem Gedächtnis ergänzt werden. Aufgrund der selektiven Wahrnehmung und Erinnerung ist bei dieser Methode die Gefahr allerdings relativ groß, die Antworten unvollständig und verzerrt wiederzugeben.

Grundsätzlich ist bei schriftlichen und besonders bei mündlichen Befragungen zu bedenken, dass die Gefahr von Verzerrungen durch Falschaussagen besteht: Im Falle der sog. *Sozialen Erwünschtheit* (vgl. MOOSBRUGGER/KELAVA 2012: 57) möchten Schüler sich nicht nega-

## Grenzen von Befragungen

tiv zum Unterricht äußern, da sie Sanktionen seitens der Lehrkraft befürchten, oder sie sind nicht motiviert, sich an der Untersuchung zu beteiligen. Liegt die sog. *Akquieszenz* vor, antworten Probanden unkritisch und stimmen unabhängig vom konkreten Inhalt immer zu (vgl. MOOSBRUGGER/KELAVA <sup>2</sup>2012: 61). Es ist deshalb wichtig, den Schülern zu verdeutlichen, welchen Zweck die Evaluation verfolgt – dass nicht sie bewertet werden sollen, sondern die Lehrkraft ihren Unterricht verbessern und dafür ihre Rückmeldung nutzen möchte. Auch kann es hilfreich sein, die Erhebung anonym (beim Fragebogen) bzw. durch eine neutrale dritte Person, der die Schüler vertrauen (z.B. einen Kollegen oder eine externe Person), durchführen zu lassen. Schließlich können Schülerinterviews auch in Kleingruppen durchgeführt werden, wenn sich die Befragten in einer Gruppensituation wohler fühlen.

### 5.5.3 Tests und Arbeitsproben

Die Überprüfung des Lernerfolgs der Schüler anhand von (lernzielbasierten) Tests und Arbeitsproben gehört zum unterrichtlichen Alltagsgeschäft. Da solche Tests immer eine Rückmeldung zum Erfolg des Unterrichts geben, sollten sie auch in die Evaluation einbezogen werden. Durch eine gezielte Stufung der Schwierigkeit der Aufgaben eines Tests und durch Formulierung von Aufgaben zu verschiedenen Teilbereichen des behandelten Lerngegenstands können Lehrkräfte ein differenziertes Leistungsbild ihrer Schüler erhalten und auf diese Weise einen wertvollen Zusatzertrag für die unterrichtliche Evaluation erzielen.

Vergleichs-  
gruppe – Vor-  
und Nachtest

Im Rahmen einer Evaluation möchte man das Lernergebnis jedoch möglichst eindeutig auf die eingesetzten Unterrichtsmaterialien und -methoden zurückführen und wissen, ob diese wirksamer waren als der herkömmliche Unterricht. Gerade Letzteres kann eigentlich nur beantwortet werden, indem eine Vergleichsgruppe herangezogen wird, von der sich die untersuchte Lerngruppe nur darin unterscheidet, dass sie einen anderen Unterricht erhalten hat.

Im Schulalltag könnten als Vergleichsgruppe Schüler einer oder mehrerer Parallelklassen, bei denen die zu evaluierende Maßnahme nicht durchgeführt wird, beteiligt werden, sofern diese vom Leistungsniveau, den Vorkenntnissen und dem Lernverhalten mit der untersuchten Gruppe vergleichbar sind. Indem beiden Gruppen vor Beginn der eigentlichen Evaluation ein Vortest und im Anschluss ein Nachtest vorgelegt wird, kann überprüft werden, ob sich die Gruppen auf

demselben Lernniveau befanden und ob sich der eingetretene Lernfortschritt unterscheidet (vgl. BORTZ/DÖRING <sup>2</sup>1995: 331 ff.). Möglich ist es auch, die Ergebnisse früherer Klassen als Vergleichsgröße heranzuziehen. In diesem Fall wird ein Test, der einer früheren Klasse vorgelegt wurde, bei der untersuchten Lerngruppe wiederholt und hinsichtlich der Ergebnisse verglichen.

Für sich genommen sind die Ergebnisse von Tests in der Regel nicht sehr aussagekräftig. Sie sollten deshalb unbedingt mit Beobachtung und/oder Befragung kombiniert werden, da erst die Kombination der Untersuchungsmethoden Rückschlüsse erlaubt, wie das Ergebnis zustande gekommen ist, wo der Unterricht erfolgreich war und wo evtl. noch Verbesserungsbedarf besteht.

#### 5.5.4 Ethische Grundsätze und Datenschutz

Bei der Evaluation von Unterricht sind neben ethischen Grundsätzen zugleich datenschutzrechtliche Bestimmungen einzuhalten (vgl. BORTZ/DÖRING <sup>2</sup>1995: 40–45): Evaluationen können fragwürdig sein, sobald Methoden oder Fragen zur Anwendung kommen, die die Probanden in eine physisch oder psychisch unangenehme Lage versetzen. Zugleich ist stets abzuwägen, ob es ethisch verantwortbar ist, dass im Rahmen einer Intervention die Testgruppe durch unterrichtsergänzende Materialien Vorteile erzielt, die einer Kontrollgruppe ohne Intervention verwehrt werden. Außerdem muss die Teilnahme an einer Studie auf Freiwilligkeit beruhen. Daher sollten die Schüler und ihre Erziehungsberechtigten umfassend über Sinn und Zweck der Evaluation informiert werden. Belohnende Anreize sind gängige Methoden zur Schülermotivation.

Schließlich ist natürlich auch auf den Datenschutz zu achten. Jegliche Auskünfte über andere Personen unterliegen dem Datenschutz. Daher empfiehlt es sich, vor Erhebungen, in denen auch persönliche Angaben (Herkunftssprache, Geschlecht, Familienstand ...) erfragt werden, zunächst die Datenschutzbestimmungen der Bundesländer einzusehen (vgl. BORTZ/DÖRING <sup>2</sup>1995: 45). Dies gilt insbesondere, wenn die Untersuchung Externe einschließt oder die Ergebnisse auch außerhalb der schulischen Einrichtung genutzt werden sollen. In einem solchen Fall bedarf es der Zustimmung schulischer Gremien, der Erziehungsberechtigten (i. d. R. bei Schülern unter 14 Jahren) sowie ggf. auch der zuständigen Schulbehörde. Für die interne Evaluation des eigenen Unterrichts sind solche Maßnahmen nicht nötig, solange sie mit den übrigen ethischen Kriterien nicht kollidieren.

## 5.6 Praxisbeispiel: Evaluation einer Unterrichtsreihe zum Tempusgebrauch

### 5.6.1 Das *CIPP-Modell* als Grundlage

#### *CIPP-Modell*

An einem Berliner Gymnasium wurde in insgesamt drei Klassen aus der 8. und 9. Jahrgangsstufe eine zweiwöchige Unterrichtseinheit (vgl. Kapitel 4.3) zur Tempusbildung und -funktion im Lateinischen und Deutschen durchgeführt (vgl. GROSSE 2011). Insgesamt nahmen 49 Schüler an der Untersuchung teil, von denen 85,7% eine andere Muttersprache als Deutsch aufwiesen. Das Untersuchungsdesign berücksichtigt alle gängigen Methoden der Evaluationsforschung, die sich auch im *CIPP-Modell* wiederfinden lassen: Auf der Kontextebene wurden mittels eines Vortests mit unterschiedlichen Frageformaten das Vorwissen der Lernenden hinsichtlich Tempusbildung und -gebrauch sowie die subjektive Einschätzung des bisherigen Lateinunterrichts (speziell hinsichtlich des sprachreflektorischen Nutzens des Unterrichts) erhoben und durch persönliche Hintergrundmerkmale (Sprache im Elternhaus, Geschlecht) in Form eines Fragebogens ergänzt. Die Berücksichtigung des Input erfolgte durch die Erhebung der spezifischen Probleme der Lernenden mit dem eingeführten Lehrbuch sowie durch die theoretische Fundierung und Dokumentation der in der Unterrichtsreihe verwendeten Materialien und Methoden. Auf der Prozessebene wurde der Umgang der Schüler mit den neu entwickelten Unterrichtsmaterialien ausgewertet. Anhand von Schätzfragen auf einer Rating-Skala wurde schließlich auch die Zufriedenheit der Schüler mit der Unterrichtsreihe gemessen. Die Produktebene wurde nicht nur durch die Auswertung von Arbeitsproben aus dem Unterrichtsgeschehen (z.B. Vorträge, Plakate) einbezogen, sondern auch durch einen Gruppenabschlussstest am Ende der Einheit sowie durch einen Nachtest. Dieser wurde parallel zum Vortest konstruiert, um Erkenntnisse über einen Lernzuwachs gegenüber dem Vortest zu gewinnen. In diesem Zusammenhang wurde auch der von den Schülern wahrgenommene subjektive Nutzen der Unterrichtsreihe für die sprachreflektorische Kompetenz abgefragt. Das methodische Repertoire wurde durch die Beobachtungen neutraler Beobachter sowie durch ein Interview mit den unterrichtenden Lehrkräften abgerundet, um andere Wahrnehmungsperspektiven einzubeziehen.



### 5.6.2 Unterrichtsbeobachtung: Gestaltung und Erkenntnisgewinn

Die Beobachtung richtet sich vor allem auf den Unterrichtsprozess und orientiert sich an folgenden Leitfragen:

Leitfragen

- Welche Probleme tauchen bei der Bearbeitung der Arbeitsmaterialien in sprachlicher Hinsicht auf? – Wann und zu welchen Aspekten suchen die Schüler Hilfe bei der Lehrkraft?
- Lassen sich Verhaltensweisen erkennen, die darauf hindeuten, dass metasprachliche Lernprozesse angestoßen werden?
- Welche Aufgaben motivieren die Schüler besonders und bereiten ihnen Freude?

Alle drei Lerngruppen wurden durch externe Beobachter begleitet. Die Beobachter erstellten zu den o. g. Fragen unstrukturierte Beobachtungsprotokolle und zeichneten die für die Beobachtung relevanten Interaktionen des Unterrichtsprozesses auf. Für die abschließende Bewertung wurden die Ergebnisse aller Beobachter über die gesamte Unterrichtseinheit hinweg zusammengefasst und typische Interaktionsmuster herausgearbeitet, aus denen eine Interpretation der Beobachtungen abgeleitet wurde.

Diese Form der Beobachtung ermöglichte grundlegende Erkenntnisse über das Unterrichtsgeschehen: So war zum Beispiel zu beobachten, dass die Lernenden die Arbeit mit einer vorgegebenen Übersetzung zunächst als wenig anspruchsvoll empfanden. Obgleich sie dieses Aufgabenformat aber nicht kannten, waren sie sehr motiviert und fanden einen leichten Zugang zum kontrastiven Sprachvergleich. Im Verlauf der Unterrichtsreihe konnte beim sprachvergleichenden Arbeiten dann allerdings gleichzeitig rasch erkannt werden, dass zahlreiche Schüler Schwierigkeiten bei der Zuordnung der lateinischen zu den deutschen Prädikaten hatten – eine wichtige Erkenntnis für den Einsatz dieses Aufgabentypus. Denn offenbar ist das kontrastive Arbeiten gerade durch die Fokussierung auf die grammatischen Strukturen, die die reine Übersetzungsarbeit ergänzen, doch in hohem Maße anspruchsvoll, da es über eine rein inhaltliche Entschlüsselung des Textes hinausgeht. Damit einhergehend war zugleich zu beobachten, dass die Bestimmung der Tempusfunktionen nach anfänglichen Verständnisschwierigkeiten auf reges Interesse bei den Schülern stieß, also keinesfalls unterfordernd war, sondern vielmehr eine neue Perspektive auf die Arbeit mit lateinischen Texten eröffnete. In detektivischer Manier wurde über die Bedeutung der Prädikate und die Funktion des Tempus nachgedacht und nach einer angemessenen Übersetzung gesucht. Zugleich zeigte sich eine zunehmend sichere

Erkenntnis-  
gewinn

Verwendung der metasprachlichen Terminologie im Unterrichtsgespräch.

### 5.6.3 Schülerbefragung: Gestaltung und Erkenntnisgewinn

#### Leitfragen

Bei der Gestaltung schriftlich gestützter Erhebungsformate standen folgende Leitfragen im Zentrum des Erkenntnisinteresses:

- Wie schätzen die Schüler ihre sprachlichen Fähigkeiten im Lateinischen und im Deutschen im Allgemeinen sowie hinsichtlich des Tempusgebrauchs ein?
- Wie bewerten sie ihren bisherigen Lateinunterricht (Vortest) bzw. die Unterrichtsreihe (Nachttest), speziell die Aspekte der Sprachreflexion?
- Wie bewerten sie das Lehrbuch (Vortest) bzw. die Arbeitsblätter und Materialien (Nachttest) allgemein (Gestaltung, Verständlichkeit) sowie speziell die zum Tempusgebrauch?

#### Tests

Die Evaluation dieser Fragen wurde durch einen Vor- und Nachttest durchgeführt. Beide Tests sollten den jeweiligen Status quo vor und nach der Unterrichtseinheit erfassen und Auskünfte über den individuellen Lernzuwachs geben. Diese Tests bedienten sich unterschiedlicher, standardisierter Frageformate (offene Fragen, Multiple Choice, Schätzfragen mittels Rating-Skalen). Beide Tests waren anonym zu beantworten. Ein Codierungssystem ermöglichte eine Zuordnung der Vor- und Nachttests, sodass die Lernprogression jedes einzelnen Schülers nachvollzogen werden konnte. Die Tests gliederten sich in drei Teile, die parallel zueinander konstruiert wurden:

- Fragen zur Person und den Erfahrungen im Lateinunterricht (Vortest) bzw. während der Unterrichtseinheit (Nachttest)
- Fragen zum Lehrbuch (Vortest) bzw. den selbstentwickelten Unterrichtsmaterialien (Nachttest)
- Wissensfragen zum Thema der Unterrichtseinheit (in beiden Tests nahezu identisch)

#### Trichtertechnik

Die Anordnung der Fragen erfolgte nach der „Trichtertechnik“ (INGENKAMP 1985: 77), d.h., dass mit allgemeineren Fragen begonnen und zunehmend zu spezielleren fortgeschritten wurde, um Antworthemmungen zu vermeiden. Ergänzt wurden diese Tests durch Fragebögen zur Person und subjektiven Einschätzung des Lateinunterrichts sowie der eigenen sprachlichen Kompetenzen in der Zweitsprache. Allgemein wurde der Großteil der Fragen in beiden Befragungen in

gleicher oder ähnlicher Form gestellt, um die Antworten aus Vor- und Nachbefragung direkt vergleichen zu können.

Bei der Formulierung der Fragen wurde darauf geachtet, ein Format zu wählen, das nicht nur dem sprachlichen Niveau der Schüler entspricht, sondern das die Schüler persönlich anspricht (Bsp.: *Heute darfst du einmal bewerten. Welche Note würdest du deinem Schulbuch geben?*). Die Formulierung der Schätzfragen, die via Rating-Skala beantwortet werden sollen, sowie die offenen Fragen sollten durch eine Formulierung aus Perspektive der Schüler gestellt werden (Bsp.: *Mir fällt das Erkennen lateinischer Tempora leicht.*). Bei den offenen Fragen erwies es sich als günstig, eine gewisse Anzahl von mindestens zu leistenden Antworten sowie einen Satzanfang zur Beantwortung vorzugeben, um möglichst aussagekräftige Ergebnisse zu erhalten. Um die Lernfortschritte vom Vor- zum Nachtest direkt zu überprüfen und die Vergleichbarkeit der erhobenen Daten zu gewährleisten, wurden in den Nachtest auch die Aufgaben des Vortests integriert.

**Teil 1: Allgemeine Fragen zu dir und den vergangenen Lateinstunden**

1.1 Ich bin ein ...  
 Junge                       Mädchen

1.2 Ich bin in der \_\_\_\_\_ Klasse. (bitte nur die Klassenstufe angeben)

1.3 In der Schule lerne ich Latein seit ...  
 der 7. Klasse             der 8. Klasse

1.4 Zu Hause spreche ich überwiegend ... (bitte nur ein Kreuz machen)  
 Deutsch    Deutsch und eine weitere Sprache    gar kein Deutsch

1.5 In meinem Freundeskreis spreche ich überwiegend ...  
 (bitte nur ein Kreuz machen)  
 Deutsch    Deutsch und eine weitere Sprache    gar kein Deutsch

Nenne bei den folgenden offenen Fragen stichwortartig jeweils mindestens zwei Dinge, die dir einfallen. Du kannst aber auch noch mehr schreiben.

1.6 Was hat dir in den letzten Lateinstunden gefallen? (Mir gefiel ...)

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

1.7 Was hat dir in den letzten Lateinstunden nicht gefallen?  
 (Mir gefiel ... nicht)

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

1.8 Welche positiven Effekte hatten die letzten Lateinstunden für dich?  
 (Ich kann besser ...)

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

1.9 Was ist dir in den letzten Lateinstunden besonders schwer gefallen?  
 (Mir fiel ... schwer)

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Abb. 4: Fragen zur Person und subjektiven Einschätzung des Lateinunterrichts

### Teil 3: Wissensfragen rund um das Thema Tempusgebrauch

3.1 Man kann das lateinische mit dem deutschen Tempussystem vergleichen. Wie kann man die folgenden Tempora des Lateinischen im Deutschen wiedergeben? (Es sind auch mehrere Antworten möglich).

Imperfekt:

Präsens     Präteritum     Perfekt

Perfekt:

Präsens     Präteritum     Perfekt

3.2 In welchem Tempus würdest du einen Aufsatz über die Geschehnisse im Trojanischen Krieg schreiben?

Präsens     Präteritum     Perfekt

3.3 In welchem Tempus würde wohl ein Römer im alten Rom einen Aufsatz über die Geschehnisse im Trojanischen Krieg schreiben?

Präsens     Präteritum     Perfekt

3.4 Ein Schüler übersetzt einen lateinischen Text wie folgt:

<i>Tandem Troiani Graecos superaverunt. Laeti equum in urbem portaverunt. Postremo dormitaverunt et Graeci ex equo ierunt.</i>	Endlich besiegten die Trojaner die Griechen. Froh zogen sie das Pferd in die Stadt.
--	---

Wie du siehst, ist der letzte Satz nicht übersetzt. Wie müsste dieser Satz lauten?

- Schließlich schlafen sie ein und die Griechen steigen aus dem Pferd.
- Schließlich schliefen sie ein und die Griechen stiegen aus dem Pferd.
- Schließlich sind sie eingeschlafen und die Griechen sind aus dem Pferd gestiegen.

Abb. 5: Fragen zum Wissen der Schüler über das Tempussystem

Diese Form des Testdesigns („Pre-Posttest-Design“) lieferte den entscheidenden Erkenntnisgewinn über die Zielrichtung und Wirksamkeit der unterrichtlichen Maßnahmen: Zunächst offenbarte die erfragte Selbsteinschätzung der Schüler einen deutlichen Widerspruch zwischen Realität und Wirklichkeit, den es dann durch einen entsprechend gestalteten Unterricht zu vermindern galt. Zahlreiche Probanden konnten nämlich bei sich selbst keine Schwierigkeiten mit der deutschen Grammatik ausmachen. Allerdings wiesen dann die Antworten

Erkenntnis-  
gewinn

auf die Frage, welche Tempora es im Deutschen gebe, in eine ganz andere Richtung, wenn z.B. Phantasiebezeichnungen wie einfache Vergangenheit, Perfekt I und II angeführt wurden oder gänzlich andere grammatische Lemmata und Kategorien zu den Tempora gezählt wurden (Akkusativ, Singular, Infinitiv) (vgl. GROSSE 2011: 62). Wie sich dann im Nachtest zeigte, wurden diese Selbsteinschätzungen im Verlauf des Unterrichtsgeschehens erheblich realitätsbezogener. Die Erkenntnis der individuellen Defizite ist ein wichtiger Schritt hin zur Bereitschaft, an den eigenen sprachlichen Kompetenzen zu arbeiten. Gerade in der Alltagssprache werden solche Defizite durch Vermeidungsstrategien zumeist kompensiert und von den DaZ-Lernenden nicht wahrgenommen. Allein durch die Frage nach der subjektiven Selbsteinschätzung im Rahmen einer solchen Befragung können sich einerseits die Schüler bewusster über ihr eigenes sprachliches Vermögen werden. Andererseits kann die Lehrkraft ein aussagekräftiges Ergebnis erhalten, ob die Schüler durch eine Intervention zu einem stärkeren Bewusstsein gefunden haben (vgl. Kapitel 2.3.1).

Dass die Schüler insgesamt deutliche Fortschritte in der Zweitsprache erreichen konnten, wurde durch den Vergleich von Vor- und Nachtest deutlich: Der Fehleranteil im Deutschen sank um 20 bis 30%! Der Nachtest bestätigte ebenfalls eine grundsätzliche Tendenz, die auch schon durch andere Untersuchungsmethoden (Beobachtungen, Arbeitsproben) aufgezeigt worden war: So zeigten die Schüler weiterhin größere Unsicherheiten im Umgang mit dem deutschen Tempussystem als mit dem lateinischen – eine Erkenntnis, die für die weitere Gestaltung des Unterrichts von zentraler Bedeutung ist. Zugleich wird hieran nachdrücklich deutlich, dass erst die Kombination verschiedener Untersuchungsmethoden die Ergebnisse valider macht.

#### 5.6.4 Auswertung von Arbeitsproben

##### Leistungstest

Will man sich als Lehrkraft nicht unbedingt der Methoden der Unterrichtsforschung bedienen, stellt auch die Auswertung von Arbeitsproben ein aussagekräftiges Werkzeug dar. Auch die hier beschriebene Unterrichtseinheit wurde ergänzt durch einen Leistungstest (Gruppenabschlussstest), wie er von allen Lehrern in Form von Grammatiktests oder Klassenarbeiten als Mittel zur Sicherung und Beurteilungsgrundlage genutzt wird. Mit dem Abschlussstest wurden die Lernergebnisse auf der Produktebene evaluiert. Der Test ergänzte damit die im vorigen Abschnitt dargestellten Selbsteinschätzungen der Schüler

und die objektiven Ergebnisse aus den Vor- bzw. Nachtests. Die Testaufgaben zielten auf folgende Aspekte ab: Kenntnis deutscher und lateinischer Tempora, Tempusfunktionen und Kenntnis der Unterschiede zwischen den Erzähltempora der lateinischen und deutschen Sprache.

Die Auswertung der Arbeitsproben erwies sich als sehr ertragreich: Ganz deutlich zeigte sich, dass die Schwierigkeiten der Schüler eben nicht im Bereich des Lateinischen zu konstatieren waren. Vielmehr offenbarten sich z.T. deutliche Schwächen bei der korrekten Formulierung und Bestimmung deutscher Prädikate. Zugleich ließ sich anhand der von den Schülern erstellten Plakate gut erkennen, dass die Anwendung von Metasprache in einem entsprechend höheren sprachlichen Register zunehmend verinnerlicht und korrekt angewendet wurde.