

CSP **Center scientific da cumpetenza per la plurilinguitad** Cogniziun Società Formation Bildung Migration Furmaziun Gesellschaft
CSP **Centro scientifico di competenza per il plurilinguismo** Scuola Arbeit Politique Communitad School Travail Ecole Community
CSP **Centre scientifique de compétence sur le plurilinguisme** Migrazione Furmaziun Societad Cognition Society scola Migration
KFM **Wissenschaftliches Kompetenzzentrum für Mehrsprachigkeit** Societé Cognizione Migraziun Schule Communauté Kognition
RCM **Research Centre on Multilingualism** Formazione Lavoro Politics Comunità Work Politik Lavur Politica Formation Gemeinschaft

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

Peter Lenz, Katharina Karges, Anna Kull,
Evelyne Berger, Daniela Kappler, Thomas Roderer

2019

Bericht des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit

Herausgeber | Publié par

Institut für Mehrsprachigkeit
www.institut-mehrsprachigkeit.ch

—
Institut de plurilinguisme
www.institut-plurilinguisme.ch

AutorInnen | Auteurs

Peter Lenz, Katharina Karges, Anna Kull, Evelyne Berger (Kompetenzzentrum für Mehrsprachigkeit), Daniela Kappler (SUPSI), Thomas Roderer (Pädagogische Hochschule St. Gallen)

Bericht zum Forschungsprojekt „Erfassung von Entwicklungsprofilen im frühen schulischen Fremdsprachenlernen“; herausgegeben von Peter Lenz und Katharina Karges, Wissenschaftliches Kompetenzzentrum für Mehrsprachigkeit; Projektleitung: Peter Lenz; für das Projekt verantwortlicher Professor: Thomas Studer

Übersetzung | Traduction

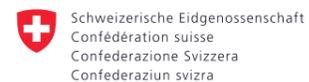
Isabelle Affolter, Mary Carozza, Mathias Piconi, pro-verbial, Priscilla Seamon

Das vorliegende Projekt wurde im Rahmen des Arbeitsprogramms 2012-2015 des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit durchgeführt. Für den Inhalt dieser Veröffentlichung sind die AutorInnen verantwortlich.

Le projet dont il est question a été réalisé dans le cadre du programme de travail 2012-2015 du Centre scientifique de compétence sur le plurilinguisme. La responsabilité du contenu de la présente publication incombe à ses auteurs.

Freiburg | Fribourg, 2019**Layout**

Billy Ben, Graphic Design Studio

Unterstützt von | avec le soutien de

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Département fédéral de l'intérieur DFI
Dipartimento federale dell'interno DFI
Departament federal da l'intern DFI
Bundesamt für Kultur BAK
Office fédéral de la culture OFC
Ufficio federale della cultura UFC
Uffizi federal da cultura UFC

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

—
Bericht zum Projekt „Erfassung von Entwicklungsprofilen im frühen schulischen Fremdsprachenlernen“
—

**Peter Lenz, Katharina Karges, Anna Kull,
Evelyne Berger, Daniela Kappler, Thomas Roderer**

2019

Bericht des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit
Rapport du Centre scientifique de compétence sur le plurilinguisme
Rapporto del Centro scientifico di competenza per il plurilinguismo
Report of the Research Centre on Multilingualism

Index

DE	Zusammenfassung	7
FR	Résumé	19
IT	Sintesi	31
EN	Summary	41
I	Einleitung Peter Lenz	52
II	Computerbasierte Beurteilung mündlicher Kompetenzen Katharina Karges	57
III	Kompetenzen im Bereich des Sprechens Katharina Karges, Evelyne Berger, Daniela Kappler	69
IV	Langage de classe Evelyne Berger, Daniela Kappler, Katharina Karges	89
V	Sprachlernstrategische Kompetenzen Thomas Roderer, Anna Kull, Peter Lenz	103
VI	Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare Daniela Kappler	129

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

Bericht zum Projekt „Erfassung von Entwicklungsprofilen
im frühen schulischen Fremdsprachenlernen“

Zusammenfassung

Das Projekt „Erfassung von Entwicklungsprofilen im frühen schulischen Fremdsprachenlernen“ ist Teil des Forschungsprogramms 2012-2015 des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit (KFM). Das Ziel bestand darin, durch die Entwicklung von geeigneten Instrumenten zur Kompetenzerfassung einen Beitrag zu einer lernorientierten und gleichzeitig computerunterstützten Beurteilungskultur im Fremdsprachenunterricht an Primarschulen (dritte bis sechste Klasse) zu leisten. Die Entwicklungsarbeiten orientierten sich an der Idee von Lernprogressionen (*learning progressions*), mit der vor allem in der Didaktik der Naturwissenschaften als Grundlage für Curricula, Unterricht und Beurteilung gearbeitet wird. Eine Lernprogression beschreibt in einem bestimmten Wissens- und Könnensbereich den typischen Entwicklungsweg vom Novizen zum Experten (z.B. in Bezug auf die Herausbildung eines angemessenen Materieverständnisses). *Learning progressions* sind nicht als natürliche Erwerbssequenzen zu verstehen, sondern als „typisches“ Ergebnis des Zusammenspiels von individuellen Voraussetzungen, zu erwerbendem Gegenstand und einem bestimmten Unterricht. Entwicklung und Einsatz von Aufgaben zur Erfassung einer Lernprogression und die Beschreibung der Lernprogression bedingen sich gegenseitig. Wenn der Verlauf einer Lernprogression genügend bekannt ist, lassen sich die Leistungen von Lernenden auf dieser Lernprogression idealerweise wie auf einer Strassenkarte verorten, sodass sich daraus optimale formative Feedbacks und „Feedforwards“ ableiten lassen.

Im Bereich des Sprachunterrichts und des Sprachenlernens sind bisher kaum *learning progressions* entwickelt worden. Das Projekt „Dynamic Language Learning Progressions“ (DLLP),¹ das im Kontext der Schulsprache Englisch situiert ist, bildet eine seltene Ausnahme. Es befasst sich mit der Entwicklung von sprachfunktionalen Kompetenzen, insbesondere der Fähigkeit, im Zusammenhang mit schulischen Aufgaben Erklärungen zu geben.

Im Projekt „Erfassung von Entwicklungsprofilen“ hat die Entwicklung exemplarischer Aufga-

ben gegenüber einer abgerundeten Beschreibung von Progressionen Priorität. Ein wichtiger Grund dafür ist das ambitionierte Ziel, über ein Internetserver-basiertes System Daten zum Leistungsstand der Schülerinnen und Schüler zu erheben und innerhalb des Systems ein Feedback zu generieren, welches für das Weiterlernen und -unterrichten von Nutzen ist. Um in technischer Hinsicht während der Projektdauer möglichst weit zu kommen, musste dieser Aspekt von Anfang an vorangetrieben werden. Ein Online-Beurteilungssystem ist u.a. deshalb so attraktiv, weil es auch dann, wenn es bereits im Gebrauch ist, aufgrund der einkommenden Daten weiterentwickelt werden kann, um die Aufgabenqualität zu verbessern und die Erkenntnisse über die Kompetenzen der Lernenden zu vertiefen.

Bei den inhaltlichen Bereichen, die im Projekt erkundet wurden, handelt es sich um die mündliche Produktion und Interaktion sowie die Sprachlernstrategien. Diese Wahl war bewusst anspruchsvoll, indem Kompetenzbereiche aufgegriffen werden, die sonst, vor allem im Hinblick auf das Beurteilen, eher vernachlässigt werden.

Für die konkrete technische Umsetzung des Projekts war mitentscheidend, dass die Erziehungsdirektorenkonferenz (EDK) im Jahr 2012 den Beschluss fasste, eine Aufgabendatenbank aufzubauen, die als Umgebung für Leistungsmessungen der Kantone, aber auch für andere Beurteilungs- und Evaluationsprojekte aus dem schulischen Bereich zur Verfügung stehen sollte. Über das Projekt Aufgabendatenbank kam eine Zusammenarbeit zwischen dem KFM und dem Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF) zustande, woraus eine Erweiterung der Testsoftware *CBA Item Builder* für die Erhebung und Beurteilung von mündlichen Leistungen hervorging.

Anspruch und Herausforderungen des Projekts „Entwicklungsprofile“ waren beträchtlich. Die fünf Beiträge des vorliegenden Berichts geben einen Überblick über die verschiedenen Arbeiten, die im Rahmen des Projekts ausgeführt wurden.

- Teil II „Computerbasierte Beurteilung mündlicher Kompetenzen“ befasst sich in erster Linie mit der technischen Umsetzung eines serverbasierten Systems, das für die forma-

tiv-diagnostische Beurteilung besonders auch der mündlichen Sprachkompetenzen eingesetzt werden kann. Im Mittelpunkt stehen die Entwicklung benutzerfreundlicher Aufgaben, die über einen Internetbrowser zur Verfügung gestellt werden können, sowie eines Moduls zur kriterienbasierten Beurteilung zuvor aufgenommener mündlicher Leistungen durch Lehrpersonen.

- Teil III „Kompetenzen im Bereich des Sprechens“ berichtet über die etappenweise Entwicklung und Validierung von Aufgaben, mittels derer diagnostische Informationen zu relevanten Teilbereichen der mündlichen Sprachkompetenz der untersuchten Primarschülerinnen und -schüler gewonnen werden können.
- Teil IV „Sprache des Unterrichts“ behandelt ein Teilprojekt, dessen Ziel es war, auf der Basis von Lehrwerkanalysen und einer Lehrpersonenbefragung sprachliche Mittel zu identifizieren, die für die mündliche Kommunikation im Unterricht (*classroom language*) wichtig sind. Mittels der Lehrpersonenbefragung wurde auch untersucht, wie gut Schülerinnen und Schüler solche sprachlichen Mittel auf den verschiedenen Klassenstufen rezeptiv und produktiv beherrschen.
- Teil V „Sprachlernstrategische Kompetenzen“ stellt die Entwicklung von Instrumenten zur Erfassung einerseits der Gebrauchshäufigkeit von Sprachlernstrategien durch die Lernenden und andererseits der Kompetenzen der Lernenden in ausgewählten Teilbereichen sprachlernstrategischer Kompetenz dar. Als Grundlage dafür wurde zunächst ein strukturiertes Inventar sprachlernstrategischer Kompetenzen erstellt.
- Teil VI „Beurteilungs- und Feedbackpraxis im Fremdsprachenunterricht in der Primar-

schule“ berichtet von den Ergebnissen einer qualitativen Untersuchung im Kanton Tessin zur Beurteilungspraxis von Primarlehrpersonen im Fach Französisch sowie zu deren potentieller Vereinbarkeit mit der Einführung einer serverbasierten Umgebung zur formativ-diagnostischen Beurteilung.

Im Rahmen des Projekts konnte im Hinblick auf ein serverbasiertes formativ-diagnostisches Beurteilungssystem eine Reihe von wichtigen Arbeiten geleistet werden. Einzelne der erarbeiteten Elemente konnten bereits auch in anderem Zusammenhang genutzt werden, insbesondere bei der Entwicklung von computerbasierten Testaufgaben.

Computerbasierte Beurteilung mündlicher Kompetenzen

Ein Teil der Projektressourcen wurde für die Erkundung und teilweise Weiterentwicklung eines Internetserver-basierten Testsystems eingesetzt, welches die Umsetzung eines breit einsetzbaren, serverbasierten formativ-diagnostischen Beurteilungssystems ermöglichen könnte. Als die EDK im Jahr 2012 die Entwicklung der sog. *Aufgabendatenbank* bekanntgab, konzentrierte sich das Projekt auf die Entwicklung eines Beurteilungssystems, das sich informationstechnisch in den Rahmen dieser Testplattform einfügt. Nach einer Erkundungsphase mit Microsoft PowerPoint und Standardsoftware zur Tonaufnahme wurden alle im Projekt entwickelten Aufgaben in der Software *CBA Item Builder* umgesetzt und durch die dazugehörige Serverapplikation EE4CBA ausgeliefert. Beides wurde vom DIPF zur Verfügung gestellt, mit dem die EDK im Hinblick auf die Aufgabendatenbank zusammenarbeitete. Diese Softwarelösung ist so konzipiert, dass schulseitig lediglich Computer mit Internetanschlüssen und aktuellen Internetbrowsern sowie Headsets benötigt werden.

Um die vorhandene Software für die Erfassung auch von mündlichen Leistungen zu erweitern, wurden über das DIPF entsprechende Ent-

1 | <https://www.dllp.org/>

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

wicklungsarbeiten in Auftrag gegeben. Einerseits wurde ein Modul zur Tonaufnahme über den Webbrowser ergänzt, andererseits ein Ratingmodul. Im Ratingmodul können vorhandene Aufnahmen (oder auch geschriebene Texte) verwaltet und anhand von Kriterien manuell bewertet werden. Die Bewertungen werden anschliessend in gleicher Weise abgespeichert wie die Antworten bei Items mit geschlossenen Formaten (z.B. Multiple-Choice).

Nicht alle informationstechnischen Leistungen, die für ein funktionsfähiges formativ-diagnostisches Beurteilungssystem benötigt würden, konnten erbracht werden. Insbesondere fehlen ein voll integriertes Modul zur Verwaltung von Personen und Aufgaben sowie ein automatisches Auswertungsmodul, das Schülerantworten und manuelle Ratings zu Skalen in Beziehung setzt und darauf basierende Berichte bzw. Feedbacks und Feedforwards (Hinweise auf nächste Entwicklungsschritte) generieren kann.

Die Hauptarbeit im Projekt bestand in der Entwicklung und Erprobung von Aufgabenformaten und -inhalten, die für die Zielgruppe junger Schülerinnen und Schüler geeignet sind und von diesen insbesondere auch selbstständig gelöst werden können.

Um zu den benötigten Informationen zu gelangen, wurde in verschiedenen Erprobungsphasen u.a. folgenden Fragestellungen nachgegangen: Vertrautheit v.a. der jüngsten Schülerinnen und Schüler mit Computer-Hard- und -Software, Eignung von unterschiedlichen Benutzeroberflächen, Bedienungselementen (Symbolen) und Inputs (Texten, Bildern), Verständlichkeit von Aufgabenstellungen (sowohl in der Schul- als auch in der Zielsprache), technische und logistische Einsetzbarkeit der Aufgaben im Klassenverband, Schwierigkeit und Angemessenheit der Aufgabenformate und -inhalte. Eine wichtige Erkenntnis besteht darin, dass der Umgang mit dem Computer offenbar auch bei den jüngsten Teilnehmenden keine Hürde darstellt. Zudem ergaben schon frühe Erprobungsphasen, dass auch nur leicht textlastige Bildschirmoberflächen sowie mehrschrittige bzw. kognitiv etwas anspruchsvollere Aufgabenformate zu erheblicher Verwirrung führen und deshalb für valides Beurteilen wenig geeignet sind. Im Laufe der Zeit wurden daher Aufgabenformate entwickelt,

deren Bedienung ein Minimum an Erklärungen erfordert und deren korrekte Lösung nicht vom Verstehen eines geschriebenen Textes abhängt (da ja mündliche Kompetenzen erfasst werden sollen). Die verschiedenen Erkenntnisse zu computerbasierten Beurteilungsinstrumenten, die für Primarschülerinnen und -schüler geeignet sind, sowie konkrete Aufgabenformate mit den darin verwendeten Bildern und Symbolen stehen für weitere Aufgaben- und Testentwicklungen zur Verfügung.

Kompetenzen im Bereich des Sprechens

Im Teilprojekt „Kompetenzen im Bereich des Sprechens“ steht die Entwicklung von computerbasierten Instrumenten (Aufgaben) zur Erfassung der noch elementaren Kompetenzen von Primarschülerinnen und -schülern im Bereich der mündlichen Produktion und Interaktion im Vordergrund. In Dokumenten aus dem Umkreis des Gemeinsamen europäischen Referenzrahmens für Sprachen (GER) sowie solchen aus dem Gebiet des Sprachentestens sind nur relativ grobe Anhaltspunkte zur Kompetenzausprägung und -entwicklung innerhalb der A-Niveaus des GER zu finden. Als charakteristisches Merkmal erscheint häufig die Abhängigkeit von wenigen gelernten Wörtern und festen Wendungen (*chunks*), welche mehr oder weniger flexibel eingesetzt werden können, um in spezifischen, vertrauten Kontexten elementare Kommunikationsaufgaben zu lösen. Im Fall von Primarschülerinnen und -schülern bedeutet dies eine starke Abhängigkeit von dem, was in Lehrwerk und Unterricht behandelt wird. Erste Erprobungen von Aufgaben in den drei Kontexten (in den Lehrplanregionen Passepartout,² Ostschweiz und Tessin) zeigten diese Abhängigkeit deutlich. Der Anspruch, kontextübergreifend relevante Entwicklungslinien zu erfassen, erwies sich schliesslich als zu grosse Einschränkung bei der Aufgabenentwicklung, sodass sich das Teilprojekt zum Spre-

2 | Die sechs Kantone entlang der deutsch-französischen Sprachgrenze: Bern, Basel-Landschaft, Basel-Stadt, Freiburg, Solothurn und Wallis.

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

chen in der abschliessenden Phase auf die Lehrplanregion Passepartout konzentrierte.

Das Teilprojekt näherte sich in insgesamt drei grösseren Etappen dem Ziel an, die mündlichen Kompetenzen von Primarschülerinnen und -schülern in der ersten Schulfremdsprache zu erfassen und zu diagnostizieren.

In der ersten Etappe lehnten sich die Aufgaben zur Erfassung der vorhandenen Kompetenzen an Inhalte und Typen von (Kommunikations-)Aufgaben an, die in den verschiedenen Lehrwerken unseres Praxisfeldes häufig vorkamen. Die noch papierbasierten Aufgaben wurden vor Ort erprobt und die Erprobungen anschliessend detailliert ausgewertet, im Hinblick einerseits auf das Funktionieren der Aufgaben und andererseits auf die Kompetenzen der jungen Lernenden. Diese von den Lehrwerken inspirierten Aufgaben waren meistens so offen gefasst, dass damit ein weites Kompetenzspektrum erfasst werden könnte.

Die Erprobungen zeigten aber, dass mit solchen Aufgaben die eben erst entstehenden Sprechkompetenzen vieler Primarschülerinnen und -schüler nicht bzw. nicht effizient erfasst werden können, da diese kaum in der Lage waren, die Aufgaben ohne Hilfe von Dritten auch nur ansatzweise zu bewältigen. Dies lag einerseits an den kognitiven Anforderungen durch die Situierung der Aufgaben in Kommunikationssituationen, andererseits an den Ansprüchen an die kommunikative Sprechkompetenz in der Fremdsprache, obschon diese zum Teil sehr gering waren.

In der zweiten Arbeitsphase wurden die sprachlichen und sprachfunktionalen *Ressourcen* eines (späteren) kommunikativen Sprachgebrauchs stärker in den Mittelpunkt gerückt als das sprachliche Handeln im Rahmen von Kommunikationsaufgaben, obschon auch solche Aufgaben, in einfachen Formaten, weiterhin eingesetzt wurden. Den Ausgangspunkt für die Bildung der neu im Mittelpunkt stehenden Konstrukte bildete eine detaillierte Analyse der sprachlich-kommunikativen Komponenten (die sog. „sprachlichen Mittel“) der Aufgaben, mit denen in den Lehrwerken Fertigkeiten im Bereich Sprechen gefördert bzw. überprüft werden.

In den Erprobungen der zweiten Phase kam eine Auswahl von Aufgaben zum Einsatz, die ent-

weder auf linguistische Ressourcen, sprachfunktionale oder kommunikative Kompetenzen zielten. Erstmals arbeiteten die Kinder an Computern. Die Erfahrungen zeigten, dass auch junge Kinder über genügend Bereitschaft und Kenntnisse verfügen, um die computerbasierten Aufgaben zu bearbeiten. Als heikle Punkte entpuppten sich schriftliche Aufgabenstellungen ab einer bestimmten Länge, Beispiele am Bildschirm sowie ganz besonders die Verwendung der Zielsprache für Aufgabenstellungen. In Bezug auf die Kompetenzen ergaben sich aus dieser Erprobungsphase insbesondere die folgenden Erkenntnisse: Ausspracheprobleme sind sehr marginal bedeutsam, während das laute Vorlesen geschriebener Sprache die jungen Lernenden durchaus vor Herausforderungen stellen kann. Die ganz konkrete Kenntnis von Wörtern und Wendungen scheint zentral zu sein, zentral vor allem auch für das Bewusstsein der Lernenden selbst. So zeigte sich im Zusammenhang mit Aufgaben zur Syntax, dass Kinder zum Teil gar keinen Lösungsversuch unternahmen, wenn sie das Vokabular der Sätze nicht genau verstanden. Mögliche syntaktische Kenntnisse blieben deshalb verborgen. Etwas umfassendere Aufgaben, in denen mündlich kommuniziert werden sollte, erwiesen sich wiederum als problematisch: Leicht anspruchsvollere Aufgabenstellungen wurden auch in der Schulsprache nicht zuverlässig zur Kenntnis genommen und in der Folge nur teilweise beachtet. Die meisten der selbstständigen mündlichen Produktionen in der Zielsprache fielen sehr bescheiden aus.

In der letzten Projektetappe wurde nur noch mit der Lehrplanregion Passepartout und damit Schülerinnen und Schülern gearbeitet, die mit dem Lehrwerk *Mille feuilles* Französisch lernten. Erstmals waren nun aber auch Fünft- und Sechstklässler dabei. Das Spektrum der fokussierten Bereiche wurde weiter verengt. Bei den verbleibenden Teilbereichen handelte es sich einerseits um zwei kommunikative Handlungsbereiche – um die Kommunikation im Unterricht (*classroom language*) und um den Austausch persönlicher Informationen – sowie um eine sprachliche Teilkompetenz, nämlich die Wortschatzkompetenz. Als neues Teilkonstrukt der Sprechkompetenz wurde das Verstehen von gesprochener Sprache hinzugenommen, dies in der

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

Erwartung, dadurch elementare Sprachkompetenzen im mündlichen Bereich differenzierter sichtbar machen zu können. Die Aufgaben zu den genannten Bereichen wurden in Struktur, Anspruch und Erscheinungsbild weiter an die Zielgruppe angepasst. Als Grundlage für die aktuelle und ggf. zukünftige Aufgabenentwicklung zu *Mille feuilles* wurde ein Korpus aus allen geschriebenen und gesprochenen Texten des Lehrwerks erstellt. Weiter wurde ein Katalog mit sprachlichen Mitteln zur Kommunikation im Unterricht zusammengestellt. Der Katalog basiert einerseits auf einer Analyse verschiedener Lehrwerke und andererseits auf einer ergänzenden grösseren Umfrage bei Lehrpersonen. Korpus und Katalog wurden genutzt, um eine Reihe von Modellaufgaben zu wichtigen Bereichen des rezeptiven, interaktiven und produktiven mündlichen Sprachgebrauchs zu entwickeln. Diese Aufgaben sind kurz und finden auf einer Bildschirmseite Platz. Um die Aufgaben möglichst intuitiv verständlich zu gestalten, wurden oft Illustrationen eingesetzt. Als einziges etwas komplexeres Format wurde für die Überprüfung der mündlichen Interaktionsfähigkeit der bereits früher erprobte „Dialog“ beibehalten. Ein solcher Dialog erstreckt sich über mehrere Bildschirmseiten und besteht aus abwechselndem Hören und Sprechen mit einer fiktiven Person.

Die Praxiserprobung erfolgte in mehreren Schritten. Dabei wurden nicht nur die Aufgaben, sondern auch die serverbasierte Auslieferung von Tests für ganze Klassen auf der schuleigenen Infrastruktur erprobt. Zudem wurde anhand der Aufnahmen aus dieser Phase die (manuelle) analytische Bewertung der mündlichen Schülerleistungen mittels Beurteilungsmasken, wie sie in *Item Builder* umgesetzt werden können, ausprobiert. Die Ergebnisse der abschliessenden Erprobungen sind insgesamt positiv. Die Testlogistik für ganze Schulklassen funktionierte, und die Aufgaben wurden den Erwartungen entsprechend bearbeitet. Die Formate können als so umfassend erprobt gelten, dass sie nun für den Einsatz in Testinstrumenten zur Verfügung stehen. Bei den rezeptiven Aufgaben haben sich die erprobten Testinhalte für eine formativ-diagnostische Verwendung als eher zu einfach erwiesen. Die produktiven Aufgaben haben

insgesamt gut funktioniert, auch bei den erstmals beteiligten Fünft- und Sechstklässlern. Sie haben grösstenteils zu der Art von Leistungen geführt, wie sie erwartet wurden. Der Versuch, zu den mündlichen Produktionen einen Typ von analytischer Bewertung zu entwickeln, der transparent genug ist, dass er von Lehrpersonen nahezu intuitiv angewandt werden kann, ist vielversprechend verlaufen. Allerdings fehlen noch die praktischen Erfahrungen einer grösseren Erprobung durch Lehrpersonen.

Sprache des Unterrichts

Wenn Lehrende und Lernende im Unterricht die Fremdsprache nutzen, um Arbeitsaufträge auszuhandeln, Hausaufgaben zu besprechen oder alltägliche Rituale wie Begrüssen und Verabschieden zu vollziehen, wird aus dem Unterrichtsinhalt „Fremdsprache“ ein authentisches Mittel zur Kommunikation. In der Fremdsprachendidaktik wird in diesem Zusammenhang oft auf Englisch von *classroom language* gesprochen – der Umgangssprache im Unterricht. Ihre Anwendung hat inzwischen als Lernziel Einzug in die schweizerischen Lehrpläne gefunden und wird in verschiedenen Lehrmitteln, darunter auch *Mille feuilles*, gezielt thematisiert. Da der Sprachunterricht ein Handlungsfeld darstellt, in dem von Beginn an authentisch kommuniziert werden kann, wurde die Sprache des Unterrichts im Projekt „Erfassung von Entwicklungsprofilen“ als Bereich aufgenommen, zu dem Aufgaben entwickelt werden sollten. Dafür wurde zunächst auf der Grundlage der beiden im Projekt relevanten Lehrwerke *Mille feuilles* (Kantone entlang der deutsch-französischen Sprachgrenze) und *Alex et Zoé* (Tessin) sowie von Weiterbildungsmaterialien eine Sammlung von Äusserungen erstellt, die für die mündliche Kommunikation im Schulzimmer relevant erschienen. Diese Sammlung wurde anschliessend Experten vorgelegt, welche die Auswahl der Äusserungen überprüften. In einem weiteren Schritt wurden Primarlehrpersonen für Französisch als Fremdsprache (N=74) in einem Onlinefragebogen gebeten einzuschätzen, wie gut ihre Lernenden (der dritten

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

bis sechsten Klasse) eine Auswahl dieser verschiedenen Äusserungen rezeptiv bzw. produktiv beherrschten. Die Lehrpersonen schlugen zudem weitere relevante Äusserungen vor, die noch nicht Teil des Fragebogens waren. Die Auswertung des Fragebogens erlaubt u.a. folgende Schlüsse:

1. Im Allgemeinen beherrschen ältere Schülerinnen und Schüler die vorgelegten Äusserungen besser; sie können auch mehr Äusserungen produzieren.
2. Die meisten Äusserungen werden zunächst rezeptiv und erst später produktiv beherrscht.
3. Im zweiten Halbjahr der 3. Klasse, nach einem halben Jahr Französischunterricht, verstehen die Lernenden eine geringe Zahl sehr häufiger Ausdrücke (z.B. „Bonjour.“ „C'est faux.“, „Ouvre la fenêtre.“), zum Teil können sie sie auch selbst produzieren.
4. Äusserungen, die in *Mille feuilles* nicht vorkommen, sind den Lernenden unbekannt.
5. Im Grossen und Ganzen verstehen die Lernenden Anweisungen recht gut. Funktionale Äusserungen, die für Gruppenspiele benötigt werden (z.B. „Jette le dé.“ bei Brettspielen), sind den Lernenden nur teilweise bekannt, obschon dieser Wortschatz durch *Mille feuilles* eigentlich gezielt aufgebaut wird.

Die detaillierten Ergebnisse dieser Umfrage sind als elektronischer Anhang des Berichts auf der Webseite des Kompetenzzentrums zugänglich (<https://bit.ly/2jTYXFM>). Sie können u.a. als Redemittelkatalog für die Kommunikation im Unterricht dienen und geben Aufschluss über die Verteilung dieser Redemittel auf die vier Schuljahre der Primarschule. Auch Redemittel, welche die Lehrpersonen im Fragebogen als ebenfalls relevant angegeben haben, sind in diesem Dokument aufgeführt.

Das Handlungsfeld Unterricht und die entsprechende Sprache dienen als Quelle für die Erstellung von Aufgaben im Teilprojekt „Kompetenzen im Bereich des Sprechens“. Die dafür ausgewählten Äusserungen stammten aus den Bereichen „Anweisungen und Anleitungen“ und „Arbeit und Spiel in der Gruppe“. Es wurden sowohl produktive als auch rezeptive Aufgaben erstellt und erprobt. Die Situierung der Wendungen in einen kommunikativen Zusammenhang geschah durch eines oder mehrere Bilder, welche die Handlung abbildeten.

Sprachlernstrategische Kompetenzen

Die Beschreibung und Erfassung der sprachlernstrategischen Kompetenzen von Primarschülerinnen und -schülern bildete neben den Arbeiten im Bereich des Sprechens einen der beiden wichtigsten inhaltlichen Schwerpunkte des Projekts „Erfassung von Entwicklungsprofilen“. Sprachlernstrategien spielen als sog. transversale Kompetenzen in der Mehrsprachigkeitsdidaktik eine Rolle. In den „Grundkompetenzen“ der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren sowie den Fremdsprachenlehrplänen der Deutschschweiz wird die Entwicklung methodischer bzw. lernstrategischer Kompetenzen unter den Zielen des Unterrichts genannt. Dies mag erstaunen, denn die Fachliteratur weist darauf hin, dass der Zusammenhang zwischen dem Einsatz von Strategien und dem Lernerfolg kein einfacher ist. Ob sich bestimmte Strategien positiv auswirken oder nicht, scheint u.a. kompetenzstands-, situations- und kontextabhängig zu sein. Immerhin aber lassen sich Lernstrategien explizit vermitteln. Zudem kann empirisch ein positiver Zusammenhang zwischen Strategieinstruktion und Sprachverwendungskompetenz gezeigt werden.

Das Teilprojekt V beschränkte sich in seinen Entwicklungsarbeiten auf eigentliche Sprachlernstrategien (Strategien im Zusammenhang mit spezifischen Lern-Handlungen, z.B. Strategien zum Wortschatzlernen) und befasste sich beispielsweise nicht mit Sprachverwendungsstrategien (z.B. Lesestrategien).

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

In einem ersten Arbeitsschritt erstellte das Projektteam auf der Grundlage von bestehenden Sammlungen sowie von Fremdsprachenlehrwerken ein strukturiertes Inventar von solchen Strategien. Nach längeren Vorarbeiten wurden schliesslich zu drei Teilbereichen Items zur Erfassung von sprachlernstrategischen Kompetenzen entwickelt. Die Teilbereiche unterscheiden sich deutlich voneinander:

1. Umfang und Vielfalt des individuellen Einsatzes von Sprachlernstrategien;
2. Wissen über Sprachlernstrategien und deren Eignung;
3. Zwei Aspekte des metakognitiven Könnens: zum einen das Einschätzen des eigenen Lernstandes, operationalisiert als die Fähigkeit, den eigenen Erfolg bei sprachlichen Aufgaben vorherzusagen; zum anderen die metakognitive Selbstüberwachung, operationalisiert als die Fähigkeit, die Korrektheit der selbst ausgeführten sprachlichen Handlungen zu evaluieren (Sicherheitsurteil).

Für die Erfassung von Daten zu jedem der Bereiche wurde jeweils ein spezifischer Itemtyp verwendet:

- zu 1. Selbstaussagen zur Häufigkeit der Verwendung von konkreten Strategien (fünfstufige Likertskala: „sehr selten“ bis „sehr oft“), vergleichbar mit Items in zahlreichen Strategiefragebogen.
- zu 2. Rankingaufgaben: Eine Sprachlernaufgabe und dazu drei mögliche, aber nach Expertenmeinung unterschiedlich geeignete Vorgehensweisen sind beschrieben. Aufgabe der Lernenden ist es, die drei Vorgehensweisen nach ihrer Eignung zu ordnen.
- zu 3. Für die Vorhersage: Beantwortung der Frage „Könntest du [eine gegebene sprachliche Aufgabe] lösen?“; Antwort auf einer vierstufigen Likertskala von „Nein, sicher nicht“ bis „Ja, ganz sicher“.

Für das nachträgliche Sicherheitsurteil: „Löse [die vorgegebene sprachliche Aufgabe].“; anschliessend Beantwortung der Frage „Hast du das richtig gemacht?“ auf der gleichen Likertskala wie zuvor. Die Lernenden lösen die beiden zusammengehörigen Aufgaben jeweils mit einem zeitlichen Abstand, sodass das Sicherheitsurteil unabhängig von der Vorhersage abgegeben wird. Für die Bewertung wird der Korrektheitsgrad der Lösung der Sprachaufgabe mit den beiden Einschätzungen verglichen.

Wie auch sonst im Projekt wurden die Instrumente zur sprachlernstrategischen Kompetenz generell für eine periodische, mehrere Monate auseinanderliegende selbstständige Nutzung durch die Schülerinnen und Schüler konzipiert. Die Ergebnisse sollten jeweils so aufbereitet werden können, dass sie im Unterricht zur Fokussierung des Weiterlernens genutzt werden können. In Bezug auf die Ergebnisse zu den Sprachlernstrategien wurden u.a. die folgenden Berichtsmöglichkeiten mit formativem Potential angedacht:

1. Selbstaussagen zu Umfang und Vielfalt des Gebrauchs von Sprachlernstrategien:
 - Häufigkeit der Verwendung der einzelnen Strategien in einer Klasse im Vergleich zur Verwendung in allen vergleichbaren Klassen (Referenzgruppe);
 - Häufigkeit der Strategienutzung durch die einzelnen Lernenden (wie oft haben die einzelnen Lernenden die Kategorien „sehr selten“ bis „sehr oft“ der Likertskala gewählt?).
2. Aufgaben zum Wissen über Sprachlernstrategien:
 - Rohdaten einzelner Lernender und ganzer Klassen zum Ranking der zur Auswahl stehenden Strategien (und

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

anschliessende Arbeit mit den Test-Items im Unterricht);

- Ergebnisse von einzelnen Lernenden und Klassen auf der Testskala zum Wissen über die Eignung von Strategien (um z.B. Fortschritte gegenüber früher zu verfolgen).
3. Aufgaben zu den beiden Aspekten des metakognitiven Könnens:
 - Ergebnisse der einzelnen Lernenden bezüglich ihrer Tendenz zur Über- oder Unterschätzung der eigenen Fähigkeiten (oder zusammengenommen zur Verschätzung);
 - Ergebnisse der einzelnen Lernenden differenziert nach der Verschätzung bei Vorhersagen und bei Sicherheitsurteilen. Die beiden Ergebnisse können auf unterschiedliche Stärken bzw. Schwächen hinweisen.

Da in diesem Teilprojekt viel Pionierarbeit zu leisten war, wurden die Instrumente in wiederholtem Kontakt mit Klassen entwickelt und erprobt. Der Entwicklungsstand der Instrumente zu den drei Bereichen unterscheidet sich. Die Selbstaussagen zum Strategieverbrauch wurden insofern erfolgreich erprobt, als die Items offenbar leicht bearbeitet werden konnten und durchaus unterschiedliche Profile der Strategieverwendung ergaben. Es zeichnen sich auch sinnvolle Möglichkeiten zur Nutzung der Ergebnisse ab. Die Rankingaufgaben zum Wissen über die Eignung von Sprachlernstrategien haben sich in der Erprobung nur bedingt bewährt. Verschiedene denkbare Gründe haben dazu geführt, dass die Items nur unbefriedigend zwischen den Lernenden diskriminierten und es im Rahmen des Projekts nicht möglich war, eine Kompetenzskala für diesen Bereich zu beschreiben. Die Aufgaben zur metakognitiven Überwachung wurden erfolgreich erprobt. Das verwendete Aufgabenformat sollte trotzdem weiter untersucht werden, denn es ist noch nicht klar genug, wie sehr

sich die Anforderungen der eingebetteten Sprachaufgabe mit den Anforderungen an die Fähigkeit zur metakognitiven Überwachung vermischen.

Mit dem Teilprojekt zur Beschreibung und Erfassung sprachlernstrategischer Kompetenz konnten in einem komplexen und v.a. für den Primarschulbereich noch wenig entwickelten Gebiet verschiedene konzeptionelle und materielle Beiträge geleistet werden. Zudem wurden, nicht zuletzt durch das Verwerfen von erfolglos erprobten Ansätzen, wichtige Erkenntnisse gewonnen, auf die weiter aufgebaut werden kann.

Beurteilungs- und Feedbackpraxis im Fremdsprachenunterricht in der Primarschule

Ein Teil des Projekts „Erfassung von Entwicklungsprofilen“, welcher vollständig im Tessin durchgeführt wurde, untersuchte mit einem qualitativen Ansatz Beurteilungspraktiken und -bedürfnisse der Lehrpersonen der Primarschule (*scuola elementare*) sowie ihre Haltung gegenüber einem diagnostisch anwendbaren, computerbasierten Beurteilungsinstrument. Zu diesem Zweck wurden neun Lehrpersonen befragt, von denen fünf bereits früher im Rahmen des Projekts an verschiedenen Aufgabenerprobungen teilgenommen hatten. Die Interviews wurden mithilfe eines Leitfadens durchgeführt, dessen Inhalte aufgrund der Erfahrungen anlässlich einer Aufgabenerprobung, einer Analyse der Unterrichtssituation im Tessin, des relevanten Lehrwerks sowie der Literatur in den Bereichen formative Beurteilung und Feedback theoriegeleitet entwickelt wurden. Befragt wurden die Lehrpersonen zunächst zu den Kompetenzen und Entwicklungswegen ihrer Lernenden sowie zu ihren eigenen Formen der Beurteilung im Unterricht. In einem zweiten Schritt wurden ihnen dann mögliche Teile eines computerbasierten Beurteilungsinstruments zur Diskussion vorgelegt. Die Interviews wurden transkribiert und mittels eines Auswertungsrasters inhaltlich ausgewertet, der aufgrund eines Vorgehens im Stil der Grounded Theory entstanden ist.

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

Bei den meisten der neun Lehrpersonen basiert die summative Beurteilung häufig auf Lehrwerkmaterialien, während formative Beurteilungselemente informell und intuitiv in den Unterricht eingebracht werden. Im Bereich der mündlichen Produktion legen die Lehrpersonen ein besonderes Augenmerk auf die Aussprache bzw. Verständlichkeit der gelernten Wörter und Wendungen. Strategien und *language awareness (éveil aux langues)* werden nur am Rande thematisiert; dies u.a., weil das Lehrwerk in diesem Bereich nichts vorsieht. Die neun befragten Primarlehrpersonen betonen, dass die Beurteilung, ähnlich wie die Auswahl der Lerninhalte, vor allem im Hinblick auf die Motivierung und das Wohlbefinden der jungen Lernenden stattfindet.

Was den periodisch wiederkehrenden Einsatz eines computerbasierten Beurteilungsinstruments angeht, welches formativ-diagnostisch einsetzbar wäre, zeigte sich in den Interviews eine gewisse Skepsis bezüglich der praktischen Umsetzbarkeit: Zu einer teilweise unzureichenden Ausstattung der Schulen mit Computern und einer sehr geringen Stundenzahl für das Fach Französisch kommen Befürchtungen wegen des organisatorischen Aufwands bei der Erfassung mündlicher Produktionen im Rahmen des Unterrichts hinzu. Die Lehrpersonen bezweifeln auch, dass der Erkenntnisgewinn den Aufwand rechtfertigen würde, den sie wegen der computerbasierten Bewertung der mündlichen Schülerproduktionen hätten. Mehrere Lehrpersonen würden sich vielmehr ein Instrument wünschen, das die Lernenden selbstständig zum Üben und Überprüfen ihrer mündlichen Kompetenzen nutzen könnten. Immerhin sind die Lehrpersonen der Ansicht, dass die Sprechaufgaben, die im Rahmen des Projekts entwickelt wurden, an sich für den vorgesehenen Verwendungszweck geeignet und dem Tessiner Kontext angemessen wären.

Récolte informatisée de données relatives aux compétences en langues étrangères à l'école primaire

Rapport du projet « Recensement des profils de développement dans le cadre de l'apprentissage précoce des langues étrangères à l'école »

Résumé

Le projet «Recensement des profils de développement dans le cadre de l'apprentissage précoce des langues étrangères à l'école» fait partie du programme de recherche 2012-2015 du Centre scientifique de compétence sur le plurilinguisme (CSP). Son objectif était de contribuer à créer, dans le cadre de l'enseignement des langues étrangères à l'école primaire (de la troisième à la sixième), une culture d'évaluation à la fois informatisée et orientée sur l'apprentissage grâce au développement d'instruments adaptés à la vérification des compétences. Les travaux de développement se sont inspirés du concept de la *progression des apprentissages (learning progressions)*, surtout appliqué en didactique des sciences naturelles, où il constitue souvent la base des programmes, de l'enseignement et de l'évaluation. La progression des apprentissages décrit les voies de développement typiques menant du statut de novice à celui d'expert dans un domaine de connaissance et de compétence donné (par ex. en lien avec l'évolution de la compréhension adéquate de la matière en science naturelle). La progression des apprentissages ne doit donc pas être comprise comme une séquence d'acquisition naturelle, mais comme le résultat «typique» de l'interaction entre les prédispositions individuelles, le sujet à acquérir et une certaine méthode d'enseignement. Le développement et l'utilisation de tâches destinées à mesurer la progression des apprentissages et la description de cette progression sont interdépendants. Lorsque la voie de progression d'apprentissages est suffisamment connue, les performances des élèves peuvent être reportées sur cette voie (à la façon d'une carte routière), de sorte qu'il est possible d'en déduire des *feedbacks* et *feedforwards* (indications concernant les étapes de développement suivantes) formatifs optimaux.

Dans le domaine de l'enseignement et de l'apprentissage des langues, très peu de progressions d'apprentissages ont été développées à ce jour. Le projet «Dynamic Language Learning Progressions» (DLLP),¹ qui se situe dans le contexte de l'apprentissage de l'anglais à l'école, est une

exception. Il aborde le développement de compétences langagières fonctionnelles, en particulier l'aptitude à donner des explications en rapport avec des tâches scolaires.

Dans le projet «Recensement des profils de développement», la priorité a été donnée au développement de modèles de tâches plutôt qu'à une description globale de la progression des apprentissages. Ceci principalement en raison d'un objectif ambitieux consistant à récolter les données relatives au niveau de performance des élèves via un système hébergé sur un serveur internet et à générer, grâce à ce système, un *feedback* utile à la poursuite de l'apprentissage et de l'enseignement. Afin de développer le plus possible la partie technique, cet aspect a été promu dès le début du projet. En effet, la grande attractivité d'un système d'évaluation en ligne repose notamment sur le fait qu'il est possible de continuer à le développer en cours d'utilisation, sur la base des nouvelles données que l'on y insère, afin d'améliorer la qualité des tâches et d'approfondir les résultats concernant les compétences des élèves.

Du point de vue du contenu, le projet s'est penché sur la production et l'interaction orales ainsi que les stratégies d'acquisition et d'apprentissage de la langue. Un choix consciemment ambitieux, dans la mesure où ces domaines de compétences sont en général délaissés, surtout lorsqu'il s'agit de les évaluer.

La décision, prise en 2012, par la Conférence suisse des directeurs cantonaux de l'instruction publique (CDIP) de mettre sur pied une banque de données de tâches a été déterminante pour la mise en œuvre technique du présent projet. Cette banque de données devait non seulement être utilisée dans le cadre de l'évaluation des compétences des cantons mais également pour d'autres projets d'évaluation dans le contexte scolaire. La collaboration née grâce à ce projet de banque de données entre le CSP et l'Institut Leibniz pour la recherche et l'information pédagogique (Leibniz-Institut für Bildungsforschung und Bildungsinformation) (DIPF) a eu pour résultat l'extension du logiciel de test *CBA Item Builder*, destiné au recensement et à l'évaluation des performances orales.

1 | <https://www.dllp.org/>

Récolte informatisée de données relatives aux compétences en langues à l'école primaire

Le niveau d'exigence et les défis du projet «Recensement des profils de développement» ont été considérables. Les cinq parties de ce rapport donnent un aperçu des différents travaux réalisés:

- Partie II «Evaluation assistée par ordinateur des compétences orales» aborde en première ligne l'aspect technique de la mise en œuvre d'un système hébergé sur serveur qui puisse tout particulièrement servir à une évaluation formative et diagnostique des compétences langagières orales. L'accent a été mis sur le développement de tâches faciles à utiliser, accessibles via un navigateur, et sur un module destiné à l'évaluation basée sur des critères de prestations orales, recueillies au préalable par les enseignantes et enseignants.
- Partie III «Compétences dans le domaine de la langue parlée» décrit le développement et la validation progressifs de tâches permettant de générer des données de diagnostic dans les sous-domaines de compétence langagière orale des élèves faisant l'objet de l'étude.
- Partie IV «Langage de classe» traite d'un sous-projet dont l'objectif était d'identifier des outils langagiers importants pour la communication orale en classe, sur la base d'analyses de moyens d'enseignement et d'un sondage réalisé auprès du corps enseignant. Ce sondage a aussi permis d'examiner dans quelle mesure les élèves maîtrisent ces outils langagiers de manière réceptive et productive à chaque niveau scolaire.
- Partie V «Compétences en stratégie d'apprentissage des langues» présente le développement d'instruments permettant de recenser d'une part la fréquence d'utilisation des stratégies d'apprentissage par les élèves, et d'autre part les compétences des élèves dans un certain nombre

de sous-domaines de compétence en stratégie d'apprentissage des langues. Un inventaire structuré des compétences en stratégie d'apprentissage des langues a été élaboré comme base de travail.

- Partie VI «Pratiques d'évaluation et de *feedback* dans l'enseignement d'une langue étrangère à l'école primaire» détaille les résultats d'une étude qualitative effectuée dans le canton du Tessin. Cette étude se penche sur les pratiques en matière d'évaluation des enseignantes et enseignants de français du primaire. Elle examine également dans quelle mesure ces enseignant-e-s accepteraient la mise en place d'un environnement informatisé destiné à une évaluation formative et diagnostique.

Dans le cadre de ce projet, une série de travaux importants a pu être menée en vue du développement d'un système d'évaluation formative et diagnostique, basé sur serveur. Certains des éléments élaborés ont déjà pu être utilisés individuellement, parfois dans d'autres contextes, en particulier pour le développement de tests informatisés.

Evaluation assistée par ordinateur des compétences orales

Une partie des ressources du projet a été engagée pour l'étude et le développement partiel d'un système de test, hébergé sur un site internet, qui pourrait permettre la mise en œuvre d'un système d'évaluation formative et diagnostique polyvalent. La CDIP ayant annoncé en 2012 l'élaboration d'une *banque de données de tâches*, le projet s'est concentré sur le développement d'un système d'évaluation informatisé pouvant justement s'intégrer dans cette plateforme de test. Après une phase d'étude avec Microsoft PowerPoint et des logiciels standards d'enregistrement sonore, toutes les tâches développées au cours du projet ont été transposées dans le logiciel *CBA Item Buil-*

Récolte informatisée de données relatives aux compétences en langues à l'école primaire

der et livrées via le serveur d'applications associé, EE4CBA. Ces deux éléments ont été mis à disposition par le DIPF, avec lequel la CDIP a collaboré au développement de la banque de données de tâches. La solution informatique a été conçue de telle sorte que les écoles aient uniquement à fournir des ordinateurs avec un accès internet, des navigateurs récents et des casques audio.

Les travaux de développement des fonctionnalités du logiciel, nécessaires au recensement des prestations orales des élèves, ont été réalisés via le DIPF. Ils ont abouti à l'ajout d'un module destiné à l'enregistrement sonore par le biais d'un navigateur et celui d'un module d'évaluation (*rating*). Ce dernier permet de gérer les enregistrements disponibles (ou les textes écrits) et de les évaluer manuellement sur la base de critères donnés. Les évaluations sont ensuite sauvegardées de la même façon que les réponses aux questions fermées (p. ex. choix multiples).

Cependant, toutes les applications garantissant que le système d'évaluation formative et diagnostique fonctionne pleinement ne sont pas fournies. Il manque notamment un module intégré, dédié à la gestion des personnes et des tâches. Il manque également un module d'évaluation automatique, qui permette de rapporter les réponses des élèves et les évaluations manuelles à des échelles et de générer, sur cette base, des rapports soit des *feedbacks* et des *feedforwards*.

Le cœur du projet a consisté à développer et à tester des formats et des contenus de tâches adaptés à de jeunes élèves, que ces derniers puissent résoudre de manière autonome.

Afin d'obtenir les informations nécessaires, les questions suivantes ont notamment été examinées lors des différentes phases de test : familiarité des élèves, surtout des plus jeunes, avec les logiciels et les ordinateurs, pertinence des interfaces utilisateurs, éléments de commandes (symboles) et de saisie (textes, images), compréhensibilité des instructions (dans la langue de l'école comme dans la langue cible), facilité d'utilisation technique et logistique des tâches en classe, difficulté et adéquation des formats et contenus des tâches. Une des principales conclusions est que l'utilisation d'un ordinateur n'est manifestement

pas un obstacle, même pour les participant-e-s les plus jeunes. D'autre part, il est apparu dès les premières phases de test que même les pages d'écran comportant peu de texte ainsi que les tâches à plusieurs étapes, donc plus exigeantes sur le plan cognitif, sont source de grande confusion et ne se prêtent par conséquent pas à une évaluation valide. Progressivement, de nouveaux formats de tâches ont donc été développés ; ces derniers ne nécessitent qu'un minimum d'explications et leur résolution correcte ne dépend pas de la compréhension d'un texte écrit (puisque ce sont les compétences orales que l'on recueille). Les observations faites dans ce projet quant à l'adéquation des instruments d'évaluation aux besoins des élèves du primaire, tout comme les formats de tâches comprenant les images et symboles utilisés sont disponibles pour le développement d'autres tâches et tests.

Compétences dans le domaine de la langue parlée

Le développement d'outils assistés par ordinateur (tâches), destinés au recensement des compétences encore élémentaires des élèves de l'école primaire dans le domaine de la production et de l'interaction orales, est au cœur du sous-projet « Langue parlée ». Les documents issus du Cadre européen commun de référence pour les langues (CECR) ou du domaine des tests de langue ne fournissent que des indications imprécises concernant les caractéristiques et le développement des compétences au sein des niveaux A du CECR. Une caractéristique souvent évoquée est que la résolution de tâches de communication dans des contextes élémentaires spécifiques et familiers repose sur quelques mots appris et expressions acquises (*chunks*) qui peuvent être utilisés de manière plus ou moins flexible. Dans le cas des élèves de primaire, cela signifie que ces derniers sont fortement dépendants du contenu du cours et des moyens d'enseignement. Cette dépendance a pu être mise en évidence au travers de premiers tests réalisés dans trois contextes différents (dans les régions du plan d'études Passepartout,²

Récolte informatisée de données relatives aux compétences en langues à l'école primaire

en Suisse alémanique et au Tessin). L'objectif, qui était de dégager des lignes de développement communes aux trois contextes, s'étant avéré trop contraignant pour le développement des tâches, le sous-projet « Langue parlée » s'est focalisé sur la région du plan d'études Passepartout dans la phase finale.

Le sous-projet a été structuré en trois grandes étapes pour procéder au recensement et diagnostic des compétences orales des élèves de primaire dans leur première langue étrangère.

Au cours de la première étape, les tâches de recensement des compétences existantes se sont appuyées sur le contenu et le type de tâches (de communication) apparaissant le plus fréquemment dans les différents moyens d'enseignement de notre domaine de pratique. Les tâches, encore sur support papier, ont été testées sur place. Ces tests ont ensuite fait l'objet d'une évaluation détaillée, d'une part du point de vue du fonctionnement des tâches et d'autre part de celui des compétences des jeunes élèves. Les tâches inspirées des moyens d'enseignement ont en général été laissées suffisamment ouvertes pour qu'un large spectre de compétences puisse être recensé.

Les tests ont toutefois indiqué que de telles tâches ne permettaient pas de recenser efficacement les compétences langagières naissantes de nombreux élèves de primaire, dans la mesure où ces derniers n'étaient pas à même de maîtriser ces tâches sans l'aide de tiers. Ceci tenait à deux facteurs : d'une part aux exigences cognitives découlant du fait que les tâches à résoudre se référaient à des situations de communication, et d'autre part au niveau d'exigences fixé en matière d'expression orale dans la langue étrangère, encore qu'il ait été généralement très faible.

Dans une deuxième phase, un accent plus important a été mis sur les *ressources* linguistiques et fonctionnelles, intervenant dans l'usage communicatif (plus tardif) d'une langue, que sur la production langagière dans le cadre de tâches de communication, même si de telles tâches ont conti-

nué à être utilisées dans des formats simples. L'analyse détaillée des composantes langagières et communicatives des tâches, définies dans les moyens d'enseignement pour stimuler et évaluer les compétences orales, a constitué le point de départ des constructions sur lesquelles le projet se focalisait à ce moment-là.

Les tests menés lors de la deuxième phase ont été réalisés au moyen d'une sélection de tâches ciblant soit les ressources linguistiques, soit les compétences langagières fonctionnelles ou communicatives. Les enfants ont travaillé pour la première fois sur des ordinateurs. Les expériences ont montré que même les jeunes enfants avaient la disposition et les connaissances nécessaires à l'exécution de tâches informatisées. Les points problématiques se sont révélés être les consignes des tâches d'une certaine longueur, les exemples à l'écran et, tout particulièrement, l'utilisation de la langue cible dans les consignes des tâches. En matière de compétences, la phase de test a permis de tirer les conclusions suivantes : les problèmes de prononciation sont très marginaux, alors que la lecture à haute voix de la langue écrite place les jeunes élèves devant un véritable défi. La connaissance concrète de mots et d'expressions semble essentielle, y compris et avant tout aux yeux des élèves eux-mêmes. Les tâches portant sur la syntaxe ont ainsi montré que les enfants n'ont parfois même pas essayé de les résoudre quand ils ne comprenaient pas tous les mots. D'éventuelles connaissances syntaxiques ont donc vraisemblablement échappé à l'évaluation. Certaines tâches plus larges, requérant la communication orale, se sont aussi révélées problématiques : les consignes des tâches légèrement plus difficiles, même formulées dans la langue scolaire, n'ont pas été qu'en partie lues et respectées. De plus, la plupart des productions orales réalisées de façon autonome dans la langue cible sont restées très modestes.

Dans sa dernière phase, le projet a uniquement porté sur la région du plan d'études Passepartout et donc seulement sur les élèves apprenant le français sur la base du manuel didactique *Mille feuilles*. Cependant, pour la première fois des élèves de cinquième et de sixième ont participé. De

2 | Les six cantons situés sur la frontière linguistique allemand-français Berne, Bâle-Campagne, Bâle-Ville, Fribourg, Soleure et Valais.

Récolte informatisée de données relatives aux compétences en langues à l'école primaire

plus, la palette des domaines ciblés a encore été réduite. Les sous-domaines traités appartenaient d'une part à deux domaines d'action communicative – la communication en classe (*classroom language*) et l'échange d'informations personnelles – et d'autre part, à une compétence langagière partielle qu'est la compétence lexicale. La compréhension de la langue parlée a été intégrée en tant que nouvel élément de la compétence langagière, dans l'espoir de faire apparaître de façon plus nuancée les compétences langagières orales élémentaires. Les tâches des domaines mentionnés ont été adaptées encore mieux au groupe cible dans leur structure, leurs exigences et leur apparence. Un corpus de tous les textes écrits et parlés du manuel didactique *Mille feuilles* a été élaboré comme base au développement actuel (voir futur) des tâches. Il a également été dressé un catalogue d'outils linguistiques, adaptés à la communication en classe. Celui-ci se base sur une analyse de différents moyens d'enseignement et sur un sondage complémentaire, réalisé auprès du corps enseignant. Le corpus et le catalogue ont été utilisés pour développer une série de tâches exemplaires, dédiées à un usage linguistique oral (réception, interaction, production). Ces tâches sont brèves et tiennent sur une page d'écran. Nombre d'entre elles ont été illustrées dans le but d'en faciliter une compréhension intuitive. Seul un format un peu plus complexe a été conservé pour l'évaluation des compétences d'interaction orale : le « Dialogue » déjà testé auparavant. Un dialogue de ce type s'étend sur plusieurs pages et consiste à alterner écoute et production orale avec une personne fictive.

Le test pratique a été réalisé en plusieurs étapes. Il a permis de vérifier les tâches elles-mêmes, mais aussi de s'assurer que la diffusion via l'infrastructure de l'école des tests hébergés sur un serveur fonctionnait correctement pour des classes entières. De plus, les enregistrements réalisés lors de cette phase ont servi de base à l'évaluation analytique (manuelle) des performances orales des élèves au moyen de masques d'évaluation, tels qu'ils peuvent être appliqués dans *Item Builder*. Dans l'ensemble, les résultats des tests finaux sont positifs. La logistique de test a fonctionné pour des classes entières et les tâches ont

été exécutées conformément aux attentes. Les formats peuvent être considérés comme ayant été pleinement validés de sorte qu'ils peuvent désormais être intégrés dans des instruments de test. Les contenus de test destinés à une application formative et diagnostique des tâches réceptives se sont révélés plutôt trop faciles. Les tâches productives ont globalement bien fonctionné, même chez les élèves de cinquième et de sixième participant pour la première fois. Dans l'ensemble, elles ont mené au genre de prestation attendue. La tentative de développer un type d'évaluation analytique pour les productions orales, qui soit suffisamment transparent pour être utilisé quasi intuitivement par les enseignant-e-s, a été prometteuse. Il manque toutefois encore l'expérience pratique d'un test à plus large échelle réalisé par le corps enseignant.

Langage de classe

Lorsque les élèves et les enseignant-e-s utilisent la langue étrangère en classe pour se mettre d'accord sur des exercices, discuter de devoirs ou pour accomplir les rituels sociaux quotidiens comme les salutations et les prises de congé, la « langue étrangère » enseignée en classe devient une forme de communication authentique. La didactique des langues étrangères parle dans ce cas de langage de classe (*classroom language*), c'est-à-dire du langage courant pratiqué en classe. Son application est devenue un objectif pédagogique intégré dans les plans d'étude suisses, largement thématiqué dans divers manuels d'enseignement, entre autres *Mille feuilles*. Puisque l'enseignement des langues constitue un champ d'action dans lequel il est possible de communiquer dès le début d'une manière authentique, le *classroom language* a été inclus dans le projet comme un domaine dans lequel des tâches doivent être développées. Dans un premier temps, un recueil des expressions qui semblent pertinentes pour la communication orale en classe a été constitué sur la base des deux moyens d'enseignement utilisés dans le projet, *Mille feuilles* (cantons suisses alémaniques à la frontière franco-

Récolte informatisée de données relatives aux compétences en langues à l'école primaire

phone) et *Alex et Zoé* (Tessin), ainsi que sur la base du matériel de formation continue. Ce corpus d'expressions a ensuite été soumis à des expert-e-s qui ont vérifié cette sélection. Au cours d'une étape ultérieure, des enseignant-e-s de français langue étrangère en primaire (N=74) ont été invité-e-s à évaluer la capacité de leurs élèves à maîtriser ces différentes expressions au niveau réceptif et productif par le biais d'un questionnaire en ligne. Les enseignant-e-s ont également proposé d'autres expressions pertinentes qui n'étaient pas encore incluses dans le questionnaire. L'analyse du questionnaire a notamment mené aux conclusions suivantes :

1. En général, les élèves plus âgés maîtrisent mieux les expressions présentées ; ils en connaissent aussi un plus grand nombre.
2. La plupart des expressions sont d'abord acquises au niveau réceptif puis, dans un deuxième temps, au niveau productif.
3. Au deuxième semestre, après six mois d'enseignement du français, les élèves de 3^e classe comprennent un petit nombre d'expressions très fréquentes (p.ex. « Bonjour. », « C'est faux. », « Ouvre la fenêtre. »). Ils sont en partie aussi capables de les produire eux-mêmes.
4. Les expressions qui n'existent pas dans *Mille feuilles* ne sont pas connues des élèves.
5. Globalement, les élèves comprennent assez bien les instructions. Les expressions fonctionnelles nécessaires aux jeux en groupe (p.ex. « Jette le dé. » pour les jeux de société) ne sont que partiellement connues des élèves bien que ce vocabulaire soit développé de façon ciblée au travers de *Mille feuilles*.

Les résultats détaillés de ce sondage sont disponibles sous forme électronique en annexe du présent rapport sur le site web du Centre de

compétence (<https://bit.ly/2kLi6dq>). Ils peuvent notamment servir de catalogue d'expressions langagières pour la communication orale en classe et donnent des indications concernant la répartition de ces expressions au cours des quatre années d'école primaire. Les expressions présentées comme pertinentes par le corps enseignant dans le questionnaire se trouvent aussi dans ce document.

Les cours en classe et la langue concernée ont servi de source pour l'élaboration de tâches dans le sous-projet « Compétences dans le domaine de la langue parlée ». Les expressions sélectionnées dans ce but sont issues des domaines « Instructions » et « Travail et jeux en groupe ». Des tâches productives et réceptives ont été élaborées et testées. La mise en situation des expressions liées à la communication a été réalisée par le biais d'une ou plusieurs images représentant l'action.

Compétences en stratégie d'apprentissage des langues

La description et le recensement des compétences en stratégie d'apprentissage des langues des élèves de l'école primaire ont constitué, avec les travaux dans le domaine du langage oral, l'un des deux axes principaux du projet « Recensement des profils de développement ». En tant que compétences transversales, les stratégies d'apprentissage de la langue jouent en effet un rôle en didactique du plurilinguisme. Le développement de compétences méthodologiques, c'est-à-dire de compétences en stratégie d'apprentissage des langues, est du reste cité comme un des buts de l'enseignement à la fois par la Conférence suisse des directeurs cantonaux de l'instruction publique, dans sa définition des compétences fondamentales, et dans les plans d'études des langues étrangères appliqués en Suisse alémanique. Ceci peut paraître étonnant car, selon la littérature spécialisée, le lien entre l'utilisation de stratégies et la réussite de l'apprentissage n'est pas direct ; il semble que l'impact, positif ou non, de certaines stratégies dépende notamment du niveau de compétence, de la situation et du

Récolte informatisée de données relatives aux compétences en langues à l'école primaire

contexte. Au moins les stratégies d'apprentissage peuvent-elles être transmises explicitement. En outre, un lien positif entre l'enseignement de stratégies et les compétences d'utilisation de la langue peut être démontré sur le plan empirique.

Dans ses travaux de développement, le sous-projet V s'est concentré sur les stratégies d'apprentissage des langues effectives (soit des stratégies en lien avec des activités spécifiques d'apprentissage, comme des stratégies d'apprentissage du vocabulaire) et n'a par exemple pas abordé les stratégies d'utilisation de la langue (comme les stratégies de lecture).

Lors d'une première étape, l'équipe de projet s'est basée sur les recueils déjà existants et sur les manuels d'enseignement de langues étrangères pour établir un inventaire structuré de ces stratégies. Après une longue phase de travaux préparatoires, des items permettant de recueillir les compétences en stratégie d'apprentissage des langues ont été développés pour 3 sous-domaines, bien distincts les uns des autres :

1. Ampleur et diversité de l'application individuelle de stratégies d'apprentissage des langues ;
2. Connaissances des stratégies d'apprentissage et de leur adéquation ;
3. Deux aspects de la compétence métacognitive : d'une part, l'appréciation de son propre apprentissage, c'est-à-dire l'aptitude à prévoir sa propre réussite à la résolution de tâches langagières ; d'autre part, l'autocontrôle métacognitif, c'est-à-dire la capacité d'apprécier l'exactitude de ses propres activités langagières (précision de jugement).

Un type spécifique d'item a été utilisé pour le recensement de données dans chacun des trois domaines cités ci-dessus :

1. Appréciation personnelle de la fréquence de l'utilisation de stratégies concrètes (sur une échelle de Likert à 5 niveaux, de

« très rarement » à « très souvent »), comparables avec les items dans de nombreux questionnaires sur les stratégies.

2. Tâches de classement (*ranking*) : description d'une tâche d'apprentissage langagier et de trois manières différentes de procéder à sa réalisation (les trois procédés n'ayant pas le même degré de pertinence, selon les expert.e.s). La tâche de l'élève est de classer ces trois procédés selon leur degré de pertinence.
3. En matière de prédiction : réponse à la question « Peux-tu résoudre [une tâche langagière donnée] ? » (sur une échelle de Likert à quatre niveaux, de « Non, sûrement pas » à « Oui, sans problème »).
En matière de précision de jugement : « Résous [la tâche langagière donnée]. » ; puis réponse à la question « L'as-tu fait correctement ? » selon la même échelle de Likert que précédemment.
Les deux tâches apparentées sont réalisées avec un intervalle suffisant pour que les évaluations des élèves en matière de précision de jugement et de prédiction soient indépendantes l'une de l'autre. Pour l'évaluation, le degré d'exactitude de la résolution de la tâche langagière est comparé aux deux appréciations.

Comme c'est le cas dans l'ensemble du projet « Recensement des profils de développement », les instruments destinés à évaluer les compétences en stratégie d'apprentissage ont été conçus pour que les élèves les utilisent périodiquement, à plusieurs mois d'intervalle et de façon autonome. Les résultats devraient être présentés de façon à pouvoir être utilisés en classe pour faire progresser les apprentissages. Il est ainsi prévu que les résultats se rapportant aux stratégies d'apprentissage des langues soient fournis sous la forme de comptes rendus à vocation formative, tels que :

1. Appréciations personnelles de l'ampleur et de la diversité de l'utilisation des stra-

Récolte informatisée de données relatives aux compétences en langues à l'école primaire

tégies d'apprentissage des langues :

- Fréquence de l'application de chaque stratégie dans une classe, par rapport à l'application dans toutes les classes comparables (groupe de référence) ;
 - Fréquence de l'utilisation des stratégies par chaque élève (combien de fois chaque élève a-t-il choisi les catégories « très rarement » à « très souvent » de l'échelle de Likert?).
2. Tâches relatives à la connaissance des stratégies d'apprentissage des langues :
 - Données brutes de chaque élève et de classes entières concernant le classement des stratégies disponibles (et travail subséquent avec les items de test en classe) ;
 - Résultats des individus et des classes, relatifs aux connaissances concernant la pertinence des stratégies sur l'échelle de test (p.ex. dans le but de suivre les progrès dans le temps).
 3. Tâches relatives aux deux aspects de la compétence métacognitive :
 - Résultats de chaque élève concernant sa tendance à surestimer ou à sous-estimer ses propres aptitudes (ou, globalement, à se tromper sur son appréciation) ;
 - Résultats différenciés de chaque élève après une erreur d'appréciation en matière de prédiction ou de précision de jugement. Les deux résultats peuvent indiquer des forces ou des faiblesses de nature différente.

Ce sous-projet ayant demandé un travail de pionnier conséquent, les instruments ont été développés et testés au contact répété des classes. Le

niveau de développement des instruments dans les trois domaines est différent. Les appréciations personnelles relatives à l'utilisation de stratégies d'apprentissage ont été testées avec succès dans la mesure où les items étaient faciles à traiter et ont généré des profils assez distincts. De plus, une utilisation judicieuse des résultats s'avère tout à fait possible. Le classement des connaissances portant sur la pertinence des stratégies d'acquisition n'a lui fait que partiellement ses preuves lors des tests. Différentes raisons peuvent expliquer pourquoi les items n'ont pas permis d'opérer une distinction suffisante entre les élèves et pourquoi il n'a pas été possible d'établir d'échelle de compétence pour ce domaine. Enfin, les tâches relatives au contrôle métacognitif ont été testées avec succès. Le format de tâches utilisé devrait malgré tout faire l'objet d'études ultérieures ; il n'est en effet pas encore suffisamment clair dans quelle mesure les exigences concernant les tâches langagières intégrées se mêlent aux exigences concernant l'aptitude au contrôle métacognitif.

Le sous-projet de description et de recensement des compétences en stratégie d'apprentissage des langues a pu être réalisé dans un domaine complexe et surtout encore peu développé dans le secteur de l'école primaire. Il a apporté diverses contributions conceptuelles et matérielles, et a permis de tirer des conclusions importantes (notamment par l'élimination d'approches infructueuses) qui pourront servir de base pour la suite.

Pratiques d'évaluation et de *feedback* dans l'enseignement d'une langue étrangère à l'école primaire

Une partie du projet « Recensement des profils de développement », menée intégralement au Tessin, a consisté à analyser les pratiques et les besoins des enseignant.e.s de l'école primaire (*scuola elementare*) en matière d'évaluation et leur attitude par rapport à un instrument d'évaluation et de diagnostic assisté par ordinateur. Neuf membres du corps enseignant ont été interrogés dans ce but, dont cinq avaient déjà pris part au projet dans le cadre

Récolte informatisée de données relatives aux compétences en langues à l'école primaire

de différentes études sur les tâches. Les entretiens ont été réalisés à l'aide d'un guide, construit à partir des expériences tirées des essais réalisés sur les tâches, d'une analyse de la situation de l'enseignement au Tessin, du manuel didactique pertinent et de la littérature spécialisée consacrée au domaine de l'évaluation formative et des *feedbacks*. Les enseignant·e·s ont d'abord été interrogé·e·s sur les compétences et trajectoires de développement de leurs élèves comme sur leurs propres formes d'évaluation en classe. Dans un deuxième temps, il leur a été demandé de s'exprimer sur les volets qui pourraient faire partie d'un instrument d'évaluation assisté par ordinateur. Les entretiens ont été transcrits et leur contenu a été évalué au moyen d'une grille d'évaluation qui a été conçue en recourant aux méthodes de la *Grounded Theory*.

Le sondage effectué auprès de ces neuf enseignant·e·s a montré que les évaluations sommatives sont souvent basées sur des moyens d'enseignement alors que les éléments d'évaluation formative sont en général intégrés à l'enseignement de manière informelle et intuitive. Dans le domaine de la production orale, les enseignant·e·s accordent une importance particulière à la prononciation et la compréhensibilité des mots et expressions appris. Les stratégies et l'éveil aux langues (*language awareness*) ne sont thématiques que de manière marginale, notamment parce que le manuel didactique ne prévoit rien dans ce domaine. Les enseignant·e·s de l'école primaire soulignent que, tout comme le choix des contenus pédagogiques, l'évaluation a pour but principal la motivation et le bien-être des jeunes élèves.

En ce qui concerne l'utilisation périodique d'un système d'évaluation formative et diagnostique assisté par ordinateur, les entretiens ont montré un certain scepticisme quant à sa mise en œuvre en pratique : les enseignant·e·s soulignent que certaines écoles ne sont pas équipées d'un nombre suffisant d'ordinateurs et que peu d'heures sont consacrées à l'enseignement du français, et elles émettent des réserves quant à l'effort organisationnel à fournir pour enregistrer les productions orales pendant les cours. Les enseignant·e·s se demandent en outre si le gain de connaissances acquis grâce à un système d'évaluation assisté par

ordinateur compenseraient effectivement l'effort requis par son implémentation. Plusieurs enseignant·e·s préféreraient largement disposer d'un instrument que les élèves pourraient utiliser pour exercer et vérifier eux-mêmes leurs compétences orales. En revanche, les enseignant·e·s estiment que les tâches orales en elles-mêmes, développées dans le cadre du projet, seraient adaptées à l'utilisation envisagée et au contexte tessinois.

Sintesi | Italiano

Raccolta computerizzata di dati sulle competenze linguistiche nella scuola primaria

Rapporto sul progetto “Profili di sviluppo nell’apprendimento delle lingue straniere nella scuola elementare”

Sintesi

Il progetto sui "profili di sviluppo nell'apprendimento delle lingue straniere nella scuola elementare" rientra nel programma di ricerca 2012-2015 del Centro scientifico di competenza per il plurilinguismo. Attraverso l'elaborazione di strumenti idonei al rilevamento di competenze si è voluto fornire, nell'ambito dell'insegnamento delle lingue straniere nella scuola elementare (dal terzo al sesto anno), un contributo a una cultura della valutazione che sia orientata all'apprendimento e che si affidi al contempo all'ausilio di mezzi informatici. Il lavoro di elaborazione era incentrato sulle progressioni di apprendimento (*learning progressions*), un concetto sovente applicato ai programmi, alle lezioni e alle valutazioni nella didattica delle scienze naturali. Si tratta di descrivere, per un determinato campo di conoscenze teoriche e pratiche, il tipico percorso che porta i novizi a progredire fino al livello di esperti (p.es. sviluppando una comprensione adeguata della materia). Le *learning progressions* non sono sequenze di apprendimento naturali, bensì il "tipico" risultato dell'interazione di prerequisiti individuali legati a un determinato oggetto di apprendimento nel quadro di un insegnamento specifico. Esiste un reciproco condizionamento fra lo sviluppo e l'impiego di compiti per rilevare una progressione di apprendimento da un lato, e la descrizione di questa progressione dall'altro. Idealmente, quando l'andamento di una simile progressione è sufficientemente noto, si possono individuare – come su una mappa stradale – le prestazioni dei discenti sul loro percorso di apprendimento. Diventa quindi possibile ottenere *feedback* e *feedforward* (indicazioni per sviluppi futuri) formativi ottimali.

Nell'ambito delle lingue, l'impiego di *learning progressions* è ancora quasi inesistente. Il progetto "Dynamic Language Learning Progressions" (DLLP)¹ per l'inglese come lingua d'insegnamento è una delle rare eccezioni e mira a sviluppare competenze linguistico-funzionali, in particolare la capacità di fornire spiegazioni legate ai compiti scolastici.

Nel nostro progetto sui profili di sviluppo, la priorità è data allo sviluppo di compiti esemplari piuttosto che a una descrizione perfetta delle pro-

gressioni. Uno dei principali motivi è l'ambizioso obiettivo di raccogliere dati sul livello dei discenti attraverso un sistema basato su internet e di generare, dentro tale sistema, un riscontro utile per migliorare l'insegnamento e l'apprendimento. Questo aspetto è stato messo in rilievo sin dall'inizio affinché nel corso della durata del progetto si potesse avanzare il più possibile sul piano tecnico. Un sistema di valutazione online ha il vantaggio di poter essere ottimizzato anche una volta in uso, in modo da migliorare la qualità dei compiti e approfondire le informazioni rilevate sulle competenze dei discenti.

I contenuti studiati nell'ambito del progetto erano incentrati sulla produzione e sull'interazione orali, e sulle strategie di apprendimento linguistico. Si tratta di una scelta fatta scientemente e volta ad approfondire aspetti che solitamente, in particolar modo per quanto riguarda le valutazioni, tendono a essere trascurati.

Per l'attuazione tecnica del progetto è stato decisivo il fatto che nel 2012 la Conferenza svizzera dei direttori cantonali della pubblica educazione (CDPE) avesse deciso di creare una banca dati dei compiti da mettere a disposizione come ambiente per le misurazioni delle prestazioni nei Cantoni, ma anche per altri progetti di valutazione scolastica. Attraverso il progetto della banca dati, è nata una collaborazione tra il Centro scientifico di competenza per il plurilinguismo e l'Istituto Leibniz di ricerca e informazione pedagogica (DIPF), dalla quale è emersa una versione elaborata del software *CBA Item Builder* per il rilevamento e la valutazione delle competenze orali.

Il progetto sui profili di sviluppo non è stato esente da difficoltà e notevoli sfide. I cinque contributi del rapporto offrono una panoramica sui vari lavori svolti nel quadro del progetto.

- La parte II "Valutazione al computer delle competenze orali" riguarda in particolare l'attuazione tecnica di un sistema online impiegabile per la valutazione formativo-diagnostica, segnatamente nel campo delle competenze linguistiche orali. Si è trattato di elaborare compiti di facile utenza da rendere accessibili su internet

Raccolta computerizzata di dati sulle competenze linguistiche nella scuola primaria

e un modulo per una valutazione delle registrazioni audio della produzione orale dei discenti basata su determinati criteri.

- La parte III "Competenze nel parlato" espone l'elaborazione e la validazione a tappe dei compiti che consentono di ottenere informazioni diagnostiche su aspetti rilevanti delle competenze orali degli allievi di scuola elementare che hanno partecipato alla prova.
- La parte IV "Lingua d'insegnamento" è dedicata a un sottoprogetto volto a identificare, partendo dall'analisi dei materiali didattici e da un questionario sottoposto ai docenti, mezzi linguistici utili per la comunicazione orale in classe. Il sondaggio ha inoltre consentito di capire quanto gli allievi dei vari anni padroneggino, a livello di ricezione e di produzione, tali mezzi linguistici.
- La parte V "Competenze strategiche nell'apprendimento delle lingue" presenta lo sviluppo di strumenti per il rilevamento sia della frequenza d'uso da parte dei discenti di determinate strategie di apprendimento sia delle competenze di cui essi dispongono in alcuni sottoambiti di queste strategie. Per svolgere il lavoro, si è innanzitutto dovuto creare un inventario strutturato delle strategie.
- La parte VI "Pratiche di valutazione e *feedback* nell'insegnamento delle lingue straniere nella scuola elementare" illustra i risultati di un'analisi qualitativa svolta in Ticino sulle pratiche valutative dei docenti di scuola elementare per quanto riguarda il francese, nonché l'eventuale compatibilità di queste pratiche con l'introduzione di uno strumento informatizzato per la valutazione formativo-diagnostica.

Il progetto ha permesso di compiere una serie di passi importanti in vista della creazione di un

sistema di valutazione formativo-diagnostica online. Alcuni degli aspetti su cui si è lavorato si sono già rivelati utili in altri ambiti, segnatamente nell'elaborazione di compiti informatizzati per i test.

Valutazione al computer delle competenze orali

Una parte delle risorse del progetto è stata impiegata per studiare e perfezionare un sistema di test informatizzato finalizzato all'approntamento di un metodo di valutazione formativo-diagnostico online. Nel 2012, quando la CDPE ha reso nota la creazione di una cosiddetta *banca dati dei compiti*, il progetto si era concentrato sullo sviluppo di un sistema di valutazione che rientrasse tecnicamente in questa piattaforma di test. Dopo una fase esplorativa con Microsoft PowerPoint e programmi standard di registrazione audio, tutti i compiti elaborati nel quadro del progetto sono stati trasferiti nel software *CBA Item Builder* con la relativa applicazione basata su server EE4CBA. Entrambi sono stati messi a disposizione dal DIPF con il quale la CDPE ha collaborato per la banca dati dei compiti. Con questa soluzione informatica, le scuole hanno bisogno soltanto di computer, una connessione internet con browser aggiornati e auricolari.

Attraverso il DIPF, è stato commissionato lo sviluppo del programma, affinché potesse essere impiegato anche per il rilevamento delle capacità di espressione orale. Vi sono stati aggiunti un modulo di registrazione audio online e un modulo di valutazione. Quest'ultimo consente di gestire le registrazioni (ma anche testi scritti) e di analizzarle manualmente in base a determinati criteri. Le valutazioni vengono poi salvate analogamente alle risposte nel caso degli item con formati chiusi (p.es. con le domande a risposte multiple).

Non è stato possibile sviluppare tutte le prestazioni informatiche necessarie per ottenere un sistema di valutazione formativo-diagnostico ottimale. Mancano in particolare un modulo perfettamente integrato per la gestione di persone e compiti, nonché un modulo di analisi automatica che associ le risposte degli allievi e le valutazioni

1 | <https://www.dllp.org/>

Raccolta computerizzata di dati sulle competenze linguistiche nella scuola primaria

manuali a scale, così da generare rapporti ossia *feedback* e *feedforward*.

Il lavoro principale del progetto è consistito nello sviluppare e valutare formati e contenuti per compiti che siano adatti agli scolari più giovani, affinché questi possano svolgerli in maniera autonoma.

Le informazioni necessarie sono emerse nel corso di più fasi, durante le quali si è riflettuto in particolare sulle questioni seguenti: dimestichezza dei bambini più piccoli con gli strumenti informatici; idoneità di varie interfacce utente, comandi (simboli) e informazioni (testi, immagini); comprensibilità delle consegne (sia nella lingua d'insegnamento sia in quella d'arrivo); possibilità tecnica e logistica di fare uso dei compiti in classe; difficoltà e adeguatezza dei formati e dei contenuti dei compiti. Un'importante constatazione riguarda il fatto che anche i partecipanti più giovani non sembrano avere alcuna difficoltà a lavorare al computer. Già nelle prime fasi, tuttavia, si è osservato che schermate con testi anche solo leggermente più consistenti, ma anche formati che richiedono più passi oppure capacità cognitive un po' più avanzate generano parecchia confusione e non si prestano quindi a un lavoro di valutazione. Pertanto, sono stati elaborati formati di compiti per i quali bastano pochissime spiegazioni e che possono essere risolti senza la necessità di capire un testo scritto, anche perché lo scopo è la valutazione delle competenze orali. I risultati di questa analisi degli strumenti di valutazione informatici idonei per allievi di scuola elementare (nonché dei formati di compiti e dei simboli ossia delle immagini in essi utilizzati) potranno servire per lo sviluppo di ulteriori compiti e di test.

Competenze nell'ambito del parlato

Il sottoprogetto "Competenze nell'ambito del parlato" è incentrato sullo sviluppo di strumenti informatizzati (compiti) per rilevare le competenze ancora rudimentali degli allievi di scuola elementare nel campo della produzione e dell'interazione orali. Sia la documentazione relativa al Quadro comune europeo di riferimento per la conoscenza

delle lingue (QCER) sia diversi documenti relativi ai test linguistici forniscono indicazioni solo sommarie riguardo agli stadi di sviluppo di competenze linguistiche o all'estensione delle stesse nei livelli A del QCER. Si riscontrano invece ricorrenti riferimenti a singole parole e a sintagmi (*chunks*) appresi da impiegare in maniera più o meno flessibile in contesti comunicativi noti. Considerando allievi di scuola elementare, è evidente che le competenze sviluppate dipendono da quanto viene trattato nel manuale e durante le lezioni. Le prime fasi di *testing* dei compiti prodotti nell'ambito del presente progetto e svolti in tre diversi contesti (nelle regioni del piano di studi Passepartout,² nella Svizzera orientale e in Ticino) hanno evidenziato tale dipendenza. Presto l'obiettivo di rilevare dati su questo insieme eterogeneo di contesti scolastici si è rivelato eccessivamente limitante poiché imponeva di concepire compiti adattati a ognuno dei tre contesti. Nella fase finale, perciò, il sottoprogetto relativo alla lingua parlata è stato svolto solo nella regione del piano di studi Passepartout.

Il sottoprogetto, suddiviso in tre tappe, ha consentito un avvicinamento all'obiettivo di rilevare e analizzare le competenze orali degli allievi di scuola elementare.

Nella prima tappa, i compiti per il rilevamento delle competenze si sono basati sui contenuti e sui tipi di compiti (comunicativi) ricorrenti nel materiale didattico impiegato. I compiti, assegnati in un primo momento su carta, sono stati testati *in loco* e valutati sia per verificarne il funzionamento, sia per rilevare e valutare le competenze dei giovani apprendenti. Ispirati al manuale in uso, i compiti sono stati spesso redatti in forma aperta, in modo tale da consentire il rilevamento di un ampio ventaglio di competenze.

La fase di sperimentazione ha tuttavia evidenziato che tali compiti non sono in grado di valutare in modo adeguato le competenze orali ancora rudimentali di molti allievi di scuola elementare. Questi ultimi, infatti, senza l'aiuto di terzi, sono

2 | I sei Cantoni lungo il confine linguistico tedesco-francese (Berna, Basilea-Campagna, Basilea-Città, Friburgo, Soletta e Vallese).

Raccolta computerizzata di dati sulle competenze linguistiche nella scuola primaria

stati a malapena in grado di affrontare una parte dei *task* assegnati. Ciò dipende sia dai requisiti cognitivi necessari per affrontare i compiti legati a determinate situazioni comunicative, sia al livello di *output* richiesto che, benché in sé già basso, si è rivelato eccessivo.

Nella seconda fase, ci si è concentrati sulle risorse linguistiche e linguistico-funzionali di un uso comunicativo (futuro) della lingua piuttosto che sulla capacità di agire nell'ambito di compiti comunicativi. Pur ricorrendo a formati semplici, non si è tuttavia rinunciato completamente a compiti relativi a quest'ambito. La scelta dei costrutti su cui si è deciso di focalizzare l'attenzione è stata compiuta sulla base di un'analisi dettagliata delle componenti linguistico-comunicative (i cosiddetti "mezzi linguistici") dei compiti che, nei manuali, sono volti a stimolare e valutare le competenze orali.

Nelle sperimentazioni della seconda fase sono stati selezionati compiti incentrati o sulle risorse linguistiche, o sulle competenze linguistico-funzionali o sulle competenze comunicative. Per la prima volta, i bambini hanno lavorato al computer. L'esperienza ha mostrato che anche i più piccoli dispongono di sufficienti nozioni e predisposizione per svolgere i compiti al computer. Gli aspetti che si sono rivelati più delicati sono stati i compiti scritti di una certa lunghezza, gli esempi a schermo e soprattutto l'impiego della lingua d'arrivo nelle consegne. In questa fase di sperimentazione è stato inoltre possibile far luce su diversi aspetti connessi alle competenze. I problemi di pronuncia si sono dimostrati poco rilevanti, mentre delle attività di lettura ad alta voce possono mettere a dura prova i giovani apprendenti. La conoscenza concreta di parole ed espressioni sembra avere un'importanza fondamentale, anche e soprattutto a livello di consapevolezza dei discenti stessi. Nei compiti di sintassi, per esempio, succedeva che i bambini non tentassero nemmeno di trovare una soluzione se non capivano il lessico delle frasi. Eventuali nozioni di sintassi non sono quindi emerse. Anche compiti più articolati che richiedevano di comunicare oralmente si sono rivelati problematici: anche se formulate nella lingua di insegnamento, le consegne leggermente più complesse non venivano comprese a sufficienza e veni-

vano quindi seguite solo in parte. Le produzioni orali autonome nella lingua d'arrivo si sono rivelate perlopiù molto modeste.

Nell'ultima tappa del progetto, si è lavorato esclusivamente con la regione del piano di studio Passepartout e quindi con gli allievi che imparano il francese con il manuale *Mille feuilles*. Questa volta erano tuttavia coinvolti anche allievi del quinto e sesto anno. Gli ambiti di studio sono stati ulteriormente ristretti rispetto alla tappa precedente, limitandosi a considerare da una parte due ambito comunicativo-operativi – il primo relativo alla comunicazione in classe (*classroom language*) e il secondo allo scambio di informazioni personali – e dall'altra una competenza linguistica parziale: la competenza lessicale. È stato considerato un nuovo costrutto parziale della competenza orale, ossia la comprensione di lingua parlata, per tentare di rendere più visibili le competenze di base in ambito orale. I compiti in questi campi sono stati ulteriormente calibrati in funzione del target sia per quanto riguarda la struttura che la difficoltà e l'aspetto grafico. Quale base per lo sviluppo attuale ed eventualmente futuro dei compiti di *Mille feuilles*, è stato creato un *corpus* di tutti i testi scritti e parlati del manuale. Si è inoltre compilato un catalogo delle strutture linguistiche utilizzate in classe, basandosi sia su un'analisi di diversi materiali didattici, sia sui risultati di un ampio sondaggio condotto tra gli insegnanti. Il *corpus* e il catalogo sono stati impiegati per sviluppare una serie di esempi di compiti sull'uso orale della lingua (ricezione, interazione, produzione) in ambiti rilevanti. Si tratta di compiti concisi, visualizzabili interamente su una schermata di computer. Allo scopo di renderli comprensibili intuitivamente sono spesso state inserite illustrazioni. Si è mantenuto un solo formato leggermente più complesso, quello del dialogo per la verifica della capacità di interazione orale, già testato in precedenza. Un dialogo del genere si estende su più pagine e consiste in un alternarsi di ascolto e di produzione orale in interazione con un personaggio fittizio.

Le prove pratiche si sono svolte in più tappe. Sono stati valutati non solo i compiti, ma anche la distribuzione via server di test per classi intere attraverso l'infrastruttura delle scuole coinvolte.

Raccolta computerizzata di dati sulle competenze linguistiche nella scuola primaria

Le registrazioni di questa fase hanno consentito di testare la valutazione analitica (manuale) delle prestazioni orali degli allievi attraverso le maschere di valutazione previste in *Item Builder*. I risultati delle sperimentazioni sono complessivamente positivi. La distribuzione dei test in classi intere ha funzionato sia dal punto di vista tecnologico che logistico, e i compiti sono stati elaborati come previsto. È quindi possibile considerare questi formati sufficientemente testati da poter essere impiegati in altri strumenti di verifica delle competenze. Per quanto riguarda i compiti ricettivi, i contenuti testati si sono rivelati un po' troppo semplici per raccomandarne un uso formativo-diagnostico. Nel complesso, i compiti produttivi hanno funzionato bene anche con allievi del quinto e sesto anno che partecipavano per la prima volta. Questi compiti hanno infatti condotto agli esiti auspicati in termini di prestazioni. Il tentativo di sviluppare una modalità di valutazione analitica della produzione orale che sia sufficientemente trasparente da poter essere utilizzata in modo pressoché intuitivo da parte degli insegnanti ha dato esiti molto promettenti. Mancano tuttavia esperienze pratiche nel quadro di una sperimentazione più ampia svolta dai docenti.

Lingua d'insegnamento

Quando gli insegnanti e i discenti utilizzano la lingua straniera in classe per discutere le consegne, i compiti o i rituali quotidiani come i saluti prima di cominciare la lezione e al momento di congedarsi, la materia "lingua straniera" diventa un autentico mezzo di comunicazione. Nel campo della didattica delle lingue straniere, si parla spesso in inglese di *classroom language* – la lingua di comunicazione in classe. Oggi, il suo uso è annoverato tra gli obiettivi didattici nei piani di insegnamento svizzeri e diversi sussidi didattici, come *Mille feuilles*, trattano esplicitamente l'argomento. Poiché l'insegnamento delle lingue è un ambito operativo in cui si può comunicare in modo autentico sin dall'inizio, la *classroom language* è stata inserita nel progetto "Profili di sviluppo" come settore per il quale elaborare compiti. Per prima cosa, si è alle-

stito un catalogo di enunciati potenzialmente rilevanti per la comunicazione orale in classe sulla base dei due manuali del progetto – *Mille feuilles* per i cantoni lungo il confine linguistico tedesco-francese e *Alex et Zoé* per il Ticino – come pure di materiali di perfezionamento. In seguito, degli esperti hanno verificato la selezione degli enunciati. In una seconda fase, è stato chiesto a insegnanti di francese come lingua straniera nella scuola elementare (n=74) di rispondere a un questionario online per valutare il livello di padronanza ricettiva e produttiva di tali enunciati da parte dei loro allievi (dalla terza alla sesta classe). Gli insegnanti hanno inoltre proposto altri enunciati rilevanti che non erano inclusi nel questionario. Dall'analisi delle risposte, si possono trarre le seguenti conclusioni:

1. In generale, gli allievi meno giovani padroneggiano meglio gli enunciati e ne conoscono anche altri.
2. La maggior parte degli enunciati viene dapprima padroneggiata in modo ricettivo e solo più tardi a livello produttivo.
3. Nel secondo semestre della terza, dopo mezz'anno di lezioni di francese, gli allievi capiscono un esiguo numero di enunciati molto frequenti (p.es. "Bonjour."; "C'est faux."; "Ouvre la fenêtre.") e talvolta sono in grado di produrli.
4. Gli allievi non conoscono gli enunciati che non compaiono in *Mille feuilles*.
5. Complessivamente, gli allievi capiscono bene le consegne. Essi conoscono solo in parte gli enunciati funzionali che si usano per i giochi collettivi (p.es. "Jette le dé." nei giochi di società), anche se in realtà in *Mille feuilles* questo vocabolario viene costruito in modo mirato.

Raccolta computerizzata di dati sulle competenze linguistiche nella scuola primaria

I risultati dettagliati del sondaggio sono accessibili dal sito del Centro di competenza come allegato elettronico al rapporto (<https://bit.ly/2IX0eL5>). Possono anche essere impiegati come catalogo di espressioni finalizzate alla comunicazione in classe e illustrano la distribuzione dei mezzi discorsivi sui quattro anni scolastici. Il documento include anche i mezzi discorsivi che gli insegnanti hanno aggiunto al questionario reputandoli rilevanti.

L'ambito operativo della lezione e la relativa lingua sono stati la fonte per l'elaborazione di compiti nel sottoprogetto sul parlato. Gli enunciati selezionati a tale scopo fanno riferimento ai due ambiti delle istruzioni e consegne, nonché del lavoro e dei giochi di gruppo. Sono stati creati e testati sia compiti produttivi sia compiti ricettivi. Le espressioni sono state contestualizzate attraverso l'uso di una o più immagini che rappresentassero l'azione in questione, così da offrire un aggancio comunicativo.

Competenze strategiche nell'apprendimento delle lingue

Due sono state le questioni fondamentali approfondite nel progetto "Profili di sviluppo": il lavoro nell'ambito del parlato nonché la descrizione e il rilevamento delle competenze strategiche nell'apprendimento delle lingue da parte degli allievi di scuola elementare. Le strategie di apprendimento delle lingue rientrano fra le cosiddette competenze trasversali, di particolare importanza nella didattica plurilingue. La Conferenza svizzera dei direttori cantonali della pubblica educazione nonché i piani d'insegnamento delle lingue straniere nella Svizzera tedesca annoverano fra le competenze di base da acquisire nella scuola dell'obbligo lo sviluppo di competenze metodiche nonché di strategie di apprendimento. Ciò può sorprendere, poiché nella letteratura specializzata si legge che il rapporto tra l'impiego di strategie e il successo dell'apprendimento è tutt'altro che semplice. L'efficacia di determinate strategie sembra dipendere non da ultimo dalle competenze già acquisite, dalla situazione specifica e dal contesto. Va tuttavia evidenziato che le strategie di

apprendimento possono essere insegnate e che è empiricamente provato un effetto positivo dell'insegnamento di strategie sulla competenza nell'uso della lingua.

Il sottoprogetto V si è concentrato sullo studio delle strategie di apprendimento linguistico propriamente dette (quindi strategie legate a operazioni di apprendimento specifiche, come quelle impiegate per memorizzare vocaboli) e non ad altri aspetti quali, per esempio, le strategie d'uso della lingua (p.es. strategie di lettura).

In una prima fase di lavoro, il team responsabile del progetto ha allestito un inventario strutturato di queste strategie sulla base di cataloghi esistenti e di manuali per l'apprendimento delle lingue.

Dopo un considerevole lavoro preliminare, sono stati sviluppati degli item per il rilevamento di competenze strategiche nell'apprendimento delle lingue tenendo conto di tre ambiti ben distinti:

1. la quantità e la varietà dell'impiego individuale di strategie di apprendimento delle lingue;
2. le conoscenze sulle strategie di apprendimento delle lingue e sulla loro idoneità;
3. due aspetti relativi alle capacità metacognitive: da un lato la valutazione del proprio livello di apprendimento nel senso operativo della capacità di prevedere il proprio successo nei compiti linguistici, dall'altro l'autosorveglianza metacognitiva nel senso operativo della capacità di valutare la correttezza della propria produzione linguistica (giudizio di sicurezza).

Per il rilevamento dei dati per ciascun ambito è stato utilizzato di volta in volta un tipo specifico di item.

Raccolta computerizzata di dati sulle competenze linguistiche nella scuola primaria

Per l'ambito 1

Autodichiarazioni sulla frequenza dell'impiego di strategie concrete (scala Likert a cinque livelli, da "Molto raramente" a "Molto spesso"), analogamente agli item di numerosi questionari sulle strategie.

Per l'ambito 2

Compiti di classificazione: un compito di apprendimento linguistico con tre possibili approcci che, secondo gli esperti, non sono però tutti ugualmente idonei. Al discente viene chiesto di metterli in ordine dal più al meno idoneo.

Per l'ambito 3

Per la previsione: formulazione della domanda "Saresti capace di risolvere [un determinato compito linguistico]?" e risposta in base a una scala di Likert a quattro livelli (da "Sicuramente no" a "Sicuramente sì").

Per il giudizio di sicurezza che segue: "Risolvi [il compito linguistico definito].", poi risposta alla domanda "Hai risolto il compito correttamente?" con la stessa scala di Likert.

Un intervallo di tempo separa di volta in volta lo svolgimento di questi due compiti da parte dei discenti, una pausa volta a impedire che la previsione influisca sul giudizio di sicurezza. La valutazione avviene paragonando la correttezza della soluzione del compito linguistico con le due autovalutazioni.

Come nel resto del progetto "Profili di sviluppo", gli strumenti relativi alle competenze strategiche nell'apprendimento delle lingue sono stati concepiti per essere utilizzati dagli allievi periodicamente e in modo autonomo nel corso di più mesi. I risultati dovevano essere analizzati in modo da poterli usare in classe mirando a un apprendimento continuo. Per quanto riguarda i risultati delle strategie di apprendimento linguistico, sono state considerate le seguenti possibilità con potenziale formativo.

1. Autodichiarazioni sulla quantità e la varietà dell'impiego di strategie di apprendimento delle lingue:

- frequenza dell'uso delle singole strategie in una classe rispetto al loro uso in tutte le altre classi comparabili (gruppo di riferimento);
- frequenza dell'utilizzo delle strategie da parte dei singoli discenti (quanto spesso hanno scelto le varie risposte della scala Likert).

2. Compiti sulle conoscenze relative alle strategie di apprendimento delle lingue:

- dati grezzi di singoli allievi e intere classi riguardo alla classificazione delle strategie selezionabili (e successivo lavoro con gli item di test durante l'insegnamento);
- risultati individuali e di classe sulla scala di test per le conoscenze sull'idoneità delle strategie (p.es. per monitorare i progressi).

3. Compiti sui due aspetti relativi alle capacità metacognitive:

- risultati dei singoli allievi per quanto riguarda la tendenza a sovrastimare o sottostimare le proprie capacità;
- risultati dei singoli allievi suddivisi secondo l'entità dell'errore di stima nelle previsioni e nei giudizi di sicurezza. I due risultati possono fornire indicazioni su vari punti di forza e debolezza.

Tale progetto comportava molto lavoro pionieristico, per cui gli strumenti sono stati sviluppati e testati in contatti ripetuti con le classi. Tali strumenti non hanno raggiunto lo stesso stadio di sviluppo nei tre ambiti. Le autodichiarazioni

Raccolta computerizzata di dati sulle competenze linguistiche nella scuola primaria

sull'impiego delle strategie sono state testate con successo: gli item sono stati elaborati con facilità e ne sono risultati profili diversi per quanto riguarda l'uso delle strategie. Si prospettano pertanto buone possibilità di fruizione dei risultati ottenuti. I compiti di classificazione sulle conoscenze relative all'idoneità delle strategie di apprendimento si sono invece rivelati essere di utilità limitata. Varie ragioni hanno fatto sì che gli item indicassero solo in modo poco soddisfacente differenze tra i discenti e che nel quadro del progetto non fosse possibile descrivere una scala di competenze in tale ambito. I compiti sulla sorveglianza metacognitiva sono invece stati testati con successo. Il formato dei compiti andrebbe comunque ancora studiato, poiché non è sufficientemente chiaro fino a che punto quanto richiesto dal compito linguistico si amalgami con quanto richiesto a livello di capacità di sorveglianza metacognitiva.

Il sottoprogetto di descrizione e rilevamento delle competenze strategiche nell'apprendimento delle lingue ha fornito contributi concezionali e materiali di vario genere in un campo complesso e ancora poco studiato, in particolare per quanto riguarda le scuole elementari. Inoltre, anche grazie all'abbandono di approcci dimostratisi inefficaci sono emerse importanti informazioni che potranno essere utili per il lavoro futuro.

Pratiche di valutazione e *feedback* nell'insegnamento delle lingue straniere nella scuola elementare

Una parte del progetto sui profili di sviluppo, interamente condotta in Ticino, era dedicata a un'analisi qualitativa delle pratiche e delle esigenze di valutazione degli insegnanti di scuola elementare, come pure al loro punto di vista in merito a uno strumento diagnostico informatizzato. A tale scopo, abbiamo intervistato nove insegnanti, cinque dei quali avevano già partecipato a varie prove volte a testare i compiti per questo progetto. Le interviste si sono svolte seguendo linee guida elaborate sulla base delle esperienze acquisite con un test dei compiti, di un'analisi della situazione

dell'insegnamento in Ticino, del materiale didattico e della letteratura sulle valutazioni e sui riscontri formativi. Gli insegnanti sono stati dapprima interrogati sulle competenze e sui percorsi di sviluppo dei discenti come pure sulle forme di valutazione da loro adottate in classe. In un secondo tempo, è stata chiesta la loro opinione su alcuni aspetti di uno strumento di valutazione informatizzato. Le interviste sono state trascritte, poi valutate mediante una griglia di analisi qualitativa, traendo spunto dalla *Grounded Theory*.

Per la maggior parte delle nove persone intervistate, la valutazione sommativa avviene spesso sulla base di materiali didattici, mentre elementi di valutazione formativa vengono integrati nelle lezioni in modo informale e intuitivo. In materia di produzione orale, gli insegnanti prestano particolare attenzione alla pronuncia e alla comprensibilità dei vocaboli e delle espressioni imparati. Strategie e *language awareness* (consapevolezza linguistica) vengono menzionate solo superficialmente, anche perché il manuale utilizzato non prevede nulla in tale ambito. Le nove persone interpellate sottolineano che la valutazione, analogamente a quanto accade per la scelta dei contenuti dell'insegnamento, avviene soprattutto tenendo conto della motivazione e delle esigenze degli allievi.

Per quanto riguarda l'impiego a intervalli regolari di uno strumento di valutazione informatizzato per scopi formativo-diagnostici, gli intervistati si sono mostrati piuttosto scettici: alla quantità insufficiente di computer a scuola si aggiungono il numero molto ridotto di ore di francese e l'onere organizzativo per il rilevamento della produzione orale durante le lezioni. Gli insegnanti temono che l'impegno richiesto sia sproporzionato rispetto alla quantità di informazioni da ottenere. Molti preferirebbero disporre di uno strumento che gli allievi possano utilizzare autonomamente allo scopo di esercitarsi e verificare le competenze orali acquisite. I compiti per la valutazione della produzione orale elaborati nel quadro del progetto sembrano comunque essere adeguati allo scopo e al contesto ticinese.

Summary | English

Computer-based assessment of foreign-language competences in primary school teaching

Report on the project “Mapping developmental profiles
in early language learning at school”

Summary

The project “Mapping developmental profiles in early language learning at school” is part of the research programme 2012-2015 of the Research Centre on Multilingualism. The overall objective of the project was to develop adequate instruments to assess language competence and thus contribute to a learning-oriented and computer-assisted assessment culture in foreign language instruction at primary schools (third to sixth class). To develop these instruments, the general work processes were guided by the notion of learning progressions, which are frequently used as the basis for developing curricula, teaching and conducting assessments in (most commonly) the natural sciences. A learning progression in a given area of knowledge and skills charts the typical learning path from the status of novice to that of expert (e.g. with reference to the formation of an appropriate understanding of the concept of matter in Physics). Learning progressions are not to be understood as natural processes of knowledge acquisition; rather, they are a “typical” result of the interplay of the aptitude of an individual, the subject at hand and the teaching method chosen. The development and implementation of tasks to measure learning progressions and the charting of learning progressions are interdependent. Once the course of a learning progression is adequately known, the performance of learners can be located on this learning progression, ideally much like using the map of a city. This then enables optimal formative feedback and feedforward to be defined.

In the area of language teaching and learning, very few learning progressions have been developed. A rare exception is the project “Dynamic Language Learning Progressions (DLLP)”,¹ which is situated in the context of English as a school language. DLLP focuses on the development of functional language competence, especially the ability to give explanations on school-related tasks.

In the project “Mapping developmental profiles”, priority was given to the development of model tasks rather than to achieving a well-rounded charting of progressions. A major factor

for this decision is the ambitious goal of using a system based on an internet server to compile data on the students’ level of performance and, within the same system, to generate formative feedback that is useful for further learning and teaching. In order to make as many technological advances as possible over the course of the project, it was important that IT aspects were given central importance from the very outset of the project. Online assessment systems are partly so attractive because they can be further developed when they are already in use – because new data are continually generated. This leads to higher-quality tasks and to better understanding of the students’ actual competences.

Regarding the various content areas that were explored in the project, the focus was on oral productions and interactions as well as on language learning strategies. A conscious decision was taken to concentrate on such demanding content because the related skill sets – and especially the assessment thereof – tend to be neglected.

For the actual technical implementation of the project, a decisive factor was the decision of the Conference of Cantonal Ministers of Education (EDK) in 2012 to develop a task database designed to function as a setting for large-scale assessments in the Swiss cantons as well as for other assessment and evaluation projects in the education sector. In the scope of developing tasks for the database, the RCM and the Leibniz Institute for Research and Information in Education (DIPF) cooperated in order to develop additions to the software CBA Item Builder that can now be used to collect and rate oral performance as well.

The demands and challenges of the project “Mapping developmental profiles” were considerable. The reports from the five sub-projects provide an overview of the various work conducted in the scope of the project.

- Part II “Computer-based assessment of oral language skills” primarily dealt with the technical implementation of a server-based system that can be used for formative diagnostic assessment, especially of oral language skills. The focus

was on the development of user-friendly tasks that are made available via an internet browser and on the development of a module for the criteria-based assessment (by the teacher) of students’ oral productions that had been recorded earlier.

- Part III “Competences in spoken language” reports on the step-by-step development and validation of tasks able to generate diagnostic data on relevant partial areas of oral language skills of the primary school children in the study.
- Part IV “Classroom language” covers the sub-project that took analyses of textbooks and a teacher survey as the basis for identifying the linguistic resources that are important in oral communication in the classroom. The results of the teacher survey were also used to find out how well students in the different classes have attained receptive and productive mastery of these linguistic resources.
- Part V “Competence in language learning strategies” describes the development of instruments for two purposes: a) finding out how frequently students use different language learning strategies; b) measuring learners’ competences in selected sub-areas of language learning strategies. In a first step, a structured inventory of language learning strategies was compiled as a basis for this study.
- Part VI “Assessment and feedback practices in foreign language instruction at primary schools” presents the findings of a qualitative study, conducted in the Canton Ticino, on assessment practices of primary school teachers of French as well as on the potential compatibility of these practices with the introduction of a server-based setting for a formative diagnostic assessment.

In the scope of the overall project, a great deal of important work was done in the area of a server-based formative diagnostic assessment system. Some products and findings from this project have already been implemented in other contexts, in particular, in the development of computer-based tests.

Computer-based assessment of oral language skills

A part of the project funding was used to investigate and further develop a server-based assessment system to enable the implementation of a formative diagnostic assessment system that can be applied in a broad range of settings. When, in 2012, the EDK announced the development of a task database (*Aufgabendatenbank*), the project focused on the development of an assessment system that can be integrated into the IT framework of this testing platform. Following an exploratory phase with Microsoft PowerPoint and standard recording software, all tasks developed in the project were implemented using the software CBA Item Builder and delivered by the accompanying server application EE4CBA. Both tools were supplied by the DIPF, which was collaborating with EDK on the task database. The software solution is designed so that schools require only computers with Internet connection, up-to-date Internet browsers and headsets.

To enable the existing software to collect oral productions as well, a corresponding project was contracted via DIPF. On the one hand, a module for sound recordings via a Web browser was added, on the other, a rating module. The rating module makes it possible to manage existing recordings (as well as written texts) and to manually rate them on the basis of pre-defined criteria. The ratings of the performances are then saved in the same way as the answers to selected-answer items (e.g. multiple choice).

Not all IT features necessary to guarantee a functioning formative diagnostic assessment system could be developed. Particularly lacking is a fully integrated module to manage persons and

1 | <https://www.dllp.org/>

tasks, as in an automatic evaluation module that relates student answers and manual ratings to scales and that can use this data to generate reports as well as feedback and feedforward (indications regarding the next developmental steps) on behalf of learners and teachers.

The major undertaking in the project lay in the development and testing of task formats and task content suitable for the target group of young schoolchildren; it was particularly important that the children are able to solve the tasks independently.

To attain the necessary information, the various test phases focused on the following points: familiarity of the youngest schoolchildren (in particular) with computer hardware and software; suitability of different user interfaces, control elements (symbols) and input (texts, images); clarity of the instructions (in the language of instruction and the target language); technical and logistic feasibility of using the tasks in class; degree of difficulty and suitability of task type and task content. A key finding is that even the youngest participants had no difficulties in using a computer. Moreover, pre-piloting immediately revealed that only slightly text-heavy interfaces and multi-step or cognitively demanding task formats led to considerable confusion and are therefore not conducive to valid assessment. To avoid such scenarios, task formats were developed that require only minimal explanation and do not demand that students first understand a written text (in an assessment of oral language skills). The various findings on computer-based assessment tools suitable for primary schoolchildren as well as the specific task formats, including images and symbols, are now available for further task and test development.

Competences in spoken language

The development of computer-based instruments (tasks) to map the as-yet rudimentary oral language skills (spoken production and interaction) of primary school children is the focus of the sub-project “Competences in spoken interaction”. Documents from the Common European Framework

of Reference and from the field of language testing provide only rough indications of the characteristics and development of skills within the CEFR A levels. One typical characteristic appears to be the dependency on a few words and fixed phrases (chunks) that have been learned and that can be more or less flexibly applied in order to perform elementary communicative tasks in specific, familiar settings. In the case of primary school students, this implies a strong dependency on the content treated in the teaching materials and in the lessons. Initial pre-piloting of tasks in three settings (in the region using the Passepartout curriculum,² Eastern Switzerland, and Ticino) clearly demonstrated this dependency. In the end, the goal of mapping cross-contextually relevant lines of development proved to be too limiting to task development; as a consequence, the final stage of this sub-project concentrated on the region using the Passepartout curriculum.

The sub-project was structured in three larger stages to achieve the goal of mapping and diagnosing the oral language skills of primary school children in the first foreign language taught at school.

In the first stage, the tasks to map the current level of skills were modelled on the content and type of (communicative) tasks that are common in the various textbooks used in the regions of the study. These still paper-based tasks were pre-piloted on site; the results and experiences were then analysed in detail in terms of how the tasks functioned and how well the young learners did. These tasks, which were designed using the teaching materials, were generally framed so openly that a broad spectrum of competences could be mapped.

The pre-piloting revealed, however, that such tasks are not able to (efficiently) map the oral language skills of many primary school students because the children were hardly able to even begin solving the tasks without outside assistance. This is partly due to the cognitive demands

2 | The six Swiss cantons located along the German-French language border (Bern, Basel-Landschaft, Basel-Stadt, Fribourg, Solothurn and Valais).

arising from situating the tasks in communicative situations, but also due to the demands placed on the communicative skills in the foreign language although these demands were often very modest.

In the second stage of the project, the linguistic and communicative *resources* for (later) communicative language use, rather than actual language use in communication tasks, were brought more clearly into focus. However, communicative tasks, in simple forms, continued to be used. A detailed analysis of the linguistic-communicative components of the tasks designed to promote oral language skills in the teaching materials (and to be used in assessments) served as the starting point for the constructs now at the centre of attention.

The pre-piloting in the second stage comprised a series of tasks that targeted either linguistic resources or functional or communicative competences. For the first time, the children worked at computers. The initial experience revealed that even young children are willing and able to solve computer-based tasks. The most problematic aspects of these tasks were slightly longer written instructions, examples on the screen and, in particular, the use of the target language in instructions. Regarding the students' competences, the pre-piloting phase yielded the following key insights: Pronunciation problems are of marginal significance; by contrast, reading written language aloud can greatly challenge young learners. Active knowledge of words and phrases appears to be of central importance; this is particularly true in the perception of the learners themselves. Regarding tasks targeting syntax, some children did not even attempt to solve the problem when they were not able to understand every word. Therefore, possible knowledge of syntax remained hidden. Somewhat more comprehensive tasks targeting oral interaction proved problematic: slightly more demanding tasks and instructions were not readily understood, even in the school language, and often neglected. Most of the independently produced oral utterances in the target language were modest in scope.

The final stage of the project incorporated only the region using the Passepartout curriculum

and therefore worked with schoolchildren who learned French with the *Mille feuilles* textbook. For the first time, fifth and sixth year students also participated in the study. The spectrum of the areas of focus was limited even more narrowly. In the remaining sub-sections, the focus was placed on two areas of communicative action – communication in the classroom (*classroom language*) and the exchange of personal information – as well as on a linguistic partial competence: vocabulary. Listening comprehension was added to the tested elements as a new partial construct of speaking skills; it was expected that this measure would render elementary oral language skills more visible. The structure, the level and the appearance of the tasks in the above-named areas were further adapted to meet the needs of the target group. As a basis for the current and possibly future task development for *Mille feuilles*, a corpus of all written and spoken texts in the textbook was built. Moreover, an inventory of expressions for classroom communication was compiled. The inventory is based on an analysis of various textbooks as well as on a supplementary, fairly comprehensive survey of teachers. Corpus and inventory were used to develop model tasks targeting a series of receptive, interactive and productive speaking skills in important areas. The tasks are short and fit well on a single computer screen. Illustrations were added frequently to ensure the tasks are intuitive to use. The only somewhat complex format retained was the “Dialogue” for the assessment of spoken interaction, which was pre-piloted in an earlier phase of the project. The dialogue is several pages long and consists of listening and speaking to a fictive person.

The field testing was done in several steps. In the process, not only the tasks but also the server-based delivery of tests for whole classes within the school's infrastructure were tested. In addition, the recordings from this phase served as the basis for a (manual) analytical rating of the students' oral productions; assessment masks were tested that could be implemented later in the Item Builder software. On the whole, the results of the final piloting are positive. The testing logistics for whole school classes functioned, and the stu-

Computer-based assessment of foreign-language competences in primary school teaching

dents handled the tasks as expected. The task formats have been so comprehensively tested that they are now available for use in testing instruments. In the case of tasks for assessing receptive skills, the test content proved to be somewhat too simple for use as a formative-diagnostic tool. The tasks for assessing productive skills were successful overall, also for the students in the fifth and sixth class who participated for the first time. For the greater part, the tasks led to the kind of performance that was expected. The attempt to develop a type of analytical assessment on oral productions that is transparent enough to be used nearly intuitively by teachers was very promising. However, practical experiences from a comprehensive piloting with teachers have not yet been gathered.

Classroom language

When teachers and students use a foreign language in the classroom to negotiate tasks, discuss assignments and perform daily rituals such as saying hello and good-bye, the teaching content “foreign language” is transformed into an authentic means of communication. In the field of language pedagogy, the language of communication during class time is known simply as *classroom language*. The use of classroom language as a learning outcome has become part of Swiss curricula and is a focus in a variety of teaching materials, including *Mille feuilles*. Because language instruction represents a setting in which authentic communication can take place from the outset, classroom language was included in the project “Mapping Developmental Profiles” as an area for which tasks should be developed. The first step was to compile a collection of utterances relevant to oral communication in the classroom from the two textbooks referenced in the project, *Mille feuilles* (used in the German-speaking cantons along the language border to the French-speaking regions) and *Alex et Zoé* (used in Ticino) as well as from teacher-training materials. The collected utterances were then presented to experts who reviewed the selection. In a further step, primary school teachers for French as a foreign language

(N=74) were invited to complete an online questionnaire in which they estimated to what degree their students mastered a selection of these utterances either receptively or productively. The teachers also recommended additional relevant utterances that did not form part of the questionnaire. The analysis of the survey yielded the following key conclusions:

1. As a rule, older students have better mastery of the utterances presented; they can also produce more utterances.
2. Most utterances are first mastered receptively and only later productively.
3. In the second half-year of the third class, after a half year of French lessons, the students understand a small number of very common expressions (e.g. “Bonjour.”, “C’est faux.”, “Ouvre la fenêtre.”), and they are somewhat able to produce them.
4. Students are not familiar with utterances that are not in *Mille feuilles*.
5. On a whole, students understand instructions quite well. They are only partly familiar with functional utterances required for games (e.g. “Jette le dé.”) although *Mille feuilles* specifically targets this vocabulary.

The detailed results of this survey are appended to the report as an electronic file and available on the website of the Research Centre (<https://bit.ly/2kLXGRG>). Their various potential uses include serving as a catalogue of useful phrases for communication in class, and they shed light on how these phrases are distributed throughout the four years of language instruction in primary schools. The phrases that the teachers added to the questionnaire are also listed in this document.

The setting of the classroom and the corresponding language formed the basis for the development of tasks in the sub-project “Competences in spoken language”. The utterances selected for

Computer-based assessment of foreign-language competences in primary school teaching

this purpose are from the sections “Instructions” and “Work and Play in Groups”. Both productive and receptive tasks were developed and tested. The placement of phrases in a communicative context was achieved by means of one or more images that illustrated the communicative act.

Competence in language learning strategies

In addition to work in the area of oral language skills, the description and mapping of competence in language learning strategies in primary school students forms one of the priorities in the project “Mapping Developmental Profiles”. Language learning strategies function as so-called transversal competences in multilingual teaching approaches. The “basic skills” defined by the Swiss Conference of Cantonal Ministers of Education and the foreign language curricula of German-speaking Switzerland both cite the development of methods or learning strategies among the outcomes for foreign language instruction. This is somewhat surprising, as specialist literature generally suggests that the correlation between the use of strategies and learning success is not straightforward; the positive effect of certain strategies is apparently dependent on level of competence, situation and context, among other factors. Nevertheless, learning strategies can be taught explicitly, and a positive correlation between the explicit teaching of strategies and the ability to use language has been empirically established.

Sub-project V limited its focus to actual language learning strategies (i.e. strategies in connection with specific learning activities such as strategies for learning vocabulary); strategies of language use, for example, were not considered (e.g. reading strategies).

In an initial phase, the project team developed a structured inventory of such strategies based on existing collections of strategies and foreign language textbooks. Following a longer preparatory stage, items for three sub-sections were developed to map competence in language

learning strategies. The sub-sections are distinctly different:

1. Frequency and diversity of the individual use of language learning strategies;
2. Knowledge about language learning strategies and their suitability;
3. Two aspects of metacognitive skill: firstly, the ability to assess one’s own learning (judgement of learning), operationalised as the ability to predict success in linguistic tasks; and secondly, metacognitive self-monitoring, operationalised as the ability to assess the accuracy of one’s own linguistic performance (confidence judgement).

A specific item type was used to map data in each sub-section:

- to 1. Self-statements on frequency of using specific strategies (on the five-point Likert scale: “very rarely” to “very often”), comparable with items in numerous strategy questionnaires.
- to 2. Ranking tasks: A language-learning task with three possible approaches is described. Experts have assigned the approaches varying degrees of suitability, and the task of the learners is to rate the approaches according to suitability.
- to 3. For prediction: Answering the question: “Could you solve [the linguistic task at hand]?” The answer is given on a four-point Likert scale from “No, definitely not” to “Yes, definitely”. For the subsequent confidence judgement: “Solve [the pre-defined linguistic task].”; then answering the question “Did you get the right answer?” using the same Likert scale as before. Students do not solve these related tasks at the same time to ensure that the confi-

dence judgement is given independent of the prediction. For the assessment, the accuracy of the student's answer to the language task is compared to both self-statements.

As in other parts of the project "Mapping Developmental Profiles", the instruments to measure competence in language learning strategies were generally designed to be used independently by the schoolchildren, and this periodically with intervals of several months. The aim is to prepare the results so that they can be used in the classroom to inform focused instruction and learning. With reference to the results on language learning strategies, the following reporting types (among others) with formative assessment potential were considered:

1. Self-statements on the frequency and diversity of using language learning strategies:
 - Frequency of using the individual strategies in a specific class in comparison to usage in all comparable classes (reference group);
 - Frequency of strategy use by individual students (i.e. how often did individual students select the categories "very rarely" to "very often" on the Likert scale?).
2. Tasks addressing knowledge about language learning strategies:
 - Raw data of individual students and whole classes on ranking the available strategies (and subsequent work with the test items in the classroom).
 - Results of individual students and whole classes on the test scale addressing knowledge about the suitability of strategies (e.g. in order to track progress).

3. Tasks regarding both aspects of metacognitive skill:
 - Results of individual students regarding their tendency to over or underestimate their own abilities (or, in sum, the tendency to *misjudge*);
 - Results of individual learners differentiated according to the errors in prediction and confidence judgement. The two types of results may indicate differing strengths or weaknesses.

Because this sub-project demanded a great deal of pioneering work, the tools were developed and tested in repeated contact with the school classes. The development stage of the instruments for each of the three sub-sections is not at the same level. The self-statements on strategy use were successfully piloted insofar as the items were apparently easy to process and also yielded quite distinct profiles of the strategy use. Ways to make use of these results in practical settings have become apparent. The piloting of the ranking tasks assessing knowledge about the suitability of language learning strategies, yielded mixed results. For a number of reasons, the items did not adequately discriminate between learners, and it was not possible to describe a competence scale for the assumed construct. The tasks regarding metacognitive skill were piloted successfully. The task types used should, however, be further investigated; it is not yet sufficiently clear to what degree the demands arising from the embedded linguistic task interact with the demands arising from the metacognitive task.

The sub-project on describing and mapping competence in language learning strategies garnered various conceptual and material findings in a complex area that is still under-developed, especially at the level of primary education. In addition, not least due to the rejection of unsuccessfully tested approaches, insights were gained that can serve as a basis for other studies.

Assessment and feedback practices in foreign language instruction at primary schools

This part of the project "Mapping Developmental Profiles" was conducted entirely in Ticino and aimed to examine the assessment practices and needs of teachers at primary schools (*scuola elementare*) as well as the teachers' attitudes toward a diagnostically applicable, computer-based assessment tool. To do so, a qualitative approach was taken: nine teachers were interviewed, five of whom had already participated in various cycles of task pre-piloting in the context of the project. The interviews were conducted using a guideline designed on the basis of task pre-piloting, an analysis of the educational setting in Ticino, the relevant textbook as well as literature on formative assessment and feedback. Teachers were first asked about the competences and developmental trajectories of their students; they also answered questions about their own methods to assess learners. In a second step, the teachers were invited to discuss potential components of a computer-based assessment tool. The interviews were transcribed and analysed using an evaluation grid. The qualitative study took the Grounded Theory as a basis.

The interviews revealed that for most of the nine teachers, summative assessments are frequently based on textbook materials and that formative assessments are most commonly introduced into the classroom setting informally and intuitively. In the area of spoken production, the teachers pay particular attention to pronunciation and clarity of the learned words and phrases. Strategies and language awareness are only addressed tangentially, partly because the textbook does not address these topics. The primary school teachers generally emphasise that an assessment – similar to the teaching content – is selected mainly in light of motivating and with regard to the well-being of their young students.

In the interviews, the teachers demonstrated a certain degree of scepticism regarding feasibility when they addressed the periodical use of a computer-based assessment tool able to function as a formative-diagnostic instrument: In addition to

the fact that some schools are poorly equipped with computers and that very few lessons are reserved for French, the teachers voiced reservations about the organisational effort required to record oral productions during lessons. The teachers also doubt that the insight gained would justify the effort required for the computer-based rating of their students' productions. Several teachers said they would prefer a tool able to help their students to independently practise and assess their oral language skills. Despite some scepticism, the teachers believe the tasks to assess oral productions that were developed in the project serve the stated purpose and are suitable for use in Ticino.

Computerbasierte Erfassung von Fremdsprachenkompetenzen im Unterricht der Primarstufe

Bericht zum Projekt „Erfassung von Entwicklungsprofilen im frühen schulischen Fremdsprachenlernen“

Herausgegeben von Peter Lenz und Katharina Karges

Wissenschaftliches Kompetenzzentrum für Mehrsprachigkeit³

I	Einleitung Peter Lenz	52
II	Computerbasierte Beurteilung mündlicher Kompetenzen Katharina Karges	57
III	Kompetenzen im Bereich des Sprechens Katharina Karges, Evelyne Berger, Daniela Kappler	69
IV	Langage de classe Evelyne Berger, Daniela Kappler, Katharina Karges	89
V	Sprachlernstrategische Kompetenzen Thomas Roderer, Anna Kull, Peter Lenz	103
VI	Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare Daniela Kappler	129

I Einleitung

Peter Lenz

Das Projekt „Erfassung von Entwicklungsprofilen im frühen schulischen Fremdsprachenlernen“ ist Teil des Forschungsprogramms 2012-2015 des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit. Das Ziel bestand darin, durch die Entwicklung von geeigneten Instrumenten zur Kompetenzerfassung einen Beitrag zu einer lernorientierten und gleichzeitig computerunterstützten Beurteilungskultur im Fremdsprachenunterricht an Primarschulen zu leisten. Ein wichtiges Merkmal des Projekts ist die Verknüpfung von *Assessment for Learning* (Assessment Reform Group, 2002) mit psychometriebasierter Kompetenzmessung auf der Grundlage eines serverbasierten Beurteilungssystems. Diese Verbindung wird insbesondere auch im *BEAR Assessment System (BAS)* angestrebt (Black u.a., 2011; UC Berkeley, Graduate School of Education, QME, 2002; Wilson & Sloane, 2000). Beim *Assessment for Learning* spielen diagnostisches Beurteilen und die formative Nutzung von Beurteilungen aller Art eine zentrale Rolle; die Verbindung mit Psychometrie (Statistik) und Servertechnologie ist an sich kein zwingender Bestandteil. Beides wird ergänzt, um die Erfassung und Beschreibung von Entwicklungswegen zu systematisieren und die Qualität der Beurteilung und damit deren Potential für Unterricht und Lernen zu steigern.

Dem gewählten psychometrischen Ansatz (Black u.a., 2011) liegt die Annahme zugrunde, dass fortgesetztes Lernen im Kern nicht erratisch ist, sondern typischen Progressionen folgt, die auch beschrieben werden können (*learning progressions*). Dabei lassen sich Progressionen in etwas komplexeren Wissens- und Könnensbereichen analytisch in verschiedene Variablen aufbrechen, in Bezug auf welche Fortschritte erzielt werden können (*progress variables*). Gut nachvollziehbar ist dies am Beispiel kommunikativer Sprachkompetenz und deren Komponenten Wortschatz, Aussprache, funktionales Wissen und Können usw. *Progress variables* sind potenziell auf unterschiedliche Weise miteinander verflochten; individuelle Entwicklungswege können sich durch unterschiedliche Kompetenzausprägungen auf verschiedenen

progress variables unterscheiden. *Progress variables* lassen sich als Skalen zu verschiedenen Teilkompetenzen beschreiben, wobei die Teil-Skalen einer Lernprogression miteinander korreliert sind.

Learning progressions oder Lernprogressionen (zur Definition vgl. Heritage, 2008, S. 4) sind seit den neunziger Jahren, vermehrt aber seit Mitte der 2000er-Jahre, vor allem in den USA, als Grundlage für Curricula, Unterricht und Beurteilung in der Diskussion. Führend ist dabei die Didaktik der Naturwissenschaften. Oft wird versucht, mittels Lernprogressionen die Entwicklung des Verständnisses von *big ideas* (z.B. Verständnis von Materie oder Auftrieb) über eine eher längere Lerndauer hinweg zu beschreiben. Der Weg vom Novizen zum Experten wird als (grundsätzlich) kontinuierliches Fortschreiten in Wissens-, Verstehens- und Könnensbereichen verstanden. Weiter geht man davon aus, dass es *typische* Lernwege gibt, die, sofern sie bekannt sind, durch Curriculum und Unterricht gezielt unterstützt werden können. *Learning progressions* sind nicht als natürliche Erwerbssequenzen zu verstehen, sondern als *typisches* Ergebnis des Zusammenspiels von individuellen Voraussetzungen, zu erwerbendem Gegenstand und einem bestimmten Unterricht. In einem Unterricht, der sich an *learning progressions* orientiert, sind Curriculum, Materialien, Unterrichten und (formativ genutzte) Beurteilung eng ineinander verzahnt. Die Ergebnisse von Beurteilungen lassen sich idealerweise auf einer bekannten Lernprogression wie auf einer Strassenkarte (*road map*) verorten, sodass sich optimale formative Feedbacks und Feedforwards ableiten lassen, dank derer Kenntnislücken gefüllt und weitere Lernschritte geplant werden können (vgl. dazu z.B. Black u.a., 2011; Heritage, 2008; Kennedy & Wilson, 2007).

Im Bereich des Sprachunterrichts und des Sprachenlernens ist die Idee der *learning progressions* bisher kaum aufgegriffen worden. Das Projekt „Dynamic Language Learning Progressions“ (DLLP)⁴ spielt eine Vorreiterrolle (Bailey & Heritage, 2014; 2016). Es ist im Kontext von Englisch als

4 | <https://www.dllp.org/>

I Einleitung

Schulsprache in den USA angesiedelt und hat zum Ziel, Lernprogressionen im Bereich von Sprachfunktionen mit besonderer Bedeutung für schulischen Erfolg (z.B. *Erklären*) zu beschreiben, empirisch zu validieren sowie deren Nutzung im Unterricht und in standardisierten Tests vorzubereiten und zu begleiten. Gemäss Bailey & Heritage (2014, S. 494) findet der Fortschritt potentiell auf mehreren Achsen statt: Quantität; Qualität; Komplexität; Präzision, Korrektheit und Flüssigkeit der Sprache und des sprachlichen Handelns. Die Beschreibung empirisch abgesicherter Lernprogressionen soll in erster Linie Unterrichtenden dabei helfen, Beobachtungen im Unterricht richtig zu verorten und dadurch den Nutzen formativer Beurteilung zu steigern. Das Projektdesign von DLLP widerspiegelt bereits Kritikpunkte und Schwierigkeiten, die im Zusammenhang mit Lernprogressionen diskutiert werden: Das Adjektiv *dynamic* trägt der Feststellung Rechnung, dass sich beobachtete individuelle Lernverläufe oft recht stark voneinander unterscheiden (Bailey & Heritage, 2016). Im Bereich der Naturwissenschaftsdidaktik haben Steedle & Shavelson (2009) beispielsweise gezeigt, dass Lernende mit „Halbwissen“ bezüglich einer *big idea* oft gar keinen kohärenten Satz von Ideen hinter ihren Lösungswegen erkennen lassen, sondern beim Lösen von Aufgaben kontextabhängig unterschiedliche Kenntnisbrocken anwenden (vgl. auch Alonzo, 2012), was die Annahme einer kohärenten, *typischen* Progression in Frage stellt. Inwiefern sprachlich-kommunikative Kompetenz mit den Kompetenzen vergleichbar ist, die bei der Bearbeitung naturwissenschaftlicher Aufgaben zum Tragen kommen, bleibt allerdings zu untersuchen. Zu den festen Bestandteilen des DLLP-Projekts gehören auch die Entwicklung von Weiterbildungen für Unterrichtende sowie die Begleitung der Arbeit mit entwickelten Lernprogressionen im Unterricht. Beides trägt der verbreiteten Feststellung Rechnung, dass die Nutzung von Lernprogressionen im Unterricht, wie kompetente formative Beurteilung überhaupt, nicht leichtfällt und gelernt werden muss (Heritage, 2011; Shavelson, 2009).

Im Projekt „Erfassung von Entwicklungsprofilen“ nimmt die Aufgabenentwicklung im Vergleich

zur Beschreibung von Progressionen eine dominante Stellung ein. Zumindest mitverantwortlich für diese Prioritätensetzung ist der gewählte psychometrische Ansatz, welcher einer systematischen Aufgabenentwicklung als unverzichtbarem Bestandteil der Kompetenzbeschreibung hohe Priorität einräumt. Hinzu kommt das ambitionöse Ziel des Projekts, Daten zum Leistungsstand der Schülerinnen und Schüler über ein Internetserver-basiertes System erheben, verarbeiten und rückmelden zu können. Um in technischer Hinsicht während der Projektdauer unter Einsatz von Projektmitteln möglichst weit zu kommen, musste dieser Aspekt von Anfang an vorangetrieben werden.

Im psychometrischen Ansatz zur Beschreibung von *progress variables* und darauf aufbauend ganzen *learning progressions*, wie ihn Wilson vorschlägt (Black u.a., 2011; Wilson, 2005), werden Aufgaben (inkl. Auswertung) und Kompetenzbeschreibung immer komplementär gedacht; Kompetenzskalen stehen Aufgabenskalen gegenüber. *Progress variables* werden durch beides zusammen beschrieben: Wer sich in einer Teilkompetenz an einem bestimmten Skalenpunkt befindet, *hat* die und die Kompetenzen und *kann* folglich Aufgaben, welche diese Kompetenzen messen, (mit einer bestimmten Wahrscheinlichkeit) so und so gut lösen.⁵ Wenn es darum geht, Kompetenzen zu erkunden, kann man vorläufige Kompetenzvorstellungen als Ausgangspunkt für die Aufgabenkonstruktion nehmen, diese Aufgaben von Lernenden bearbeiten lassen und anschliessend aufgrund der Ergebnisse (bzw. des *outcome space*) vertiefte Erkenntnisse zu den Kompetenzen gewinnen. Aufgabenentwicklung und Skalenbeschreibung sind miteinander verknüpfte iterative Prozesse. Ein serverbasiertes Beurteilungssystem ist u.a. deshalb

5 | Die PISA-Skalen zum Lesen (*reading literacy*) sind ein Beispiel dafür (Organisation for Economic Co-operation and Development, 2010, Kap. 2). Neben einer Gesamtskala gibt es drei Teilskalen zu verschiedenen Lesetypen (z.B. Reflektieren und Evaluieren) sowie je eine zum Lesen von kontinuierlichen und nicht-kontinuierlichen Texten. In den PISA-Skalen sind Kompetenz- und Aufgabenaspekte miteinander verbunden, indem pro Skalensegment beschrieben wird, was Lesende leisten müssen, um Aufgaben zu lösen, die sich in diesem Skalensegment befinden.

I Einleitung

so attraktiv, weil es auch dann, wenn es bereits im Gebrauch ist, aufgrund der einkommenden Daten weiterentwickelt werden kann, um die Aufgabenqualität zu verbessern und die Erkenntnisse über die Kompetenzen der Lernenden zu vertiefen.

Bei den inhaltlichen Bereichen, die im Projekt „Erfassung von Entwicklungsprofilen“ erkundet werden sollten, handelt es sich um die mündliche Produktion und Interaktion sowie die Sprachlernstrategien. Diese Wahl war bewusst anspruchsvoll: Obschon das Sprechen in Curricula und im Unterricht ein wichtiges, oft das wichtigste, Lernziel ist, kommt es in grossen Leistungsmessungen oder in computerbasierten Beurteilungen oft zu kurz. Sprachlernstrategien sind erst in den letzten Jahren im Zuge der Mehrsprachigkeitsdidaktik und der Kompetenzorientierung als eigenständiger Lernzielbereich in den Vordergrund gerückt (v.a. im Lehrplan der deutschschweizerischen Pässepartout-Kantone; Grossenbacher u.a., 2012; Pässepartout, 2015), sodass da in verschiedener Hinsicht Grundlagenarbeit zu leisten war.

Wenn an Lernprogressionen gearbeitet wird, ist es günstig, wenn die Lernenden, die Aufgaben bearbeiten, insgesamt ein breites Kompetenzspektrum abdecken (vom Novizen zum Experten), damit Entwicklungsverläufe deutlich zutage treten. Noch besser wären echte Längsschnittdaten, die über eine lange Zeit hinweg bei individuellen Lernenden erhoben werden. An Letzteres wäre jedoch erst nach der Umsetzung eines serverbasierten Systems zu denken. Im Projekt wurde mit Schülerinnen und Schülern aus dritten, vierten und fünften Primarklassen gearbeitet, sowie einigen Sechstklässlerinnen und Sechstklässlern. Das Kompetenzspektrum, das beleuchtet werden konnte, war also recht eng, zumal die Stundendotierungen für den Unterricht in einer Fremdsprache mit zwei bis drei Wochenlektionen eher bescheiden sind.

Für die konkrete technische Umsetzung des Projekts war mitentscheidend, dass die Erziehungsdirektorenkonferenz (EDK) und die Kantone ungefähr zur selben Zeit, als die Planung des Projekts bereinigt wurde, den Beschluss fassten, eine Aufgabendatenbank aufzubauen, die als Umgebung für die Tests zur Überprüfung der Grundkompetenzen dienen und darüber hinaus auch für

andere Beurteilungs- und Evaluationsprojekte zur Verfügung stehen sollte (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK), 2012). Weil das Projekt „Erfassung von Entwicklungsprofilen“ grundsätzlich auf eine Nutzung durch die Volksschule ausgerichtet ist, erschien es sinnvoll, die Instrumente in den technischen Rahmen einzupassen, der durch die Aufgabendatenbank zumindest in den Grundzügen vorgegeben war. Über das Projekt Aufgabendatenbank konnte ein Kontakt mit dem Leibniz-Institut für Bildungsforschung und Bildungsinformation (DIPF) hergestellt werden, woraus u.a. ein gemeinsames Projekt zur Weiterentwicklung der Testsoftware *CBA Item Builder* entstand. Das Hauptziel der Weiterentwicklung bestand darin, eine Möglichkeit zur Aufnahme, Speicherung und Beurteilung von mündlichen Leistungen (durch menschliche Rater) bei internetbasierten Tests programmtechnisch umzusetzen.

Anspruch und Herausforderungen des Projekts waren beträchtlich. Die fünf Beiträge des vorliegenden Dokuments berichten alle von Arbeiten, die als Teil des Projekts ausgeführt wurden. Die Teilberichte II, III und V handeln dabei direkt von der Entwicklung von Aufgaben zu bestimmten Kompetenzbereichen. Der Teilbericht IV beschreibt Grundlagenarbeiten zur Sprache des Unterrichts, während sich Teilbericht VI vor allem mit Fragen der Implementierung eines computerbasierten, formativ nutzbaren Beurteilungssystems in einem spezifischen Kontext befasst. Die Teilberichte im Überblick:

- Teil II „Computerbasierte Beurteilung mündlicher Kompetenzen“ befasst sich in erster Linie mit der technischen Umsetzung eines serverbasierten Systems, das für die formativ-diagnostische Beurteilung besonders auch der mündlichen Sprachkompetenzen eingesetzt werden kann. Im Mittelpunkt stehen die Entwicklung benutzerfreundlicher Aufgaben, die über einen Internetbrowser zur Verfügung gestellt werden können, sowie eines Moduls zur kriterienbasierten Beurteilung zuvor aufgenommener mündlicher Leistungen durch Lehrpersonen.

I Einleitung

- Teil III „Kompetenzen im Bereich des Sprechens“ berichtet über die etappenweise Entwicklung und Validierung von Aufgaben, mittels derer diagnostische Informationen zu relevanten Teilbereichen der mündlichen Sprachkompetenz der untersuchten Primarschülerinnen und -schüler gewonnen werden können.
- Teil IV „Langage de classe“ handelt von einem Teilprojekt, dessen Ziel es war, auf der Basis von Lehrwerkanalysen und einer Lehrpersonenbefragung sprachliche Mittel zu identifizieren, die für die mündliche Kommunikation im Unterricht (Sprache des Unterrichts, *classroom language*) wichtig sind. Mittels der Lehrpersonenbefragung wurde auch untersucht, wie gut die Schülerinnen und Schüler solche sprachlichen Mittel auf den verschiedenen Klassenstufen rezeptiv und produktiv beherrschen. Innerhalb des Projekts dienten die Ergebnisse als Grundlage für Aufgaben zur mündlichen Sprache des Unterrichts.
- Teil V „Sprachlernstrategische Kompetenzen“ stellt die Entwicklung von Instrumenten zur Erfassung einerseits der Gebrauchshäufigkeit von Sprachlernstrategien durch die Lernenden und andererseits der Kompetenzen der Lernenden in ausgewählten Teilbereichen sprachlernstrategischer Kompetenz dar. Als Grundlage dafür wurde zunächst ein strukturiertes Inventar sprachlernstrategischer Kompetenzen erstellt.
- Teil VI „Pratiche di valutazione e *feedback* nell'insegnamento della lingua straniera alla scuola elementare“ berichtet von den Ergebnissen einer qualitativen Untersuchung im Kanton Tessin zur Beurteilungspraxis von Primarlehrpersonen im Fach Französisch sowie zu deren potentiellen Vereinbarkeit mit der Einführung einer serverbasierten Umgebung zur formativ-diagnostischen Beurteilung.

Obschon viel geleistet wurde, steht das „grosse System“, das im Unterricht mit modellhaften Anwendungen zur formativ-diagnostischen Beurteilung direkt eingesetzt werden könnte, (natürlich) noch nicht bereit. Einzelne der erarbeiteten Elemente konnten aber bereits in anderem Zusammenhang genutzt werden oder werden dies bald.

Referenzen

- Alonzo, A. C. (2012). Learning progressions: significant promise, significant challenge. *Zeitschrift für Erziehungswissenschaft*, 15(1), 95-109.
- Assessment Reform Group (2002). *Assessment for learning: 10 principles. Research-based principles to guide classroom practice*. Abgerufen von http://www.hkeaa.edu.hk/DocLibrary/SBA/HKDSE/Eng_DVD/doc/Afl_principles.pdf [25.07.2019]
- Bailey, A. L. & Heritage, M. (2014). The role of language learning progressions in improved instruction and assessment of English language learners. *TESOL Quarterly*, 48(3), 480-506.
- Bailey, A. L. & Heritage, M. (2016). *DLLP – Dynamic Language Learning Progressions*. Abgerufen von <http://www.dllp.org/> [28.03.2017]
- Black, P., Wilson, M. & Yao, S.-Y. (2011). Road maps for learning: a guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspectives*, 9(2-3), 71-123.
- Grossenbacher, B., Sauer, E. & Wolff, D. (2012). *Mille feuilles. Neue fremdsprachendidaktische Konzepte. Ihre Umsetzung in den Lehr- und Lernmaterialien*. Bern: Schulverlag Plus.

I Einleitung

- Heritage, M. (2008). *Learning progressions: supporting instruction and formative assessment*. Washington, DC: Council of Chief State School Officers. Abgerufen von https://www.csai-online.org/sites/default/files/Learning_Progressions_Supporting_2008.pdf [02.09.2019]
- Heritage, M. (2011). Commentary on road maps for learning: a guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspectives*, 9(2-3), 149-151.
- Kennedy, C. A. & Wilson, M. (2007). *Using progress variables to interpret student achievement and progress* (BEAR Technical Report No. 2006-12-01). Berkeley: Berkeley Evaluation & Assessment Research (BEAR) Center, University of California, Berkeley. Abgerufen von http://bearcenter.berkeley.edu/sites/default/files/Kennedy_Wilson2007.pdf [30.10.2018]
- Organisation for Economic Co-operation and Development (2010). *PISA 2009 results: what students know and can do: student performance in Reading, Mathematics and Science* (Volume I). OECD Publishing. Abgerufen von <http://dx.doi.org/10.1787/9789264091450-en> [02.09.2019]
- Passepartout (Hrsg.) (2015). *Lehrplan Französisch und Englisch*. Abgerufen von <http://www.passepartout-sprachen.ch/services/downloads/download/533/get> [09.03.2016]
- Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) (2012). *Überprüfung der Erreichung der Grundkompetenzen; Konzept und Finanzierung der Aufgabendatenbank: Beschlussfassung*. Abgerufen von http://edudoc.ch/record/105010/files/PB_aufgabendatenbank_d.pdf [09.02.2017]
- Shavelson, R. J. (2009). *Reflections on learning progressions*. Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA. Abgerufen von <http://education.msu.edu/projects/leaps/proceedings/Shavelson.pdf> [28.03.17]
- Steedle, J. T. & Shavelson, R. J. (2009). Supporting valid interpretations of learning progression level diagnoses. *Journal of Research in Science Teaching*, 46(6), 699-715.
- UC Berkeley, Graduate School of Education, QME (2002, 2017). *The BEAR assessment system*. Abgerufen von <http://bearcenter.berkeley.edu/page/bear-assessment-system> [20.03.17]
- Wilson, M. (2005). *Constructing measures: an item response modeling approach*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Wilson, M. & Sloane, K. (2000). From principles to practice: an embedded assessment system. *Applied Measurement in Education*, 13(2), 181-208.

II Computerbasierte Beurteilung mündlicher Kompetenzen

Katharina Karges

1. Eine ideale Beurteilungsplattform	58
2. Aufgabenerstellung und -erprobung	59
2.1. Herausforderungen	59
2.2. Erprobungen	60
2.2.1. Erste Erkundungen	60
2.2.2. Erprobungen mit <i>CBA Item Builder</i>	61
3. Ergebnisse	62
3.1. Bedienung des Computers	62
3.2. Testoberfläche: Klarheit von Inhalt und Form	63
3.3. Ein mögliches Ratingmodul	65
4. Ausblick	65
5. Referenzen	67

II Computerbasierte Beurteilung mündlicher Kompetenzen

1. Eine ideale Beurteilungsplattform

Neben der Beschreibung möglicher Entwicklungsprofile für die fremdsprachlichen Kompetenzen junger Lernender (Teil I „Einleitung“) und der Entwicklung entsprechender Aufgaben (Teil III „Kompetenzen im Bereich des Sprechens“ und Teil V „Sprachlernstrategische Kompetenzen“) war es ein erklärtes Ziel des Projekts, die Möglichkeiten zu untersuchen, wie das so entstehende Beurteilungsinstrumentarium computerbasiert für den Einsatz im Unterricht umgesetzt werden könnte. Als Vorbild diente das frühere *Formative Assessment Delivery System (FADS⁶)*, welches vom BEAR Center der University of California, Berkeley speziell dafür entwickelt wurde, Entwicklungsprofile als Teil des Unterrichtsgeschehens zu erfassen und ein entsprechendes Feedback bzw. Feedforward⁷ auszugeben. Das System richtet sich typischerweise an Lehrpersonen und soll es ihnen ermöglichen, auf den Lehrplan abgestimmte Aufgaben formativ im Unterricht einzusetzen und so den Lernfortschritt ihrer Schülerinnen und Schüler zu verfolgen und diesen bei Bedarf zu beeinflussen. Den Kern von FADS bilden die auf dem jeweiligen Lehrplan basierenden Entwicklungsprofile sowie ihnen zugeordnete Aufgaben, durch deren Bearbeitung die Lernenden statistisch einer bestimmten Stufe im Lernprozess zugeordnet werden können. Aufgrund dieser Zuordnung wiederum kann FADS Hinweise zu wahrscheinlichen nächsten Lern- und Entwicklungsschritten ausgeben.

Ein System, welches eine ähnliche Funktionalität hätte, müsste also gleichzeitig die Verwaltung von Aufgaben bzw. Items,⁸ die webbasierte Durch-

führung von Tests, die Aufzeichnung der Resultate sowie die Aufbereitung und Ausgabe von Lernfortschrittsinformationen unterstützen, die auf einer Skalierung der Leistungen beruhen. Dabei ist es wünschenswert, dass die Items Teil eines regelmässig durch Klassen genutzten, serverbasierten Aufgabenpools sind und so im Prinzip mit jedem Einsatz genauer kalibriert werden.

Angesichts der inhaltlichen Ausrichtung eines Teils des Projekts auf mündliche Kompetenzen⁹ (vgl. Teil III „Kompetenzen im Bereich des Sprechens“) müsste es das System ausserdem erlauben, aus einem herkömmlichen Internetbrowser heraus das Mikrofon eines Computers anzu-steuern und die Audioaufnahme zeitnah auf einem Server zu speichern. Diese müssten dann entweder durch automatische Spracherkennung ausgewertet oder aber einem Rating zugeführt werden, in welchem die Aufnahmen der Lernenden gezielt angehört und durch eine Person bewertet würden.

Schliesslich sollten Lehrpersonen die Möglichkeit haben, ihre Schülerinnen und Schüler im System zu verwalten, ihnen einzeln oder in Gruppen gezielt Aufgaben bzw. vordefinierte Bündel von Aufgaben zuzuweisen und nach Wunsch zusammengestellte Leistungsberichte abzurufen (zu einer bestimmten Aufgabe; zu einem Lernbereich; zur Entwicklung über einen gewissen Zeitraum hinweg; einzelne Lernende und/oder Klassen im Vergleich zu anderen Lernenden und/oder Klassen betreffend usw.). Ein System mit diesen Eigenschaften wäre optimal auf die Nutzung im Zusammenhang mit schulischem Unterricht zugeschnitten. Angesichts der grossen Unterschiede bezüglich der Informatikkenntnisse der Lehrpersonen einerseits und der technischen Ausstattung der Schulen andererseits müsste das System darüber hinaus hochgradig benutzerfreundlich, kompatibel und ressourcenschonend funktionieren. Die Entscheidung, mit einem Internetbrowser-ba-

6 | FADS wurde inzwischen durch BASS ersetzt (*Berkeley Assessment System Software*): <http://bearcenter.berkeley.edu/software/berkeley-assessment-system-software-bass>.

7 | Informationen zum aktuellen Lernstand und zu den Massnahmen, welche den Fortschritt erleichtern könnten.

8 | Zwischen Aufgaben und Items wird oft nicht systematisch unterschieden. Eine Aufgabe besteht oft aus nur einem Item, d.h. aus einer Antwortmöglichkeit, die zu einer einzelnen Bewertung führt. Eine Aufgabe kann aber auch mehrere Items umfassen.

9 | Die Aufgaben zur Erfassung lernstrategischer Kompetenzen wurden zunächst auf Papier entwickelt und erprobt und erst in einem zweiten Schritt inhaltlich äquivalent am Computer umgesetzt. Die hier beschriebenen Überlegungen und Erkenntnisse beziehen sich daher weitgehend auf die Überprüfung mündlicher Kompetenzen.

II Computerbasierte Beurteilung mündlicher Kompetenzen

sierten System zu arbeiten, ist in diesem Zusammenhang zu sehen.

2. Aufgabenerstellung und -erprobung

Mit Blick auf die zukünftige schweizerische Testlandschaft wurde im Jahr 2013 die Entscheidung getroffen, die im Projekt entstehenden Aufgaben technisch so umzusetzen, dass sie in der neu entstehenden Aufgabendatenbank der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren (EDK) funktionieren würden. Damit war die Hoffnung verbunden, durch Synergieeffekte mit dem Projekt „Aufgabendatenbank“ letztlich ein besseres und weiterreichendes Beurteilungsinstrument zu entwickeln. Die Wahl der Auslieferungsoftware, die wir im Projekt benutzten, war Teil der Entscheidung für die Aufgabendatenbank. Es handelt sich dabei um *CBA Item Builder* (DIPF/TBA & Nagarro, 2010ff.). Um das Aufnehmen und Bewerten mündlicher Produktionen im Browser zu ermöglichen, musste *Item Builder* um die entsprechenden Funktionalitäten ergänzt werden. Die notwendigen programmtechnischen Ergänzungen wurden im Rahmen des Projekts „Erfassung von Entwicklungsprofilen“ ausgeführt.

Für die Aufgabenerprobung wurde die speziell für *CBA Item Builder* vorgesehene Auslieferungsoftware EE4CBA¹⁰ verwendet, da die Aufgabendatenbank während der Laufzeit des Projekts noch nicht einsatzbereit war. Die Erkenntnisse aus der Erprobung erlauben deshalb zwar Schlüsse in Bezug auf Form und Inhalt der Aufgaben, nicht aber bezüglich deren Nutzung im Rahmen der Aufgabendatenbank. Auf die spezifische Entwicklung und Erprobung eines Verwaltungsmoduls, in dem Lehrpersonen ihre Lernenden und Aufgaben verwalten können, wurde verzichtet, da diese Funktionen in den Bereich der Aufgabendatenbank fallen sollten.

10 | Diese Software wird unterdessen unter der Bezeichnung *CBA Execution Environment* (CBA EE) geführt (<https://tba.dipf.de/de/infrastruktur/softwareentwicklung/cba-execution-environment>).

Die Arbeit des Projektteams konzentrierte sich auf die Entwicklung von Aufgaben, welche die Lernenden möglichst selbstständig in Einzelarbeit bearbeiten können sollten. So sollte ein formativ-diagnostisches Beurteilungssystem entstehen, welches die Lehrpersonen ohne grösseren Aufwand in ihrem Unterricht nutzen könnten. Zusätzlich sollte praktisch erkundet werden, inwiefern die Erfassung mündlicher Kompetenzen über ein computerbasiertes System (d.h. allein auf Internetserver und -browser zurückgreifend, ohne Beteiligung der Lehrperson während der Aufnahme selbst) möglich ist. So sollte unter anderem auch der Weg hin zu kantonalen oder nationalen *Large-Scale*-Erhebungen im Bereich Sprechen geebnet werden. Solche Leistungsmessungen umfassen bisher keine Elemente zum Sprechen, obschon gerade dieses in den Lehrplänen hohe Priorität geniesst.

2.1. Herausforderungen

Um den Einsatz der zu entwickelnden Aufgaben im Kontext der Primarschule zu ermöglichen, musste eine Reihe von Aspekten besonders beachtet werden, die nicht direkt mit den Fremdsprachenkenntnissen zu tun haben: das Alter und der Erfahrungshintergrund der Lernenden, ihre Fertigkeiten im Umgang mit Computern sowie die Computer- und Netzwerkinfrastruktur an den Schulen.

Eine besondere Herausforderung stellte die Gruppe der jüngsten Lernenden dar, die Drittklässlerinnen und Drittklässler. Im Alter von 8-9 Jahren sind sowohl die Lesefertigkeiten wie auch der Umgang mit dem Computer und die Fähigkeit, selbstständig zu arbeiten, bei manchen Lernenden noch relativ schwach entwickelt. Die Gestaltung und die Funktionalität der Testoberfläche mussten diesen Einschränkungen Rechnung tragen, damit die Aufgaben auch von diesen Schülerinnen und Schülern bearbeitet werden konnten.

Eine wichtigere Rolle als in anderen Kontexten spielt auch das im Unterricht eingesetzte Lehrwerk. Einerseits beeinflusst das Lehrwerk auf dem sehr tiefen Sprachniveau der Zielgruppe die möglichen Inhalte von Testaufgaben erheblich, da

II Computerbasierte Beurteilung mündlicher Kompetenzen

diese nur beinhalten können, was die Lernenden in der einen oder anderen Form bereits zumindest angetroffen haben. Andererseits hat das Lehrwerk auch eine Bedeutung für die Wahl der Aufgabenformate: Lernende können vertraute Aufgabenstellungen schneller und leichter erfassen, was dazu führt, dass die Wahrscheinlichkeit unerwünschter Einflüsse durch Aspekte der Aufgaben, die nichts mit der zu überprüfenden Kompetenz zu tun haben, sinkt.

Schliesslich musste bei der Erstellung der Aufgaben und allgemein beim Testdesign darauf geachtet werden, dass eine Durchführung vor Ort, d.h. in Primarschulhäusern, überhaupt möglich ist. Dies bedeutete konkret, dass die webbasierten Aufgaben auch bei langsamer Internetanbindung und veralteter Hardware technisch funktionieren und in der Umgebung eines möglicherweise engen Computerraums ohne Zugriff auf die üblichen Lernmaterialien des Fremdsprachenunterrichts (z.B. Wörterbücher) grundsätzlich lösbar sein mussten.

2.2. Erprobungen

Die verschiedenen beschriebenen Herausforderungen wurden in mehreren Erprobungsrunden erkundet. Mithilfe von systematischen Beobachtungen, zusätzlichen Befragungen und detaillierten Auswertungen der Schülerantworten wurden die Bedürfnisse der Zielgruppe näher bestimmt und sukzessive in den Aufgaben berücksichtigt.

2.2.1. Erste Erkundungen

Im Frühjahr 2013 wurde noch vor der Übernahme des *CBA Item Builder* als Testsoftware eine relativ umfangreiche Erprobung von Aufgaben zum Sprechen durchgeführt, die als PowerPoint-Präsentation (Microsoft, 2010b) erstellt worden waren. Die Antworten der Lernenden wurden mithilfe der Audio-Software *Audacity* (The Audacity Team, 2013) über das Computermikrofon aufgezeichnet. Für die Schülerinnen und Schüler entstand so der

Eindruck eines am Computer durchgeführten Tests. Neben den Antworten der Kinder wurden gleichzeitig auch ihre Fragen, zusätzliche Erläuterungen durch die Aufsichtspersonen und lange Denkpausen aufgezeichnet. Diese lieferten im Nachhinein zusätzliche Hinweise auf das Funktionieren der Aufgaben.

Die Erprobung wurde auf mitgeführten Laptops des Projektteams in den Räumen der beteiligten Schulen durchgeführt. Es nahmen insgesamt 114 Mädchen und Jungen aus den Kantonen Freiburg und Tessin teil (7 Klassen in 9 Schulhäusern, Klassenstufen 3-5). Die Kinder wurden mit einer Auswahl von Aufgaben konfrontiert, bei denen sie in ihrer ersten Fremdsprache Französisch u.a. Begriffe benennen oder einzelne Wörter oder Sätze laut lesen mussten oder auch komplexere Aufgaben wie das Rekonstruieren eines Dialogs lösen mussten (vgl. für inhaltliche Informationen den Teilbericht III zum Sprechen).

Nach jeder Erprobung verfassten die beteiligten Projektmitarbeitenden Erfahrungsberichte, in denen Auffälligkeiten bei der Testdurchführung notiert wurden. Sie waren neben den aufgezeichneten Schülerleistungen eine wichtige Grundlage für die Weiterentwicklung der Aufgaben. So konnte zum Beispiel schon während der Erprobungen festgestellt werden, dass Aufgaben, bei denen die Lernenden sich in eine in der Schulsprache beschriebene Situation hineinversetzen und darin eine zielsprachliche Äusserung produzieren sollten (vgl. Abbildung II-1), kaum jemals ohne Hilfe gelöst wurden.

Weitere Beobachtungen und die daraus gezogenen Schlüsse werden im Abschnitt 3 „Ergebnisse“ zusammengefasst.

II Computerbasierte Beurteilung mündlicher Kompetenzen

Abbildung II-1: Eine kognitiv zu anspruchsvolle Aufgabenstellung. Erprobungsaufgabe von 2013. Erstellt in MS PowerPoint 2010.

2.2.2. Erprobungen mit *CBA Item Builder*

Die Erprobung der mit *CBA Item Builder* erstellten Aufgaben fand im Jahr 2014 in vier grossen Abschnitten statt, wobei insgesamt rund 565 Kinder der Klassenstufen 3-5 in den Kantonen Bern, Freiburg und Solothurn involviert waren. Bei den Aufgaben handelte es sich einerseits um Hörverstehensaufgaben, bei denen auf die eine oder andere Weise der Zusammenhang zwischen einer bildlichen Darstellung und einer Hördatei hergestellt werden musste, und andererseits um Aufgaben zur mündlichen Produktion, bei denen aufgrund eines Bildes oder einer Hördatei etwas gesagt werden musste.¹¹ In beiden Fällen gab es

11 | Neben den Aufgaben zum Sprechen wurden auch Aufgaben zu den lernstrategischen Kompetenzen miterprobt. Da diese sich nicht erheblich von ihren Vorbildern auf Papier unterschieden, werden sie hier nicht diskutiert.

sowohl Aufgaben zu einzelnen lexikalischen Elementen als auch solche, die das Verstehen oder Produzieren von einzelnen Sätzen oder kurzen Texten erforderten.¹²

Bei den ersten beiden Erprobungen im März und April 2014 bekamen die Kinder die Aufgaben offline als sog. Preview-Versionen des *CBA Item Builder* vorgelegt. Dabei wurde das Antwortverhalten der Kinder wiederum aufgezeichnet, allerdings anders als im Jahr zuvor nun als Bildschirmvideos (mithilfe der Software CamStudio: Smith, 2007ff.). Dadurch wurde es möglich, die Lernenden auch im Nachhinein bei der Bedienung der Testoberfläche zu beobachten und diese bei Bedarf aufgrund der Beobachtungen zu verbessern.

Als zusätzlicher Schritt in der Qualitätssicherung wurden die Verständlichkeit und Eindeutigkeit der als Teil der Items eingesetzten Bilder

12 | Zu inhaltlichen Aspekten der Aufgaben und ihrer Begründung vgl. den Teilbericht III zum Sprechen.

II Computerbasierte Beurteilung mündlicher Kompetenzen

überprüft, indem die Lernenden diese auf einem Arbeitsblatt in der Schulsprache benennen mussten. Ausserdem wurde das in den Aufgaben verwendete Vokabular schriftlich abgefragt, um ggf. abschätzen zu können, ob aussergewöhnliche Lösungsraten bei einer Aufgabe auf ihren Inhalt oder ihr Format zurückzuführen waren. Die Erprobungen wurden durch Gruppengespräche ergänzt, in welchen u.a. Fragen zur Computernutzung der Kinder innerhalb und ausserhalb des Fremdsprachenunterrichts sowie zur Ausstattung des Familienhaushalts mit Computern gestellt wurden. Sofern ausreichend Zeit zur Verfügung stand, wurden direkt nach dem Lösen der Aufgaben die Eindrücke der Lernenden mehr oder weniger systematisch erfragt und aufgezeichnet. Die so gesammelten Daten flossen in die Überarbeitung der Testoberfläche ein und hatten Einfluss auf die inhaltlichen Aspekte der zukünftigen Aufgaben. Die Daten zeigten, dass die erprobten Aufgaben zu einem grossen Teil bereits gut auf die Zielgruppe abgestimmt waren.

Bei den grösseren Erprobungen im Juni/Juli und Oktober/November 2014, welche online auf den Rechnern der Schulen stattfanden, wurde weitgehend auf begleitende und unterstützende Massnahmen verzichtet. So war es möglich, zu erkunden, wie die Aufgaben unter nah-authentischen Bedingungen funktionieren würden. In dieser Phase konnten wichtige Erkenntnisse zu technischen und logistischen Aspekten des Einsatzes des Beurteilungsinstrumentariums im Schulalltag gewonnen werden. Die zu diesem Zweck formulierten Fragestellungen konnten weitgehend aufgrund der Beobachtungen der Projektmitarbeitenden beantwortet werden, welche die Tests vor Ort durchführten. Inhaltlich ging es ausserdem darum, die Aufgaben mit einer grösseren Zahl von Lernenden auszuprobieren, um so wenigstens explorativ Informationen zu verschiedenen Aufgabenmerkmalen, wie z.B. ihrer Schwierigkeit oder ihrer Eignung, zwischen Schülerinnen und Schülern unterschiedlicher Klassenstufen zu unterscheiden, zu gewinnen (vgl. den Teilbericht III zum Sprechen).

3. Ergebnisse

3.1. Bedienung des Computers

Bei allen Erprobungen konnte durchgehend beobachtet werden, dass der Umgang mit dem Computer für Primarschülerinnen und -schüler ab der 3. Klasse grundsätzlich ohne zusätzliche Einweisung möglich war: Die Benützung von Mäusen oder Touchpads stellte kein Problem dar und auch die z.T. sehr grossen Headsets wurden häufig mit Begeisterung aufgenommen („Lueg, i bin e Pilot!“) und führten nicht zu sichtbaren Berührungssängsten. Lediglich gegen Ende der Primarstufe zeigten einzelne Lernende eine gewisse Schüchternheit bei der Verwendung des Mikrofons.

Eine wichtige Erkenntnis war, dass das Schreiben am Computer von der Mehrheit der Lernenden erst ab Klasse 5 erwartet werden kann: Jüngere Schülerinnen und Schüler taten sich dort, wo sie auch etwas tippen sollten, sehr schwer. Das Problem wurde durch die für Französisch typischen Sonderzeichen noch verschärft. Aufgaben, welche von Kindern unter 10 Jahren gelöst werden können, sollten daher möglichst geringe Schreibkompetenzen (am Computer) voraussetzen bzw. deren Einsatz vermeiden.

Schliesslich ging aus den Erprobungen auch hervor, dass die Bedeutung der technischen Kompetenzen der Personen, welche die Tests durchführen, nicht zu unterschätzen ist: Insbesondere dann, wenn die technische Infrastruktur vor Ort ohne Hilfe durch eine lokale Fachperson verwendet werden soll, sind Kenntnisse der vorhandenen Betriebssysteme, ihrer Systemeinstellungen und der verschiedenen Hardwarekomponenten unerlässlich. So konnte z.B. die Konfiguration der mitgebrachten Headsets oder das Auffinden und Starten der nötigen Browser bei unbekanntem Schulcomputern schnell zur Herausforderung werden.

II Computerbasierte Beurteilung mündlicher Kompetenzen

3.2. Testoberfläche: Klarheit von Inhalt und Form

Die Gestaltung einer Testoberfläche, welche eine möglichst fehlerlose Bedienung der Aufgaben durch die Lernenden ermöglicht, wurde vor allem in den Erprobungen von 2013 und Frühjahr 2014 untersucht. Dabei wurden zunächst der Aufbau der einzelnen Aufgaben vereinfacht und die Anzahl der auf einer Bildschirmseite vorhandenen Elemente auf ein Minimum reduziert. Bei den grösseren Erprobungen im Sommer und Herbst 2014 wurden schliesslich durch das Blockieren bzw. Verstecken von Schaltflächen die Aktionsmöglichkeiten der Nutzenden so stark eingeschränkt, dass Fehler und Sackgassen bei der Bedienung fast ausschliesslich auf externe Faktoren zurückzuführen waren.¹³ Die geradlinige und wiedererkennbare Struktur der Aufgaben hat sich in der Folge auch in anderen Projekten unseres Instituts als nützlich erwiesen.

Aufgrund der umfassenden Erfahrung kann festgehalten werden, dass die üblichen Symbole zum Abspielen, Aufnehmen oder Stoppen eines Audioelements auch von sehr jungen Lernenden ohne Probleme bedient werden können. Das Gleiche gilt für die Schaltfläche, welche zur nächsten Aufgabe führt. Im Laufe der Erprobungen zeigte sich allerdings, dass es Sinn macht, die Bedienung dieser Schaltflächen einfürend kurz am Beispiel zu zeigen und gemeinsam auszuprobieren. Dafür wurde eine Startseite konzipiert, welche jeweils im Plenum mündlich eingeführt wurde (vgl. Abbildung II-2).

Hingegen war es wenig zielführend, den Aufgabenseiten zusätzliche erklärende Elemente beizufügen: Die Mehrheit der Lernenden nahm weder gutgemeinte schriftliche Hilfestellungen wahr (z.B. Satzanfänge, welche die Lösung einleiteten) noch nutzten sie eine Schaltfläche mit einem grossen roten Fragezeichen, welche zu einer Hilfeseite führte. Die Präsenz einer Person, welche bei Unsi-

13 | Möglich waren zum Beispiel noch immer das versehentliche Abschalten des Mikrofons mittels einer Tastenkombination oder der Abbruch des Tests durch das Betätigen der „Zurück“-Taste des Browsers.

cherheiten bezüglich der Formate oder des Funktionierens der Testoberfläche half, war in vielen Fällen die effizientere Lösung und meist auch nur solange notwendig, wie sich die Lernenden mit den Aufgaben vertraut machten. Es ist daher davon auszugehen, dass diese Rolle bei einem formativ eingesetzten Beurteilungsinstrument von der Lehrperson übernommen werden kann und nach wenigen Einsätzen des Instruments unnötig wird. Denkbar ist auch eine videobasierte oder interaktive Präsentation der verwendeten Formate zu Beginn der Arbeit mit der Beurteilungsplattform. Diese Präsentation müsste allerdings gründlich erprobt werden: Im Laufe der Erprobungen bewährten sich Erklärungen, die direkt auf dem Computerbildschirm gegeben wurden, in der Regel nicht. Insbesondere Bildschirmvideos, bei denen sich ohne Zutun der Lernenden „die Maus bewegt“, führten zu grosser Verwirrung. Zu überprüfen wäre hingegen auf jeden Fall, wie effizient entsprechende Hilfen wären, die sich an die betreuende Lehrperson anstatt an die Lernenden richten.

Die Notwendigkeit einer klar strukturierten Aufgabendarstellung lässt sich auch auf die Wahl der Aufgabenformate übertragen. So wurde bereits bei den ersten Erprobungen deutlich, dass Bildschirmseiten mit viel Text zu mehr Nachfragen führten. Ebenso war zu beobachten, dass vor allem jüngere Lernende noch über zu wenig Leseerfahrung verfügen, um aus der Anordnung oder Form von Textelementen Schlüsse auf ihre Wichtigkeit zu ziehen.

Weiter konnte beobachtet werden, dass Primarschülerinnen und -schüler am besten mit Aufgabenbildschirmen zurechtkamen, welche eindimensional strukturiert waren, deren Elemente also auf nur einer Seite von oben nach unten angeordnet waren (vgl. Abbildung II-3).

Damit fielen alle Aufgabenformate, für deren Bearbeitung das Lesen einer längeren Erklärung oder Einleitung nötig war, weg (u.a. das in Abbildung II-1 dargestellte Beispiel). In diesem Zusammenhang kann ebenfalls festgehalten werden, dass Aufgabenstellungen in der Zielsprache bei den jüngsten Lernenden ein unnötiges Hindernis darstellten. Umgekehrt konnten keine betont negativen Reaktionen auf die Aufgabenstellungen

II Computerbasierte Beurteilung mündlicher Kompetenzen

Willkommen!

In den nächsten Minuten bekommst du einige kurze Aufgaben. Du hörst kurze Sätze oder einzelne Wörter. Manchmal musst du auch selber etwas ins Mikrofon sprechen.

Es steht immer genau da, was du machen musst.

In den Aufgaben gibt es drei wichtige Knöpfe:

-  **Klicke hier, um dir etwas anzuhören.**
-  **Klicke hier, um selber etwas aufzunehmen.**
-  **Klicke hier, um zur nächsten Aufgabe zu kommen.**



Abbildung II-2: Screenshot der Startseite, wie sie während der Erprobungen im Sommer/Herbst 2014 verwendet wurde. Die Schaltflächen Abspielen und Aufnahmen waren funktionstüchtig und dienten gleichzeitig der Kontrolle der Audiofunktionen des Computers. Üblicherweise wurde die Startseite zu Beginn einer Testsession mündlich gemeinsam mit den Schülerinnen und Schülern eingeführt.

Was sagt die Lehrerin?

Höre, was die Lehrerin sagt und kreuze das passende Bild an!



Abbildung II-3: Typische Bildschirmaufteilung (Beispiel: CL_mpc3i1a_decoupleimage)

II Computerbasierte Beurteilung mündlicher Kompetenzen

in der Schulsprache beobachtet werden.¹⁴

Aufgrund all dieser Feststellungen wurden die Aufgaben in *CBA Item Builder* von 2014 um bildliche Darstellungen herum aufgebaut und die Formate so vereinfacht, dass ihre Bearbeitung nicht auf das Lesen einer längeren Erklärung oder eines Inputtextes angewiesen ist. Im Allgemeinen genügten in diesem Fall die Bilder in Verbindung mit den Audioschaltflächen, um die Aufgabenstellung deutlich zu machen. Unterstützend wurden zudem kurze, standardisierte Anweisungen und Kontextualisierungen in der Schulsprache formuliert. Sie wurden vor allem von den älteren Schülerinnen und Schülern bei selteneren Formaten gut genutzt. Es bleibt aber festzuhalten, dass sich eine starke Begrenzung bzw. eine Vermeidung von Textpassagen mit Anweisungen, welche für das Lösen der Aufgaben *unbedingt gelesen* werden müssen, als zielführend erwiesen hat. Die Notwendigkeit, die Lernenden bei der Arbeit zu unterstützen, nahm dadurch erheblich ab.

3.3. Ein mögliches Ratingmodul

Ein entscheidender Faktor bei der Entwicklung von Aufgaben zur mündlichen Produktion ist die Entwicklung eines Ratingmoduls.¹⁵ Da *CBA Item Builder* diese Funktion zunächst nicht mitbrachte, wurde sie zusammen mit der Entwicklung der Aufnahmefunk-

14 | Zunächst stand die Vermutung im Raum, dass die Kinder durch die Mischung von Schul- und Zielsprache verwirrt sein könnten. Um dies zu überprüfen, wurden einige Lernende mit rein zielsprachlichen Aufgaben konfrontiert, während andere die gleichen Aufgaben, aber mit Aufgabenstellungen in ihrer Schulsprache erhielten.

15 | Generell wurde im Projekt davon ausgegangen, dass freie Antworten der Lernenden manuell von einer Lehrperson bewertet, und nicht etwa durch ein Modul für automatische Spracherkennung gescort würden. Dies ist einerseits darauf zurückzuführen, dass die dafür nötige Technologie bislang nur mit grossem Aufwand entwickelt werden kann. Andererseits ist es aber auch im Sinne eines formativ-diagnostischen Instruments, dass die Lehrperson selber die Produktionen ihrer Lernenden begutachtet und so unabhängig von der statistischen Auswertung der Resultate einen Überblick über deren Lernfortschritt erhält.

tion für mündliche Produktionen beim DIPF¹⁶ in Auftrag gegeben. Bei Projektabschluss ist die dabei entstandene Bewertungsumgebung zwar einsatzbereit, das Handling aber noch nicht ausgereift: Um eine einzelne Produktion eines einzelnen Lernenden zu bewerten, muss diese aus einer Liste gewählt und angeklickt werden, woraufhin für das Aufrufen der eigentlichen Beurteilungsmaske mindestens drei weitere Klicks mit der Maus gemacht werden müssen. Ist die Bewertung dieser einzelnen Produktion dann abgeschlossen, sind wiederum drei Klicks notwendig, bevor wieder die Liste aufgerufen wird, in der die nächste Einzelproduktion ausgewählt werden kann.

Abgesehen vom Vorschlag einer Beurteilungsmaske mit den genannten Nachteilen in der Handhabung ist momentan keine Umsetzung der Bewertungsfunktion in der Aufgabendatenbank angedacht. Im Rahmen des hier beschriebenen Projektes wurde daher nur eine einzelne Bewertungsmaske probenhalber konkret umgesetzt (vgl. Teil III zum Sprechen). Die Beurteilung der mündlichen Schülerleistungen aus den Erprobungen wurde daher extern in Excel (Microsoft, 2010a) vorgenommen.

4. Ausblick

Durch die erst im Laufe des Projektes „Erfassung von Entwicklungsprofilen“ beschlossene Anbindung an das Projekt „Aufgabendatenbank“ der EDK veränderte sich die technische Arbeit des Projektteams v.a. durch die Nutzung der dort vorgesehenen Software erheblich. Bestimmte allgemeinere Funktionen wie die eines Verwaltungsmoduls, in dem die Lehrpersonen (oder Testanbieter) potentiell Aufgaben verwalten, Tests zusammenstellen und Resultate auslesen können, müssten Teil einer übergeordneten Ebene der Aufgabendatenbank sein und sollten nicht im Rahmen eines einzelnen Projekts wie des vorliegenden entwi-

16 | Leibniz-Institut für Bildungsforschung und Bildungsinformation. Das DIPF ist für die technische Bereitstellung der Aufgabendatenbank und damit auch die Entwicklung ihrer Komponenten zuständig. Der *CBA Item Builder* gehört dazu.

II Computerbasierte Beurteilung mündlicher Kompetenzen

ckelt werden müssen. Die Weiterentwicklung eines formativ-diagnostischen Beurteilungsinstrumentariums hängt in diesem Sinne direkt damit zusammen, wie rasch die geplante Aufgabendatenbank weiterentwickelt und schulsystemnahen Nutzern für Anwendungen zur Verfügung gestellt wird. Zum aktuellen Zeitpunkt sind die Serverkapazitäten der Aufgabendatenbank noch zu gering, als dass ein Zugang für Lehrpersonen vorgesehen werden könnte, über den sie Lernende, Aufgaben und Ergebnisse verwalten oder Tests erstellen und durchführen könnten. Ein Hindernis für Aufgabenentwicklungen im grösseren Stil könnte die Software *CBA Item Builder* darstellen: Gerade, weil das Programm ein grosses Spektrum an Möglichkeiten bietet, erfordert die Erstellung einzelner Aufgaben viele Arbeitsschritte und ein nicht unerhebliches technisches Know-how. Selbst kleinste Änderungen müssen ausserdem manuell vorgenommen werden.¹⁷ Das Programm eignet sich daher gut für die Erstellung einer festen Anzahl von Aufgaben für *Large-Scale-Assessments* oder, zusammen mit seiner Auslieferungssoftware EE4CBA, für Forschungsprojekte,¹⁸ beispielsweise jedoch kaum für das selbstständige Ergänzen von Aufgaben durch Lehrpersonen. Auch das Auslesen der Resultate von Testdurchführungen ist bislang nicht voll integriert möglich, wodurch wiederum der Bezug auf Referenzskalen oder Entwicklungsprofile und damit die Ausgabe eines formativen Feedbacks oder Feedforwards ausgeschlossen ist. Schliesslich wird eine integrierte Durchführung von Tests zur mündlichen Produktion, obwohl auf der Seite der Lernenden möglich, durch das noch zu wenig ausgereifte Beurteilungsmodul erschwert.

Das Projekt „Erfassung von Entwicklungsprofilen“ konnte zweifellos einen Beitrag leisten zur Erkundung der spezifischen Bedürfnisse von Pri-

marschülerinnen und -schülern im Bereich des computerbasierten Testens. Am Ende des Projektes liegen Erfahrungen zum Bildschirmdesign und zur Testdurchführung vor, die es möglich machen, dass die Zielgruppe im Klassenverband und ohne grössere Unterstützung durch die Lehrperson Aufgaben zur mündlichen Produktion und Rezeption löst. Das Kompetenzzentrum für Mehrsprachigkeit verfügt zudem dank des Projektes über den Zugang zu einer Testerstellungs- und Auslieferungssoftware mit erheblichen technischen Möglichkeiten, die weit über die Aufnahme mündlicher Produktionen hinausgehen. Die Nutzung und der Ausbau dieser Ressourcen, wie auch die weitere Erkundung der Möglichkeiten und Grenzen von computerbasiertem Testen wären ohne Zweifel eine lohnende Investition zugunsten der schweizerischen Bildungslandschaft.

17 | Zum Beispiel wäre es nicht möglich, die Hintergrundfarbe einer Reihe von Aufgaben auf einmal zu ändern. Stattdessen müsste jede einzelne Bildschirmseite separat geändert werden.

18 | Die Erfahrungen aus dem Projekt „Erfassung von Entwicklungsprofilen“ konnten direkt auf andere Projekte unseres Instituts sowie die Erstellung der Aufgaben für die Überprüfung der Erreichung der Grundkompetenzen (ÜGK) durch die EDK übertragen werden.

II Computerbasierte Beurteilung mündlicher Kompetenzen

5. Referenzen

- DIPF/TBA & Nagarro (2010ff.). *CBA Item Builder*. Frankfurt (Main): DIPF. Abgerufen von <https://tba.dipf.de/de/infrastruktur/softwareentwicklung/cba-item-builder/cba-itembuilder> [04.10.18]
- Microsoft (2010a). *Microsoft Excel* (Version 2010) [Microsoft Windows 7]. Redmond: Microsoft Corporation.
- Microsoft (2010b). *Microsoft PowerPoint* (Version 2010) [Microsoft Windows 7]. Redmond: Microsoft Corporation.
- Smith, N. (2007ff.). *CamStudio*. Free streaming video software (Version 2.7.2) [Windows 7]. Abgerufen von <http://camstudio.org/> [12.05.14]
- The Audacity Team (2013). *Audacity* (Version 2.0.3) [Microsoft Windows 7].

III Kompetenzen im Bereich des Sprechens

Katharina Karges, Evelyne Berger, Daniela Kappler

1. Anhaltspunkte zur Kompetenzentwicklung im Bereich des Sprechens	70
1.1. Kompetenzentwicklung gemäss den Niveaubeschreibungen des Europäischen Referenzrahmens	70
1.2. Progressionsvorstellungen für den mündlichen Bereich in der Fachliteratur zum Testen	70
2. Möglichkeiten der Erfassung von Sprechkompetenz	73
2.1. Arbeitsphase 1: Lehrwerkorientierung	73
2.1.1. Konstrukte	73
2.1.2. Aufgaben und Aufgabenerprobung	74
2.2. Arbeitsphase 2: Ressourcenbezogenes Testen	75
2.2.1. Konstrukte	75
2.2.2. Aufgaben und Aufgabenerprobung	76
2.2.3. Allgemeine Erkenntnisse aus der Erprobung	76
2.2.4. Ergebnisse der Aufgabenerprobung	77
2.3. Arbeitsphase 3: Erfassung einer gezielten Auswahl von mündlichen Kompetenzen	78
2.3.1. Konstrukte	78
2.3.2. Formate	79
2.3.3. Entwicklungsschritte	80
2.3.4. Erkenntnisse	80
3. Schluss und Ausblick	86
4. Referenzen	88

III Kompetenzen im Bereich des Sprechens

1. Anhaltspunkte zur Kompetenzentwicklung im Bereich des Sprechens

1.1. Kompetenzentwicklung gemäss den Niveaubeschreibungen des Europäischen Referenzrahmens

Im Teilprojekt „Kompetenzen im Bereich des Sprechens“ wurde als Erstes versucht, für den interessierenden Bereich der elementaren Sprachkenntnisse von Primarschülerinnen und -schülern literaturbasiert Entwicklungslinien mündlicher interaktiver und produktiver Kompetenz zu beschreiben. In der Schweiz orientieren sich die schulischen Lehrpläne und Lehrwerke grundsätzlich an den Kompetenzvorstellungen des Gemeinsamen europäischen Referenzrahmens (Europarat, 2001, im Folgenden „Referenzrahmen“). Die illustrativen Deskriptoren des Referenzrahmens für die mündliche Interaktion und Produktion auf den interessierenden Niveaus A1 und A2¹⁹ erlauben nur grobe Rückschlüsse auf die erwarteten Schritte bei der Kompetenzentwicklung. Wesentliche Beschreibungselemente sind im „Raster zur Selbstbeurteilung“ (vgl. Tabelle III-1) und im „Beurteilungsraster zur mündlichen Kommunikation“ (vgl. Tabelle III-2) des Referenzrahmens zu finden.

Diese und ähnliche Niveaubeschreibungen des Referenzrahmens bieten nur wenige konkrete Anhaltspunkte für die Formulierung von Lernprogressionen innerhalb der beiden Niveaubänder. Deutlich wird v.a., dass kurze, oft als Ganzes gelernte Wendungen, für das Gelingen von (elementarer) Kommunikation von zentraler Bedeutung sind.

19 | A2 muss im Primarschulunterricht als Obergrenze angesehen werden, sofern das Kind keine ausserschulischen Kontakte mit der Fremdsprache hat.

1.2. Progressionsvorstellungen für den mündlichen Bereich in der Fachliteratur zum Testen

Auch in der Literatur zum Beurteilen und Testen sind Versuche zu finden, sprachliche Entwicklungen im Bereich elementarer Fremdsprachenkenntnisse zu beschreiben. McKay (2006) stellt ein Modell der Entwicklung der mündlichen Kommunikationsfähigkeit (*language ability*) junger Lernender vor, das sich an Bachman & Palmer (1996) anlehnt. Die Autorin beschreibt darin, was an Entwicklung geschehen muss, bis ein Kind in der Lage ist, in der Fremdsprache erfolgreich zu kommunizieren. Zur Beschreibung der Entwicklungsdimensionen nutzt sie Bachman & Palmers Klassifikation der sprachlichen Ressourcen (*language knowledge*):

Grammatical knowledge: Children's knowledge of vocabulary, syntax and phonology needs to grow and deepen. Their syntax needs to increase in accuracy. [...]

Textual knowledge: Children need to be increasingly able to speak in ways that are cohesive and well-organized; [...]

Functional knowledge: Children's ability to use language for many different functions needs to grow. They need to learn how to use language to get what they want, to learn, to imagine things, to think about things. They need to learn how to understand the purposes behind the language that is spoken to them, even when those purposes are not directly stated. [...]

Sociolinguistic knowledge: Children need to learn to use oral language appropriate to the language use situation that they are in. [...]

(McKay, 2006, S. 183ff.)

III Kompetenzen im Bereich des Sprechens

	A1	A2
An Gesprächen teilnehmen	Ich kann mich auf einfache Art verständigen, wenn mein Gesprächspartner bereit ist, etwas langsamer zu wiederholen oder anders zu sagen, und mir dabei hilft zu formulieren, was ich zu sagen versuche. Ich kann einfache Fragen stellen und beantworten, sofern es sich um unmittelbar notwendige Dinge und um sehr vertraute Themen handelt.	Ich kann mich in einfachen, routinemäßigen Situationen verständigen, in denen es um einen einfachen, direkten Austausch von Informationen und um vertraute Themen und Tätigkeiten geht. Ich kann ein sehr kurzes Kontaktgespräch führen, verstehe aber normalerweise nicht genug, um selbst das Gespräch in Gang zu halten.
Zusammenhängendes Sprechen	Ich kann einfache Wendungen und Sätze gebrauchen, um Leute, die ich kenne, zu beschreiben und um zu beschreiben, wo ich wohne.	Ich kann mit einer Reihe von Sätzen und mit einfachen Mitteln z.B. meine Familie, andere Leute, meine Wohnsituation, meine Ausbildung und meine gegenwärtige oder letzte berufliche Tätigkeit beschreiben.

Tabelle III-1: Deskriptoren aus dem „Raster zur Selbstbeurteilung“ zum Kompetenzbereich Sprechen auf den Niveaus A1 und A2 (Auszug aus Europarat, 2001, S. 36)

	A1	A2
Spektrum	Hat ein sehr begrenztes Repertoire an Wörtern und Wendungen, die sich auf Informationen zur Person und einzelne konkrete Situationen beziehen.	Verwendet elementare Satzstrukturen mit memorierten Wendungen, kurzen Wortgruppen und Redeformeln, um damit in einfachen Alltagssituationen begrenzte Informationen auszutauschen.
Korrektheit	Zeigt nur eine begrenzte Beherrschung von einigen wenigen einfachen grammatischen Strukturen und Satzmustern in einem auswendig gelernten Repertoire.	Verwendet einige einfache Strukturen korrekt, macht aber noch systematisch elementare Fehler.
Flüssigkeit	Kann ganz kurze, isolierte, weitgehend vorgefertigte Äußerungen benutzen; braucht viele Pausen, um nach Ausdrücken zu suchen, weniger vertraute Wörter zu artikulieren oder um Verständigungsprobleme zu beheben.	Kann sich in sehr kurzen Redebeiträgen verständlich machen, obwohl er/sie offensichtlich häufig stockt und neu ansetzen oder umformulieren muss.
Interaktion	Kann Fragen zur Person stellen und auf entsprechende Fragen Antwort geben. Kann sich auf einfache Art verständigen, doch ist die Kommunikation völlig davon abhängig, dass etwas langsamer wiederholt, umformuliert oder korrigiert wird.	Kann Fragen stellen und Fragen beantworten sowie auf einfache Feststellungen reagieren. Kann anzeigen, wann er/sie versteht, aber versteht kaum genug, um selbst das Gespräch in Gang zu halten.
Kohärenz	Kann Wörter oder Wortgruppen durch einfache Konnektoren wie und oder dann verknüpfen.	Kann Wortgruppen durch einfache Konnektoren wie und, aber und weil verknüpfen.

Tabelle III-2: Deskriptoren aus dem „Beurteilungsraster zur mündlichen Kommunikation“ zu vorwiegend qualitativen Aspekten mündlichen Sprachgebrauchs (Auszug aus Europarat, 2001, S. 38)

III Kompetenzen im Bereich des Sprechens

In einem zweiten Schritt fügt McKay diesen Aspekten weitere Kenntnisse hinzu, zu denen auch Lern- und Sprachgebrauchsstrategien und die Fähigkeit zur Selbstregulierung gehören. Letztlich beschreibt sie ein umfassendes Modell von Sprechkompetenz, in welchem sich Lernfortschritt durch qualitativen und quantitativen Zuwachs in sämtlichen Bereichen kommunikativer und strategischer Kompetenz ausdrückt. Sie betont zwar z.B. im Zusammenhang mit grammatischer Kompetenz, junge Lernende hätten ganz zu Beginn des Spracherwerbs „very little conscious awareness of the grammar within the language, both because of the use of chunks and because of limits in metalinguistic awareness“ (McKay, 2006, S. 194). Sie folgert aus dieser Feststellung aber lediglich, dass die Beurteilung grammatischer Kompetenz in dieser Zeit nicht oder nur mit grosser Vorsicht vorzunehmen ist, um die Entwicklung nicht zu hemmen. Bei fortgeschrittenen Lernenden sei es dann, im Sinne diagnostischen Testens, sinnvoll, spezifische grammatische Phänomene neben anderen Aspekten mündlicher Kompetenz zu erfassen. Wodurch genau sich sehr geringe, beginnende Sprechkompetenz auszeichnet, wird auch bei McKay nicht näher diskutiert.

Einen umfassenderen Überblick über die Diskussion um die Entwicklung von Sprechkompetenz gibt Bygate (2009). Dabei macht er, vor allem im Zusammenhang mit Beurteilungsskalen und -kriterien, ähnliche Beobachtungen wie die oben beschriebenen. Zwar nennt er eine Reihe von Merkmalen anhand derer die Sprechkompetenz in einer Fremdsprache beschrieben und eingestuft werden könnte, stellt aber zugleich fest, dass diese von den Bedingungen beeinflusst werden, unter welchen die gesprochene Sprache konkret entsteht. Er beschreibt ein komplexes Zusammenspiel von Faktoren,²⁰ welche mehr oder weniger direkt für die Qualität und Quantität der produzierten Sprache verantwortlich sind (Bygate, 2009, S. 415ff.).

Für den Entwurf eines Entwicklungsprofils schlägt er vor, zwischen deklarativem (dem linguistischen und diskursiven Repertoire) und prozeduralem Wissen (der Fähigkeit, dieses anzuwenden) zu unterscheiden (u.a. basierend auf Alderson, 2007; Hulstijn, 2007; Johnson, 1996). Die Zusammensetzung, Interaktion und Entwicklung dieser beiden Bereiche könnten dann als Ausgangspunkt für die Erforschung und Beschreibung der Entwicklung von Sprechkompetenz dienen. Gleichzeitig kritisiert er, dass existierende Beurteilungsskalen und Referenzniveaus wie jene des Referenzrahmens (Europarat, 2001) oder des ACTFL (z.B. ACTFL, 2012a; 2012b) nur bedingt auf empirischer Forschung zur Entwicklung von Sprachkompetenz basieren. Um dem entgegenzuwirken, plädiert er dafür, konkrete Lernprogressionen stärker zu erforschen und die bisher geltenden Vermutungen empirisch zu untermauern oder zu widerlegen. Das Projekt „Erfassung von Entwicklungsprofilen“ kann in dieser Tradition verstanden werden, stellt aber vorerst eher die Entwicklung eines Instrumentariums für formativ nutzbare Kompetenzdiagnosen in den Mittelpunkt als die Konzipierung und empirische Stützung bzw. Überprüfung von Aussagen zur progressiven Entwicklung der Sprechkompetenz. Wenn ein solches Instrumentarium erst einmal im Einsatz wäre, könnte die über längere Zeiträume hinweg entstehende Sammlung von zahlreichen individuellen Sprachdaten und Testleistungen rasch zu einem grossen, empirisch fundierten Erkenntnisgewinn bezüglich der tatsächlichen Kompetenzentwicklung bei den Deutschschweizer Primarschülerinnen und Primarschülern führen.

20 | Linguistische und diskursive Merkmale, sprachliche Fragmentierung und emotionale Involvierung, Zeitdruck und Kooperation, Automatisierung und Planung.

III Kompetenzen im Bereich des Sprechens

2. Möglichkeiten der Erfassung von Sprechkompetenz

Insgesamt wurden im hier beschriebenen Projekt mehrere Ansätze, mündliche Kompetenz zu erfassen und zu diagnostizieren, verfolgt; teilweise mussten sie wieder verworfen werden. So orientierten sich die ersten entworfenen Aufgaben stark an Aufgabenformaten und Beurteilungsrastern der verschiedenen Lehrwerke (vgl. Abschnitt 2.1.), dies mit dem Ziel, einen Eindruck von den Sprechkompetenzen von Primarschülerinnen und -schülern zwischen der dritten und sechsten Klasse zu bekommen. Da lehrwerknahe Aufgaben kaum über das jeweilige Einsatzgebiet des Lehrwerks hinaus einsetzbar sind, wurde in der Folge ein allgemeinerer Ansatz gewählt, bei welchem handlungs- und sprachressourcenorientierte Aufgaben nebeneinanderstanden, die ergänzend eingesetzt wurden (Näheres dazu in Abschnitt 2.2.). Dabei fokussierten die Aufgaben auf Bereiche, die auf dem geringen Sprachniveau der Zielgruppe am wahrscheinlichsten Entwicklungen zeigen würden. Zum einen waren dies die Wortschatzkompetenzen, zum anderen zwei klar eingrenzbar kommunikative Handlungsbereiche: die Sprache des Unterrichts und der Austausch persönlicher Informationen (vgl. Abschnitt 2.3.).

Im Laufe des Projektes wurden nicht nur die Inhalte, sondern auch die eingesetzten Aufgabenformate schrittweise verfeinert und an die Zielgruppe junger Schülerinnen und Schüler sowie das Medium Computer angepasst. Im Folgenden sollen vor allem die inhaltliche Arbeit und die dabei gewonnenen Erkenntnisse skizziert und erläutert werden. Die technische Seite der Arbeiten wird in einem eigenen Teilbericht dargestellt (vgl. Teil II zur computerbasierten Beurteilung).

2.1. Arbeitsphase 1: Lehrwerkorientierung

2.1.1. Konstrukte

Erste Versuche, ein Testkonstrukt zu entwerfen, welches die zu erfassenden Sprechfertigkeiten beschreibt, orientierten sich in erster Linie an den Beurteilungsskalen der eingesetzten Lehrwerke sowie verschiedener existierender Testinstrumente (darunter auch den Testbeilagen der Lehrwerke). Die darin vorgeschlagenen Bewertungskriterien wurden gesammelt und zu drei Gruppen zusammengefasst, welche als (vorläufige) Teilkonstrukte definiert wurden:

- Flüssiges Sprechen,
- Richtiges Sprechen (Wortschatz, Grammatik, Aussprache),
- Aufgabenerfüllung (funktional-kommunikativer Erfolg, Kreativität).

Aufgrund der vorhandenen Literatur wurden dann für jedes Konstrukt Entwicklungserwartungen formuliert, welche eine grosse Spanne sprachlicher Kompetenz abdecken sollten. Dabei wurde auf die aus der Psychometrie stammende Idee der *construct map* (Wilson, 2005) zurückgegriffen.

III Kompetenzen im Bereich des Sprechens

TALKING WITH EASE	6	Speech ²¹ is spontaneous and independent, pauses are made at appropriate moments, conversational fillers are used to soften, structure and complete speech, hesitations and reformulations are rare, listening is very easy ²²	when producing long turns and detailed utterances	spontaneously
	5	Showing ease/fluency ²²	when producing longer turns	with almost no time to prepare
	4	Showing ease/fluency/confidence ²²	when producing longer turns	with enough time to prepare
	3	Showing some ease/fluency	when expressing short utterances	with almost no time to prepare
	2	Showing some ease/fluency	only when expressing short utterances	only with enough time to prepare
	1	Speech is hesitant, false starts and repetitions happen often	even when only producing short utterances	even with time to prepare
	0	No data		

Tabelle III-3: Entwurf einer *construct map* für das angenommene Teilkonstrukt Flüssiges Sprechen

Die Produktionen der Lernenden sollten schliesslich pro Aufgabe in Bezug auf jedes der drei Teilkonstrukte bewertet werden.

2.1.2. Aufgaben und Aufgabenerprobung

Die in diesem Projektabschnitt verwendeten Teilkonstrukte waren inhaltlich sehr offen gehalten. Die Konkretisierung der Aufgaben erfolgte in Anlehnung an Inhalte und Aufgaben in den Lehrwerken, mit denen die Schülerinnen und Schüler unseres Praxisfeldes arbeiteten und die im Wesentlichen deren sprachlich-kommunikativen Erfahrungsbereich darstellten.

Wie auch in den späteren Entwicklungsrunden wurde eine relativ kleine Anzahl Aufgaben ent-

wickelt, welche die Spanne möglicher Formate und Themen exemplarisch repräsentieren sollte. Allerdings wurde die Auslieferung am Computer in dieser Phase nur mitgedacht: Sämtliche Aufgaben wurden auf Papier entwickelt und erprobt. Dafür wurde der Versuch unternommen, möglichst viele der in den Zielregionen verwendeten Lehrwerke miteinzubeziehen, indem aufgrund einer Analyse die Themen und Formate ausgewählt wurden, welche dort häufiger vorkommen. In den meist auf Kärtchen basierenden Aufgaben mussten die Lernenden zum Beispiel:

- die Uhrzeit nennen,
- Wörter und Sätze lesen und nachsprechen,
- Bilder benennen,
- Fragen zu einem bestimmten Thema beantworten (Haustiere oder Schule),
- aus vorgegebenen Wörtern Sätze bilden, um eine bestimmte Äusserung zu machen,
- in einem Restaurant ein Verkaufsgespräch führen (in Zweiergruppen),
- aufgrund von vorgegebenen Informationen eine kurze Geschichte erzählen.

21 | Zu Beginn des Projektes wurde ein Teil der konzeptuellen Arbeit in englischer Sprache durchgeführt, um die Kommunikation zwischen den verschiedenen Partnerinstitutionen zu erleichtern. Auf eine nachträgliche Übersetzung der verschiedenen Entwürfe wurde verzichtet.

22 | Es wurde davon ausgegangen, dass Schülerinnen und Schüler der Zielgruppe, welche keine speziellen Kontakte zur Fremdsprache aufweisen, nicht über Stufe 3 hinauskommen würden.

III Kompetenzen im Bereich des Sprechens

Um aus erster Hand einen Einblick in das Funktionieren der Aufgaben zu erhalten, wurden diese durch die Projektmitarbeitenden mit einzelnen Schülerinnen und Schülern erprobt. Das Geschehen wurde mit Diktiergeräten aufgenommen und später detailliert ausgewertet. Dabei konnten auch die Fragen der Lernenden, Hilfestellungen durch die Testleitenden, Pausen, Denkzeiten usw. erfasst werden, sodass wichtige Erkenntnisse zur Gestaltung der Aufgaben sowie Anhaltspunkte zu den erwartbaren Kompetenzen der Primarschülerinnen und -schüler gewonnen werden konnten. Die Erprobung der Aufgaben fand in einer dritten Klasse im Raum Passepartout (Französisch als Fremdsprache) und in einer vierten Klasse in der Ostschweiz (Englisch als Fremdsprache) statt. Eine einzelne Aufgabe, bei der die Lernenden eine mündliche Präsentation zunächst vorbereiten sollten, wurde darüber hinaus schriftlich in einer weiteren Klasse im Raum Passepartout eingesetzt, um zu erkunden, inwiefern die Lernenden diese Vorbereitungsphase meistern würden.

Die Ergebnisse dieser ersten Erprobungen deckten eine Reihe von Problemen auf, die dazu führten, dass weder die Konstrukte noch die Aufgaben für die weiteren Untersuchungen beibehalten wurden. Zwei Punkte erwiesen sich als besonders kritisch: Zum einen ergaben sich deutliche Hinweise darauf, dass die z.T. komplexen und mehrschrittigen Aufgabenstellungen für den selbstständigen Gebrauch am Computer ungeeignet sein würden. Zum anderen wurde bei der Erprobung v.a. in den unteren Klassenstufen deutlich, dass sich Erwartungen an eine spontane Sprachverwendung auf sehr wenige spezifische Bereiche beschränken mussten. Aufgabenstellungen, die auch fortgeschritteneren Lernenden eine Herausforderung bieten wollten, erwiesen sich als wenig geeignet, um frühe Entwicklungen und subtile Unterschiede zwischen den zahlreichen Lernenden mit sehr geringen mündlichen Ausdrucksmöglichkeiten effizient zu erfassen.

Aufgrund dieser und anderer Erkenntnisse wurde der Entwicklung der Instrumente in der Folge ein Modell von Sprechkompetenz zugrunde gelegt, welches stärker auf die Phase beginnender Sprachkompetenz ausgerichtet ist und Kompe-

tenzelemente berücksichtigt, die zu Beginn des Lernprozesses, unabhängig von den jeweiligen Lehrwerken, eine entscheidende Rolle spielen.

2.2. Arbeitsphase 2: Ressourcenbezogenes Testen

2.2.1. Konstrukte

Den Ausgangspunkt für die neu zu fokussierenden Konstrukte bildete eine detaillierte Analyse der sprachlich-kommunikativen Komponenten der Aufgaben, mit denen in den Lehrwerken die Fähigkeit „Sprechen“ gefördert bzw. überprüft werden sollte. Dabei wurden Elemente wie allgemeiner und thematischer Wortschatz, grammatische Strukturen und Sprachfunktionen kodiert²³ und ggf. speziell thematisierte Aussprachephänomene bzw. Flüssigkeitsaspekte festgehalten. Aufgrund der dabei entstandenen Übersichten konnten Wissens- und Könnensbereiche identifiziert werden, welche im Fremdsprachenunterricht der Primarschulen offenbar besonders grosse Relevanz haben.²⁴

Vor dem Hintergrund der Analyse der Lehrwerke wurden drei hierarchische Ebenen festgelegt, zu denen Aufgaben entwickelt werden sollten:

- Linguistische Ressourcen: Entsprechende Aufgaben sollten gezielt einen bestimmten sprachlichen Kenntnis- und Fertigungsbereich (z.B. Wortschatz oder Aussprachefähigkeit) erfassen;

23 | Die Kategorien basierten auf dem Referenzrahmen (Europarat, 2001) bzw. den Listen in *Threshold Level* (Ek u.a., 1980).

24 | Zum Beispiel die kommunikativen Funktionen „Erfragen von Informationen“ und „Identifizieren/Benennen“, die Wortschatzfelder „Zahlen“, „Farben“ oder „Sport und Hobbys“ oder grammatische Phänomene wie das Konjugieren von Verben oder die Anwendung von Pronomen.

III Kompetenzen im Bereich des Sprechens

- b. Sprachfunktionen: Diese Aufgaben sollten auf die Realisierung einer spezifischen Handlung mit einer eng umgrenzten Funktion abzielen (z.B. Begrüssen oder Auffordern);
- c. Kommunikatives Handeln: Im Rahmen von quasi-kommunikativen Aufgaben sollten verschiedene vorhandene linguistische und funktionale Ressourcen und Funktionen kombiniert werden müssen, um ein kommunikatives Ziel zu erreichen.

2.2.2. Aufgaben und Aufgabenerprobung

Für eine erste Erprobung wurden beispielhaft je einige auf linguistische Ressourcen bezogene, funktionsorientierte und kommunikative Aufgaben entwickelt. Die an den Erprobungen beteiligten Kinder wurden daher mit einer Auswahl von Aufgaben konfrontiert, bei denen sie:

- Objekte und Handlungen benennen,
- einzelne Wörter laut lesen oder nachsprechen,
- aus durcheinander geratenen Wörtern Aussagen oder Fragen bilden,
- aufgrund von Situationsbeschreibungen oder Sätzen in der Schulsprache kurze mündliche Reaktionen in der Fremdsprache zeigen,
- oder einen einfachen Dialog rekonstruieren sollten.

Die Anweisungen zu diesen Aufgaben wurden sowohl in der Schulsprache als auch in der Fremdsprache entwickelt, um zu erkunden, ob dies einen sichtbaren Einfluss auf das Verhalten der Lernenden haben würde.

Die Aufgaben wurden den Lernenden mittels PowerPoint-Präsentationen (Microsoft, 2010) vorgelegt. Die mündlichen Leistungen wurden mit der Computersoftware Audacity (The Audacity Team, 2013) aufgenommen und später inhaltlich transkribiert und ausgewertet. Die Erprobung wurde auf mit-

geführten Laptops des Projektteams in den Räumen der beteiligten Schulen durchgeführt. Es nahmen insgesamt 114 Mädchen und Jungen aus neun Klassen in sieben Schulhäusern teil, wobei jeweils etwa die Hälfte aus der Deutschschweiz (Kanton Freiburg) und dem Tessin stammte. Die Aufgaben wurden für die erste Fremdsprache Französisch gelöst. Nur eine kleine Anzahl Sechstklässlerinnen und Sechstklässler bearbeitete sie in englischer Sprache.

2.2.3. Allgemeine Erkenntnisse aus der Erprobung

Da bei diesen Erprobungen erstmals Computer im Einsatz waren, konnten erste Erfahrungen zum Umgang der jungen Schülerinnen und Schüler mit der Technik und mit computerbasierten Testaufgaben gesammelt werden.

Zunächst ist festzuhalten, dass alle beteiligten Kinder ausreichend mit Computern umgehen konnten, um die Aufgaben zu handhaben: Die Benützung von Mäusen oder Touchpads an Laptops stellte kein Problem dar; die z.T. sehr grossen Headsets wurden häufig mit Begeisterung aufgenommen und führten nicht zu sichtbaren Berührungängsten.

Die schriftlichen Aufgabenstellungen waren hingegen für eine Reihe von Schülerinnen und Schülern auch in dieser Arbeitsphase noch zu anspruchsvoll bzw. wurden nicht oder nur nach Aufforderung gelesen; mündliche Anweisungen durch die Aufsichtspersonen wurden deutlich besser aufgenommen. Dies galt umso mehr, je jünger die beobachteten Kinder waren.

Anweisungen in der Zielsprache wirkten störend: Insbesondere jüngere Lernende konnten den Unterschied zwischen Aufgabenstellung und eigentlicher Aufgabe nicht erfassen und verbrachten viel Zeit damit, die fremdsprachigen Anweisungen zu „übersetzen“, selbst wenn sie die Aufgabenstellung eigentlich schon erkannt hatten. Es wurden hingegen keine wesentlichen negativen Reaktionen beobachtet, wenn die Arbeitsanweisungen in der Schulsprache abgefasst waren.

Beispiele am Bildschirm führten z.T. zu Verwirrung bzw. halfen nur wenigen Kindern, sich in der

III Kompetenzen im Bereich des Sprechens

Aufgabe besser zurechtzufinden. Dasselbe gilt für sprachliche Hilfestellungen, z.B. Satzanfänge, welche die eigene Äusserung einleiten sollten. Sie halfen starken Lernenden nicht weiter und verunsicherten schwache.

2.2.4. Ergebnisse der Aufgabenerprobung

Allgemein konnte festgestellt werden, dass die deutschsprachigen Kinder im Kanton Freiburg mit dem Lösen der Aufgaben erheblich grössere Probleme hatten als die Schülerinnen und Schüler im Tessin. Dies hat sicher ganz wesentlich mit der unterschiedlichen Distanz der beiden Schulsprachen Deutsch bzw. Italienisch zum Französisch zu tun. Allerdings ist es auch möglich, dass das Lehrwerk *Mille feuilles*, welches zum Zeitpunkt der Untersuchung im Kanton Freiburg gerade neu eingeführt wurde, andere Fertigkeiten und Inhalte in das Zentrum des Unterrichtsgeschehens stellt, als in den Aufgaben erwartet wurde.

Der Einfluss des Lehrwerks wurde bei den Wortschatzaufgaben, sowie einer Aufgabe, bei der sich die Schülerinnen und Schüler selbst vorstellen sollten, besonders deutlich: Während die italienischsprachigen Kinder keine Probleme hatten, die Fragen „Comment tu t'appelles?“ oder „Quel âge as-tu?“ zu beantworten, stellten diese für die Deutschfreiburger Kinder unüberwindbare Hürden dar, wohl weil sie in dieser Form nicht zum Grundwortschatz ihres Lehrwerks gehören. Umgekehrt lösten die deutschsprachigen Kinder verschiedene Wortschatzitems aus dem Bereich *Schule und Klassenzimmer* offensichtlich deshalb besser, weil ihnen die Wörter aufgrund des Lehrwerks aus dem Unterricht vertraut waren.

Bei den Aufgaben zur Aussprache wurden in beiden Regionen vor allem Probleme beim Vorlesen von Wörtern und Sätzen beobachtet, während sich beim Nachsprechen die Abweichungen von der standardmässigen Aussprache auf sehr spezifische Phänomene beschränkten. Ersteres lässt auf Schwächen bei der Verbindung von Laut und Schrift schliessen, welche wohl auch auf das Alter der Lernenden zurückzuführen sind. Das Nachsprechen der franzö-

sischen Wörter offenbarte so wenige Probleme, dass in einem diagnostischen Testsystem auf eine Erhebung der entsprechenden Kompetenzen entweder generell verzichtet oder diese auf einige wenige Problemlaute beschränkt werden könnte.

Aufgaben und Items, für deren Lösung zunächst eine längere Aufgabenstellung gelesen und verstanden werden musste, wurden häufig schlechter gelöst als solche in anderen Formaten. Ähnliches gilt für Formate, bei denen sich die Kinder in andere Personen oder fiktive Situationen hineinversetzen sollten. Dies war insbesondere für die jüngeren Schülerinnen und Schüler offenbar sehr anspruchsvoll und konnte auch mit mündlicher Unterstützung durch die Testleitenden nicht kompensiert werden. Dort, wo dennoch verwertbare Antworten entstanden, musste festgestellt werden, dass längere, weitgehend spontane Äusserungen kaum jemals entstanden. Sie scheinen in aller Regel noch nicht in Reichweite der jungen Lernenden zu liegen.

Schliesslich offenbarten die Aufgabenerprobungen auch, dass viele Kinder sehr klare Vorstellungen darüber haben, was sie wissen und was nicht. Dort, wo sie den Eindruck hatten, die in der Aufgabe enthaltenen Wörter *nicht kennen zu können*, versuchten sie auch nicht zu verstehen. So erzeugten z.B. Aufgaben, welche die Fähigkeit überprüfen sollten, aus einzelnen Wörtern (syntaktisch richtige) Aussagen oder Fragen zu bilden, zwar trennscharfe Ergebnisse,²⁵ es blieb aber unklar, welchen Anteil die hauptsächlich angezielten syntaktischen Kompetenzen am Erfolg oder Misserfolg hatten. Die Erprobungen legen den Schluss nahe, dass die Kenntnis der vorgegebenen Wörter eine wichtige Rolle spielte und dass sprachstrukturelle Kenntnisse (z.B. die Unterscheidung von Nomen und Verben anhand ihrer äusseren Merkmale oder das Wissen über die Reihenfolge von Artikeln und Nomen) Wortschatzlücken nicht auszugleichen vermochten.

Die Auswertungen zu allen Aufgaben führten insgesamt zum Schluss, dass Wortschatz ein sehr zentrales Element im sprachlichen Kompetenzge-

25 | Nur auffallend starke Schülerinnen und Schüler konnten (kürzere) Sätze bilden. Kinder mit einem privaten Hintergrund in Französisch erzielten die mit Abstand besten Ergebnisse. Alle anderen lösten nur sehr wenige Items erfolgreich.

III Kompetenzen im Bereich des Sprechens

füge der Primarschülerinnen und -schüler ist: Nur wenn die Bedeutung der vorgegebenen Wörter und Sätze zumindest erschlossen werden konnte, zeigten sich die Schülerinnen und Schüler bereit und in der Lage, auswertbare Antworten zu produzieren.

2.3. Arbeitsphase 3: Erfassung einer gezielten Auswahl von mündlichen Kompetenzen

2.3.1. Konstrukte

Aufgrund der oben beschriebenen Erkenntnisse, v.a. der oft unerwartet bescheidenen kommunikativen Leistungen anlässlich der Erprobungen wurde im Hinblick auf die letzte Runde der Aufgabenentwicklung die Entscheidung getroffen, das Sprechen-Konstrukt enger zu fassen und kommunikative Aufgaben nur noch auf zwei Bereiche zu beziehen, nämlich die Sprache im Klassenzimmer und den Austausch von Informationen zur eigenen Person. Zudem sollte die lexikalische Kompetenz als sichtlich entscheidende Teilkompetenz weiterhin erfasst werden. Als neues Teilkonstrukt der Sprechkompetenz wurde das Verstehen von gesprochener Sprache hinzugenommen. Damit war die Hoffnung verbunden, auch auf dem sehr bescheidenen Sprachniveau eines Teils der Primarschülerinnen und -schüler Ansätze einer Entwicklung sichtbar machen zu können, die nach und nach zu eigenen Sprachproduktionen führen würde.

Im Folgenden werden sowohl Aufgaben zum Wortschatz als auch solche zur mündlichen Kommunikationsfähigkeit in den ausgewählten Bereichen vorgestellt und kurz besprochen.

Wegen der Lehrwerkabhängigkeit einerseits der Kenntnisse der Lernenden und andererseits der möglichen Aufgabenstellungen sowie auch aus logistischen Gründen konzentrierten sich die Arbeiten in dieser Phase auf die Lehrplanregion Passepartout und damit auf die Inhalte des Lehrwerks *Mille feuilles*. Die weniger spezifischen Erkenntnisse aus den Erprobungen lassen sich

aber mit hoher Wahrscheinlichkeit auf andere Lehrplanregionen übertragen.

Wortschatz

Mit „Wortschatz können“ ist insbesondere bei den jüngsten Lernenden durchaus die klassische Vorstellung verbunden, ein in der Erstsprache (bzw. in der Schulsprache) vorhandenes Konzept mit einem fremdsprachlichen Wort ausdrücken zu können. Der Umstand, dass lexikalische Elemente in aller Regel zunächst rezeptiv „gekonnt“ werden und erst mit der Zeit vielleicht auch produktiv zur Verfügung stehen, ist vielen weniger bewusst. Mit steigender fremdsprachlicher Kompetenz wachsen sowohl die Quantität als auch die Qualität des Wortschatzes: Die Lernenden sind fähig, *mehr* fremdsprachliche Begriffe zu gebrauchen, sie können diese aber auch schneller, sicherer und formell korrekter verwenden und können nach und nach den vorhandenen Wortschatz stärker ausdifferenzieren. So werden lexikalische Elemente zunehmend leichter verknüpft, Bedeutungsunterscheidungen zielsprachlich vorgenommen und auch Konzepte gemeistert, die in der L1 so nicht vorhanden sind.

Die Entwicklung von lexikalischer Kompetenz vollzieht sich sowohl für konkrete Begriffe, die sich mit einem oder wenigen Wörtern ausdrücken lassen (*écrire, l'éléphant, jouer au foot*), als auch für längere Wendungen oder ganze Sätze (z.B. „Je m'appelle ...“, „C'est ton tour.“ oder „Cinq chiens chassent six chats.“²⁶). Welche Wörter und Wendungen die Schülerinnen und Schüler jedoch erlernen und zu welchem Grad sie diese rezeptiv bzw. produktiv verwenden können, ist in erheblichem Masse vom Lehrwerk, der Lehrperson und den sonstigen Sprachkenntnissen einzelner Lernender abhängig. Teilweise mögen auch Faktoren wie das persönliche Interesse an bestimmten Themen oder andere individuelle Umstände mit hineinspielen.

Für die Erstellung der Aufgaben wurden aufgrund der Resultate einer Korpusanalyse des Lehrwerks *Mille feuilles* und weiterer Analysen der darin vorkommenden Aufgabenstellungen prominente Wort-

26 | Ein Zungenbrecher aus *Mille feuilles* 3.1.1.1, S. 54.

III Kompetenzen im Bereich des Sprechens

schatzelemente aus drei Bereichen ausgewählt: Tiere, häufige Verben und lokale Präpositionen. Für die grösseren Erprobungen im Juni und Oktober 2014 wurden gezielt Parallelwörter, Begriffe aus dem Schulalltag sowie das Thema Zirkus hinzugenommen. Bei all dem handelt es sich um Wortschatzbereiche, mit denen die Primarschülerinnen und -schüler in den vier Lernjahren immer wieder konfrontiert sind. Die getroffene Auswahl hat trotzdem bloss Beispielcharakter und ist nicht etwa als repräsentative Stichprobe aus dem Wortschatz von *Mille feuilles* zu verstehen. Ein tatsächlicher, repräsentativ-diagnostischer Wortschatztest müsste erheblich mehr Sprachmaterial umfassen und dieses gezielt aus bestimmten Schul- bzw. Schulhalbjahren selektionieren, um diagnostisch relevant zu sein und auch stärker zwischen den einzelnen Lernenden zu differenzieren. Zudem sollte der Bestand an Testitems so gross sein, dass damit wiederholt – und progressiv – individuelle Tests erzeugt werden könnten. Für die Erprobung der allgemeinen Funktionalität des Tests und seiner Formate eigneten sich aber die im Projekt ausgewählten Wortschatzaufgaben durchaus.

Handlungsorientierte Aufgaben I: Sprache des Unterrichts

Das Handlungsfeld „Sprache im Unterricht“ nimmt im Lehrplan Passepartout (Passepartout, 2015) eine Sonderstellung ein. Da die Verwendung von mehr oder weniger ritualisierten Äusserungen im Unterrichtsgeschehen stark von der Lehrperson abhängt, konnte die Auswahl der Wendungen für die Aufgabenerprobung nicht nur aufgrund des Lehrwerks vorgenommen werden. Stattdessen wurde eine relativ umfangreiche Befragung von Lehrpersonen durchgeführt. Diese wird in einem weiteren Teilbericht näher beschrieben (vgl. Teil IV „Langage de classe“). Eines der Ergebnisse dieses Projektteils sind Listen von Äusserungen, die Auskunft darüber geben sollen, welche Elemente der Sprache des Unterrichts die Lernenden eines Schuljahres rezeptiv und/oder produktiv aus Sicht von Lehrpersonen bereits beherrschen sollten. Diese Listen dienten der hier beschriebenen Aufgabenerstellung als Grundlage. Wiederum wurden

sowohl Aufgaben zum Hörverstehen als auch zur mündlichen Produktion entworfen.

Handlungsorientierte Aufgaben II: Austausch persönlicher Informationen

Ein Überbleibsel aus den Aufgabenentwicklungen von 2013 ist das Aufgabenformat „Dialog“, in welchem die Lernenden einem virtuellen Kind Fragen zu ihrer Person beantworten bzw. diesem entsprechende Fragen stellen sollen. Dieses wurde beibehalten, da verschiedene Sprachaktivitäten vom Typ „persönliche Informationen austauschen“ zu den Grundkompetenzen der Lernenden gehören. Weiter war aber auch der computerbasierte Dialog als Aufgabenformat an sich von Interesse. Er hatte vom Prinzip her schon 2013 gut funktioniert und stellt einen wichtigen Schritt hin zur Simulation „echter“ Kommunikation dar.

In der Aufgabe „Dialog“ wurde den Schülerinnen und Schülern der Region Passepartout, anders als 2013, insofern speziell Rechnung getragen, als auf sprachlich offenbar intransparente Wendungen wie „Quel âge as-tu?“ verzichtet wurde und stattdessen Fragen im Stil „Moi, j'ai 10 ans. Et toi?“ gestellt wurden.

2.3.2. Formate

Aufgrund des Projektziels, der Literaturanalyse und der Erfahrungen aus den früheren Erprobungen wurden folgende Bedingungen für die Auswahl der Formate formuliert:

- computerbasiert;
- kognitiv wenig anspruchsvoll:
 - (möglichst) keine geschriebenen Prompts,
 - kleines Spektrum an unterschiedlichen Formaten bzw. einander ähnliche Formate,
 - spielerische Elemente (v.a. Bilder);
- eindeutig interpretierbar:
 - Prompts und Aufgabenstellungen mit

III Kompetenzen im Bereich des Sprechens

minimalem Interpretationsspielraum für die Schülerinnen und Schüler,

- Antwortmöglichkeiten, die eine einfache Auswertung der mündlichen Leistungen durch die Lehrperson zulassen (richtig oder falsch, wenig dazwischen).

Diese Kriterien führten letztlich zu einer eher kleinen Anzahl unterschiedlicher Formate, welche zudem nur wenig innovativ wirken mögen. Mitverantwortlich für die bescheidene Anzahl an Formaten ist auch die aufwändige und technisch anspruchsvolle Umsetzung der Aufgaben zum Sprechen. Die Möglichkeit, gesprochene Sprache mittels eines herkömmlichen Internetbrowsers aufzunehmen, ist erst seit kurzem Teil der Software *Item Builder*, weshalb die Einbindung noch relativ umständlich ist (vgl. Teil II zur computerbasierten Beurteilung).

Mit Ausnahme der Interaktionsaufgabe „Dialog“ sind alle Aufgabenformate auf einer einzelnen Bildschirmseite dargestellt. Sie enthalten ein Minimum an geschriebenen Anweisungen (und zwar in der Schulsprache) und basieren weitestgehend auf Bildern als Stimuli und Antwortoptionen. Die meisten Aufgabenformate können in produktive und rezeptive unterteilt werden: Bei produktiven Aufgaben müssen die Kinder selbst etwas in der Fremdsprache sagen, bei rezeptiven Aufgaben lediglich etwas verstehen. Bei den rezeptiven Aufgaben handelt es sich im Kern um Multiple-Choice-Aufgaben in verschiedenen Varianten, welche aus einer oder mehreren Play-Schaltflächen zum Starten der relevanten Audioaufnahme(n) und einem oder mehreren Bildern bestehen, die diesen Audioaufnahmen zuzuordnen sind. Produktive Aufgaben sind üblicherweise um ein Bild oder eine kurze Bildersequenz sowie eine Schaltfläche zum Starten der Aufnahme herum aufgebaut. Eine Ausnahme bildet die Aufgabe „Dialog“, welche sich über mehrere Seiten erstreckt und bei der die Lernenden abwechselnd hören und sprechen müssen. Alle Aufgaben sind so gestaltet, dass die Aufgabenstellung möglichst auf den ersten Blick klar wird bzw. keine alternativen Lösungswege eingeschlagen werden können.

2.3.3. Entwicklungsschritte

Insgesamt wurde die letzte Aufgabenserie im Laufe des Jahres 2014 in vier Schritten entwickelt, wobei ein nicht unerheblicher Teil der Arbeit auch technischen und logistischen Fragestellungen gewidmet war, dies hauptsächlich in Verbindung mit dem Einsatz von *CBA Item Builder* und seiner Auslieferungsoftware EE4CBA (vgl. wiederum Teil II zur computerbasierten Beurteilung).

Die Vorerprobungen im März und April 2014 dienten vor allem einer ersten Validierung der Formate und dem Erproben des Aufgabendesigns mit einer kleinen Anzahl von Schülerinnen und Schülern. Im Juni wurden dann vor allem die Funktionalität der serverbasierten Auslieferung über die Infrastruktur der Schulen sowie die Durchführbarkeit der Tests in grösseren Gruppen bzw. im Klassenverbund erprobt. Erst im Oktober und November 2014 konnte der Fokus wieder stärker auf inhaltliche Aspekte der Aufgaben gelegt werden, wobei vor allem anspruchsvollere Versionen bereits früher erprobter Aufgaben und das oben beschriebene Aufgabenformat „Dialog“ von Interesse waren. Ausserdem wurde der Frage nachgegangen, welchen Effekt das Zulassen der Schulsprache bei produktiven Aufgaben als Alternative beispielsweise bei fehlendem Wortschatz auf das Antwortverhalten haben würde.

2.3.4. Erkenntnisse

Über alle vier Erprobungen und die relativ hohe Zahl einzelner Aufgaben hinweg lässt sich eine Reihe von Aussagen über das Funktionieren der Formate, die Angemessenheit der Bilder und Themen, Erweiterungsmöglichkeiten sowie die Bewertung der Schülerantworten treffen. Auch war es möglich, die Ergebnisse zu den rezeptiven Aufgaben, die im Sommer und Herbst des letzten Projektjahres eingesetzt wurden, zumindest grundlegend statistisch auszuwerten, sodass Aussagen vor allem zur Schwierigkeit der einzelnen Aufgaben sowie zu Unterschieden zwischen den Schülerinnen und Schülern der verschiedenen Schuljahre gemacht werden können.

III Kompetenzen im Bereich des Sprechens

Allgemeine Beobachtungen

Wie bereits zuvor, bestätigten die Erprobungen von 2014 die Feststellung, dass Lernende aller vier Klassenstufen rund eine Lektion am Computer arbeiten können, ohne dass merklich Unruhe aufkame. Auch die Reduzierung der Aufgabenstellungen auf das Allernötigste, sowie die Vereinfachung der Formate erwies sich als zielführend: Im Vergleich zu früheren Erprobungen brauchten die Schülerinnen und Schüler bei der Testdurchführung deutlich weniger Erklärungen und Hilfestellungen. Auch die Zahl der klar auf missverständliche Aufgabenstellungen zurückzuführenden Antworten war deutlich geringer als zuvor. Da im Jahr 2014 erstmals auch fünfte und sechste Klassen an den Erprobungen teilnahmen, konnte festgestellt werden, dass diese etwas älteren Lernenden geschriebene Anweisungen bereits deutlich besser lesen und verstehen können als die jüngeren Schülerinnen und Schüler. Aus diesem Grund konnten im Oktober auch etwas komplexere, nicht rein intuitiv verständliche Aufgabenstellungen erfolgreich erprobt werden.

Bei den Erprobungen im Oktober und November 2014 konnte beobachtet werden, dass sich insbesondere Mädchen in den sechsten Klassen offenbar in einem Alter befanden, in dem die spielerische Komponente der Aufgaben keinen Eindruck auf sie machte und das Sprechen der Fremdsprache im Beisein von Mitschülerinnen und Mitschülern sie peinlich berührte. Es wäre zu untersuchen, wie sehr und während welcher Zeitdauer die Aussagekraft von formativ-diagnostischen Tests zum Sprechen dadurch tangiert wird.

Weiterhin sei festgehalten, dass die Formate im Laufe des Jahres 2014 so umfassend erprobt werden konnten, dass eine inhaltliche Erweiterung des Instruments möglich ist, ohne dass das Funktionieren der Aufgaben grundsätzlich neu überprüft werden muss. Möglicherweise problematisch ist, dass die meisten Aufgabenstellungen entscheidend auf Bildern basieren: Deren Gestaltung ist sehr aufwändig; zudem können nicht alle Inhalte bildlich dargestellt werden. So erwies sich zum Beispiel bereits die Überprüfung der Kenntnis der lokalen Präpositionen (*auf, unter, neben* usw.) als

relativ schwierig. Immerhin ist zu vermuten, dass dies auch auf fehlende Wortschatzkenntnisse der Lernenden, und nicht nur auf Mängel in der Präzision der Illustrationen zurückzuführen ist.

Schliesslich ergaben die Erprobungen vor allem auch in technischer Hinsicht wichtige Erkenntnisse, die für zukünftige Projekte von grosser Bedeutung sind (mehr Informationen dazu sind im Teilbericht II zur computerbasierten Beurteilung zu finden).

Ergebnisse zu den rezeptiven Aufgaben

Aufgrund der relativ geringen Probandenzahlen sind die statistischen Aussagen zu den Aufgaben mit Vorsicht zu interpretieren.

Da ein relativ grosser Anteil der rezeptiven Aufgaben mit dem Ziel entwickelt wurde, selbst geringste Kenntnisse in der Fremdsprache aufzuzeigen, ist es wenig verwunderlich, dass sowohl im Sommer als auch im Herbst 2014 ein deutliches Übergewicht (zu) leichter Aufgaben festzustellen war. Zu beiden Zeitpunkten bewegt sich die mittlere Schwierigkeit der skalierten Items unter der mittleren Fähigkeit der Lernenden. Das bedeutet, dass die mittlere Lösungswahrscheinlichkeit bei vielen Aufgaben über 50% lag, was grundsätzlich für das Potential der Aufgaben, stärkere von schwächeren Schülern zu trennen, nicht optimal ist.²⁷ Im Juni stellten v.a. einige wenige Wörter wie *travailler, choisir, le pion* oder *rêver* eine Herausforderung dar. Es kann nicht ganz ausgeschlossen werden, dass diese Wörter auch deshalb auffällig waren, weil die Bilder nicht absolut eindeutig identifizierbar waren. Parallelwörter wie *éléphant* und einzelne, im Lehrwerk sehr häufig auftretende

²⁷ | Wenn allerdings, wie im Unterrichtskontext üblich, überprüft werden soll, wie weitgehend das beherrscht wird, was eigentlich beherrscht werden sollte, dann ist eine nach oben verschobene Personenverteilung erstrebenswert. Ein anderes Ergebnis würde darauf hinweisen, dass die Lernziele von vielen nicht erreicht wurden. Beim spezifisch diagnostischen Testen liegt der Fall wiederum etwas anders: Dort sollen vor allem auch Defizite aufgedeckt werden, um anschliessend gezielt eingreifen zu können. Diesem Zweck könnten die vorliegenden Items nur eingeschränkt dienen.

III Kompetenzen im Bereich des Sprechens

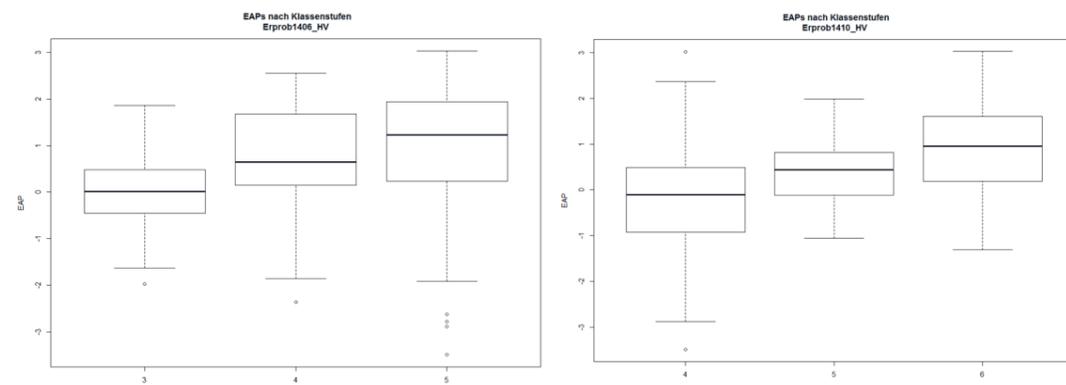


Abbildung III-1: Boxplots der Lernerfähigkeiten (EAP-Personenwerte) nach Klassenstufe

Wörter (z.B. *vache*, *ordinateur*, *école*) bewegten sich im untersten Bereich der Schwierigkeit. Entsprechende Items wurden von einer sehr grossen Mehrheit der Lernenden korrekt gelöst und differenzieren daher selbst bei sehr schwachen Lernenden kaum. Mit dem Ziel einerseits einer stärkeren Diagnoseorientierung und andererseits einer stärkeren Differenzierung zwischen den Lernenden wurden für die letzte Erprobung im Oktober/November 2014 Aufgaben entworfen, die zusätzliche Schwierigkeiten aufwiesen (z.B. wurden die gesuchten Wörter in Sätzen oder kurzen Texten präsentiert). In der Tat entsprach die Verteilung der Itemschwierigkeiten im Oktober 2014 besser der Verteilung der Lernenden, blieb aber, v.a. für eine diagnostische Aufgabenverwendung, weiterhin etwas tief.

Trotz der relativ geringen Zahl an Lernenden pro Schuljahr wurden auch die Ergebnisse der Schülerinnen und Schüler verschiedener Schuljahre verglichen. Damit sollte festgestellt werden, ob die Aufgaben den zu erwartenden Lernfortschritt der Schülerinnen und Schüler abbilden könnten. Zusätzlich wurde überprüft, ob die Aufgaben, welche in mehreren Schuljahren eingesetzt wurden, dafür auch geeignet waren.

Nach Ausscheidung einiger weniger ungeeigneter Items lässt sich feststellen, dass die Unterschiede zwischen den Klassen zwar sichtbar werden und der Median der Personenfähigkeiten (fette Linie in den Kästen) von den jüngeren zu den älteren Schülerinnen und Schülern steigt (vgl.

Abbildung III-1). Wegen der bereits erwähnten kleinen Stichproben (v.a. denjenigen der Fünft- und Sechstklässler/innen im Oktober/November 2014, vgl. Tabelle III-4) und der eher geringen Anzahl Aufgaben ist der gemessene Zuwachs aber statistisch nicht signifikant. Die vorliegenden Ergebnisse zu den Items erlauben also die Zuordnung der Aufgaben zu den verschiedenen Schuljahren bzw. zu einem bestimmten Fähigkeitsniveau, was natürlich für ein einsatzbereites Beurteilungsinstrument wichtig ist.

Ergebnisse zu den produktiven Aufgaben

Die Aufgaben zur mündlichen Produktion haben insgesamt relativ gut funktioniert. So war unter den Antworten der Lernenden zwar die eine oder andere abwegig, die grosse Mehrheit der Schülerinnen und Schüler hat die Aufgabenstellungen und Stimuli aber erwartungsgemäss interpretiert. Geht man davon aus, dass ein tatsächliches Beurteilungsinstrument im Unterricht wiederholt eingesetzt würde und den Lernenden die Handhabung und die Formate dadurch vertraut wären, kann angenommen werden, dass die eingesetzten Aufgabenformate so funktionieren wie vorgesehen.

Bei einzelnen Aufgaben müsste allerdings das Spektrum der akzeptierten Antworten erweitert und/oder die zugrundeliegende Aufgabenstellung (inkl. Bildmaterial) überarbeitet werden. So

III Kompetenzen im Bereich des Sprechens

wurde zum Beispiel bei der Aufgabe „CL_prod_dessinelmage“ (vgl. Abbildung III-2) eine relativ grosse Bandbreite verschiedener Antworten beobachtet. Einzelne Lernende antworteten schlicht mit „le crayon“, andere sagten „colorie le [Bild]“ oder „crayez une palm“, während wieder andere sich an „dessine sur la [Blatt]“ versuchten. Keine dieser Antworten kann als gänzlich falsch angesehen werden („Du bittest [jemanden], ein Bild zu zeichnen.“). Das oben dargestellte Bewertungsraster kann der Vielfalt der tatsächlich gemachten Äusserungen aber nur unzureichend gerecht werden, was zu einem erheblichen Informationsverlust führt.

Bei der Erprobung im Oktober/November 2014 wurde auch untersucht, welchen Einfluss es haben könnte, wenn den Schülerinnen und Schülern erlaubt wäre, im Fall von fehlenden sprachlichen Mitteln im Französisch auch in ihrer Schul-

sprache Deutsch zu antworten. Zu diesem Zweck wurden sämtliche produktiven Aufgaben in zwei Fassungen erstellt: eine mit der Erlaubnis, Deutsch zu sprechen (siehe Abbildung III-3), eine ohne. Die beiden Testversionen wurden zufällig auf die Lernenden verteilt. Prinzipiell ist festzustellen, dass das Zulassen des Deutschen keinen erheblichen Einfluss auf das Antwortverhalten der Lernenden hatte. Zwar sind in dieser Testversion etwas mehr deutschsprachige Antworten zu beobachten als bei der anderen, ihre Zahl hält sich aber in Grenzen. Dort, wo die Antwort sehr leicht war (z.B. wenn die Kinder sagen sollten, wie sie heissen, oder beim Benennen einzelner Objekte), hat fast kein Kind zum Deutschen gegriffen. Auch die Zahl der Schülerinnen und Schüler, die „Ich weiss nicht.“ angeklickt haben, ist dort, wo es eigentlich möglich gewesen wäre, Deutsch zu sprechen, nicht kleiner.

	3. Klasse	4. Klasse	5. Klasse	6. Klasse	gesamt
Juni/Juni 2014	102	113	100		315
	32.4%	35.9%	31.7%		
Oktober/November 2014		124	47	50	221
		56.1%	21.3%	22.6%	

Tabelle III-4: Anzahl der Schülerinnen und Schüler nach Klassenstufe bei den abschliessenden Erprobungen (2014)



Abbildung III-2: Screenshot der Aufgabe CL_prod_dessinelmage



Abbildung III-3: Erster Bildschirm der Aufgabe SP_dial_m_3qu_de. Der grüne Balken enthält die Erlaubnis, Deutsch zu sprechen.

III Kompetenzen im Bereich des Sprechens

Aus den Beobachtungen kann daher geschlossen werden, dass die Lernenden in Aufgaben, die auf die produktive Verwendung der Fremdsprache abzielen, kein besonderes Bedürfnis verspüren, ihre Schulsprache einzusetzen, bzw. es vorziehen, keine Antwort zu geben, wenn sie die Antwort nicht in der Fremdsprache geben können. Interessant sind in diesem Zusammenhang allerdings Aufgaben, bei denen mehrere sprachliche Elemente kombiniert werden mussten (z.B. „Écoute le texte et numérote les images.“). Bei diesen liess sich vermehrt beobachten, dass die Lernenden schwierige Elemente (im Beispiel v.a. das Wort „image“) auf Deutsch sagten und so immerhin die drei anderen Teile der Antwort produzierten. Es wäre wohl aus diagnostischer Sicht wünschenswert, dieses Phänomen näher zu beleuchten, um festzustellen, ob das Zulassen der Schulsprache in diesen Fällen tatsächlich zu mehr auswertbarem Material in der Fremdsprache führt.

Aufgrund der geringen zur Verfügung stehenden Testzeit und der Vielzahl an Aufgaben, welche erprobt wurden, war eine statistische Auswertung der produktiven Aufgaben wenig zielführend

und so der Vergleich der produktiven und der rezeptiven Ergebnisse nur bedingt möglich. In Abbildung III-4 wurde der Versuch unternommen, die Ergebnisse der Lernenden auf rudimentäre Weise vergleichbar zu machen, indem bei jedem Kind für beide Fertigkeitsbereiche die Anzahl der als korrekt kodierten Aufgaben durch die Anzahl der insgesamt bearbeiteten Aufgaben dividiert wurde. Bei den produktiven Aufgaben wurde dabei die Angabe zum kommunikativen Erfolg einer Antwort als Kriterium verwendet und auf 0 (nicht erfolgreich bzw. teilweise erfolgreich) und 1 (erfolgreich) reduziert. Die so entstandenen Werte geben für jedes Kind Auskunft über den jeweiligen Anteil der erfolgreich gelösten Aufgaben. Die in Abbildung III-4 dargestellten Ergebnisse sind insofern unterschiedlich zuverlässig, als die Anzahl der gelösten produktiven Aufgaben zwischen 4 und 11 und diejenige der rezeptiven Aufgaben zwischen 8 und 40 liegt und diese wiederum unterschiedlich anspruchsvoll waren. Offensichtlich ist, dass ein Teil der rezeptiven Aufgaben sehr einfach war (die meisten Schülerinnen und Schüler lösten mehr als die Hälfte der Aufgaben korrekt). Die Resultate der

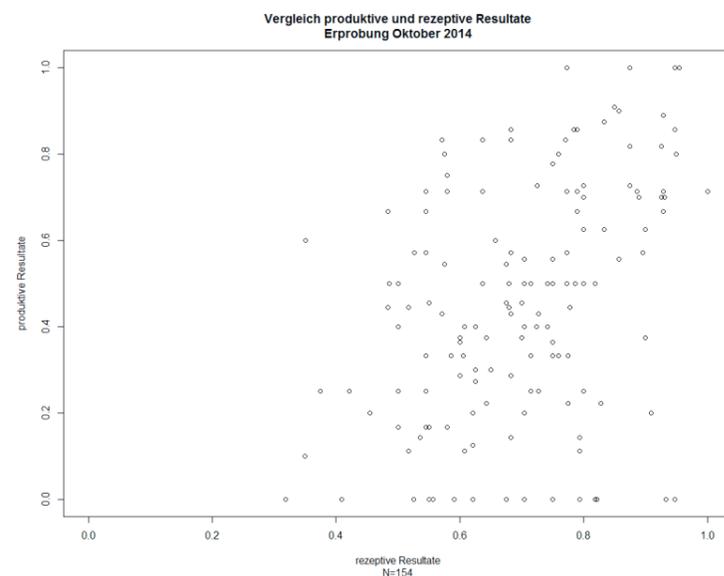


Abbildung III-4: Vergleich der Ergebnisse in den produktiven und rezeptiven Aufgaben pro Testteilnehmer/in. Die Resultate entsprechen dem Anteil korrekt gelöster Aufgaben in jedem der Fertigkeitsbereiche.

III Kompetenzen im Bereich des Sprechens

produktiven Aufgaben weisen hingegen eine breite Streuung über das gesamte Spektrum auf und zeigen eine leichte Tendenz zu höheren Werten bei Lernenden mit guten Resultaten bei den rezeptiven Aufgaben. Die Korrelation der auf wenige Informationen reduzierten Ergebnisse in den beiden Fertigkeiten liegt immerhin bei einem als moderat einzuschätzenden $r(152) = .43, p < .001$.

Es bliebe zu klären, inwiefern die Ergebnisse von Lernenden mit hohem Score in der einen, aber tiefem Score in der anderen Fertigkeit auf technische Probleme zurückzuführen sind. Zu den Zielen dieses Projektteils gehörte auch die Ausarbeitung eines Beurteilungs- bzw. Bewertungsinstrumentes zu den Aufgaben zur mündlichen Produktion.²⁸ Dieses Instrument sollte so gestaltet sein, dass es von den Lehrpersonen möglichst selbstständig und möglichst zeitsparend eingesetzt werden könnte. Deshalb wurden die Aufgaben so gefasst, dass die Produktionen dazu möglichst eindeutig als richtig oder falsch eingeschätzt werden können. Für die Bewertung von Äusserungen, bei denen eine differenziertere Beurteilung angewendet werden muss, wurde der Versuch unternommen, jeweils eine Reihe von Bewertungskriterien zu erstellen, bei denen lediglich die Entscheidung *zutreffend* oder *nicht zutreffend* gefällt werden muss. Mit einer derartig dichotomen Bewertung ist die Hoffnung verbunden, dass sie eindeutiger anzuwenden ist als z.B. Ratingskalen und damit auch weniger Zeit in Anspruch nimmt bzw. bei gleichem Zeitaufwand detailliertere Informationen liefert. Schliesslich eignen sich Ergebnisse zu mehreren eindeutig identifizierbaren Bewertungskriterien besser für eine statistische Auswertung und die anschliessende Interpretation als etwa eine einzelne, verschiedene Aspekte integrierende (holistische) Skala.

Diese Form der Bewertung wurde erst im Zusammenhang mit der letzten Aufgabenerprobung in Angriff genommen. Dabei wurde darauf verzichtet, die Bewertungsformulare für die Online-Auslieferung und die Nutzung durch externe Rater bereitzustellen. Stattdessen wurde beispielhaft

eine Bewertungsmaske mit diagnostischer Ausrichtung entworfen und in die Testsoftware integriert. Abbildung III-5 zeigt, wie eine solche Maske zum Gebrauch durch Lehrpersonen aussehen könnte. Die darin enthaltenen Bewertungskriterien wurden auf der Grundlage einer Analyse tatsächlich erhobener Lernerleistungen ermittelt. Die Maske umfasst fünf Elemente, welche in dieser Form zur Kodierung der meisten „einfachen“²⁹ produktiven Aufgaben geeignet sein dürften:

1. ein Item für ein holistisches Urteil über die Erfüllung der Kommunikationsaufgabe;
2. Items zu relevanten Aussprachephänomenen, bei denen einerseits angegeben werden kann, ob sie überhaupt vorkamen, und andererseits, ob sie zielsprachlich bzw. „verständlich“³⁰ produziert wurden (je nach gewählter Konvention), sowie weiter, wie flüssig artikuliert wurde;
3. freie Texteingabefelder für Notizen oder Transkriptionen (wobei Letzteres wohl vor allem im Rahmen eines Forschungsprojektes interessant ist, vielleicht aber auch im Zusammenhang mit formativem Feedback);
4. ein globales Urteil zu Vorhandensein, Vollständigkeit und Korrektheit der einzelnen inhaltlichen Elemente (unabhängig von der konkreten Wortwahl);
5. detaillierte Angaben zur sprachlichen Korrektheit der zu erwartenden grammatischen Elemente. Die verschiedenen Beurteilungskriterien wurden so gewählt, dass sie einerseits schnell und eindeutig bear-

29 | Gemeint sind jene Aufgaben, bei denen die Antwort inhaltlich und sprachlich gut vorhersehbar ist und aus relativ wenigen Worten besteht.

30 | Die konkrete Auswertung dieser Ausspracheitems müsste zunächst erkundet werden. Angedacht ist vor allem, dass das diagnostische Beurteilungssystem aufgrund dieser Information bei jedem einzelnen Lernenden ausgeben kann, welche Aussprachephänomene geübt werden sollten.

28 | Die Entwicklung automatischer Spracherkennung lag weit ausserhalb der technischen und personellen Möglichkeiten des Projektes.

III Kompetenzen im Bereich des Sprechens

beitet werden können und andererseits eine gewisse Bandbreite unterschiedlicher, relevanter Informationen über die Antworten der Lernenden zur Verfügung stellen. Dadurch sollte sich ein differenziertes Bild der Sprechkompetenz herauskristallisieren, sobald die Schülerinnen und Schüler eine Reihe solcher Aufgaben gelöst haben. Es ist allerdings zu betonen, dass diese Hoffnung erst dann sinnvoll überprüft werden kann, wenn diese und ähnliche Aufgaben über einen längeren Zeitraum hinweg von verschiedenen Lernenden bearbeitet werden. Aufgrund der Erprobungen im Projekt kann lediglich festgestellt werden, dass die Kriterien generell anwendbar sind. Ob sie tatsächlich brauchbare diagnostische Informationen liefern, d.h. relevante Entwicklungen erfassen, in die auch fördernd eingegriffen werden kann, müsste aber näher erforscht werden.

3. Schluss und Ausblick

Im Teil „Kompetenzen im Bereich des Sprechens“ des Projektes „Erfassung von Entwicklungsprofilen“ wurden verschiedene Ansätze zur Entwicklung von formativ-diagnostisch einsetzbaren Aufgaben verfolgt, welche geeignet erscheinen, die Entwicklung auch sehr geringer mündlicher produktiver Sprachkompetenzen sichtbar zu machen. Dabei wurden letztlich vor allem Aufgaben zum funktionalen und thematischen Wortschatz sowie solche zu zwei klar eingrenzenden kommunikativen Handlungsbereichen (Sprache im Unterricht, Austausch persönlicher Informationen) entwickelt und erprobt. Die Gründe für diese Beschränkung hängen v.a. mit den noch geringen Sprachkenntnissen der Primarschülerinnen und -schüler zusammen. Sie stehen ganz am Anfang ihres Sprachlernprozesses und können daher in der Fremdsprache nur wenige, sehr spezifische kommunikative Handlungen ausführen.

Scoringtask:
CL_2prod_jeuDes_prendsCarte

1 + 2

Erwartete Antwort:
Jette le dé et prends une carte!

0:00 0:50

Das kommunikative Ziel wurde ... **1**

... vollständig erreicht. ?

... teilweise erreicht. ?

... nicht erreicht. ?

Transkription
(freilassen bei leerer Aufnahme) **3**

Jette le dé ...

vollständig vorhanden teilweise vorhanden **4** falsch oder unverständlich dieses Element fehlt

"jeter" oder Alternative "dé" oder Alternative

Alles Zutreffende anklicken:

Verbform korrekt Artikel korrekt **5**

... et ... **4**

Konjunktion "et" oder Alternative falsches Wort dieses Element fehlt

... prends une carte. **4**

vollständig vorhanden teilweise vorhanden falsch oder unverständlich dieses Element fehlt

"prendre" oder Alternative "carte" oder Alternative

Alles Zutreffende anklicken:

Verbform korrekt Artikel korrekt **5**

Aussprache

j/ch da ok u-Laute da ok Wortenden da ok

e-Laute da ok Nasale da ok flüssig stockend **2**

Bemerkungen (hier eintragen) **3**

Abbildung III-5: Entwurf einer Bewertungsmaske zu einer produktiven Aufgabe. Die roten Zahlen beziehen sich auf die Erklärungen im Text (Aufzählung).

III Kompetenzen im Bereich des Sprechens

Die im Projekt gewonnenen Erkenntnisse betreffen in erster Linie die Bandbreite und die Ausgestaltung der Aufgaben, die für die Erfassung mündlicher Sprachkompetenzen junger Schülerinnen und Schüler geeignet sind. Damit existiert nun ein Fundament, auf dem der Aufbau einer einsatzbereiten Datenbank von formativ-diagnostischen Aufgaben vorangetrieben werden könnte. Zu diesem Zweck müsste der Bestand an Aufgaben erheblich erweitert und noch konkreter an die Lerninhalte der verschiedenen Kontexte und Schuljahre angepasst werden. Parallel dazu müssten die Bewertungsinstrumente zu den mündlichen Produktionen der Lernenden zuhanden der Lehrpersonen weiterentwickelt und im Schulalltag pilotiert werden. Die bei einer breiteren empirischen Validierung der Instrumente gewonnenen Ergebnisse zu Entwicklungsstand und Entwicklungswegen der Schülerinnen und Schüler in ausgewählten Bereichen sollten zur Skalenbildung und -beschreibung genutzt werden. Skalen bilden, zusammen mit den Ergebnissen bei konkreten Aufgaben, die Grundlage für lernbezogene Feedbacks und Feedforwards. Aus der Sicht der formativ-diagnostischen Beurteilung wäre es zudem interessant und wichtig, den Effekt von Lehr- und Lernentscheidungen, die auf diagnostischen Ergebnissen beruhen, zu untersuchen.

Insbesondere für die Nutzung des formativ-diagnostischen Instrumentariums auf der Sekundarstufe I müssten weitere Bereiche der Sprechkompetenz erschlossen und entsprechende Aufgabenformate und -inhalte entwickelt werden. Die im Projekt „Erfassung von Entwicklungsprofilen“ gemachten Erfahrungen im Bereich der computerbasierten Beurteilung können auch hierfür einen Rahmen bieten.

Schliesslich bleibt zu erwähnen, dass auch die Akzeptanz und der praktische Bedarf für ein serverbasiertes Instrument zur begleitenden, formativ-diagnostischen Beurteilung im Bereich des Sprechens noch näher abgeklärt werden muss (vgl. dazu Teilbericht VI zu Befunden aus dem Tessin). Insbesondere wäre abzuklären, inwieweit die Lehrpersonen bereit sind, die mündlichen Leistungen ihrer Schülerinnen und Schüler mithilfe einer Computerplattform manuell zu beurteilen. Auch die Ausstattung der Schulen (Internetanbindung,

Leistungsfähigkeit der Computer, Headsets usw.) spielt in diesem Zusammenhang eine Rolle. Auf der technischen Seite bleibt darüber hinaus die automatische Spracherkennung ein wichtiger Faktor: Wenn es gelingt, diese im Kontext des Fremdsprachenlernens, insbesondere von Kindern und Jugendlichen, so weit zu entwickeln, dass sie allgemein einsetzbar wird, werden sich neue Möglichkeiten ergeben, um Kompetenzen zu diagnostizieren, die weniger personellen Aufwand erfordern. Es ist damit zu rechnen, dass dann der Durchbruch für computerbasierte, lernbegleitende Beurteilungssysteme auch im Bereich der produktiven Kompetenzen bedeutend wahrscheinlicher würde als er aktuell noch ist.

III Kompetenzen im Bereich des Sprechens

4. Referenzen

- ACTFL (Hrsg.) (2012a). *Performance descriptors for language learners*. Abgerufen von <https://www.actfl.org/sites/default/files/pdfs/ACTFLPerformance-Descriptors.pdf> [09.03.16]
- ACTFL (Hrsg.) (2012b). *Proficiency Guidelines 2012*. ACTFL. Abgerufen von http://www.actfl.org/sites/default/files/pdfs/public/ACTFLProficiencyGuidelines2012_FINAL.pdf [09.03.16]
- Alderson, J. C. (2007). The CEFR and the need for more research. *The Modern Language Journal*, 91(4), 659-663.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice: designing and developing useful language tests*. Oxford; New York: Oxford University Press.
- Bygate, M. (2009). Teaching and testing speaking. In: M. Long & C. J. Doughty (Hrsg.), *The handbook of language teaching*. Chichester etc.: Wiley-Blackwell, 412-40.
- Ek, J. A. van, Alexander, L. G. & Council of Europe (1980). *Threshold level English in a European unit/credit system for modern language learning by adults*. Oxford; New York: Pergamon Press (published for and on behalf of the Council of Europe).
- Europarat (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. (J. Quetz & G. Schneider, Übers.). Berlin: Langenscheidt.
- Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91(4), 663-667.
- Johnson, K. (1996). *Language teaching and skill learning* (Bd. 5). Oxford; Cambridge: Blackwell.
- McKay, P. (2006). *Assessing young language learners*. Cambridge: Cambridge University Press.
- Microsoft (2010). *Microsoft PowerPoint* (Version 2010) [Microsoft Windows 7]. Redmond: Microsoft Corporation.
- Passepartout (Hrsg.) (2015). *Lehrplan Französisch und Englisch*. Abgerufen von <http://www.passepartout-sprachen.ch/services/downloads/download/533/get> [09.03.16]
- The Audacity Team (2013). *Audacity* (Version 2.0.3) [Microsoft Windows 7].
- Wilson, M. (2005). *Constructing measures: an item response modeling approach*. Mahwah, N.J.: Lawrence Erlbaum Associates.

IV Langage de classe

Evelyne Berger, Daniela Kappler, Katharina Karges

1. Introduction	90
2. Objectifs	91
3. Méthodes	91
3.1. Analyse des moyens d'enseignement	91
3.2. Sondage auprès des enseignant-e-s	94
4. Résultats du sondage	96
4.1. Participation au sondage	96
4.2. Analyse des réponses	96
5. Implications pratiques	98
5.1. Brochure d'objectifs d'apprentissage	98
5.2. Conception des tâches	99
6. Conclusion	100
7. Références	100

IV Langage de classe

1. Introduction

Le langage de classe est une notion largement utilisée dans le domaine de la didactique des langues, ainsi qu'en attestent les nombreux ouvrages et sites web, destinés aux enseignants, et offrant des listes d'expressions habituelles pour la vie en classe. Cette notion prend ses racines dans la recherche sur les pratiques communicatives et langagières propres au contexte de la classe (Wilkinson & Silliman, 2000). Le langage de classe se réfère à l'utilisation de la langue cible, tant par l'enseignant que par les élèves, pour la gestion des activités et de la vie en classe. Ainsi, il est par exemple utilisé pour donner des instructions sur un exercice, organiser l'agenda, mais aussi pour effectuer des rituels sociaux (salutations, prises de congés) ou encore pour réguler la participation. Une étude empirique, menée au sein d'une école primaire en Suisse romande en 2009-2010 dans trois classes de langue étrangère, a permis de recueillir, sur la base d'enregistrements audio des échanges en classe, un large éventail d'expressions clés ayant trait au langage de classe (Gauthier & Egli Cuenat, 2010).

La langue étrangère n'est alors pas seulement une langue « cible », ou objet d'enseignement, servant à exprimer des contenus et produire des « formulations grammaticalement correctes » ; elle est utilisée dans un but interpersonnel de réalisation d'actions avec autrui. Ainsi, le langage de classe est une forme de communication authentique connue des élèves, c'est-à-dire qu'il est utilisé en réponse à des tâches réelles et pertinentes pour ce contexte (cf. « authenticité intrinsèque », van Lier, 1988). Il constitue, par là même, un vecteur central d'apprentissage de la langue cible. En effet, l'élève aura l'occasion de développer ses compétences orales en L2, à travers l'usage de la langue cible pour accomplir des actions « authentiques » au sein de la classe. Le langage de classe peut aussi représenter une source de motivation à apprendre la L2, puisque la langue est utilisée quotidiennement à des fins pratiques.

L'usage du langage de classe est d'autant plus significatif dans le contexte de l'école pri-

maire, qu'à ce stade, les élèves n'ont que peu d'occasions effectives de pratiquer la L2 en-dehors de la classe (Domínguez, 2000). Les résultats de notre courte enquête sur la biographie langagière des élèves ayant participé aux phases pilotes de 2014 confirment ce constat : la majorité des enfants interrogés, entre la 4^{ème} et la 6^{ème} année de l'école primaire, affirment ne pas utiliser le français en-dehors de l'école, sauf pour effectuer leurs devoirs de français avec leurs parents. Ainsi, Domínguez (2000) souligne la pertinence d'évaluer le langage de classe chez les apprenants précoces d'une langue étrangère :

Young learners of foreign languages do not use the [foreign language] at all outside the language class, thus, it would be more useful in this case to base the test tasks exclusively upon language use in the classroom [...] [and] it would be better for young FL learners, who speak the target language only in the classroom to be evaluated against domains that are usual to the FL classroom. (Domínguez, 2000, p. 11)

Le langage de classe est d'ailleurs largement encouragé dans les milieux éducatifs, ainsi que le prévoient les plans d'études des deux régions concernées par le projet (DECS-DS, 2015 ; Passepartout, 2015) et les moyens d'enseignement actuellement utilisés dans ces contextes (Bertschy u.a., 2011ff. ; Samson, 2010). Dans *Mille feuilles* par exemple, le développement du langage de classe constitue un objectif pédagogique sur lequel est mis un accent particulier dès la première année d'apprentissage, avec des exercices visant explicitement à acquérir un vocabulaire utile pour la communication en L2 entre élèves. Il n'est cependant que très peu évalué en tant qu'objet d'apprentissage.

Dans la présente étude, nous prenons le langage de classe comme domaine exemplaire pour l'évaluation de la compétence communicative orale en français L2 (cf. la partie III « Kompetenzen im Bereich des Sprechens »). Cette démarche concorde avec la recommandation formulée par Domínguez (2000) d'évaluer des domaines d'usage

IV Langage de classe

de la L2 correspondant aux situations de communication en classe. Prendre le langage de classe permettra en outre de réfléchir aux contraintes et aux défis qu'impliquent l'évaluation, et notamment l'évaluation assistée par ordinateur, de formes de communication authentique et spontanée. Il permettra aussi un effet de *washback* (effet de reflux) sur les pratiques d'enseignement, en soulignant la pertinence du langage de classe.

2. Objectifs

Nous prenons le langage de classe comme un domaine exemplaire dans lequel peut être évalué le développement de la compétence communicative des jeunes apprenants, le langage de classe étant un des rares contextes d'usage authentique de la L2 à des fins communicatives à ce stade de l'apprentissage des langues étrangères.

Sur la base des indications données dans diverses ressources didactiques consacrées à la communication en classe (notamment les divers documents reliés à Egli Cuenat, 2010, ainsi que les manuels concernés), nous concevons le langage de classe comme se composant d'expressions conversationnelles habituelles, issues de routines quotidiennes et utilisées tant dans les échanges entre enseignant-e et élève(s) qu'entre élèves, et servant à la gestion de la vie en classe (régulation de la participation et des comportements telles que les réprimandes ou les instructions, maintien du lien social tels que les salutations ou les encouragements, gestion des imprévus telle qu'une demande d'autorisation pour aller aux toilettes), ainsi qu'à l'accomplissement du travail (consignes pour l'exécution des activités pédagogiques, leur encadrement et leur évaluation, planification du travail et de la leçon, etc.).

L'objectif de la présente recherche est de développer une base linguistique, empiriquement fondée, pour un instrument d'évaluation des connaissances des élèves en langage de classe. La démarche va consister à :

1. établir un corpus d'expressions propres au langage de classe et proposer une collection d'expressions constitutives du répertoire langagier produit oralement aux différents stades de l'apprentissage de la L2 ;
2. établir un catalogue d'objectifs d'apprentissage pouvant servir au développement de tâches dans le cadre d'une évaluation formative.

3. Méthodes

Un premier volet du sous-projet concernant le langage de classe consiste à établir une collection d'expressions langagières clés dans le domaine du langage de classe. Cette démarche se divise en deux étapes successives : l'analyse des moyens d'enseignement et un sondage auprès d'un certain nombre d'enseignants.

3.1. Analyse des moyens d'enseignement

Dans un premier temps, les moyens d'enseignement ont fait l'objet d'une analyse qui vise à identifier, de manière exhaustive, les expressions langagières en rapport avec le langage de classe contenues dans le matériel didactique, mais utilisées (également) à l'oral. Le matériel didactique constitue un point de départ fiable et pertinent pour esquisser le potentiel répertoire langagier commun à tous les élèves à ce stade de l'apprentissage. En effet, il est la source principale d'input linguistique en français pour la majorité des élèves du primaire, qui n'ont pas accès systématiquement au français en-dehors de l'école. Même si le repérage de ces expressions langagières se base sur l'analyse de supports avant tout écrits, les expressions retenues circulent également dans l'environnement de la classe de manière orale, dans la mesure où l'enseignant-e et les élèves lisent à haute voix, discutent et font référence

IV Langage de classe

oralement aux contenus du matériel didactique (p.ex. consignes ou éléments de vocabulaire comme aide à un travail de groupe).

Ainsi, la totalité des expressions du langage de classe ont été répertoriées par année scolaire, pour les deux méthodes *Mille feuilles* et *Alex et Zoé*, sur la base du livre de l'enseignant-e, du livre et du cahier d'activités de l'élève et des supports audio.

Nous avons retenu à la fois les expressions faisant explicitement l'objet d'un apprentissage en tant que langage de classe (p.ex. exercice visant l'apprentissage d'expressions typiques pour jouer à un jeu, cf. Figure IV-1) et les expressions présentes de manière implicite dans le texte accompagnant les activités (p.ex. introduction à l'activité, consignes, explications ou exemples, cf. Figure IV-2). Cette identification s'est parfois avérée problématique dans la mesure où il était difficile de définir exactement ce qui relevait du langage de classe à proprement parler et les expressions qui, sans être typiques du contexte de la classe, avaient des fonctions communicatives, comme des rituels d'ouverture de leçon ou des consignes autour d'activités pédagogiques (p.ex. «c'est étonnant...» ou des expressions présentées dans le cadre d'un parcours autour de la BD: «Drôle d'idée!» «Zut!» pouvant cependant être utilisées dans le contexte d'un jeu ou d'un travail en groupe).

L'analyse des moyens d'enseignements révèle une différence importante entre *Mille feuilles* et *Alex et Zoé*. Dans le premier, le développement du langage de classe est explicitement un objectif pédagogique, annoncé dans l'introduction au moyen d'enseignement. Ainsi, un certain nombre d'activités se centrent sur l'apprentissage et l'utilisation d'expressions pertinentes pour le langage de classe, notamment pour les activités de jeu entre élèves ou le travail en groupe. Il existe aussi des activités permettant de travailler spécifiquement sur la compréhension des consignes et le vocabulaire lié au contexte scolaire (les objets dans la salle de classe et le monde de l'école en général). *Mille feuilles* propose également des grilles d'évaluation pour les activités entre pairs qui contiennent l'usage des expressions pertinentes. De plus, la «Revue» de l'élève, comportant

tous les éléments de vocabulaire vus au cours de l'année, organise ceux-ci selon leur contexte d'usage, y inclus les activités de jeu, de travail en groupe ou les consignes.

Par contraste, *Alex et Zoé* ne met pas un accent explicite sur le langage de classe. Certes, le moyen d'enseignement propose quelques activités entre pairs pour lesquelles des expressions utiles sont données pour faciliter la communication en français (p.ex. «Tu veux jouer?», «Qu'est-ce que tu fais?»), ainsi que des routines d'ouverture de leçon et d'encadrement d'activités pédagogiques (p.ex. consignes), mais ces expressions sont en nombre relativement réduit dans l'ensemble du matériel didactique. De la même manière, le vocabulaire apparaissant dans les textes accompagnant les activités pédagogiques est relativement restreint et circulaire, à l'exemple des consignes, souvent très courtes.

La collection issue de ces deux méthodes a été complétée par des expressions langagières du langage de classe, absentes du matériel didactique en tant que tel mais fréquemment utilisées dans le contexte de la classe (p.ex. «Est-ce que je peux aller aux toilettes?»). Les expressions choisies figurent fréquemment sur les listes de vocabulaire portant sur le langage de classe, contenues dans les supports didactiques destinés aux enseignant-e-s de langue. L'étude empirique récente évoquée par Gauthier & Egli Cuenat (2010), présentant un corpus de langage de classe recueilli sur la base d'enregistrements des pratiques communicatives dans des classes francophones de 1^{ère} à 4^{ème} années dans le canton de Neuchâtel, corrobore la pertinence de ces expressions langagières dans le contexte de la communication en classe.

Cette collection d'expressions pour le langage de classe a donc été constituée à partir des moyens d'enseignement, utilisés dans les classes de primaires pour l'enseignement précoce du français en Suisse alémanique et au Tessin, et complétée au moyen de diverses ressources pédagogiques pour la formation des enseignants eux-mêmes. Les expressions recueillies ont ensuite été regroupées au sein d'une série de situations d'usage prototypiques du langage de classe, tels

IV Langage de classe

□ Lisez les règles du jeu et jouez.
Lest die Spielregeln und spielt.

RÈGLES DU JEU

Matériel
des cartes de l'activité F

Qui gagne?
Celui qui devine en premier cinq cartes du partenaire.
Gewonnen hat, wer zuerst fünf Karten des Partners / der Partnerin erraten hat.

Spick-Notizen
Notiere Sätze, die du verwenden kannst. Wurf beim Sprechen wenn nötig einen Blick auf die vorbereiteten Sätze und wende sie an.

J'aimerais un dessert. Tu as un dessert? Oui! Voici un dessert.
Non, désolé. Non, pas de chance.

□ Habt ihr weitere Ideen für Spiele mit den Karten aus dem fichier?

Découvrir un monde merveilleux
Du durchstreifst wunderbare Landschaften und entdeckst Tiere, die hier leben.

□ Travaillez à l'ordinateur. Cherchez des animaux dans les paysages.
□ Cliquez sur les animaux et écoutez.
□ Notez les animaux trouvés.

les paysages die Landschaften

Figure IV-2: Consignes données en français pour une activité portant sur la nature et les animaux (Mille feuilles 4.1.1, p. 20)

Figure IV-1: Activité de jeu avec travail explicite sur la langue de classe (Mille feuilles 3.1.2, p. 83)

Nous parlons français!

TRAVAILLER EN GROUPE

Fragen stellen

Tu travailles avec moi?
Arbeitest du mit mir?

Antworten geben

→ Oui, volontiers.
← Ja, gern.
Non, merci.
Nein, danke.

On travaille ensemble?
Arbeiten wir zusammen?

→ Oui, d'accord.
← Ja, einverstanden.
Non, désolé/e.
Nein, tut mir leid.

Qu'est-ce qu'il faut faire?
Was muss man tun?

Il faut ...
Man muss ...

Qu'est-ce que c'est en allemand ...?
Was heisst ... auf Deutsch?

→ C'est ...
← Das ist ...

Qu'est-ce que c'est en français ...?
Was heisst ... auf Französisch?

→ C'est ...
← Das ist ...

Qu'est-ce que tu fais?
Was machst du?

→ Je ...
← Ich ...

On parle français?
Sprechen wir französisch?

→ Oui, d'accord.
← Ja, einverstanden.

C'est juste?
Ist das richtig?

→ Oui, c'est juste.
← Ja, das ist richtig.
Non, c'est faux.
Nein, das ist falsch.

Tu as fini?
Bist du fertig?

→ Oui, ça y est.
← Ja, das wär's.
Non, pas encore.
Nein, noch nicht.



Figure IV-3: Vocabulaire pour le travail en groupe (Mille feuilles 4, Revue, p. 6)

IV Langage de classe

qu'ils apparaissent également souvent dans les supports didactiques présentant le langage de classe :

- rituels d'ouverture et de clôture de la leçon,
- activités en groupe (jeu, travail en groupe),
- résolution de problèmes de compréhension des tâches et des contenus, de langue, etc.,
- instructions et consignes,
- gestion de la discipline,
- gestion des déplacements dans l'espace et gestion du temps.

Dans un deuxième temps, la collection d'expressions langagières ainsi établie a été soumise au regard critique des acteurs du terrain. En demandant aux enseignant·e·s de langue d'estimer la difficulté d'utilisation de ces expressions langagières, l'enquête visait à identifier les expressions pouvant être réalistement incluses dans l'instrument d'évaluation en cours de développement et à définir à quelle étape du parcours d'apprentissage cela s'avérait pertinent.

3.2. Sondage auprès des enseignant·e·s

Après cette première validation par un petit nombre d'experts, nous avons soumis notre collection à une vérification de plus large envergure. Pour ce faire, nous avons réalisé un sondage auprès des enseignant·e·s de français langue étrangère à l'école primaire, dans trois cantons de la région Pässepartout en Suisse alémanique (BE, SO, BL) et au Tessin. Il a été demandé aux enseignant·e·s d'estimer, sur la base de leur expérience personnelle en classe, la capacité des élèves à utiliser et comprendre oralement le langage de classe, et ceci, pour les années scolaires allant de la 3^{ème} à la 6^{ème} primaire. Les résultats obtenus doivent par conséquent être compris comme présentant une *appréciation basée sur l'expérience* qu'ont les enseignant·e·s du degré de familiarité

de leurs élèves avec le langage de classe et donc une représentation indirecte de l'utilisation et la compréhension effective de ces expressions en classe.

Cette enquête a été réalisée parallèlement en Suisse alémanique et en Suisse italienne, moyennant deux versions linguistiques du même questionnaire (une en allemand et une en italien). Une participation suffisamment importante de plusieurs régions linguistiques aurait pu permettre d'identifier une différence dans l'acquisition du langage de classe en fonction de la région linguistique. Différence due aux moyens d'enseignement divergents et à la proximité ou la distance linguistique entre la langue première et la langue étrangère (allemand – français et italien – français). Cependant, dû au nombre limité de participant·e·s tessinois·e·s, seul un aperçu global des pratiques en matière de langage de classe a pu être obtenu pour ce contexte.

Afin d'atteindre un panel plus large d'enseignant·e·s du français et de faciliter le recueil de données, le sondage a été réalisé à l'aide d'un outil de création de questionnaires (LimeSurvey Project Team & Schmitz, 2014), hébergé sur un serveur de l'Institut de plurilinguisme. L'accès au questionnaire a été diffusé largement à travers différents réseaux de contacts dans les milieux éducatifs. L'identification personnelle des participant·e·s, pouvant permettre un éventuel contact ultérieur, était facultative et permettait une participation anonyme si celle-ci était souhaitée. Le sondage est resté actif en ligne pendant près de 4 mois (de mars à juillet 2014).

Deux questionnaires différents ont été élaborés, l'un destiné aux enseignant·e·s de classes de 3^{ème} et 4^{ème} année, l'autre destiné aux enseignant·e·s de classes de 5^{ème} et 6^{ème} année. Les deux versions du questionnaire comprennent des items spécifiques selon un possible ordre d'apprentissage du langage de classe (déterminé par les contenus des moyens d'enseignement et de la complexité grammaticale de l'expression), mais également 17 items identiques, dans le but d'observer une éventuelle progression à travers les quatre années scolaires.

IV Langage de classe

Seule une sélection d'expressions clés de la collection initiale a été incluse dans le questionnaire. Les deux questionnaires étaient cependant de taille à peu près égale : le questionnaire pour les 3^{ème} et 4^{ème} comptabilisait 75 items et le questionnaire pour les 5^{ème} et 6^{ème} en comptabilisait 74.

Afin de pouvoir identifier au mieux où se situe l'apprenant·e dans le parcours d'apprentissage (fonction diagnostique), le choix des items s'est porté d'une part, sur des expressions extrêmement routinières et fréquentes, dont le potentiel d'acquisition était très élevé (p.ex. « bonjour », « au revoir ») et d'autre part, sur des expressions moins fréquentes ou plus complexes, dont le potentiel d'acquisition était plus incertain. De manière générale, le genre n'a pas été différencié dans le choix des items (p.ex. « Tu es prêt/prête ? »). Par contre, dans le cas des expressions interrogatives, le format syntaxique a été différencié (l'inversion sujet-verbe, p.ex. « Avez-vous fait vos devoirs ? », le marqueur interrogatif p.ex. « Est-ce que vous avez fait vos devoirs ? », la phrase interrogative SVO, p.ex. « Vous avez fait vos devoirs ? »), de façon à identifier laquelle de ces expressions alternatives est plus à même d'être utilisée et maîtrisée par les élèves, à ce stade de l'apprentissage. La forme à la 2^{ème} personne du singulier et celle à la 2^{ème} personne du pluriel ont été différenciées (p.ex. « répète » et « répétez »), dans la mesure où ces expressions ne sont pas utilisées avec une fréquence équivalente, car elles relèvent de routines spécifiques à un contexte donné (p.ex. « Comment vas-tu ? » est plus souvent utilisé pour s'adresser à son interlocuteur dans un échange à deux que « Comment allez-vous ? » pour s'adresser à la classe toute entière).

Les items ont couvert l'ensemble des situations d'usage identifiées lors de l'établissement du corpus et étaient présentés sous cette forme dans le questionnaire ; autrement dit, les items ont été regroupés sous des labels contextualisant la situation dans laquelle ils étaient utilisés. Après chaque groupe d'items interrogés pour une même situation d'usage, le questionnaire prévoyait une case vide dans laquelle il était possible d'ajouter volontairement des expressions jugées pertinentes par les enseignant·e·s pour la situation d'usage,

mais qui ne figuraient pas dans le questionnaire, ceci afin d'élargir notre corpus.

L'objectif du questionnaire étant de recueillir une estimation des compétences productives et réceptives à l'oral en matière de langage de classe, il a été demandé, pour chaque item, de répondre à la question suivante : « Un élève moyen est-il capable, en ce moment, de produire et comprendre à l'oral l'expression suivante ? » Etant donné que les compétences des élèves au sein d'une même classe pouvaient être très variables, avec des élèves très forts et d'autres très faibles, et en fonction du stade d'avancement dans le programme scolaire annuel, la question était formulée avec précision, dans le but d'amener les participant·e·s à répondre à l'enquête de la manière la plus convergente et comparable possible. Les questions sur certains items étaient parfois accompagnées d'indications complémentaires sur la manière de les comprendre. Par exemple, pour l'item « Donne-moi ton crayon. », un commentaire indiquait que « crayon » peut être remplacé par tout autre nom d'objet que les élèves connaîtraient, dans la mesure où l'intérêt de l'expression concernait en fait le syntagme verbal et non pas l'objet.

Chaque item a été évalué sur une échelle rendant compte du degré de facilité/difficulté avec lequel l'item pouvait être compris ou produit par un élève moyen :

- oui, bien et facilement (c.a.d. peut l'utiliser spontanément et sans aide),
- oui, plutôt bien (c.a.d. la plupart du temps, mais pas toujours ou pas toujours juste),
- oui, avec peine (c.a.d. avec des répétitions, lentement, avec l'aide de notes ou une traduction),
- non (c.a.d. si l'expression n'est pas connue ou si elle n'est jamais utilisée en classe).

L'échelle était basée sur quatre niveaux de façon à éviter que les participant·e·s cochent par défaut la case du milieu en l'absence d'avis clair.

IV Langage de classe

4. Résultats du sondage

4.1. Participation au sondage

Au cours des quatre mois d'activité du questionnaire en ligne, nous avons recueilli 65 contributions pour la version allemande contre 9 pour la version italienne (cf. Tableau IV-1). Cette différence importante de participation entre les trois cantons alémaniques et le canton du Tessin s'explique par une diffusion différente de l'appel à participation à notre sondage. En effet, pour les cantons alémaniques, l'information ayant été diffusée par le biais des coordinations cantonales de l'instruction publique auprès des directions d'établissements scolaires et des centres de formation continue, l'appel a reçu un accueil positif auprès des enseignant-e-s. Au Tessin, même si les autorités cantonales ont offert leur soutien, la prise de contact s'est effectuée sur une base individuelle par le biais de l'équipe de recherche, moyennant, dès lors, des efforts importants pour convaincre les enseignant-e-s et susciter leur intérêt, ce qui a fortement limité le panel de participant-e-s.

La répartition des contributions en fonction de l'année scolaire se présente de la manière suivante :

Année scolaire	Sondage en allemand	Sondage en italien
3 ^{ème} année	17	5
4 ^{ème} année	21	1
5 ^{ème} année	23	3
6 ^{ème} année	4	0
Total	65	9

Tableau IV-1 : Nombre de participant-e-s au sondage

Alors que la participation d'enseignant-e-s de classes de 3^{ème}, 4^{ème} et 5^{ème} est d'importance relativement comparable (près d'un tiers des résultats pour chacune des trois années scolaires), les contributions pour les 6^{ème} année sont très peu nombreuses. Ceci s'explique par le fait que le

matériel didactique *Mille feuilles* pour les 6^{ème} année était en cours de pilotage au moment de ce sondage. Par conséquent, les enseignant-e-s ayant utilisé ce moyen d'enseignement, qui met au cœur de sa démarche le développement de compétences communicatives à travers le langage de classe, étaient encore peu nombreux.

4.2. Analyse des réponses

Les réponses données par les enseignant-e-s concernant les compétences de leurs élèves pour chaque item du questionnaire ont fait l'objet d'une analyse statistique simple et de valeur de preuve limitée, étant donné le nombre modeste de réponses. Les différentes réponses de l'échelle à quatre niveaux (cf. plus haut) ont reçu les points suivants :

- oui, bien et facilement : 3 pts,
- oui, plutôt bien : 2 pts,
- oui, avec peine : 1 pts,
- non : 0 pts.

La valeur numérique obtenue constitue ainsi la moyenne des réponses obtenues pour chaque item.

La valeur numérique indique donc un degré de facilité/difficulté dans l'utilisation de l'expression cible. Plus la valeur est haute (le maximum étant 3), plus l'item sera jugé comme étant bien

IV Langage de classe

maitrisé par les élèves. A l'inverse, plus la valeur est basse (le minimum étant 0), plus l'item sera jugé comme étant utilisé avec difficulté.

Les valeurs moyennes obtenues montrent ainsi les tendances suivantes :

- La valeur numérique d'une expression tend à croître d'une année scolaire à une autre (de la 3^{ème} à la 4^{ème}, de la 5^{ème} à la 6^{ème} et de la 3^{ème} à la 6^{ème} selon les items). Cela signifie que la capacité des élèves à utiliser ces items augmente au fil du temps.
- Pour la plupart des items, les résultats montrent une différence flagrante entre les valeurs moyennes au niveau réceptif et au niveau productif, par année scolaire (une exception constitue par exemple « Joyeux anniversaire » dont la valeur moyenne décroît entre la 3^{ème} et la 4^{ème} année tant au niveau réceptif que productif). Ainsi, les items sont, la plupart du temps, d'abord acquis au niveau réceptif avant de l'être au niveau productif.
- En 3^{ème} année, c'est-à-dire en début d'apprentissage, un nombre restreint d'items semblent être véritablement acquis, et cela même au niveau productif. Ces items relèvent de routines conversationnelles et pédagogiques extrêmement fréquentes :
 - Salutations : *bonjour, au revoir, ça va ?*
 - Consignes : *ouvre/ferme le livre/la porte, répète/répétez, regarde/regardez.*
 - Evaluation : *bravo, c'est juste, c'est faux.*
- Les items qui ont reçu les valeurs les plus basses sont des items qui ne figurent pas dans le répertoire de *Mille feuilles* (p.ex. « Retourne à la case départ » ou « Nicole, tu es en retard ») et dont l'insertion dans la collection d'expressions du langage de classe se justifie par une autre source d'information (c.a.d. provenant notamment d'autre matériel didactique).

- En ce qui concerne les groupes d'items, les expressions en lien avec la situation de jeu entre pairs ou de travail en groupe n'ont pas obtenu des valeurs très élevées, alors même que ces situations de communication sont explicitement un objet d'enseignement par de nombreuses activités pédagogiques prévues dans *Mille feuilles* qui implémentent ces contextes d'usage. Quant aux consignes/instructions, les enseignant-e-s en général considèrent qu'elles sont bien maîtrisées par les élèves.
- Puisque nous avons soumis à l'évaluation différents formats syntaxiques pour les énoncés interrogatifs, les résultats montrent une préférence pour le format SVO (p.ex. « Tu as compris ? »), plus typique du français parlé, par opposition au format interrogatif avec inversion sujet-verbe ou avec la structure en « est-ce que ». Ces deux derniers formats présentent en général des valeurs d'aisance plus basses que le premier format pour les différents items.
- Les formes verbales à la 2^{ème} personne du singulier ou à la 2^{ème} personne du pluriel qui existent pour un même item ne présentent en général pas d'écart important entre les valeurs de chacune des alternatives. Ceci n'est pas surprenant dans la mesure où il s'agit d'expressions se référant à des consignes ou instructions qu'un-e enseignant-e ou un-e élève peut adresser à un individu seul ou à un groupe – formes verbales dont la fréquence est comparable dans le contexte de la classe. Il serait possible qu'une différence soit observée quant à l'acquisition des mêmes formes verbales pour la 3^{ème} personne ou à l'infinitif.

Pour les résultats du questionnaire en italien, la valeur moyenne des réponses n'a pas été calculée, puisque le nombre de participant-e-s pour chaque

IV Langage de classe

année est très réduit. Par contre, les rares réponses obtenues montrent les résultats suivants, en rapport avec les réponses du questionnaire en allemand :

- Un écart semblable entre les compétences réceptives et productives, les premières étant généralement mieux évaluées que les deuxièmes.
- Une évaluation du degré d'aisance similaire, voire même légèrement plus favorable.
- On peut observer cependant quelques disparités dans les réponses, dans le domaine du jeu en particulier (p.ex. « J'ai gagné », « Mélange les cartes » sont mieux évalués pour le questionnaire en allemand, alors que « Retourne à la case départ » et « Attends ton tour » sont mieux évalués dans le questionnaire en italien) ; ces différences peuvent découler de la différence des moyens d'enseignement qui ne focalisent pas sur le même éventail de ressources langagières utiles aux activités entre élèves.

Il va de soi que ces résultats ne peuvent pas, en soi, être comparés aux résultats du sondage en allemand, puisqu'ils se basent sur un nombre très restreint de réponses ; mais ils peuvent néanmoins esquisser très grossièrement une tendance.

De nombreuses expressions ont également été signalées volontairement par les participants au sondage en allemand et en italien dans les emplacements prévus à cet effet. Une partie des propositions se recoupe avec des expressions interrogées ailleurs dans le questionnaire ou dans le questionnaire parallèle, mais ces contributions libres recensent aussi beaucoup de nouvelles propositions qui n'étaient pas présentes dans la collection initiale d'expressions du langage de classe.

5. Implications pratiques

5.1. Brochure d'objectifs d'apprentissage

Les informations récoltées sur l'usage d'items prédéfinis ainsi que par les propositions additionnelles des enseignant-e-s sont synthétisées dans un tableau récapitulatif des résultats du sondage. Ce tableau fait l'objet d'une annexe à ce document, publiée en ligne à l'adresse <https://bit.ly/2kLi6dq>. La valeur moyenne des réponses pour chaque item y est indiquée au moyen de trois catégories : « grande aisance d'utilisation réceptive et productive », « aisance émergente d'utilisation réceptive et productive », « difficulté d'utilisation réceptive et productive ». Ainsi, le tableau donne une vue d'ensemble approximative de l'utilisation probable des expressions par les élèves à travers les années. Par contre, nous n'avons pas fait de distinction entre compétence réceptive et productive, dans la mesure où les résultats plus détaillés (cf. ci-dessus) ont montré une tendance à une compétence réceptive dans l'utilisation de ces items avant la compétence productive. À ces résultats s'ajoutent également les contributions libres pour lesquelles la catégorie « acquis » a été signalée pour l'année scolaire indiquée par le/la participant-e. Pour chaque item, nous mentionnons également, dans la mesure du possible, à quel endroit l'item est introduit dans le moyen d'enseignement (*Mille feuilles* ou *Alex et Zoé*) ou s'il provient d'une autre source. Ainsi, ce tableau offre un aperçu du degré d'utilisation des expressions clés du langage de classe et un échelonnage en fonction de l'année scolaire. Il (dé)montre également l'ancrage de ces items dans les moyens d'enseignement.

A partir de ce document récapitulatif, il est possible d'établir des objectifs d'apprentissage et de concevoir des instruments d'évaluation du développement du langage de classe en repérant aisément à quel moment les items de la liste peuvent faire l'objet d'une évaluation à des fins diagnostiques.

IV Langage de classe

5.2. Conception des tâches

Du fait que le langage de classe constitue une forme de communication authentique, il s'avère pertinent de mesurer la production ou la compréhension de ces expressions en contexte. Par contexte est entendue une situation spécifique dans laquelle l'usage d'une expression donnée est fréquent (p.ex. « J'ai gagné » dans une situation de jeu ou de compétition sportive). Ainsi, les tâches développées au sein du projet, destinées à éprouver la fiabilité d'un instrument diagnostique, ont porté sur les situations d'usage suivantes : les instructions données par un élève à un autre dans le cadre d'un jeu ou d'un travail en groupe³¹ ainsi que les consignes et ordres donnés par l'enseignant-e à la classe ou à un élève.

Les tâches consistent à évaluer la maîtrise d'une expression toute entière, ancrée dans une situation d'usage spécifique, à la manière des tours de parole d'un court dialogue, au lieu d'éléments lexicaux isolés. Ceci dit, les tâches ont été conçues pour être réalisées individuellement sur ordinateur, en raison du contexte de ce projet, lequel vise à développer un instrument d'évaluation informatisé. Une mise en situation réelle pourrait poser un certain nombre de défis pour le groupe d'apprenants étudiés : il serait très laborieux et contraignant d'obtenir les productions ciblées et homogènes pour l'ensemble des élèves (p.ex. l'utilisation spontanée et rapide de ces expressions face à des camarades peut s'avérer difficile pour certains élèves ; ce qui aurait pour conséquence qu'ils ne parlent pas du tout en langue cible) et de les tester individuellement (p.ex. en situation d'interaction réelle, plusieurs élèves sont évalués simultanément). Par contre, un instrument informatisé permet d'obtenir facilement et rapidement des réponses, tout en assurant des conditions semblables pour tous les élèves, ce qui rend l'analyse de leurs réponses plus fiable.

A part certaines tâches qui demandent des expressions simples comme « J'ai gagné ! », une bonne partie des tâches relevant du langage de classe demande la production ou la réception de consignes. Ces tâches se composent de deux parties : une première partie où l'instruction est donnée et une deuxième partie où l'instruction est réalisée. Selon qu'il s'agisse d'une tâche productive ou réceptive, l'élève est amené à produire ou identifier l'expression langagière correspondant à l'instruction donnée initialement. Afin de permettre une meilleure contextualisation des items, la tâche se base sur une série d'images, à la façon d'une BD, ou d'une image simple qui illustrent l'action devant être réalisée par la consigne. Dans le cas d'une tâche réceptive (Figure IV-4), l'élève écoute trois items audio reproduisant une instruction (soit avec une voix d'adulte, soit avec une voix d'enfant selon la situation ciblée) et doit choisir l'item qui correspond à l'image. Dans le cas d'une tâche productive (Figure IV-5), l'élève doit se mettre lui-même à la place du personnage et formuler l'instruction qui correspond à l'action à réaliser.



Figure IV-4 : Tâche CL_mpc3a1i_rangerLivre. Instruction : « Que dit l'enseignant ? Choisis la bonne phrase. »

31 | Seulement les expressions identifiées comme bien acquises dans le questionnaire ont été choisies pour le développement des tâches.

IV Langage de classe



Figure IV-5: Tâche CL_prod_marquer. Instruction (en allemand): «Ta collègue doit marquer un mot dans le texte. Qu'est-ce que tu dis?»

6. Conclusion

Cette partie du projet « Recensement de profils de développement » investiguait spécifiquement le domaine du langage de classe. Une analyse de la littérature et des manuels en utilisation a abouti à une liste d'expressions utiles pour la communication quotidienne dans la salle de classe. Cette liste a été soumise à 74 enseignants et enseignantes en Suisse alémanique et au Tessin, à qui nous avons demandé, dans un questionnaire en ligne, de juger si et à quel point les élèves dans leurs cours de français maîtrisaient les différentes expressions de manière réceptive et/ou productive. Ces jugements issus de la pratique ont donné des indications concernant la difficulté des expressions pour des élèves de différentes années scolaires. Le questionnaire a aussi permis d'élargir la collection des expressions pertinentes grâce aux propositions libres des enseignant·e·s. Une version commentée de toutes les expressions est publiée sous la forme d'un tableau d'objectifs d'apprentissage qui contiendra les expressions attendues par les praticien·ne·s pour les différentes années scolaires.

Dans la partie III « Kompetenzen im Bereich des Sprechens » du projet, la liste des expressions a directement servi à la création de tâches. Les

différents essais de ces tâches en classe ont pourtant démontré que les jugements des enseignant·e·s ne sont qu'une source d'information entre autres : une partie des expressions, notamment celles liées au jeu et aux travaux de groupes, n'ont été maîtrisées que par une minorité des élèves – contrairement aux déclarations des enseignant·e·s et à la structure des manuels. L'enquête auprès des personnes concernées et l'analyse de la littérature ne nous donnent donc pas toutes les informations nécessaires sur les pratiques et les compétences effectives des élèves dans le domaine du langage de classe.

7. Références

- Bertschy, I., Grossenbacher, B. & Sauer, E. (2011ff.). *Mille feuilles. Material für Lernende, Lehrmaterial.* (Vol. 1-4). Bern: Schulverlag Plus.
- DECS-DS (Éd.) (2015). Piano di studio della scuola dell'obbligo. Consulté à l'adresse <http://www.pianodistudio.ch/> [09.03.16]
- Domínguez, R. (2000). Oral tests for children learning a foreign language. *Melbourne Papers in Language Testing*, 9(2), 1-17.
- Egli Cuenat, M. (2010). *Konzept Berufsspezifische Sprachkurse C1* für Lehrpersonen der Primarstufe (Französisch und Englisch)*. Passepartout. Consulté à l'adresse <http://www.passepartout-sprachen.ch/services/downloads/download/1942/get> [11.03.16]
- Gauthier, S. & Egli Cuenat, M. (2010). *Gérer la communication en classe : interagir et soutenir, activer les élèves*. Passepartout. Consulté à l'adresse https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwji2sLVkd3kAhVmlosKHZV6AmsQFjAAeQIAC&url=https%3A%2F%2Fwww.fremdsprachenunterricht.ch%2Fservices%2Fdownloads%2Fdownload%2F1950%2Fget&usg=AOvVaw08_4YGBWwGzkKPPWajSLM3 [19.09.2019]

IV Langage de classe

fremdsprachenunterricht.ch%2Fservices%2Fdownloads%2Fdownload%2F1950%2Fget&usg=AOvVaw08_4YGBWwGzkKPPWajSLM3 [19.09.2019]

- LimeSurvey Project Team & Schmitz, C. (2014). *LimeSurvey: An open source survey tool, v. 2.05* (Unix). Hamburg: LimeSurvey Project. Consulté à l'adresse <http://limesurvey.org> [11.03.16]
- Passepartout (Éd.) (2015). *Lehrplan Französisch und Englisch*. Consulté à l'adresse <http://www.passepartout-sprachen.ch/services/downloads/download/533/get> [08.02.16]
- Samson, C. (2010). *Alex et Zoé et compagnie. Méthode de Français.* (Nouvelle édition, Vol. 1-3). Paris: CLE International.
- van Lier, L. (1988). *The classroom and the language learner: ethnography and second-language classroom research*. London & New York: Longman.
- Wilkinson, L. C. & Silliman, E. R. (2000). Classroom language and literacy learning. In: M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Éd.), *Handbook of reading research* (Vol. 3). New York & London: Routledge, 337-60.

V Sprachlernstrategische Kompetenzen

Thomas Roderer, Anna Kull, Peter Lenz

1. Einleitung und Kontext	104
2. Ausgangslage	105
2.1. Sprachlernstrategien und Sprachkompetenz	105
2.2. Vermittlung von Sprachlernstrategien	106
2.3. Entwicklung im Bereich der strategischen Kompetenzen	106
3. Ziele des Teilprojekts	107
4. Entwicklungsarbeiten	107
4.1. Arbeitsdefinition, Eingrenzung	107
4.2. Instrumente	110
4.2.1. Drei Typen von Instrumenten	110
4.2.2. Zuordnung der Instrumente zu einer hierarchischen Prozessdimension	114
4.3. Erprobung der Instrumente	116
4.4. Nutzung der Instrumente im Unterrichtskontext	118
4.4.1. Umfang des Einsatzes von Sprachlernstrategien	118
4.4.2. Wissen über Sprachlernstrategien und deren Eignung für bestimmte Lernaufgaben	120
4.4.3. Metakognitives Wissen über die eigene sprachliche Leistungsfähigkeit	120
5. Diskussion und Schluss	122
6. Referenzen	124

V Sprachlernstrategische Kompetenzen

1. Einleitung und Kontext

Das Projekt „Erfassung von Entwicklungsprofilen“ hatte insgesamt zum Ziel, für den Fremdsprachenunterricht in der Primarschule (3. bis 6. Klasse bzw. 5. bis 8. Schuljahr gemäss HarmoS-Zählung) ein Instrumentarium zur Beschreibung und Erfassung von Kompetenzen in Bereichen des Fremdsprachenlernens zu entwickeln, die oft als schwer zu beurteilen gelten. Die sprachlernstrategischen Kompetenzen stellten einen der hauptsächlichsten Arbeitsbereiche dar.

Sprachlernstrategische Kompetenzen zählen zu den sog. transversalen Kompetenzen, die im Rahmen einer umfassenden Mehrsprachigkeitsdidaktik entwickelt werden sollen (Beacco u.a., 2010; Neuner, 2005). Zur Entwicklung transversaler Kompetenzen leisten im Prinzip alle Fächer einen mehr oder weniger spezifischen Beitrag. Ebenso profitieren sie potenziell alle von den Synergien, die sich aus transversalen Kompetenzen ergeben. Durch die gezielte Förderung von transversalen Kompetenzen im Bereich der Mehrsprachigkeit will die Schule auch eine Basis für ein lebenslanges, selbstbestimmtes (autonomes) (Weiter-)Lernen von Sprachen schaffen (z.B. Passepartout, 2015).

In der Schweiz, wo gemäss der Sprachenstrategie der Erziehungsdirektorenkonferenz (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK), 2004) bereits auf der Primarstufe neben der Schulsprache eine weitere Landessprache sowie Englisch unterrichtet werden (sollen), wird in den relevanten strategischen Dokumenten wiederholt auf die Wichtigkeit von Sprachlernstrategien hingewiesen: „Frühes Sprachenlernen [...] schafft günstige Voraussetzungen für das Erlernen weiterer Sprachen und fördert das Entwickeln von Strategien für das Sprachenlernen“ (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK), 2004). In den „Grundkompetenzen für die Fremdsprachen“ (Dokument zu den schulischen Minimalstandards) wird die Bedeutung von strategischen und interkulturellen Kompetenzen wie folgt erläutert:

Dies [d.h. die eher tiefen Sprachhandlungsziele für die erste Fremdsprache auf der Primarstufe] ermöglicht es, dem Aufbau der Sprachmittlung und der überfachlichen Kompetenzen (interkulturelle und methodische Kompetenzen) in der ersten Fremdsprache Zeit einzuräumen. Diese Kompetenzen spielen eine zentrale Rolle, damit die Schülerinnen und Schüler beim Lernen der zweiten Fremdsprache von den ersten Lernerfahrungen profitieren (z.B. durch Transfers von Strategie- und Sprachlernwissen). Im Rahmen der Umsetzung der Sprachenstrategie fliesst die Förderung dieser überfachlichen Kompetenzen in die Lehrmittel und die didaktische Ausbildung der Lehrpersonen ein, wodurch die Schülerinnen und Schüler beim Lernen beider Fremdsprachen (sowie der Schulsprache) angeleitet werden, diese Transfers zwischen den Sprachen zu machen, was die Entwicklung entsprechender überfachlicher Kompetenzen fördert. (Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK), 2011)

Im Deutschschweizer Lehrplan 21 (LP21) (D-EDK, 2014) haben Sprachlernstrategien v.a. den Charakter von Kompensationsstrategien im Zusammenhang mit der Sprachverwendung bzw. deren Planung und Vorbereitung (Kommunikationsstrategien: Hör-, Lese-, Sprech- und Schreibstrategien). Zusätzlich werden spezifisch Strategien zum Wortschatzerwerb und zur Wortbildung genannt. Zu den strategischen Kompetenzen wird auch ein individualisiertes, erfahrungsbasiertes Wissen über die Eignung von Strategien für bestimmte Zwecke gezählt (D-EDK, 2014, Kapitel FS1F.5.B.2).

Im Passepartout-Lehrplan (spezifischer Fremdsprachenlehrplan der sechs Kantone an der deutsch-französischen Sprachgrenze) bilden die lernstrategischen Kompetenzen sogar einen der drei hauptsächlichsten Kompetenzbereiche und werden im Lehrplandokument recht ausführlich charakterisiert. Dabei wird zwischen dem Aufbau von Strategien in verschiedenen Bereichen des Unterrichts unterschieden: kommunikative Sprachver-

V Sprachlernstrategische Kompetenzen

wendung (Fertigkeiten), Wortschatzlernen, Entwicklung von (positiver) Sprachemotion (z.B. Motivierung) und Sprachlernreflexion (Passepartout, 2015).

In den neueren Fremdsprachen-Lehrwerken der Deutschschweiz werden sprachlernstrategische Kompetenzen zwar kaum systematisch entwickelt, aber immer wieder thematisiert. Ein wichtiger Faktor für das tatsächliche Entstehen strategischer Kompetenzen dürfte ein „erwerbsfreundlicher“ Umgang damit im Unterricht sein, d.h. Einsatz von genügend Unterrichtszeit, Verbalisierung und Reflexion strategischen Handelns usw. Für die Kompetenzausprägung bei den einzelnen Schülerinnen und Schülern dürften zudem individuelle Faktoren wesentlich sein.

Neben der Betonung der strategischen Kompetenzen in aktuellen Konzepten und Richtlinien für den Fremdsprachenunterricht war auch das Vorhandensein von Fachliteratur (Studien, Inventare, Erhebungsinstrumente usw.) mitverantwortlich dafür, dass im vorliegenden Projekt ein Schwerpunkt auf die Entwicklung von Instrumenten zur Erfassung von sprachlernstrategischen Kompetenzen gelegt wurde.

2. Ausgangslage

2.1. Sprachlernstrategien und Sprachkompetenz

Sprachlernstrategien lassen sich als Bindeglied zwischen Lernmotivation und Lernerfolg verstehen (Credé & Phillips, 2011). Die Anfänge der Forschung zur Verknüpfung von Strategiegebrauch und Fremdsprachenlernen werden mit Naimans Studie von 1978 zu guten Sprachlernenden (Naiman u.a., 1996) in Verbindung gebracht, aus denen mehrere Strategiekategorien abgeleitet wurden. Diese Strategien führten zu weiterer Forschung darüber, wie Strategien im Bereich des Fremdsprachenlernens mit dem Lernerfolg zusammenhängen. Dabei wurde neben der Art der Stra-

tegien auch die *Häufigkeit* des Strategiegebrauchs betrachtet. Eine einfache Beziehung konnte jedoch nicht hergestellt werden. Geschuldet ist dieser Umstand zum einen der Komplexität des Gegenstandes: Der Gebrauch von Lernstrategien wird von mehreren, schwer kontrollierbaren Faktoren beeinflusst, die ihrerseits starken Einfluss auf den Lernerfolg haben können (z.B. die Persönlichkeit oder Motivation der Lernenden). Zum anderen zeigt sich immer wieder, dass nicht nur erfolgreiche, sondern auch weniger erfolgreiche Lernende Strategien verwenden. Die Frage nach „generell guten“ Strategien lässt sich nicht einfach beantworten; zumindest lassen sich kaum einzelne Erfolgsstrategien benennen, die allgemein empfohlen werden könnten. Ob eine Strategie gut ist oder nicht, ist in der Regel situations- und kontextabhängig.

Ellis (2008) skizziert die Verbindung zwischen Sprachenlernen und Sprachlernstrategien wie folgt: Die Verwendung von Strategien führt *nicht generell* zu besserem Lernen, sondern ist mit dem Lernstand verknüpft. So finden sich metakognitive Strategien eher bei fortgeschrittenen Lernenden, die Erfahrung mit dem Lerngegenstand gesammelt haben und diese Erfahrung für ihre metakognitiven Einschätzungen heranziehen können. Zudem verwenden fortgeschrittene Lernende *mehr* Strategien oder qualitativ *andere* Strategien. Strategien müssen flexibel eingesetzt werden, weil verschiedene Kompetenzbereiche nach unterschiedlichen Strategien verlangen. Schliesslich gibt es Hinweise darauf, dass sich auch nach Altersgruppe unterscheidet, was die jeweils optimale Strategie ist. Insbesondere für Erwachsene ist der Einsatz von metakognitiven Strategien, z.B. für die Planung und Überwachung des eigenen Lernens, besonders relevant, während jüngere Lernende eher von sozialen und interaktionalen Strategien (Einbezug von anderen) profitieren können.

Ähnlich fällt das Ergebnis von Oxfords (2011b) Übersicht zum Zusammenhang zwischen Strategiegebrauch und Sprachlernerfolg in verschiedenen Bereichen aus. Interessant ist, dass Untersuchungen teilweise auf einen linearen und teilweise auf einen kurvilinearen Zusammenhang zwischen Sprachkompetenz und Strategiege-

V Sprachlernstrategische Kompetenzen

brauch hinweisen. Ein kurvilinearere Zusammenhang wurde insofern beobachtet, als Lernende mit mittleren Sprachkompetenzen mehr Strategien verwenden als solche mit tiefen oder hohen Kompetenzen. Das könnte darauf zurückzuführen sein, dass mittelgute Lernende dank Strategiegebrauch die vorhandenen Ressourcen tatsächlich besser nutzen können, während schwächere Lernende zu grosse Lücken aufweisen, um diese mittels Strategien einfach zu kompensieren, und bereits gute Lernende je nach Kompetenzbereich kaum kompensatorische Strategien benötigen oder die relevanten Strategien automatisiert haben und nicht als solche bemerken.

2.2. Vermittlung von Sprachlernstrategien

In Anbetracht des komplexen, nur teilweise geklärten Zusammenhangs zwischen Strategiegebrauch und sprachlichen Kompetenzen mag es erstaunen, dass Strategien heute in viele Curricula und Lehrwerke Eingang finden. Zumindest ein Erklärungsansatz ist wohl darin zu finden, dass sich Strategien im Unterschied etwa zu Sprachbegabung, Motivation oder Massnahmen gegen Lernangst explizit vermitteln lassen. Strategiekennnisse können auch überprüft und beurteilt werden.

Den Sinn von Strategieinstruktion sehen etwa Rubin u.a. (2007) vor allem darin, auch weniger guten Lernenden die Instrumente in die Hand zu geben, über die gute Lernende bereits verfügen. Die Forschungsergebnisse zum Erfolg von Strategievermittlung sind allerdings nicht eindeutig. Dies hat unter anderem mit der grossen Vielfalt an Strategien und Settings zu tun, die bisher untersucht wurden. Während z.B. Ellis (2008) ein eher kritisches Fazit bezüglich der Effizienz und Effektivität von Strategievermittlung zieht, verweist Oxford (2011b) auf Studien mit Experimental- und Kontrollgruppen, die einen Effekt der Strategievermittlung auf das Sprachenlernen zeigen. Die Übersicht von Rubin u.a. (2007) wiederum weist auf unterschiedliche Effekte der Strategievermittlung je nach Kompetenzbereich hin: Beim Lesen und Schreiben in einer Fremdsprache zeigt sich eine

Wirkung, während weniger eindeutige Effekte auf das Hören, Sprechen und Vokabellernen zu beobachten sind. In einer Metastudie, die insgesamt 95 Stichproben mit Experimental-Kontrollgruppen-Designs einbezieht, findet Plonsky (2011) über die verschiedenen Studien hinweg immerhin einen als mittelstark einzustufenden durchschnittlichen Effekt (Cohens $d = 0.49$) von Strategieinstruktion auf den Erfolg bei der Sprachverwendung oder auf die Strategiekennnisse selbst. Spezifisch auf jüngere Lernende (< 12 Jahre) bezogen können wegen der geringen Anzahl von Studien bei dieser Altersgruppe nur beschränkt gültige Aussagen gemacht werden. Die vorhandenen Ergebnisse weisen aber auf einen stärkeren Effekt in dieser Altersgruppe hin.

2.3. Entwicklung im Bereich der strategischen Kompetenzen

Untersuchungen zur individuellen Entwicklung strategischer Kompetenz liegen vor allem aus anderen Gebieten als dem Fremdsprachenlernen vor. Aus der Entwicklungspsychologie ist bekannt, dass bereits Vorschulkinder Strategien verwenden können (Bjorklund, 2011), dass deren Transfer auf neue Aufgaben aber oft eine Hürde darstellt, die sie nicht selbstständig überwinden können (sog. *mediational deficiency*). Auch bei älteren Kindern werden noch alterstypische Einschränkungen beobachtet: Sie wenden bereits erworbene Strategien in neuen Situationen nicht ohne spezifische Aufforderung (z.B. durch eine Lehrperson) an (sog. *production deficiency*). Dies ist ein Defizit, das in allen Altersgruppen beobachtet werden kann. Es hängt damit zusammen, dass die ersten Anwendungen neu erworbener Strategien viele kognitive Ressourcen benötigen, solange die Abläufe nicht automatisiert sind, sodass es nicht gelingt, aus dem Strategieeinsatz einen Gewinn an Effizienz und Effektivität zu gewinnen (sog. *utilization deficiency*).

Nur in klar umrissenen Lerngebieten lassen sich Entwicklungsschritte im strategischen Handeln identifizieren, z.B. beim Zählen und Addieren,

V Sprachlernstrategische Kompetenzen

beim Lernen von Bilderlisten oder beim Kategorisieren von Informationen bei Aufgabenstellungen zur selektiven Aufmerksamkeit (Unterscheidung von Wichtigem und Unwichtigem). Auch da ist aber kein einheitliches Bild zu beobachten: Abhängig von Kontext und kognitiver Belastung setzen die Lernenden nämlich einmal einfachere und sicherere, und einmal komplexere und effizientere Strategien ein (Bjorklund & Rosenblum, 2002; Siegler, 1996).

Im Bereich der Sprachlernstrategien zeichnen Untersuchungen zum tatsächlichen Gebrauch von Strategien in unterschiedlichen Altersgruppen insgesamt ein noch weniger einheitliches Bild: So ist es möglich, dass ein und dieselbe Person je nach Umständen unterschiedliche Typen von Strategien einsetzt, um eine bestimmte Aufgabe anzugehen. Bei Sprachlernstrategien scheint es, anders als etwa im eng umrissenen Gebiet des Zählens und Addierens, mehrere, zum Teil frei kombinierbare Alternativen zu geben, und es scheint, dass eine „gute“ Verwendung von Sprachlernstrategien auf unterschiedlichen Wegen und mit unterschiedlichen Mitteln erreicht werden kann.

3. Ziele des Teilprojekts

Im Einklang mit den Zielen des Gesamtprojekts wollte das Teilprojekt zu den sprachlernstrategischen Kompetenzen letztlich einen Beitrag zur Entwicklung von computerbasierten Instrumenten leisten, mit denen im Fremdsprachenunterricht formativ-diagnostisch relevante Aspekte sprachlernstrategischer Kompetenz erfasst werden können. Wie bei der Entwicklung von Instrumenten, die psychologische Konstrukte messen, üblich (vgl. Wilson, 2005), wurde angestrebt, die Items und die Beschreibung des zu messenden Konstrukts parallel zueinander zu entwickeln. Da für unseren Anwendungsbereich auf nur wenig Vorhandenes zurückgegriffen werden konnte, sollte zuerst eine operationalisierbare Definition und eine Strukturierung des Bereichs der Sprachlernstrategien vorgenommen werden. Die darauf aufbauende Entwicklung von Items zu ausgewählten

Teilbereichen (den zu erfassenden *progress variables* nach Wilson, 2005) hatte notwendigerweise explorativen Charakter und konnte nur durch ein iteratives Vorgehen im engen Kontakt mit der Zielgruppe angegangen werden.

4. Entwicklungsarbeiten

4.1. Arbeitsdefinition, Eingrenzung

In einer frühen Projektphase wurde der Arbeitsbereich eingegrenzt und dabei der Bereich der Strategien definiert, um den es gehen sollte. Eine wichtige, auf dem Gebiet der Fremdsprachendidaktik übliche Abgrenzung ist diejenige zwischen Sprachlernstrategien und Sprachgebrauchsstrategien (Cohen, 1998; Cohen & Macaro, 2007). Bei dieser Unterscheidung ist das Ziel des Strategieeinsatzes das wesentliche Kriterium. Inhaltlich gibt es eine Überlappung zwischen den beiden Strategietypen. So kann etwa das Verwenden einer Sprache selbst auch eine Lernstrategie sein; Abrufstrategien wie das Bauen von Eselsbrücken können sowohl beim gezielten Sprachenlernen (Informationsspeicherung) als auch bei der Sprachverwendung in einem Kommunikationskontext (Abruf von etwas Gelerntem) genutzt werden. Im Projekt wurde der Schwerpunkt jeweils auf Strategien als *Sprachlernstrategien*, nicht als *Sprachverwendungsstrategien* gelegt. Für die Zwecke des Projekts wurde „Sprachlernstrategie“ – in Anlehnung an die Definition von Cohen (1998) – folgendermassen definiert:

Sprachlernstrategien können als Gedanken und Handlungen definiert werden, welche die Lernenden mit unterschiedlichen Bewusstseitsgraden (auf einem Kontinuum) aktivieren, um das Lernen einer Sprache zu unterstützen. Sprachlernstrategien können verschiedene kognitive und metakognitive Prozesse und Verhaltensweisen umfassen und allein oder in sozialer Interaktion aus-

V Sprachlernstrategische Kompetenzen

geführt werden. Die Angemessenheit einer bestimmten Sprachlernstrategie variiert je nach Kontext, Aufgabe und individuellen Voraussetzungen. Dadurch kann die Effektivität einer Strategie über verschiedene Situationen oder Lernende hinweg variieren.

Diese recht allgemein gehaltene Definition wird im Projekt auf das Alter (Primarschülerinnen und -schüler) und den institutionellen Kontext (Schule) der Zielgruppe bezogen interpretiert. Aus diesem Grund werden auch wenig komplexe Strategien, die nur wenige Handlungsschritte umfassen, bereits als solche bezeichnet, und nicht etwa als Techniken oder Taktiken, wie das in der Literatur verschiedentlich getan wird. Affektive Strategien werden nicht berücksichtigt, weil möglichst Strategien berücksichtigt werden sollen, die direkt auf schulische Lernhandlungen ausgerichtet sind.

Aus der theoretischen Literatur und bestehenden Strategieinventaren (Bialystok, 1978; Chamot, 1996; Cohen, 1998; Ellis, 1994; Gu u.a., 2005; Gürsoy, 2010; 2013; Naiman u.a., 1996; O'Malley & Chamot, 1990; Oxford, 1990; 2011a; 2011b; Oxford & Cohen, 1992; Rubin, 1981; 1987; Stern, 1983; Ting, 2009) sowie auf der Basis von aktuellen Lehrwerken (*Young World, Alex et Zoé, Mille feuilles*)³² und Curricula wurde ein umfassendes Inventar von Sprachlernstrategien zusammengestellt, das es erlaubt, Strategien auf verschiedenen Ebenen zu kategorisieren. Das Inventar wurde grundsätzlich für alle Strategien erstellt, die im Bereich des Lernens und Kommunizierens in einer Fremdsprache relevant sind. In diesem Bericht sind allerdings die Sprachgebrauchsstrategien ausgeblendet. Abbildung V-1 zeigt die Struktur des Inventars der Sprachlernstrategien auf drei Ebenen. Auf der ersten Gliederungsebene werden die Sprachlernstrategien nach grundlegenden kognitiven oder sozialen Merkmalen unterschieden, wobei drei Gruppen entstehen: kognitive Strategien, metakognitive Strategien und soziale Strategien.

32 | Im Teilprojekt zu den Sprachlernstrategien kam dabei *Mille feuilles* eine besondere Rolle zu, weil dort den lernstrategischen Kompetenzen eine prominente Rolle eingeräumt wird als in den anderen Lehrwerken.

Kognitive Strategien betreffen Informationsverarbeitungsprozesse, d.h. Informationsaufnahme, -speicherung, -verarbeitung, -abruf, sowie deren Ergebnisse. Besonders die Taxonomien von Gürsoy (2010) und Oxford (2011b) umfassen ein breites Spektrum von kognitiven Strategien und waren eine wichtige Grundlage für das vorliegende Projekt.

Auch auf der Ebene kognitiver Prozesse höherer Ordnung, der Metakognition, gibt es viele relevante Lernstrategien. Diese metakognitiven Strategien betreffen die Überwachungs-, Beurteilungs- und Kontrollprozesse des eigenen Denkens sowie das Wissen über das eigene Wissen und Denken.

Das deklarative metakognitive Wissen besteht aus verbalisierbaren Wissenskomponenten über das eigene Gedächtnis und seine Funktionsweisen (Wissen über die eigene Person sowie Merkmale von Aufgaben und Strategien). Die prozedurale Metakognition überwacht und kontrolliert Wahrnehmung, Gedanken, Wissen, Erinnerung sowie das Handeln und dient insbesondere der Selbstregulation des Handelns (Shimamura, 2008). Im Folgenden werden metakognitives Wissen und Prozesse, die nicht auf die Lerninhalte, sondern auf die Bedingungen und Eigenschaften des Lernens ausgerichtet sind, recht weit gefasst als metakognitive Lernstrategien verstanden.

Soziale Strategien stehen mit der Kontaktaufnahme zu anderen Lernenden und dem gemeinschaftlichen Lernen im Zusammenhang (Conrad, 2006). Sie sind dem Lernen nicht direkt förderlich wie kognitive Strategien, sondern wirken sich vermittelt über soziale Kontakte auf das Lernen aus. Soziale Strategien scheinen v.a. für jüngere Lernende relevant zu sein (vgl. Cohen & Macaro, 2007). Im Inventar stehen soziale Strategien, die kognitive Lernprozesse stützen und begleiten, im Vordergrund. Soziale Strategien, die z.B. über Motivationssteigerung wirken können, werden hingegen ausgeblendet. Die durch das Inventar abgedeckten Aspekte der Informationsverarbeitung beziehen sich vor allem auf den Austausch von Informationen durch das Ersuchen bzw. Anbieten von Hilfe sowie gemeinsame Aktivitäten zur verbesserten Speicherung oder dem Abruf von sprachlichem Wissen.

V Sprachlernstrategische Kompetenzen

Die drei Hauptkategorien von Sprachlernstrategien sind auf der zweiten Ebene (vgl. Abbildung V-2) weiter aufgegliedert. Die Bezeichnungen geben in der Regel kurz die Ziele der Strategieverwendung an. Sie stammen zum Teil aus der Literatur, andere wurden in der Absicht, Strategien aus verschiedenen Quellen zusammenzubringen, neu formuliert. Die dritte Ebene stellt eine weitere Konkretisierung

der zweiten Ebene dar. Das Inventar wurde im Projekt darüber hinaus noch um eine vierte Ebene ergänzt (nicht dargestellt), auf der insbesondere verschiedene Realisierungsvarianten aufgeführt sind. So umfasst die Nutzung von Informationsquellen (dritte Ebene) zur Informationssuche (zweite Ebene) auf der vierten Ebene die Varianten „Personen“ und „Medien“ als Informationsquellen.

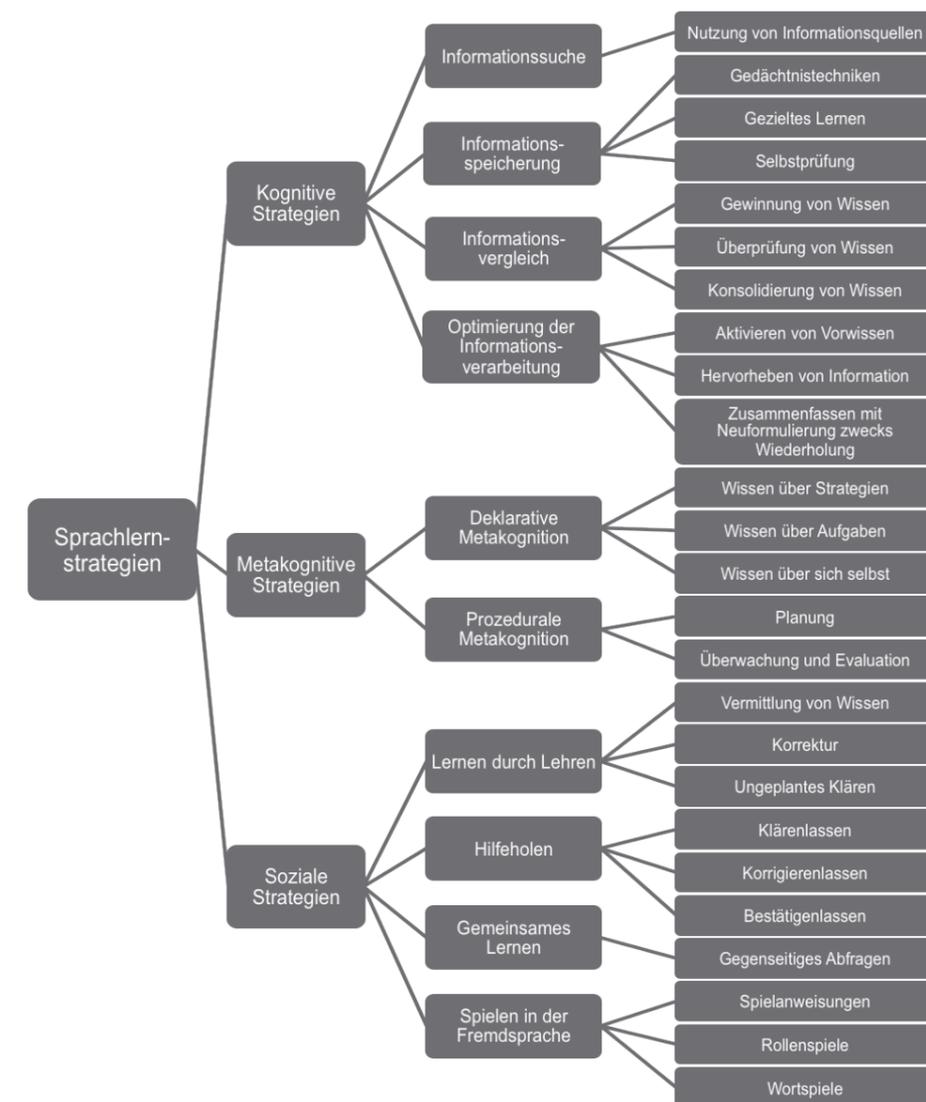


Abbildung V-1: Struktur des Inventars der Sprachlernstrategien (drei Ebenen)

V Sprachlernstrategische Kompetenzen

4.2. Instrumente

4.2.1. Drei Typen von Instrumenten

Nach längeren Vorarbeiten wurden schliesslich zu drei sehr unterschiedlichen Aspekten lernstrategischer Kompetenz Items erarbeitet.

1. Umfang und Vielfalt des individuellen Einsatzes von Sprachlernstrategien;
2. Wissen über Sprachlernstrategien und deren Eignung für bestimmte Lernaufgaben;
3. Zwei Aspekte des metakognitiven Könnens: zum einen das Einschätzen des eigenen Lernstandes, operationalisiert als die Fähigkeit, den eigenen Erfolg bei sprachlichen Aufgaben vorherzusagen (*judgement of learning, JoL*); zum anderen die metakognitive Selbstüberwachung, operationalisiert als die Fähigkeit, die Korrektheit der selbst ausgeführten sprachlichen Handlungen zu evaluieren (Sicherheitsurteil: *confidence judgement, CJ*).

Auf den nächsten Seiten wird zu jedem der drei Bereiche ein Aufgabenbeispiel gegeben, an dem das Aufgabenformat ersichtlich wird. Gleich im Anschluss daran werden jeweils formale und inhaltliche Aspekte erläutert.

1. Bereich: Umfang des individuellen Einsatzes von Sprachlernstrategien

Um Umfang und Vielfalt des Strategieeinsatzes durch die Primarschülerinnen und -schüler zu erfassen, wird auf das bewährte Format der Selbstaussage (SA) zurückgegriffen. Zu den bekannten Instrumenten, die ebenfalls dieses Format verwenden, gehören insbesondere das „Stra-

tegy Inventory for Language Learning“ (SILL) (Oxford, 1989; 1990) und der „Motivated Strategies for Learning Questionnaire“ (MSLQ) (Pintrich u.a., 1991). Neben diesen beiden Inventaren wurden im Projekt insbesondere auch Items aus dem für die Primarstufe adaptierten „Metacognitive Awareness Inventory for Children (MAI)“ (Sperling u.a., 2002) sowie Items aus der PISA-Studie von 2009 (Konsortium PISA.ch, 2011; Organisation for Economic Co-operation and Development, 2010) als Steinbruch für das eigene Inventar genutzt. Es wurden weiter auch Items formuliert, die an Sprachlernstrategien anschliessen, die in den Lehrwerken der Schülerinnen und Schüler vermittelt werden. Insgesamt wurden 70 Items dieses Formats aufgenommen. Es konnte aber nur eine Auswahl mit Schülerinnen und Schülern erprobt werden. Wie beim SILL wird in unserem Instrument eine fünfstufige Likert-Skala verwendet, um angeben zu lassen, wie häufig ein bestimmtes strategisches Verhalten praktiziert wird.

Die Nutzung von Selbstaussagen zum eigenen Verhalten ist eine direkte und effiziente Methode, um an grundlegende Informationen über das Lernverhalten zu kommen. Die Direktheit der Erhebungsmethode und die Selbsteinschätzung bringen eine gewisse Gefahr von Verzerrungseffekten bei der Beantwortung der Fragen mit sich. Trotzdem erfreuen sich solche Instrumente in der Forschung (z.B. zum SILL), bei der diagnostischen Beurteilung und in der Lernberatung gerade wegen ihrer Einfachheit grosser Beliebtheit. Für die Bewertung und Benotung von Lernenden wären sie dagegen wegen ihrer Manipulierbarkeit kaum geeignet.

2. Bereich: Wissen über Sprachlernstrategien

Für die Überprüfung des Wissens über Lernstrategien (deklaratives metakognitives Wissen) wären Selbstaussagen wenig geeignet (Schlagmüller u.a., 2001). Besonders auf der Primarstufe dürften Kinder rasch überfordert sein, wenn sie ihr eigenes strategisches Handeln über mehrere Lernerfahrungen hinweg abstrahieren und reflek-

V Sprachlernstrategische Kompetenzen

Abbildung V-2: Beispielitem zum 1. Bereich „Umfang des Einsatzes von Sprachlernstrategien“. Die Lernenden geben an, wie häufig sie die in weisser Schrift abgedruckte Strategie anwenden (fünfstufige Skala).

Abbildung V-3: Beispielitem zum 2. Bereich „Wissen über Sprachlernstrategien und deren Eignung für bestimmte Lernaufgaben“; Drag-and-Drop-Format zum Ranken der vorgeschlagenen Strategien.

V Sprachlernstrategische Kompetenzen

tieren sollten oder wenn sie aufgefordert würden, sich über bis dahin nicht reflektierte Lernhandlungen zu äussern und dabei metakognitives Wissen zu aktivieren. Dagegen bestehen bei Verfahren, bei denen die Effektivität verschiedener Strategiealternativen beurteilt werden soll, kaum Gedächtnis- und Verbalisierungsanforderungen dieser Art, weshalb sie auch mit Kindern durchgeführt werden können. Ein solches Verfahren sind Rankings. Dieses Format hat sich verschiedentlich bewährt, um bei Kindern und Jugendlichen lernstrategisches und metakognitives Wissen zu erfassen (vgl. Konsortium PISA.ch, 2011; Neuenhaus, 2011; Organisation for Economic Co-operation and Development, 2010; Schlagmüller u.a., 2001). Inhaltlich konnten nur beim Instrument von Neuenhaus Anleihen gemacht werden, weil die anderen nicht auf den Fremdspracherwerb ausgerichtet sind. Im vorliegenden Projekt sind nur Aufgaben zu einem relativ engen inhaltlichen Bereich entstanden, vorzugsweise solche zur Überprüfung des metakognitiven Wissens in Bezug auf die kognitiven Strategien zur Informationssuche und Informationsspeicherung. Diese Beschränkung hat damit zu tun, dass unser Inventar für die schulische Praxis relevant sein soll und die meisten Strategien, die in der Literatur, den Lehrwerken und den Lehrplänen erwähnt werden, auf diese Aspekte der Informationsverarbeitung ausgerichtet sind.

In Rankingaufgaben unterscheiden sich die vorgegebenen Antwortoptionen, anders als bei üblichen Multiple-Choice-Aufgaben, nur graduell.³³ Der Auftrag besteht darin, zu beurteilen, wie gut sich die beschriebenen strategischen Handlungen für einen vorgegebenen Zweck eignen. Im abgedruckten Beispiel (Abbildung V-3) geht es darum, die am besten, am zweitbesten und am wenigsten gut geeignete Strategie zum Lernen einer Reihe von Nomen, die Möbel bezeichnen, zu identifizieren und in die richtige Rangfolge zu ziehen. Zur Bewertung wird die gewählte Reihenfolge mit einer Referenzreihenfolge verglichen, die auf Expertenurtei-

33 | Briggs u.a. (2006) nennen diesen Itemtyp, der besonders geeignet sein soll, um Entwicklungsstände zu erfassen, Ordered Multiple-Choice (OMC).

len basiert. Die vorgegebenen Alternativen unterscheiden sich qualitativ hinsichtlich verschiedener Faktoren voneinander, z.B. bezüglich der Tiefe der Elaboration der zu lernenden Information, der Anzahl der Wiederholungen, der Qualität der Informationsquelle oder der Eignung der genannten Lernmaterialien. Mit der Frage (im Beispiel) „Wie lernt sie [Nina] am besten?“ anstatt z.B. „Wie würdest du die Wörter am liebsten lernen?“ wird versucht, rein individuelle Vorlieben beim Wortschatzlernen in den Hintergrund zu rücken und allgemeines Wissen über Strategien zu aktivieren. Das Vorgeben von Antwortoptionen (anstelle des Erfragens eigener Vorschläge) trägt ebenfalls zu einer „Objektivierung“ bei.

3. Bereich: Vorhersage der eigenen sprachlichen Leistungsfähigkeit (*judgement of learning*) sowie Sicherheitsurteil in Bezug auf eigene sprachliche Handlungen (*confidence judgement*)

Mit dem dritten Instrumententyp werden zwei Aspekte des metakognitiven Könnens in Bezug auf die eigene sprachliche Leistungsfähigkeit erfasst, nämlich deren Vorhersage (*judgement of learning* – *JoL*) und deren Überwachung (*confidence judgement* – *CJ*). Beim *JoL* werden die Lernenden mit einer (nicht trivialen und nicht komplett überfordernden) eingebetteten sprachlichen Aufgabe konfrontiert, im obigen Beispiel mit einer Übersetzungsaufgabe; daneben kommen auch Multiple-Choice-Aufgaben mit vier Optionen vor. In einem zweiten Schritt werden sie aufgefordert, ihre Erfolgchancen bei der Aufgabe einzuschätzen, ohne vorher die Aufgabe wirklich zu lösen. *JoL* sind Urteile darüber, wie gut früher gelernte Inhalte zu einem späteren Zeitpunkt reproduziert werden könnten. Durch ein solches Instrument wird überprüft, wie exakt der eigene Kenntnis- und Lernstand eingeschätzt werden kann. Obwohl das Urteil nicht direkt im Zusammenhang mit dem Lernen erhoben wird, sondern eine unbestimmte Zeit später, handelt es sich um die Überwachung eines Informationsverarbeitungsprozesses, also

V Sprachlernstrategische Kompetenzen

The screenshot shows a digital interface for a 'judgement of learning' task. On the left, a yellow sidebar contains instructions: 'Lies zuerst die Aufgabe gut durch.', 'Löse die Aufgabe nicht!', 'Überlege dir nur, ob du sie lösen könntest.', 'Kreuze an:', and 'Könntest du diese Aufgabe lösen?'. The main area has a yellow header with 'Du liest einen Satz. Könntest du diesen Satz auf Deutsch aufschreiben, wenn du müsstest?'. Below is a blue box with the text 'Ouvrez le livre.' and a white input field. A yellow footer asks 'Könntest du das?' with four radio button options: 'Nein, sicher nicht.', 'Eher nicht.', 'Eher ja.', and 'Ja, ganz sicher.' A right arrow icon is in the bottom right corner.

Abbildung V-4: Beispielitem zum 3. Bereich, Typ Vorhersage (*judgement of learning*), zum metakognitiven Wissen über die eigene sprachliche Leistungsfähigkeit mit Angabe des Grades der Sicherheit, mit der die eingebettete sprachliche Aufgabe richtig gelöst würde (vierstufige Skala).

The screenshot shows a digital interface for a 'confidence judgement' task. The left sidebar is identical to the previous one, but the main instruction is 'Du liest einen Satz. Wie lautet dieser Satz auf Deutsch? Schreibe den Satz auf Deutsch in das Feld.' The blue box contains 'Ouvrez le livre.' with a red arrow pointing to a white input field. The footer and options are the same as in the previous screenshot.

Abbildung V-5: Beispielitem zum 3. Bereich, Typ Sicherheitsurteil (*confidence judgement*), in Bezug auf die Überwachung der eigenen sprachlichen Leistung; Angabe des Grades der Sicherheit, mit der die eingebettete sprachliche Aufgabe richtig gelöst wurde (vierstufige Skala).

V Sprachlernstrategische Kompetenzen

um eine prozedurale metakognitive Fertigkeit. Die Fähigkeit, das eigene Lernen einzuschätzen, ist in der Praxis immer dann wichtig, wenn es um die Selbststeuerung des Lernens geht und damit um die Entscheidung, ob, wofür und in welchem Umfang weitere Ressourcen für Lernhandlungen eingesetzt werden sollen (Thiede u.a., 2003).

Auf einer anderen Prozessebene setzen die Sicherheitsurteile (CJ) an. Im Fall der CJ bearbeiten die Lernenden die eingebettete Sprachaufgabe tatsächlich. Parallel dazu überwachen sie den Aufgabenbearbeitungs- und -lösungsprozess und beurteilen ihren Erfolg. Auf Englisch wird dieser Prozess meist als *monitoring* bezeichnet. Beim gewählten Aufgabenformat geben die Lernenden im Anschluss an die Bearbeitung der Sprachaufgabe auf einer vierstufigen Skala an, wie sicher sie sind, dass sie die Aufgabe erfolgreich gelöst haben. Solche Sicherheitsurteile haben in der Praxis einen Einfluss darauf, ob und wann eigene Lösungen geäußert werden, ob auf einer Lösung aufgebaut oder eine Aufgabe noch einmal bearbeitet wird (Dunlosky & Metcalfe, 2009; Koriat u.a., 2001; Nelson & Narens, 1990). Sicherheitsurteile vermitteln gewissermassen zwischen den Lernstrategien und lösen deren Einsatz selektiv aus.

Der Bereich der metakognitiven Selbstüberwachung ist Teil wichtiger Taxonomien zu Lernstrategien (Ellis, 1994; Gürsoy, 2010; O'Malley & Chamot, 1990; Oxford, 2011b). Aufgaben zur Erfassung dieser Kompetenz konnten aber aus der Literatur keine übernommen werden und mussten deshalb selbst erstellt werden.

Der Einsatz der Aufgaben des gewählten Typs ist etwas komplizierter als der Einsatz anderer Instrumente, weil er in zwei zeitlich auseinanderliegenden Schritten erfolgen muss. In einem ersten Schritt schauen die Lernenden die eingebettete Aufgabe an, ohne sie zu lösen, und machen ihre Vorhersage, und erst später, wenn sie sich (voraussichtlich) nicht mehr genau an ihre Vorhersage erinnern können, werden sie aufgefordert, die sprachliche Aufgabe tatsächlich zu bearbeiten und ihr Sicherheitsurteil dazu abzugeben. Die Auswertung beider Teile basiert auf dem Vergleich zwischen dem Erfolg bei der Sprachaufgabe und jeder der beiden Einschätzungen. Es wird davon ausge-

gangen, dass die metakognitive Kompetenz bei der Vorhersage bzw. dem Sicherheitsurteil umso besser ist, je kleiner der Abstand zwischen dem sprachlichen Ergebnis und den beiden Einschätzungen ausfällt. Damit dieses Aufgabenformat funktioniert, müssen zwei Bedingungen erfüllt sein: Erstens müssen die Lernenden kooperieren, denn sobald das Prinzip durchschaut wird, ist es ein Leichtes, ein maximales Ergebnis zu erreichen, indem die eingebettete Sprachaufgabe nicht oder ganz falsch gelöst und dazu bei Vorhersage und Sicherheitsurteil der tiefste Wert angekreuzt wird. Zweitens muss die Sprachaufgabe zwar im Prinzip gemeistert werden können, aber für die Lernenden auch eine Herausforderung bieten. Wenn eine Aufgabe trivial oder aber unlösbar ist, fällt es sehr leicht, ein präzises Urteil über den Erfolg abzugeben. Diese Bedingungen sind eine Schwäche des Instruments. Ihr kann teilweise entgegengewirkt werden, indem die individuellen Lernenden gezielt mit eingebetteten Sprachaufgaben konfrontiert werden, deren Schwierigkeitsspektrum mehrheitlich eine Herausforderung darstellt. Dadurch entsteht zwar kein exakt messendes Instrument, aber doch eines, das für eine formativ-diagnostische Nutzung im Unterrichtskontext Anhaltspunkte bietet.

4.2.2. Zuordnung der Instrumente zu einer hierarchischen Prozessdimension

Ein strukturiertes Inventar der Sprachlernstrategien, wie es in Abschnitt 4.1 vorgestellt wird, ist eine wichtige, aber noch keine ausreichende Grundlage für die Entwicklung eines Instrumentariums, das unterschiedlich hoch entwickelte sprachlernstrategische Kompetenzen erfassen soll. Es gibt vielmehr Hinweise zur möglichen Breite sprachlernstrategischer Kompetenzen und macht keine expliziten Aussagen zu einer vertikalen Anordnung auf einer Entwicklungsdimension. Theoriegestützte Aussagen zum Verlauf von Lernfortschritten in den betroffenen Bereichen wären aber durchaus wünschenswert. Wie in Abschnitt 2.3 kurz beschrieben, können aus der bisherigen Forschung zu Sprachlernstrategien keine umfassen-

V Sprachlernstrategische Kompetenzen

den Aussagen zu einer solchen Erwerbshierarchie gemacht werden. Deshalb wurde im Projekt versucht, auf allgemeinere Grundlagen zurückzugreifen und dort Anhaltspunkte für eine Hierarchisierung des sprachlernstrategischen Wissens und Könnens zu finden.

Eine solche Grundlage, die Hinweise auf eine mögliche Hierarchisierung metakognitiven, also auch lernstrategischen Wissens und Könnens enthält, ist die überarbeitete Fassung von Blooms „Taxonomy of Educational Objectives“ (Bloom u.a., 1956).³⁴ In der überarbeiteten Fassung (Anderson u.a., 2001) ist zu den Wissensarten *Faktenwissen*, *konzeptuelles Wissen* und *prozedurales Wissen* das *metakognitive Wissen* neu hinzugekommen (Krathwohl, 2002; Pintrich, 2002). Das metakognitive Wissen umfasst die Unterarten a) strategisches Wissen, b) Wissen über kognitive Aufgaben und c) Wissen über die Stärken und Schwächen der eigenen Wissens- und Könnensbasis – drei Berei-

che also, die für das vorliegende Projekt offensichtlich relevant sind. Die Taxonomie umfasst neben der Klassifikation der Wissensarten auch eine quer dazu verlaufende Prozessdimension, die jeder der vier Wissensarten zugeordnet werden kann (vgl. Abbildung V-6). Sie besteht aus den Prozessen 1. Erinnern, 2. Verstehen, 3. Anwenden, 4. Analysieren, 5. Evaluieren und 6. Kreieren.

Dabei implizieren Prozesse mit höherer Nummer jeweils die früher genannten, sodass die später genannten Prozesse jeweils mindestens so anspruchsvoll sind wie die früher genannten. Die Prozessdimension stellt somit ein mögliches Gerüst für eine hierarchische Einordnung von Aufgaben dar. Wenn Lernende beispielsweise kritisch überprüfen (d.h. evaluieren) sollen, ob sich eine Lernstrategie bei einer bestimmten Lernaufgabe eignet, müssen sie sich an die Strategie erinnern, sie verstehen, sie anwenden können und das strategische Vorgehen auch analysieren können.

Wissensdimension	Kognitive Prozessdimension					
	1. Erinnern	2. Verstehen	3. Anwenden	4. Analysieren	5. Evaluieren	6. Kreieren
A Faktenwissen	■	■	■	■	■	■
B Konzeptuelles Wissen	■	■	■	■	■	■
C Prozedurales Wissen	■	■	■	■	■	■
D Metakognitives Wissen	■	■	■	■	■	■
	■	■	■	■	■	■
Legende	■ 1. KB: Umfang Strategienegebrauch		■ 2. KB: Wissen über Eignung von Strategien		■ 3. KB: Metakognitive Überwachung	

Abbildung V-6: Einordnung der Kompetenzbereiche (KB), die mittels der Instrumente erfasst werden sollen, in die überarbeitete Taxonomie von Bloom (Anderson u.a., 2001).

34 | Die primäre Motivation zur Schaffung dieser Taxonomie bestand darin, einen Referenzrahmen bereitzustellen, in dem Testitems aus verschiedenen Quellen zu denselben curricularen Zielen verortet und dadurch besser ausgetauscht werden können.

V Sprachlernstrategische Kompetenzen

Selbstaussagen und Rankingaufgaben (1. und 2. Bereich) erfordern gemäss der überarbeiteten Taxonomie von Bloom metakognitives Wissen der Unterart a) „strategisches Wissen“, während die Aufgaben zur metakognitiven Überwachung (3. Bereich) metakognitives Wissen der Unterart c) „Wissen über die Stärken und Schwächen der eigenen Wissens- und Könnensbasis“ implizieren. Auf der Prozessdimension reichen die Anforderungen bei den Selbstaussagen bis zu Stufe 4 „Analysieren“ (Durchschauen des eigenen, tatsächlich ausgeführten lernstrategischen Handelns als Voraussetzung für Aussagen zum eigenen strategischen Handeln). Bei den Rankingaufgaben reichen sie bis zu Stufe 5 „Evaluieren“ (Beurteilen der Eignung und Güte unterschiedlicher strategischer Handlungen). Bei den Überwachungsaufgaben wird beim metakognitiven Wissen Stufe 3 („Anwenden“) erreicht. Beim Sicherheitsurteil (aber auch ansatzweise bei der Vorhersage) zu eigenen sprachlichen Leistungen spielt ausserhalb des Bereichs des metakognitiven Wissens auch die Fähigkeit eine Rolle, die eigene sprachliche Leistung mithilfe der vorhandenen Sprachkenntnisse zu analysieren und zu evaluieren. Entsprechend werden durch diese Aufgabentypen alle Arten von Wissen abgedeckt, die je nach Aufgabenstellung für sprachliche Aufgaben relevant sind. Dabei handelt es sich um Evaluationsprozesse (Prozessdimension 5) bei den Wissensarten Faktenwissen, konzeptuelles Wissen und prozedurales Wissen.

Ein Scheitern bei den Überwachungs- oder Rankingaufgaben kann durch Mängel bei einem der beteiligten Prozesse verursacht werden. Beim formativen Einsatz der Instrumente im schulischen Kontext wird es darum gehen, herauszufinden, an welchem Punkt angesetzt werden muss, um die lernstrategischen Kompetenzen zu verbessern. Ein tiefer Wert bei den Selbstaussagen weist wegen der verwendeten Häufigkeitsskala vorerst nur auf eine seltene Verwendung der betreffenden Strategien hin. Ob deshalb Handlungsbedarf auch in Bezug auf die Beherrschung der Strategien besteht, gilt es im persönlichen Kontakt mit den Lernenden abzuklären.

4.3. Erprobung der Instrumente

Die Instrumente zur Erfassung sprachlernstrategischer Kompetenzen wurden in mehreren Etappen entwickelt. Dabei kam es in insgesamt fünf Phasen zu Kontakten mit dem Praxisfeld Schule: anlässlich von drei Aufgabenerprobungen mit qualitativem Schwerpunkt und von zwei Pilotierungen mit quantitativem Schwerpunkt.

In den ersten Kontakten³⁵ ging es darum, mit den kognitiven und sprachlichen Möglichkeiten des Zielpublikums vertraut zu werden und verschiedene Arten von Aufgaben zu erkunden. Ausgangspunkt für die Aufgaben waren hauptsächlich Lernziele zu den Bereichen Lernstrategien und Sprachbewusstheit aus dem Französischlehrwerk *Mille feuilles*. Ebenfalls aus *Mille feuilles* entlehnt war die Idee von relativ offenen Aufgabenstellungen, für deren Bearbeitung Hilfsmittel zur Verfügung gestellt werden. Im Zusammenhang mit anspruchsvolleren Aufgaben wurden auch Möglichkeiten des *Scaffolding* ausprobiert, d.h. der schrittweisen Anpassung des Anspruchsniveaus einer Aufgabe, bis es dem jeweiligen Lernenden angemessen ist. Die Lernenden wurden bei diesen Aktivitäten von Projektmitarbeitenden individuell begleitet. Für eine vertiefte Auswertung der Erfahrungen wurden Audioaufnahmen und ethnografische Notizen angefertigt. Die Erfahrungen der ersten Erprobungen waren vor allem insofern wertvoll, als Möglichkeiten und Grenzen der Arbeit mit dem Zielpublikum intensiv erfahren wurden. In Bezug auf Aufgabentypen zur Erfassung lernstrategischer Kompetenzen zeigte sich vor allem, dass offenere Aufgabenstellungen, bei denen die Lernenden neben der Arbeit an sprachlichen Aufgaben auch Strategien einsetzen, zeitlich aufwändig sind und nur wenige Datenpunkte ergeben. Aus diesem Grund wurde ab diesem Zeitpunkt nur noch an Aufgabenformaten gearbeitet, mit denen gezielt Wissens- und Könnenselemente erfasst werden können.

Im nächsten Erprobungsschritt wurden einigen Schülergruppen bereits solche Aufgaben in

35 | Die Erprobungen der drei ersten Phasen fanden in den Kantonen Solothurn und St. Gallen statt.

V Sprachlernstrategische Kompetenzen

neuen Formaten vorgelegt. Im Kern handelte es sich um Vorläufer der Selbstaussagen zur Häufigkeit strategischen Handelns, der metakognitiven Überwachung und des Rankings. In dieser frühen Form der Rankingaufgaben vergaben die Schülerinnen und Schüler noch Noten für die Effektivität oder Effizienz einer Reihe von Alternativen strategischen Handelns. Die Analyse der Audioaufzeichnungen und der Ergebnisse bei den Aufgaben zeigte, dass diese Aufgabenformate dem jungen Zielpublikum grundsätzlich zugänglich waren. Probleme bereiteten zu komplexe Anweisungen sowie Items, die genaues Verstehen verlangten (in der Schulsprache Deutsch).

Im folgenden Erprobungsschritt wurde der Übergang zu computerbasierten Aufgaben vollzogen. Bei der Erprobung standen zwei Ziele im Vordergrund, erstens die Überprüfung der *usability* der Aufgaben am Computer, zweitens detaillierte Hinweise zur Verständlichkeit und Lösbarkeit der neu formulierten Strategieitems. Methodisch wurde ein Think-aloud-Verfahren eingesetzt (Boren & Ramey, 2000; Gu u.a., 2005), um möglichst unverfälschte Einblicke in das Vorgehen und die Gedanken der Lernenden während des Bearbeitungsprozesses zu erhalten. Bei der Konzipierung des Vorgehens wurden bekannte Besonderheiten der Anwendung dieser Methode bei Kindern berücksichtigt. So wurde darauf geachtet, dem Kind ausdrücklich eine Expertenrolle zuzuteilen. Gleichzeitig wurde eine Smiley-Lampe gebaut, mit der das Kind ohne persönlichen Eingriff eines Erwachsenen, aber doch gut sichtbar zum Weitermachen ermuntert werden konnte.

Jeweils vier Kindern aus der dritten, fünften und sechsten Klasse wurde jeweils eine Auswahl von Überwachungs-, Selbstaussage- und (zukünftigen) Rankingitems unterbreitet. In Bezug auf die Rankingitems wurde die Erprobung genutzt, um Alternativen lernstrategischen Vorgehens zu sammeln, die den Schülerinnen und Schülern selbst plausibel erscheinen. Nach einer Einführung in die Aufgabe wurden den Lernenden lediglich Lernsituationen (ohne Handlungsalternativen) vorgestellt; die Schülerinnen und Schüler wurden dann aufgefordert, selbst Vorschläge für ein sinnvolles Vorgehen beim Lernen zu machen. Auf diese Weise

entstand eine ansehnliche Sammlung von Lernstrategien, die als Steinbruch für die Ergänzung und Anpassung der Rankingitems genutzt werden konnte. Damit wurde erreicht, dass die Sammlung der Rankingitems eine grössere Nähe zur Lern- und Lebenswelt der Zielgruppe aufweist.

Die erste von zwei Aufgabenpilotierungen fand 2013 in Französischklassen der Lehrplanregion Pässepartout und des Tessins sowie Englischklassen der Ostschweiz statt. Insbesondere wegen der oft mangelhaften Computerinfrastruktur wurde auf Papierversionen der Aufgaben zurückgegriffen. Insgesamt konnten Daten von über 400 Schülerinnen und Schülern erhoben werden. Die Lehrpersonen, welche die Pilotierung nach den Vorgaben des Projektteams durchführten, erfassten zudem den Zeitbedarf ihrer Klassen und notierten wiederkehrende Fragen zu den Aufgaben. Aufgrund der Auswertung wurden bei den Überwachungsaufgaben mit eingebetteten Sprachaufgaben im Multiple-Choice-Format die Distraktoren überarbeitet und ihre Anzahl auf vier erhöht. Bei eingebetteten Übersetzungsaufgaben wurden die zu übersetzenden Elemente stärker schuljahrspezifisch formuliert, um extreme Schwierigkeit bzw. Einfachheit zu vermeiden.

Die abschliessende Aufgabenpilotierung fand in zwei Etappen im Sommer und Herbst 2014 ausschliesslich in Französischklassen des Pässepartout-Raumes statt. Beide Male waren über 200 Schülerinnen und Schüler beteiligt. Diesmal wurden die Aufgaben über das Internet ausgeliefert, sodass sich nur Schulen beteiligen konnten, die auch über die entsprechende Infrastruktur verfügten bzw. Zugang zu entsprechender Infrastruktur hatten. In dieser Erprobung genossen die Aufgaben zum mündlichen Sprachgebrauch Priorität. In der ersten Etappe konnten aus dem Teilprojekt zu den Sprachlernstrategien lediglich einige Selbstausageitems mitausgeliefert werden. In der zweiten Etappe wurden immerhin sechs Rankingitems und 15 zweiteilige Überwachungsaufgaben aufgenommen.³⁶

36 | Den einzelnen Schülerinnen und Schülern wurden lediglich die sechs Rankingitems und zusätzlich eine Auswahl von 5 Überwachungsitems unterbreitet.

V Sprachlernstrategische Kompetenzen

4.4. Nutzung der Instrumente im Unterrichtskontext

Die Instrumente, die im Rahmen des Projekts zu entwickeln waren, sind grundsätzlich für eine formativ-diagnostische Nutzung in grösseren Abständen (d.h. in Intervallen von mehreren Monaten) gedacht. Die Ergebnisse sollten daher in einer Form ausgegeben werden, die im Kontext des Sprachunterrichts relevant und v.a. durch die Lehrpersonen leicht nutzbar ist. In einem serverbasierten Beurteilungssystem, wie es für die Auslieferung der Instrumente letztlich vorgesehen ist, könnten grundsätzlich sehr flexibel Berichtsfunktionen umgesetzt und von den Nutzern aufgerufen werden. Im Folgenden werden für die drei Typen von Instrumenten kurz einige Berichtsmöglichkeiten skizziert, die uns nützlich erscheinen. Die unter anderem vorgeschlagenen (normorientierten) Vergleiche einzelner Lernender mit ihrer Klasse und darüber hinaus mit einer grösseren Referenzgruppe, z.B. aller Schülerinnen und Schüler auf derselben Klassenstufe, sind allerdings nur möglich, wenn eine entsprechende Datenbank aufgebaut wird.³⁷

4.4.1. Umfang des Einsatzes von Sprachlernstrategien

Wie in Abschnitt 2.1. beschrieben, stehen die Häufigkeit der Nutzung von Sprachlernstrategien und die Sprachkompetenzen der Lernenden in einem komplexen Verhältnis zueinander. Eine Lehrperson kann die Aussagen zu Art und Häufigkeit der Strategienutzung durch Klassen und Individuen vor allem dann sinnvoll interpretieren und formativ nutzen, wenn sie diese zu ihrem gesamten Wissen über die betroffenen Schülerinnen und Schüler in Beziehung setzt, insbesondere mit deren sprachlicher Leistungsfähigkeit. Sowohl

37 | Die Daten, die hinter den folgenden exemplarischen Illustrationen stehen, stammen aus den Pilotierungen der Instrumente.

Überforderung als auch Leichtigkeit beim Fremdsprachenlernen können beispielsweise zu geringem Strategieeinsatz führen.

In Abbildung V-7 sind beispielhaft Informationen auf der Klassenebene zuhanden einer Lehrperson einer 5. Klasse dargestellt. Für 20 verschiedene Strategien wird gezeigt, mit welcher Häufigkeit diese (auf der fünfstufigen Skala) im Klassendurchschnitt verwendet werden (hellgraue Säulen) und wie häufig diese im Vergleich dazu von allen Schülerinnen und Schülern der entsprechenden Referenzgruppe verwendet werden (dunkelgraue Säulen). Aus dieser Abbildung kann die Lehrperson ableiten, dass die Strategie 14 „jemanden um Hilfe bitten, der die Zielsprache besser beherrscht, wenn man nicht weiterkommt“ in der ausgewählten 5. Klasse und in den 5. Klassen insgesamt am häufigsten benutzt wird. Es handelt sich dabei offensichtlich um eine einfache und weit verbreitete Strategie. Die Strategie 6 „Filme und Sendungen in der Zielsprache ansehen, um Sprachen zu lernen“ wird laut den Schülerinnen und Schülern am seltensten verwendet.

Anstelle der Verwendungshäufigkeit einzelner Strategien könnte es auch interessant sein zu wissen, wie viele der erfragten bzw. erhobenen Strategien (hier 20) alle Schülerinnen und Schüler einer Klasse im Durchschnitt sehr selten, selten, manchmal usw. verwenden (Abbildung V-8).

Ein Vergleich dieser Zahlen mit den entsprechenden Werten der Referenzgruppe könnte beispielsweise als Anhaltspunkt für die Frage dienen, ob Sprachlernstrategien insgesamt den richtigen Stellenwert haben oder auch, ob das Spektrum der Strategien, mit denen die Klasse vertraut ist, eher weit oder eher eng ist.

Analoge Informationen können auf der Ebene der einzelnen Schülerinnen und Schüler gewonnen werden. In einer Darstellung der Häufigkeit des Gebrauchs einzelner Strategien (vgl. Abbildung V-7) kann zu diesem Zweck für jeden einzelnen Schüler eine dritte Säule ergänzt werden, um die individuelle Verwendungshäufigkeit der 20 Strategien den beiden durchschnittlichen Häufigkeiten entgegenzustellen. Aufgrund der Ergebnisse können einzelne Lernende bei Bedarf gezielt mit bestimmten Strategien vertraut gemacht werden.

V Sprachlernstrategische Kompetenzen

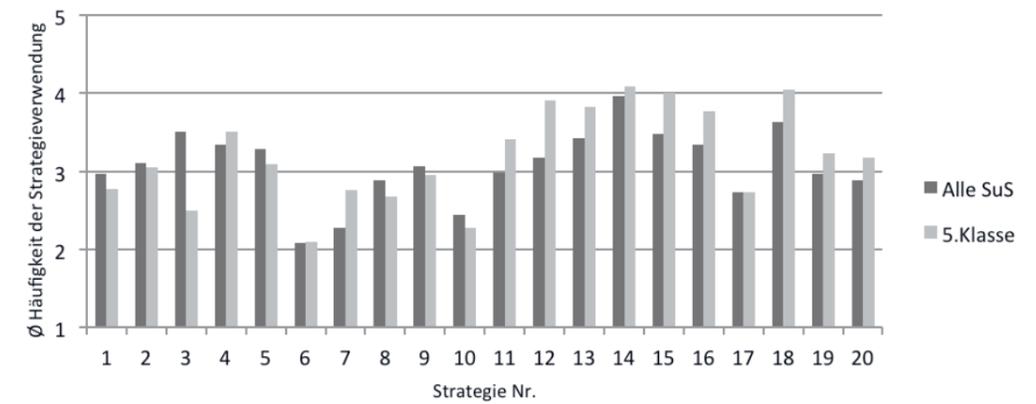


Abbildung V-7: Häufigkeit des Gebrauchs verschiedener Strategien in einer Klasse im Vergleich zur Referenzgruppe



Abbildung V-8: Strategiegebrauch in einer Klasse im Vergleich zu einer Referenzgruppe, dargestellt nach Häufigkeiten (1 = sehr selten; 5 = sehr oft). Lesehilfe (Bsp.): In der 5. Klasse werden im Mittel rund zwei Strategien sehr selten verwendet, fünf Strategien dagegen oft.

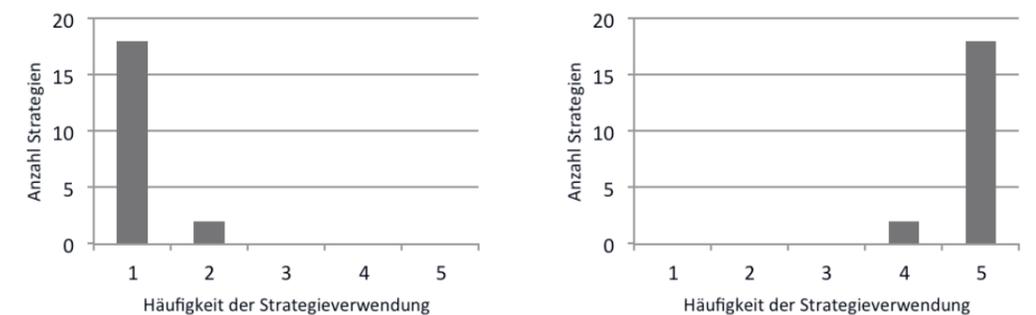


Abbildung V-9: Häufigkeit der Strategieverwendung durch einzelne Schülerinnen und Schüler

V Sprachlernstrategische Kompetenzen

Eine grafische Darstellung (vgl. Abbildung V-9) der Häufigkeit der individuellen Strategienutzung über alle Strategien hinweg kann rasch und anschaulich über die Strategienutzung durch die einzelnen Schülerinnen und Schüler informieren.

Abbildung V-9 zeigt zwei Extreme. Die Schülerin links braucht im Französisch kaum Sprachlernstrategien (im konkreten Fall wohl weil sie zweisprachig aufgewachsen ist); der Schüler rechts gibt bei allen Strategien an, diese oft oder sehr oft einzusetzen. Unabhängig davon, wie genau die Information in einer solchen Darstellung tatsächlich zutrifft, ist sie nützlich, weil sie die Lehrperson darauf hinweist, dass zumindest eine Klärung mit den betreffenden Lernenden angezeigt ist. Eine Darstellung, die zugleich die entsprechenden Informationen zur ganzen Klasse oder zur ganzen Referenzgruppe in Beziehung setzt, könnte zusätzlich hilfreich sein.

4.4.2. Wissen über Sprachlernstrategien und deren Eignung für bestimmte Lernaufgaben

Für die Erfassung des (metakognitiven) Wissens über Sprachlernstrategien werden im vorliegenden Projekt Rankingaufgaben vorgeschlagen, bei denen die Schülerinnen und Schüler jeweils drei Lernstrategien daraufhin beurteilen sollen, wie gut sie relativ zueinander geeignet sind, eine bestimmte Lernaufgabe zu lösen (vgl. Abschnitt 4.2.1.). Zur Auswertung werden bei jeder Aufgabe drei Paarvergleiche vorgenommen. Die Frage dabei ist jeweils: Entspricht die Rangreihenfolge, in der zwei Strategien zueinanderstehen, der Reihenfolge, die von Experten auf dem Gebiet empfohlen würde? Steht also die beste Strategie vor der zweitbesten und der drittbesten; steht die zweitbeste vor der drittbesten? Aufgrund der drei Vergleiche können pro Aufgabe null bis drei Punkte erreicht werden. Wer eine hohe Punktzahl erreicht, kann die Faktoren, welche die Effizienz und Effektivität der Sprachlernstrategien beeinflussen, richtig einschätzen – so der Ausgangspunkt für den Einsatz dieses Aufgabentyps.

Ähnlich wie bei den Selbstaussagen zum Strategieggebrauch kann auch hier die Interpretation und Nutzung der Ergebnisse über verschiedene Aufgaben hinweg oder in Bezug auf einzelne Aufgaben erfolgen, zudem können für die Nutzung im Unterricht neben individuellen auch gruppenbezogene Ergebnisse interessieren. Weiter können die Aufgaben und die jeweils drei Handlungsoptionen, die Teil der Aufgaben sind, selbst als Lerninput verwendet werden. Die unterschiedlichen Qualitätsmerkmale der verschiedenen Strategien können im Unterricht bewusstgemacht und die jeweils bevorzugten Optionen den Lernenden als Beispiele für gutes strategisches Handeln empfohlen werden. Bei der Beschäftigung mit vorgefertigten Strategien ist allerdings immer auch darauf zu achten, dass sich unter den gegebenen Umständen und für individuelle Lernende Varianten der empfohlenen Strategien durchaus noch besser eignen können.

Insgesamt wurden den Primarschülerinnen und -schülern, die an den Aufgabenpilotierungen teilnahmen, elf Rankingaufgaben vorgelegt. Diese erwiesen sich im Mittel als eher einfach. Zudem zeigte die psychometrische Analyse, dass sie nur unbefriedigend differenzieren, weil die Items offensichtlich viel individuelle Varianz erzeugen, die nicht der angezielten Kompetenzdimension (metakognitives Wissen über Sprachlernstrategien) zugeschrieben werden kann. Was an Items vorliegt, genügt daher nicht, um ein „Testinstrument“ im engeren Sinn zu generieren. Gerade bei formativ-diagnostischer Beurteilung muss das Ziel darin bestehen, breit gestreute Ergebnisse zu bekommen, damit auch in einer relativ homogenen Gruppe (gleiches Alter, gleicher Unterricht usw.) Lernende mit Nachholbedarf gut identifiziert werden können.

4.4.3. Metakognitives Wissen und Können in Bezug auf die eigene sprachliche Leistungsfähigkeit

Um die Fertigkeiten der Schülerinnen und Schüler bei der Überwachung der eigenen Sprachverwendung zu erfassen, wurden sie mit einer für sie

V Sprachlernstrategische Kompetenzen

relevanten Aufgabe mit sprachlichem Inhalt konfrontiert und mussten dazu in einem ersten Schritt die Erfolgswahrscheinlichkeit (*JoL*) und in einem zweiten Schritt den tatsächlich erzielten Erfolg (*CJ*) einschätzen (vgl. Abschnitt 4.2.1.). Metakognitive Überwachungskompetenzen sind eine wichtige Voraussetzung auch für einen gezielten Strategieinsatz (z.B. als Entscheidungsgrundlage für sprachliche „Reparaturen“) oder für das weitere Lernverhalten. Aus den erhobenen Daten kann berechnet werden, ob die Schülerinnen und Schüler bezüglich ihrer eigenen Leistungen eher eine Tendenz zur Überschätzung oder zur Unterschätzung haben. Beides zusammengenommen ist ein Mass für die Tendenz zur *Verschätzung*. Die Daten werden durch einen einfachen Vergleich der Einschätzung (*JoL* oder *CJ*) mit dem Ergebnis bei der Sprachaufgabe erzeugt. In den eingesetzten Aufgaben konnten im Sprachteil jeweils 0, 1, 2 oder 3 Punkte erzielt werden,³⁸ während der Erfolg auf einer vierstufigen Skala eingeschätzt werden sollte. Die Verschätzung beträgt jeweils dann null Punkte, wenn sich Ergebnis und Skalenstufe entsprechen, also z.B. 3 Punkte (korrekte Lösung) bei der Sprachaufgabe und „Ja, ganz sicher“ in der Selbsteinschätzung. Bei der Kombination 3 Punkte und „Nein, sicher nicht“ würde mit -3 Punkten beispielsweise eine maximale Unterschätzung vorliegen. Wenn die Ergebnisse bei mehreren Items für einen Schüler addiert werden, wird ersichtlich, wie gross die Tendenz zur Unter- bzw. Überschätzung ist. Werden die Absolutwerte addiert, ergibt sich ein Mass für die Verschätzung, das unabhängig ist von der Richtung.

Psychologische Untersuchungen haben gezeigt, dass die Qualität metakognitiver Überwachungsfertigkeiten mit der Gedächtnisleistung in einem positiven Zusammenhang steht. Dies gilt auch für den schulischen Kontext (vgl. Renner & Renner, 2001; Roderer & Roebbers, 2014; Schneider & Pressley, 1997; Thiede u.a., 2003). Schülerinnen

und Schüler mit besseren schulischen Fähigkeiten können diese auch genauer einschätzen. Kinder mit weniger guten metakognitiven Fertigkeiten neigen eher zur Überschätzung als zur Unterschätzung ihrer Kompetenzen, was sich auf das weitere Lernverhalten negativ auswirken kann. Ein schwaches Abschneiden im Bereich der metakognitiven Überwachung kann auf grundlegende Probleme hinweisen, die eventuell bei den schulischen Leistungen noch gar nicht offenkundig geworden sind.

Bei der Interpretation der Ergebnisse sollte im Auge behalten werden, dass moderat zu positive Selbsteinschätzungen, gerade in Bezug auf Bereiche, in denen die Schülerinnen und Schüler noch nicht viel Erfahrung haben, eher positiv zu werten sind, weil sie ein Zeichen von Zuversicht und Motivation sein können (Bjorklund, 2011). Zu negative Selbsteinschätzungen können das Symptom eines zu negativen Selbstkonzepts sein und die Entwicklung im fachlichen Bereich behindern (Dörnyei, 2001; Liu & Huang, 2011). Lehrpersonen sollten auf die Verschätzungswerte achten und, wenn diese ein bestimmtes Mass übersteigen, das Muster der Verschätzung genauer analysieren. Dabei ist grundsätzlich nicht nur zwischen den beiden Verschätzungsrichtungen, sondern auch zwischen den Ergebnissen für Vorhersage (*JoL*) und Sicherheitsurteil (*CJ*) zu unterscheiden. Verschätzungen bei den *JoL* lassen auf Probleme bei der Überwachung des eigenen Lernens und der Einschätzung des Lernstandes schliessen, während Verschätzungen bei den *CJ* eher auf ein problematisches Antwortverhalten hinweisen: Bei einer Tendenz zu Überschätzung werden Antworten vorschnell, ohne zu hinterfragen geäußert oder beispielsweise bei einer Prüfung vor der Abgabe nicht noch einmal überprüft. Dem Typ der Verschätzung entsprechend können das Feedback der Lehrperson sowie die weiterführenden Aktivitäten unterschiedlich ausfallen. Generell gilt, dass durch wiederholtes Üben von metakognitiven Urteilen die zugrundeliegende Kompetenz aufgebaut und auch mehr Sicherheit entwickelt werden kann.

Bei den Pilotierungen in den Klassen wurden einmal sechs und einmal 15 dieser zweiteiligen Aufgaben eingesetzt. Dazu kurz einige Ergebnisse, die für die Entwicklung eines Instruments zur

38 | Bei der Aufgabe in Abbildung V-5 sollte die Aufforderung „Ouvrez le livre“ auf Deutsch übersetzt werden. Bei der Beurteilung konnte für die folgenden Merkmale jeweils ein Punkt erreicht werden: 1) Imperativ bzw. Aufforderung; 2) Bedeutung *öffnen*; 3) Bedeutung *Buch*.

V Sprachlernstrategische Kompetenzen

metakognitiven Überwachung interessant sind:

Aus statistischer Sicht könnten die Ergebnisse bei Vorhersage und Sicherheitsurteil auf eine einzige Skala bezogen werden. Die entsprechenden Items haben generell einen befriedigenden Fit in Bezug auf eine gemeinsame psychometrische Dimension.³⁹ Sicherheitsurteile sind im Mittel präziser ausgefallen als Vorhersagen. Das war zu erwarten, handelte es sich doch um einen zweiten Einschätzungsversuch, der zudem direkt mit der tatsächlich gelösten Sprachaufgabe in Verbindung steht. Eine interessante Tendenz zeigt sich bei schwierigen Sprachitems im Vergleich zu mittelschweren: Die Sicherheitsurteile fallen zwar erwartungsgemäss bei schwierigen Sprachitems deutlich präziser aus als bei solchen, die von vielen Lernenden zumindest teilweise gelöst werden können. Bei den Vorhersagen ist jedoch eher das Gegenteil zu beobachten: Die Erfolgsaussichten werden gerade bei schwierigeren Items tendenziell überschätzt. Bei einer Weiterentwicklung der Diagnoseinstrumente zur metakognitiven Überwachung der sprachlichen Leistung sollte aus diesem Grund und auch, weil die Genauigkeit der Einschätzung generell klar weniger gut ist als bei den Sicherheitsurteilen, bei den Vorhersagen (*judgments of learning*) angesetzt werden.

5. Diskussion und Schluss

Im Projektteil zur Erfassung von Kompetenzprofilen von Primarschülerinnen und -schülern im Bereich der Sprachlernstrategien war viel Pionierarbeit zu leisten. Mit dem strukturierten Strategieinventar wurde eine wertvolle Ausgangsbasis für gezielte Entwicklungsarbeiten zu ausgewählten Bereichen sprachlernstrategischer Kompetenz geschaffen. Im Rahmen des Projekts wurden mit den Selbstaussagen zur Häufigkeit der Strategieverwendung, den Rankingaufgaben zum Wissen über Strategien sowie den Vorhersagen und

39 | Eine Analyse der Residuen zeigt aber, dass immerhin etwa 15% der verbleibenden Varianz dem Merkmalspaar *JoL/CJ* zugeschrieben werden können.

Sicherheitsurteilen zur eigenen sprachlichen Leistungsfähigkeit Instrumente zu drei recht unterschiedlichen Bereichen geschaffen, die aber alle schon auf der Primarstufe relevant sind.

Die vorliegende Sammlung von Selbstaussagen zur Strategieverwendung deckt das Inventar der Sprachlernstrategien am breitesten ab. Items zu Bereichen, die zurzeit noch unterrepräsentiert sind, sind relativ leicht zu ergänzen. Dabei geht es im Wesentlichen darum, relevantes strategisches Verhalten zu beobachten und in Form eines gut verständlichen Items zu beschreiben. Items mit einem konkreten Bezug zum Unterricht sind besonders relevant. Wie sinnvoll und dringend der Ausbau der Sammlung an Selbstaussagen ist, hängt insbesondere auch von deren zukünftiger Nutzung ab. Ein häufig wiederkehrender Einsatz eines Instruments zur Ermittlung des Strategiegebrauchs über verschiedene Lernjahre hinweg erfordert beispielsweise eine relativ grosse Sammlung.

Mit den Rankingaufgaben liegt ein Versuch vor, das Wissen über Strategien und deren Verwendung im Kontext „testbar“ zu machen. Wichtig ist daran, dass durch die Aufgabenstellung versucht wird, die Einschätzung der Qualität der vorgegebenen Lernstrategien von den subjektiven Bedingungen und Präferenzen des Schülers zu lösen, sie also gewissermassen zu objektivieren, sodass das Wissen über geeignetes und weniger geeignetes strategisches Vorgehen überprüft werden kann. Ähnlich wie bei den Selbstaussagen ist auch bei den Aufgaben zum Wissen über Strategien und deren Leistungsfähigkeit der Anwendungsbereich sehr breit. Grundsätzlich können solche Rankingaufgaben zu allen Sprachlernstrategien formuliert werden. Allerdings ist das Format der Rankingaufgaben wesentlich anspruchsvoller als das der Selbstaussagen. Die Erfahrungen im Projekt haben gezeigt, dass es keineswegs einfach ist, sinnvolle Strategievarianten zu finden, die ähnlich genug sind, um überhaupt vergleichbar zu bleiben, und gleichzeitig objektivierbare Unterschiede aufweisen, die ein festes Ranking erlauben. Bei jüngeren Lernenden stellt auch die notwendigerweise relativ differenzierte Sprache zur Unterscheidung der Alternativen ein Problem dar. Die bisher vorliegenden Aufgaben haben sich in

V Sprachlernstrategische Kompetenzen

den Pilotierungen nur mit Vorbehalt bewährt. Sie weisen wenig gemeinsame und viel individuelle Varianz auf, sodass sich daraus noch kein befriedigendes Erhebungsinstrument mit einer Testskala bilden lässt, auf der die individuellen Ergebnisse abgetragen werden könnten. Dieses Ergebnis mag auch damit zu tun haben, dass sich die Lernenden, mit denen die Items erprobt wurden, mit Lernstrategien so wenig auskennen, dass sie nicht in der Lage waren, in den vorgegebenen möglichen Lernstrategien die Merkmale zu identifizieren, die eine gute Lernstrategie kennzeichnen.

Aufgaben zur metakognitiven Überwachung der eigenen sprachlichen Leistungsfähigkeit sind wiederum einfacher zu erstellen. Die eingebetteten Sprachaufgaben sind leicht zu finden. Am besten handelt es sich dabei um Aufgaben zu sprachlichen Wissens- oder Könnenselementen, die vor einiger Zeit unterrichtet bzw. gelernt wurden, sodass die Lösung nicht einfach durch Wiedererkennen abrufbar ist. Das Skalenformat für die beiden Einschätzungen wird vorzugsweise über die verschiedenen Aufgaben hinweg unverändert verwendet. Noch weiter zu untersuchen ist allerdings der Einfluss der Aufgabenschwierigkeit der eingebetteten Sprachaufgabe auf das Ergebnis in den eigentlich interessierenden Kompetenzbereichen, nämlich der Fähigkeit zu Vorhersagen bzw. Sicherheitsurteilen in Bezug auf die eigene sprachliche Leistungsfähigkeit. Grundsätzlich sollte die Einschätzung der Erfolgsaussichten und des Erfolgs dann am schwierigsten sein, wenn die vorhandenen Sprachkompetenzen einen Erfolg und einen Misserfolg bei der eingebetteten Sprachaufgabe etwa gleich wahrscheinlich machen. Da sich die Sprachkompetenzen der Einzelnen aber unterscheiden, dürfte die Schwierigkeit der Items zur metakognitiven Überwachung nicht nur von den Kompetenzen in diesem Bereich abhängen, sondern auch von der Sprachkompetenz der Lernenden. Dies wiederum ist keine günstige Eigenschaft von Items, die nur die Fähigkeit im einen Bereich – nämlich der metakognitiven Überwachung – messen sollten.

Wie aufgrund des Forschungsstandes zu den Sprachlernstrategien nicht anders zu erwarten war, konnte im Rahmen des Projekts die „vertikale Dimension“ der fokussierten Kompetenzbereiche

nicht umfassend beschrieben, sondern höchstens punktuell beleuchtet werden. Aus gewissen Items lassen sich immerhin konkrete Einzelheiten zu Können und Nichtkönnen ablesen. Deshalb scheint uns, dass die Items und die mit ihnen generierten Ergebnisse im konkreten Schulkontext durchaus sinnvoll, aber eher spezifisch und mit einem klar formativen inhaltlichen Fokus eingesetzt werden können. Noch nicht möglich ist es, Kompetenzentwicklungen und Lernwege vor dem Hintergrund gesicherter Lernprogressionen zu beschreiben, die über längere Zeit als Orientierungshilfe für Lernen, Förderung und Beurteilung dienen könnten.

V Sprachlernstrategische Kompetenzen

6. Referenzen

- Anderson, L. W., Krathwohl, D. R. & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives*. New York: Longman.
- Beacco, J.-C., Byram, M., Cavalli, M., Coste, D., Egli Cuenat, M., Goullier, F. & Panthier, J. (2010). *Guide pour le développement et la mise en œuvre de curriculums pour une éducation plurilingue et interculturelle*. Strasbourg: Conseil de l'Europe. Abgerufen von <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806ae64a> [03.09.19]
- Bialystok, E. (1978). A theoretical model of second language learning. *Language Learning*, 28(1), 69-83.
- Bjorklund, D. F. (2011). *Children's thinking: cognitive development and individual differences* (5. Aufl.). Wadsworth: Cengage Learning.
- Bjorklund, D. F. & Rosenblum, K. E. (2002). Context effects in children's selection and use of simple arithmetic strategies. *Journal of Cognition and Development*, 3(2), 225-242.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H. & Krathwohl, D. R. (Hrsg.). (1956). *Taxonomy of educational objectives. The classification of educational goals. Handbook I: cognitive domain*. New York: David McKay.
- Boren, T. & Ramey, J. (2000). Thinking aloud: reconciling theory and practice. *IEEE Transactions on Professional Communication*, 43(3), 261-277.
- Briggs, D. C., Alonzo, A. C., Schwab, C. & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment*, 11(1), 33-63.
- Chamot, A. U. (1996). The Cognitive Academic Language Learning Approach (CALLA): theoretical framework and instructional applications. In: J. E. Alatis (Hrsg.), *Georgetown University roundtable on languages and linguistics*. Georgetown: Georgetown University, 108-115.
- Cohen, A. D. (1998). *Strategies in learning and using a second language*. London; New York: Longman.
- Cohen, A. D. & Macaro, E. (Hrsg.). (2007). *Language learner strategies: 30 years of research and practice*. Oxford; New York: Oxford University Press.
- Conrad, J. C. (2006). *Lernstrategien als Voraussetzung eines modernen, schülerorientierten Fremdsprachenunterrichts*. Duisburg-Essen: Universität Duisburg Essen.
- Credé, M. & Phillips, L. A. (2011). A meta-analytic review of the motivated strategies for learning questionnaire. *Learning and Individual Differences*, 21(4), 337-346.
- D-EDK (Hrsg.) (2014). *Lehrplan 21*. D-EDK. Abgerufen von <https://www.lehrplan21.ch/> [19.09.2019]
- Dörnyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.
- Dunlosky, J. & Metcalfe, J. (2009). *Metacognition*. Los Angeles: Sage.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2008). *The study of second language acquisition* (2. Aufl.). Oxford: Oxford University Press.
- Gu, P. Y., Hu, G. & Zhang, L. J. (2005). Investigating language learner strategies among lower primary school pupils in Singapore. *Language and Education*, 19(4), 281-303.
- Gürsoy, E. (2010). Investigating language learning strategies of EFL children for the development of a taxonomy. *English Language Teaching*, 3(3), 164-75.
- Gürsoy, E. (2013). The development of a children's inventory for language learning strategies (CHILLS). *Journal of Educational and Social Research*, 3(7), 263-272.
- Konsortium PISA.ch (Hrsg.) (2011). *PISA 2009: regionale und kantonale Ergebnisse*. Bern; Neuchâtel: BBT/EDK und Konsortium PISA.ch.
- Koriat, A., Goldsmith, M., Schneider, W. & Nakash-Dura, M. (2001). The credibility of children's testimony: can children control the accuracy of their memory reports? *Journal of Experimental Child Psychology*, 79(4), 405-437.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: an overview. *Theory Into Practice*, 41(4), 212-218.
- Liu, M. & Huang, W. (2011). An exploration of foreign language anxiety and English learning motivation. *Education Research International*, 2011, 1-8.
- Naiman, N., Fröhlich, M., Stern, H. H. & Todesco, A. (1996). *The good language learner* (2. Aufl.). Clevedon: Multilingual Matters.
- Nelson, T. O. & Narens, L. (1990). Metamemory: a theoretical framework and new findings. *The Psychology of Learning and Motivation*, 26, 125-173.
- Neuenhaus, N. (2011). *Metakognition und Leistung: eine Längsschnittuntersuchung in den Bereichen Lesen und Englisch bei Schülerinnen und Schülern der fünften und sechsten Jahrgangsstufe*. (Dissertation). Bamberg: Otto-Friedrich-Universität Bamberg.
- Neuner, G. (2005). Mehrsprachigkeitskonzept und Tertiärsprachendidaktik. In: B. Hufeisen, G. Neuner & European Centre for Modern Languages (Hrsg.), *Mehrsprachigkeitskonzept - Tertiärsprachenlernen - Deutsch nach Englisch* (2. Aufl.). Strasbourg: Council of Europe Publishing, 13-34. Abgerufen von <http://archive.ecml.at/documents/pub112G2003.pdf> [07.01.15]
- O'Malley, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge; New York; Melbourne: Cambridge University Press.
- Organisation for Economic Co-operation and Development (2010). *PISA 2009 results: what students know and can do: student performance in Reading, Mathematics and Science* (Volume I). OECD Publishing. Abgerufen von <https://www.oecd.org/pisa/pisaproducts/48852548.pdf> [19.09.2019]
- Oxford, R. L. (1989). *Strategy inventory for language learning* (SILL), Version 7.0 (ESL/EFL). Abgerufen von <http://richarddpetty.files.wordpress.com/2010/03/sill-english.pdf> [12.10.14]
- Oxford, R. L. (1990). *Language learning strategies: what every teacher should know*. New York: Newbury House Publisher.
- Oxford, R. L. (2011a). Strategies for learning a second or foreign language. *Language Teaching*, 44(02), 167-180.

V Sprachlernstrategische Kompetenzen

- Oxford, R. L. (2011b). *Teaching and researching language learning strategies*. New York: Taylor & Francis.
- Oxford, R. L. & Cohen, A. D. (1992). Language learning strategies: crucial issues of concept and classification. *Applied Language Learning*, 3(1&2), 1-35.
- Passepartout (Hrsg.) (2015). *Lehrplan Französisch und Englisch*. Abgerufen von <http://www.passepartout-sprachen.ch/services/downloads/download/533/get> [09.03.16]
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Into Practice*, 41(4), 219-225.
- Pintrich, P. R., Smith, D. A. F., Garcia, T. & McKeachie, W. J. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)* (Technical report No. 91-8-004). Ann Arbor, MI: National Center for Research to Improve Postsecondary Teaching and Learning. Abgerufen von <http://eric.ed.gov/?id=ED338122> [15.10.14]
- Plonsky, L. (2011). The effectiveness of second language strategy instruction: a meta-analysis. *Language Learning*, 61(4), 993-1038.
- Renner, C. H. & Renner, M. J. (2001). But I thought I knew that: using confidence estimation as a debiasing technique to improve classroom performance. *Applied Cognitive Psychology*, 15(1), 23-32.
- Roderer, T. & Roebbers, C. M. (2014). Can you see me thinking (about my answers)? Using eye-tracking to illuminate developmental differences in monitoring and control skills and their relation to performance. *Metacognition and Learning*, 9(1), 1-23.
- Rubin, J. (1981). Study of cognitive processes in second language learning. *Applied Linguistics*, 11(2), 117-131.
- Rubin, J. (1987). Learner strategies: theoretical assumptions, research history and typology. In: A. Wenden & J. Rubin (Hrsg.), *Learner strategies in language learning*. Englewood Cliffs, N.J.; London etc.: Prentice-Hall International, 15-30.
- Rubin, J., Chamot, A. U., Harris, V. & Anderson, N. J. (2007). Intervening in the use of strategies. In: A. Cohen & E. Macaro (Hrsg.), *Language learner strategies: 30 years of research and practice*. Oxford: Oxford University Press, 141-60.
- Schlagmüller, M., Visé, M. & Schneider, W. (2001). Zur Erfassung des Gedächtniswissens bei Grundschulkindern: Konstruktionsprinzipien und empirische Bewährung der Würzburger Testbatterie zum deklarativen Metagedächtnis. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 33(2), 91-102.
- Schneider, W. & Pressley, M. (1997). *Memory development between 2 and 20* (2. Aufl.). Mahwah, NJ: Lawrence Erlbaum.
- Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) (2004). Sprachenunterricht in der obligatorischen Schule: Strategie der EDK und Arbeitsplan für die gesamtschweizerische Koordination. Abgerufen von http://edudoc.ch/record/30008/files/Sprachen_d.pdf [27.11.15]
- Schweizerische Konferenz der kantonalen Erziehungsdirektoren (EDK) (Hrsg.) (2011). Grundkompetenzen für die Fremdsprachen. Nationale Bildungsstandards. EDK. Abgerufen von https://edudoc.ch/record/96780/files/grundkomp_fremdsprachen_d.pdf?version=1 [04.02.16]

V Sprachlernstrategische Kompetenzen

- Shimamura, A. P. (2008). A neurocognitive approach to metacognitive monitoring and control. In: J. Dunlosky & R. A. Bjork (Hrsg.), *Handbook of memory and metacognition*. New York: Psychology Press, 373-90.
- Siegler, R. S. (1996). *Emerging minds: the process of change in children's thinking*. Oxford; New York: Oxford University Press.
- Sperling, R. A., Howard, B. C., Miller, L. A. & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology*, 27(1), 51-79.
- Stern, H. H. (1983). *Fundamental concepts of language teaching: historical and interdisciplinary perspectives on applied linguistic research*. Oxford: Oxford University Press.
- Tiede, K. W., Anderson, M. C. M. & Theriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology*, 95(1), 66-73.
- Ting, L. (2009). Language learning strategies – the theoretical framework and some suggestions for learner training practice. *English Language Teaching*, 2(4), 199-206.
- Wilson, M. (2005). *Constructing measures: an item response modeling approach*. Mahwah, N.J.: Lawrence Erlbaum Associates.

VI Pratiche di valutazione e *feedback* nell'insegnamento della lingua straniera alla scuola elementare

Indagine sulla valutazione del francese alla scuola elementare in previsione di una possibile adozione di una piattaforma informatizzata per la valutazione formativa

Daniela Kappler

1. Introduzione	130
1.1. Il progetto "Profili di sviluppo"	130
1.2. L'indagine sulla valutazione	130
1.3. L'insegnamento del francese in Ticino	131
1.4. I docenti	132
2. Valutare e dare <i>feedback</i>	132
2.1. I risultati delle sperimentazioni	132
2.2. Il quadro teorico	132
3. Domande di ricerca	134
4. Metodologia	135
5. Risultati	136
5.1. Competenze linguistiche e metacognitive	136
5.2. Pratiche di valutazione e ipotesi sulle progressioni linguistiche	136
5.3. Valutazione informatizzata: ipotesi sul valore aggiunto e sul lavoro in classe	138
6. Discussione e conclusione	139
7. Bibliografia	141

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

1. Introduzione

La presente indagine è legata al progetto “Profili di sviluppo nell'apprendimento precoce delle lingue straniere a scuola” (qui di seguito “Profili di sviluppo”) organizzato dal Centro scientifico di competenza per il plurilinguismo di Friburgo e iniziato nel 2012 in collaborazione con la SUPSI-DFA (Dipartimento Formazione e Apprendimento) e l'Alta Scuola Pedagogica di San Gallo (PHSG). L'obiettivo del progetto consiste nel proporre spunti di riflessione per la creazione di strumenti informatizzati utilizzabili dai docenti di scuola elementare a sostegno della valutazione delle competenze produttive in lingua francese e delle abilità metacognitive. L'indagine qui descritta, che si è svolta nel 2014, rappresenta uno dei sotto-progetti di “Profili di sviluppo” e ha l'obiettivo di raccogliere dati tramite interviste semi-strutturate ad un gruppo di docenti ticinesi relative alle loro pratiche di valutazione, alle loro visioni sulla progressione nell'apprendimento del francese e alle loro ipotesi su una adeguata fruizione di una piattaforma informatizzata.

1.1. Il progetto “Profili di sviluppo”

Considerando i programmi educativi attuali, gli obiettivi dei nuovi piani di studio svizzeri e le correnti innovative in ambito della valutazione tramite piattaforma informatizzata, all'inizio del 2012 il Centro scientifico di competenza per il plurilinguismo ha avviato un progetto triennale in collaborazione con la SUPSI-DFA e l'Alta scuola pedagogica di San Gallo (PHSG). Il progetto “Profili di sviluppo” s'ispira ai concetti sulle progressioni nell'apprendimento di Black, Wilson & Yao (2011). Il progetto è uno studio esplorativo sulle progressioni nell'apprendimento della lingua straniera, in particolare nella comprensione e produzione orale, nonché sullo sviluppo di strategie d'apprendimento. Nel progetto si parte dal presupposto che i piani di studio cantonali per le lingue, l'insegnamento e la valutazione costituiscano l'insieme fon-

damentale per formare un costrutto sulle progressioni linguistiche e creare le basi per uno strumento diagnostico su una piattaforma informatizzata. La piattaforma potrebbe mostrare la progressione degli allievi in base ad un'analisi sistematica dei dati e sostenere i docenti nella valutazione formativa dei loro allievi. In particolare, si ipotizza che l'uso di uno strumento di valutazione diagnostico informatizzato favorisca l'accertamento di difficoltà degli allievi nell'apprendimento della lingua straniera.

Gli strumenti potrebbero esser usati al termine di un ciclo di insegnamento per verificare una o più competenze e procedere a un bilancio. Essi permetterebbero inoltre di seguire l'andamento durante l'anno scolastico sia di singoli allievi, sia di classi e di metterli in relazione. L'attuazione tecnica, e le caratteristiche dei compiti adattati per una valutazione diagnostica informatizzata e pensati per uno specifico gruppo di allievi sono state analizzate in diversi sotto-progetti (conf. le parti II a V del rapporto).

1.2. L'indagine sulla valutazione

Il sotto-progetto descritto in questo rapporto ha analizzato qualitativamente la conciliabilità degli strumenti con le pratiche dei docenti: si tratta di un'indagine per capire meglio in che misura questi strumenti informatizzati possono rappresentare un supporto utile per il rilevamento, l'analisi e la valutazione in ottica formativa delle competenze degli allievi interpellando i diretti interessati. Attraverso un'intervista semi-strutturata un gruppo di docenti di scuola elementare si è espresso in merito alle proprie pratiche didattiche e ad un'eventuale modalità d'uso autonomo e adeguato di una piattaforma informatizzata per la valutazione. Sono state analizzate alcune pratiche di valutazione in relazione agli obiettivi, agli interessi, alle priorità, ai bisogni e alle aspettative dei docenti. Questa analisi persegue l'obiettivo di individuare degli elementi rilevanti e pratici per la (ri)costruzione dei compiti e della piattaforma. Inoltre con la presente indagine si è cercato di

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

approfondire alcuni quesiti sorti dalle sperimentazioni, sia in ambito tecnico sia in ambito didattico, mettendoli in relazione con le pratiche di valutazione di cui sopra. I temi di fondo, insieme ai quesiti, confluiscono in domande di ricerca, dalle quali a loro volta scaturiscono le domande per l'intervista. Nello specifico, i temi affrontati sono: importanza dell'insegnamento del francese per un docente generalista (in termini di motivazione, competenze linguistico-didattiche, adesione ai principi e approcci pedagogici attuali); impostazione del manuale (compresa la valutazione). Nella preparazione dei contenuti delle interviste e nell'analisi ed interpretazione dei dati ci si è orientati alla metodologia della *Grounded Theory* (Tarozzi, 2008). Questa metodologia flessibile permette un adattamento dei primi assunti e domande di ricerca, una ciclicità nella raccolta e nell'analisi dei dati e correzioni nel processo di ricerca: permette inoltre di muoversi fra induzione e deduzione. L'indagine non ha lo scopo di creare una teoria, né di validare o confutare delle ipotesi, bensì, avendo carattere esplorativo, di riportare i risultati in modo descrittivo e di trovare degli elementi che indichino una vicinanza (o distanza) concettuale e metodologica nei confronti della valutazione diagnostica informatizzata (*computer-based assessment*), in modo tale da giungere ad alcune indicazioni pratiche e a spunti per ulteriori approfondimenti.

1.3. L'insegnamento del francese in Ticino

Qui di seguito si descrivono alcuni elementi che caratterizzano l'insegnamento del francese in Ticino. Ciò permette di contestualizzare e di mettere in relazione una serie di dati dell'indagine. Il Programma cantonale per la lingua francese (DECS-DS, 2004) raccomanda un approccio comunicativo orientato all'azione. Il manuale in uso da 10 anni *Alex et Zoé (AZ)* soddisfa gli orientamenti pedagogico-didattici cantonali: la finalità consiste in agire linguisticamente, all'interno di atti comunicativi vicini agli interessi e al mondo degli allievi. Dal 3° al 5° anno della scuola elementare si

usano consecutivamente i tre libri di testo di AZ e l'insegnamento del francese avviene in tre lezioni a settimana di circa 40 minuti. Per un passaggio graduale da un anno all'altro e poi alla scuola media, il cantone raccomanda ai docenti di svolgere circa 10 unità su 15 di AZ all'anno e le restanti unità l'anno scolastico successivo prima di cominciare con il nuovo libro. La struttura del manuale di AZ è uniforme: ogni unità è suddivisa in quattro capitoli e introduce un nuovo campo lessicale; AZ ripropone dopo varie unità gli stessi elementi linguistico-comunicativi insieme ad alcuni elementi nuovi, in particolare nei primi due libri. Ogni libro di testo per gli allievi è accompagnato anche da tre quaderni, un CD-ROM con esercizi, una guida pedagogica (GP) e un CD-ROM per l'insegnante. Il CD-ROM per l'insegnante favorisce lo sviluppo della capacità d'ascolto e del discernimento dei suoni, propone vocaboli da imparare, brevi enunciati e dialoghi in lingua francese. La guida pedagogica (GP) di AZ è ricca di indicazioni programmatico-didattiche e di attività opzionali che possono esser utilizzate per raggiungere i livelli previsti dal programma cantonale alla fine della scuola elementare.⁴⁰ D'altro canto la GP di AZ contiene pochi compiti e poche indicazioni per la valutazione. Gli item presenti sono per lo più destinati alla valutazione della comprensione orale, usando formati come l'associazione parola-immagine, simili alle attività del libro per gli allievi.

40 | A1 per la produzione orale, A1+ per le competenze ricettive orali e scritte; la correttezza formale nella produzione non è prioritaria. Nella versione più attuale del manuale si indica una preparazione per un livello A1.1 nel primo libro, e un livello A1.2 per i due libri successivi; in linea con il programma, come pure con il nuovo piano di studio ticinese (DECS-DS, 2015, pag. 113ff.).

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

1.4.
I docenti

In Ticino, dal 2012 cinque docenti e classi di scuola elementare hanno partecipato alle sperimentazioni in classe dei compiti per l'oralità e per le strategie (si veda parti II e V del rapporto). A loro nel 2014 si sono aggiunti quattro docenti disponibili per l'intervista sulle pratiche di valutazione del francese, per un totale di nove docenti. Durante il progetto una maestra ha cambiato sede, dal luganese al mendrisiotto e ha quindi potuto parlare di entrambe le esperienze. Sei docenti sono attivi nel luganese, due in comuni suburbani e tre in piccoli comuni, con pluriclassi. Due docenti lavorano con la stessa pluriclasse. Due docenti sono invece attivi nella parte settentrionale del Cantone. Una lavora in un comune di valle con pluriclasse, mentre l'altra lavora da anni con classi di prima, seconda, terza e quarta media. Quest'ultima è anche coordinatrice cantonale per l'insegnamento del francese alla scuola elementare e porta con sé esperienze legate al rapporto con il territorio e al passaggio fra scuola elementare e scuola media. Gli anni d'esperienza dei docenti (incaricati, con nomina o d'appoggio) nell'insegnamento del francese spaziano da pochi anni ad una ventina.

2.
Valutare e dare *feedback*2.1.
I risultati delle sperimentazioni

In seguito alle sperimentazioni con classi in diverse parti della Svizzera, i ricercatori hanno rilevato delle potenzialità e delle difficoltà di sviluppo della valutazione informatizzata, nonché delle problematiche di applicabilità del dispositivo per i docenti. Gli interrogativi sorti dalle sperimentazioni dei compiti nelle classi e raccolti qui di seguito in un quadro problematico, ruotano intorno ad alcune constatazioni.

- Le risorse informatiche delle scuole elementari in Ticino non permettono l'uso autonomo e adatto dello strumento. L'indagine si interroga su quali debbano essere le competenze organizzative e informatiche dei docenti, e di riflesso, delle scuole, per svolgere i compiti e approfittare positivamente del supporto.
- Per rendere i compiti accessibili agli allievi, sia i contenuti che i formati sono abbastanza semplici e simili a quelli dei manuali (*Mille feuilles* e *AZ*). La presente indagine permetterà di aggiungerci le modalità di rilevamento e valutazione scelte dai docenti.
- Le ipotesi sulle progressioni sono state elaborate analizzando la letteratura scientifica e le progressioni dei manuali in uso. L'indagine investiga se i criteri o le componenti linguistiche elaborati nel progetto si allineano con quelli ritenuti prioritari dai docenti.
- Infine, l'indagine dovrà affrontare la comprensione d'uso degli strumenti di valutazione e del *feedback* informatizzato da parte delle persone coinvolte: quali supporti o quali informazioni permettono ai docenti e agli allievi di capire funzione e funzionamento degli strumenti? L'uso adeguato degli strumenti di valutazione permette di determinare quale azione didattica e linguistica sia efficace e di adeguare l'insegnamento non solo in funzione agli obiettivi dei libri e dei curricula, ma anche ai progressi degli allievi.

2.2.
Il quadro teorico

Gli apporti teorici sono stati utili per precisare le domande di approfondimento della presente ricerca. La metodologia della *Grounded Theory* (Tarozi, 2008) permette poi di ricalibrare il qua-

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

dro problematico, inserendovi interrogativi nuovi o più precisi risultanti man mano dalle interviste. Gli assunti teorici che ispirano il progetto e il presente sotto-progetto sono proposti nel seguente quadro teorico e si incentrano sui concetti della valutazione formativa in relazione alla creazione di item per la valutazione informatizzata.

Di norma per svolgere una valutazione che si possa definire formativa, i dati dovrebbero essere ripresi alla fine dei compiti per informare l'allievo sul grado di padronanza raggiunto e interagire su possibili difficoltà diagnosticate a vari livelli e nelle varie competenze. Nel progetto e nella presente indagine si fa riferimento alla definizione di valutazione formativa compresa nei 10 principi esposti dal gruppo di riforma della valutazione del Regno Unito (Assessment Reform Group, 2002) definita "valutazione per l'apprendimento" (*assessment for learning*):

Assessment for learning

- ... should be part of effective planning of teaching and learning
- ... should focus on how students learn
- ... should be recognised as central to classroom practice
- ... should be regarded as a key professional skill for teachers
- ... should be sensitive and constructive because any assessment has an emotional impact
- ... should take account of the importance of learner motivation
- ... should promote commitment to learning goals and a shared understanding of the criteria by which they are assessed
- ... should receive constructive guidance about how to improve
- ... develops learners' capacity for self-assessment so that they can become reflective and self-managing
- ... should recognise the full range of achievements of all learners

La definizione concisa della valutazione formativa spesso chiamata in causa dagli studiosi del settore, la quale sembra tener conto anche dei 10 principi, è quella secondo Black & William (2009, pag. 9):

Practice in a classroom is formative to the extent that evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited.

Da queste due definizioni si possono desumere molteplici piste di approfondimento. Nel nostro caso ci atteniamo ad una variante specifica dei principi per l'insegnamento delle lingue straniere. Per la valutazione diagnostica informatizzata, ci interessa l'importanza del *feedback* sulle performance, il grado di trasparenza tra docenti e allievi e l'attenzione data allo sviluppo dell'autonomia degli allievi stessi.

Secondo Hill & McNamara (2012) e Ur (1999), il nocciolo della questione risiede nell'atteggiamento, nella consapevolezza e nell'abilità ad impostare correttamente la valutazione da parte dei docenti, così come nella loro trattazione dell'errore e nella loro trasparenza nei confronti degli alunni. Alcune ricerche empiriche qualitative sulla valutazione in classe (*classroom-based assessment*) di lingua straniera (Hill & McNamara, 2012; Torrance & Pryor, 2004), segnalano che non è sempre facile per i docenti rendere tutti i passaggi trasparenti e chiari; in particolare nelle classi di scuola elementare, in cui l'attribuzione di senso alle cose da parte degli allievi non coincide sempre con quella degli adulti. Il discorso sulla trasparenza viene perciò ripreso nell'intervista, per capire quali siano le possibili impostazioni per la valutazione formativa.

Nella letteratura di genere, quando si parla di valutazione, o *assessment*, ci si concentra maggiormente su contenuti, modalità di rilevamento e criteri diagnostici, mentre non di rado si trascurava quello che la stessa comporta a livello di *feedback*, lasciando il compito ai docenti di formulare (o meno) un *feedback*. Secondo Ur (1999) il *feedback* è un'informazione che viene data al discente in merito alla sua prestazione in un compito/attività, in questo caso in lingua straniera, con l'obiettivo di migliorarne la prestazione. Il *feedback* formativo

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

implica qualcosa di più ampio rispetto a ciò che succede nella valutazione “convergente” o “di controllo”, la quale consiste in un input dal docente e una risposta dell'allievo, a cui segue una valutazione e un *feedback* sulla prestazione.

Il *feedback* è legato ad un giudizio sulla prestazione rispetto ad un compito, può riferirsi alla persona o ad una competenza e può essere orale o scritto, di una o più parole, ma anche non verbale. Può anche contenere puntuali suggerimenti o riformulazioni appropriate (come la *modélisation*) per migliorare la prestazione (Hill & McNamara, 2012; Ur, 1999). L'obiettivo principale del *feedback* formativo è proprio quello di rendere l'allievo attento a determinati elementi della sua performance, implicare delle modifiche comportamentali e accrescere la consapevolezza e l'autonomia nell'apprendimento e nell'auto-valutazione (Brookhart, 2011; Hill & McNamara, 2012), verosimilmente anche nella valutazione dei pari. I *feedback* tra compagni sono definiti *peer-feedback*. Alcuni studiosi avvertono comunque anche di eventuali rischi: dati empirici indicano che il *feedback*, se fornito immediatamente dopo una risposta può influenzare anche negativamente l'atteggiamento di chi sta ancora rispondendo ad item successivi (Hill & McNamara, 2012).

Nel progetto europeo sulla valutazione informatizzata delle competenze linguistiche di una lingua straniera, *Dialang* (Alderson, 2005), il *feedback* è immediato, quindi fornito poco dopo la prestazione, perché la valutazione diagnostica non ha senso né valore se viene data a distanza di tempo (settimane o più) durante il quale il discente difficilmente ricorda la tipologia di compito e la prestazione. Dare *feedback* significa sostenere l'apprendimento e favorire la consapevolezza dell'allievo riguardo ai processi coinvolti. Se s'intende far usare degli strumenti informatizzati per una valutazione formativa, il progetto deve affrontare una riflessione sul *feedback* formativo con i diretti interessati: l'indagine permette dunque di delineare le tipologie di *feedback* che i docenti utilizzano e quali *feedback* possano utilizzare in seguito con il supporto degli strumenti di valutazione informatizzata.

3. Domande di ricerca

Le domande di ricerca affrontano il quadro problematico sorto dall'assegnazione dei compiti, dai dati empirici e dalle riflessioni tratte dal quadro teorico, in un approccio qualitativo e descrittivo. Non è intenzione di questa indagine valutare o giudicare le pratiche del campione di docenti che si è reso disponibile. Il quesito di fondo consiste nel capire in che misura questi strumenti informatizzati possano diventare parte integrante, calibrata ed utile ai docenti per il rilevamento, l'analisi e la valutazione formativa delle competenze degli allievi, ossia un valore aggiunto.

Qui di seguito proponiamo i tre macro argomenti o tematiche che sono stati analizzati (le singole domande sono state delineate in una traccia per l'intervista):

1. le interpretazioni dei docenti in merito alle competenze e alle progressioni linguistiche e alle competenze informatiche proprie e quelle dei loro allievi;
2. i comportamenti dei docenti nel rilevamento, nella valutazione formativa dei loro allievi, *feedback* incluso e nell'adattamento del programma in base ai dati;
3. le modalità di lavoro con uno strumento informatizzato di valutazione e la sua utilità nella pratica didattica.

Per ognuna delle tre tematiche si sono prese in considerazione le competenze linguistiche ricettive e produttive e una serie di componenti mirate dell'insegnamento delle lingue straniere, ossia il ruolo del lessico, della morfosintassi, delle funzioni, della pronuncia, della *langue de classe*, delle strategie e della motivazione. Nella presente indagine si considera anche la componente *éveil aux langues* (l'interesse, la sensibilizzazione alle lingue e la consapevolezza dei fenomeni linguistici, Beacco et al., 2011), in quanto questo aspetto era stato considerato all'inizio del progetto “Profili di sviluppo” e, come le strategie, trova più spessore

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

e rilevanza nei nuovi piani di studio regionali. Per quanto concerne le strategie, finora in Ticino non si richiede formalmente ai docenti di verificarne lo sviluppo, pertanto anche nell'intervista si tende ad approfondire principalmente la valutazione dell'oralità.

Per esplorare il primo argomento sintetizzato sopra si sono formulati dei quesiti sull'insegnamento e l'apprendimento del francese all'inizio, durante e verso la fine dei tre anni (motivazione e apertura, priorità, progressione e padronanza delle varie abilità linguistiche, sviluppo di strategie). Questo proposito ci porta poi a ragionare sull'origine delle ipotesi degli insegnanti relative alle progressioni. Ci soffermiamo dunque sulla seconda tematica incentrata sulla modalità con cui loro svolgono il rilevamento, la valutazione e il *feedback* e ne studiamo la preparazione (differenziazione, materiali, scelta di competenze da valutare), lo svolgimento (incluso l'autovalutazione e dai *peer*), e l'analisi (criteri, riferimenti a standard, alla classe o *ad personam*, tipologia di *feedback*, condivisione e adattamenti). La discussione verte anche sul ruolo e sul valore che questo gruppo di docenti attribuisce alla valutazione e sugli aspetti critici della valutazione.

Per affrontare la terza tematica, è stato mostrato un modello cartaceo⁴¹ di una piattaforma di valutazione comprensivo di diverse interfacce possibili, alcune con esempi di item, altre per l'analisi e la valutazione. Le domande riguardano quindi l'idoneità e la chiarezza delle interfacce sia per gli allievi che per i docenti. Si è cercato anche di capire, se un tale modello possa essere utilizzato dai docenti in classe, e in caso affermativo con quale frequenza, in che contesto (dentro o fuori dall'aula) e con quale obiettivo (a livello di risultati). E non da ultimo, si è cercato di rilevare quale *feedback* sarebbe utile che un dispositivo informatizzato offrisse loro. Un altro aspetto affrontato riguarda l'utilità di ulteriori funzioni di *feedback* da includere nel dispositivo.

41 | In seguito a difficoltà riscontrate in diversi ambiti coperti dal progetto, non è stato possibile mostrare ai docenti una versione concreta della piattaforma informatizzata, bensì una bozza cartacea.

4. Metodologia

L'indagine è modellata intorno ad un disegno di ricerca che comprende diverse procedure operative, tra cui: formulazione del quadro problematico dal progetto; sviluppo del quadro teorico; creazione di domande di ricerca; scelta della metodologia; sviluppo di temi, domande ed esempi per le interviste; svolgimento delle interviste; trascrizione delle interviste in un corpus; trasposizione delle risposte in uno schema di monitoraggio; selezione dei dati rilevanti; riflessione sui dati.

L'obiettivo dell'indagine risiede nell'individuare uno spettro di elementi pertinenti per la (ri)costruzione dei compiti e della piattaforma. La traccia dell'intervista è suddivisa nei tre macro-argomenti riportati nel capitolo 3, ai quali abbiamo subordinato ulteriori categorie. Queste a loro volta incorporano una serie di domande semi aperte e aperte. Per questa indagine si è ritenuto indispensabile cercare di formulare delle domande e dei commenti, che non diano ai docenti la sensazione di essere giudicati. Per il resto, la tipologia di domande scelta permette di elicitarle delle esperienze, delle riflessioni, dei bisogni e delle opinioni e favorisce inoltre l'aggiunta, e la creazione di nuove domande nel corso dell'intervista. La traccia serve, insieme allo schema di monitoraggio, a sistematizzare e codificare lo spettro di risposte.

Le interviste durano all'incirca 50 minuti e sono tutte registrate e trascritte con l'ausilio di un programma specifico (express scribe; NCH Software, s.d.). I dati tratti dalle risposte sono poi stati inseriti in uno specifico riquadro in corrispondenza ai due principali parametri dello schema di monitoraggio; ovvero laddove le dimensioni tematiche e le sotto-categorie si incrociano con le singole componenti mirate e le competenze. Questo procedimento ha permesso di verificare prima di tutto se sussiste (almeno) una risposta/un dato per i criteri seguenti: dimensione, sotto-categorie, competenze e componenti esplorate. In un secondo momento i dati sono stati analizzati e interpretati seguendo indicazioni metodologiche di studi socio-educativi (Guidicini & Castrignano, 1997; Hill & McNamara, 2012; Tarozzi, 2008; Trincherò, 2002).

5. Risultati

Le risposte riconducibili a dati significativi sono riportate qui di seguito seguendo la chiave di lettura dei tre macro-argomenti presentati nelle domande di ricerca. Il tema delle progressioni linguistiche è stato inglobato nel secondo macro-argomento, in quanto risultante dalla riflessione che i docenti fanno sulle loro pratiche di valutazione. Per le medesime ragioni, si sono accorpate le competenze informatiche alle ipotesi relative alla valutazione informatizzata.

5.1. Competenze linguistiche e metacognitive

Per quanto concerne l'interpretazione, da parte dei docenti, delle proprie competenze linguistiche e didattiche si sono riscontrate risposte diverse. Ad ogni modo i docenti non di madrelingua fanno spesso uso del supporto audio (CD-ROM) del manuale AZ per esporre gli allievi ad un modello sicuramente corretto, ma anche per esporli ad una gamma di voci e intonazioni diverse, nonché per svolgere la valutazione. Questo non significa che i docenti non parlino in francese in classe, fin da subito oppure progressivamente o che non abbiano la percezione di migliorare linguisticamente e didatticamente con gli anni d'insegnamento. La maggioranza degli insegnanti cerca di sviluppare la *langue de classe*, offrendo spesso attività ludiche (extra) o materiale autentico e cerca di stimolare la voglia di comunicare in francese, convinta che permetta a tutti di progredire nella produzione linguistica.

La competenza ritenuta prioritaria è l'oralità, in primis l'ascolto, inteso come apertura e scoperta di suoni e parole nuove, come discernimento di suoni e come comprensione orale. Quest'ultima sembra facilitata dall'uso di gesti, immagini e di una pronuncia corretta. Progressivamente si passa alla produzione orale. L'attenzione si focalizza sulla pronuncia (ripetendo), in modo tale che non solo il docente ma anche i compagni possano

capire ciò che un allievo dice. I docenti associano al lavoro sull'ascolto anche la lettura e il riconoscimento delle parole scritte facendo ad esempio leggere dei fumetti agli allievi dopo aver più volte ascoltato ripetutamente determinate parole con il CD. Secondo diversi docenti, l'uso del *Dico-memento* (dizionario "per campo lessicale" costruito man mano dagli allievi con ogni unità) e del *Cahier d'exercice*, favorisce lo sviluppo delle competenze scritte degli allievi; permette loro di "fissare" la "strana" trasposizione fonema-grafema della lingua francese e permette di svolgere di tanto in tanto una riflessione sugli aspetti grammaticali. Una riflessione metalinguistica che coinvolga altre lingue tramite attività di *éveil aux langues* non è ancora diffusa. Simile è il discorso sullo sviluppo mirato delle strategie. Più volte abbiamo rilevato che le strategie vengono tematizzate ogni tanto, in occasione di riflessioni (spontanee), di discussioni collettive o individuali, oppure incluse in attività di apprendimento del lessico in vista di un test.

5.2. Pratiche di valutazione e ipotesi sulle progressioni linguistiche

Entrando in merito alle pratiche e al significato della valutazione sommativa e formativa notiamo che il valore e il ruolo, ma anche la metodologia e le conseguenze della valutazione soggiacciono a interpretazioni e usanze diverse. Si sono rilevati anche svariati dubbi in proposito. Le risposte su cosa significhi un allievo "medio" o "bravo" sono molteplici. Le opinioni divergono anche rispetto a quali siano i criteri sui quali basare le valutazioni sommativa e/o formative: su criteri specifici, sugli standard rispetto alla classe nel suo complesso (ad. es. "in media, segue abbastanza bene"), sull'impegno o sulla progressione del singolo allievo in generale o in merito a singole competenze o componenti mirate. Nello specifico, alcuni docenti che lavorano con gli stessi allievi tutti i giorni, agiscono e differenziano le attività in base a osservazioni e intuizioni, senza avvalersi necessariamente di strumenti sistematici o di una for-

male valutazione *evidence-based*. I momenti di rilevamento formale sono realizzati usando in primis i test presenti nella GP di AZ, i cui formati si concentrano sull'associazione di un'immagine e una parola pronunciata (e poi scritta). Si tratta per lo più di compiti cartacei. Seppur non di rado ritenuti superficiali, in quanto gli allievi riescono a trovare la risposta giusta senza impegnarsi troppo o non appropriati per rilevare tutte le competenze, oppure un po' *old school*, i docenti li usano per una questione di tempo o perché non tutti pensano di riuscire ad inventare cose nuove e diversificate. Alcuni docenti ritengono anche che la strutturazione del pensiero diventi difficile per i bambini se l'input non è accessibile e ben circoscritto, breve e con una struttura data (del tipo *chunk*). Nel complesso, per gli insegnanti intervistati, è importante assegnare compiti fattibili, che gli allievi riescono a risolvere anche da soli sviluppando l'autonomia e l'autostima.

La produzione orale, su cui l'indagine si sofferma in particolare, viene rilevata e valutata con la lettura di parole o dialoghi o in momenti di interazione, con attenzione alla pronuncia. A volte i docenti esaminano anche campi tematici proposti tempo addietro e funzionali all'introduzione di una nuova *unité* di AZ (che riprende l'argomento) per vedere cosa rimane a lungo termine. La *langue de classe* o le strategie non sono oggetto di valutazione di per sé, ma durante l'intervista hanno suscitato interesse. A volte il momento e la modalità di rilevamento viene decisa dal docente, a volte negoziata con la classe, indice di un grado non indifferente di condivisione e trasparenza. Per la produzione orale, guidata o autonoma, i docenti tendono a dare importanza principalmente al contenuto. A seconda del compito, i criteri per valutare la competenza orale includono la realizzazione positiva della funzione comunicativa, l'uso appropriato del lessico e di una sintassi accettabile o comprensibile e, molto spesso, una buona pronuncia. Prima delle interviste, nessuno dei docenti aveva mai registrato le *performance* degli allievi.

L'esperienza di osservazione e valutazione permette ai docenti di fornire la propria opinione riguardo al quesito sulle progressioni (sull'arco dei tre anni). Secondo alcuni, le difficoltà che possono

sorgere dipendono da un lato dall'attribuzione di senso alle parole nuove da parte degli allievi, dall'altro da condizioni cognitive innate. Secondo i docenti la comprensione orale sembra progredire piuttosto linearmente, mentre la produzione orale sembrerebbe più altalenante. I docenti si sentono abbastanza liberi nella programmazione iniziale ed essa può dunque subire variazioni in corso d'opera se ad esempio ciò è suggerito dai colleghi, oppure se si osservano cambiamenti radicali nell'andamento o nella motivazione della classe. Quest'ultimo caso è quello che si verifica più comunemente: i docenti modificano il proprio programma quando è a rischio la motivazione. Meno prioritarie sembrano le alternanze nelle progressioni, in primo luogo perché a loro dire la maggioranza degli errori, considerati delle "insicurezze", fanno comunque parte del processo di apprendimento e si risolvono senza adattare troppo le lezioni. In fondo, e qui si presenta una seconda giustificazione, ufficialmente "una vera padronanza" non è neanche richiesta nella scuola elementare. La terza ragione di questo agire risiede nelle limitazioni temporali dettate dal programma cantonale.

Quando – durante l'intervista – i docenti hanno dovuto spiegare come convertono i principi della valutazione formativa nella pratica, sono sorti diversi dubbi sul valore socio-pedagogico della valutazione in sé e degli strumenti utilizzati o previsti a tale scopo. Il discorso sugli obiettivi minimi raggiungibili da tutti gli allievi presuppone che i docenti decidano cosa insegnare e come, e di riflesso cosa valutare e come – anche perché il manuale non offre indicazioni complete in merito. Su come valutare le competenze degli allievi, in particolare alla fine di un quadrimestre o dell'anno, alcuni docenti si attengono maggiormente ad una norma di gruppo e altri invece adottano piuttosto dei criteri non dipendenti dal livello della classe. Alla fine però, pochissimi docenti si dicono sicuri del sistema adottato. Troviamo in ogni caso docenti che associano in modo indissolubile la valutazione formativa al *feedback*. Il *feedback* dev'essere immediato per essere efficace, qualunque caratteristica esso abbia (correttivo, esplicativo, relativo al compito o alla competenza). Il metodo considerato più efficace e motivante è quello individuale,

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

orale e interattivo, in quanto permette agli allievi di chiedere chiarimenti, ed entrambi (docente e allievi) possono accertarsi del senso degli enunciati, sviluppare delle riflessioni e avviare una modifica comportamentale.

5.3. Valutazione informatizzata: ipotesi sul valore aggiunto e sul lavoro in classe

Per introdurre l'eventuale uso futuro di una piattaforma informatizzata nella valutazione delle competenze degli allievi in lingua francese, ci si sofferma brevemente sulle competenze informatiche di allievi e docenti. I docenti affermano di possedere delle competenze informatiche generalmente buone o soddisfacenti nell'uso di programmi di base, ma non le usano nell'insegnamento del francese. Due docenti seguono, con le loro classi, un progetto ticinese di insegnamento della matematica tramite supporto informatizzato chiamato *CabriElem*, e possiedono dunque delle competenze specifiche per questo programma e un'aula attrezzata (adiacente) per svolgere le attività. Tra gli allievi, taluni raccontano in classe di saper usare dei programmi informatici, altri invece non dimostrano interesse. I docenti si rendono conto che dovrebbero informarsi meglio almeno su come gli allievi usano il CD-ROM di *AZ* per esercitarsi a casa.

Partendo dall'esempio delle interfacce cartacee che i docenti hanno potuto vedere e "interrogare", i docenti hanno sviluppato alcune ipotesi d'uso. I docenti nel complesso non sentono un urgente bisogno di un dispositivo informatizzato per la valutazione diagnostica e formativa delle competenze in francese. Alcuni docenti eventualmente utilizzerebbero gli strumenti (così come sono) perché hanno l'infrastruttura e per trarne dei bilanci sulle competenze della classe e di singoli allievi; ammesso che ricevano spiegazioni sul funzionamento del ciclo di valutazione completo. Altri docenti lo utilizzerebbero se avessero l'infrastruttura per farlo, ma non tutti userebbero il PC in classe, né per l'insegnamento, né per la valutazione.

Nel complesso gli item e compiti per gli allievi sono comunque considerati adatti ad un pubblico giovane (vedi il capitolo II); sono semplici, esposti uno per pagina e chiari nella consegna e sembrano accessibili anche agli allievi che non hanno una conoscenza informatica specifica. Questo vale in particolare per gli item che riguardano le competenze linguistiche, compresi gli item sulla *langue de classe*; mentre l'item sulle strategie ha generato più perplessità sia per il contenuto che per il formato della domanda. Le insicurezze sulle strategie d'apprendimento risalgono, a loro dire, alla conoscenza delle strategie in sé, nonostante ciò, vedendo l'esempio, la maggioranza degli intervistati non saprebbe come sfruttarlo.

Un discorso simile è stato notato anche per ulteriori aspetti: la presentazione dei risultati organizzati nell'interfaccia in base ai seguenti filtri: 1. allievo, 2. item e 3. competenza, sembra molto complicata da gestire e impegnativa a livello di tempo. Quando poi i docenti capiscono che, secondo questo schema, dovrebbero cominciare la valutazione riascoltando le risposte degli allievi e analizzarle in base a criteri dati, per poi giungere a dei risultati inglobati in schemi psicometrici, le perplessità aumentano notevolmente.

Il tempo a disposizione per lo svolgimento del programma di francese e dunque per la valutazione delle competenze dei singoli allievi è spesso ritenuto troppo ridotto. Il rilevamento, l'analisi e il *feedback* individualizzato della produzione orale richiedono sicuramente tempo e una buona capacità di gestione della classe (i docenti sottolineano che la classe comprendono tra i 20 e i 24 allievi). Pertanto i docenti vogliono focalizzare le attività di valutazione su ciò che pare loro prioritario. A questo punto la questione del tempo e delle conseguenti priorità si ripropone anche in vista di un dispositivo informatizzato; essi si domandano dove trovare ulteriore tempo per svolgere il ciclo completo della valutazione informatizzata. C'è chi sostiene che forse con l'applicazione effettiva e con la crescente dimestichezza nell'uso degli strumenti, la problematica potrebbe diminuire. D'altra parte, rimane aperta un'altra questione: tra i docenti c'è chi si interroga anche su come svolgere la fase di rilevamento, non solo quando manca l'in-

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

frastruttura necessaria, quale un minimo di 5-7 PC con relativi microfoni e auricolari, ma anche quando questa fosse presente, ma situata in un'altra aula distante dalla propria. Questa disposizione implicherebbe di sdoppiarsi per seguire gruppi separati di allievi, in quanto gli allievi non possono essere soli in classe, tanto meno per svolgere una valutazione. Questa problematica si collega inoltre ad una successiva questione sollevata anche da chi i computer li ha in classe: siccome gli allievi dovranno rispondere a voce alta alle domande, essi non solo si possono ascoltare, ma anche intralciare reciprocamente e, non da ultimo, interferire con altri allievi che allo stesso momento svolgono altre attività e che ancora dovranno svolgere i compiti.

Tra i docenti c'è chi propone delle ipotesi di lavoro: forse basterebbe che gli allievi si esercitino nella produzione orale e nelle strategie grazie a questi item ben fatti, a patto che i contenuti rispecchino i temi presenti in *AZ*. Un docente prende spunto dall'esperienza con *CabriElem*, grazie al quale può osservare seduta stante (oppure dopo la lezione) le esercitazioni di matematica svolte dagli allievi e può risalire ai procedimenti utilizzati per risolvere le incognite o rielaborare i risultati (da solo o in coppia, presupponendo quindi un processo di *peer-feedback*). In questo modo non solo i risultati ma anche i processi sarebbero trasparenti sia per il docente che per gli allievi e permetterebbero di sviluppare diverse strategie. Sarebbe interessante, secondo questo docente ed altri colleghi, se gli allievi in ambito della produzione orale potessero anche solo registrarsi e ascoltarsi ed eventualmente registrarsi e riascoltarsi di nuovo. Questo sistema stimolerebbe gli allievi all'ascolto e alla valutazione attiva (eventualmente con l'ausilio di un modello corretto) delle proprie *performance*. Alla domanda posta ai docenti su quale potrebbe essere il *feedback* che lo strumento informatizzato fornisce agli allievi, i docenti propongono come riferimento il *feedback* prodotto dal CD-ROM per le esercitazioni di *AZ*, perché già conosciuto dagli allievi. Si tratta di simboli grafici e di un *feedback* orale breve in francese, tipo "bravo, tu as réussi, parfait, c'est juste", ecc. In questo modo gli allievi potrebbero riconoscere facilmente il *feedback* risultante dalla loro

performance. Tale supporto, secondo un docente, lo farebbe diventare uno strumento innovativo. Altri docenti menzionano la possibilità per le classi, per le scuole e per i docenti di confrontarsi sui risultati, il che potrebbe servire su larga scala anche a livelli dipartimentali. Un'altra idea consiste nel preparare questo strumento per l'uso alle scuole medie, credendo che possa essere più utile a docenti specialisti. Questi ultimi possiedono maggiori competenze disciplinari e insegnano in molte classi con tanti allievi. Il dispositivo potrebbe aiutarli a monitorare le competenze e le progressioni di ciascuno. Infine, c'è chi vede l'utilità di una piattaforma informatizzata con compiti, criteri e analisi sistemiche principalmente per il lavoro di ricercatori, formatori e studiosi del settore.

6. Discussione e conclusione

L'indagine presenta una panoramica qualitativa di alcune pratiche e opinioni presenti sul territorio. Le risposte riportate in modo sintetico nei risultati sono state selezionate perché ritenute dati pertinenti ad una riflessione per la (ri)costruzione degli strumenti e della piattaforma. Tali dati vengono qui di seguito anche correlati e discussi in relazione ai concetti approfonditi nel quadro teorico, agli orientamenti pedagogici del Programma cantonale, del manuale *AZ* e del progetto, cercando di creare una visione d'insieme ma anche di sviluppare considerazioni specifiche.

L'analisi dei dati ha permesso di capire che i docenti intervistati fanno riferimento agli orientamenti del *Programma* di francese ancora in vigore, come pure ai principi dell'*assessment for learning* (Assessment Reform Group, 2002) che ispirano il progetto. Ciò è visibile nell'integrazione costante della valutazione nell'insegnamento, nell'approccio trasparente con gli allievi, nella condivisione e negoziazione in classe di alcuni aspetti legati alla valutazione, nel valore attribuito al *feedback* e alla promozione dell'autonomia degli allievi. Il campione di docenti intervistato si rapporta in modo abbastanza stretto anche ai libri di testo (di *AZ*). In particolare per la valutazione usano e integrano i test

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

semplici proposti dalla GP. Questo permette di pensare che gli item semplici proposti nel modello di interfaccia del dispositivo informatizzato siano apprezzati dai docenti in quanto simili ai formati e ai contenuti con cui si esercitano gli allievi. Ciò suggerisce al gruppo di ricercatori di “Profili di sviluppo” di continuare a creare degli item legati a contenuti e formati conosciuti e di mantenere la loro semplicità e chiarezza.

Considerando che la valutazione nella loro pratica è supportata da compiti specifici per la comprensione, produzione e interazione orale, ma non da una registrazione delle performance, le funzioni del dispositivo per la registrazione e l'archiviazione sistematica delle risposte potrebbero essere considerate utili. La possibilità di poter poi usufruire di dati sulla produzione orale sia puntuali, sia longitudinali interessa coloro che potrebbero andare oltre una valutazione di tipo “interpretativo e mnemonico”, ossia docenti che desiderano ricalibrare la propria programmazione in modo *evidence-based*, e insegnanti che devono fornire un riscontro ai genitori che non condividono sempre la loro valutazione. Per quanto concerne il progetto “Profili di sviluppo”, le conoscenze e competenze sollecitate con gli item (elementi minimi, lessico, funzioni e comunicazione-interazione) e i criteri individuati per la valutazione informatizzata (la correttezza fonetica, l'appropriatezza lessicale, di contenuto e funzionale), paiono simili a quelli usati complessivamente dai docenti intervistati. Questo vale anche per i criteri che definiscono la progressione: l'evoluzione dell'ampiezza lessicale, l'articolazione delle frasi, del grado di correttezza richiesta. Questo risultato si può ritenere positivo in quanto rispecchia un certo grado di compatibilità e, di conseguenza, di applicabilità degli strumenti informatizzati.

Nel complesso, se da una parte c'è interesse (altrimenti i docenti non avrebbero partecipato al progetto), dall'altra i docenti sollevano una serie di punti critici, legati alle loro conoscenze e competenze attuali, agli strumenti stessi e alla rilevanza di questi ultimi con la pratica.

Tra i fattori critici, si è rilevata la logistica, per la quale attualmente non si riesce a trovare una soluzione, visto che raramente le classi ticinesi

possiedono un'aula adiacente adatta per lo svolgimento del rilevamento della produzione orale. Inoltre è importante anche la questione temporale, la quale, già complessa nella didassi quotidiana, parrebbe acuirsi con l'ausilio di uno strumento così impegnativo. Chi ha sperimentato forme nuove di rilevamento e valutazione, riflette anche sulle proprie competenze in materia di gestione della classe, portando alla nostra attenzione ciò che Ur (1999) definisce “competenze di *testing* e di *rater*”. I docenti sono dunque consapevoli che le nuove proposte, malgrado tutte le buone intenzioni, saranno usate poco e poi abbandonate. Inoltre, poiché i docenti esprimono una generale necessità di semplificazione del sistema di ascolto e di analisi delle risposte, sarebbe indispensabile per i ricercatori continuare ad approfondire la questione per raggiungere un'applicabilità reale. D'altra parte è necessario considerare che, anche quando i docenti un giorno potranno adoperare efficientemente gli strumenti e utilizzarli durante le lezioni (*feedback* compreso), potrebbe subentrare comunque il rischio che si soffermino sull'errore (nella valutazione diagnostica) in quanto deficit e non sugli indicatori di progressione (Torrance & Pryor, 2004). Considerando l'aderenza del gruppo di docenti ad una trattazione formativa dell'errore, è necessario pensare a formulare *feedback* (sia per i docenti che per gli allievi), in cui l'errore abbia chiaramente un carattere costruttivo e nel quale la correttezza formale sia inserita in un processo di sviluppo delle competenze. Si dovrebbe poter accompagnare i docenti nell'uso efficiente dei *feedback* e incoraggiare questo *washback effect*, evitando così un uso diverso degli strumenti e dei dati (Ur, 1999).

Se sulla valutazione della lingua straniera i docenti generalisti in Ticino possono trarre pochi riferimenti didattico-pedagogici dal manuale di AZ, sarebbe utile pensare ad uno strumento che offra delle istruzioni ai docenti e delle considerazioni di ordine pedagogico-didattico nonché degli esempi di *feedback* basati sull'analisi dei dati. Un *feedback* “descrittivo” generato da un dispositivo informatizzato dovrebbe in ogni caso essere programmato ad hoc. Dalle caratterizzazioni di *feedback* (di Hill & McNamara, 2012) esposte sopra, lo strumento

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

potrebbe, o meglio, dovrebbe per definirsi diagnostico (Alderson, 2005) e in supporto alla valutazione formativa, offrire diverse tipologie di *feedback*, tra cui quello di tipo esplicativo, descrittivo e riferito ad un *task*; rispecchiando la pratica dei docenti. Secondo Alderson (*ibid.*) il *feedback* informatizzato è possibile, anzi indispensabile, ma c'è anche da considerare che il progetto *Dialang* di cui Alderson (*ibid.*) è responsabile, è indirizzato ad un pubblico adulto. Le variabili per gli allievi e il ruolo dell'insegnante di scuola elementare portano invece a pensare che il sistema sarebbe più adatto se il *feedback* fosse dato principalmente dal docente, almeno fino a quando determinate competenze e usanze nella valutazione non siano state elaborate, negoziate e assicurate con la classe. Probabilmente sarebbe anche utile poter avere una modalità in cui accendere o spegnere l'ausilio del *feedback*, visto che alcuni studiosi sostengono che potrebbe interrompere i processi in atto per la produzione orale. Infine, se i *feedback* fossero sempre gli stessi, ciò aiuterebbe anche gli allievi a orientarsi con più profitto, garantendo la trasparenza dei processi in corso (ad es. come si è svolta la performance).

L'indagine qualitativa ha messo in risalto perché il valore aggiunto di una piattaforma informatizzata sia ancora a livello di stima, e pertanto in discussione. Secondo i dati attuali, sembrerebbe che il suo valore e il suo ruolo nella prassi dipendano anche dall'importanza percepita dai docenti ad avere una serie di dati concreti sullo sviluppo dell'oralità e delle competenze strategiche. Gli strumenti dovrebbero rappresentare un sostegno per i docenti, ma questi ultimi devono essere consapevoli del fatto che per un uso efficiente della piattaforma bisogna formare i discenti per utilizzare il *feedback* in modo corretto. Dai dati dell'indagine si nota che oltre ai chiarimenti sulla piattaforma o ai corsi di formazione possibili su *assessment* e *rating* per i docenti, rimangono da affrontare le problematiche inerenti il tempo e la logistica. Si suggerisce di sviluppare ulteriori ricerche, utilizzando strumenti più vicini alla metodologia della valutazione nella classe (*classroom-based assessment*) per rilevare l'impatto degli strumenti su docenti e allievi e la sua reale applicabilità nel contesto classe (di lingua straniera).

Infine, considerando anche le differenze tra i manuali per l'insegnamento delle lingue straniere in uso nelle diverse regioni svizzere, rimane ancora da approfondire la possibilità di armonizzare o adattare i compiti per più regioni, sia per la produzione linguistica che per lo sviluppo delle strategie. Quando si desidera creare degli strumenti che si possano usare anche in Ticino, i contenuti e i formati dovrebbero essere accessibili anche agli allievi ticinesi per permetter loro di farne buon uso. Infine, il legame con il contesto e con gli obiettivi e orientamenti dei piani di studio regionali, ai quali i docenti e di riflesso anche gli allievi sono abbastanza ancorati, implica diversi adattamenti del sistema informatizzato di valutazione.

7. Bibliografia

- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: the interface between learning and assessment*. London; New York: Continuum.
- Assessment Reform Group (2002). *Assessment for Learning: 10 principles. Research-based principles to guide classroom practice*. Recuperato da http://www.hkeaa.edu.hk/DocLibrary/SBA/HKDSE/Eng_DVD/doc/Afl_principles.pdf [30.10.18]
- Beacco, J.-C., Byram, M., Cavalli, M., Coste, D., Egli Cuenat, M., Goullier, F. & Panthier, J. (2011). *Guida per lo sviluppo e l'attuazione di curricoli per una educazione plurilingue e pluriculturale*. Strasbourg: Conseil de l'Europe; Milano: Università degli Studi. Recuperato da <https://rm.coe.int/guida-per-lo-sviluppo-e-l-attuazione-di-curricoli-per-una-educazione-p/16805a028d> [31.03.16]
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31.

VI Pratiche di valutazione e feedback nell'insegnamento della lingua straniera alla scuola elementare

- Black, P., Wilson, M. & Yao, S.-Y. (2011). Road maps for learning: a guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspectives*, 9(2-3), 71-123.
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3-12.
- DECS-DS (2004). *Programma sperimentale di francese per la scuola elementare*. Recuperato da http://www.supsi.ch/dfa/dms/dfa/docs/dipartimento/20110203_Programma_sperimentale_francese.pdf [30.10.18]
- DECS-DS (a c. di) (2015). *Piano di studio della scuola dell'obbligo*. Recuperato da <http://www.pianodistudio.ch/> [09.03.16]
- Guidicini, P. & Castrignano, M. (1997). *L'utilizzo del dato qualitativo nella ricerca sociologica*. Milano: Franco Angeli.
- Hill, K. & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395-420.
- NCH Software (s.d.). *Express Scribe*. Greenwood: NCH Software.
- Tarozzi, M. (2008). *Che cos'è la grounded theory*. Roma: Carocci.
- Torrance, H. & Pryor, J. (2004). Investigating formative classroom assessment. In: L. Poulson & M. Wallace (a c. di), *Learning to read critically in teaching and learning*. London: SAGE, 129-150.
- Trincherò, R. (2002). *Manuale di ricerca educativa*. Milano: Franco Angeli.
- Ur, P. (1999). *A course in language teaching. Trainee book*. Cambridge: Cambridge University Press.

