

BABYLONIA

2|2016



*Rivista per l'insegnamento e
l'apprendimento delle lingue*

*Zeitschrift für Sprachenunterricht
und Sprachenlernen*

*Revue pour l'enseignement et
l'apprentissage des langues*

*Revista per instruir ed emprender
linguas*

*A Journal of Language Teaching and
Learning*



WWW.BABYLONIA.CH

Evaluation

Valutazione

Évaluation

Evaluaziun

idT

FRIBOURG
FREIBURG
2017

BRÜCKEN GESTALTEN –
MIT DEUTSCH VERBINDEN



XVI. Internationale Tagung der Deutschlehrerinnen und Deutschlehrer

31. Juli - 4. August 2017

Freiburg, Schweiz

**Seien Sie dabei!
Melden Sie sich
frühzeitig an.**

www.idt-2017.ch

EINREICHUNG DER ABSTRACTS

Alle Informationen für die
Einreichung eines Beitrags finden
Sie auf der Webseite.

Early-bird Einschreibungen und
Anmeldung von Beiträgen

bis 30. September 2016

TAGUNGSGEBÜHREN

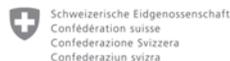
(Abendempfang inbegriffen)

- Ganze Tagung: CHF 440.- / 240.-*
- Dienstag-Freitag: CHF 360.- / 220.-*
- Nur Montag: CHF 200.- / 100.-*

* Studierende / Doktorierende

DIE IDT 2017 AUF EINEN BLICK

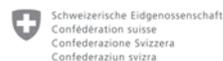
- 5 Kongresstage
- Teilnehmende aus über
100 Ländern
- Reiches fachliches und
fachwissenschaftliches Angebot
- Mehr als 20 Hauptvorträge
- 36 Sektionen
- Internationale Aussteller
- Vielfältiges Rahmenprogramm
- Ausflugsmöglichkeiten



Eidgenössisches Departement des Innern EDI
Bundesamt für Kultur BAK



Eidgenössisches Departement für
Wirtschaft, Bildung und Forschung WBF
Staatssekretariat für Bildung,
Forschung und Innovation SBF



Eidgenössisches Justiz- und Polizeidepartement EJPD
Staatssekretariat für Migration SEM

Evaluation
Valutazione
Évaluation
Evaluaziun

Responsabili della parte tematica:
Laura Loder Büchel & Mathias Picenoni

Babylonia
La rivista per l'insegnamento
delle lingue

Trimestrale plurilingue
edito dalla
Fondazione Lingue e Culture
cp 120, CH-6949 Comano
ISSN 1420-0007
no 2/anno XXV/2016

Con il sostegno di

- Dipartimento Educazione, Cultura e Sport del Canton Ticino



- Ambassade de France en Suisse



AMBASSADE DE FRANCE EN SUISSE

- Ufficio Federale della Cultura



- Fondazione Oertli



- Segreteria di Stato della migrazione



- Dipartimento dell'educazione, cultura e protezione dell'ambiente dei Grigioni



Erziehungs-, Kultur- und Umweltschutzdepartement Graubünden
Departament d'educaziun, cultura e protecziun da l'ambient dal Grischun
Dipartimento dell'educazione, cultura e protezione dell'ambiente dei Grigioni

4

Editoriale della redazione

6

Introduction | Einleitung

Laura Loder Büchel

10

Von
Fremdsprachenlernen,
Lehrplan 21 und
Beurteilung

Christian Amsler

13

Leseverstehen
überprüfen: Welche
Sprache für die Fragen
und Antworten in den
Testitems?

Malgorzata Barras, Katharina
Karges & Peter Lenz

19

Pour une évaluation sous
contrôle didactique

Jean-François de Pietro &
Murielle Roth

24

Analisi di prove di ascolto
in esami di certificazione
italiani (livello B2)

Paolo Torresan

30

Descrittori letterari
e acquisizione di
competenze: strumenti,
obiettivi e passione!

Eleonora Rothenberger Barbaro

34

Worauf die Profis von
Hörverständnisübungen
achten! Checkliste

Alberta Novello
Überstzung von Mathias
Picenoni

36

Nützliche(re)
Sprachtests

Thomas Studer

44

More than testing the water
Beurteilung von
Englischkompetenzen im
Verlauf der Volksschule

Daniel Stotz

50

L'épreuve cantonale de
référence d'allemand:
une nouveauté en Pays de
Vaud

Claudia Bartholemy, Olivier
Bolomey & Pascal Schweitzer

53

Early reading and writing
in the foreign language
classroom – maybe not
as controversial as one
might think

Jutta Rymarczyk

58

Assessing Modern
Languages in Swedish
Upper Secondary Schools

Per Snoder

63

Interkulturelle Kompe-
tenzen testen – Gra-
duierungskriterien und
Beurteilungsraster für
den Fremdsprachenun-
terricht auf der Sekun-
darstufe I

Jan-Olivier Eberhardt

69

Un test a livello
nazionale di produzione
scritta: vengono
raggiunti gli obiettivi a
fine del secondo ciclo
della scuola secondaria?

Anne Beeker-Schoenmakers,
Daniela Fasoglio, Kim de Jong, Jos
Keuning & Alma van Til

BABYLONIA

2|2016

Didattica

Bloc-notes

73

Testvalidität im Unterricht: eine Checkliste für reflektierte(re) Beurteilungen

Katharina Karges

76

Lassoing in languages at Manpower

An interview with Romain Hofer

L'intervista

79

Transition from primary to secondary language learning: using an Assessment for Learning approach to unpack pupils' 'suitcases of skills and abilities'

Janes Jones

84

App Review: The Benefits Of Using ClassDojo For Assessment Data Collection In The Foreign Language Classroom

Caleb Howard

86

Standards-based Grading Tools

Laura Loder Büchel

88

Integrated Performance Assessment in English: How to assess and develop several aspects of a debate on animal testing

Esmerelda Meili

90

Oldies but goodies: Bloom's taxonomy

Laura Loder Büchel

94

Using the Oxford Young Learners Placement Test to Keep an Eye on Progress

Denise Kündig

97

Student generated rubrics to guide and assess

Patrick Büchel

100

Unterrichtsvorbereitung

Andreas Pfister

Il racconto

102

4 Sprachen zum Dessert

SRF mySchool

Finestra

104

Italiano e romancio in Svizzera

106

Recensioni

108

l'Opinione

109

Info

110

Agenda & impressum



Auszuschöpfendes Potenzial

Den Handschlag aus religiösen Gründen verweigern - der Einzelfall im Baselbiet sorgt für eine Welle der Empörung in den Medien und Leserkommentaren.

Die NZZ greift das Thema aus einer anderen Perspektive auf mit der Frage, ob die Schule nicht zu schweizerisch sei, also ob nicht mehr Lehrpersonal mit Migrationshintergrund helfen würde, den verschiedenen Kulturen in der Schule offen und angstfrei zu begegnen. Sie verweist dabei auf die Empfehlung der EDK, „die Heterogenität, 'das Exotische' in die Schulen zu holen“ (3.6.2016).

Die Wortwahl mag erstaunen, doch hebt sie positiv hervor, dass im unterschiedlichen kulturellen Hintergrund und in der aus der Migration gewonnenen Erfahrung ein Vorteil erkannt wird, ein *Potenzial*, das den Schulalltag verbessern und bereichern kann.

Ins gleiche Horn bläst die Erziehungswissenschaftlerin Margrit Stamm, die in den Migranten die Bevölkerungsgruppe „mit den wohl am häufigsten unentdeckten Begabungsreserven“ entdeckt, was daran liege, dass „der Blick auf sie ein fast durchgängig negativer“ sei (*Terra Cognita* 27, S. 26). Demnach könne das Leistungspotenzial von Jugendlichen mit Migrationshintergrund oder unterer Sozialschichten gar nicht zum Tragen kommen, weil man das Potenzial bei ihnen gar nicht erst erwarte, es infolgedessen nicht erkannt und auch nicht gefördert werde.

Die gleiche Ansicht scheint auch der Schweizer Lehrerverband (LCH) zu vertreten, wenn er bei der Durchführung des „Schweizer Pisa-Tests“ – der Überprüfung des Erreichens der Grundkompetenzen (ÜGK) – dem Datenschutz höchste Priorität gibt, auf dass Rankings nicht gerade jene Schulen und Schülergruppen treffen, die „benachteiligt“ sind. Benachteiligt? Widerspiegeln die Tests das soziale Gefüge an Schweizer Schulen? Lassen sie sich nicht vielmehr so gestalten, dass sie nicht nur nicht die Defizite, sondern gleichsam das Potenzial jener Schülerinnen und Schüler aufzeigen, bei denen es gar nicht erst erwartet wird?

Die vorliegende Nummer diskutiert aus wissenschaftlicher und praktischer Perspektive, was valide Tests ausmachen, ob professionelle Tests gut kalibriert sind und wie Tests das 'Exotische' – interkulturelle, mehrsprachige und literarische Kompetenzen – notenwirksam valorisieren können.

Gute Lektüre!

Ihr Redaktionsteam



Potenziale nascosto

Rifiutarsi di stringere la mano alla propria insegnante per motivi religiosi – il caso di Therwil scatena indignazione nei mass media e nei blog.

Contribuendo all'accesso dibattito, la NZZ cambia prospettiva: che non sia la scuola a essere troppo svizzera? Forse una maggiore presenza di insegnanti provenienti da un contesto migratorio contribuirebbe a gestire con maggiore apertura e serenità la realtà multiculturale nelle classi? La testata svizzero-tedesca rinvia alla raccomandazione della CDPE di “portare a scuola l'eterogeneità, l'esotico” (3.6.2016). Per quanto possa stupire quest'ultima formulazione, essa evidenzia il potenziale che diversità culturale e esperienze di migrazione possono avere in ambito scolastico, a vantaggio di un migliore e più ricco vissuto quotidiano.

Margrit Stamm, esperta in scienze dell'educazione, giunge a una conclusione simile: la connotazione negativa attribuita ai migranti impedisce di riconoscerne i talenti nascosti (*Terra Cognita* 27, p. 26). Ai giovani cresciuti in contesti svantaggiati, sia migranti sia appartenenti a ceti sociali bassi, non viene data la possibilità di dimostrare il proprio potenziale – poiché nessuno se lo aspetta, non viene né riconosciuto né valorizzato. Anche l'associazione svizzera degli insegnanti (LCH) sembra assumere una posizione analoga quando chiede che i dati rilevati tramite i “test PISA svizzeri” (volti alla “Verifica delle competenze fondamentali” VCF) siano trattati con la massima discrezione, onde evitare che classifiche possano andare a discapito di sedi scolastiche e scolari già di per sé “svantaggiati”. Svantaggiati? Tali test sono in grado di considerare il potenziale dell'eterogeneità sociale presente nelle scuole svizzere? Non potrebbero essere calibrati in modo tale da non evidenziare solo i deficit, ma appunto le risorse di quegli scolari da cui sembra che nessuno se le aspetti? Il presente numero discute, dalla prospettiva della ricerca scientifica e della prassi, quali fattori rendono valido un test, se i test sviluppati professionalmente sono ben calibrati e come si possa valutare l'“esotico”, tra cui le competenze interculturali, plurilingui e letterarie. Buona lettura!



Il vostro team redazionale

Un potentiel à exploiter

Un élève de Therwil, dans le canton de Bâle-Campagne, a refusé de serrer la main à son enseignante pour des motifs religieux – l'incident a soulevé l'indignation dans les médias et les blogs. La NZZ, cependant, est intervenue dans le débat en prenant une perspective différente et en nous interrogeant: l'école n'est-elle pas «trop» suisse? Une présence plus forte d'enseignant-e-s issu-e-s de la migration ne contribuerait-elle pas à une gestion plus ouverte et sereine de la réalité multiculturelle des classes? Le journal se réfère à la recommandation de la CDIP «d'amener l'hétérogénéité, l'exotique, dans l'école». Cette dernière formulation peut surprendre, mais elle ne fait que souligner le potentiel d'amélioration et d'enrichissement du quotidien scolaire que recèlent la diversité culturelle et les diverses expériences de migration des élèves.

Margrit Stamm, chercheuse en sciences de l'éducation, parvient à une conclusion similaire: les connotations négatives attribuées aux migrants nous empêchent de reconnaître leurs talents cachés (*Terra Cognita* 27, p. 26). Pour celles et ceux qui viennent d'un autre pays ou d'une classe sociale «défavorisée», il n'est que trop rarement donné de démontrer leur propre potentiel, qui reste ignoré et peu valorisé car... il n'est pas attendu!

L'Association suisse des enseignant-e-s (LCH) semble prendre une position analogue lorsqu'elle demande que les données recueillies au moyen des «tests PISA suisses» (censés vérifier l'atteinte des compétences fondamentales définies dans le projet HarmoS) soient traitées avec un maximum de précautions et de discrétion, afin d'éviter que des classements hâtifs puissent conduire à discréditer et pénaliser encore des écoles et des élèves qui sont déjà «désavantagés». Désavantagés?... De tels tests sont-ils en mesure de prendre en compte l'hétérogénéité présente dans les classes? Ne pourraient-ils pas être conçus de manière à ne pas faire ressortir uniquement les déficits mais également les ressources dont disposent ces élèves dont on n'attend rien?...

Le présent numéro porte sur l'évaluation. Il discute, dans une perspective scientifique et pratique, les facteurs qui rendent un test valide, la qualité des tests professionnels... mais aussi comment l'exotique – les compétences interculturelles, plurilingues, littéraires... – pourrait être mieux mis en valeur.

Bonne lecture!

La rédaction

LO

Potenzial zuppà

Refusar da dar il maun a la magistra u al magister per motivs religius – il cas da Therwil a Basilea-Champagna procura per indignaziun en las medias e per bleras brevs da lecturs.

La NZZ contribuescha a la debatta d'ina autra perspectiva: ella sa dumonda, schebain la scola na saja betg memia svizra e sch'ina preschientscha pli gronda da persunas d'instrucziun cun in'experientscha migratorica na gidass betg a scuntrar las differentas culturas a moda pli averta e paschavla en scola. Ella renviescha a la recumandaziun da la CDEP/EDK da "purtar l'eterogenitad, l'exotic' en las stanzas da scola" (3-6-2016). Questa formulaziun po far smirvegliar, ella metta dentant en evidenza ils avantatgs che diversitads culturalas ed experientschas migratoricas pon avoir, in potenzial che permetta da meglierar ed enritgir il mintgadi en scola.

Dal medem avis è la scenziada d'educaziun Margrit Stamm: la connotaziun negativa attribuida a las migrantas ed als migrants impedescha da renconuscher ils talents zuppads (*Terra Cognita* 27, p. 26). Il potenzial da prestaziun da giuvenils cun biografias da migraziun u da gruppas socialas inferiuras na possia pia betg sa manifestar, perquai ch'ins na spetgia dad els gnanc in tal potenzial che na vegnia consequentamain betg renconuschì e pia era betg promovì.

In'opiniun sumeglianta para d'avair era l'Associazion svizra da magistras e magisters (MCH/LCH): ella pretenda numnadamain ch'ils resultats dals "tests da Pisa svizzers" (verificaziun da las cumpetenzas da basa) vegnian tractads cun discreziun maximala, per ch'ils rankings na gajan betg a disfavur da scolas e gruppas da scolaras e scolaras gia "dischavantagiadas". Dischavantagiadas? Reflecteschan ils tests la structura sociala en las scolas svizras? Na pudessan els betg vegnir concepids da maniera a revelar betg sulettamain ils deficits, mabain era il potenzial da quellas scolaras e da quels scolaras, dals quals in tal potenzial na vegn insumma betg spetgà?

Il numer preschent discuta d'ina optica scientifica e pratica quals facturs che rendan tests valavels e sche tests sviluppads a moda professiunala en bain calibrads e co ch'els permettian da valorisar l'exotic', tranter il qual las cumpetenzas interculturalas, plurilinguas e litteraras.

Buna lectura!

Voss team da redacziun

NIA

INTRODUCTION

EINLEITUNG

● Laura Loder Büchel | Zürich

Laura Loder Büchel is a lecturer at Zurich University of Teacher Education and a member of Babylonia's editorial staff.



I am constantly struggling to make meaning of how my own children are being assessed in Swiss primary and secondary schools and I am sure I am any teacher's nightmare when I come with a question about such practices. How can an American who only learned her first foreign language at the age of 20 even possibly fathom the depth and breadth of foreign language assessment in the multilingual Swiss context? There are many things going right when my Swiss-German speaking children come home motivated to learn, to study, to use their French when we visit Geneva, to scream a French poem at me from another room, to add a good piece to their portfolio, to rewrite something for their portfolio. So how can I possibly even consider the following dark thoughts (see Table 1 on the next page) when so much is right? Luckily, I am not the only one having these thoughts; the mass of literature found will attest that is a hot topic with mismatches between what teachers say they are doing in terms of assessment and what they are actually doing (Pat-El, Tillema, Segers & Vedder, 2015). There

Täglich stelle ich mir tausend Fragen darüber, wie meine Kinder in der Primar- und Sekundarstufe in den Fremdsprachen Englisch und Französisch - und auch in der Schulsprache Deutsch - beurteilt werden. Wie kann eine Amerikanerin, die ihre erste Fremdsprache erst im Alter von 20 gelernt hat, die vielfältigen Dimensionen von Beurteilungen im mehrsprachigen Schweizer Kontext nachvollziehen? Vieles läuft gut, wenn meine deutschsprachigen Kinder motiviert von der Schule nach Hause kommen, um die „Ufzgi“ ohne Jammern zu machen, oder wenn sie versuchen, bei einem Besuch in Genf ihr Französisch zu sprechen, wenn sie lustige Gedichte auf Französisch laut vorsagen, wenn sie stolz etwas Neues im Portfolio eintragen und ihre Texte mit Freude bearbeiten. Also: Wie kann es sein, dass ich immer wieder finstere Gedanken habe (vgl. *Table 1* auf der nächsten Seite), auch wenn so vieles richtig läuft? Glücklicherweise bin ich überhaupt nicht alleine mit meinen finsternen Gedanken: Eine umfassende Literatur bestätigt, dass Beurteilen und Bewerten ein echtes Spannungsfeld ist, mit Diskrepanzen

Table 7

Dark Thoughts....		Finstere Gedanken	
Mommy me	Lecturer me	die Mutter in mir	die Dozentin in mir
Why were points taken off on my daughter's English test when she spelled a German word incorrectly?	Why are children even being given vocabulary translation tests when most kids will certainly not become translators/interpreters? What is actually being measured through these tests?	Warum hat meine Tochter weniger Punkte für einen Englischtest erhalten nur wegen ihren deutschen Rechtschreibfehlern?	Warum machen Lernende Übertzungstests, obschon nur die wenigsten Übersetzer werden? Was sind die Konstrukte hinter Übertzungstests?
Why do they have to learn old and even incorrect words for things like professions (scheduler? warehouseman?)? Why is the writing test on the use of the present perfect and not on some specific text type (like an informal letter)? What else can they do in the foreign language?	Are schools actually measuring the skills that learners will need in the future? How can we better teach functional literacy skills in English that might even transfer to general skills (writing letters, giving instructions, etc...)	Warum lernen meine Kinder veraltete Wörter wie "warehouseman"? Warum konzentrieren sich Schreibtests auf Grammatik (wie "Present Perfect") und nicht auf Textsorten (wie Einladungen) oder Satzstrukturen? Was können sie sonst noch in der Fremdsprache?	Beurteilen Schulen die nötigen Fertigkeiten, welche Lernende in Zukunft und im täglichen Leben tatsächlich einsetzen werden? Wie können schriftliche Kompetenzen im Englischunterricht besser unterrichtet werden, damit sie vielleicht sogar übertragbar werden (Briefe schreiben, Anweisungen geben)?
Why did I receive a sneak peek at my daughter's report card grades from the school's data collection tool and there was no separation of subskills (reading, writing, speaking, listening)?	How can one actually separate the four skills for testing purposes when even a gap fill measures reading and writing? How is interaction measured? How can teachers have good systems for recording development in each skill?	Kurz vor der Zeugnisabgabe gewährte mir die Lehrperson Einblick ins Notenraster meiner Tochter. Warum habt sie nur die Durchschnitte aller Prüfungen genommen und nicht die vier Fertigkeiten berücksichtigt?	Wie könnten die vier Fertigkeiten (Lesen, Schreiben, Sprechen, Zuhören) isoliert bewertet werden, zumal beispielsweise ein Lückentext sowohl Lese- wie Schreibkompetenz verlangt? Wie lässt sich „Interaktion“ bewerten? Wie könnten Lehrende die sprachliche Entwicklung einzelner Fertigkeiten dokumentieren?
When I asked the teacher what evidence s/he had collected about my daughters' performances, I got "here are her tests" and no other indications of having met any standards.	Where are the models for data collection (such as a grade being 20% average of 4 tests, 40% average of 8 performance assessments such as essays, 20% average of 20 pop quizzes...) and why are they not being used for grades?	Frage ich die Lehrerin nach Beispielen für die Leistungen meiner Tochter, erhalte ich als Antwort "Hier sind die Prüfungen" und keine anderen Hinweise darauf, ob sie Lernziele erreicht hat.	Was sind nützliche Modelle für evidenzbasierte Zeugnisnoten (z.B. dass 4 Prüfungen durchschnittlich 20% zählen, dass Leistungen wie Essays 40% ausmachen usw.) und warum werden sie nicht für die Beurteilung genutzt?
My daughter came home with a "True/False/Not in the text" test that even I got wrong.	Why are test-designers not aware of "True/False/Justify" and better test item design?	Meine Tochter zeigt mir eine "Richtig/Falsch/Keine Aussage" Englischprüfung, die selbst ich als Muttersprachlerin nicht richtig beantworten könnte.	Warum haben Lehrmittelentwickler oft beschränkte Repertoires von Test Items, die verwendet werden können? Wäre die Formel „Richtig/Falsch/Rechtfertigen“ nicht die bessere Alternative zu „Richtig/Falsch/Keine Aussage“?
Where are the stories, the arts, the intercultural skills, the strategic competencies that my daughters might have to learn about or to use? I mean, both are pretty good at talking with their hands!	How can we measure such skills when there is often no right or wrong? How can the ability to "negotiate meaning" and the development of strategies be assessed and reported?	Wo sind die Geschichten, die Kunst, die interkulturellen Fertigkeiten, die strategischen Kompetenzen, die meine Töchter lernen und einsetzen könnten? Schliesslich können sich beide gut mit Gesten ausdrücken!	Wie können wir Fertigkeiten evaluieren, für die es kein "gut" oder "unzureichend" gibt? Wie kann das „Aushandeln von Bedeutungen“ und die Entwicklung von Strategien beurteilt und bewertet werden?
Despite all the "hoopla" around "Kompetenzorientierung", has standards-based grading reached most classrooms – and should it?		Ist standardbezogene Bewertung an Schweizer Schulen schon angekommen, trotz des Tamtams rund um "Kompetenzorientierung"?	

The mass of literature found will attest that is a hot topic with mismatches between what teachers say they are doing in terms of assessment and what they are actually doing.

is a clear need for improvement and development and assessment-literacy is an often-overlooked aspect of teacher training as well as continuing practice (Popham, 2011). DeLuca, Chavez, Bellara & Cao (2013) found four explicit constructs that aid in teacher assessment literacy development:

- › perspective-building conversations,
- › praxis activities,
- › modeling, and
- › critical reflection and planning for learning.

These constructs are alive and kicking in this issue of *Babylonia* as we take a look at perspectives and models from the US (see *Howard on Class Dojo*), from England (see *Jones on Assessment for Learning*), from Sweden (see *Snoder on the new Swedish curriculum*) from Germany (see *Rymarczyk on the importance of writing*), from Holland (see *Fasoglio et al.*) and return back home to Swiss turf (see *Stotz on the Stellwerk and Multicheck tests*, *Bartholemy, Bolomey & Schweitzer on the new German language test in the canton of Vaud*, *Hofer on the importance of language at Manpower*) whilst travelling through the ever-changing turbulent landscape of grading and assessment policies and practices. This issue of *Babylonia* can be supportive in light of recent suggestions for new school report cards (as *Amsler* mentions) and a turn towards standards-based grading (see *Kuendig or Büchel or Meili for some concrete ideas*) and provides myriad ideas for rethinking and refining practices and promoting discussion. It can help classroom teachers to examine how to measure intercultural skills (see *Eberhardt*), test listening and reading comprehension (see *de Pietro & Roth on testing gist* or *Barras, Karges & Lenz on testing reading* or *Rothenberger on literature*), write valid tests (see *Studer on two useful test item types* or *Torresan on testing Italian language skills as well as Karges who provides a checklist for teachers when writing tests*) and it provides numerous praxis activities. It can encourage instructors to critically think about what messages are being passed on to future and in-service teachers.

So, enjoy the roller-coaster ride through Tema 1 (*assessing receptive skills*), Tema 2 (*testing in general*) and Tema 3 (*practical ideas*) and feel free to use the Reader's Guide questions to promote discussion.

zwischen dem, was Lehrkräfte zu machen scheinen, und dem, was tatsächlich gemacht wird (e.g. Pat-El, Tillema, Segers & Vedder, 2015). Es gibt Spielraum für Verbesserungen und Entwicklung bei den Beurteilungskompetenzen von Lehrpersonen, was leider oft in der Aus- und Weiterbildung übersehen wird (vgl. Popham, 2011). DeLuca, Chavez, Bellara & Cao (2013) definierten vier explizite Konstrukte, die bei der Entwicklung von Beurteilungskompetenzen helfen können:

1. Austausch zwischen Lehrpersonen über Beurteilungspraxis (*perspective-building conversations*);
2. Praktische Aufgaben (*praxis activities*);
3. Modellierung (*modeling*); und
4. Kritische Reflexion und Planung von Unterrichtseinheiten (*critical reflection and planning for learning*).

Diese vier Modelle sind in dieser Nummer gut vertreten, indem wir auf Perspektiven und Modelle aus den USA (*Howard zu Class Dojo*), aus England (*Jones zu Assessment für Learning*), Schweden (*Snoder zum neuen schwedischen Lehrplan*), Deutschland (*Rymarczyk zum Stellenwert des Schreibens*) und Holland (*Fasoglio et al.*) eingehen, um dann auf dieser Reise durch die sich ständig veränderten Landschaften der Beurteilungs- und Prüfungspraxis in die Schweiz zurückzukehren: *Stotz* berichtet über die Stellwerk- und Multicheck-Tests; *Bartholemy, Bolomey & Schweitzer* über die neue Deutschprüfung im Kanton Waadt und *Hofer* über Sprachen bei *Manpower*. In dieser Ausgabe von *Babylonia* finden sich Empfehlungen für neue Schulzeugnisnoten (*Amsler*), Vorschläge zu standardbasierter Beurteilung (vgl. *Kuendig oder Büchel oder Meili*) und eine Vielzahl weiterer Ideen, die dazu anregen, die Praxis zu überdenken und zu verfeinern. Klassenlehrpersonen werden von einer Vielfalt an Praxisideen profitieren, wie zum Beispiel zur Bewertung von interkulturellen Fertigkeiten (vgl. *Eberhardt*), zur Beurteilung von Lese- und Hörverständnis (siehe *de Pietro & Roth oder Rothenberger*) und zur Testvalidierung (vgl. *Studer, der zwei neue Testformate beschreibt, oder Torresan, zu Sprachkompetenztests für Italienisch und Karges, die eine Checkliste zur Testerstellung für Lehrende vorstellt.*). Mit dieser Nummer werden Dozierende ermutigt, kritisch darüber nachzudenken, welche Botschaften

sie zukünftigen und bereits im Beruf tätigen Lehrenden vermitteln.

Geniessen Sie die Achterbahnreise durch Tema 1 („Beurteilung von rezeptiven Kompetenzen“), Tema 2 (Sprachtests allgemein) und Tema 3 („Praxis“). Zur Förderung der Diskussion dient bestimmt unser „Reader’s Guide“.

Reader’s Guide

You can use this issue of Babylonia for discussing the topic with fellow teachers and students. Here are some questions to discuss before or after reading this issue of Babylonia!

- > What does testing in a competence-oriented way mean? How can tests be written to reflect WHAT learners should be able to do in the language?
- > What systems are you (teachers) using to have an evidence-based approach to assessing both receptive and productive skills? How do you gather evidence? Did you pick up on any new tools or ideas that you’d like to try out?
- > Teachers in some cantons are asked to report on their learners’ reading, writing, speaking and listening skills in isolation on school report cards. How, thus, can they measure “interaction” and also separate the skills for these reports?
- > What constructs can you find mentioned for: listening comprehension, intercultural skills, literature studies, a good test? When you are focusing on a specific skill in the classroom, how do you make learners take ownership for what is being measured?
- > One could argue that testing in schools is not even necessary, and that performance assessments (presentations, essays, dialogues), good note-taking habits and smaller quizzes would be enough – do you agree?
- > What makes a good reading, writing, speaking or listening test?
- > What do vocabulary-translation tests measure? From this issue of Babylonia, can you make a list of other techniques that might provide a more in-depth picture of learner knowledge and skill?

References

- DeLuca, C., Chavez, T., Bellara, A. & Cao, C.** (2013). Pedagogies for preservice assessment education: Supporting teacher candidates’ assessment literacy development. *The Teacher Educator*, 48: 2, 128-142.
- Pat-El, R., Tillema, H., Segers, M. & Vedder, P.** (2015). Multilevel predictors of differing perceptions of Assessment for Learning practices between teachers and students. *Assessment in Education: Principles, Policy & Practice*. 22, 2.
- Popham, W. J.** (2011). Assessment literacy overlooked: A teacher educator’s confession. *The Teacher Educator*, 46(4), 265-273.

VON FREMDSPRACHENLERNEN, LEHRPLAN 21 UND BEURTEILUNG

The introduction of Lehrplan 21 encourages certain changes to how language skills are assessed in Swiss public schools and new report card models are being developed that support these changes. In describing what might be considered as “standards-based grading”, Christian Amsler provides an overview of the main considerations for classroom teachers as they plan their assessment practices.

● Christian Amsler | Schaffhausen

Christian Amsler, Regierungsrat (FDP), ist als Vorsteher des Erziehungsdepartements des Kantons Schaffhausen zuständig für die Bereiche Bildung, Jugend, Familie, Sport und Kultur. Der 52jährige Politiker ist mit einer Lehrerin verheiratet, Vater von 3 Kindern und präsidiert die Erziehungsdirektoren-Konferenz der Deutschschweiz (D-EDK). Damit ist er auch Schirmherr des Lehrplans 21. Eher Sprachen- denn Zahlenmensch, ist er ein grosser Anhänger der viersprachigen Schweiz, liebt besonders Französisch und Englisch.



Beurteilung ist ein Brennpunktthema, das losgelöst von den einzelnen Fächern im pädagogischen Umfeld heiss diskutiert wird. Neue ganzheitliche Perspektiven sind in den vergangenen Jahrzehnten in der Fachdiskussion, in der Ausbildung der Lehrerinnen und Lehrer wie auch in kantonalen Entwicklungsprojekten thematisiert worden und in kantonale Konzepte, Regelungen und Handreichungen eingeflossen (Beurteilen und Fördern). Auch bezüglich des gemeinsam von 21 Kantonen der Deutschschweiz erarbeiteten Lehrplan 21 stellen sich diese Fragen unverändert. Hingegen erleichtert und fördert der Lehrplan 21 die Umsetzung von Beurteilungskonzepten, welche die Förderung der Schülerinnen und Schüler in den Vordergrund stellen, und die sich verstärkt an der sachlich, kriterialen aber weniger an der sozialen Bezugsnorm orientiert. Dies ist eine Folge der Beschreibung des Kompetenzaufbaus in der Form von Könnens-Beschreibungen. Wissen und Können werden zur Kompetenz verknüpft. Die Begrifflichkeit Kompetenzorientierung besagt lediglich ausdrücklich, dass der Lehrplan nicht

bereits erfüllt ist, wenn der im Lehrplan aufgelistete Stoff im Unterricht behandelt wurde, sondern erst dann, wenn die Kinder und Jugendlichen über das nötige Wissen verfügen und dieses in einer konkreten Situation anwenden können. Dieses Verständnis entspricht zudem sowohl der heutigen pädagogischen Praxis wie auch den Anforderungen der Lehrbetriebe und der weiterführenden Schulen an die Schulabgänger. Es gilt ganz besonders für das Fremdsprachenlernen.

Die Freude an der Sprache und die Freude am Kommunizieren stehen im Zentrum. Auch in den Fremdsprachen werden die zu erreichenden Grundansprüche und deren Beurteilung im Lehrplan 21 klar definiert. Die Grundansprüche des Lehrplans 21 basieren in den Fachbereichen Schulsprache (Deutsch), den Fremdsprachen Französisch und Englisch, in Mathematik sowie in den Naturwissenschaften auf den von der EDK beschlossenen und gesamtschweizerisch geltenden Grundkompetenzen (den nationalen Bildungszielen). Damit wird der Auftrag der Bundesverfassung zur Harmonisierung der Ziele der Bildungsstufen umgesetzt.

PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > SCHEDE [6](#) E [14](#)

Wie gut diese nationalen Bildungsziele in den Kantonen erreicht werden, wird die EDK im Rahmen des Bildungsmonitorings gesamtschweizerisch überprüfen. Jeder Kanton wird daran mit einer Stichprobe von Schulklassen teilnehmen. In diesen werden alle 3 bis 4 Jahre Tests in den Fachbereichen Schulsprache (Deutsch), Mathematik, Französisch / Englisch sowie in den Naturwissenschaften durchgeführt. Die Auswertung erfolgt auf der Ebene der Kantone und lässt keine Rückschlüsse auf die Ergebnisse einzelner Schülerinnen und Schüler, Klassen oder Schulen zu. Die Ergebnisse bilden eine Grundlage für die Qualitätsentwicklung auf gesamtschweizerischer Ebene.

Generell erscheinen mir folgende Aspekte beim Thema Beurteilung und Fremdsprachenlernen wichtig und bedenkenswert:

- > Die ganzheitliche Beurteilung umfasst kriterienorientiertes Beobachten und Beschreiben von Lernleistungen, sowohl formative als auch summative Lernkontrollen, Selbstbeurteilung der Lernenden und Fremdbeurteilung durch die Lehrperson und ein ausgewogenes Verhältnis der Beurteilung von produktiven und rezeptiven Fertigkeiten.

> Grundsätzlich soll sich die Beurteilung an Kompetenzen statt an Defiziten orientieren. Die Beurteilung fokussiert, was die Lernenden können und nicht, was sie falsch machen. Im Zentrum der Beurteilung stehen Kompetenzen in den folgenden kommunikativen Fertigkeiten: Hörverstehen, Lesen, Sprechen (monologisch und dialogisch) und Schreiben. Die Kompetenzbereiche Wortschatz, Aussprache und Grammatik werden nicht isoliert, sondern in die Fertigkeiten und in einen Kontext integriert beurteilt. Die Beurteilung ist ausgewogen und konzentriert sich nicht einseitig auf formale Aspekte. Vom weit verbreiteten isolierten, kontextunabhängigen Wörterprüfen zum Beispiel ist abzusehen und die Beurteilung enthält Hinweise auf die Verbesserung von inhaltlichen und formalen Aspekten.

> Ein fertigungsorientierter Unterricht bedingt auch fertigungsorientierte Formen der Beurteilung und somit fertigungsorientierte Tests. Die Inhalte aus dem Unterricht werden mitbeurteilt. Ganz nach dem Motto „Test what you teach and teach what you test!“

> Die Schülerinnen und Schüler können sich auf völlige Transparenz verlassen: Die Lernziele der Beurteilungsse-

quenzen, die Themen sowie die Inhalte und Formen der Lernkontrollen sind bekannt.

Beurteilungsblatt 1. Fremdsprache				Französisch			
Name: Felix Muster			Klasse: 5				
Datum	Beurteilungsanlass	Note/Prädikat		Bemerkungen			
Hören				2. Zyklus		3. Zyklus	
				A1.1	A1.2	A2.1	A2.2
Lesen				2. Zyklus		3. Zyklus	
				A1.1	A1.2	A2.1	A2.2
Sprechen				2. Zyklus		3. Zyklus	
				A1.1	A1.2	A2.1	A2.2
Schreiben				2. Zyklus		3. Zyklus	
				A1.1	A1.2	A2.1	A2.2
Sprache im Fokus							
	Wortschatz						
	Aussprache						
	Grammatik						
	Rechtschreibung						
Kulturen im Fokus							
Zeugnisnote 1. Semester:				2. Semester:			
Entwurf AG Beurteilen				27.08.2015			

Zuständig für den Bereich Beurteilung sind die Kantone. Auf dem Weg zur Implementierung des Lehrplans 21 ist selbstverständlich auch der Bereich Beurteilung / Notengebung ein wichtiges Thema. Der Fachbericht Beurteilen der D-EDK stellt im Sinne von Hilfen den Kantonen gewisse Unterlagen zur Verfügung; so z.B. Beurteilungsblätter für die einzelnen Fächer. Das Beurteilungsblatt für die 1. und 2. Fremdsprache ist nach den Kompetenzbereichen des Lehrplans 21 strukturiert. Es werden alle dort definierten Kompetenzbereiche abgebildet.

Fachbericht Beurteilen der Arbeitsgruppe der Kommission Volksschule der D-EDK (D-KV), 2015.

Die Begrifflichkeit Kompetenzorientierung besagt lediglich ausdrücklich, dass der Lehrplan nicht bereits erfüllt ist, wenn der im Lehrplan aufgelistete Stoff im Unterricht behandelt wurde, sondern erst dann, wenn die Kinder und Jugendlichen über das nötige Wissen verfügen und dieses in einer konkreten Situation anwenden können.

Quellenangaben

Grundsätze der Beurteilung im Englischunterricht der Primarstufe, erarbeitet von der Projektgruppe Englisch Primarschule der BKZ.

In den Kompetenzbereichen Hören, Lesen, Sprechen und Schreiben liegt den Kompetenzstufen die Systematik des Gemeinsamen Europäischen Referenzrahmens für Sprachen zugrunde. Das Beurteilungsblatt bietet für diese Kompetenzbereiche die Möglichkeit zu vermerken, welche Kompetenzstufe die Schülerin oder der Schüler am Ende des Schuljahrs erreicht hat. Es wird sichtbar gemacht, welche Kompetenzstufen dem Grundanspruch des Zyklus und welche dem Auftrag des Zyklus entsprechen. Die Einstufung erfolgt mit Hilfe der Deskriptoren des Europäischen Sprachenportfolios (ESP) und kann als Selbst- und Fremdbeurteilung vorgenommen werden.

Im Bereich „Sprache im Fokus“ sollten die vier Kompetenzen Wortschatz, Aussprache, Grammatik und Rechtschreibung beurteilt werden; diese sind daher auf dem Beurteilungsblatt bereits vorgegeben.

Die Beurteilungsblätter sind auf die Verwendung über ein ganzes Schuljahr ausgerichtet. Eine Verwendung am Ende des 1. Semesters bedeutet, dass es sich um eine vorläufige Zwischenbilanz handelt. Die Anzahl der Beurteilungsanlässe pro Schuljahr muss genügend gross sein, um eine ausgewogene Beurteilung aller Teilbereiche an mehreren Beurteilungsanlässen zu ermöglichen. Das Beurteilen soll jedoch nicht den Unterrichtsalltag dominieren. Daher wird nicht jede Beurteilungssituation auch als Leistungssituation ausgestaltet: Das Lernen soll im Zen-

trum stehen. Um die nötigen Grundlagen für die Beurteilung im Zeugnis mit vertretbarem Aufwand erstellen zu können, nutzt die Lehrperson insbesondere auch prozessbegleitende Beurteilungsanlässe. Die Beurteilungsblätter dienen der Transparenz. Daher sollte den Eltern im Rahmen von Beurteilungs- oder Standortgesprächen Einblick in die Blätter der eigenen Kinder gewährt werden.

Die Ableitung der Zeugnisnoten aus den Prädikaten der einzelnen Beurteilungsanlässe ist ein professioneller Ermessensentscheid. Das heisst, dass sich Zeugnisnoten nicht aus einem arithmetischen Mittel von Noten einzelner Beurteilungsanlässe errechnen lassen. Das heisst aber nicht, dass das Setzen von Zeugnisnoten zu Willkür und Beliebigkeit führen darf. Die Ableitung der Zeugnisnoten erfolgt in zwei Schritten. In einem ersten Schritt werden die Beurteilungsanlässe eines Kompetenzbereichs zusammengefasst. Dabei wird berücksichtigt, ob die Tendenz der einzelnen Beurteilungen auf- oder absteigend ist. Zudem ist zu entscheiden, wie mit ‚Ausreissern‘ umgegangen werden soll. Im zweiten Schritt werden die Beurteilungen der einzelnen (Kompetenz-) Bereiche zu einer Zeugnisnote zusammengefasst.

Beurteilung, Fördern, Feedback, Notengebung sind und bleiben anspruchsvolle Kerntätigkeiten, bilden gleichzeitig aber auch eine der wichtigsten Fähigkeiten von Lehrerinnen und Lehrern – auch im Fremdsprachenlernen!

LESEVERSTEHEN ÜBERPRÜFEN

WELCHE SPRACHE FÜR DIE FRAGEN UND ANTWORTEN IN DEN TESTITEMS?

Notre article présente les premiers résultats du projet de recherche *Task Lab* mené au Centre scientifique de compétence sur le plurilinguisme. Le projet examinait entre autres une question qui se pose souvent dans le contexte de l'évaluation des langues étrangères à l'école: le choix de la langue des questions et des réponses. Si, par exemple, des élèves suisse-allemands passent un test dans lequel ils lisent un texte en français et répondent à des questions, faudrait-il leur poser ces questions – y compris les options dans le cas des items à choix multiple – en allemand ou en français? Et, s'ils doivent répondre à ces questions par écrit, devraient-ils écrire en allemand ou en français?

L'analyse des données, principalement qualitative, démontre que les questions et options de réponse dans la langue étrangère (ici: français) n'ont souvent pas été comprises de manière suffisante. Par ailleurs, les compétences d'écriture en français n'étaient pas encore assez développées pour exprimer ce que les élèves voulaient répondre sur la base de leur compréhension du texte. Nous concluons que l'utilisation de la langue de scolarisation (ici: allemand) permet d'évaluer de manière plus fiable les compétences de compréhension écrite en langue étrangère.

Tema

1. Hintergrund: Projekt *Task Lab*

Im Frühling 2017 wird von der Schweizerischen Konferenz der kantonalen Erziehungsdirektoren (EDK) zum ersten Mal das Erreichen der Grundkompetenzen in der ersten Fremdsprache am Ende der Primarstufe überprüft. Die Tests werden computerbasiert stattfinden und umfassen Aufgaben zum Lese- und Hörverstehen. Mit der Entwicklung dieser Testaufgaben wurde das Kompetenzzentrum für Mehrsprachigkeit (KFM) beauftragt. Bereits im Vorfeld der Aufgabenentwicklung führte das KFM das Projekt *Task Lab* durch, das wichtige Vor- und Grundlagenarbeiten für diese und geplante weitere Untersuchungen im Bereich des computerbasierten Testens von Leseverstehenskompetenzen leisten sollte. Das allgemeine Ziel des Projekts bestand darin, einen besseren Einblick in das Faktorengewebe zu gewinnen, welches beim Lösen der Leseverstehensaufgaben eine Rolle spielt. Dadurch soll sowohl die Erstellung zukünftiger Aufgaben verbessert als auch ein stärkeres empirisches Fundament für die Interpretation von Testergebnissen und -skalen geschaffen werden.

Im Rahmen von *Task Lab* wurde schwerpunktmässig das Funktionieren von Aufgaben mit unterschiedlichen Formaten (*Multiple Choice*, Kurzantwort und *Matching*¹) und mit unterschiedlichen sprachlichen Merkmalen überprüft, insbesondere bezüglich der Sprache der Fragen und Antworten. Das methodische Vorgehen war vorwiegend quantitativ, jedoch wurde bei der Vorerprobung der Aufgaben auch qualitativ gearbeitet. Die Hauptuntersuchung fand im Frühjahr 2015 unter Beteiligung von insgesamt 35 sechsten Klassen (8. Schuljahr gemäss HarmoS) mit über 500 Schülerinnen und Schülern der Passepartout-Region statt. Diese Lernenden lernen seit der dritten Klasse Französisch und sollten im Leseverstehen laut dem für sie geltenden

¹ *Matching*: Zuordnungsaufgaben, bei denen aus 5-7 kurzen Texten je einer aufgrund einer inhaltlichen Frage ausgewählt werden musste. Wie die anderen Formate bestanden auch die *Matching*-Aufgaben aus je drei inhaltlichen Fragen zu den gleichen 5-7 Kurztexen.

Malgorzata Barras,
Katharina Karges &
Peter Lenz | Freiburg

Malgorzata Barras
wissenschaftliche
Mitarbeiterin



Katharina Karges
wissenschaftliche
Mitarbeiterin



Peter Lenz
Projektleiter



arbeiten am Wissenschaftlichen Kompetenzzentrum für Mehrsprachigkeit (Universität Freiburg und PH Freiburg/Schweiz), u.a. zu den Themen kompetenzorientierte Beurteilung, computerbasiertes Testen, fremd- und zweitsprachliche Kompetenzprofile, Zweitspracherwerb bei Erwachsenen.

- 2 Jede „Aufgabe“ besteht aus einer Textgrundlage und drei Items.
- 3 Bei den sechs Matching-Aufgaben gab es jeweils nur Fragen. Als Antwort musste ein zielsprachlicher Text ausgewählt werden.

Lehrplan das Niveau A2.1 erreicht haben. In der Hauptuntersuchung lösten die Schülerinnen und Schüler eine Reihe von Leseverstehensaufgaben in Französisch und bearbeiteten zusätzlich weitere schriftliche und mündliche Tests, die der Erfassung von Kompetenzkomponenten (z.B. des phonologischen Bewusstseins oder der Gedächtnisspanne) dienen. Zudem füllten die Schülerinnen und Schüler einen Fragebogen zu ihrem sprachlichen Hintergrund, ihrer Motivation für den Fremdsprachenunterricht sowie ihren Lese- und Lerngewohnheiten aus. In einer frühen Projektphase wurden die computerbasierten Aufgaben zum Leseverstehen mit einer Reihe von Kindern vorerprobt. Dabei lösten diese ein Testitem nach dem anderen und gaben jeweils direkt im Anschluss daran im Rahmen von sog. retrospektiven Interviews Auskunft über ihren Lösungsweg.

2. Sprache der Testfragen und -antworten

Wie eingangs erwähnt, betraf eine der zentralen Forschungsfragen des Projekts *Task Lab* die Wahl der Sprache für die Fragen, Antwortoptionen und Kurzantworten in den Aufgaben. Untersucht wurde, welche Kombination aus Sicht der Testvalidität besser ist und deswegen in Leistungsmessungen (und bei anderen Beurteilungsanlässen) zum Einsatz kommen sollte: Fragen und Antworten in der Schulsprache (d.h. in der Deutschschweiz Deutsch) oder in der Zielsprache (also in der *Passepartout*-Region Französisch bzw. Englisch)?

Quantitative wie auch qualitative Methoden sollten zur Klärung dieser Frage beitragen. Quantitative Informationen zu diesem Aspekt wurden durch die systematische Variation der beiden Sprachen in den Fragen und Antworten der Leseverstehensaufgaben erhoben. Konkret heisst das, dass jede der insgesamt 18 in der Studie eingesetzten Aufgaben (mit insgesamt 54 Items²) jeweils in einer Version mit deutschsprachigen und einer ansonsten identischen Version mit französischsprachigen Fragen und Antworten³ vorlag. Bei 12 Texten wurde darüber hinaus auf die gleiche Weise das Antwortformat variiert (vgl. Abbildung 1 für ein Beispiel). In den verschiedenen Testheften, die für die Lernenden zusammengestellt wurden, kamen jeweils ungefähr gleich viele Aufgaben jeder Sprachversion vor.

Qualitative Daten zur Frage der Sprachenwahl lieferten einerseits die bereits erwähnten retrospektiven Interviews mit 34 Schülerinnen und Schülern und andererseits Antworten aus einem Kurzfragebogen, den die Lernenden direkt nach dem Lösen der Aufgaben in der quantitativen Hauptuntersuchung ausfüllten. In diesem Artikel werden vor allem die Ergebnisse des qualitativen Untersuchungsteils vorgestellt. Sie werden durch einen ersten Blick auf die quantitativen Ergebnisse ergänzt.

3. Beobachtungen aus den retrospektiven Interviews

Bei den retrospektiven Interviews arbeiteten Projektmitarbeitende mit einzelnen Schülerinnen und Schülern zusammen.

Abbildung 1: Unterschiedliche Varianten des gleichen Items (kommentiertes Beispiel).

Die Abbildung zeigt vier Varianten eines Testitems, die auf dem Text eines E-Mails von Alicia basieren. Die Varianten unterscheiden sich durch die Sprache der Fragen und die Art der Antwortmöglichkeiten:

- Variante 1 (Französisch/Französisch):** Die Frage lautet 'Quel est le thème du mail d'Alicia ?'. Die Antwortmöglichkeiten sind 'Kurzantwort Französisch' und 'Multiple Choice Französisch'.
- Variante 2 (Deutsch/Deutsch):** Die Frage lautet 'Über welches Thema schreibt Alicia in ihrem Mail?'. Die Antwortmöglichkeiten sind 'Kurzantwort Deutsch' und 'Multiple Choice Deutsch'.
- Variante 3 (Französisch/Deutsch):** Die Frage lautet 'Quel est le thème du mail d'Alicia ?'. Die Antwortmöglichkeiten sind 'Kurzantwort Deutsch' und 'Multiple Choice Französisch'.
- Variante 4 (Deutsch/Französisch):** Die Frage lautet 'Über welches Thema schreibt Alicia in ihrem Mail?'. Die Antwortmöglichkeiten sind 'Multiple Choice Französisch' und 'Multiple Choice Deutsch'.

Diese bearbeiteten jeweils ein Leseverstehensitem und gaben gleich anschliessend über ihr Vorgehen und ihre Gedanken Auskunft. Nach Bedarf wurden vertiefende Fragen gestellt. Die Auswertung der Interviews hat gezeigt, dass in den Aufgaben etwa die Hälfte der auf Französisch gestellten Fragen, und im Fall von Multiple-Choice-Fragen auch viele der Antwortoptionen, von den Schülerinnen und Schülern entweder nicht oder falsch verstanden wurden. Die Verständnisprobleme verunsicherten viele der Kinder, da sie sich nicht sicher sein konnten, ob sie die Frage richtig interpretiert hatten, und so daran gehindert wurden, genügend Anhaltspunkte für die Suche nach der richtigen Antwort im Text zu finden. Lagen die Fragen hingegen auf Deutsch vor, wurden diese nur in seltenen Fällen falsch verstanden, und zwar vor allem dann, wenn zu schnell und/oder zu ungenau gelesen wurde.

Die Gründe für die Probleme der Lernenden mit den auf Französisch formulierten Fragen waren vielfältig. In manchen Fällen waren die Lernenden gänzlich ratlos. Wurde die Frage nicht verstanden, führte dies oft dazu, dass bei Multiple-Choice-Items irgendeine Antwort angekreuzt und bei den Kurzantwort-Items irgendeine Antwort geschrieben oder auch einfach irgendetwas vom Text abgeschrieben wurde.

In anderen Fällen konnten die Kinder in den Fragen (und beim Multiple-Choice-Format in den Antwortoptionen) nur einzelne (Parallel-)Wörter entschlüsseln (Beispiel 1).

Beispiel 1:

Frage: *Qu'est-ce qu'Alicia a aimé le plus?*

- a) *Essayer de faire du jonglage.*
- b) *Manger les spaghettis.*
- c) *Voir la petite girafe.*

Schüler/in (S1): Ich habe die Frage nicht so verstanden.

Interviewer/in (I): Was hast du verstanden?

S1: Eine kleine Giraffe und etwas von Spaghetti essen. Bei der ersten Antwort habe ich nur „jonglage“ verstanden.

I: Was heisst die Frage?

S1: Das weiss ich eben nicht.

Es kam aber auch häufig vor, dass der Sinn einer Frage nur deswegen nicht richtig erfasst wurde, weil ein einzelnes Fragewort bzw. Schlüsselwort nicht oder falsch verstanden wurde (Beispiele 2 und 3).

Beispiel 2:

Frage: *Quel est le problème avec C dans cette scène?*

I: Wie lautet die Frage?

S2: Wer löst die Probleme in dieser Szene.

Beispiel 3:

Frage: *Pourquoi le chemin de l'école de Vidal est-il spécial?*

I: Weisst du, wie die Frage lautet?

S3: Warum ist das Chemielabor von Vidal so speziell.

Einige Kinder berichteten von sich aus, dass sie den Lesetext eigentlich gut verstehen würden, nur die französische Frage nicht.

Die Schülerinnen und Schüler wurden im Verlauf der Interviews auch gefragt, ob sie die Fragen und Antworten in der Schulsprache oder in der Zielsprache einfacher fänden. Wenig überraschend gaben die meisten von ihnen an, die Fragen in ihrer Schulsprache vorzuziehen. Darüber hinaus wurden sie auch gefragt, ob es für sie in diesem Fall schwierig sei, zwischen Deutsch und Französisch zu wechseln. Die grosse Mehrheit verneinte dies aber.

Wie bereits erwähnt, wurden bei Multiple-Choice-Items, die in der Zielsprache verfasst waren, neben den Fragen auch die Antwortoptionen auf Französisch präsentiert. Das gleiche Prinzip galt bei den Kurzantwort-Items: Wurde die Frage auf Französisch gestellt, musste sie auch in dieser Sprache beantwortet werden. Auf Französisch eine Antwort zu schreiben, stellte für viele Kinder ein besonders grosses Problem

dar. Die Schülerinnen und Schüler gaben an, dass ihnen beim Schreiben Wörter fehlten und dass es für sie schwierig sei, sich gleichzeitig auf den Inhalt, auf die Syntax und auf die Rechtschreibung zu konzentrieren (obwohl ihnen am Anfang klar kommuniziert wurde, dass sprachliche Fehler bei diesem Test keine Rolle spielten). Ausserdem konnte in den Gesprächen festgestellt werden, dass die Lernenden nicht selten etwas anderes schreiben wollten, als was dann tatsächlich auf Französisch zu lesen war:

Beispiel 4:

S4: Ich habe geschrieben: „Pierre Dumont en danger.“

I: Was würdest du auf Deutsch schreiben, wenn du diese Antwort auf Deutsch schreiben könntest?

S4: Pierre Dumont ist sehr gefährlich.

Beispiel 5:

I: Kannst du mir deine Antwort kurz vorlesen?

S5: Also Karusu verliert seinen Papa im Zoo.

I: Und geschrieben hast du: „Karusu devenu papa“.

Einige Kinder gaben an, dass es für sie einfacher war, wenn die Fragen auf Französisch waren, weil sie dann die Stelle im Text, auf die sich die Frage bezog, einfacher identifizieren konnten. Diese Strategie führte jedoch aufgrund der Konzeption der Fragen verschiedentlich ins Leere: Bei den Kurzantwort-Fragen schrieb so manches Kind wegen eines einzelnen Schlüsselwortes eine unpassende Stelle aus dem Text ab und bei den Multiple-Choice-Aufgaben wurde jene Antwort ausgewählt, zu der die Lernenden zuerst Wörter im Text fanden. Ein Kind berichtete, dass es viel schwieriger sei, die Fragen auf Deutsch zu beantworten, wenn man den Text nicht verstanden habe. Erkenntnisse wie diese sprechen dafür, die Fragen in der Schulsprache zu stellen, wenn es darum gehen soll, das Textverständnis in der Fremdsprache zu überprüfen.

4. Welche Sprache finden die Schülerinnen und Schüler leichter? Warum? Antworten aus dem Fragebogen

Im Fragebogen, der im Anschluss an die Haupterhebung im Leseverstehen auszufüllen war, mussten die Lernenden ankreuzen, ob sie die Fragen und Antworten auf Deutsch oder auf Französisch einfacher fanden (mit „Antworten“ wa-

ren sowohl offene als auch geschlossene Formate gemeint). Zudem sollten sie ihre Wahl kurz begründen. Die Ergebnisse fallen sehr deutlich aus: 528 Schülerinnen und Schüler gaben an, die Fragen und Antworten auf Deutsch seien für sie einfacher gewesen, und nur 50 fanden es leichter, mit den Fragen und Antworten in französischer Sprache umzugehen. Weitere 19 Lernende konnten sich offenbar nicht entscheiden und kreuzten entweder beides oder nichts an. Die Begründungen der 528 Schülerinnen und Schüler, welche die Fragen und Antworten in der Schulsprache Deutsch leichter fanden, lassen sich wie folgt zusammenfassen:

- > Deutsch kann ich besser./Deutsch kann ich besser verstehen und schreiben (162 Schüler/innen). Die Fragen auf Deutsch kann ich (besser) verstehen (158 Schüler/innen). Auf Deutsch kann ich besser schreiben/kann ich die Antworten besser formulieren (86 Schüler/innen).
- > Deutsch ist meine Muttersprache (29 Schüler/innen).
- > Ich bin nicht gut in Französisch (49 Schüler/innen).
- > Ich weiss nicht warum (44 Schüler/innen).

Die Antworten der 50 Schülerinnen und Schüler, welche die Fragen auf Französisch leichter fanden, können wie folgt kategorisiert werden:

- > Auf Französisch kann man die Antworten direkt vom Text abschreiben (25 Schüler/innen).
- > Es ist einfacher, auf Französisch zu antworten, weil auch der Text in dieser Sprache vorliegt (16 Schüler/innen).
- > Ich bin (sehr) gut in Französisch (4 Schüler/innen).
- > So kann ich mehr lernen (1 Schülerin).
- > Ich weiss nicht warum (4 Schüler/innen).

Von den 19 Lernenden, die sich nicht entscheiden konnten, welche Sprache der Fragen und Antworten für sie leichter war, gaben zwölf an, dass sie beide Sprachen als gleich einfach oder auch gleich schwierig empfanden. Die übrigen sieben Kinder sahen bei beiden Sprachvarianten erschwerende und erleichternde Faktoren (z.B.: *Weil man die Frage nicht übersetzen muss. Aber man kann dann auch keine Französischen (sic!) Wörter übernehmen.*)

Erkenntnisse wie diese sprechen dafür, die Fragen in der Schulsprache zu stellen, wenn es darum gehen soll, das Textverständnis in der Fremdsprache zu überprüfen.

Die Ergebnisse der Befragung zeigen klar, dass es für eine grosse Mehrheit der Schülerinnen und Schüler leichter war, wenn die Fragen zu den fremdsprachigen Lesetexten in der Schulsprache verfasst waren und auch in der Schulsprache beantwortet werden konnten. Sie nannten dafür eine ganze Reihe plausibler Argumente. Auch eines der Hauptargumente der Kinder für die Fragen und Antworten in der Fremdsprache – das Abschreiben vom Text – spricht gerade für die Formulierung der Fragen und Antworten in der Schulsprache.

5. Vergleich der Testresultate bei den Fragen und Antworten in der Schul- und in der Zielsprache

Ein erster, grober Vergleich der Schülerleistungen in der Hauptuntersuchung in Bezug auf die Testresultate in den beiden Sprachvarianten zeigt, dass die Lernenden bei den Aufgaben mit den Fragen und Antworten in der Schulsprache deutlich besser abgeschnitten haben⁴. Bei allen drei Itemformaten wurde die richtige Antwort häufiger gefunden, wenn die Sprache der Fragen und der Antworten die Schulsprache war.

Bei der Auswertung der Antworten auf die Kurzantwort-Fragen wurde auch die Anzahl der abwegigen⁵ und der direkt vom Text übernommenen Antworten⁶ erfasst (vgl. Tabelle 1). Die Ergebnisse sind frappierend: Jede dritte Antwort auf Französisch wurde als abwegig kodiert, während dies auf Deutsch nur bei jeder zehnten Antwort der Fall war. Bei den abwegigen Antworten auf Deutsch liegt die Vermutung nahe, dass die Schülerinnen und Schüler den entsprechenden Lesetext nicht bzw. ungenügend verstanden. Dank der Interviews bei der Vorerprobung können wir davon ausgehen, dass die Ursache für die bedeutend höhere Zahl der abwegigen französischsprachigen Antworten oft darin lag, dass die Lernenden die Frage nicht verstanden und/oder ihre Antwort auf Französisch nicht formulieren konnten. Bemerkenswert ist ausserdem die Tatsache, dass mehr als ein Drittel aller Antworten auf Französisch direkt aus dem Text abgeschrieben wurde. Bei diesen bleibt oft unklar, aus welchem Grund die Lernenden den Textausschnitt auswählten: Haben sie ihn tatsächlich verstanden (und damit die Frage richtig oder falsch beantwortet)? Oder haben sie den Textausschnitt zufällig z.B. aufgrund eines Wortes ausgewählt (ohne die Frage

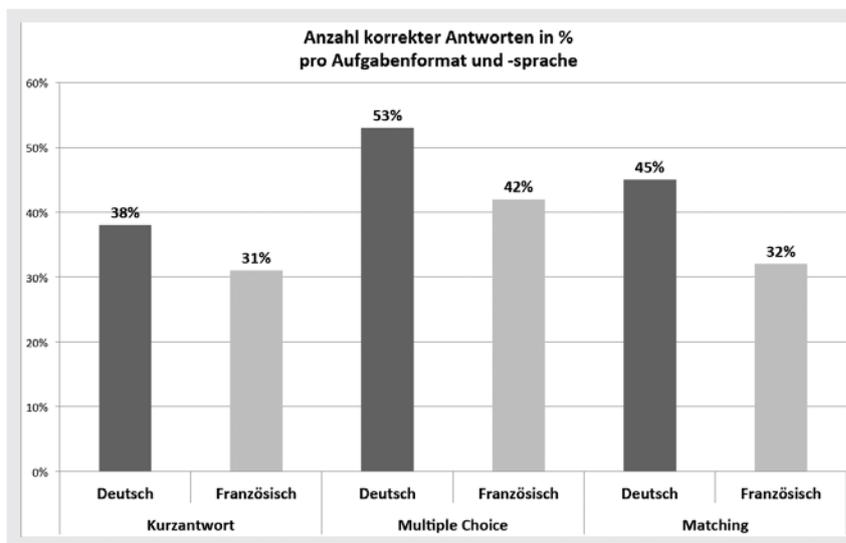


Abbildung 2: Anteil korrekter Antworten bei Deutsch oder Französisch pro Aufgabenformat.

Tabelle 1: Anteil abwegiger und abgeschriebener Antworten bei Kurzantworten auf Deutsch bzw. Französisch.

oder die Antwort darauf zwingend verstanden zu haben)? - Der hohe Anteil von Antworten, die Zweifel aufkommen lassen, spricht gegen die Validität von Items, die ganz in der Zielsprache gehalten sind.

Sprache der Fragen und Antworten	Anteil abwegiger Antworten	Anteil vom Text übernommener Antworten
Französisch	33%	37%
Deutsch	10%	1%

6. Schluss

Die dargestellten Resultate zeigen, dass in unserer Untersuchung bei Leseverstehensitems Fragen und Antworten in der Schulsprache zu valideren Ergebnissen geführt haben als solche in der Zielsprache. Dies geht aus der Analyse sowohl der retrospektiven Interviews als auch der Antworten im Fragebogen deutlich hervor. Viele Sechstklässlerinnen und Sechstklässler gaben an, Fragen und Antwortoptionen auf Deutsch (besser) zu verstehen. Auch können sie in dieser Sprache Antworten schreiben, die besser das ausdrücken, was sie ausdrücken wollen. Bei Items, die eine französischsprachige Kurzantwort verlangten, gaben Kinder verschiedentlich an, vom Text abgeschrieben zu haben, und zwar auch dann, wenn sie die betreffende Stelle nicht wirklich verstanden hatten. Schliesslich zeigt der deutlich tiefere Anteil richtiger Antworten bei Items, die ganz in der Zielsprache gehalten sind,

- 4 Die statistische Auswertung der Resultate ist zum Zeitpunkt der Entstehung dieses Artikels noch nicht abgeschlossen.
- 5 Eine Antwort wurde als abwegig kodiert, wenn zwischen der Frage und der Antwort kein inhaltlicher Zusammenhang erkennbar war. Zum Beispiel: Frage: *Où sont les trois personnes?*; Antwort: *elle intéressante ce livre, imageset parler. (Sic!)*
- 6 Eine Antwort wurde als direkt vom Text übernommen kodiert, sobald sie drei oder mehr Wörter umfasste, welche in der genau gleichen Form und Reihenfolge im Text vorkamen.

dass die Sprache der Fragen und Antworten im Verstehensprozess eine wichtige Rolle spielt.

Die Deutlichkeit, mit welcher ein unerwünschter Einfluss von Fragen und Antworten in der Zielsprache sichtbar wird, hat offensichtlich mit dem noch sehr tiefen Sprachniveau der Lernenden zu tun. Nicht-triviale Aufgabenstellungen zum Leseverstehen lassen sich aber sprachlich nicht beliebig vereinfachen. Es wäre interessant, herauszufinden, ab welchem Sprachkompetenzniveau die Umsetzung von vergleichbaren Aufgaben ganz in der Zielsprache möglich wäre, ohne dass dabei nicht mehr vertretbare Abstriche einerseits beim Testinhalt und andererseits bei der Testvalidität gemacht werden müssten.

Gegen die Entscheidung für Fragen und Antworten in der Schulsprache wird oft argumentiert, dass bekannte Prüfungsanbieter von internationalen Sprachtests wie z.B. Cambridge, Pearson, CIEP (DELF-DALF) oder telc nicht auf die Erstsprache der Testteilnehmenden zurückgreifen. Die Einsprachigkeit solcher Prüfungen ohne weitere Überlegung als anzustrebendes Ideal zu sehen, wäre

aber falsch, denn sie ergibt sich wohl als einzige praktikable Möglichkeit aus den Gegebenheiten. Die Zahl der „Schulsprachen“ der potenziellen Kandidatinnen und Kandidaten ist nämlich gross, was zu enormen Kosten für Testentwicklung, Qualitätssicherung und Logistik führen müsste. Damit die Aufgaben auch auf den untersten Niveaus zu verstehen sind, werden sprachlich einfache, oft bildgestützte Formate eingesetzt, die durch Übungstestsätze bekannt gemacht werden und von den Kandidatinnen und Kandidaten oft in spezifischen Kursen eingeübt werden. Bei schulbezogenen Tests sind die Bedingungen in zweifacher Hinsicht anders: Zum einen gibt es innerhalb einer Sprachregion eine gemeinsame Schulsprache, womit der Anlass fehlt, die Tests auf einfache, oft auch inhaltlich reduktionistische Formate zu reduzieren. Zum anderen besteht keine realistische Möglichkeit, das ganze Schulsystem intensiv auf die Bewältigung der geeigneten Formate vorzubereiten. Unter diesen Umständen spricht nichts gegen, aber – wie die Ergebnisse unserer Untersuchung zeigen – vieles für Fragen und Antworten in der Schulsprache.

Empfehlung

Aufgrund der klaren Ergebnisse der Auswertung der qualitativen Daten kann zumindest für Lernende auf einem tiefen Sprachkompetenzniveau Folgendes als Empfehlung formuliert werden: Beim Testen von fremdsprachlichen Kompetenzen im Leseverstehen sollten Fragen und Antworten in der Schulsprache vorgesehen werden. Mithilfe von Fragen und Antworten in der Schulsprache kann zuverlässiger das gemessen werden, was auch tatsächlich gemessen werden soll: die Fähigkeit, zielsprachliche Texte zu verstehen. Wenn dafür die Zielsprache verwendet wird, besteht die Gefahr, dass die Schülerinnen und Schüler an nicht oder falsch verstandenen Fragewörtern, unvollständigem Verstehen von Antwortoptionen oder an ungenügenden Schreibfertigkeiten scheitern, anstatt dass sie ihre Kompetenzen in dem Bereich unter Beweis stellen können, um den es im Test gehen soll: dem Textverstehen.

POUR UNE ÉVALUATION SOUS CONTRÔLE DIDACTIQUE

QU'EST-CE QUI PERMET DE DIRE QU'UN ÉLÈVE A «COMPRIS» UN TEXTE ORAL ?

La quantità e la diversità dei dispositivi di valutazione degli apprendimenti solleva numerose domande, soprattutto dal punto di vista didattico. Affinché l'allievo possa andare avanti nel suo percorso di apprendimento è primordiale che sia in grado di posizionarsi rispetto agli obiettivi. Questo presuppone però che ci sia una vera e propria riflessione, per esempio sugli scopi e i contenuti delle valutazioni: a che scopo si fa una valutazione? Per chi? E soprattutto, cosa si valuta esattamente?

Basato su delle osservazioni della valutazione della comprensione orale, il nostro contributo si orienta verso la definizione di un quadro per la valutazione che sia in grado di distinguere chiaramente le diverse dimensioni della comprensione orale in relazione alle tipologie di attività che possono evidenziare quelle dimensioni. Le riflessioni proposte trattano principalmente della lingua di scolarizzazione, ma certi punti possono presentare anche un interesse per le lingue seconde.

Tema

1. L'évaluation, oui mais...

Qu'elle soit réalisée par soi-même ou par autrui, l'évaluation – autrement dit la possibilité de situer des connaissances et capacités d'action en lien à des objectifs – est au cœur de l'apprentissage. Ceci vaut pour tout domaine de connaissance, et pour les langues bien évidemment. L'évaluation s'inscrit d'ailleurs dans le déroulement même de l'activité langagière, lorsqu'on évalue et, si on l'estime nécessaire, modifie – ou «corrige» – ce qu'on a écrit ou dit, ou lorsqu'on réagit à ce qu'autrui a écrit ou dit, en demandant une précision, en reformulant ses dires, en les corrigeant. L'évaluation, par conséquent, est également au cœur de l'activité enseignante, pour *aider* les élèves à se situer. Elle est formative et représente une dimension fondamentale de toute didactique.

Cependant, sous la pression de l'économie et de la politique, l'évaluation tend aujourd'hui à être confondue avec la mesure: il s'agit de *rendre compte* à tout prix, de *montrer* les progrès, ou leur absence, de les *certifier*. Pour l'enseignement des langues comme pour d'autres disciplines, on assiste ainsi à une multiplication des

dispositifs de mesure – des épreuves internationales, type PISA, aux tests de référence nationaux, régionaux ou cantonaux et aux tests en ligne qui prolifèrent aujourd'hui –, si possible standardisés, sans qu'il y ait nécessairement une véritable réflexion didactique en amont: pourquoi évalue-t-on? Pour qui? Et, surtout, qu'évalue-t-on vraiment? La (volonté de) rigueur de la mesure tend trop souvent, nous semble-t-il, à prendre le dessus sur les motivations didactiques. Et le risque est grand, dès lors, que l'évaluation reste un pur moyen de classement, voire de sélection, ou qu'elle ne soit qu'un alibi pour le discours politique, sans guère d'effets sur les apprentissages des élèves. L'évaluation est donc un élément constitutif de l'enseignement/apprentissage d'une langue, en particulier lorsqu'elle est «formative». Mais il importe de rester particulièrement prudent dès qu'on vise à établir un bilan des connaissances et compétences des élèves – une évaluation *sommative* –, d'en user avec modération et rigueur (de Pietro *et al.*, à paraître).

Le but de notre contribution, liée à des travaux réalisés pour la *Conférence*

Jean-François de Pietro
& Murielle Roth |
IRDP Neuchâtel

Jean-François de Pietro est collaborateur scientifique à l'Institut de recherche et de documentation pédagogique (IRDP) et membre du comité de rédaction de la revue *Babylonia*.

Murielle Roth est collaboratrice scientifique à l'IRDP, à Neuchâtel. Chercheuse en didactique du français, ses travaux actuels portent sur l'évaluation des apprentissages des élèves, en lien avec l'analyse du *Plan d'études romand* (PER).



PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > [SCHEDE 6 E 20C](#)

L'élève ne peut pas tout retenir à l'écoute, et il ne peut pas non plus y rechercher après-coup une information qui lui aurait échappé; dès lors, il paraît indispensable de faire porter prioritairement l'évaluation sur les éléments «cruciaux» du texte — éléments qui ne peuvent être définis qu'en fonction de la finalité de l'écoute, elle-même déterminée en terme de tâche.

intercantonale de l'instruction publique de la Suisse romande et du Tessin (CIIP), est de mettre en évidence – sur la base de difficultés rencontrées dans ces travaux – la nécessité de fonder l'évaluation sur une réflexion didactique solide et critique qui permette, en lien avec la manière de «rendre visibles» les connaissances et compétences, de savoir ce qu'on veut/peut mesurer (cf. Roth et al., 2014).

Afin de mettre en discussion ces fondements didactiques, nous nous appuyons sur des observations concernant l'évaluation de la compréhension orale (désormais CO) dans la langue de scolarisation – le français en l'occurrence¹. Nous nous interrogeons sur ce qui, à l'oral, doit/peut faire l'objet d'une évaluation, en suggérant notamment que l'évaluation gagnerait à mieux définir et clairement distinguer des dimensions différentes de la compréhension (orale), liées à des types de tâches à même de les rendre visibles. Nos réflexions sont basées sur des données concernant la langue première, de scolarisation, et non les L2. Nous pensons cependant que certaines de nos observations peuvent intéresser également les enseignant-e-s et didacticien-ne-s des L2 dans la mesure où elles mettent en évidence des dimensions constitutives de la compréhension, dépassant en partie la seule compréhension orale et la L1, qui ont des conséquences sur la manière d'évaluer.

2. L'évaluation de la compréhension orale

Il est aujourd'hui reconnu que la CO, en tant que l'une des quatre composantes de la compétence langagière, doit être enseignée (cf. PER)². Cela n'est guère discuté pour les L2. En ce qui concerne la L1, c'est moins évident et la CO reste bien souvent encore un parent pauvre de l'enseignement (Dolz & Mabilard, à paraître). Il a d'ailleurs été observé que, en matière de compréhension, c'est bien souvent l'évaluation – réalisée en général sous forme de questionnaire – qui tient lieu d'enseignement (Rodriguez-Smith & Goasdoué, 2015). Parmi les diverses manières possibles d'évaluer la CO, le questionnaire semble bien constituer la forme la plus fréquemment utilisée. En examinant divers exemples, nous avons ainsi été amenés à nous interroger sur la manière dont ces questionnaires conçoivent la CO et l'évaluent: sur quoi portent les questions posées? Que permettent-elles *vraiment* de dire de la compétence de CO

des élèves? Quels sont, finalement, les éléments à évaluer dans la CO?

Cet examen nous a amenés aux constats suivants:

1. La CO ne fait pas systématiquement l'objet d'une évaluation.

2. Lorsqu'elle est testée, sous forme de questionnaire portant sur un texte entendu, on constate souvent que les questions portent sur des objets de niveaux et/ou de nature très différents, sans que ces différences soient rendues explicites, par exemple en séparant le questionnaire en parties distinctes. Par exemple, elles mêlent fréquemment des questions «de détail», de compréhension locale, et des questions portant sur ce que le nouveau *Plan d'étude romand* (CIIP, 2010; désormais PER) nomme parfois «le sens général» du texte³.

3. On observe essentiellement trois types de questions:

a) Des questions qui n'ont pour ainsi dire aucun rapport ni avec le texte – qui n'est alors qu'un *prétexte* – ni avec l'oral et qui visent à mesurer des connaissances linguistiques générales, par exemple la maîtrise d'un objectif grammatical (vocabulaire, conjugaison, accords...).

b) Des questions «de détail», de portée locale, portant sur des aspects du texte qui ne sont pas indispensables mais qui participent néanmoins à la compréhension du texte.

c) Des questions qui, de diverses manières, nous semblent toucher à la compréhension globale du texte. Celles-ci peuvent porter sur des éléments divers: éléments du contexte qui définissent la situation de communication, éléments cruciaux d'un récit ou renvoyant à l'interprétation qu'on peut en faire, arguments décisifs d'un débat, etc.

Le premier type de question n'a rien à voir avec la CO et devrait être clairement séparé des autres. Les deux derniers types nous ont amenés à réfléchir à ce qu'on doit évaluer. Ils portent en fait sur des niveaux différents de la compréhension: ce qu'on pourrait considérer comme la *compréhension d'aspects locaux du texte* pour (b) et la *compréhension globale du texte* pour (c). Ces deux formes de compréhension peuvent selon nous être légitimes. Mais il nous semble important de les distinguer clairement, de savoir ce que l'on évalue: des ressources langagières d'un côté, qui relèvent de la compétence linguistique générale, et la compétence de compréhens-

¹ Nous avons principalement analysé des épreuves communes ayant pour but d'évaluer les acquisitions des élèves à l'échelle d'un canton.

² Contrairement à ce qu'on observe dans l'enseignement des L2, la *compétence interactionnelle* n'est pas considérée en tant que composante autonome en L1; elle est toutefois, à des degrés divers, prise en compte de manière intégrée lorsque l'enseignement et l'évaluation concernent des genres interactifs tels le débat.

³ Selon le PER (cf. lexique Langues), le texte, oral ou écrit, peut être d'ampleur et de nature très diverses. Il recouvre toute «trace matérielle autonome d'une activité langagière, orale ou écrite, qui fait sens...».

sion textuelle de l'autre – qui mobilise des ressources de divers niveaux et de diverses natures, linguistiques mais aussi encyclopédiques, logiques... Et il nous semble surtout important, si on souhaite évaluer vraiment la compétence de compréhension d'un texte, d'inclure dans les moyens d'évaluation utilisés des questions à même de rendre visible ce type de compréhension: choisir le bon résumé, reconstituer la chronologie d'un récit, définir la position argumentative d'un débatteur, etc. C'est, évidemment, ce qui est le plus difficile et... qui est parfois laissé de côté.

3. Spécificités de l'oral et contraintes pour l'évaluation

La distinction entre ces différentes formes de compréhension vaut aussi bien pour l'écrit que pour l'oral. Cependant, en raison même des spécificités de l'oral, elle apparaît en quelque sorte exacerbée pour l'oral, en particulier dans des épreuves communes sommatives où l'on prétend mesurer ce qui est présenté comme un objectif d'apprentissage de haut niveau dans le PER, à savoir la compétence «comprendre des textes oraux variés propres à des situations de la vie courante» (PER, 2010, 42): qu'entend-on ici par «comprendre»?...

Deux des caractéristiques de l'oral s'avèrent selon nous décisives:

- > L'ancrage situationnel généralement plus marqué dans les documents oraux, qui tend à imposer plus encore qu'à l'écrit une claire détermination de certains paramètres de l'écoute; autrement dit, alors que, dans une perspective scolaire, on peut *dans une certaine mesure* décontextualiser les textes écrits étudiés, cela nous paraît plus problématique à l'oral.
- > Le caractère éphémère et la linéarité de la parole, qui rendent bien plus difficile tout «retour» sur le texte entendu, toute mise à distance et décontextualisation de celui-ci.

Selon nous, l'évaluation de la compréhension d'un texte à l'oral doit absolument tenir compte de ces spécificités, notamment en définissant une *perspective d'écoute*, une finalité liée à une tâche⁴ qui oriente le questionnement du texte – par l'écouter-élève, d'abord, et par l'évaluation censée rendre visible cette «compréhension orientée» de l'élève. Pour le dire plus simplement, l'élève ne peut pas/ne doit pas tout retenir à l'écoute d'un texte

oral, et il ne peut pas non plus y revenir, rechercher après-coup une information qui lui aurait échappé⁵; dès lors, il paraît indispensable, plus encore qu'à l'écrit, de faire porter prioritairement l'évaluation sur les éléments «cruciaux» du texte – éléments qui ne peuvent être définis qu'en fonction de la finalité de l'écoute, elle-même déterminée en terme de tâche. De fait, à l'oral, le comportement «expert» de l'auditeur compétent consiste précisément à sélectionner de manière drastique – du fait de nos possibilités limitées de mémorisation de la parole – les éléments qui lui permettent d'accomplir la tâche d'écoute à laquelle il est confronté. Ceci, d'ailleurs, est plus vrai encore lorsque la communication se déroule dans une L2. Il apparaît ainsi nécessaire de s'interroger sur ce qu'on vise par les questions qu'on soumet aux élèves, sur ce qu'on attend par exemple qu'ils montrent lorsqu'on les interroge à propos d'un détail qui n'aurait pas d'impact sur cette compréhension ciblée.

4. A la recherche du «sens général» du texte oral...

Le test ci-après, utilisé à titre d'illustration, nous semble aller dans une bonne direction, même s'il est certainement perfectible. Elaboré en relation étroite avec un «modèle de compétence», il vise à cerner cette compétence selon ses diverses composantes⁶.

La tâche est présentée ainsi:

Vous allez écouter une émission enregistrée d'une durée de dix minutes environ, où il est question du loup. Vous ne l'entendrez qu'une seule fois. Écoutez attentivement, afin de saisir le sens général, mais aussi de comprendre les interventions et les prises de position des différents participants. Après l'écoute, vous pourrez lire les questions et y répondre. Lisez attentivement les consignes indiquant comment vous devez répondre. Pendant l'écoute, vous ne pouvez pas prendre de notes.

Ce test, qui nécessiterait une analyse détaillée, comporte 13 questions basées sur cette consigne. Certaines concernent le vocabulaire, comme la question 8 qui ne semble pas indispensable pour comprendre le débat mais dont le contenu participe tout de même à la compréhension du texte: *“Jean-Lou Déforêt affirme qu'il faut arrêter de “faire porter le chapeau” au loup. Que signifie ici cette expression?”* Mais la plupart, de diverses manières, visent des éléments «essentiels» à la compréhension du texte – qui ressortent d'autant plus clairement qu'on les définit en fonction d'une perspective d'écoute et du genre dont ce document relève, en

4 Il importe à ce propos de souligner que ces tâches peuvent être très diverses: recherche d'informations, accomplissement d'une action, interprétation d'un texte, etc.

5 C'est d'ailleurs pour cette raison qu'on recourt parfois, un peu artificiellement, à une deuxième écoute dans l'évaluation de l'oral, en particulier en L2.

6 Ce test provient des épreuves HarmoS (Langue de scolarisation, 6, H07) testées lors de l'élaboration des standards de compétence. Pour une présentation du modèle de compétence: http://edudoc.ch/record/96790/files/grundkomp_schul-sprache_f.pdf

l'occurrence un débat –, et portent sur ses finalités (*Quel est le but de cette émission?*), son contexte (*Quelle image du loup, qu'on trouve dans de nombreuses légendes, a-t-on aujourd'hui tendance à oublier?*), son contenu comme par exemple la question 2, qui demande d'identifier la question – nécessairement polémique – qu'on trouve à la base de tout débat, ou la question 10, qui concerne une des positions qui structurent le débat:

2. Sur quelle question porte le débat? Une seule bonne réponse.

- Faut-il modifier la convention de Berne?
- Faut-il réintroduire le loup?
- Faut-il abattre le loup?
- Faut-il créer des jardins zoologiques?
- Le loup est-il de retour dans notre pays?

10. Parmi les propos suivants, lequel exprime le mieux la position d'Amédée Vaurez, le berger? Une seule bonne réponse.

- Il me semble qu'il n'est pas possible que le loup cohabite avec le mouton
- Je pense que le loup peut très bien cohabiter avec le mouton
- Il est évident que le loup ne peut pas cohabiter avec le mouton
- Le loup n'a pas d'intérêt à cohabiter avec le mouton

Comme on le voit, il est possible d'élaborer des questionnaires de CO clairement orientés vers des éléments cruciaux du texte, en limitant les questions de compréhension de détail, appelées à être «oubliées» dans le processus expert de compréhension orale. Dans l'exemple examiné, la contextualisation gagnerait à être précisée et la finalité de l'écoute explicitée car l'évaluation aurait plus de sens encore si les questions posées aux élèves étaient définies en fonction d'une tâche explicite, liée au genre dont relève le texte⁷. Par exemple, il pourrait être suggéré ici, comme c'est souvent le cas quand on écoute un débat, que l'élève doit se forger sa propre opinion quant à une éventuelle réintroduction du loup. Certes, d'autres questions peuvent porter sur des détails du débat ou, même, sur des aspects de fonctionnement de la langue mais elles devraient à notre avis faire l'objet d'un traitement clairement séparé. *Comprendre un texte oral* (ou écrit) et en dégager le «sens général» dépend de la tâche et du genre textuel concerné. D'une certaine manière, le «sens général» n'existe pas pour lui-même: il n'existe que des «sens» en contexte. Cette notion de «sens général», qu'on trouve notamment dans le PER, gagnerait à être concrétisée dans des tâches ayant une finalité autre que de répondre à des questions – questions qu'elles permet-

traient d'ailleurs, en retour, d'orienter de façon cohérente et qui donneraient à voir si les élèves ont réussi la tâche grâce à leur compréhension orientée du texte. Par ses spécificités, la CO plus encore que la CE met ainsi en évidence ce qu'est la compréhension d'un texte et comment on devrait l'évaluer.

Cet exemple nous permet de faire un certain nombre de constats qui peuvent être élargis à l'ensemble de nos analyses. De manière générale, il en ressort une conception de la CO selon laquelle l'élève qui «comprend» est celui qui sait aller chercher dans le texte des informations, souvent disparates, souvent «locales», et qui n'ont pas forcément d'impact sur la compréhension globale du texte – en contradiction par conséquent avec l'objectif qui apparaît en tout premier dans le PER (*cf. supra*).

5. Une modélisation de la CO pour l'évaluation

La question de l'évaluation est à l'évidence des plus complexes, notamment pour la CO. Nous avons vu que, s'il est facile de proposer des questions qui concernent des aspects locaux de la compréhension, il s'avère bien plus difficile de concevoir des tâches évaluatives qui rendent visible la compréhension par les élèves du «sens général» du texte – notion présente dans le PER –, en tenant compte du genre dont il relève et de la perspective d'écoute définie par ces tâches.

Cette notion de *sens général* doit, selon nous, être clarifiée, notamment pour souligner l'importance d'une contextualisation des tâches. Elle est fondamentale, non seulement parce qu'elle figure en tête des objectifs dans le PER mais, aussi, parce qu'elle nous interroge sur ce que nous voulons évaluer: des compétences, de haut niveau, ou des ressources langagières de toute nature, plus ou moins locales? Dans le but d'explicitier les enjeux sous-jacents à une telle distinction, nous avons tout d'abord tenté de clarifier le statut des contenus figurant dans le PER: visées prioritaires, objectifs d'apprentissage, progression des apprentissages, attentes fondamentales... Il nous est ainsi apparu que, là aussi, des contenus de niveaux très différents apparaissaient ensemble, sans que leur statut soit clairement distingué. Nous nous sommes attelés, dès lors, à élaborer – en complément au PER – qui est un référentiel conçu pour l'enseignement/apprentissage et non pour l'évaluation –

⁷ En l'état, l'élève ne sait pas vraiment – si ce n'est, bien sûr, pour répondre aux questions – pourquoi il écoute ce débat, ce qu'il y cherche... On peut également s'interroger sur le fait que les élèves ne peuvent pas prendre de notes pendant l'écoute.

⁸ Pour une présentation détaillée du modèle *cf. Roth et al., 2014.*

une modélisation de la compréhension orale (et écrite) qui met en relation, de manière privilégiée mais non nécessairement exclusive, des objets figurant dans le PER avec des formats d'activités, en leur conférant en même temps un statut dans la perspective de l'évaluation.

Cette modélisation a pour point de départ l'activité de l'élève⁸. Dans la perspective des compétences, l'accès au «sens général» dépend de ressources mobilisées pour comprendre les éléments cruciaux du texte, en lien à une tâche qui définit la finalité de l'écoute, elle-même étroitement liée au genre dont relève le texte. Dans notre exemple, en contexte scolaire, d'un débat, autrement dit d'un texte «qui argumente», on va généralement l'écouter pour se forger une position ou pour comprendre la position de quelqu'un. Et on va devoir saisir la question polémique qui, explicitement ou non, est à la base du débat, le statut et les prises de position des intervenants, certains des arguments auxquels ils recourent, etc. Pour ce faire, on devra maîtriser diverses ressources langagières telles que: l'orientation argumentative portée par certains connecteurs, les modalisations d'énoncés qui colorent les prises de position, etc.

Ainsi, si nous voulons cerner la CO en tant que compétence, nous considérons, en nous appuyant sur notre modélisation, qu'il faut viser en priorité à poser des questions qui touchent aux éléments *constitutifs* de la tâche – car c'est en fait la tâche, telle qu'elle est définie dans la consigne, et le genre de texte, sa visée en particulier, qui vont déterminer alors ce que signifie «comprendre le texte». «Bien écouter» un texte qui argumente, c'est bien en «[identifier] la visée et le sujet», y «[identifier et reformuler] quelques arguments, une opinion (...)» (PER, p. 29) en lien à un projet d'écoute.

6. L'évaluation, de la politique éducative à la didactique

L'évaluation comporte une composante didactique et une composante politique. En effet, c'est bien la société, par les autorités chargées de la politique éducative, qui décide finalement ce que les élèves doivent savoir et savoir faire. Mais il est nécessaire aussi de prévenir toute dérive purement techniciste et économiste et de garder l'évaluation – notamment les questions de mesure – *sous contrôle didactique*. Ce qui signifie savoir ce qu'on fait lorsqu'on évalue telle chose de telle manière.

Il importe ainsi, finalement, que l'évaluation – qu'elle soit conçue dans une perspective formative ou sommative, qu'elle porte sur les acquis d'apprentissage ou qu'elle serve à vérifier l'efficacité et l'efficacité du système scolaire – soit adaptée aux objectifs définis par les autorités et formulées dans un référentiel tel que le PER; mais il importe en même temps qu'elle prenne appui sur une réflexion didactique approfondie à propos des objets à évaluer et des formats d'activité à utiliser pour cette évaluation (de Pietro, 2015). Il en va assurément de même pour l'évaluation en L2. Car, là aussi, il s'agit de savoir ce qu'on vise, de ne pas se contenter d'évaluer des éléments épars, de nature et de statut différents. Les objectifs doivent être plus modestes, les textes, oraux ou écrits, qu'il s'agit de comprendre peuvent être plus simples, certes, mais cela ne doit pas faire oublier que la visée de l'apprentissage est bien d'accéder au sens des textes auxquels les élèves seront confrontés dans leurs pratiques langagières futures et non simplement de maîtriser des éléments formels sans liens entre eux. Il s'agit donc, dans tous les cas, de trouver un équilibre entre l'évaluation de ressources et l'évaluation de compétences plus globales, telles que la compétence à comprendre des textes oraux.

Références bibliographiques

Conférence intercantonale de l'Instruction publique de la Suisse romande et du Tessin (CIIP). (2010). *Plan d'études romand (PER)*. Neuchâtel: CIIP.

[Page Web]. Accès: <http://www.plandetudes.ch/> (Page consultée le 9.3.2016).

De Pietro, J.-F., Roth, M. & Sánchez Abchi, V. (à paraître). Vers un référentiel pour l'évaluation: choix politiques et enjeux didactiques. In: E. Falardeau et al., *L'évaluation en classe de français, outil didactique et politique*. Recherches en didactique du français, Volume 8.

De Pietro, J.-F. (2015). Formats d'activités et statut des objets à enseigner/évaluer. *Formation et pratiques d'enseignement en questions*, 19, 43-53.

Dolz, J. & Mabilard, J.-P. (à paraître). Enseigner la compréhension orale: un projet d'ingénierie didactique. In: J.-F. de Pietro, C. Fisher & R. Gagnon, *Approches didactiques de l'oral*. Recherches en didactique du français, Volume 9.

Rodriguez-Smith, S. & Goasdoué, R. (2015). L'évaluation de la compréhension par des questionnaires: des questions pour apprendre, pour comprendre, pour attester? In: *L'évaluation à la lumière des contextes et des disciplines, Actes du colloque de l'ADMEE-Europe*. http://admee2015.sciencesconf.org/conference/admee2015/admee15_Actes_du_colloque4incp.pdf

Roth, M., De Pietro, J.-F. & Sánchez Abchi, V. (2014). Comment évaluer la lecture en référence au Plan d'études romand? *Forumlecture.ch*, 3.

ANALISI DI PROVE DI ASCOLTO IN ESAMI DI CERTIFICAZIONE ITALIANI (LIVELLO B2)

Italia weist bezüglich Sprachzertifizierung eine Besonderheit auf: Gleich vier Zertifikatsprüfungen sind vom italienischen Bildungsministerium anerkannt: CELI, CILS, PLIDA und .IT. Diese Vielfalt bietet Gelegenheit, die Prüfungen zu vergleichen und die einzelnen Items auf ihre Validität hin zu überprüfen.

161 Maturanden aus Schweizer und Südtiroler Gymnasien haben die vier Hörverständnisübungen auf dem Niveau B2 durchgeführt und zudem die Prüfungen bezüglich ihrer Attraktivität bewertet. Die folgende, eigens für Babylonia durchgeführte Untersuchung legt die Resultate der vier Prüfungen dar und zeigt die Schwachstellen der jeweiligen Prüfungen auf. Die Lektüre dieses Beitrags empfiehlt sich besonders für jeden, der selbst Multiple-Choice-Prüfungen durchführt und um angemessene und unmissverständliche Fragestellungen und Distraktoren ringt.

Die Prüfungen sind auf der Babylonia-Homepage abrufbar:

babylonia.ch > Nummer 2/2016 > Artikel Torresan

● Paolo Torresan | Venezia

Paolo Torresan ha svolto attività di insegnamento presso l'Università dello Stato di Rio de Janeiro (*Visiting Professor*), il Santa Monica College (*Fulbright Visiting Scholar*) e l'Università di Catania (sede di Ragusa). Ha svolto attività di ricerca presso l'Università Ca' Foscari di Venezia e la Lancaster University.



1. Il sistema plurale delle certificazioni italiane¹

I *proficiency test* nell'ambito dell'italiano come lingua straniera riconosciuti dal Ministero degli Affari Esteri italiano sono rappresentati dalle 4 certificazioni ufficiali:

- > CELI, Università per Stranieri di Perugia
- > CILS, Università per Stranieri di Siena
- > PLIDA, Società Dante Alighieri
- > .IT, Università degli Studi Roma Tre

Altre certificazioni, come gli esami AIL dell'Accademia di Firenze², o gli esami CLIDA³ del Centro Linguistico Italiano Dante Alighieri (con sedi a Firenze e Roma), pur essendo implementate in alcuni contesti (es. in Svizzera nel primo caso; in vari paesi, nel secondo), ricadono fuori dal riconoscimento ministeriale.

Dinanzi a questa pluralità di proposte, domande ricorrenti sono:

- > Perché quattro certificazioni e non una, per una lingua, qual è l'italiano, che non riscuote un interesse internazionale al pari dell'inglese?
- > Esistono dei confronti tra una certificazione e l'altra?

Alla prima domanda non siamo in grado di rispondere, visto che comporta uno sguardo di tipo storico e sociologico esterno alle nostre competenze. Alla seconda rispondiamo che non esistono studi comparativi in termini di risultati, quanto descrizioni, alcune delle quali molto raffinate (Novello, 2014). C'è da riconoscere, ad ogni modo, che una comparazione in termini di risultati tra le 4 certificazioni non è praticabile in sé e di per sé. Si dovrebbe supporre che uno stesso campione si sottoponga alle varie prove; i tempi di gestione di una impresa del genere sono enormi (almeno un giorno per certificazione); inoltre esiste un 'effetto-accumulo', tale per cui, passando dalle prove della prima certificazione alle prove della quarta, il candidato acquisirebbe maggiore competenza (o perlomeno guadagnerebbe in *expertise* nello svolgimento di esercizi simili) e tale fattore inciderebbe a favore della certificazione somministrata per ultima.

In luogo di una comparazione 'in blocco', esiste però la possibilità di un'analisi parallela di singole prove omologhe. Tale indagine ci risulta non essere mai sta-

PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
 ARCHIVIO TEMATICO > [SCHEDA 6](#)

ta eseguita nell'ambito dell'italiano ed è l'oggetto del nostro saggio.

2. Item Analysis parallelo: presentazione della ricerca

In seno a tre corsi di formazione, svoltisi in tempi assai ravvicinati (dicembre 2015-gennaio 2016), tra la provincia di Bolzano (corso promosso dall'Ispettorato delle scuole di lingua tedesca) e la Svizzera (corsi tenuti rispettivamente presso l'Università di Friburgo e l'Alta Scuola Pedagogica di Zurigo), abbiamo raccolto, grazie alla disponibilità di colleghi, le risposte di un campione omogeneo (studenti maturandi) a prove omolo-

ghe (ascolto, livello B2, quesiti a scelta multipla)⁴, tratte dalle 4 certificazioni di italiano, liberamente consultabili *online*⁵. Il lettore può scaricare le prove dai siti riportati nella sitografia. Ai candidati è stato inoltre somministrato un questionario attraverso il quale, oltre a segnare le preferenze per le singole prove attraverso una scala Likert, ciascuno era libero di esprimere i propri giudizi in merito alla sperimentazione (§ 2.4).

Nella tab. 1 rappresentiamo le caratteristiche delle prove.

Tab. 1. Caratteristiche generali delle prove

	CELI	CILS	PLIDA	.IT
Sessione	estiva 2007	estiva 2012	non dich.	non dich.
Minutaggio soundfile, senza istruzioni	9'12"-12'46"	10'56"-14'09"	28'03"-31'42"	1'05"-3'45"
Durata	3'34"	2'13"	3'45"	2'50"
Numero di ascolti	2	2	2	2
Tipologia	intervista	intervista	monologo	intervista
Dominio	sport	un'offerta formativa	storia dell'urbanistica	un'attività commerciale
Numero item	5	7	5	5
Introduzione al tema	✓			
Item-esempio	✓			✓
Layout chiaro ⁶	✓	✓		
Testo autentico ⁷	✓		✓	✓
Citazione della fonte			✓	✓
Ortoepia naturale	✓	✓	✓	✓
Testo interessante	✓		✓	✓
Testo chiaro	✓		✓	✓
Punteggio solo positivo		✓		non dich.
Trascrizione		✓		
Soluzioni	✓	✓	✓	✓

1 Ringraziamo Erika Saccuti e Maria Simona Morosin per la lettura e il confronto. Ringraziamo i colleghi che ci hanno aiutato nella raccolta dei dati.

2 Gli esami ALL, nell'ambito della valutazione delle abilità di comprensione, constano soprattutto di prove di *vero/falso*, senza peraltro provvedere ad alcuna misura che consenta di arginare l'arbitrarietà che il *vero/falso* comporta. Si tratta, quindi, di un esame che presenta margini di criticità sin dalla scelta dei *format*.

3 Gli esami CLIDA, nell'ambito della valutazione delle abilità di comprensione, presentano soprattutto *domande aperte*, con numerose ingenuità nella confezione degli *item*, come l'*interdipendenza*, tale per cui si può rispondere a una domanda leggendo una successiva (es. prova CLIDA P3 2014: domanda 1 – "Dove è stato il Campionato Mondiale di Ciclismo 2013?"; Domanda 4 – "Quanti «fan» sono stati presenti lungo le strade tra Firenze e Fiesole?") o domande a cui è facile rispondere grazie a preconcoscienze (es. prova CLIDA P3 2011: domanda 3 – "Dov'è nato il melodramma?").

4 Una leggera differenza è data dal fatto che la prova CILS consta di 7 *item* anziché 5, come le altre; ma ciò non crea difficoltà in termini di comparazione.

5 La prova tratta dal PLIDA corrisponde al vecchio formato; il centro certificatore sta rivedendo l'impianto della certificazione; per ora sono uscite le specifiche del nuovo formato del B1 (Cardillo, Vecchio & Menzinger, 2015), nel corso del 2016 usciranno quelle relative al B2 e al C1.

6 Si tratta di un giudizio soggettivo. Nel caso del PLIDA, la distribuzione della prova su due fogli (fronte/retro) rende poco agevole, secondo noi, l'esecuzione. Nel caso di .IT l'*item*-esempio è stato inteso, da una buona parte del campione, come facente parte della prova.

7 Anche questo è un nostro giudizio: il testo pare essere studiato a tavolino.

Nella tab. 2 riportiamo il tempo a disposizione per leggere i quesiti secondo le indicazioni fornite nei file.

Le *research question* che ci siamo posti sono state:

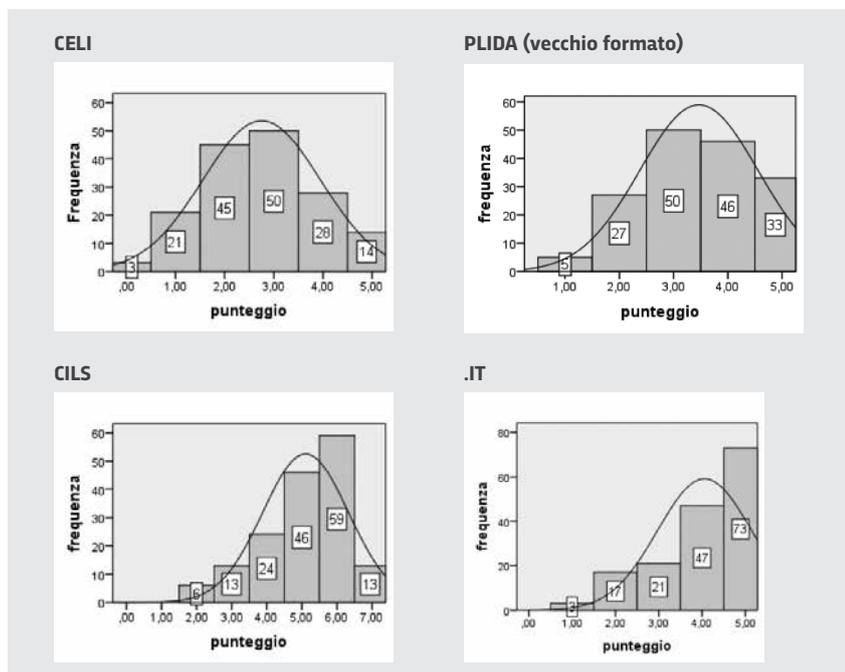
- > I risultati del campione variano *significativamente* da prova a prova?
- > Esistono casi evidenti di sotto- e/o sovracalibrazione?
- > Esistono casi di cattiva scrittura degli item?
- > Qual è il *feedback* dei candidati?

Tab. 2. Tempi di lettura dichiarati nei soundfile prima, tra e dopo gli ascolti

Consideriamo ciascuna nei paragrafi a seguire.

	prima del primo ascolto	tra il primo e il secondo ascolto	dopo il secondo ascolto	tot.
CELI	20" (60" per tutte e tre le prove di ascolto)	25"	35"	1'20"
CILS	-	60"	120"	3'
PLIDA	60"	-	60"	2'
.IT	indicazioni non presenti nell'audio			

Fig. 1. Frequenza dei punteggi



2.1. Variazione tra le prove

Da un'analisi degli istogrammi alla fig. 1. si evince che i risultati variano significativamente da prova a prova.

Agli estremi abbiamo due prove:

- > .IT, la più facile, con un profilo fortemente disallineato a destra
- > CELI, con un profilo compatibile con una distribuzione normale.

La relativa facilità della prova .IT è evidente se si considera la percentuale degli studenti che hanno raggiunto l'eccellenza, e cioè che hanno risposto correttamente a 5 item su 5: 45,3%, cifra ben superiore rispetto a quelle relative a quanti hanno raggiunto l'eccellenza nelle altre prove (tab. 3).

La relativa difficoltà della prova CELI, come dicevamo, è immediatamente inferibile considerato il profilo della curva che presenta una distribuzione pressoché normale (con un leggerissimo disallineamento a destra). Si tratta di un profilo che non ci si attende con un campione rappresentativo come il nostro, cioè con studenti che si stanno preparando sistematicamente ad affrontare una prova di questo livello.

Il profilo della prova CILS è simile a quello della prova .IT, salvo presentare una percentuale molto bassa in riferimento all'eccellenza (solo 13 candidati hanno risposto a tutti gli item: 8,1%). Ciò si spiega con la forte sovracalibrazione di un item (#3), come avremo modo di vedere nel paragrafo successivo.

Nel complesso pare che alla prova PLIDA corrisponda il profilo più armonico e atteso.

2.2. La calibrazione degli item

Nella tab. 3 rappresentiamo le percentuali di studenti che hanno risposto correttamente ai singoli item (*Facility Value*, d'ora in poi *FV*).

A quali valori attenerci, sotto i quali o oltre i quali si danno, rispettivamente, casi di sotto- e di sovracalibrazione? Nella letteratura i pareri sono discordanti, con oscillazioni significative (per un prospetto: Green, 2013). A nostro giudizio, un ente certificatore, nella più flessibile delle ipo-

Tab. 3. Facility Value
n=161

tesi, si può orientare attorno all'intervallo 35-90, e cioè ritenere che con un $FV \leq 35$ vi sia *evidente* sovracalibrazione (vale a dire che se solo un candidato su 3 indovina la risposta, l'item è *evidentemente* difficile per il livello *target*), così come, al contrario, con un $FV \geq 90$ vi è *evidente* sottocalibrazione (se oltre 9 studenti su 10 indovinan la chiave, l'item è *evidentemente* facile in riferimento al livello *target*). In termini di ricerca, sempre nell'ottica di un contesto certificatorio, suggeriamo valori più ristretti: 40-85; e cioè con item con

- > $FV \leq 40$ abbiamo una sovracalibrazione che merita di essere indagata
- > $FV \geq 85$ abbiamo una sottocalibrazione che, al pari, costringe a un'analisi accurata.

Alla domanda, dunque, se si diano *item* sotto- o sovracalibrati, la risposta è affermativa. Abbiamo, in effetti:

- > due *item* sovracalibrati nella prova CELI
- > due *item* orientati in maniera opposta nella prova CILS, uno sotto- e uno sovracalibrato
- > un *item* sottocalibrato nella prova PLIDA
- > tre *item* sottocalibrati nella prova .IT

In alcuni casi, come avremo modo di appurare, i casi di sotto- e sovradimensionamento si spiegano per via di problemi legati alla confezione degli item stessi.

2.3. La qualità degli item

La tab. 4 riprende la tab. 3: riportiamo ancora una volta gli FV dei singoli item. In aggiunta, rappresentiamo con un doppio asterisco ed evidenziati in grassetto i casi di cattiva fattura degli item. Come anticipato, in alcuni frangenti, l'errato dimensionamento è attribuibile a una scrittura poco sorvegliata dei quesiti.

In sintesi, più casi richiederebbero una riformulazione dei quesiti e un conseguente riadattamento delle prove:

- > CELI, tre casi: *item* ## 1,2,4
- > CILS, cinque casi: *item* ## 1,2,3,6,7
- > PLIDA, un caso: *item* # 2
- > .IT, tre casi: *item* ## 1,2,3

	1	2	3	4	5	6	7
CELI	55,9	27,3	72,7	37,9	82,6		
CILS	75,8	76,4	37,9	71,4	83,9	93,2	72
PLIDA	88,2	54,7	60,9	73,9	70,2		
.IT	90,1	88,8	71,4	67,2	87,6		

Tab. 4. Item di cattiva fattura
n=161

	1	2	3	4	5	6	7
CELI	55,9**	27,3**	72,7	37,9**	82,6		
CILS	75,8**	76,4**	37,9**	71,4	83,9	93,2**	72**
PLIDA	88,2	54,7**	60,9	73,9	70,2		
.IT	90,1**	88,8**	71,4**	67,2	87,6		

Tab. 5. Mappa della criticità.

	CELI	CILS	PLIDA	.IT
Distrattore parzialmente vero	1,2,4			
Assenza di una chiave		3		3
Ambiguità della chiave			2	
Ambiguità del testo di riferimento		1		
Risalto dalla chiave		6		
Dipendenza tra gli item		2		2
Inadeguata distribuzione tra gli item		7		1

Ancora una volta, la prova PLIDA pare essere la meno approssimativa, mentre la prova CILS comporta il numero maggiore di criticità.

Nella tab. 5 illustriamo i tipi di criticità. I numeri riportati fanno riferimento ai singoli item delle prove (es. il numero 3 riferito alla prova CILS rappresenta il terzo item dell'esercizio).

Illustriamo a seguire la natura delle criticità:

- > *Distrattore parzialmente vero*: si tratta di un distrattore che, per via inferenziale, può essere colto come possibile – non è cioè totalmente smentito dal testo.
- > *Assenza di una chiave*: nessuna tra le opzioni presentate dal certificatore è

perfettamente sovrapponibile alle informazioni veicolate dal testo; di conseguenza non vi è una risposta corretta.

> *Ambiguità della chiave*: la chiave è formulata in maniera tale da ammettere margini di interpretazione.

> *Ambiguità del testo di riferimento*: il testo presenta aspetti di incoerenza logica (contraddizioni) o tematica (annuncia tratterà un certo argomento, ma passa a trattare tutt'altro).

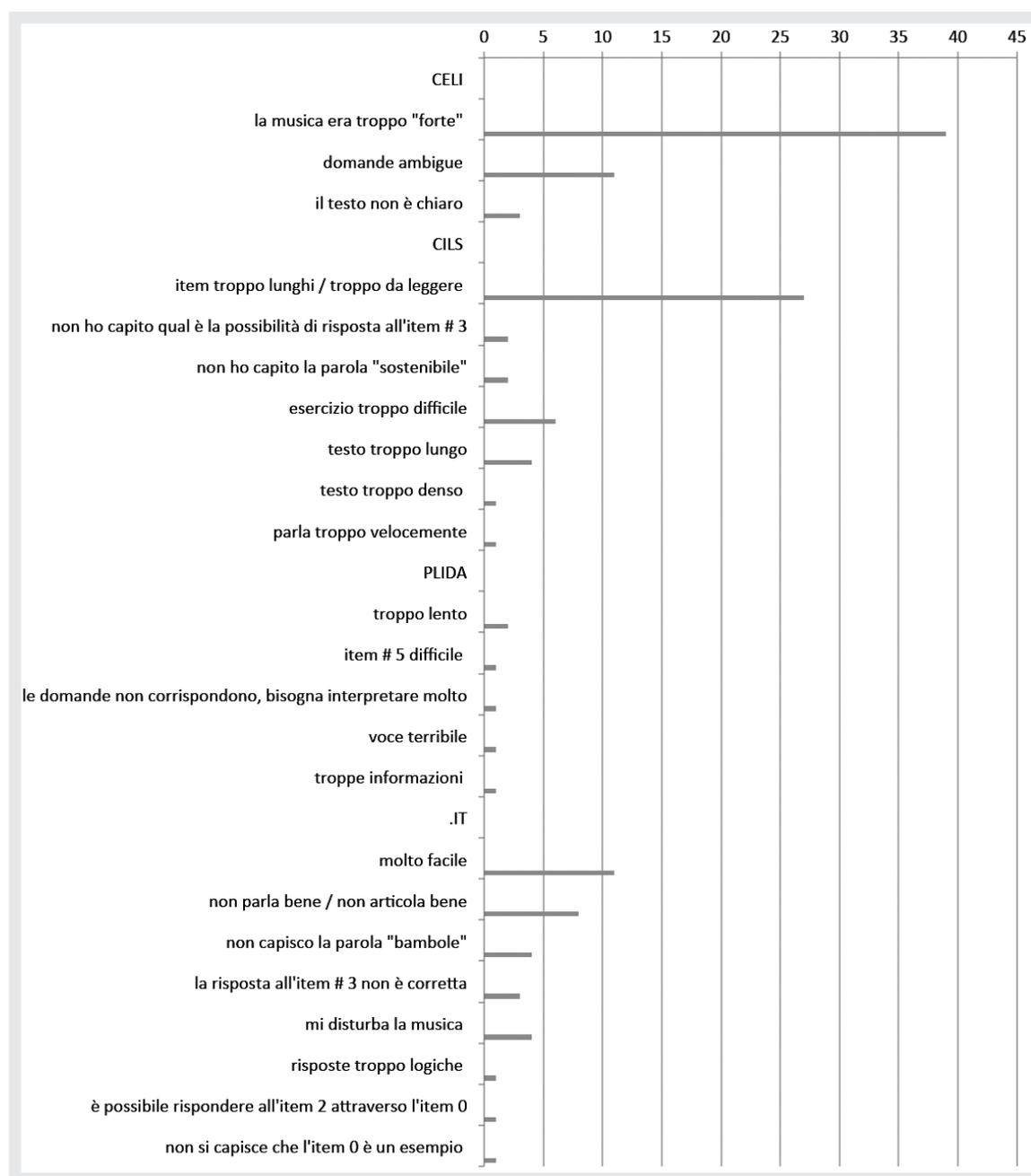
> *Risalto della chiave*: la chiave si distingue graficamente dai distrattori perché più lunga e/o più elaborata; è quindi

facile la sua individuazione da parte del candidato.

> *Dipendenza tra gli item*: si può rispondere a un item attraverso la presa visione di un item precedente o di uno successivo.

> *Inadeguata distribuzione degli item*: gli item rimandano a informazioni troppo ravvicinate tra loro, se non addirittura 'incastrate' l'una nell'altra; oppure l'ordine degli item non rispetta la progressione delle informazioni presenti nel testo, o infine gli item si riferiscono a informazioni trasmesse nei primi se-

Fig. 2. Osservazioni critiche di una parte del campione



condi del brano, sicché l'ascoltatore fa fatica a individuarle.

Per il lettore interessato a una disamina attenta, caso per caso, rimandiamo alla versione estesa online di questo nostro articolo, sul sito della rivista.

2.4. Il *feedback* del campione

Mentre i dati ottenuti tramite la scala Likert circa il grado di preferenza non risultano significativi, data la dispersione delle risposte, i commenti critici alle prove (fig. 2) ci illuminano su aspetti sfuggiti all'indagine, oltre a ribadire questioni già trattate durante l'analisi degli item. In particolare, quali punti finora non trattati, rileviamo:

Per la prova CELI

- > Il disagio provocato dalla musica di sottofondo, che, oltre ad essere presente negli stacchi si sovrappone spesso al parlato; per alcuni, la musica catalizza l'attenzione ("se sento la musica non mi concentro più sulle parole")

Per la prova CILS

- > La generale complessità del testo (denso, lungo, intricato, con velocità di eloquio eccessiva)
- > La lunghezza degli item e la complessità, in particolare, di alcuni. Ciò costituisce un fattore di invalidità, data la presenza di una variabile esterna rispetto al costrutto (*construct-irrelevant variance*): la lettura, cioè, rispetto all'ascolto
- > L'incomprensione della parola-chiave "sostenibile"

Per la prova PLIDA:

- > La lentezza del monologo
- > La qualità della voce
- > La densità delle informazioni

Per la prova .IT:

- > La difficoltà a capire il parlato di una delle donne intervistate
- > L'incomprensione della parola-chiave "bambole"
- > La predittività delle risposte ("risposte troppo logiche"), anch'esso fattore di invalidità, in quanto a variabile esterna
- > Difetti di *layout* (non è chiaro che l'item 0 costituisca un esempio).

3. Conclusioni

Nel nostro studio abbiamo analizzato prove di ascolto di livello B2 (nella forma del quesito a scelta multipla), tratte dalle certificazioni di italiano LS. Sono emerse difficoltà di vario tipo, legate in particolare alla confezione degli item. Le differenze in merito sono abbastanza rilevanti; ciò significa che gli esiti che un candidato ipotetico avrebbe potuto raggiungere – nell'ambito della valutazione dell'abilità di ascolto, a un livello B2 – sarebbero potuti variare se costui si fosse sottoposto alla prova di una certificazione piuttosto che a quella di un'altra.

Lo scopo di questo contributo non è, comunque, quelli di auspicare prove 'perfette', dall'oggettività assoluta, assodato del resto che ogni misurazione delle abilità linguistiche equivale pur sempre a un'istantanea sfocata e parziale (Douglas, 2011); piuttosto è quello di stimolare a un maggiore impegno, affinché i margini di approssimazione si possano ridurre.

Va precisato, in ogni caso, che poiché l'analisi è limitata a una singola prova e dato il numero ristretto di item, lo studio non è rappresentativo, non consente cioè di formulare un giudizio in generale sul modo con cui le singole certificazioni valutano la comprensione orale a livello B2. Ciò che emerge ha, altrimenti, un valore circoscrivibile ai soli dati a disposizione. L'impianto metodologico dell'indagine può valere comunque da riferimento per analisi più estese.

Bibliografia

Cardillo, G., Menzinger, C. & Vecchio, P. (ed.) (2015). *Livello B1 PLIDA. Quaderno delle specifiche*. Roma: Società Dante Alighieri. <http://plida.it/certificazione-plida/documenti.html> (consultazione: 18.01.2016)

Douglas, D. (2011). *Understanding Language Teaching*. Londra: Hodden.

Green, R. (2013). *Statistical analyses for language teachers*. Basingstoke: Palgrave MacMillan.

Novello, A. (2014). *La valutazione delle lingue straniere e seconde nella scuola. Dalla teoria alla pratica*. Venezia: Edizioni Ca' Foscari.

Sitografia

Prova ascolto B2 CELI
<http://www.cvcl.it/categorie/categoria-104>
(consultazione: 18.01.2016)

Prova ascolto B2 CILS
http://cils.unistrasi.it/89/198/Prove_Liv_B2.htm (consultazione: 18.01.2016)

Prova ascolto B2 PLIDA (vecchio formato)
<http://plida.it/certificazione-plida/che-cos-e-la-certificazione-plida/esempi-di-prove-d-esame/esempi-sessioni-esame-plida.html>
(consultazione: 18.01.2016)

Prova ascolto B2 .IT
<http://www.certificazioneitaliano.uniroma3.it/B2-intlT.aspx> (consultazione: 18.01.2016)

DESCRITTORI LETTERARI E ACQUISIZIONE DI COMPETENZE: STRUMENTI, OBIETTIVI E PASSIONE!

Die Deskriptoren für den Literatur-Unterricht und für die Aufgaben, die das ESP-Zusatzteam für Italienisch, Deutsch, Französisch und Englisch vorbereitet hat, sollen literarische Kompetenzen beschreiben und einen Beitrag zur Förderung der Selbstevaluation und der metakognitiven Kompetenz der Lernenden leisten. Hier werden einige Teile der Aufgaben für den Italienisch-Unterricht präsentiert – im Bewusstsein, dass die Motivation der Lernenden steigt, wenn sie wissen, mit welchen Instrumenten sich welche Ziele erreichen lassen.

● Eleonora Rothenberger Barbaro | Thurgau

Eleonora Rothenberger Barbaro, 1964, romana, ha studiato lingue straniere e letteratura comparata. Insegna italiano come lingua straniera alla Kantonsschule am Burggraben e alla Interstaatliche Maturitätsschule für Erwachsene, St. Gallen e la Didattica dell'italiano alla Pädagogische Hochschule Thurgau. Ha promosso diverse manifestazioni culturali per la Società Dante Alighieri di San Gallo, di cui è presidente.



La passione letteraria è una competenza? Una premessa fondamentale!

Da diversi anni ormai, dopo la pubblicazione del Quadro Comune Europeo di Riferimento per le lingue (QCER) e del Portfolio europeo delle Lingue (PEL), in molti circoli scolastici, si è andata manifestando la tendenza a voler delimitare, a seconda del punto di vista, campi e settori dell'apprendimento linguistico per rendere misurabile e quantificabile l'apprendimento. L'obiettivo si è spostato da una conoscenza della materia a un 'saper fare' con la materia in un'ottica che in parte ha ridotto lo studio di una lingua ad una acquisizione utilitaristica della stessa. Si volevano soprattutto equiparare saperi nell'ambito della mobilità degli studenti e della valutazione dei livelli dell'apprendimento linguistico, si è spostato invece l'obiettivo sul prodotto finale e sulle cosiddette competenze riducendo l'apprendimento a qualcosa di sempre e comunque quantificabile. Qualcosa è però sfuggito e sfugge a questa logica, o meglio, ne è la premessa spesso sommersa o sommersa: la moti-

vazione e il piacere. Se si dimenticano questi importanti fattori di apprendimento, ma soprattutto di acquisizione, non può esserci appunto acquisizione. Sapere dove andare e perché muoversi è sempre stato importante da che mondo è mondo, in montagna come a scuola. Per questo non bisogna eludere "la domanda come habitus mentale, come posizione intellettuale e morale della persona che cerca un significato nelle faccende della vita."¹ Tutti gli strumenti di equipaggiamento per una buona camminata in montagna, come a scuola l'aver un *tablet*, una piattaforma o un sito dedicato per un buon percorso di studio, sono mezzi che facilitano la strada, ma che non possono sostituire le fondamentali premesse del porsi domande di partenza e di arrivo, di stabilire l'obiettivo e della coscienza delle ragioni del fare. Aggiungo che se si incontrano come compagni di strada insegnanti appassionati e innamorati della letteratura consapevoli del fatto che per osmosi "si può trasmettere la passione della letteratura, della scrittura e magari aiutare a scoprire un piccolo e grande

PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > [SCHEDA 6](#)

talento in chi ignorava di possederlo”² si ha una grande fortuna da far festa!

La banca dei descrittori e il lavoro di squadra

In questa ottica può essere utilissimo sia per apprendenti che per insegnanti lo ‘strumentario’ o ‘armamentario’, che dir si voglia, dei descrittori letterari, che sarà a breve pubblicato online come supplemento ai materiali del PEL III nel portale www.sprachenportfolio.ch per la *Schulverlag plus*. Si tratta di una preziosa banca di sessantadue descrittori letterari in lingua francese, tedesca, italiana e inglese che ha elaborato il gruppo di lavoro guidato da Hans Peter Hodel della Scuola Pedagogica di Lucerna, coadiuvato scientificamente da Thomas Studer e dalla supervisione scientifica del KOGS (n.d.r. Gruppo di coordinamento per le lingue del CDPE) di H.P. Schneider e D. Little, e composto da Brigitte Gerber, Béatrice Leonforte, Marisa Rossi, Patrick Blum, Raffael D’Amaro, Michèle Knuchel e dalla sottoscritta Eleonora Rothenberger Barbaro. Al progetto hanno preso parte anche diversi insegnanti di licei svizzeri nella messa in prova e a punto dei materiali didattici con i descrittori. Il gruppo di lavoro, denominato ESP Lit. Zusatz, PEL Complément lit. o supplemento lett., che si è costituito alla fine dell’anno 2013 e ha cominciato il lavoro concretamente nei primi mesi dell’anno 2014, ha creato, vagliato e definito una banca di descrittori letterari. Questi ultimi sono da intendersi come frutto di un’elaborazione basata sull’esperienza pratica d’insegnamento e su diversi presupposti teorici (dalla tassonomia di Bloom al QCER), nonché da una ricerca di catalogazione che ne permettesse subito la comprensione e il facile accesso e uso: sapere letterario, sapere contestuale interno ed esterno al testo, posizione e comportamento del lettore/studente/insegnante. Sono stati prodotti anche diversi strumenti didattici, dall’elaborazione di tabelle d’uso per il lavoro individuale per la verifica dell’appreso da parte degli apprendenti e per il lavoro di accompagnamento degli insegnanti, a concreti materiali o fogli didattici creati da ciascun insegnante partecipante al progetto per le lingue francese, italiano, tedesco e inglese. I materiali didattici, in particolare, nonché alcune schede di lavoro sono stati sperimentati in diver-

se classi di licei svizzeri da colleghi che hanno aderito alla sperimentazione e in base al feedback si è arrivati alla versione definitiva che sarà pubblicata online. L’accesso alla banca dati e ai materiali sarà in parte libero in parte protetto.

I materiali con i descrittori per l’insegnamento dell’italiano: alcuni esempi

Si presentano quindi e si spiegano qui alcuni esempi di materiali per l’insegnamento dell’italiano come lingua straniera.

Si prendano i seguenti descrittori:

- > so riconoscere particolarità estetiche e stilistiche di un racconto breve e portarne esempi;
- > so riconoscere in un racconto breve alcune modalità stilistiche (per es. figure retoriche) e descriverne funzione e effetto.

Simili tra loro questi descrittori sono stati ordinati nella categoria ‘testo-lingua’, sia perché si tratta di riconoscere elementi di lingua all’interno del testo che ne determinano la peculiarità letteraria appunto, sia nella categoria ‘formulare ipotesi’ e ‘identificare’ perché si tratta per lo/la studente/essa di mettere in moto capacità cognitive che evidenziano quegli elementi del testo che lo rendono appunto letterario e quindi diverso da un articolo giornalistico e da un manuale di istruzioni. Inoltre il secondo descrittore implica un livello di prenosce anche letterarie per poter appunto ‘fare con la letteratura’ e conseguentemente acquisirne pienamente i meccanismi per poter fino in fondo scoprirne il piacere. Nel mio lavoro di produzione di materiali didattici con i descrittori, nel ‘foglio materiali 2’ collegato a una delle unità didattiche da me create, ‘Letterarietà vs cronaca: *Viollica la bambina albanese*’ basata sul testo ‘*Viollica la bambina albanese*’ di Dacia Maraini, tratto da ‘Buio’ (1999), Milano: Rizzoli, pag. 23-29, si chiede allo/la studente/essa di lavorare con delle immagini di animali e altre immagini collegate a prenosce possibili. I due esempi qui presi in considerazione sono un orsetto collegabile alla tenerezza dell’infanzia e un cane che morde collegabile alla paura e alla violenza.

Considerando i compiti richiesti, le soluzioni che si potrebbero avere sono le seguenti, laddove le frasi del testo originale

1 Davide Rondoni (2010), *Contro la letteratura. Poeti e scrittori una strage quotidiana a scuola*, Milano: Il Saggiatore: 71. Rondoni è molto critico sulle domande retoriche e sul gioco finto che può porsi anche in un’aula scolastica in attività che non sembrano riguardare la persona (del discente come dell’insegnante) con un effetto di straniamento e di perdita di motivazione. Si veda a riguardo anche l’interessante testo di Massimo Recalcati (2014), *L’ora di lezione*, Torino: Einaudi.

2 Fabio Pierangeli (2009). *In attesa della festa. Letture ed esercizi dei laboratori di creatività. Nota introduttiva*. Roma: Universitalia. Il libro di Pierangeli è uno strumento prezioso per capire come da una passione, in comune, quasi per gioco, possano nascere esperienze creative irripetibili che lasciano il segno tra docenti e discenti, partendo dalla lettura di testi condivisa. D’altronde capolavori come *Frankenstein* di Mary Shelley sono nati proprio da un mettersi in gioco tra appassionati alla letteratura.

sono qui ridotte in quanto sono molte altre le espressioni che rimandano ai due animali, soprattutto all'orso:

Letterarietà vs cronaca: 'Violca la bambina albanese' Foglio materiali 2 ©RoE

DOPO LA LETTURA DELLA PRIMA PARTE (fino a pag.5)

Conosci delle espressioni linguistiche nella tua lingua che associ a questi animali e cose? Di che sentimenti si tratta? Scrivile/li nella colonna 1

DOPO LA LETTURA DI TUTTO IL TESTO

Cerca nel testo *Violca* le espressioni che hanno a che fare con queste immagini e scrivile nella colonna 2.

	1. Espressioni nella tua lingua/sentimenti	2. Espressioni nel testo
 © Eleonora Rothenberger	<i>zärtlich, kuschelig, tapsig wie ein kleiner Bär Liebe, Zärtlichkeit</i>	-La madre [...] fa per togliere l'orso dalle mani, ma la bambina lo stringe al petto con impeto. Nessuno, per nessuna ragione, deve separarla dal suo Malek, con cui dorme da quando era piccola; -Violca [...] stringendo l'orsacchiotto Malek al suo petto; - Violca si siede sul letto e aspetta cullando l'orsacchiotto Malek.
 © Eleonora Rothenberger	<i>gefährlich, aggressiv wie ein Wolfshund Angst, Gewalt</i>	-[...] lo vede sopra di sé [...] Dal petto nudo colano gocce di un sudore che sa di cane bagnato. Forse è un cane trasformato in uomo. -Il cane ha morso, il cane ha morso.

Letterarietà vs cronaca: "Violca la bambina albanese" Fogli materiali 5 ©RoE

Leggi l'articolo tratto dal Corriere della Sera e poi fai un confronto tra il testo di Violca e la cronaca del giornale dal punto di vista della letterarietà. Che cosa li divide e che cosa li unisce? Indica almeno 6 punti per ogni settore.

Violca	Articolo di giornale	Elementi comuni
• ...	• ...	• ...
• ...	• ...	• ...

Che conclusione ne puoi trarre? Discuti con un/a compagno/a che cosa ottiene la Maraini con la sua scrittura. Poi scrivi qui in breve il risultato della discussione!

.....

.....

Nel testo la scrittrice Dacia Maraini tratta il tema dell'avviamento alla prostituzione di bambine provenienti dai paesi dell'Est europeo in Italia, della violenza sui bambini e della pedofilia, che è stato di grande attualità sulle cronache dei giornali italiani, in particolare negli anni tra il 1980 e il 2000. La scrittrice, da sempre in prima fila contro la violenza e lo sfruttamento di donne e bambini, utilizzando alcune tecniche letterarie in maniera magistrale, espone il problema con grande delicatezza. Chi legge il testo non si accorge quasi della gravità dei fatti per la leggerezza alla maniera 'Calviniana'

della scrittura della Maraini: uso sapiente della metonimia, delle metafore, delle similitudini, del lessico, della tensione interna al testo, della brevità del testo (racconto breve), della cornice del 'giallo', e della conclusione per un lettore, certo non 'happy' alla fine, vista la gravità dei fatti fittivamente accaduti. Questo ne fa un testo prezioso per la sua letterarietà, per il trattare in letteratura un tema che potrebbe altrimenti suscitare scandalo e subire censure a priori.

Attraverso i due punti, presi esemplarmente, dal 'foglio materiali 2', qui in parte riprodotto, si porta gli/le apprendenti a cercare i luoghi nel testo che sviluppano l'antinomia amore-violenza, vittima-carnefice e a vedere che nel testo ogni qualvolta la bambina sta per subire o subisce una violenza si tenta di strapparle il suo orsacchiotto o le viene strappato del tutto ed è distante da lei, quando la violenza sta per avvenire il violentatore è identificato con un cane bagnato, quindi negativamente, quando la violenza è piena il violentatore è paragonato ad un cane che ha morso. Alla fine di diverse attività di analisi del testo, quindi, si arriva alla lettura di un articolo di cronaca, crudo e oggettivo, tratto dal *Corriere della Sera* del 17 gennaio del 2000, che riporta gli stessi fatti narrati nel testo - e non escludo che la Maraini sia stata ispirata dallo stesso articolo, anche se in lei la sensibilità verso il delicato tema della violenza, e non solo sulle donne, è sempre stata presente. Il testo giornalistico affiancato a quello letterario permette di vedere i due generi nelle loro caratteristiche peculiari mettendo senza dubbio in risalto la bellezza letteraria in contrasto con il dramma della cronaca! La riflessione dello/a studente/essa sul confronto dei due testi viene guidata dal 'foglio materiali 5', qui accanto riportato, e deve avvenire prima osservando e rimarcando elementi dell'uno e dell'altro testo, comuni o non comuni, poi sviluppandone la conclusione che viene condotta in lavoro di coppia.

La modalità di lavoro di coppia è stata scelta in modo da avere, nel confronto con il partner, un primo feedback della comprensione dei testi.

L'attività di sintesi finale di tutto il lavoro è un'attività di transfer dell'appreso, in cui si chiede allo/la studente/essa di ribaltare l'atmosfera drammatica, da

cronaca nera in cronaca rosa, trasformandosi in giornalista di cronaca rosa che produce un articolo giornalistico su un tema mondano prima e poi in scrittore/ scrittrice che produce un testo letterario ispirato dall'articolo. L'attività è incentrata sull'ambizione dello/la studente/essa che deve mettere a frutto la propria competenza attraverso un'attività complessa, non immediata, ma che può far uscire dalla dinamica del semplice svolgimento di un esercizio. Il controllo e l'approfondimento critico del lavoro di scrittura creativa avviene da parte dell'apprendente sia riflettendo con nuovi descrittori letterari applicati al racconto breve come:

- > so individuare climax, punti di svolta nonché tensione narrativa
- > so mettere diverse parti (per es. la scena iniziale e quella finale) in relazione tra loro e con il testo intero
- > so riconoscere la situazione o la condizione di un personaggio (per es. il conflitto cognitivo o emotivo) e so metterlo in relazione con la mia esperienza
- > so riconoscere ambiguità e polisemie e distinguere tra uso letterale e metaforico della lingua,

sia con la lettura dei testi prodotti dagli altri che vengono appesi in classe e sui quali si produce una graduatoria di merito

scegliendo per votazione il testo letterario più riuscito. Il prodotto finale è quindi da intendersi come sintesi dell'appreso. Lo studente è invitato ad autovalutarsi nel confronto dei suoi prodotti con quelli degli altri.

Consapevolezza acquisita: aumento di motivazione

In conclusione il lavoro di creazione, valutazione, ordinamento dei descrittori letterari nel gruppo di lavoro menzionato all'inizio dell'articolo e l'applicazione attraverso i materiali da me preparati e modificati nel corso della messa in prova ha acquisito validità attraverso tutto il percorso fatto, permettendo la messa in luce dei punti di forza come la possibilità per gli/le apprendenti di essere maggiormente consapevoli dei mezzi per arrivare agli obiettivi e, aumentandone la consapevolezza, creare gusto, piacere e motivazione nel loro apprendimento. I punti deboli dei descrittori possono vedersi nell'applicazione per cui il loro uso si potrebbe ridurre a semplice strumento per la quantificazione dell'appreso o ancora peggio nella riduzione dell'apprendibile. 'Ai posteri ardua sentenza' e se non ai posteri ai contemporanei che vorranno avvalersi degli strumenti presto fruibili online.

Il lavoro di creazione, valutazione e ordinamento dei descrittori letterari ha acquisito validità attraverso tutto il percorso fatto, permettendo agli/alle apprendenti di essere maggiormente consapevoli dei mezzi per arrivare agli obiettivi e, aumentandone la consapevolezza, di creare gusto, piacere e motivazione nel loro apprendimento.

WORAUF DIE PROFIS VON HÖRVERSTÄNDNISÜBUNGEN ACHTEN! CHECKLISTE

Übersetzung: Mathias Piconi

Mit dem Einverständnis der Autorin angepasst aus:

<http://www.almaedizioni.it/it/officina/2015/officina-25/>

● Alberta Novello | Venezia

Alberta Novello ist Dozentin an der Universität Ca' Foscari, Venezia, und forscht im Bereich Zertifizierung von Fremdsprachen.



Die Anleitung...

- > ist klar und allenfalls mit Beispielen verdeutlicht
- > gibt Rahmeninformationen zur Hörverständnisübung
- > stellt von Anfang an klar, ob nur eine Lösung richtig ist oder mehrere

Die Stimmen im Audio sind...

- > nicht zu viele – im Optimalfall zwei: eine männliche und eine weibliche
- > nicht Teil der Prüfung – es wird also nicht danach gefragt, wer was sagt. Das ist bei vielen Stimmen für die Zuhörer schwer zu ermitteln

Der Ablauf des Tests gestaltet sich so, dass...

- > sich die Zuhörer an die Stimmen gewöhnen können (keine Fragen zu den ersten 10-15 Sekunden)
- > für die Abfolge Antwort – neue Information – darauffolgende Antwort genug Zeit vorhanden ist

Die Items...

- > folgen chronologisch der Hörverständnisübung
- > sind kurz und bündig
- > enthalten nicht zu viel Information
- > lassen sich nicht vom Kontext oder vom Allgemeinwissen ableiten
- > lassen sich nicht von grammatikalischen Merkmalen ableiten (männlich / weiblich, usw.)
- > haben einen niveaugerechten Wortschatz
- > vermeiden komplizierte Verneinungen und Doppelverneinungen
- > stehen für sich – ohne Bezug auf das vorhergehende oder folgende

Das Multiple-Choice...

- > evita opzioni troppo simili
- > evita opzioni assurde
- > ripete le parole chiave non in tutte le opzioni (altrimenti si concentra l'ascoltatore troppo su queste)

Generell...

- > non vengono usati nelle domande figure retoriche
- > le domande non presuppongono conoscenze culturali
- > tiene conto che le emozioni variano in base alla cultura
- > si riferisce ai contenuti rilevanti dell'audio

Officina n° 25 – Dicembre 2015 – La valutazione

Segnaliamo il n. 25 della rivista Officina.it, scaricabile dal sito Alma Edizioni (www.almaedizioni.it/it/officina/2015/officina-25) e dedicato alla valutazione.

I diversi contributi, molto agili, riguardano diversi aspetti del tema.

Mirella Pederzoli, nell'intervista, mette in evidenza l'atteggiamento positivo che si è venuto formando negli ultimi anni attorno alla valutazione, dopo decenni di sospetto, in cui veniva spesso colta dai docenti come una dimensione riservata agli addetti ai lavori, o comunque a rischio di minare la motivazione ad apprendere da parte dell'alunno. Costanza Menzinger tratta la valutazione del parlato. Di particolare rilievo, è il suo confronto tra due modalità di interazione in sede di esame: candidato-valutatore e candidato-candidato, evidenziando per ciascuna pro e contro.

Segue il compendio di Alberta Novello sulla scrittura di item per la valutazione dell'ascolto.

Chiude, infine, il numero una batteria di test (livello B1) tratta dal libro Da zero a cento di Cinzia Ciulli e Anna Lia Proietti (Alma, 2005).



officina.it
la rivista on line di ALMA Edizioni

Officina.it è la rivista che ALMA Edizioni dedica a tutti coloro che si occupano di italiano per stranieri: insegnanti, glottodidatti, linguisti, appassionati, studenti, amanti della lingua italiana e dell'Italia.

Officina.it è una piccola rivista virtuale a tema che gli iscritti al sito Alma riceveranno nella loro casella e-mail.

È sufficiente cliccare sui collegamenti qui sotto per accedere ai numeri in archivio.

Responsabile dei contenuti:
Paolo Torresan

Redazione e coordinamento:
Euridice Orlandino

ISSN 2384-9371

NÜTZLICHE(RE) SPRACHTESTS

BEISPIELE FÜR DIE FORMATIVE BEURTEILUNGS-PRAXIS VOR DEM HINTERGRUND EINER NEUEREN PERSPEKTIVE AUF TEST-GÜTEKRITERIEN

Practicable? – Definitely! Reliable? – As much as possible. Valid? – Yes, at least to some degree! Practicable, reliable, valid: language assessment - from standardized language testing to informal classroom assessment - must be all of this, and even more. Language assessments need to suit their purpose and be *useful* to those concerned by them, to teachers and learners in particular. This is how recent conceptualizations of test quality can be summarized (cf. Green, 2014; Bachman & Palmer, 2010; Grotjahn & Kleppin, 2015). With this in mind, first the concept of usefulness will be presented. This is then used as a point of departure for a closer look at two very different forms of assessment: C-Tests and Dynamic Assessments. Both forms are still not commonly found in practice but can be very useful additions to a teacher's assessment toolkit.

● Thomas Studer | Fribourg

Thomas Studer ist assoziierter Professor für Deutsch als Fremd- und Zweisprache an der Universität Freiburg/Fribourg und Direktionsmitglied des Freiburger Instituts für Mehrsprachigkeit. Zu seinen Forschungs- und Arbeitsschwerpunkten gehören die Sprachlehr- und Sprachlernforschung, die Fremdsprachendidaktik und das Testen und Prüfen.



Praktikabel? Unbedingt! Zuverlässig (reliabel)? Soweit möglich. Gültig (valide)? Ja, mindestens teilweise! – Praktikabel, zuverlässig, gültig: All diese Eigenschaften und noch einige mehr sollten Sprachtests aufweisen. Das gilt nicht nur für standardisierte Tests, sondern eingeschränkt auch für informelle Assessments im Unterricht. Es kommt aber noch etwas dazu: Über die Standard-Merkmale hinaus oder sogar in erster Linie sollten Sprachtests, gleich welchen Kalibers, das Kriterium der Nützlichkeit erfüllen. Tests sollten nützlich sein, und zwar für Lehrende *und* Lernende. So lässt sich eine neuere Perspektive auf Qualitäts-Kriterien für Tests umreißen (u.a. Green, 2014; Bachman & Palmer, 2010 und auch Grotjahn & Kleppin, 2015). Im Folgenden soll diese Nützlichkeits-Forderung zunächst dargestellt (Kap. 1) und dann zum Anlass genommen werden, exemplarisch zwei sehr verschiedene Tests näher zu beleuchten: Varianten des C-Tests einerseits (Kap. 2) und sog. dynamische Assessments andererseits (Kap. 3). Beide Verfahren sind noch wenig verbreitet, bringen aber ein beachtliches Nützlichkeits-Potenzial mit.

Das macht sie für die schulische Praxis interessant, insbesondere für eine auf (Förder-)Diagnostik ausgerichtete Test- und Beurteilungspraxis.

1. Qualitäts-Kriterien für Tests

Green (2014) stellt die Qualitäts-Kriterien für einen Test in Form eines Kegels dar und bringt damit eine Wichtigkeits-Reihenfolge dieser Kriterien zum Ausdruck (Siehe Abb. 1).

Nützlichkeit als zentrales Qualitäts-Kriterium

An der Spitze der Pyramide stehen *nützliche Konsequenzen*, weil es sich dabei um das ultimative Ziel eines jeden Testsystems handelt (Green, ebd.: 59). Die Spitzenposition der „*beneficial consequences*“ ist so zu verstehen, dass diesem Kriterium durch Berücksichtigung der anderen drei Kriterien zwar zugearbeitet werden kann und soll, dass sich nützliche Konsequenzen aber nicht automatisch aus diesen anderen Kriterien ergeben, sondern schon bei der Testauswahl resp. bereits zu Beginn der Testerstellung ‚offensiv‘ angegangen werden sollten. Um-

PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > [SCHEDA 6](#)

gekehrt werden nützliche Konsequenzen kaum zu erreichen sein, wenn es einem Test (weitgehend und gleichzeitig) an Praktikabilität, Reliabilität und Validität mangelt.

Unbestritten ist, dass ohne *Praktikabilität* – bei Bachman & Palmer (2010: 262) durch die Differenz zwischen den für die Testerstellung nötigen und den dafür vorhandenen Ressourcen bestimmt – gar nichts geht. Bei den hierarchisch höher stehenden Kriterien wird es dann komplexer und zunehmend kontrovers: Während es für die *Zuverlässigkeit* oder *Verlässlichkeit* von Messinstrumenten und Prüfungsergebnissen noch handhabbare Prüf-Mechanismen gibt (z.B. Zweitkorrektur einer Schreibaufgabe), sind mit *Validität* eine Reihe von schwierigen Fragen verbunden. Veranschaulicht anhand eines Hörtests aus dem Lehrbuch – etwa ein Interview, das mit einigen Richtig-Falsch-Items verbunden ist –, geht es z.B. um folgende Fragen: Kann ich mit diesem einen Test gültige Aussagen über die Hörverstehens-Kompetenz der Lernenden machen? Oder brauche ich dafür mehrere Texte mit unterschiedlichen Textfunktionen, die von verschiedenen SprecherInnen gesprochen werden, sowie eine Mindest-Anzahl von Items? Nach welchen Regeln müssen diese Items konstruiert sein, damit sie – einzeln und insgesamt – etwas Relevantes über die Hörverstehens-Kompetenz aussagen? Und wie umfangreich müssen die Vor-Tests sein, damit ich das alles feststellen oder zumindest besser abschätzen kann? Unter *nützlichen Konsequenzen* eines Tests schliesslich lässt sich ganz Verschiedenes subsumieren; primär wohl das, was andernorts *Impact*, *Washback* oder *Backwash* heisst: Gemeint ist die Tatsache, dass jeder Test Auswirkungen auf verschiedenen Ebenen hat, darunter ‚gut gemeinte‘, d.h. intendierte und in der Regel positive ebenso wie nicht-intendierte (meist negative). Positive Auswirkungen von Tests im schulischen Kontext (Mikro-Ebene) sind z.B. gezielte Hörverstehens-Übungen als Folge von grösseren Tests, in denen das Hörverstehen berücksichtigt oder aufgewertet wird, oder Flüssigkeits-Trainings als Folge von Tests zum Sprechen. Negative Wirkungen können sich u.a. einstellen, wenn durch eine hohe Kadenz summativer Tests Lerngegenstände zurückgedrängt werden, die nicht prüfungsrelevant sind¹.

Konsequenzen für die Testpraxis

Je weiter nach oben man also in Greens Kegel geht, desto komplexer und anspruchsvoller werden die Anforderungen an qualitativ gute Tests und desto kontroverser sind die Positionen. Doch was bedeutet das für die Testpraxis? Zweierlei: Zum einen wird es um ein Abwägen der Qualitäts-Kriterien gehen, seien es nun die vier Gross-Kriterien nach Green, die sechs Kriterien nach Bachman & Palmer oder die insgesamt elf Aspekte der Testqualität bei Grotjahn & Kleppin². So wird man bei der Überprüfung des Hörverstehens Abstriche bei der Reliabilität in Kauf nehmen und keine aufwändigen Abklärungen im Bereich der Validität machen können. Unverzichtbar hingegen ist immer die zur Validität zählende Frage, „ob und in welchem Umfang die Aussagen und Entscheidungen, die wir anhand der beobachteten Leistungen der Prüflinge treffen, gerechtfertigt sind.“ (Grotjahn & Kleppin, 2015: 50) Zu den leicht realisierbaren Proben in diesem Bereich sollte mindestens die Überprüfung der folgenden beiden, vermeintlich einfachen Fragen gehören: Können die Items eines Hörverstehens-Tests auch ohne den Text gelöst werden? Sind die Items tatsächlich eindeutig lösbar und voneinander unabhängig, sodass Testteilnehmende, die ein Item falsch lösen, eine echte zweite Chance haben? Treten bei diesen Proben Mängel auf, sind gültige Aussagen über die Leistungen der ProbandInnen nicht möglich; der Test geht an der Hörverstehens-Kompetenz vorbei bzw. ist nicht valide.

Das gerade angeführte Zitat von Grotjahn & Kleppin verweist noch auf etwas anderes, nämlich auf eine Dynamisierung des Validitäts-Begriffs und damit auf eine Umorientierung in der Validitäts-Theorie, wie sie bereits Messick (1989: 13) gefordert hatte:

Hence, what is to be validated is not the test or observation device as such but the inferences derived from test scores or other indicators – inferences about score meaning or interpretation and about the implications for action that the interpretation entails.

Nach Messick kommt es letztlich ganz zentral auf die Schlussfolgerungen an, die man auf Basis der Testresultate macht, nicht (so sehr) auf den Test selbst. Dieses Messick-Zitat lässt sich, das ist der zweite Aspekt, sehr wohl auch ganz praktisch

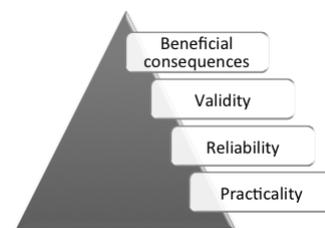


Abb. 1: Vier Qualitäts-Kriterien für Tests (Green, 2014: 58; hier als Pyramide reproduziert)

- 1 Wash-Back-Effekte werden manchmal auch als Aspekt von Validität angesehen (vgl. den Ausdruck *consequential validity*) und würden dann mit zu dem gehören, was einen gültigen Test ausmacht.
- 2 Im Unterschied zu Green (2014) sind bei Bachman & Palmer (2010) auch *Authentizität* (Inwieweit korrespondieren Merkmale der Testaufgabe mit Merkmalen von Aufgaben ausserhalb der Testsituation?) und *Interaktivität* (Sind die Testaufgaben so gestaltet, dass die Testteilnehmenden tatsächlich das zeigen, was sie können?) separate Kriterien. Grotjahn & Kleppin (2015) halten ausserdem am Objektivitäts-Kriterium fest und weisen zusätzlich Fairness und Transparenz sowie die (Item-bezogenen) Merkmale *Trennschärfe* und *Schwierigkeit/Leichtigkeit* separat aus. Im Wesentlichen sind alle drei Ansätze miteinander kompatibel, unterschiedlich weit gefasst wird v.a. das Validitäts-Kriterium. Fairness z.B. – u.a. umschrieben als Anspruch, niemanden z.B. aufgrund des Geschlechts oder durch ‚Tabuthemen‘ zu benachteiligen – ist zweifellos ein zentrales Güte-Kriterium, kann aber, ebenso wie der Back-Wash-Effekt, als Aspekt von Validität angesehen werden. Kurz gefasst ergibt sich die folgende Maxime: Als gültig erweist sich ein Test erst dann, wenn sich Fairness und intendierte Auswirkungen nachweisen lassen.

verstehen: ‚Magere‘ Tests (praktikable, aber unzuverlässige und bezüglich Gültigkeit nicht wenigstens minimal geprüfte) wird es immer geben, aber solche Tests sind keine Basis für Schlussfolgerungen auf die Kompetenz der Lernenden, und Anschlusshandlungen wie z.B. die Vergabe von Abschlussnoten sind auf dieser Grundlage nicht gerechtfertigt. Im konzeptuellen Bereich schliesst sich mit Messick ein Kreis: Wenn Schlussfolgerungen ausschlaggebend sind und entscheidend ist, ob die Anschlusshandlungen in der Folge eines Tests gerechtfertigt sind, ist man wieder an der Spitze der Hierarchie von Test-Gütekriterien angelangt, wie sie Green vorschlägt (vgl. Abb. 1): Das Zentrale beim Testen sind die Konsequenzen, die ein Test haben sollte. Zusammenfassend kann man also sagen: Möglichst nützliche Konsequenzen sind das Leitkriterium für die Qualität von Tests im Dickicht der Tests-Gütekriterien.

Abb. 2: Beispiele für drei C-Test-Varianten, Exzerpte aus Baur et al. (2013: 5, 14 und 15);

- 1) C-Test zur Überprüfung der allgemeinen Sprachfähigkeit und Sprachrichtigkeit und des Textverständnisses
- 2) C-Test zur Überprüfung der Morphologie in Nominalphrasen
- 3) C-Test zur Überprüfung von Fachwortschatz

1) Steinzeitmenschen

Schon vor Tausenden von Jahren lebten Menschen auf der Erde. Sie benutzten Werkzeuge _____ aus Stein. Die _____ Zeit heißt Steinzeit _____. Die Steinzeitmenschen jag _____ oft kleine Ti _____ ...

2) Die Fernbedienung

Ich bin komplett durcheinander. Ich habe bei mir i _____ mei _____ Zimmer ferngesehen und wollte v _____ d _____ ers _____ Programm, vom Tennis, auf das Zweite schalten. ...

3) Das Gebiss des Wolfes

Das Gebiss des Wolfes ist besonders gut zum Fleischfressen geeignet. Die dolchartigen _____ kähne dienen zum _____ sthalten und Töten der _____ te. ...

3 Zu präzisieren ist, dass es sich bei diesem Feedback vorerst nur um ein Lernangebot handelt. Wie dieses Angebot von den Lehrpersonen arrangiert wird und ob und wie es dann von den Lernenden genutzt wird, sind andere, weitgehend offene Fragen (vgl. u.a., mit einiger Skepsis, Dunn & Mulvenon, 2009; optimistischer ist z.B. Davin, 2013, im Kontext des *dynamic assessment*, vgl. unten, Kap. 3).

4 Zum Forschungsstand mit besonderer Berücksichtigung der Validitäts-Diskussion vgl. u.a. Grotjahn (2011).

rende Informationen über Lernstände im Anschluss an eine längere Lernphase zu bekommen (*summative assessments*), sondern auch Informationen, die auf Lernprozesse und Lernfortschritte bezogen sind, denn solche Informationen kommen dem aktuellen Lerngeschehen näher und zielen auf den Entwicklungshorizont der Lernenden (*formative assessments*). Wichtig ist dabei, dass Informationen aus summativen Tests durchaus auch für formative Zwecke genutzt werden können (vgl. etwa den GER (2001: 81); Dunn & Mulvenon (2009) erwägen weitergehend auch die von ihnen so genannten „Evaluation“ formativer Befunde für summative Zwecke).

Vor diesem Hintergrund sollen nun zwei noch weniger bekannte Test-Verfahren mit intakter Praktikabilität und gutem Nützlichkeits-Potenzial besprochen werden, die für formative Zwecke eingesetzt werden können: Varianten des C-Tests (Kap. 2) und dynamische Assessments (Kap. 3).

2. C-Tests: Überraschend vielfältig

Beim C-Test, einer besonderen Form von Lücken-Test („C“ steht für *cloze*), gab es in letzter Zeit neuere Entwicklungen u.a. für Deutsch als Zweitsprache in der Schule (siehe Baur et al., 2013). Diese Entwicklungen lassen sich auch für die Beurteilung im Fremdsprachenunterricht nutzen, und zwar etwa ab der sechsten Klasse (der achten Klasse nach HarmoS), weil für diese Tests eine basale Lese- und Schreibkompetenz in der Zielsprache Voraussetzung ist. Ausgehend vom klassischen C-Test⁴ lassen sich zwei Varianten unterscheiden: a) Eine Variante, mit der sich einerseits die allgemeine Sprachfähigkeit und Sprachrichtigkeit und andererseits das Textverständnis überprüfen lässt und b) eine Variante zur Beurteilung von sprachlichen Teilfertigkeiten, namentlich von linguistisch-grammatischen Kompetenzen (z.B. Beherrschung von Präteritums-Formen) und von Fachwortschatz. Beide Varianten können von Lehrpersonen gemäss den eigenen Bedürfnissen und mit kleinerem Aufwand selbst erstellt und auch selbst eingesetzt und ausgewertet werden. Die folgende Darstellung ist auf die Beispiele in Abb. 2 bezogen und an Baur et al. (2013) orientiert.

Wann ist ein Test nützlich?

Nützlich ist ein Test für Lehrende dann, wenn er z.B. dabei hilft, Lernziele zu überprüfen und ggf. anzupassen, eine einzelne Lektion zeitlich und/oder inhaltlich zu justieren oder Massnahmen der Binnendifferenzierung zu treffen. Lernende wiederum können *dann* Nutzen aus einem Test ziehen, wenn sie ein reichhaltiges, möglichst individuelles Feedback zum Testergebnis bekommen, das sie beim Weiterlernen unterstützt³. Der allgemeine Zweck des Testens und Prüfens besteht darin, Informationen über Lernstände und -entwicklungen zu beschaffen (vgl. z.B. Studer, 2012). Beim Testen und Prüfen *im Unterricht* kommt es zentral darauf an, nicht nur bilanzie-

a) C-Tests zur Überprüfung der allgemeinen Sprachfähigkeit und Sprachrichtigkeit und des Textverständnisses

Bei C-Tests gemäss a) wird ein Text im Umfang von höchstens etwa 200 Wörtern nach der Methode einer 3er-Tilgung bearbeitet, d.h. in jedem dritten Wort wird die zweite Hälfte gelöscht und durch eine durchgehende Linie ersetzt (vgl. Abb. 2.1: DiESE, jagTEN, TiERE). Bei ungerader Buchstaben-Anzahl wird ein Buchstabe mehr getilgt. Komposita werden so bearbeitet, dass nur die Hälfte der letzten bedeutungstragenden Einheit gelöscht wird (z.B. Werk-zeUGE, Stein-zeIT in Abb. 2.1). Eingerahmt wird der so bearbeitete Text von je einem unveränderten Einleitungs- und Schlusssatz, damit das Thema des Textes trotz Lücken erfasst werden kann. Üblicherweise werden in dieser Art vier bis fünf thematisch verschiedene Texte à je 20 Lücken bearbeitet, sodass sich insgesamt 80 bis 100 Lücken bzw. Items ergeben. Die Texte sollten den SchülerInnen inhaltlich bekannt sein, kommen also z.B. aus dem Lehrbuch, können aber gegenüber dem Original gekürzt und leicht verändert (i.d.R. vereinfacht) werden. Pro Text wird ca. fünf Minuten Lösungszeit veranschlagt, sodass der gesamte Test in 20-30 Minuten durchführbar ist. Die Instruktion für die Lernenden ist einfach („ergänze die Lücken“; Hinweis auf zügiges Arbeitstempo und, während der Testdurchführung, auf noch verfügbare Zeit). Wichtig ist, dass das ungewohnte Verfahren anhand eines Beispiels geübt wird. Dabei sollte den Lernenden auch das Auswertungsverfahren erklärt werden (gilt insbesondere auch für den WE-Wert, s. den übernächsten Abschnitt).

Verglichen mit dem klassischen C-Test (Löschung der zweiten Hälfte von jedem zweiten Wort, Punkte für jeden fehlenden Buchstaben) bietet diese für SchülerInnen entwickelte Variante mehr Anhaltspunkte für die inhaltliche Texterschliessung, was technisch auch durch die Wahl von Linien anstelle von Punkten unterstützt wird; zu diesen und weiteren Einzelheiten vgl. Baur *et al.* (2013: 3ff.). Gegenüber dem klassischen C-Test unverändert bleibt dagegen das theoretische Konzept der reduzierten sprachlichen Redundanz als Grundidee dieser Test-Form.

Das Besondere der C-Testvariante a) besteht darin, dass bei der Auswertung zwei Werte ermittelt werden: Ein Richtig/Falsch-Wert (R/F-Wert) und ein Worterkennungswert (WE-Wert). Der

R/F-Wert ergibt sich aus der Anzahl der vollständig richtig (d.h. semantisch UND grammatikalisch UND orthografisch korrekt) ergänzten Lücken, z.B. SteinzeIT: der WE-Wert ergibt sich aus der Anzahl der semantisch korrekt ergänzten Lücken, d.h. neben SteinzeIT wäre z.B. auch SteinzeIHT akzeptabel, nicht aber SteinzeBRA. R/F- und WE-Wert erlauben nützliche Einblicke in das individuelle Sprachvermögen der SchülerInnen, entlang folgender drei Interpretationsregeln (Baur *et al.*, ebd.: 9):

1. Keine oder eine geringe Differenz zwischen R/F- und WE-Wert im oberen Punktbereich sprechen für eine gute allgemeine Sprachkompetenz.
2. Eine eher kleine Differenz zwischen R/F- und WE-Wert im unteren Punktbereich lässt auf fehlendes Textverständnis schliessen.
3. Höhere WE- als R/F-Werte im und ab dem mittleren Punktbereich deuten auf intaktes Textverständnis und gleichzeitig auch auf Schwierigkeiten im formalsprachlichen Bereich hin.

b) C-Tests zur Überprüfung von Teilfertigkeiten

C-Tests zur Überprüfung von Teilfertigkeiten funktionieren im Prinzip gleich wie die unter a) besprochenen Tests, mit folgenden Modifikationen: Bei der Ermittlung *linguistisch-grammatischer Kompetenzen* wird zuerst diejenige zielsprachliche Form oder Struktur bestimmt, deren Beherrschung überprüft werden soll, z.B. Präteritums-Formen, Endungen in Nominal- und Präpositionalphrasen, Pronomina. In Abb. 2.2 ist die Morphologie in Nominalphrasen das Zielphänomen. Mit diesem Fokus wurde der Text an *den* Stellen bearbeitet, an denen das Zielphänomen realisiert ist (z.B. iN meiNEM Zimmer, vON dEM ersTEN Programm in Abb. 2.2). Bei dieser Variante des C-Tests reicht es in der Regel, nur etwa zehn Lücken pro Text zu bestimmen; wollte man mehr Lücken erreichen, müssten die Texte zu stark bearbeitet werden und würden unnatürlich. Unter diesen Bedingungen resultieren für einen gesamten Test 40 bis 50 Items (vier bis fünf Texte à je zehn Lücken).

Wichtig bei Teilfertigkeitstests zur Grammatik ist, dass jeweils nur *ein* Phänomen pro Gesamttest überprüft wird, ansonsten ist die Validität stark eingeschränkt. Werden mehrere Phänomene gleichzeitig überprüft, ist der Blick auf Stärken und Schwächen der SchülerInnen

Praktikabel, zuverlässig,
gültig: All diese
Eigenschaften und noch
einige mehr sollten
Sprachtests aufweisen.
Das gilt nicht nur für
standardisierte Tests,
sondern eingeschränkt
auch für informelle
Assessments im
Unterricht.

- 5 Für belastbare förderdiagnostische Aussagen mit eigenen C-Test-Varianten bräuchte es Standardisierungen, in deren Rahmen auch Referenzwerte (Durchschnittswerte der Zielgruppen) ermittelt werden. Entsprechende Untersuchungen sind für einzelne Lehrpersonen natürlich nicht mehr praktikabel, sie sollten aber an die Hand genommen werden, wenn ein selbst entwickelter C-Test z.B. als Übertritts-Prüfung verwendet werden soll.
- 6 Wesentlich geprägt wird das Forschungsfeld DA durch die Arbeiten von Lantolf und Poehner, vgl. u.a. Poehner (2008), Lantolf & Poehner (2011). Eine praktisch ausgerichtete Darstellung ist Grotjahn & Kleppin (2015), Kap. 6.1.2; ein *computerized dynamic assessment* haben Poehner *et al.* (2015) entwickelt.

bei *einer* Zielsprachen-Struktur verstellt; an diesem Problem kranken viele selbst gemachte Lückentexte. Bei der Auswertung solcher Teilfertigkeitstests macht die Unterscheidung zwischen R/F-Wert und WE-Wert natürlich keinen Sinn: Da es um Grammatik geht, werden nur die (vollständig) richtigen Lösungen gezählt. Der R/F-Wert gibt recht differenziert Auskunft über die Beherrschung (genauer: Reproduzierbarkeit) der ausgewählten grammatischen Struktur oder Form. Eine weitere Art von C-Tests überprüft den erarbeiteten *Fachwortschatz*, was insbesondere für den zweisprachigen Sachfachunterricht interessante Perspektiven bietet. Im Unterschied zu den bisher besprochenen Testvarianten wird beim Teilfertigkeitstest zum Fachwortschatz Stammtilgung der zu überprüfenden Wörter empfohlen. Abb. 2.3 zeigt den Anfang eines bearbeiteten Texts aus dem Fachbereich Biologie. Gelöscht und durch einen Strich ersetzt wurde dort jeweils die erste Hälfte der Zielwörter, z.B. Eck-zähne und BEute. Überprüfungen des Fachwortschatzes können pro Text zehn bis 20 Lücken umfassen, wobei die Textinhalte und Inhaltswörter den SchülerInnen bekannt sein müssen. Bei der Auswertung dieses Teilfertigkeitstests stehen wieder die R/F-Werte im Vordergrund.

Vorteile und Einschränkungen

Die drei vorgestellten C-Test-Varianten, insbesondere die Teilfertigkeitstests, können nicht nur als (formative und summative) Tests, sondern auch als Übungsformen, etwa als Gruppenarbeit, eingesetzt werden, um die SchülerInnen für den korrekten Sprachgebrauch zu sensibilisieren und die Aktivierung von Lese- und Lösungsstrategien zu fördern (Baur *et al.* 2013: 13, 15). Zusammengefasst sind die Vorteile dieser Tests folgende: gute Anpassbarkeit an die unterrichtete Zielgruppe (Schulstufe, Alter und Erfahrungshorizont der SchülerInnen) und spezifische Bedürfnisse (z.B. gerade behandelte Lernstoff), hohe Testökonomie bei der Konstruktion, Durchführung und Auswertung und somit gute Praktikabilität. Um eine befriedigende Validität zu erreichen, sollten selbst entwickelte C-Tests an erwachsenen Muttersprachlern erprobt werden. Können Muttersprachler die Tests nicht eindeutig und problemlos lösen, müssen die betroffenen Lücken revidiert werden. Geht der mit dem Test verbundene Anspruch über

formative Zwecke hinaus in Richtung eines grösseren Abschluss-tests, ist darüber hinaus eine minimale Erprobung in ein bis zwei Klassen unentbehrlich⁵. Hinsichtlich Nützlichkeit vermögen C-Tests der besprochenen Art zu überzeugen: Sie liefern je nach Schwerpunkt differenzierte Anhaltspunkte entweder zur (allgemeinen schriftsprachlichen) Sprachkompetenz und dem Textverständnis der SchülerInnen oder zu sprachlichen Teilkompetenzen, an denen im Unterricht gerade gearbeitet wird.

Zentral ist, dass C-Tests nicht als Ersatz für andere, z.B. kommunikative Verfahren, verwendet werden, sondern als Ergänzung dazu. So ist z.B. offen, ob das in einem C-Test unter Beweis gestellte Form- und Strukturwissen auch beim freien Schreiben verwendet wird, und generell sollten C-Tests nicht dazu (ver-)führen, den Schwerpunkt des Unterrichts doch wieder auf die Grammatik und schriftsprachliche Korrektheit zu legen (siehe dazu oben, Kap. 1, negativer Backwash).

3. Dynamische Assessments: Lernpotenziale im Blick

Dynamische Beurteilung (*dynamic assessment*, DA) spielt in der Pädagogik und besonders bei der Arbeit mit Lernenden, die so genannte ‚besondere Bedürfnisse‘ haben, seit langem eine Rolle, ist aber in der Fremdsprachenforschung ein relativ neues Phänomen. Zum fremdsprachenspezifischen DA gibt es inzwischen eine Reihe von theoretischen Auseinandersetzungen, erst wenige empirische Studien (mit kleiner Probandenzahl) und eine Vielfalt verschiedener Anwendungen, die von Lehr-Lern-Gesprächen mit unterschiedlichen Graden der (Vor-)Strukturierung über systematische Arten der Fehlerkorrektur beim Schreiben bis hin zu standardisierten, computerbasierten Tests zum Hören und Lesen reichen (Übersichten, an denen sich die folgende Darstellung hauptsächlich orientiert, bieten Grotjahn, 2015, und Kley, 2011)⁶. Im Unterschied zu den bekannten, oft statisch genannten Tests kommunikativer Sprachkompetenzen, wie sie etwa auch *lingualevel* bereit stellt (vgl. Lenz & Studer, 2007), zielt DA zwar auch auf den aktuellen Stand der Kompetenzentwicklung, strebt aber darüber hinaus und vor allem eine Diagnose des Lernpotenzials an und ist direkter auf die Förderung der Lernenden ausgerichtet. Das zentrale Element des DA sind Hilfestellungen,

welche die Lernenden erhalten resp. in Anspruch nehmen können, wenn sie bei der Bearbeitung von Aufgaben ‚anstehen‘. Beim Sprechen z.B. bestehen die Hilfestellungen aus spezifischen Feedbacks der Lehrperson, wenn die Lernenden Fehler machen, vgl. Abb. 3.1.

Erläuterungen zum Transkript: Die Lernerin Molly bildet die Frage ¿Cómo le gusta hacer en Buenos Aires?, braucht aber das falsche Fragewort (cómo statt qué). Die Lehrperson führt Molly mit vier Hinweisen zum korrekten Fragesatz. Die ersten drei Hinweise – eine Pause und ein skeptischer Blick (prompt 1), die Wiederholung der gesamten falschen Äusserung (prompt 2) und die Wiederholung des nicht korrekten Frageworts (prompt 3) – führen nicht zum Ziel. Erst der vierte Hinweis, eine Entscheidungsfrage (¿Cómo o qué?), kann von der Lernerin zur Verbesserung ihrer Äusserung genutzt werden.

Typischerweise sind die Rückmeldungen beim DA zunächst implizit (z.B. eine nonverbale, mimische Reaktion) und werden dann immer expliziter, wobei die Richtigstellung des Fehlers (der explizite Pol des Hinweis-Spektrums) nur dann erfolgt, wenn die Lernenden nicht vorher selbst reagieren. Die Idee ist also, dem/der Lernenden quantitativ und qualitativ genau die Hilfen anzubieten, die er oder sie braucht, um den Fehler zu erkennen und selbst zu korrigieren. Und genau diese Idee ist dann auch ausschlaggebend für die Bestimmung des Lernpotenzials: Die Indikatoren für das Lernpotenzial sind die Art und der Umfang der beanspruchten Hilfen. Beispielsweise ist die gerade gemachte Feststellung, dass die Lernerin Molly auf die impliziteren Hinweise der Lehrperson nicht reagiert (und relativ viel Hilfe beansprucht, bis die Selbstkorrektur gelingt), ein Indiz für ein noch eher kleines Lernpotenzial, was die behandelten WH-Fragen betrifft. Beides, Qualität und Quantität der beanspruchten Hilfen, kann sich von Lerner/in zu Lerner/in unterscheiden – auch, und das ist hervorzuheben, bei gleichem Kompetenz-Stand. Diese Perspektiven-erweiterung, der Blick auf individuelle Lernpotenziale, ist der Mehrwert des DA gegenüber ‚statischen‘ Tests.

Das theoretische Fundament des DA sind die soziokulturelle Theorie Vygotskij's und dessen Konzept einer Zone der nächsten Entwicklung (zone of proximal development, ZPD), definiert als Differenz zwischen dem, was Lernende selbständig

1. Molly: ¿Cómo le gusta hacer en Buenos Aires? [How do you like to do in Buenos Aires?]
2. Teacher: (Prompt 1—paused and looked skeptically at student)
3. Molly: ¿Cómo... [How...]
4. Teacher: (Prompt 2) ¿Cómo le gusta hacer en Buenos Aires? [How do you like to do in Buenos Aires?]
5. Molly: ¿Dónde? [Where?]
6. Teacher: (Prompt 3) ¿Cómo? [How?]
7. Molly: (silence)
8. Teacher: (Prompt 4) ¿Cómo o qué? [How or what?]
9. Molly: ¿Qué le gusta hacer? [What do you like to do?]
10. Teacher: Sí, ¿Qué significa qué en inglés? [Yes, what does qué mean in English?]



Abb. 3.1: Classroom DA: Einzelgespräch Lehrperson-Molly, während die Klasse zuhört; Kontext: Spanisch als Fremdsprache für Kinder, ca. A2-Niveau, Thema WH-question formulation; Exzerpt aus Davin (2013: 16).

leisten können und dem, was sie zu leisten im Stande sind, wenn sie unterstützt werden⁷.

Interventionistische und interaktionistische Ansätze

Die Vielfalt von DA-Konzepten lässt sich grob zweiteilen in interventionistische und interaktionistische Ansätze (Grotjahn, 2015: 472f.; Kley, 2011: 68f.): Der *interventionistische Ansatz* folgt einem *Prätest-Intervention-Posttest-Design* und wird bisher v.a. in der Forschung praktiziert. Es bieten sich aber im Rahmen der Handlungs- und Aktionsforschung auch Umsetzungsmöglichkeiten im Unterricht an. Beim interventionistischen Ansatz geht es darum, die Lernstände zu Beginn der Untersuchung und nach der Intervention zu messen, und zwar mit gleichartigen, in der Regel ‚statischen‘ Tests und mit der Erwartung, dass sich die Leistungen beim zweiten Mess-Zeitpunkt verbessert haben⁸. Die Intervention selbst erfolgt nach den Prinzipien des DA, kann mehr oder weniger stark strukturiert sein und sich im Rahmen des Klassengesprächs auf jeweils eine/n Lernende/n beziehen oder in Gruppen realisiert werden. Abb. 3.1 ist ein Beispiel eines stark strukturierten Lehrer-Schüler-Gesprächs, in dessen Verlauf ein definiertes Set von Hinweisen eingesetzt wird. Davin (2013: 10) weist das folgende Hinweis-Set aus, das aus fünf nach Expliztheit geordneten Hinweisen besteht (Abb. 3.2)⁹.

⁷ Genaueres und Weiteres zum theoretischen Fundament des DA u.a. bei Grotjahn (2015: 471f.). Den Grundsatz, dass Lernende mit Hilfen mehr und/oder anderes können als ohne Hilfen, kennt man auch aus den Can-do-Checklisten des Europäischen Sprachenportfolios.

⁸ Weiter ausgebauten Forschungs-Designs würden zusätzlich eine Kontrollgruppe einbeziehen und einen dritten Mess-Zeitpunkt ansetzen, um die Nachhaltigkeit des erwarteten Kompetenz-Zuwachses zu prüfen.

⁹ Ein illustratives Beispiel für den interventionistischen Ansatz des DA, bei dem in der Interventionsphase sowohl mit Lehrer-Schüler-Dialogen als auch mit Gruppen-Gesprächen gearbeitet wird, ist Davin (2013; 2011).

Abb. 3.2: Festgelegte Mediations-Hinweise der Lehrperson im Rahmen eines interventionistischen Ansatzes des DA (Davin 2013: 10)



Level of Explicitness	Mediation Prompt
Prompt 1	Pause with skeptical look
Prompt 2	Repetition of entire phrase by teacher with emphasis on location of error
Prompt 3	Repetition of specific site of error
Prompt 4	Forced choice option (i.e., when or where?)
Prompt 5	Correct formulation of question accompanied by explanation

10 Die Lösungshinweise bei schriftlichen Tests können beim Durchsprechen der Aufgaben mündlich gegeben werden, es sind aber auch schriftliche und sogar standardisierte Rückmeldungen möglich, wie das im von Poehner *et al.* (2015) entwickelten Computer-basierten DA zum Hörverstehen realisiert ist.

Beim *interaktionistischen Ansatz*, einem sog. *Train- Within-Assessment-Design*, erfolgt grundsätzlich auf jedes „Item“ eine Rückmeldung an die Lernenden. Beim Mündlichen geschieht dies in der Art der geschilderten Feedbacks bei Fehlern, wobei die Rückmeldungen weit vielfältiger und individueller sein können als in Abb. 3.2 dargestellt. Bei Tests zu den rezeptiven Fertigkeiten, z.B. bei Hörverstehens-Aufgaben mit Auswahl-Antworten, erfolgen die Rückmeldungen in Form von immer konkreteren Lösungshinweisen, wenn ein Item zuerst falsch gelöst wurde¹⁰.

Bereits deutlich geworden sein sollte, dass insbesondere beim interaktionistischen DA Testen, Lehren und Lernen eng miteinander verzahnt sind. Testen und Unterrichten, sonst ja oft als Gegensatz dargestellt, werden hier zusammengeführt.

Deutlich geworden sein sollte weiter auch, dass den Lehrpersonen bei dynamischen Beurteilungen eine aktive und ganz entscheidende Rolle zukommt. Dies schlägt sich auch in der Begrifflichkeit nieder: Im Rahmen des DA sind Lehrpersonen nicht einfach Feedback-Lieferanten, sondern *Mediatoren*, womit gut zum Ausdruck kommt, dass es um die Vermittlung zwischen zielsprachlichen Strukturen/Aufgaben und den Lernenden geht, nicht bloss um Instruktionen oder Korrekturen der bekannten Art.

Aspekte der Auswertung

Was die Auswertung des DA angeht, eignen sich Beobachtungsprotokolle, und zwar auch solche für die Selbstbeobachtung der Lehrpersonen. Lehrerseitig wird es z.B. nützlich sein, sich anhand von Protokollen der eigenen Hilfestellungen und der Reaktionen der Lernenden darauf bewusst zu werden, sodass das Repertoire an Hilfestellungen ggf. angepasst werden kann (vgl. Abb. 3.3; Grotjahn & Kleppin, 2015: 134) führen für den interaktionistischen Ansatz des DA zehn verschiedene

Typen von Hilfen an, die die Selbstkorrektur der SchülerInnen unterstützen). Werden stärker strukturierte Formen des DA gemäss Abb. 3.2 eingesetzt, bieten sich zur Diagnose des Lernpotenzials der Lernenden Tabellen an, in die man für die jeweils beobachteten SchülerInnen Anzahl und Art der in Anspruch genommenen Hinweise einträgt. Idealerweise werden solche Eintragungen – beim gleichen Lernstoff! – mehrmals, mind. aber zu zwei Zeitpunkten gemacht, sodass sich auch Veränderungen des Lernpotenzials abschätzen lassen.

Vorteile und Einschränkungen

Die enge Verzahnung von Unterrichten und Testen bei dynamischen Assessments hat auch Folgen für die Interpretation der Test-Gütekriterien. Für den interaktionistischen Ansatz des DA lässt sich in etwa festhalten: Hinsichtlich Praktikabilität erweist sich das Verfahren einerseits als sehr flexibel, was Individualisierungs-Möglichkeiten und Lerngegenstände angeht, andererseits aber auch als recht aufwändig in Bezug auf Planung, Organisation und nachbereitende Reflexion. Validität ist dann gegeben, wenn das Ziel des DA erreicht wird, d.h. wenn es gelingt, das Lernpotenzial der Lernenden zu erfassen und die (weiteren) Lernprozesse zu unterstützen. Wie dies gelingt, d.h. mit welchen und wie vielen Hinweisen man die nächste Entwicklungszone ausleuchtet, kann von Lerner/in zu Lerner/in sehr verschieden sein und wird gerade deshalb kaum je eins zu eins wiederholbar sein, sodass Reliabilität bei diesem Verfahren kein sinnvolles Kriterium ist. Das Nützlichkeits-Potenzial schliesslich ist als gross einzuschätzen: Kenntnisse von „Leistungsreserven“ (Grotjahn, 2015: 470) sind nicht nur für Lehrpersonen relevant (u.a. Hilfe für das Abstimmen des Unterrichts auf individuelle Kompetenzen und für das Setzen realistischer Lernziele), sondern nützen auch den SchülerInnen (Bewusstmachung von Lernvorgängen und Lernstrategien, auch als zentraler Bedingung für autonomes Lernen).

Kritisch anzumerken ist, dass dynamische Assessments zwar für eine breite Palette von Lerngegenstände in Frage kommen, aber erfassen und weiter entwickeln lassen sich doch nur Fähigkeiten

Abb. 3.3: Protokoll zur Selbstbeobachtung der (Wirksamkeit der) Unterstützung von Lernenden beim interaktionistischen DA (Grotjahn & Kleppin, 2015: 136)



Hilfen zur Selbstkorrektur, die ich ausprobiert habe	Reaktionen der Lernenden, die Hinweise auf schon vorhandenes Potenzial geben

und Fertigkeiten, die dem Bewusstsein zugänglich sind, z.B. Strategie-Kenntnisse oder grammatisches Regelwissen. Nicht erfassen lassen sich dagegen unbewusste kognitive Prozesse, und unberücksichtigt bleibt der gesamte Bereich des impliziten Lernens. Klar ist schliesslich auch, dass sich der (erwartete) Erfolg des DA nicht über Nacht einstellen wird. Gerade der sehr flexible interaktionistische Ansatz des DA muss ausprobiert, reflektiert und trainiert werden, insbesondere auch, was den Umgang mit dem Hinweis-Repertoire angeht. Das aber – der Weg hin zu reflektierteren Beurteilungspraktiken – kann sich lohnen. Auf diesem Weg können mehr Sicherheit in der eigenen Beurteilungspraxis und mehr Klarheit über die Nützlichkeit der eingesetzten Verfahren erreicht werden.

Fazit

Jeder Test, auch ein unterrichtsnaher, kann und sollte noch vermehrt unter dem Aspekt beurteilt werden, ob er den beteiligten Lehrenden und Lernenden etwas nützt und worin genau der Nutzen besteht. Mit dieser Frage-Haltung lässt sich nicht zuletzt ein Auswahl- und Analysegesichtspunkt für bestehende Tests gewinnen *und* man bekommt eine Richtschnur für das Erstellen eigener Tests in die Hand. Bei den beiden hier diskutierten Tests, Varianten des C-Tests und dynamischen Assessments, handelt es sich um praktikable Verfahren mit beachtlichem Nützlichkeits-Potenzial. Das privilegiert sie für Try-Outs im eigenen Unterricht – als Ergänzung zu den üblichen Tests im Rahmen eines erweiterten Konzepts von Beurteilung.

Literatur

Bachman, L. & Palmer, D. (2010). *Language Assessment in Practice*. Oxford: Oxford University Press.

Baur, R.S., Goggin, M. & Wrede-Jackes, J. (2013). *Der C-Test: Einsatzmöglichkeiten im Bereich DaZ*. Universität Duisburg-Essen: pro-DaZ. Online (1.4.16): www.uni-due.de/imperia/md/content/prodaz/c_test_einsatzmoeglichkeiten_daz.pdf

Davin, K. J. & Donato, R. (2013). Student Collaboration and Teacher-Directed Classroom Dynamic Assessment: A Complementary Pairing. *Foreign Language Annals* 46/1, 5-22.

Davin, K. J. (2011). *Group dynamic assessment in an early foreign language learning program: Tracking movement through the zone of proximal development* [Unpublished doctoral dissertation]. University of Pittsburgh. Online (1.4.16): http://d-scholarship.pitt.edu/7269/1/DAVINKJ_ETD.pdf

Dunn, K. E. & Mulvenon, S. W. (2009). A Critical Review of Research on Formative Assessment: The Limited Scientific Evidence of the Impact of Formative Assessment in Education. *Practical Assessment, Research & Evaluation* 14/7, 1-11.

Europarat (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin u. a.: Langenscheidt.

Green, A. (2014). *Exploring language Assessment and Testing*. New York: Routledge.

Grotjahn, R. (2015). Dynamisches Assessment: Grundlagen, Probleme, Potenzial. In: J. Bäcker & A. Stauch (Hrsg.), *Konzepte aus der Sprachlehrforschung - Impulse für die Praxis*. Frankfurt am Main: Lang, pp. 469-488.

Grotjahn, R. (2011). C-Tests – Aspekte der Validität. *Deutsch als Fremdsprache* 48/3, 131-137.

Grotjahn, R. & Kleppin, K. (2015). *Prüfen, Testen, Evaluieren*. München: Klett-Langenscheidt.

Kley, K. (2011). Dynamic Assessment. Zusammenführung von Unterricht und Leistungsmessung. *Deutsch als Fremdsprache* 48/2, 67-74.

Lenz, P. & Studer, T. (2007). *lingualevel. Instrumente zur Evaluation von Fremdsprachenkompetenzen*. Bern: Schulverlag blmv.

Lantolf, J. P. & Poehner, M. E. (2011). *Dynamic assessment in the foreign language classroom: a teacher's guide, second edition*. The Pennsylvania State University: CALPER Publications.

Messick, S. (1989). Validity. In: R. L. Linn (Ed.), *Educational measurement*, 3rd ed. New York: American Council on Education and Macmillan, pp. 13-103.

Poehner, M. E. (2008). *Dynamic Assessment: A Vygotskyan Approach to Understanding and Promoting L2 Development*. Berlin: Springer.

Poehner, M. E., Zhang, J. & Lu, X. (2015). Computerized dynamic assessment (C-DA): Diagnosing L2 development according to learner responsiveness to mediation. *Language Testing* 32/3, 337-357.

Studer, T. (2012). Leistungsbeurteilung: Testfunktionen als Orientierungshilfe. Positiver Washback: Mit *lingualevel* Sprachkompetenzen beurteilen. *Grundschulmagazin Praxis* 5, 7-12.

MORE THAN TESTING THE WATER

BEURTEILUNG VON ENGLISCHKOMPETENZEN IM VERLAUF DER VOLKSSCHULE

This contribution reports about sites and occasions for diagnostic, summative and high-stakes assessment in the normal course of English language learners' careers in primary and secondary school in the Canton of Zurich, Switzerland. It describes and discusses widespread diagnostic tests such as "Stellwerk" and "Multicheck" as well as an entry test for vocational education. In the absence of proper graduation or certificate exams, these assessments represent key moments for the planning of further learning and the transition to next higher schools. The article discusses the issue of whether the various assessment procedures occur at the right times and with the required degree of care and quality. It concludes that the Stellwerk test is well-suited to diagnose partial competences and to foster further learning, but it should be used as a certificate of achievement only with caution. Yet Stellwerk is still a far better option than ineptly designed commercial packages, which are often sought after by employers.

● Daniel Stotz | Eglisau

Daniel Stotz lehrt Fremdsprachendidaktik an der PH Zürich. Sein forschendes Augenmerk gilt dem inhaltsbezogenen und aufgabenorientierten Englischunterricht sowie curricularen und sprachenpolitischen Fragestellungen. Nebenamtlich ist er Schulpfleger in Eglisau; in dieser Rolle hat er auch Zugang zum Stellwerk-Test und dem im Text beschriebenen Klassenzimmer gefunden. Dem Schulleiter Stephan Passerini und den Lehrerinnen Tanja Wolff und Carolyn Lutterbeck sei hiermit herzlich gedankt!



Pssst! In diesem Klassenzimmer herrscht vollkommene Stille, zu hören sind nur leise Tippergeräusche. Die 16 Schülerinnen und Schüler einer Sekundarschulklasse im zürcherischen Eglisau sitzen in gespannter Körperhaltung vor ihren Laptops und stellen sich den Englisch-Aufgaben, die ihnen das Testprogramm vorlegt. Ab und zu dringen aus den Ohrhörern gedämpfte Stimmen. Die Stirn runzelt sich, die Hand bewegt die Maus, ein Kästchen wird angekreuzt – oder nein, doch eher die andere Lösung? Die Schülerinnen und Schüler arbeiten äusserst konzentriert an ihren Tests, und in die Pause gehen nur diejenigen, die ihn schon beendet haben. Wie ist es möglich, dass eine Gruppe von lebendigen, oft lauten Teenagern sich bis zu einer Stunde lang „mega“ konzentriert einer anspruchsvollen Sache widmet – und dies in einer Fremdsprache? Eine Antwort auf diese Frage soll dieser Beitrag geben, jedoch muss das beobachtete Phänomen zunächst in einen Kontext eingebettet werden. Die Klasse in Eglisau stellt sich einer besonderen Beurteilungssituation, dem in der 8. Klasse erstmals durchgeführten Stellwerk-Test. Dessen Bedeutung

kann nur verstanden werden als Glied einer Kette von Lernstandserhebungen und Prüfungen, denen sich Schülerinnen und Schüler der Volksschule während ihrer obligatorischen Schulzeit zu unterziehen haben. Im Brennpunkt stehen im Folgenden die Beurteilungssituationen an Übergängen und zu Zeitpunkten, wenn Weichen für weiteres Lernen und für das Ausbügeln von Defiziten gestellt werden. Der Knackpunkt bei dieser kontrovers diskutierten Geschichte ist die Tatsache, dass die Volksschule keine Abschlussprüfungen in Analogie zur Maturität kennt, und dass Schülerinnen und Schüler ihr letztes Schulzeugnis meist erst dann in Empfang nehmen, wenn die nächsten Schritte ihrer weiteren Schul- oder Berufslaufbahn bereits eingeleitet sind¹. Zu diskutieren ist deshalb die Frage, in wie weit die verschiedenen Beurteilungsprozesse von Fremdsprachkompetenzen auf dem Normalparcours eines Volksschülers den Erfordernissen und Qualitätsansprüchen der Schul- und Berufsbildungspraxis genügen und ob die bestmöglichen Lernstandserhebungen auch zu geeigneten Zeitpunkten erfolgen.

PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
 ARCHIVIO TEMATICO > [SCHEDE 6 E 7](#)

Als Vorbereitung für die Beurteilungssituationen gegen Ende der Schulzeit widmet sich der erste Teil den formativen und summativen Leistungsbewertungen im Schulalltag der Primarstufe. In Abschnitt 2 geht es dann um diagnostische Beurteilung und deren Konsequenzen auf der Sekundarstufe I. Der dritte Teil beleuchtet die einzige umfassende Prüfung am Ende der Sekundarstufe I, die Aufnahmeprüfung für die Berufsmaturitätsschulen. In Abschnitt 4 steht der Test „Multicheck“ zur Debatte, der in Betrieben der Berufsbildung oft den schulischen Leistungserhebungen vorgezogen wird.

1. Leistungserhebung und -bewertung im Englischunterricht der Primärschule

Zu beschreiben ist hier in knapper Form der *Courant normal* der Progressionsmessung aus der Perspektive einer beispielhaften Phantomschülerin – nennen wir sie Lea, und konzentrieren wir uns auf die Fremdsprache Englisch. In der Unterstufe (2. und 3. Klasse) wird das Fach nicht benotet, Lea wird aber eine Einschätzung ihrer Hör- und Sprechkompetenzen in ihrem Zeugnis finden². Ihre Lehrerin kann sich auf Mustertests abstützen, die den offiziellen Lehrmitteln zugeordnet sind (z.B. Achermann & Ramsey 2012). Die Organisation von summativen mündlichen Leistungsmessungen ist erfahrungsgemäss anspruchsvoll – es wäre nicht untypisch, wenn Lea von ihrer Lehrerin anlässlich eines Kurzvortrags zu einem Lieblingstier mit einem groben Raster fast unbemerkt beurteilt würde. In der Mittelstufe (ab 4. Klasse) erfordert das Zeugnis des Kantons Zürich das Ausstellen einer Note und die Beurteilung der vier Teilbereiche Hörverstehen, Leseverstehen, Sprechen und Schreiben³. Auch wenn die Assessment Packs und Lernkontrollen der Lehrmittel *Explorers* und *Young World* nur an Schulen verkauft werden, wird die Lehrerin gut daran tun, die Aufgaben abzuändern und anzupassen, für den Fall, dass die Tests bereits in der Öffentlichkeit zirkulieren.

Beim Übertritt in die Sekundarstufe I – Lea wird in die lokale Schule gehen und nicht ins Langzeitgymnasium – erfolgt die Zuteilung zu einem der Leistungszüge A (erweiterte Anforderungen) oder B (allenfalls C) mittels einer Gesamtbeurteilung und einem Elterngespräch. Dabei haben Leistungen und Potenzial

in einzelnen Fächern keine besondere Bedeutung. Da jedoch Leas Sekundarschule den Englischunterricht in drei Niveaus führt, sollte ihre Primarlehrerin eine entsprechende Empfehlung abgeben. Nicht wenige Lehrpersonen projizieren ein eher veraltetes Bild von den Leistungsanforderungen der Sekundarstufe auf den zu erwartenden Unterricht. Es könnte sein, dass Lea trotz sehr guten Bewertungen in den rezeptiven Kompetenzen und guten Leistungen im Sprechen im Niveau II landet, weil angeblich in der Sek die Grammatik und das Schreiben stärker gewichtet würden. Der Lehrplan der Sekundarstufe (Lehrplan Englisch EDK-Ost, Lehrplan 21) orientiert sich aber genauso an Kompetenzdeskriptoren und den Niveaubeschreibungen des Europarats (Europarat 2001) und von *lingualevel* (Studer & Lenz 2007) wie derjenige der Primarstufe. Hier rächt es sich nun, dass in diesem Kanton das Europäische Sprachenportfolio kaum eingesetzt wird; die flächendeckende Einführung ist am Widerstand der Lehrerverbände gescheitert (Beschluss des Bildungsrats vom 31.8.2009). Eine Durchsicht der im Portfolio abgelegten schriftlichen Arbeiten Leas liesse vielleicht eine positive Tendenz erkennen; eine Zuteilung zum Niveau I könnte motivierend wirken. Die Gesamtbeurteilung im Sinne einer prognostischen Einschätzung beim Übertritt in die Sekundarstufe I bietet zwar mit dem Einbezug von Schüler- und Elternperspektive viele Vorteile, sie schränkt aber die Kontinuität und Kohärenz des Kompetenzaufbaus in einem einzelnen Fach wie Englisch als Fremdsprache möglicherweise ein. Als Ergänzung zur prognostischen Beurteilung über die Stufen hinweg sollten Schulen und Lehrpersonen Massnahmen ins Auge fassen, die mehr Transparenz und Zusammenhalt zwischen den Schulstufen schaffen. Ein Beispiel sind Brückenaufgaben, wie sie im Projekt „*Primary and Secondary Continuity in Language Learning (Pri-Sec-Co)*“ erarbeitet wurden (Megias Rosa & Stotz 2010)⁴.

2. Diagnostische Beurteilung im Verlaufe der Sekundarstufe I

Bis zum Zeitpunkt des Stellwerttests am Ende des ersten Semesters im 8. Schuljahr bewegt sich die Leistungsüberprüfung der Schülerin Lea im Fach Englisch in ähnlichem Rahmen wie auf der Mittelstufe. In die Zeugnisnote und die Bewertung der Teilbereiche fliessen

- 1 Im Kanton Zürich existiert seit Jahren die Möglichkeit für Schulabbrecher und spät Zugezogene, als Erwachsene den Sekundarschluss gemäss Lehrplan nachzuholen. Diese „Sekundarabschlussprüfung für Erwachsene“ ist die einzige umfassende Schulabschlussprüfung. Informationen und Musterprüfungen können über www.vsa.zh.ch (Suchbegriff Sek I für Erwachsene) abgerufen werden.
- 2 In einzelnen Kantonen werden schon ab der 3. Klasse alle Fertigkeiten beurteilt, in anderen wird für das Fach Englisch nur eine Note erteilt.
- 3 Um die Lehrpersonen von der zeitraubenden Arbeit für diese Erhebungen zu entlasten, strich der Bildungsrat des Kantons Zürich die entsprechenden Skalen aus dem Zeugnis des ersten Semesters. Im Beschluss vom 19.3.2012 heisst es: „Die Beurteilung der Teilbereiche in den Sprachfächern berücksichtigt die Leistungen des ganzen Schuljahrs und wird im Zeugnis des 2. Semesters abgebildet.“
- 4 Mit sogenannten *bridging tasks* lösen Schülerinnen und Schüler sprachlich-kommunikative Aufgaben, die am Ende der Primarstufe in der Fremdsprache vorbereitet werden und deren Ergebnisse sie zu Beginn der Sekundarstufe präsentieren. Damit wird ihnen der Einstieg erleichtert und Sekundarlehrpersonen können sich anhand der mündlichen und/oder schriftlichen Produkte ein diagnostisches Bild über die Kompetenzen machen. Beispiele für solche Aufgaben finden sich im Modul „Moving on“ des Lehrmittels *Explorers* (Achermann & Stauer, 2009) und auf der Website des Projekts *Pri-Sec-Co* (Megias Rosa & Stotz, 2010).

Die Crux mit dem Stellwerk-Test zeigt sich nicht etwa in seiner Anwendung als Instrument der Diagnostik und Lernförderung, sondern bei seiner Nutzung als Wegweiser in der Laufbahnplanung.

auch die Erkenntnisse der formativen Beurteilung ein; bis zu einem Fünftel der Bewertung kann sich aus Beobachtungen im Unterricht und aus der Begutachtung von schriftlichen Hausaufgaben ergeben. Die oft noch anzutreffende Praxis des wöchentlichen „Voci-Tests“ müsste eigentlich in diese formative Bewertung einfließen, denn summative Prüfungen sollten nicht auf isolierte Wortschatztests abstellen – diese vermögen die anzustrebenden Kompetenzen und Wissensstände kaum abzubilden.

Die verschiedenen publizierten Muster-tests (z.B. Keller-Bolliger & Haller, 2010) umfassen Aufgabenstellungen zu Hör- und Leseverstehen, Sprechen in Interaktion und/oder zusammenhängendem Sprechen (Produktion) sowie Schreiben. Grammatik und Wortschatz sind, wie im Lehrplan 21 vorgesehen, in diese Aufgabenstellungen eingebettet. Lea lernt am besten auf solche kompetenzorientierten Tests indem sie dem Unterricht konzentriert folgt, die Lernaufgaben ausführt und mit den angebotenen Tools intelligent übt.

Der Stellwerk-Test wird in vielen Kantonen der Zentral- und Ostschweiz eingesetzt (in Zürich flächendeckend), um den Lernstand von Schülerinnen und Schülern in mehreren Fächern festzustellen und daraus abzuleiten, an welchen Themen und Kompetenzen sie in der verbleibenden Schulzeit verstärkt arbeiten sollen. Das Testsystem ist webbasiert und adaptiv, das heisst, es passt die Aufgaben sukzessive der Leistungsfähigkeit der Schülerin oder des Schülers an, bis es für jede Teilkompetenz das Niveau eruiert hat, auf dem sie/er die Aufgaben gerade noch lösen kann. Gemäss Eigendeklaration ermöglicht Stellwerk mit dem individuellen „Leistungsprofil einen schultypenunabhängigen Kompetenznachweis und leistet damit einen Beitrag zu einer objektivierten Beurteilung, aber auch zu einer Einschätzung der Fähigkeiten nach einem förderorientierten Kriterienraster“ (Stellwerk-Website, www.stellwerk-check.ch, Zugriff 29.3.2016)⁵. Ein Probelauf mit dem Stellwerk-Test durch den Autor und dessen Beobachtung der Schülerinnen und Schüler in Eglisau zeigt Folgendes:

Die Testaufgaben ...

- > sind interessant und abwechslungsreich;
- > sind unabhängig von Lehrmitteln und trotzdem inhaltsbezogen und meist kontextualisiert;

- > erfordern genaues Lesen der Anweisungen und Aufgabenstellungen, welche auf Deutsch gehalten sind;
- > berücksichtigen verschiedene Lese- und Hörmodalitäten (*Skimming, Scanning*).

Die Qualität der Hörstücke ist gut, diejenige der wenigen Videos eher unbefriedigend. Die Ergonomie lässt zu wünschen übrig (kleine Schriften, Bildschirmgrösse wird schlecht genutzt). Auch kann die Tatsache, dass die Anweisungen deutsch gegeben werden, Jugendlichen Mühe bereiten, die nicht in der Schweiz eingeschult werden, z.B. Asylsuchenden, deren Englisch allenfalls besser ist.

Das Testsystem als Ganzes (vgl. Moser, 2006)...

- > passt sich relativ rasch dem Wissen und der Leistungsfähigkeit der Kandidatin an – schwankende Leistungen können den Test aber verlängern;
- > verwendet viele unterschiedliche Prüfformate, was für einige Lernende verwirrend sein kann;
- > erstellt nach den Teilbereichen (Hören/Lesen/Sprachliche Mittel) differenzierte individuelle Profile;
- > stellt die Ergebnisse als Punktzahlen (Range 200 bis 800, Skalenmittelwert = 500) dar und bietet eine Interpretationshilfe mit Kann-Beschreibungen, die aber nicht genau mit denjenigen des GER oder von *lingualevel* (Lenz & Studer, 2007) übereinstimmen;
- > erlaubt keine umfassenden Aussagen zu Fremdsprachenkompetenzen, da produktive Fertigkeiten nicht getestet werden.

Zu vermuten ist, dass die Schülerinnen und Schüler gut auf den Stellwerktest ansprechen, weil sie sich in erster Linie mit sich selbst messen und weil die Aufgaben sich rasch in dem Bereich einpendeln, wo sie nicht zu hohen Ansprüchen ausgesetzt sind, aber auch nicht unterfordert werden.

Nach der Vermittlung der Ergebnisse nach Fächern und Fertigkeiten durch die Lehrerin geht es für Lea darum, die festgestellten Defizite zu kompensieren und bekannte Stärken auszunutzen (z.B. strategische Kompetenzen auf andere Fertigkeiten zu transferieren). Sie kann dies mit dem massgeschneiderten Instrument „Lernpass“ (www.lernpass.ch) tun, welches für jeden der drei Stellwerk-Testbereiche in Englisch drei Module auf verschiedenen Anspruchsniveaus bietet. Die

⁵ Im Bildungsraum Nordwestschweiz gibt es mit „Check-dein-Wissen“ ein ähnliches System; es ist mit der Aufgabensammlung „Mindsteps“ verknüpft. Interessant dabei ist, dass Schreibkompetenzen von speziell geschulten externen Personen beurteilt werden. Die Check-Ergebnisse sollen ausschliesslich für die individuelle Förderung und die Unterrichtsentwicklung genutzt werden – nicht also als Teil eines Leistungsprofils für Lehrbetriebe.

Förderung mit Lernpass wird von einem Dock aus gesteuert, die Lehrerin erhält Einblick in das Lernjournal und kann Lea bei der Planung unterstützen. Dennoch, die Arbeit kann mühsam werden, die Anleitungen sind lange und in einem etwas verwirrenden Mix von Deutsch und Englisch gehalten, und das Layout der Hefte ist nicht besonders ansprechend. Zudem könnte die Zeit fehlen, um neben der Arbeit mit dem Lehrmittel und der ersten Lektüre von *Graded Readers* die individuelle Förderung seriös zu betreiben. Es kann sich für die Lehrperson lohnen, auf Grund der Ergebnisse der ganzen Klasse Schwergewichte zu setzen oder Schülergruppen zusammen an bestimmten Fertigkeiten arbeiten zu lassen. Im 9. Schuljahr wird der Stellwerk-Test wiederholt, und in den meisten Fällen werden Fortschritte sichtbar gemacht werden können. Wichtig ist, dass in der Zwischenzeit die von Stellwerk nicht erfassten Kompetenzen laufend gefördert und diagnostiziert werden.

3. Eintrittsprüfung in die Berufsmaturitätsschule (BMS)

Die Crux mit dem Stellwerk-Test zeigt sich nicht etwa in seiner Anwendung als Instrument der Diagnostik und Lernförderung, sondern bei seiner Nutzung als Wegweiser in der Laufbahnplanung. Die Ergebnisse des Tests können nämlich auf sogenannte Job-Skills-Profil abgebildet werden. Angenommen Lea strebt eine kaufmännische Lehre an, dann findet sie auf www.jobskills.ch idealtypische Profilvergaben für den Einstieg in den entsprechenden Beruf. In Englisch müsste sie bereits für die Basis-Ausbildung (Kauffrau EFZ-Profil B) für alle Fertigkeiten eine Punktzahl zwischen 550 und 650 erzielen. Für das erweiterte Profil (Kauffrau EFZ-Profil E) bräuchte sie zwar nicht mehr Punkte, müsste aber in Französisch dasselbe Niveau erreichen. Die Stellwerk-Tests erlangen so eine Triage-Funktion. Die Verantwortlichen der Lehrbetriebe werden explizit dazu eingeladen, neben den Zeugnisnoten die Stellwerk-Profile der Bewerberinnen und Bewerber zu beachten und sie mit den Vorgaben des Lehrberufs vergleichen. Damit stehen diese ursprünglich als Instrumente der Diagnostik und Förderplanung konzipierten adaptiven Tests in Konkurrenz mit den von kommerziellen Anbietern vermarkteten „Checks“ (vgl. Abschnitt 4). Wenn Lea bei ihrem Ziel des kaufmännischen Berufs bliebe, sich jedoch während

der Berufslehre auf die höhere Weihe der Berufsmaturität vorbereiten wollte, würde sie auf den einzigen Fall in der Karriere einer Volksschülerin treffen, in der eine Prüfung im Fach Englisch mit darüber entscheidet, ob der Übertritt in eine weiterführende Schule gelingt⁶.

Die 30-minütige schriftliche BMS-Prüfung besteht aus drei Teilen: 1. Leseverstehen inkl. Wortschatz; 2. Verfassen eines Texts von ca. 60 Wörtern; 3. Grammatik mit 2 Aufgaben. Die Musterprüfungen zeigen auf, dass es in weiten Teilen weniger um den Nachweis von funktionalen Kompetenzen denn um das Vervollständigen von Lücken und Umformen von Strukturen geht. Beispiele:

- > *In autumn most _____ (leaf) fall from the trees.*
- > *She waved _____ (cheerful).*
- > *We cut the grass every Saturday. The grass _____ every Saturday.*

Im Vergleich zu diesen BMS-Aufgaben sind die Stellwerk-Items besser kontextualisiert und die Anleitungen leichter verständlich. Einige der Verstehensfragen wirken eher skurril: „*Ms Kondo believes that socks should get a rest. True or False?*“ Die Antwort lässt sich leicht im Text finden, dessen Alltagsrelevanz und Altersgerechtigkeit indes in Frage zu stellen wäre. Grundsätzlich bildet diese Art von Prüfung den Lehrplan und die unterrichtlichen Lerngelegenheiten der Sekundarstufe I ungenügend ab. Es ist bedauerlich, dass sich ausgerechnet die einzige „*high-stakes*“-Prüfung am Ende der obligatorischen Schule wenig an den Pfeilern der zeitgemässen Fremdsprachendidaktik und den ihr zugrundeliegenden Lerntheorien orientiert.

Weit ärgerlicher noch ist allerdings die Diskrepanz zwischen den Desideraten der Didaktik- und Testwissenschaften und den „*Multicheck*“-Tests, der in der Welt der Berufsbildung gehandelten „*Leitwährung*“.

4. Testsysteme im Dienst der Lehrlingsselektion

Gemäss einem Artikel in der Wochen-Zeitung vom 17.9.2015 haben innerhalb eines Jahres 30'000 Stellensuchende (zumeist Jugendliche) einen Test der Firma Multicheck absolviert. Viele Lehrlingsverantwortliche in Schweizer Firmen verlangen diesen kostenpflichtigen Test, der mehrfach wiederholt werden kann, „weil die Vergleichbarkeit der Schulzeugnisse nicht gegeben“ sei. Die

⁶ Beim Übertritt von der 2. oder 3. Sekundarstufe ins Kurzgymnasium legen die Zürcher Schülerinnen und Schüler eine Prüfung in Französisch ab.

Kritisch zu vermerken bleibt die Tatsache, dass bei wichtigen Übergängen, wo viel auf dem Spiel steht, teilweise ungeeignete Beurteilungsinstrumente eingesetzt werden.

WoZ schreibt dazu: „... der Test wird von den Firmen oft genutzt, um diejenigen ausscheiden zu können, die nicht zu einem Vorstellungsgespräch eingeladen werden. Entsprechend ist der Multicheck zu einem wichtigen Selektionsinstrument geworden und die Vorbereitung darauf zu einem lukrativen Markt für Lernstudios und Privatschulen.“ Man mag die sozialkritische Ausrichtung des zitierten Artikels teilen oder nicht, die Macht des Gate-Keeping-Werkzeugs „Multicheck“ wäre wohl weniger umstritten, wenn minimale Anforderungen an die fachliche Unterfütterung erfüllt wären. Wenn man von der Relevanz und Vergleichbarkeit der Musteraufgaben auf der Website der Firma ausgehen kann, kommt der Sprachdidaktiker nicht umhin, die fachliche Qualität des Tests grundsätzlich in Frage zu stellen. Einige Beispiele genügen. Unter der Rubrik „Richtige Grammatik erkennen“ muss der einzig korrekte Satz in der Auswahl gefunden werden (siehe Abb. 1), oder das Nomen „City“ (im Test gross geschrieben) muss in die Mehrzahl gesetzt werden. Leseverstehen beschränkt sich auf das Finden der richtigen Textaussagen in Einzelsätzen wie „Bonny and Clyde have a dog called Susie“ (Abb. 2). Handlungsorientierte Textaufgaben sucht man vergeblich. Von einer Ausrichtung an Lehrplänen oder am Referenzrahmen kann keine Rede sein. Der „Check“ entspricht in keiner Weise den Anforderungen eines Kompetenznachweises auf der Sekundarstufe I. Auch wer hier ein gutes Resultat erzielt, kann damit keineswegs demonstrieren, dass er/sie über praktisch einsetzbare Englischkenntnisse verfügt. Schülerinnen wie Lea, die 100 Franken für den Test ausgegeben haben, müssen sich vollkommen veräppelt vorkommen.

Abbildung 1

Beispielaufgabe ohne Zeitbeschränkung

Richtige Grammatik erkennen.

Dir werden Sätze vorgegeben. Nur einer dieser Sätze ist grammatikalisch korrekt. Klicke an, welcher der grammatikalisch korrekte ist.

- When begins the race?
- When do the race begin?
- When does the race begin?
- When does the race begun?

Abbildung 2

Beispielaufgabe ohne Zeitbeschränkung

Richtige Textaussage finden.

Zu einem Text findest du jeweils vier Aussagen, von denen nur eine zutrifft. Markiere die zutreffende Aussage.

Bonny and Clyde have a dog called Susie

- Bonny's dog is called Clyde.
- Their dog is calling Susie
- Susie is a dog
- Clyde's dog is called Bonny

Der Bildungsökonom Michael Siegenthaler (2010) kommt nach empirischer Analyse von 300 Lernenden der Migros zur ernüchternden Schlussfolgerung, dass die prädiktive Validität des Tests sehr klein ist; so korrelieren Ergebnisse im Multicheck nicht mit der Wahrscheinlichkeit einer Lehrvertragsauflösung und der Test trage auch wenig dazu bei, die Noten der Lernenden in der Berufsschule vorherzusagen. Er leiste somit nur einen geringen Beitrag für den Lehrbetrieb, die Kompetenzen der Schulabgänger besser zu signalisieren als dies Schulnoten und Berichte aus Schnupperlehren bereits tun.

Was die Fremdsprachenfächer betrifft, ist der Aussagewert eines Diagnostikinstrumentes, das sich nicht am Referenzrahmen GER und somit an den Lehrplänen und den bekannten internationalen Zertifikaten orientiert, ohnehin nahe bei Null. Besorgt fragt man sich, wie es kommt, dass Lehrmeister und Berufsbildungsverantwortliche diesen und ähnliche „Checks“ ernst nehmen können.

5. Schlussbemerkungen

Unser lernbiografischer *Tour d'horizon* hat gezeigt, dass Volksschüler und -schüle-

rinnen im Verlaufe ihrer Lernkarriere immer wieder Situationen der Leistungsbeurteilung ausgesetzt sind, und dass, wenn Lehrpersonen und Prüfungsbehörden die Instrumente bedarfsgerecht einsetzen, Gewähr für eine laufende Einschätzung des Lernstandes geboten wird. Kritisch zu vermerken bleibt die Tatsache, dass bei wichtigen Übergängen, wo viel auf dem Spiel steht, teilweise ungeeignete Beurteilungsinstrumente eingesetzt werden. Beim Übertritt zwischen Primar- und Sekundarstufe fehlt in unserem Fall die lernerfreundliche Abstützung auf das sprachenübergreifende ESP. Bei der einzigen prüfungsgestützten Eintrittshürde, der BMS-Prüfung während dem 3. Sekundarschuljahr kommt ein reduzierter, grammatiklastiger Kurztest zum Einsatz, und die Validierung der schulisch erworbenen Kompetenzen durch Lehrbetriebe bedient sich allzu oft eines gänzlich ungeeigneten Tests.

Es geht an dieser Stelle nicht darum, den einen gegen den anderen Testkoffer auszuspielen. Die informellen Beobachtungen im Klassenzimmer am Anfang dieses Beitrags, gepaart mit der kritischen Betrachtung gängiger Evaluationspraktiken, lassen jedoch Raum für Zukunftsperspektiven. Mit dem Konzept „*Assessment for Learning*“ (z.B. Chappuis, 2009; Clarke, 2005) kann eine Basis geschaffen werden, um nicht in erster Linie Aussagen über das Passieren von oder das Scheitern an einer Eintrittsschwelle zu treffen, sondern um das Lernen zu verbessern. Im Fall des Stellwerk-Tests und dessen Auswertung stehen die Zeichen gut, dass die wertvollen Informationen des Testings auch nutzbringend in den weiteren Lernprozess eingebracht werden können, vorausgesetzt bestimmte Bedingungen werden erfüllt:

Die Kompetenzbeschreibungen und Aufgabenstellungen müssen teilweise an den Lehrplan 21 angepasst werden.

Der bestehende Test sollte zumindest mit dem Teilbereich „Schreiben“ ergänzt werden (für Deutsch als Schulsprache gibt es bereits Schreibaufgaben); lehrwerk-

unabhängige mündliche Tests könnten aus *lingualevel* bezogen werden.

Für den Einsatz von Lehrmitteln als diagnostikgesteuerte Förderinstrumente braucht es didaktische Konzepte und eine bessere Koordination mit den gängigen Klassenlehrmitteln.

Die Förderplanung sollte nicht einseitig auf das Kompensieren von Defiziten („Wissenslücken stopfen“, Originalton Stellwerk) ausgerichtet sein, sondern auch Wege aufzeigen, um Stärken und Strategien gezielt zu transferieren.

Was die Validierung der Stellwerk-Testergebnisse betrifft: auch wenn es ein wenig unpassend anmutet, dass diagnostische Befunde für das weitere Lernen in der Schule als Teil eines Bewerbungsdossiers an Lehrbetriebe verschickt werden, müsste diese Möglichkeit weiter verfolgt und systematisiert werden – nur schon wegen der verfälschten Konkurrenzsituation mit den kommerziellen „Checks“. Für den Lehrmeister oder die Berufsschullehrerin wäre es ohne Zweifel wertvoll zu wissen, wie sich das Leistungsprofil der Schülerin Lea, die sich für eine KV-Lehrstelle beworben hat, innerhalb des letzten Schuljahres der Sekundarstufe entwickelt hat. Mit den zusammengefassten Ergebnissen aus den Stellwerktests 8 und 9 bekämen diese Schlüsselpersonen statt eines wertlosen Schnappschusses einen Einblick in das realisierte Entwicklungspotenzial der Schülerin.

Literatur

Achermann, B. & Ramsey, G. (2012). *first choice – Assessment Pack*. Zürich: Lehrmittelverlag Zürich.

Achermann, B. & Stauffer, K. (2009). *Moving on. Explorers 3*. Zürich: Lehrmittelverlag Zürich.

Chappuis, J. (2009). *Seven Strategies for Assessment for Learning*. Boston: Pearson.

Clarke, S. (2005). *Formative Assessment in the Secondary Classroom*. London: Hodder Education.

Europarat, Rat für kulturelle Zusammenarbeit (2001). *Gemeinsamer Europäischer Referenzrahmen für Sprachen: Lernen, lehren, beurteilen*. München: Langenscheidt.

Keller-Bolliger, R. & Haller, K. (2010). *Assessment Pack voices 1*. Zürich: Lehrmittelverlag Zürich.

Megías Rosa, M. & Stotz, D. (2010). Moving on, broadening out. A European perspective on the transition between primary and secondary foreign language learning. *Babylonia* 1/10, 30-34.

Moser, U. (2006). *Wie werden die Ergebnisse in den Stellwerk-Tests interpretiert?* Zürich: Kompetenzzentrum für Bildungsevaluation und Leistungsmessung an der Universität Zürich. http://www.stellwerk-check.ch/Uploads_Cy mos/UploadedDocuments/67_Document.pdf (Zugriff am 30.3.2016).

Siegenthaler, M. (2010). *Was kann ein standardisierter Berufseignungstest, was die Schulnoten nicht können? Eine quantitative Fallstudie zum Signalwert zweier Instrumente der Lehrlingsselektion*. Universität Bern: Masterarbeit. <http://edudoc.ch/record/103442/?ln=de> (Zugriff am 30.3.2016).

Studer, T. & Lenz, P. (2007). *lingualevel. Instrumente zur Evaluation von Fremdsprachenkompetenzen*. Bern: Schulverlag blmv plus.

L'ÉPREUVE CANTONALE DE RÉFÉRENCE D'ALLEMAND: UNE NOUVEAUTÉ EN PAYS DE VAUD

Der vorliegende Artikel schildert, wie im Kanton Waadt anhand eines kantonalen Deuschtests am Ende der Primarstufe die sprachlichen Kompetenzen der Schülerinnen und Schüler evaluiert werden (*Epreuve cantonale de référence d'allemand*, 8H). Die aufwändige Konzeption der Tests (Niveaubestimmung, Auswahl der Themen und Aktivitäten, Redaktion der Prüfung durch ein Expertenkomitee, Durchführung eines Probestests, statistische Analyse der einzelnen Test-Items; Gender-Konformität, Redaktion von Korrekturanweisungen) wird genau beschrieben und begründet. Die doppelte Funktion der ECRs, einerseits zur Kontrolle der Lernstandards, andererseits zur Verwendung als Orientierungsinstrument wird im Anschluss daran anhand von Reaktionen aus der Praxis dargestellt und problematisiert.

● Claudia Bartholemy,
Olivier Bolomey |
HEP-VD
Pascal Schweitzer |
Unil

Claudia Bartholemy est chargée d'enseignement en didactique de l'allemand langue étrangère à la HEP-Vaud et MER à la Faculté des Lettres de l'Université de Lausanne.



Olivier Bolomey est chargé d'enseignement en didactique de l'allemand langue étrangère à la HEP-Vaud et enseignant d'allemand dans l'Etablissement primaire de Gimel-Etoy.



Pascal Schweitzer est enseignant d'allemand au Centre de langues de l'Université de Lausanne et chargé d'enseignement en didactique de l'allemand langue étrangère à la HEP-Vaud.



Dans le canton de Vaud¹, suite à l'entrée en vigueur de la nouvelle Loi sur l'enseignement obligatoire à la rentrée scolaire 2013, l'évaluation externe a été renforcée en introduisant une nouvelle épreuve cantonale de référence (ECR) pour l'allemand en fin de 8P².

Depuis 2002, seuls le français et les mathématiques faisaient l'objet d'une ECR conçue par la Direction générale de l'enseignement obligatoire (DGEO). Ces épreuves interviennent à intervalle régulier tout au long de la scolarité obligatoire (uniquement le français en 4P, le français et les mathématiques en 6P, 8P et 10S). Avec l'introduction d'une ECR d'allemand en 8P, le législateur a souhaité renforcer le statut de cette discipline. L'ECR dure 90 minutes, porte sur un total minimum de 60 points et a lieu en général début mai. Dans les autres cantons romands, l'allemand fait l'objet d'une évaluation externe depuis plusieurs années déjà. Fribourg ou Genève, par exemple, effectuent des épreuves de référence en 6P, en 8P et au cycle 3.

Objectifs d'apprentissage visés

L'ECR d'allemand teste l'atteinte des standards minimaux définis par la CDIP (2011). Elle est basée sur trois activités langagières: *écouter, lire et écrire*. Le niveau A1.2 du Cadre européen commun de référence pour les langues (CECR) est visé, conformément aux objectifs du Plan d'études romand (PER). Les éléments du fonctionnement de la langue (telles la maîtrise de l'orthographe et la conjugaison) sont évalués en contexte.

Conception

A l'instar des autres épreuves communes vaudoises, l'ECR d'allemand est réalisée par un groupe de travail composé d'enseignants³ du cycle concerné et d'un expert didactique de la Haute école pédagogique du canton de Vaud (HEP). Afin de garantir le niveau de langue visé (niveau A1.2 du CECR) et de contrôler la qualité de celle-ci, deux enseignants du Centre de langues de l'Université de Lausanne (CDL) complètent l'équipe. Les activités sont rédigées par un enseignant du degré

Lukas écrit un e-mail à un ami pour lui parler de Lisa, une camarade de sa classe qui habite son quartier.

- Tu dois présenter Lisa.
- Dans chaque thème proposé ci-dessous, choisis un dessin ou une information.
- Entoure le dessin ou l'information que tu choisis.
- Lis ensuite les consignes du guide de production.

concerné (bénéficiant de 3 périodes de décharge hebdomadaire). Le groupe de travail est coordonné par un collaborateur pédagogique de la DGEO.

La réalisation de l'ECR est un processus relativement long et complexe: près de deux ans de travail et de nombreuses étapes de validation jalonnent le parcours de la future épreuve. Les thématiques choisies pour le «fil rouge» de l'épreuve et les différentes activités sont liées au moyen d'enseignement officiel⁴. Elles ont nécessairement un lien avec la culture des pays germanophones.

Lors de la rédaction de l'ECR, les thèmes sont également analysés au travers de l'approche genre et doivent éviter de prêter à polémique. Il est aussi primordial que les thèmes et le fil rouge proposés soient proches de la réalité des élèves. A titre d'exemple, le fil rouge de l'ECR 2014 est consacré à un enfant de 12 ans qui déménage de Bâle à Berlin. Les différentes activités proposées permettent ensuite à cet enfant – et à tous les élèves qui réalisent l'épreuve – de découvrir quelques aspects de Berlin et de la vie dans la capitale allemande.

Le groupe de travail veille à concevoir des textes proches de situations de communication authentiques. Pour la compréhension de l'oral, le débit est, par exemple, adapté au niveau des apprenants (A1). Les enregistrements des textes sont réalisés à la HEP par des personnes de langue maternelle allemande.

La réalisation d'un cahier de correction complète l'épreuve. Il doit contenir toutes les réponses possibles – et plausibles – pour chaque activité. Il met également en exergue des exemples de productions écrites d'élèves avec une proposition de corrigé et d'attribution des points pour chaque critère. Ce document, destiné aux correcteurs, contribue à un traitement égalitaire de tous les élèves.

Enfin, pour garantir la qualité de l'épreuve, l'ECR est pré-testée près d'une année à l'avance dans plusieurs classes. Ces pré-tests, effectués dans une douzaine d'établissements sur environ 250 élèves, permettent de récolter des informations pertinentes sur la qualité des items et la justesse des activités proposées. Les résultats sont minutieusement analysés par un statisticien. Ces différentes données permettent d'ajuster l'épreuve et de la finaliser.

Outre les pré-tests et leur analyse statistique, les ECR de 8P (français, mathématiques et allemand) sont construites sur la base de la théorie de la réponse aux items (TRI) qui postule notamment que chaque réponse donnée par l'élève équivaut à un point. Dans le cadre de l'expression écrite, la rigueur statistique imposée par la méthode TRI ne permet pas d'effectuer une évaluation globale du texte produit par l'élève. Ceci conduit à une évaluation critériée pour chaque phrase. De plus, le lexique utilisé est lié au moyen d'enseignement et est forcément limité en raison du niveau visé.



Meine Freundin Lisa

Thema 1



27. Juli

Thema 2




Exemples de productions d'élèves

- Les phrases supplémentaires rédigées par l'élève ainsi que les éléments hors sujet ne sont pas pris en compte ! Ces phrases sont grisées.
- Les autres erreurs sont en **caractère gras**.
- Les mots manquants sont signalés par un trait _____.

Texte 1

Lisa ist 12. Sie hat **ein** Freund. Sie hat **ein** Hund. Sonntag **sie** trifft seine Freunde.
Lisa mag radfahren. Meine Freundin mag nicht kochen.

Bis bald Dein Lukas

	Prod.	FdL.
Thème 1 Lisa ist 12.		
Message compréhensible et adapté à la situation	1 / 1 pt	
Le verbe est conjugué correctement		1 / 1 pt
Thème 2 Sie hat ein Freund.		
Message compréhensible et adapté à la situation	1 / 1 pt	
Le verbe est conjugué correctement		1 / 1 pt
Le déterminant est adéquat		1 / 1 pt
Thème 3 Sie hat ein Hund.		
Message compréhensible et adapté à la situation	1 / 1 pt	
Le verbe est conjugué correctement		1 / 1 pt
Le déterminant est adéquat		1 / 1 pt

- 1 Les analyses et avis émis dans cet article engagent les auteurs et non l'institution scolaire.
- 2 Nous employons la numérotation HarmoS. La 8P correspond donc à la dernière année de l'école primaire.
- 3 Le masculin est employé dans cet article comme genre générique.
- 4 Geni@! A1 des éditions Langenscheidt. Ce moyen sera remplacé à la rentrée scolaire 2017 par Junior 8 des éditions Klett-Langenscheidt.

5 Association of Language Testers in Europe: www.alte.org. Le Goethe Institut et ÖSD (Österreichisches Sprachdiplom Deutsch) qui ont, entre autres, inspiré le groupe de travail font partie de cette association.

6 En fin de 8P, lors de la procédure d'orientation, la moyenne finale en français, mathématiques et allemand prend en compte les résultats des ECR à hauteur de 30%.

Références

CDIP (2011). *Compétences fondamentales pour les langues étrangères. Standards nationaux de formation adoptés par l'Assemblée plénière de la CDIP.*

CIIP (2010). *Plan d'études romand (PER).* Consultable sur le site Internet: www.plandetudes.ch

Conseil de l'Europe (2005). *Cadre européen commun de référence pour les langues.* Paris: Editions Didier.

Endt, E., Koenig, M., Ritz-Udry, N. & Pistorius, H. (en cours d'édition). *Junior – Deutsch für die Romandie – Kursbuch 8. Klasse.* München: Klett-Langenscheidt.

Funk, H., Koenig, M., Koithan, U. & Scherling, T. (2002). *Geni@I – Deutsch als Fremdsprache für Jugendliche – Kursbuch A1.* München: Langenscheidt.

Mottier Lopez, L. (2015). *Evaluations formative et certificative des apprentissages. Enjeux pour l'enseignement.* Louvain-la-Neuve: De Boeck.

Studer, T. (2012). Leistungsbeurteilung: Testfunktionen als Orientierungshilfe. – Positiver Washback: Mit lingualem Sprachkompetenzen beurteilen. *Grundschulmagazin Praxis* 5, 7-12.

Tessaro, W. & Ntamakiliro, L. (2010). *Les évaluations externes en Suisse romande: enjeux, pratiques et effets.* Université de Genève: FAPSE.

Exemple d'ECR d'allemand de fin 8P

Un exemple peut être téléchargé sur le site internet suivant:

<http://www.vd.ch/themes/formation/scolaire-obligatoire/evaluation/ecr/>

Le groupe de travail «ECR d'allemand 8P» veille également à un certain équilibre entre les différentes parties de l'épreuve qui testent à chaque fois une activité langagière différente. En outre, à l'intérieur de chacune des parties, il est également important de respecter une certaine gradation dans le degré de difficulté des activités. A titre d'exemple, dans les parties écouter et lire, les premières activités sont en général basées sur la compréhension globale avant de se diriger vers une compréhension plus sélective. Les concepteurs de l'épreuve s'appuient sur des exemples proposés par d'autres cantons romands ainsi que sur des examens internationaux⁵.

Quid de la production de l'oral?

Pour l'heure, la DCEO, pour des raisons organisationnelles, renonce à tester cette activité langagière lors de la passation des ECR alors qu'elle constitue l'objectif central de l'enseignement des langues étrangères au deuxième cycle primaire. Néanmoins, afin d'encourager les établissements à entraîner la production orale (en interaction et en continu) et à la tester, un matériel d'évaluation «clé en main» est proposé aux équipes pédagogiques qui peuvent ainsi l'employer comme une évaluation formative ou sommative.

Effet «washback»

Par ailleurs, la manière d'évaluer a des répercussions tangibles sur la manière d'enseigner: on parle de l'effet «washback». L'un des grands dangers du «washback» est que l'enseignement se limite à ce qui est testé, d'autant plus si les résultats comptent pour l'orientation (Tessaro & Ntamakiliro, 2010). Cependant, l'effet «washback» peut aussi être bénéfique pour certaines pratiques enseignantes: il permet potentiellement à un grand nombre d'enseignants de se familiariser avec des démarches didactiques visant l'atteinte des standards. Une telle épreuve

peut moduler progressivement les habitudes d'enseignement (par exemple en testant séparément les activités langagières et en proposant des situations de communication transférables à la vie réelle).

Bilan

L'évaluation est omniprésente dans le monde de l'éducation et de la formation. Il semble même que l'on tende, aujourd'hui, à vouloir plus évaluer et tout évaluer: des apprentissages, des enseignements, des dispositifs, des programmes, des établissements, des systèmes éducatifs, des politiques, etc. (Mottier Lopez, 2015: 7).

Le canton de Vaud n'échappe pas à ce constat. L'ECR d'allemand de fin de 8P doit, dans ce contexte, relever un double défi: elle doit assumer des fonctions de pilotage du système éducatif, tout en étant employée à des fins d'orientation vers les voies et niveaux du degré secondaire⁶. Cependant, le groupe de travail a tenté de relever ces défis. Il a fait de l'ECR d'allemand, non pas seulement un outil de sélection, mais avant tout un outil pédagogique favorisant l'enseignement-apprentissage. La variété et la pertinence des activités proposées soutiennent le développement des différentes activités langagières testées.

Pour le futur de cette ECR, les concepteurs devront s'adapter à un nouveau public scolaire. En effet, les élèves du cycle 2 auront suivi un enseignement formalisé durant quatre ans (et non plus des élèves ayant reçu une sensibilisation durant deux ans, puis un enseignement formalisé durant deux ans). Cette nouvelle répartition horaire des périodes d'allemand permettra toujours d'atteindre les objectifs d'apprentissage du PER et le niveau A1 du CECR. De plus, les élèves devraient davantage développer leurs compétences en compréhension et production de l'oral, ce qui à terme pourrait permettre d'intégrer l'évaluation de cette dernière.

EARLY READING AND WRITING IN THE FOREIGN LANGUAGE CLASSROOM – MAYBE NOT AS CONTROVERSIAL AS ONE MIGHT THINK

Es gibt offenbar keinen Grund, warum man Erstklässler (6-7-Jährige) nicht im Schreiben in der Fremdsprache, insbesondere in Rechtschreibung, unterrichten sollte. Tatsächlich kann ein Fokus auf orthographische Strukturen und das einfache Bewusstsein der Idee, dass die deutsche und die englische/französische Rechtschreibung sich unterscheiden, Lernen vereinfachen und Grundkenntnisse zu einem Zeitpunkt vermitteln, zu dem Lernende extrem motiviert sind. Der Artikel berichtet über Forschungsergebnisse aus Deutschland und ermutigt Lehrkräfte mit dem Schreiben und der Rechtschreibung direkt zu Anfang des Fremdsprachenunterrichts zu beginnen.

When early literacy acquisition in the foreign language classroom was discussed in Germany about ten years ago, one usually heard opposing voices. At the beginning there were mostly those (parents and teachers alike) who believed that the children would be asked too much of, that it was hard enough for many children to learn German orthography, and that, consequently, the young learners would end up overtaxed and confused. This opinion was especially held in those federal states (e.g. Baden-Württemberg and Rhineland-Palatinate) where English and French classes started right at the beginning, in grade 1. However, you also found some voices, researchers mainly, who took a rather sceptical stance towards two years of oral instruction only which was the educational policy in those states. These researchers were convinced that the children would want to read and even to write in the foreign language, too, as soon as they started their literacy classes in the language of schooling in the first grade (cf. Rymarczyk, 2008). In the years to follow, more and more teachers became qualified to teach for-

eign languages at primary level, and with the resulting more professional mindsets, expectations and conceptions changed as far as the character and goals of foreign language learning outcomes at primary level were concerned. It was realized and appreciated that primary school learners were able and willing to do more in the English or French classroom than to colour in mandalas in order to learn some colour terms or to learn some basic phrases to tell their names and how old they were. Teachers started to have children work on their language awareness, encouraged them to expand their vocabulary knowledge and allowed them to tackle the foreign language in a more analytical way every now and again as the situation arose (cf. Diehr & Rymarczyk, 2012). As more research on the process of simultaneous literacy acquisition in German and English or French in the first two years of schooling was conducted, it became obvious that even children who were considered poor learners transferred their knowledge of how to read and write in German onto their ideas of how to do

Jutta Rymarczyk | ●
Heidelberg, DE

teaches at Heidelberg University of Education, Germany, in the Department of English as a Foreign Language. Her current research interests focus on early literacy development and diagnostics, immersion programmes at primary and secondary level (Bilingualer Sachfachunterricht/CLIL) and the role of visual literacy in this context.



PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > [SCHEDE 6 E 20B](#)



The mistakes prove that the children can already do quite a lot. Instead of regarding their spellings as flawed, teachers might rather want to appreciate the children's achievements.

this in the foreign language. While this is basically a positive finding, it followed that the children need support in their transfer so that the results correspond to the orthographic norms of the target language English or French (cf. Rymarczyk, 2010).

Survey findings: Nothing wrong with writing

In order to find out how widespread the change of paradigms in primary foreign language classrooms actually was and whether educational policies corresponded to experts' conviction of what was beneficial for the young learners, two surveys were carried out in Germany in 2015. Experts working e.g. at various institutes for school education and advanced teacher training (so-called "Landesinstitute") or at universities and representing all 16 German federal states were first asked to report on how (among other issues) early literacy was dealt with in their state, and in a second survey they were then asked which of the ways of proceeding, which had been reported to be the modus operandi in one or more of the various states, they would actually want to implement or continue to use, as the case may be (cf. Hempel, Kötter, Rymarczyk & Steinlen, forthcoming).

The results of the first survey drew a very heterogeneous picture which is still valid at present: While the majority of states (ten) started to teach reading and writing in grade 3, there are also four (from 2016 on, five) states with an immediate start in grade 1 and one state (Bremen) where reading and writing is not taught before grade 4. As far as the methods are concerned, there are also differences among the states due to the individual state curricula: Whereas some states uses a holistic approach, the so-called whole-word method, and employ the visibility of written forms as a mnemonic device only, there are other states which first import orthographic rules, i.e. make use of the so-called phonics approach. Further regional differences are found in the levels of difficulty and complexity of what the children are expected to write. Here we find the full range from one-word text productions to complete sentences and from copying from text models to the free production of complete sentences without any models (ibid.).

Interestingly, the results of the second survey showed a far more homogeneous picture than the first survey. The discrepancies between educational policies and individual

votes were e.g. striking in reply to the question on the start of literacy teaching. While the curricula of eleven states prescribe a late start in grade 3 or even 4, with 66 % the majority of experts (23

out of a total of 35), voted for the earliest start possible, namely to start reading and writing instruction in grade 1.

As far as the expected goals are concerned, there was an even more homogeneous outcome. 77% of the experts opted for the most demanding goal to be set, namely

that children should be expected to write complete sentences by the end of grade 4 and that these text productions should be used for assessment.

The strongest agreement, however, was found in the answer to the question of whether one should use a phonics approach by pointing out differences in the German and English sound-letter-relationships and thus foster the children's orthographic awareness. 88% of the experts agreed that children should be made aware of the differences between the English and German orthography. It is to be noted, however, that a gentle way of using the phonics approach is favoured as there

were comments added like "only as discovery learning"; "not systematically" (ibid.). In order to take a stance in the debate about early reading and writing and to inform about findings of current research on early literacy acquisition, the following pages take up some central issues of the topic and finally present two ways of how children can assess their own spelling skills and practice reading and writing at an early level.

Should reading and writing instruction begin in grade 1?

Strongly agree	Agree	Undecided	Disagree	Strongly disagree
15	8	2	7	3

Should complete sentences be copied and written without any models at the end of grade 4 and be assessed?

Strongly agree	Agree	Undecided	Disagree	Strongly disagree
13	10	2	7	2

Should learners be made aware of sound-letter-relationships which differ between German and the L2?

Strongly agree	Agree	Undecided	Disagree	Strongly disagree
19	11	1	2	1

Arguments for an early start

First, we want to refer to the still wide-spread notion that reading and writing should not be taught before the children master these skills in the language of schooling. There are strong arguments against such a delayed start of reading and writing instruction. The grapheme-phoneme-correspondences of both English and French are not overly transparent and very irregular. Because of this there is a long so-called logographic phase (cf. Frith, 1998) at the beginning of learners' literacy acquisition process, especially in English where orthography is particularly irregular. In this phase, the learners perceive the words as a whole, similar to how they see a picture, i.e. they do not connect the letters to sounds. They do not really read the word but rather recognize a given picture. As it is, the whole idea of teaching sight words at the beginning of the literacy acquisition process is based on this "reading" strategy.

In contrast to this, German has a transparent and regular grapheme-phoneme-correspondence. This enables the learners to link the letters to the corresponding sounds fairly easily. Because of this learners reach the next phase of the literacy acquisition process, the so-called alphabetic phase (cf. Frith, 1998), relatively early. They are able to assign the individual letters or letter groups to sounds.

Most young children who mainly learn how to read and write in their German classrooms readily transfer their knowledge about German orthography to English as they are generally very eager to write in the foreign language, too. Of course, this transfer leads to mistakes in most cases due to the differences between the German so-called shallow orthography and the English deep orthography. Erroneous spellings like the following are typical for children who apply the German sound-letter correspondence to the English system.

Teachers should be aware that the children who make mistakes like this are by no means at the beginning of their literacy acquisition. To produce spellings like the ones in the example, they need to be in

the alphabetic phase of their literacy acquisition in German, the second phase of the acquisition process. Actually, the mistakes prove that the children can already do quite a lot. Instead of regarding their spellings as flawed, teachers might rather want to appreciate the children's achievements:

1. The children are able to hear all sounds that are contained in the word, i.e. their phonemic awareness seems to be high.
2. They know that these sounds need to be linked to letters – they just choose the wrong ones, namely the German set.
3. Children who write words like <Kwin> (green) and <Jelou> (yellow) link all of the sounds to letters. This proves that they are already at an advanced stage of the alphabetic phase as there is also an earlier stage when learners just write some of the letters that are contained in a word, the so-called "partial alphabetic phase" (cf. Ehri, 1999). Often, these letters are the consonants of the word as they are more salient. The vowels are missing – or to take up the metaphor of the technical term of this phenomenon, "skeleton writing" – there is no flesh on the bones so to speak: "My Mt Bt" for "my motor boat" (cf. Gentry, 2006) or "HS" for "Haus" (cf. Scheerer-Neumann, 2003).
4. They know that capital letters are used for objects and colours (in their nominal forms), however, they wrongly transferred the rule from German to English.

Of course, looking at the achievements of a young writer does not mean that teachers should restrain themselves from correcting the children's mistakes. A positive stance which appreciates the child's efforts rather than just looking out for deficits allows the teacher to diagnose the texts in a more differentiated way. Only such a close diagnosis enables the teacher to distinguish children whose mistakes are based on transference from German (which testifies the achievements 1-4 mentioned above) from those children who lag behind, make random mistakes or might actually suffer from dyslexia.



First grader, 6 years old.

Literature

Diehr, B. & Rymarczyk, J. (2012). Current developments in Teaching English as a Foreign Language in German Primary Schools. *Anglistik*, Volume 23, Issue 1, 13-23.

Ehri, L. (1999). Phases of Development in Learning to Read Words. In: J. Oakhill & R. Beard (Hrsg.). *Reading Development and the Teaching of Reading*. Oxford: Blackwell, pp. 79-108.

Frith, U. (1998). Psycholinguistische Aspekte orthographischen Wissens. In: E. Augst (Hrsg.). *New Trends in Graphemics and Orthography*. Berlin, New York: De Gruyter, pp. 218-233.

Gentry, R. J. (2006). *Breaking the Code. The New Science of Beginning Reading and Writing*. Heinemann.

Hägele, G. (2008). Let's read and write! *Grundschulmagazin Englisch*. H.1|2008. München: Oldenbourg, pp. 9-12.

Hempel, M., Kötter, M., Rymarczyk, J. & Steinlen, A. (forthcoming). Fremdsprachenunterricht in der Grundschule in den Bundesländern Deutschlands: Eine Bestandsaufnahme. In: J. Appel, S. Jeuk & J. Mertens (Hrsg.). *Sprachen lehren*. Dokumentation zum 26. Kongress für Fremdsprachendidaktik der Deutschen Gesellschaft für Fremdsprachenforschung (DGFF). Baltmannsweiler: Schneider Verlag Hohengehren, pp. 60-78.

Rymarczyk, J. (2008). Früher oder später? Zur Einführung des Schriftbildes in der Grundschule. In: H. Böttger (Hrsg.), *Fortschritte im Frühen Fremdsprachenlernen. Ausgewählte Tagungsbeiträge*. München: Domino, pp. 170-182.

Rymarczyk, J. (2010). Früher Schriftspracherwerb in der ersten Fremdsprache Englisch bei Kindern mit Migrationshintergrund. *Forum Sprache. Die Online-Zeitschrift für Fremdsprachenforschung und Fremdsprachenunterricht*. 4/2010, 60-78. http://www.hueber.de/seite/pg_hefte2010_fos#id191606.

Scheerer-Neumann, G. (2003). Rechtschreibschwäche im Kontext der Entwicklung. In: I. Naegle & R. Valtin (Hrsg.). *LRS – Legasthenie in den Klassen 1-10*. Bd. 1 Weinheim, Basel: Beltz, pp. 45-65.

Logographic strategies (e.g. word-picture cards, picture cards, and word cards attached to realia) can be used in the foreign language classroom, even though the alphabetic phase for German might not be completed and alphabetic strategies might not be coped with completely, respectively. Children at this stage might e.g. still produce the so-called skeleton-writing in German (<HS> for *Haus*), but be absolutely capable of assigning English word cards to pictures or realia. (Cf. Activity 1a.)

There will not be any disturbances by such holistically oriented, logographic exercises as first graders will be leaving behind the logographic phase in German after only a few weeks of instruction. As it is, many first graders are already in the alphabetic phase in German upon entering school due to story reading at home or kindergarten and print in their daily environment.

Second, children have to be made aware of the differences between the grapheme-phoneme-correspondences between German and English or French and be able to recognize them. The creation and usage of idiosyncratic orthographic rules ("invented spelling") turn out less severe if children are aware of the differences. Please note that the children do not have to know any explicit phonics rules at first. They just need to be alert, i.e. know that they cannot simply transfer everything they learned about letters and sounds from writing in German to writing in English or French.

Third, young learners need orthographically correct input which allows them to see possible differences between their hypotheses as far as the assumed writing of a word is concerned and its actual spelling. The input corrects the mistakes in the invented spelling and helps the children to write the words according to the orthographic norm. No child can be kept away from his or her own ideas about written forms – including poor learners. The later children become acquainted with the written forms of the foreign language, the greater is the danger of the fossilization of these mostly inevitably wrong forms. Also, the loss of motivation when the children are to learn the orthographically correct forms as late as grade three or even four can be enormous as the children happily and also effectively used their idiosyncratic rules up to this point in time. An informed and professional use of early literacy certainly serves the children best in

their actual phase of language acquisition and facilitates smoother roads later on.

Activities to practice and (self-)assess spelling competence with young learners

Both activities (1+2) can be used in such a way that they either support learners in their logographic phase, i.e. at the very beginning of their literacy acquisition process when they still perceive words as a whole, or at the alphabetic stage, the more advanced stage, when learners already work with the connections between sounds and letters.

Activity 1: Wordmonster

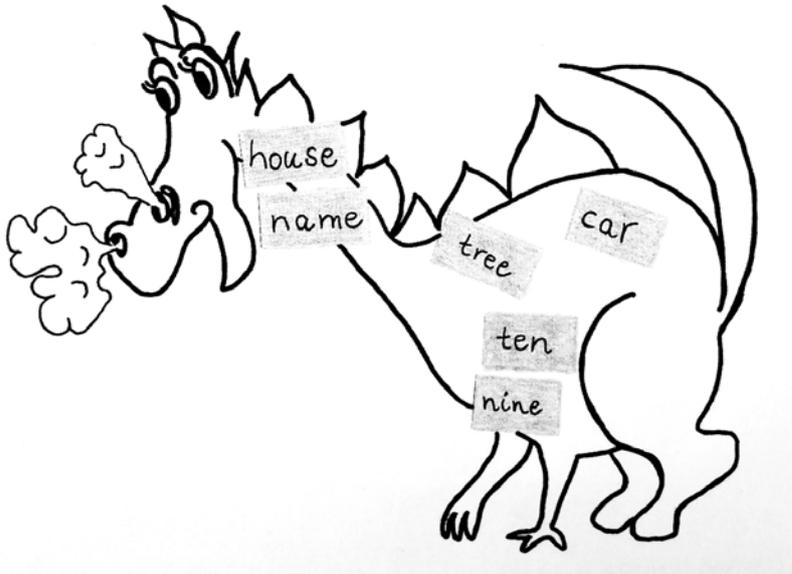
The "Wordmonster" is an animal that lives exclusively on written English words. Put the Wordmonster as a poster on the wall of your classroom. Every day the pupils need to feed it: Words that can be connected with a picture can be put into the monster's mouth. If the pupils still remember a certain word-picture-connection a few days later, the words can be put forward into the monster's belly. Words which seem to be more difficult for the children to remember, stay in the monster's throat (cf. Hägele, 2008).

> a) There are special days on which the monster only wants to eat toys, school things or only round objects, etc. (Holistic activity to be used at the logographic stage.)

> b) There are special days on which the monster only wants to eat words starting with , which have a silent <e>, or which contain the diphthong / aʊ/. (Phonics activity to be used at the alphabetic stage.)

Activity 2: Spelling-Bandolo

The "Spelling-Bandolino" is based on the learning game "Bandolo" (or "Bandolino") and needs to be re-designed by the teacher (or by advanced students!). A "Bandolo" is a fan consisting of about ten cards, each of which presents an activity on its own. A card with a vocabulary activity shows e.g. five pictures of objects on the left side and five words on the right side. There is an indentation next to each picture or word, respectively. The string which is attached at the point where the cards of the fan are held together is wound around the single card (which is fanned out from the stack) in such a way that the pictures get connected to the matching words using the indentations to be fixed.



Wordmonster



The intriguing part is the opportunity for the students to check their solutions easily by themselves: All they have to do is turn around the card and see whether the string is following the trace on the card. If so, they were right, if the string follows a path different from the trace on the card, they were wrong.

When you create your own “Spelling-Bandolo”, the pages of the original fan are pasted over with one’s own sheets which take up items which are of current interest for the students.

> a) You can put together pictures of objects and the corresponding words like e.g.: *cat, duck, banana, bird, apple, orange*. These words can be chosen from the vocabulary a child is struggling with or –as in this example – from an easy reader. (Here: Roderick Hunt and Alex Brychta: *Silly Races*. Oxford Univ. Press, 2005). (Holistic activity to be used at the logographic stage.)

> b) For more advanced and very advanced learners you can focus on orthographic rules like capital letters vs. lower case letters or on particularly complex sound-letter relationships. Again, the words can be chosen freely or be taken from a picture book. The following examples are chosen from Jon Blake and Axel Scheffler: *Daley B.* Walker Books (1994), “Orthographic and phonics activity to be used at the alphabetic stage”.



Original Bandolo game



Spelling-Bandolo: model of front (= exercise) and back (= trace of string to check one’s answers).

This Bandolino template was found here: <http://thomas-ebinger.de/2015/10/water-unser-bandolino-als-bastelvorlage/>



Fold here

Capital or lower case letters?

Start!			
1.	N	aley B.	
2.	n	og	
3.	P	orcupine	
4.	p	ew York	
5.	D	eter	
6.	d	est	

Fold here

Find the matching sounds of the letter $\langle a \rangle$!

Start!			
1.	weasel	ask	
2.	day	head	
3.	rabbit	saw	
4.	was	eat	
5.	answer	Jazzy	
6.	said	face	

ASSESSING MODERN LANGUAGES IN SWEDISH UPPER SECONDARY SCHOOLS

L'article présente les modalités d'évaluation du nouveau plan d'étude pour les langues étrangères en Suède. Partant d'une présentation générale des nouveaux plans d'étude pour les langues modernes, le propos se concentre ensuite sur le cas particulier de l'italien L2 pour débutants au secondaire II: le lecteur est invité à suivre les étudiants du lycée de Kärntorp, à la périphérie de Stockholm, dans leur première approche de l'italien L2 jusqu'à leur premier examen, neuf mois plus tard, lorsque l'évaluation des connaissances acquises a été faite selon le nouveau plan d'études.

● Per Snoder | Stockholm

Per Snoder has taught English, French and Italian for ten



years in secondary schools in Sweden and in the UK. He has also worked as a lecturer in Language Education in pre- and in-service courses at Stockholm University for four years. He is now a PhD student with a keen interest in formulaic language and language pedagogy.

The purpose of this report is to provide an example of how Italian for beginners at a Swedish upper secondary school (students aged 16-19) was assessed in light of the new Swedish foreign language syllabus. After an introduction of the new syllabi for modern languages, the reader is invited to follow my learners of Italian at Kärntorps gymnasium outside Stockholm from their first encounters with the new language at the start of the course in August to the final examination period nine months later, when the knowledge requirements from the syllabus were assessed and summarized for the course grade.

The new national curriculum

In 2011 a new National Curriculum was introduced in Sweden with new syllabi for the school subjects and yet another new grading system, the third in just 18 years. The new concise syllabi consist of three parts:

- > the aim of the subject;
- > the core content; and
- > the knowledge requirements (a new term which in most English-speaking contexts are referred to as 'standards').

Similar to the previous curriculum, a communicative approach to language teaching is prescribed in the aim for modern languages, giving learners the opportunity to "develop all-round communicative skills". Furthermore, in line with communicative methodologies, the suggestion is that foreign language should be taught in the foreign language to the greatest extent possible, which in many cases is asking much of teachers. The following three components of communicative competence are emphasized in the aim of the subject:

1. *Linguistic competence* in reception and production (i.e. understanding target language (TL) speech/text and speaking/writing it) which is relatively unproblematic from an assessment point of view as it is specified in the knowledge requirements.
2. *Strategic competence* is given a more prominent role in the new syllabi. However, assessing it is fraught with problems as strategies are difficult to define and may constitute non-observable data: avoidance is a possible and presumably common strategy, in-

PIÙ ARTICOLI SU QUESTO TEMA:

WWW.BABYLONIA.CH >

ARCHIVIO TEMATICO > [SCHEDE 6 E 7](#)

deed difficult to assess! It is explicitly mentioned in the knowledge requirements but not sufficiently specified.

3. *Sociolinguistic competence* involving the ability to adapt language to different purposes, recipients and situations. This competence is only implicitly subsumed in the knowledge requirements for linguistic competence.

In addition, there is a TL cultural part where learners are to comment on phenomena from TL contexts and make comparisons with their own experiences and knowledge. This component is included in the knowledge requirements but insufficiently specified.

Another new feature is found in the core content of the syllabus, detailing what factual content the teaching in the course should cover, e.g. that signs and advertisements are relevant objects of study for promoting receptive skills and that politeness and fixed expressions should be integrated in production-oriented lesson activities.

Ongoing Assessment

There is something special about teaching a foreign language to complete beginners – you really have the opportunity to isolate and gauge your own effectiveness as a language teacher given the fact that your learners know nothing at the outset, at least in theory. My being an Italian teacher at upper secondary school in Sweden alters this picture in an important way as I know that Italian is their third foreign language learned in school, having started formal instruction of English around the age of nine (this being compulsory), and then choosing a second foreign language (Spanish, French and German) at the age of 12. My learners thus came to the classroom with different experiences of and attitudes towards formal language learning, as well as different proficiency levels in the languages they had learned.

Right from the first lesson I taught, I made an effort to convey the idea that their knowledge of other languages was resourceful for them when learning Italian, e.g. by introducing the concept of cognatedness, i.e. that the same or similar

word form in two languages often have related meanings – as in the Italian word *dialogo* and *dialog* in Swedish – and I encouraged educated guessing as a powerful metalinguistic strategy. I also focused on fostering native-like pronunciation by highlighting consonant quality in Italian, with e.g. the ‘thick’ sound of the initial sound in the Italian word *cinema*. The third feature I prioritized from day one was routines with massive comprehensible exposure to and pushed output of Italian. I emphasized how much them joining in these routines, and accepting the language showers they were exposed to, improved the chances of their scoring high on the final examinations. The three central components of the routines were:

1. *roll call*: using attendance checking as a learning opportunity. The first time they responded *Sì (yes)* themselves or classmates said *No, è assente (No, he/she is absent)*, using words that were written on the whiteboard. This practice grew in complexity so that eventually learners provided elaborate answers such as *Io preferisco l'estate (I prefer summer)*. The lesson plan was also written on the whiteboard and I went through it, in Italian, every lesson. Though time-consuming this roll call routine proved immensely educational, particularly its repetitive nature.

2. *text pre-activities*: the introduction of a new text included setting learners up for success by various techniques to give them a pre-understanding of it, e.g. my choosing five core sentences from the text which they analysed in pairs and then guessed what it would be about it (in their L1).

3. *weekly written tests*: learners were informed of the final examination and how it was to summatively assess their communicative competence in Italian. They were also told that weekly (marked but not recorded) tests would be administered to make sure that they were “on track” learning-wise, a consistent poor result of which would generate a written notice.

Grading

I now turn to the central topic of this report: how I assessed my learners of Italian

Assessing strategic competence is fraught with problems as strategies are difficult to define and may constitute non-observable data: avoidance is a possible and presumably common strategy, indeed difficult to assess!

and the grading system. The knowledge requirements reflect the competences in the aim of the subject and are specified for three grade levels (Figure 1 and Figure 2): E (the lowest passing grade), C, and A (the highest passing grade). For the unspecified grade D, all knowledge requirements for grade E should be met plus more than half of the ones for grade C. The same procedure applies to the grade B, involving the requirements for C and A.

Figure 1: The Swedish grading system (as of 2011)

For a grade of:	The student has:
A	achieved all the defined knowledge requirements
B	achieved all the knowledge requirements for Grade C and the bulk of the requirements for Grade A.
C	achieved all the knowledge requirements for Grade C
D	achieved all the knowledge requirements for Grade E and the bulk of the requirements for Grade C.
E	achieved all the knowledge requirements for Grade E
F	failed to achieve any or all of the knowledge requirements for Grade E

Continuous assessment is not part of the Swedish grading system as what the student knows at the end of the course is what matters. However, should teachers have reason to believe that a student risks getting a failing grade (F), he/she is entitled to receive a written notice during the parent/teacher meeting, specifying on what grounds the warning is issued, what needs to be done and the role each party is expected to play for a passing grade to be reached.

Final examinations

The final examinations of my Italian course took place in May and required a great deal of planning, practical arrangements and innovation. The examinations were divided into three parts (see the complete test on www.babylonia.ch > No 2/2016 > Article Snoder):

- > (1) a written test assessing listening, reading, and writing;
- > (2) an oral test in which learners were put in pairs and interacted for around 15 minutes; and
- > (3) the cultural part for which learners had prepared two topics related to Italian civilization, a test they took directly after the oral test.

In April learners were given detailed information as to the contents of these assessments (vocabulary lists, grammar, etc.) and also, importantly, the opportunity to take mock exams, i.e. taking the tests in the exact same form, but with other contents, to practice the procedure.

The written test

The written test consisted of three parts that were assessed qualitatively with explicit reference to the knowledge requirements for levels E, C, and A. Let us look at examples of all three parts for the E level. The listening instructions tell learners that they are to hear a dialogue in Italian, once, and then answer questions in Swedish (after having read them through). **The choice of language was deemed important to me as they were to be assessed on how well they understood Italian and not how well they expressed themselves in Italian, referred to as the validity of the assessment, or that the measurement measures what it is intended to measure (the choice of language for the listening comprehension is not specified in the syllabus).** As can be seen in the test itself, the knowledge requirements are found in the box at the bottom of the page (Figure 2) and specify qualities for each level of performance, e.g. that at the E level the “Learner shows his/her understanding by [...] acting upon messages with an **acceptable result**”. Here quantitative measures were judged necessary as a distinction between “acceptable” (“godtagbart”), “satisfactory” (“tillfredsställande”) and “good” (“gott”) results needed to be quantified. Acceptable was thus, admittedly somewhat arbitrarily, defined as a score of three out of seven points, satisfactory four to five points and good six to seven points and how to interpret these adjectives is decided by individual or teams of teachers.

For the reading and writing parts, the E, C, and A levels were put in separate sections so that learners were allowed to 'target' their grade level, e.g. only taking the E level reading part if they so desired. To be awarded a grade C learners need to pass both the E and C level, and for an A learners need to pass all three levels. In the reading part, the knowledge requirement for each level was stated first and then a short text followed with questions, again, to be answered in Swedish. The threshold for pass for the E level was set at 60% correct answers, meaning that students either passed or failed that task.

The knowledge requirements are specified for reading and e.g. for grade E the student "understands common words and simple phrases in short, simple texts about everyday activities and familiar topics", followed by the same subcriteria as for the reading part. For the writing part learners were asked to write a short coherent text in Italian (introducing a fictitious female friend) and were provided with necessary information in Swedish. The knowledge requirements for writing specify that the learner is "capable of expressing him-/herself comprehensibly with single words and phrases". The threshold for pass was set at 50% comprehensible answers, which in this case did not require complete grammatical accuracy, but instead comprehensibility was the key. So, if a learner wrote the female personal pronoun *Lei* (for *She*) and included all the necessary content words, but failed to make subject-adjective agreements, it still counted as a correct answer from a comprehensibility point of view (for those who don't speak Italian: without *Lei* and agreements it could have referred to a man, which is not acceptable in this situation).

The oral exam

In the oral test learners performed in self-selected pairs and were scheduled a 20-minute appointment. They were provided with test materials which included two grids (different for person A and B) with contexts and prompts in the L1 on content matter from the course and instructed how to perform the intended procedure. Then they interacted for 10-15 minutes, sitting at desks facing each other and with myself as a non-participant assessor with forms to fill out in which I put comments and ticks in relevant boxes according to the knowledge requirements (I also audio-recorded the dialogues). Strategic competence is, as mentioned above, emphasized in the syllabi but I did not register a single (observable) instance of strategy use during all these interactions, apart from one instance of looking up words on a mobile phone. To remedy this situation I asked them to individually write down examples of possible strategies. The specifications in the knowledge requirements regarding strategies are limited (no examples are given) and the few details that are provided require interpretation. For grade E it says "some strategies", for C "some different strategies" and for A "several different strategies". I adopted a pragmatic approach to it and two similar strategies sufficed for an E, two distinctively different strategies were enough for a C and more than two distinctively different strategies met the A grade criterion.

The cultural part

For the third part the knowledge requirements are unspecific and confusing as they do not state which language to use and the exact same wording is used for both the E and C grades while for the A level only the word "very" is left out: "Students comment in very simple forms on some phenomena in different contexts and areas where the language is used, and make simple comparisons with their own experiences and knowledge" (for grades E and C; emphasis in original). My learners had prepared

Kunskapskrav för hörförståelse från Gy11:

E: Eleven visar sin förståelse genom att [...] med **godtagbart** resultat agera utifrån budskap [...] i innehållet – eleven fick tre rätt av sju på ovanstående hörförståelseuppgift _____

C: Eleven visar sin förståelse genom att [...] med **tillfredsställande** resultat agera utifrån budskap [...] i innehållet – eleven fick fyra-fem rätt av sju på ovanstående hörförståelseuppgift _____

A: Eleven visar sin förståelse genom att [...] med **gott** resultat agera utifrån budskap [...] i innehållet – eleven fick sex-sju rätt av sju på ovanstående hörförståelseuppgift _____

More in-service training is needed, run by assessment experts who can explain the rationales behind the syllabus and provide concrete examples of learning activities and assessment tasks.

to discuss two aspects of Italian culture and civilization that I had suggested – e.g. football and food – and relate them to their own experiences and knowledge, in writing. The pragmatic yet unsatisfactory approach to go about assessing these short coherent texts written in Swedish was to simply categorize them as pass or non-pass performances (i.e. similar to a course requirement).

It took me four weeks in May to gather enough data to grade all my learners. Let us now look at examples of two learners, referred to as *Daniella* and *Bianca* (fictitious names). *Daniella* received an E at the oral test, she passed the culture test and received three C's on the listening-reading-writing tests. Her final course grade was a D. Her classmate *Bianca* received an A on the oral test, a pass on the culture test, A on the listening and writing tests and a C for the writing test. Her final course grade was a B.

This was the first batch of students I graded with the new curriculum. Fortunately, I was not fresh out of university, but had taught modern languages for ten years, and also worked three years as a teacher educator at the university. I was therefore fully aware of the importance of keeping everything "by the book", i.e. sticking very close to the curricular documents (to the extent that it was feasible), which is the hallmark of a professional and competent teacher. I tried to cooperate with my colleagues (teachers of other foreign languages such as French and Spanish), but being the only Italian teacher at the school I was largely left to my own devices.

If I were to teach the same course again I would definitely keep the routines, but

I think I would show the students, at a rather early stage, an example of the final examinations to boost motivation and explain the grading system. Of the three competences, I found that linguistic competence was the easiest one to score, whereas scoring strategic competence was more challenging. For lack of a better test instrument, I resorted to assigning students to simply enumerate strategies, but I do not know how else to go about it. Sociolinguistic competence was even more problematic as it is not explicitly specified in the knowledge requirements. However, it was embedded, though implicitly, in the writing part where students needed to show awareness of politeness markers and also use other textual devices to adapt it to the particular genre. The cultural part was clearly the most problematic aspect, for reasons already specified.

Conclusion

As a result of these flaws and inconsistencies I do not endorse the Swedish National Curriculum 2011 and its new syllabus for modern languages and the grading system. The knowledge requirements need to be revised so that the aims may be validly assessed. Furthermore, more in-service training is needed, run by assessment experts who can explain the rationales behind the syllabus and provide concrete examples of learning activities and assessment tasks. Moreover, school management should also impose collegial cooperation for same subject teachers in syllabus design and not naively believe that it will occur on a voluntary basis.

INTERKULTURELLE KOMPETENZEN TESTEN — GRADUIERUNGSKRITERIEN UND BEURTEILUNGS- RASTER FÜR DEN FREMDSPRACHENUNTERRICHT AUF DER SEKUNDARSTUFE I

Come considerare e valutare a scuola le competenze interculturali? Il presente contributo espone in quale misura tali competenze sono ancorate nel *Lehrplan 21*, offrendo pertanto buone premesse per introdurre elementi interculturali nelle lezioni ‚convenzionali‘ e valutarli. A questo scopo l'articolo propone una griglia che permette di testare tali competenze distinguendo fra ‚astrazione‘, ‚complessità‘ e ‚precisione‘.

La photo représente des Français parce que c'est importante pour tous les Français qu'ils mangent ensemble avec la famille. (...) En outre, il y a un verre de vin et Baguette au la table. (...) Je voudrais savoir si tous les familles se comportent même antéposé.

Diesen Text hat eine Achtklässlerin im dritten Französischlernjahr verfasst. Wie sollte die Lehrperson damit umgehen? Deren Fokus richtet sich nicht selten auf die sprachliche Korrektheit – im vorliegenden Fall also auf die fehlerhafte Verwendung der weiblichen Form des Adjektivs ‚important‘, die Grossschreibung des Nomens ‚Baguette‘, die Präposition ‚au‘ sowie den nicht sinnhaften Gebrauch des Adjektivs ‚antéposé‘. Interessant wäre aber zu erfahren, wie Französischlehrkräfte mit dem Inhalt solcher Zitate umgehen. Denn die Beurteilung schriftlicher Produktionen im Hinblick auf ihre soziokulturelle Differenziertheit oder auf das darin geäußerte Interesse an fremden Kulturen kann dazu beitragen, nicht nur die sprachlichen, sondern auch die in sämtlichen Fremdsprachen-

curricula verankerten interkulturellen Kenntnisse und Fertigkeiten zu beurteilen. Die Bedeutsamkeit interkultureller Kompetenzen im Fremdsprachenunterricht – verstanden als Oberbegriff für all diejenigen Fähig- und Fertigkeiten eines Menschen, die im Kontext interkultureller Begegnungssituationen neben den kommunikativen Kompetenzen zum Tragen kommen – ergibt sich insbesondere aus der Multikulturalität unserer Gesellschaft. Dementsprechend gelten interkulturelle Fähig- und Fertigkeiten auf nationaler wie auf kantonaler Ebene als zentrales Ziel schulischen Fremdsprachenlernens, wie beispielsweise im *Lehrplan 21* oder im Lehrplan *Passepartout – Fremdsprachen an der Volksschule*.

Dass der Umgang mit interkulturellen Kompetenzen viele Fremdsprachenlehrpersonen vor grosse Herausforderungen stellt, ist sicherlich unbestritten. Eine mögliche Ursache hierfür ist die nach wie vor herausragende Rolle kommunikativer Kompetenzen im Fremdsprachenunterricht, wie auch Saudan (2015: 12) hervorhebt: „En classe de langue étrangère, les objectifs actionnels, c'est-à-dire la

Jan-Oliver Eberhardt | ●
Basel

Jan-Oliver Eberhardt ist seit 2015 Dozent für Fachdidaktik Französisch an der Pädagogischen Hochschule FHNW. Zuvor war er mehrere Jahre als Sekundarschullehrer mit den Fächern Französisch und Chemie sowie als wissenschaftlicher Mitarbeiter im Bereich der Didaktik der romanischen Sprachen tätig. Seine Promotionsschrift widmet sich der Modellierung interkultureller Kompetenzen für den Französischunterricht der Sekundarstufe I.



1 Dabei bilden kulturspezifisches Wissen über eigene und fremde Kulturen sowie kulturallgemeines Wissen über persönliche und gesellschaftliche Interaktionsprozesse die Wissenskomponente interkultureller Kompetenz (*savoirs*). Eine weitere Komponente beinhaltet Einstellungen wie Neugier, Offenheit sowie die Bereitschaft, eigene Überzeugungen zu relativieren (*savoir être*). Des Weiteren spricht Byram interkulturell kompetenten Menschen die Fähigkeit zu, Dokumente oder Ereignisse anderer Kulturen zu interpretieren, zu erklären und miteinander in Beziehung zu setzen (*savoir comprendre*). Ebenso versteht sich als Bestandteil interkultureller Kompetenz die Fähigkeit, neues Wissen über andere Kulturen zu erwerben und interkulturelle Begegnungssituationen zu bewältigen (*savoir apprendre/faire*). Die fünfte Teilkompetenz, *savoir s'engager*, impliziert den kritischen Umgang mit kulturellen Produkten und Praktiken.

2 Unter Berufung auf kompetenztheoretische Forschungsdiskurse wird hier unter dem Graduierungsparameter 'Präzision' die Fülle an Details beziehungsweise die Genauigkeit gelieferter Informationen in einer Äusserung verstanden. Unter der 'Komplexität' eines Teilaspekts interkultureller Kompetenz versteht man den Grad der Vernetzung und Dimensionalität, und eng verbunden damit auch die Kreativität und Weiträumigkeit. Der 'Abstraktions'-Grad einer interkulturellen Reflexion spiegelt schliesslich das Ausmass wider, in dem von konkret wahrnehmbaren kulturellen Phänomenen abstrahiert wird, diese unter einem allgemeineren Blickwinkel betrachtet und somit auch tiefer liegende, abstrakte kulturelle Ebenen beschritten werden (Eberhardt, 2013)

construction de capacités communicatives, sont clairement dominants et le travail sur le culturel et l'interculturel reste marginal". Hinzu kommt, dass die interkulturelle Kompetenz komplex und vielschichtig ist. Verglichen mit kommunikativen Fertigkeiten wie dem Lese- und Hörverstehen gilt sie zudem als schwer messbar (Eberhardt, 2013). Dies hat zur Folge, dass interkulturelle Lernziele in vielen fremdsprachlichen Curricula nicht mit ausreichender Präzision beschrieben und infolgedessen von Seiten der Schulbuchverlage bis dato weder Lehrmittel noch Evaluationsinstrumente bereitgestellt werden, welche den Fremdsprachenlehrern eine gezielte und differenzierte Förderung und Überprüfung interkultureller Kompetenzen ermöglichen. Beispielsweise tauchen in den auf Volksschulebene in der Nordwestschweiz durchgeführten Leistungstests sowie in der daran gekoppelten Aufgabensammlung zum kompetenzorientierten Üben, Prüfen und Testen ausschliesslich kommunikative Kompetenzen auf. Ebenso wenig spielt interkulturelles Lernen in den *lingualevel-Instrumenten zur Evaluation von Fremdsprachenkompetenzen* für den Englisch- und Französischunterricht vom 5. bis zum 9. Schuljahr in der deutschsprachigen Schweiz eine Rolle.

Graduierungskriterien und Beurteilungsraster zur Evaluation interkultureller Kompetenzen

Mitverantwortlich für das Fehlen konkreter Lern- und Testmaterialien zur Ausbildung interkultureller Kompetenzen ist ohne Frage der bisherige Mangel an differenzierten Kompetenzmodellierungen für den standardorientierten Fremdsprachenunterricht (Eberhardt, 2013). Denn innerhalb der Fachdidaktik besteht darüber Einigkeit, dass erst auf Grundlage präziser und bewährter Standardformulierungen auch kompetenzorientierte Förder- und Prüfungsaufgaben mit passenden Beurteilungsrastern konstruiert werden können. Im Idealfall sollten fachspezifische Kompetenzmodelle neben der horizontalen Untergliederung einer fachbezogenen Kompetenz auch eine Graduierung in unterschiedliche Kompetenzniveaus vornehmen. Während im Gros nationaler Curricula des institutionellen Fremdsprachenlernens keine differenzierte horizontale und vertikale Aufschlüsselung interkultureller Kompetenzen vorliegt, werden diese im

Lehrplan 21 - unter der Bezeichnung „Kulturen im Fokus“ (Deutschschweizer Erziehungsdirektoren-Konferenz, 2014: 72) – auf horizontaler Ebene in interkulturelle Kenntnisse, Haltungen und Handlungen untergliedert, in Form von *can-do*-Deskriptoren präzisiert und darüber hinaus spezifische Kompetenzniveaus festgelegt. Zwar gewähren die Kompetenzdeskriptoren im *Lehrplan 21* Fremdsprachenlehrkräften einen guten Überblick über die zu erreichenden Mindestanforderungen. Allerdings sollten Beschreibungen von Kompetenzstufen für die Entwicklung von Testaufgaben und Beurteilungsrastern stets einheitliche, und nicht vermischte, Teilaspekte einer Kompetenz widerspiegeln (Eberhardt, 2013; Harsch, 2006), zudem wäre angesichts der Komplexität interkultureller Kompetenz ein noch höherer Differenzierungsgrad der horizontalen Dimensionierung wünschenswert. Ein möglicher erster Schritt hin zur Generierung eines differenzierteren und zur Evaluation interkultureller Fertigkeiten im schulischen Fremdsprachenunterricht einsetzbaren Kompetenzniveaumodells wird in meiner Forschungsstudie „Interkulturelle Kompetenzen im Fremdsprachenunterricht“ (Eberhardt, 2013) unternommen. Im Zentrum steht dabei die empirische Ausdifferenzierung des europaweit anerkannten Modells interkultureller Kompetenz des Fremdsprachendidaktikers Byram (1997), welcher die horizontale Dimension interkultureller Kompetenz in 5 Teilkomponenten untergliedert, in die sogenannten *savoirs*¹. In der qualitativ ausgerichteten Studie wird die interkulturelle Kompetenz von Französischlernenden am Ende der Sekundarstufe I aufgrund deren Überlegungen zu kulturell aufgeladenen Fotos und zu interkulturellen Kommunikationssituationen erhoben. Dabei geht es beispielsweise darum, den Gesprächspartner sowie Gemeinsamkeiten und Unterschiede auf der Ebene des Denkens, Fühlens und Handelns zu beschreiben, Vorwissen in Bezug auf interkulturelle Kommunikationssituationen zu aktivieren oder Bereiche des Nicht-Verstehens zu nennen. Als Erhebungsinstrumente kamen dabei offene Fragebögen (N=106) und fokussierte Interviews (N=15) zum Einsatz. Ausgewertet wurden die Schülerantworten mittels Qualitativer Inhaltsanalyse (Mayring, 2003), im Sinne eines Kompetenzverstehens durch Interpretation von Sprache. Die Analyse der interkulturellen Reflexionen brachte so-

wohl eine sehr differenzierte horizontale Untergliederung der *savoirs* hervor als auch eine vertikale Aufschlüsselung der ermittelten Teilaspekte interkultureller Kompetenz in Form einer einheitlichen 3-Stufung. In Anlehnung an bildungswissenschaftliche Diskurse der Kompetenzforschung werden diese Niveaus als basales, intermediäres und elaboriertes Niveau bezeichnet und anhand der – auch in Kompetenzmodellen anderer Unterrichtsfächer zum Tragen kommenden – Graduierungskriterien Abstraktion, Komplexität und Präzision² erklärt. Am Beispiel des Teilaspekts Wissen über Gegenstände, Institutionen, Ereignisse

und Persönlichkeiten anderer Kommunikationsgemeinschaften der Teilkompetenz *savoirs* soll im Folgenden zunächst die kriterienbasierte 3-Stufung der Studie empirisch veranschaulicht werden (Tab. 1). Die empirischen Repräsentationen interkultureller Kompetenz spiegeln dabei exemplarische interkulturelle Reflexionen der untersuchten Sekundarschülerinnen und -schüler wider. Ein weiterer Schritt bei der Entwicklung eines empirisch abgesicherten domänenspezifischen Kompetenzmodells ist stets die Transformation empirischer Kompetenzrepräsentationen in deskriptorenbasierte can-do-Skalen. Dabei un-

Tab. 1: Empirische Repräsentationen interkultureller Kompetenz

Teilkompetenz: <i>savoirs</i>			
Teilaspekt: Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften			
Graduierungsparameter	Basales Niveau	Intermediäres Niveau	Elaboriertes Niveau
Abstraktion	Bild 7 geht es sicherlich um ein Fussballspiel, wo Frankreich gerade spielt. Das sehe ich einmal durch die Flaggen von Frankreich. Ausserdem kann ich noch erkennen, dass die Männer jeweils ein Baguette hoch halten. Das gibt mir auch das Gefühl, dass diese Personen aus Frankreich kommen. Baguette ist ein typisches Brot für Frankreich, und als ich in Frankreich war, habe ich jeden Morgen Baguette bekommen. (Ges3_Ahmed)	Auf jeden Fall hat das Bild 7 etwas mit Frankreich zu tun, Männer, angezogen in den Nationalfarben Frankreichs, jubeln begeistert. Sie heben die Hände hoch, in denen sie Baguettes halten, die wieder-um immer in Verbindung mit Frankreich gebracht werden und somit charakteristisch für Frankreich sind. Auf diesem Bild bemerkt man sehr doll das Nationalgefühl, welches die Franzosen miteinander verbindet, und den grossen Stolz auf ihr Land, den sie in sich tragen. (Gym2_Vera)	Oft spricht man ja „die Franzosen machen das so und so, bei den Franzosen ist das ja so, dass die Frauen sich nicht rasieren“ solche Sachen einfach. Das sind Vorurteile, die es bei jeder Nation gibt und die man, wenn man nicht in dem Land selbst gelebt hat für eine Zeit, dann auch einfach glaubt, weil irgendwie möchte man sich ja schon ein Bild machen von einer Gesellschaft. Und da kommen solche Vorurteile natürlich ganz gelegen, sage ich jetzt mal, aber es ist halt relativ schwierig zu sagen, ob die jetzt richtig oder falsch sind, wenn man keine persönlichen Eindrücke bekommen hat. Ich denke, es gibt mit Sicherheit Frauen, die sich nicht rasieren oder mit Sicherheit gibt es Menschen, die mal mit einem Baguette rumlaufen, weil sie einfach ein Baguette essen möchten. Aber es ist ja bestimmt nicht so, dass da jeder mit einem Baguette rumläuft und das jeden Tag gegessen wird. (Gym3_Franziska)
Komplexität	Bild 8 zeigt Hitler und andere nationalsozialistische Anhänger vor dem Pariser Eiffelturm, also dem Einmarsch der NS in Frankreich. (Gym3_Nastassja)	Foto 8 hat sicherlich was mit Frankreich zu tun. Während des 2. Weltkriegs besetzten die Deutschen von 1940-1944 das Land, und die Franzosen mussten in den Süden, in der „Freien Zone“ leben. Die Regierung war nicht länger in Paris, sondern in Vichy. Staatschef war Pétain, der mit Hitler zwangsweise zusammenarbeiten musste. (Gym2_Nicole)	Bild 8 möchte ich nicht miterlebt haben, da die dort abgebildeten NS-Funktionäre und der Führer das Sinnbild einer Schreckensherrschaft sind. Trotzdem würde ich Paris im Jahre 1940 gerne gesehen haben wegen den Veränderungen zu dem Paris, das ich kenne. In der Deutschen Schule Paris hatten wir mal ein Gespräch mit einem Résistant aus der Zeit zwischen 1940 - 1944. Mein Gedanke war, wie er sich jetzt wohl gefühlt hat, von Deutschen, den ehemaligen Aggressoren, befragt zu werden. (Gym2_Julian)
Präzision	Auf dem letzten Bild sehe ich den Eiffelturm. Das hat ganz klar mit Frankreich zu tun. Wenn ich etwas über Frankreich erzählen würde, würde ich sofort an den Eiffelturm denken. (Ges3_Ahmed)	Das andere Bild habe ich genommen, weil dort der Eiffelturm steht. Der ist ja das Wahrzeichen dort und das zeigt ja schon alles. (Ges2_ Stefan)	Also Bild 8 hat in jedem Fall etwas mit Frankreich zu tun. Man sieht auf dem Bild den Eiffelturm, das Wahrzeichen Frankreichs. Viele Touristen besteigen ihn, und abends, wenn es dunkel wird, ist er schön beleuchtet und ein totaler Blickfang. (Ges3_Sarah1)



Quelle: http://www.focus.de/fotos/mit-baguette-power-zum-sieg-hof-fentlich-nicht_id_3967361.html



Quelle: Bernlochner, L., Boesenberg, L. & Braun, M. (2006). Histoire/Geschichte – L'Europe et le monde du Congrès de Vienne à 1945. Paris: Éditions Nathan, p. 314.

terscheidet man stets benutzerorientierte und beurteilungsorientierte Skalen (Eberhardt, 2013). Benutzerorientierte Skalen haben eine Berichtsfunktion und dienen der allgemeinen Feststellung des Leistungsstands im Rahmen von Fremd- oder Selbstbeurteilung und „beschreiben nicht eine erwartete oder tatsächlich beobachtete Performanz in einem bestimmten Kontext, sondern geben in generalisierter Form [kontextfreie] Beschreibungen des Könnens wieder“ (Harsch, 2006: 139). Dagegen werden beurteilungsorientierte Skalen bei der Evaluation von Lernerleistungen eingesetzt und beschreiben häufig „konkrete Merkmale der zu beobachtenden oder erwarteten Performanz auf einen bestimmten Stimulus hin“ (Harsch, 2006: 140).

In der folgenden Tabelle (Tab. 2) ist ein Beispiel für eine benutzerorientierte Skala aufgeführt, welche zum Zwecke allgemeiner Aussagen über Lernerleistungen generalisierte, kontextfreie Beschreibungen von Lernerkompetenzen. Zur Entwicklung benutzerorientierter Deskriptoren werden stets exemplarische Schüleräusserungen – die sogenannten Ankerbeispiele – in allgemeine, kontextfreie, für ein Niveau prototypische Kann-Beschreibungen eines Teilaspekts interkultureller Kompetenz umformuliert, getrennt nach den Graduierungskriterien Präzision, Komplexität und Abstraktion. Um den Transformationsprozess von empirischen Repräsentationen in allgemeine Kann-Beschreibungen eines Teilaspekts interkultureller Kompetenz

Tab. 2: Benutzerorientierte Skala interkultureller Kompetenz

Teilkompetenz: <i>savoir comprendre</i>			
Teilaspekt: Fähigkeit, Bereiche des Missverstehens und Störungen in interkulturellen Interaktions- und Kommunikationssituationen zu identifizieren und zu erklären			
	Basales Niveau	Intermediäres Niveau	Elaboriertes Niveau
Abstraktion	Die Schüler dokumentieren ein konkret-oberflächliches Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften, welches in der Regel von stereotypem Charakter ist.	Die Schüler dokumentieren ein vom Bezugsgegenstand bedingt abstrahierendes Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften, welches zum Teil stereotyp und in Ansätzen differenziert ist.	Die Schüler dokumentieren ein vom Bezugsgegenstand abstrahierendes Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften, welches in der Regel differenziert ist.
Komplexität	Die Schüler dokumentieren ein isoliertes und ein-dimensionales Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften, welches in der Regel von stereotypem Charakter ist.	Die Schüler dokumentieren ein bedingt vernetztes und mehrdimensionales Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften, welches zum Teil stereotyp und in Ansätzen differenziert ist.	Die Schüler dokumentieren ein vernetztes und mehrdimensionales Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften, welches in der Regel differenziert ist.
Präzision	Die Schüler dokumentieren ein vages und detailarmes Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften, welches in der Regel von stereotypem Charakter ist.	Die Schüler dokumentieren ein bedingt präzises und detailreiches Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften, welches zum Teil stereotyp und in Ansätzen differenziert ist.	Die Schüler dokumentieren ein präzises und detailreiches Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften, welches in der Regel differenziert ist.

nachvollziehen zu können, bezieht sich die benutzerorientierte Skala erneut auf den Teilaspekt *Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften*.

Indem sich benutzerorientierte Skalen ausdrücklich auf die generelle, aufgabenunabhängige Beschreibung der bei Fremdsprachenlernenden vorhandenen interkulturellen Kompetenzen konzentrieren, können sie auf andere Testaufgaben und Lerngruppen angewendet werden. Demzufolge ist mit benutzerorientierten Skalen zum Beispiel die Möglichkeit verbunden, Schülerinnen und Schülern ein formatives Feedback zu geben, welches sich nicht auf eine spezifische Lernaufgabe oder Prüfung, sondern auf mehrere Evaluationsmomente bezieht.

Im Gegensatz zu benutzerorientierten Skalen enthalten die Kompetenzformulierungen beurteilungsorientierter Skalen keine generalisierten Kann-Beschreibungen, sondern geben die einer konkreten Testaufgabe entstammenden prototypischen Performanzen wieder. Demzufolge finden dort auch die in einer Aufgabe angesprochenen spezifischen kulturellen Inhalte Erwähnung. Beurteilungsorientierte Skalen können deshalb beispielsweise als Kodierleitfaden zur Auswertung schriftlicher Sprachpro-

duktionen im Kontext formativer oder summativer Beurteilungen im Fremdsprachenunterricht eingesetzt werden. Dadurch ist es also möglich, interkulturelle Testaufgaben kriterienorientiert zu beurteilen, indem interkulturelle Reflexionen von Fremdsprachenlernenden mit spezifischen Teilaspekten interkultureller Kompetenz versehen und hinsichtlich ihres Präzisions-, Komplexitäts- und Abstraktionsgrades auf ein basales, intermediäres oder elaboriertes Niveau eingestuft werden. Im Sinne einer praktikablen und zeitökonomischen Leistungsbeurteilung von Fremdsprachenlernenden bietet es sich bei beurteilungsorientierten Skalen an, die Graduierungskriterien Abstraktion, Komplexität und Präzision zusammenzufassen (Eberhardt, 2013). Zum Zwecke der Vergleichbarkeit der Skalentypen ist im Folgenden (Tab. 3) eine solche beurteilungsorientierte Skala zum Teilaspekt *Wissen über Gegenstände, Institutionen, Ereignisse und Persönlichkeiten anderer Kommunikationsgemeinschaften* aufgeführt. Die Kann-Beschreibungen basieren dabei auf prototypischen Schülerreflexionen in Bezug auf eine im französischen Spielfilm „2 Tage Paris“ (Julie Delpy, 2007) zu beobachtende interkulturelle Kommunikationssituation in einem Fast-Food-Restaurant.

Tab. 3: Beurteilerorientierte Skala interkultureller Kompetenz

Teilkompetenz: <i>savoir être</i>		
Teilaspekt: Bereitschaft zur Perspektivendarstellung und Perspektivenübernahme im Rahmen kulturell aufgeladener Situationen		
Basales Niveau	Intermediäres Niveau	Elaboriertes Niveau
Die Schüler nennen im Rahmen der Kommentierung einer interkulturellen Kommunikationssituation einzelne konkret wahrnehmbare Gegenstände, Institutionen, Ereignisse und Persönlichkeiten und ordnen diese als typisch für sämtliche Teilnehmer einer spezifischen Kommunikationsgemeinschaft ein. Dabei werden die genannten kulturellen Phänomene weder präzisiert, noch werden Abstraktionen oder Vernetzungen mit anderen kulturellen Phänomenen vorgenommen.	Die Schüler nennen im Rahmen der Kommentierung einer interkulturellen Kommunikationssituation einzelne abstrakte oder konkret wahrnehmbare Gegenstände, Institutionen, Ereignisse und Persönlichkeiten, beschreiben diese in Ansätzen näher und assoziieren sie mit Teilnehmern einer spezifischen Kommunikationsgemeinschaft. Dabei werden einfache Abstraktionen von den kulturellen Bezugsgegenständen sowie einfache Vernetzungen der genannten Gegenstände, Institutionen, Ereignisse und Persönlichkeiten oder weiterer kultureller Phänomene vorgenommen.	Die Schüler beschreiben im Rahmen der Kommentierung einer interkulturellen Kommunikationssituation einzelne abstrakte oder konkret wahrnehmbare Gegenstände, Institutionen, Ereignisse und Persönlichkeiten sehr präzise, und assoziieren diese mit Teilnehmern einer oder mehrerer spezifischer Kommunikationsgemeinschaften. Dabei werden weitreichende Abstraktionen von den kulturellen Bezugsgegenständen sowie vielfältige Vernetzungen der genannten Gegenstände, Institutionen, Ereignisse und Persönlichkeiten oder weiterer kultureller Phänomene vorgenommen.

Bibliographie

Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Clevedon: Multilingual Matters.

Deutschschweizer Erziehungsdirektoren-Konferenz (2014). *Lehrplan 21 Sprachen*. Kanton Basel-Stadt. Luzern: D-EDK.

Eberhardt, J.-O. (2013). Interkulturelle Kompetenzen im Fremdsprachenunterricht. Auf dem Weg zu einem Kompetenzmodell für die Bildungsstandards (*Studien zur Fremdsprachendidaktik und Spracherwerbsforschung Band 1*). Trier: Wissenschaftlicher Verlag Trier.

Harsch, C. (2006). *Der Gemeinsame europäische Referenzrahmen für Sprachen: Leistung und Grenzen. Die Bedeutung des Referenzrahmens im Kontext der Beurteilung von Sprachvermögen am Beispiel des semikreativen Schreibens im DESI-Projekt*. Augsburg: Universität Augsburg.

Mayring, P. (2003). *Qualitative Inhaltsanalyse. Grundlagen und Techniken*. Weinheim: Beltz UTB.

Saudan, V. (2015). Didactiques des langues, didactique de l'(inter-)culturel et du culinaire: le projet *Les Mets et les Mots*. *Babylonia* 1/ 2015, 10-19.

Kompetenzorientierung als Chance zur Aufwertung interkulturellen Lernens

Im vorliegenden Beitrag sollte deutlich geworden sein, dass es erst mit der Konzeption von benutzer- und beurteilungsorientierten Skalen möglich sein wird, kompetenzbasierte Lern- und Testaufgaben für den Bereich interkultureller Kompetenzen einsetzen und kriterienorientiert beurteilen zu können. Mit dem Ziel der Aufwertung interkulturellen Lernens im Fremdsprachenunterricht der Volksschule müsste in Zukunft zuvorderst die Konzeption einer horizontal hochdifferenzierten, jahrgangsübergreifenden und eventuell an die

GeR-Niveaus anschlussfähigen 3- oder 6-Stufung angestrebt werden. Ebenso scheint es notwendig zu sein, für formative Tests und summative Klassenarbeiten Beurteilungsraster zur Evaluation der unterschiedlichen Dimensionen und Teilaspekte interkultureller Kompetenz zu entwickeln, welche für Fremdsprachenlehrerinnen und -lehrer einen dem Unterrichtsalltag angemessenen Korrekturaufwand mit sich bringen. Meines Erachtens stellt die mit dem HarmoS-Konkordat eingeläutete Outputorientierung eine einmalige Chance dar, die Förderung interkultureller Kompetenzen aus dem Schatten fremdsprachlicher kommunikativer Kompetenzen treten zu lassen.

UN TEST A LIVELLO NAZIONALE DI PRODUZIONE SCRITTA: VENGONO RAGGIUNTI GLI OBIETTIVI A FINE DEL SECONDO CICLO DELLA SCUOLA SECONDARIA?

Tema

LA SITUAZIONE IN OLANDA — FORSE EMBLEMATICA PER LA SVIZZERA

Im Schuljahr 2012-13 ist in Holland eine quantitative Untersuchung zu den Schreibkompetenzen auf Maturitätsniveau in den Sprachen Englisch, Deutsch und Französisch durchgeführt worden. Eine Lehrergruppe hat zuerst die mehreren hundert Texte mithilfe eines besonderen Punktesystems bewertet, derweil in der Folge eine Expertengruppe die Punkte den GER-Niveaus zugewiesen hat. Etwas mehr als 50% der Probanden hat das im Lehrplan festgesetzte Niveau erreicht, einzig in Englisch lag die Quote wesentlich höher.

Introduzione

In Olanda esistono due tipi di scuola secondaria che al termine della scuola elementare (4-12 anni) preparano gli allievi all'istruzione universitaria. *havo* è un corso quinquennale che prepara gli allievi a corsi di laurea quadriennali in scuole universitarie a sbocco professionale; *vwo* è un corso di sei anni equiparabile al liceo che prepara gli allievi a tutti gli studi universitari.

In generale, si notano delle differenze di livello tra *havo* e *vwo* riguardo alle competenze di inglese, tedesco e francese. Nei programmi di esame *havo* e *vwo* gli obiettivi linguistici per la produzione scritta in lingua straniera sono descritti nel modo seguente:

Il candidato è in grado di:

- > reagire in modo adeguato nei contatti scritti con persone che utilizzano la lingua straniera come lingua di comunicazione;
- > chiedere e fornire informazioni;
- > presentare in modo adeguato le informazioni ottenute tenendo conto di scopo e destinatari;
- > descrivere persone ed esprimere sentimenti e punti di vista;

> scrivere una relazione in lingua straniera (inglese *havo* e *vwo*, francese e tedesco *vwo*).

Nel 2007 i programmi d'esame di lingua straniera *havo* e *vwo* sono stati correlati al QCER. Per le abilità valutate attraverso l'esame scolastico sono stati formulati degli obiettivi di competenza orientati al QCER (Meijer & Fasoglio, 2007). A questo scopo i piani di studio sono stati correlati ai descrittori del QCER. La tabella 1 illustra gli obiettivi per la produzione scritta *havo/vwo*.

La correlazione al QCER non ha potuto tuttavia basarsi su dati empirici, in quanto fino al 2007 la formazione scolastica in Olanda non era orientata al Quadro Comune. L'*istituto olandese per lo sviluppo dei programmi di studio* (SLO) è stato incaricato dal Ministero dell'Educazione olandese di fornire dati empirici per rispondere alla seguente domanda: **Quale livello QCER raggiungono gli allievi nella produzione scritta in inglese, tedesco e francese alla fine dei cicli di formazione *havo* e *vwo*?** Il presente articolo riassume le fasi più

Anne Beeker-Schoenmakers,
Daniela Fasoglio, Kim de Jong,
Jos Keuning & Alma van Til |

Amsterdam

Anne Beeker-Schoenmakers, anglista, Master of Arts African Literature (Sheffield). Insegnante, ha lavorato presso l'Istituto neerlandese di sviluppo curriculare (SLO).



Daniela Fasoglio, germanista, Master of Arts in linguistica italiana (Amsterdam). Lavora presso SLO.



Kim de Jong, Lic. lingua e cultura francese (Utrecht). Ha lavorato presso SLO.



Jos Keuning, Dr. in scienze dell'educazione. Lavora come metodologo nella sezione di ricerca psicometrica presso Cito.



Alma van Til, MA Français Langue Etrangère (Lille III). Lavora presso Cito.



Tabella 1. Livelli QCER per la produzione scritta

Lingua e tipo di scuola secondaria	Obiettivo QCER
inglese havo	B1
inglese vwo	B2
tedesco havo	A2+
tedesco vwo	B1
francese havo	A2+
francese vwo	B1

Figura 1: Esempio di un singolo compito di scrittura per il francese

<p>Situazione</p> <p>Lavori per quattro settimane nel negozio di un campeggio in Spagna. La mole di lavoro risulta essere maggiore del previsto. Il proprietario del campeggio vuole appendere alla bacheca del negozio un annuncio per assumere temporaneamente una persona in più. Poiché ci sono parecchi turisti francesi, ti ha chiesto di scrivere per lui l'annuncio anche in francese. Qui di fianco trovi le informazioni che l'annuncio deve contenere.</p> <p>Compito</p> <p>Scrivi un annuncio di almeno 50 parole che contenga le informazioni che ti ha fornito il proprietario del campeggio.</p>	<p>Età minima: 16 anni Tariffa oraria: € 8 12 ore / settimana suddivisi in più giorni Possibilità di lavorare più ore assistenza in negozio inizio da subito informazioni / domanda di assunzione : Frey, proprietario del campeggio</p>
--	---

Figura 2. Esempio di frasi di ancoraggio – inglese vwo

Inglese vwo		Com	Cor
Frammento	Frasi di ancoraggio		
Introduci l'argomento	Basso: Deforestation is a main problem in the whole world. For the environment it's a disaster, for some it's their livelihood.		
	Alto: I am writing with reference to an article that I have read from WWF and Greenpeace about deforestation. It happens more often and it destroys our world.	0/1/2	0/1/2

importanti di uno studio effettuato sul campo al fine di fornire una risposta alla suddetta domanda.

Innanzitutto, si è stabilito che il 75% degli allievi dovrebbe essere in grado di raggiungere l'obiettivo di competenza indicato nella tabella 1. Si tratta della stessa percentuale adottata dall'*Expertgroep doorlopende leerlijnen taal*, il gruppo olandese di esperti in curricula L1.

Per la conduzione dello studio si è fatto riferimento – sia per la scelta metodologica che per il reclutamento di esperti – ad una precedente ricerca effettuata dall'*istituto per la valutazione in ambito educativo* (Cito) nelle abilità di lettura e ascolto (cfr. Feskens, Keuning, van Til & Verheyen, 2014).

Per questo studio sono stati raccolti e analizzati 1'735 elaborati scritti di probandi rappresentativi per età, sesso, livello di istruzione dei genitori e distribu-

zione geografica. A ogni elaborato scritto sono stati assegnati dei punti. Successivamente è stata sviluppata specificamente per questo studio una procedura di *standard-setting* che permette di determinare la percentuale di lavori scritti in ognuna delle tre lingue che soddisfano i criteri del livello QCER selezionato come standard. Ogni test era costituito da tre o quattro compiti di scrittura differenziati, per quanto possibile, secondo tipologia testuale, genere, lunghezza, argomento, ambito (secondo la suddivisione del QCER: ambito personale, pubblico, occupazionale ed educativo) nonché categorie quali *corrispondenza, appunti, messaggi e moduli e relazioni*. I livelli del QCER previsti dai test di inglese e francese corrispondevano ai livelli fissati dagli obiettivi didattici (Meijer & Fasoglio, 2007 – vedi tabella 1). Per il tedesco, invece, si è optato per un livello più alto rispetto a quello degli obiettivi didattici, partendo dal presupposto che chi sa l'olandese abbia più facilità a scrivere in tedesco – lingua appartenente alla stessa famiglia linguistica – che in francese.

Modalità di valutazione

I lavori scritti sono stati suddivisi in frammenti – nell'esempio riportato nella figura 1 ognuna delle informazioni dell'annuncio costituisce un frammento. A ciascun frammento sono stati assegnati dei punti distinguendo fra *efficacia comunicativa e correttezza*.

I punti sono stati assegnati da un gruppo di insegnanti mediante una procedura di attribuzione basata sull'ancoraggio a frasi modello. Alcuni studi precedenti hanno dimostrato l'affidabilità dei risultati ottenuti con questo metodo di valutazione (vedi ad es. Melse & Kuhlemeier, 2000; Van Gelderen, Oostdam & Van Schooten, 2011; Feenstra, 2014). Nel nostro studio i valutatori hanno confrontato ciascun passaggio negli elaborati scritti con due frasi modello precedentemente selezionate dai ricercatori costituenti due punti di ancoraggio: un cosiddetto 'ancoraggio basso' (un frammento di qualità relativamente bassa) e un ancoraggio 'alto' (un frammento di qualità relativamente buona).

Ad ogni frammento di qualità inferiore rispetto all'ancoraggio basso sono stati assegnati 0 punti. Se la qualità del frammento è stata giudicata migliore dell'ancoraggio basso e peggiore di quello alto, gli è stato assegnato 1 punto. Ai frammenti di qualità superiore a quella dell'ancoraggio alto sono stati assegnati 2 punti. Ai valutatori è stato fornito per ciascuno dei sei test un libretto contenente le frasi di ancoraggio basso e alto per ogni frammento. Nella figura 2 riportiamo a titolo illustrativo un esempio tratto dalle istruzioni per inglese *vwo*.

In questo modo, ad ogni esame è stato assegnato un punteggio complessivo. Successivamente, in una procedura di *standard-setting*, un gruppo di esperti internazionali ha stabilito quanti punti un esame deve raggiungere per soddisfare le richieste minime di un determinato livello QCER.

Tali esperti hanno operato secondo il cosiddetto *metodo dei gruppi contrastanti* (Sanders & Verstralen, 2011). La procedura per ciascun settore educativo (*havo/vwo*) e per ciascuna lingua comprendeva 4 fasi:

> **Fase 1.** Durante questa fase gli esperti hanno esaminato i sei test di scrittura per stabilire quale livello QCER rappresentasse lo standard per ciascuno.

> **Fase 2.** 52 elaborati scritti sono stati sottoposti alla valutazione degli esperti secondo un cosiddetto *balanced incomplete block design*: ciascun esperto ha valutato 16 dei 52 esami e ciascun esame è stato valutato da 4 esperti. Per ogni esame gli esperti hanno valutato se i requisiti minimi per il livello QCER individuato come standard sono stati soddisfatti.

> **Fase 3.** Ogni esperto ha giustificato su una piattaforma digitale la propria valutazione di sei lavori scritti.

> **Fase 4.** Gli stessi 52 esami sono stati nuovamente sottoposti alla valutazione degli esperti. Il consenso fra gli esperti è sensibilmente aumentato in quest'ultima fase.

I risultati dell'analisi dei risultati di questa procedura sono riportati nella tabella 2 che indica il livello QCER stabilito come standard, il punteggio minimo e massimo ottenibile per ciascun test, lo standard di prestazione (numero minimo di punti necessario per raggiungere il livello QCER standardizzato) e la percentuale di affidabilità (livello di consenso fra gli esperti).

Risultati

Sulla base degli standard di prestazione individuati dagli esperti si sono potute stabilire le quote di alunni che hanno raggiunto un determinato livello QCER. Tali percentuali sono riportate nella Tabella 3. Al termine della formazione *havo* oltre il 77% degli allievi raggiunge il livello B1 in inglese, mentre circa la metà degli allievi *vwo* consegue il livello B2. Per il francese

Tabella 2. Risultati della definizione degli standard di abilità scritta

Lingua	Scuola	Livello QCER standardizzato	Min.	Max.	Standard di prestazione	% di affidabilità
Inglese	Havo	B1	0	96	38	0.89
	Vwo	B2	0	92	43	0.82
Tedesco	Havo	B1	0	88	32	0.89
	Vwo	B2	0	84	37	0.82
Francese	Havo	A2	0	64	21	0.80
	Vwo	B1	0	80	37	0.73

Tabella 3. Prestazioni allievi *havo* e *vwo*

Lingua	Scuola	Livello QCER standardizzato	N	Raggiunto	95%-BI
Inglese	Havo	B1	371	77.6	(73; 82)
	Vwo	B2	298	50.7	(44; 57)
Tedesco	Havo	B1	221	55.7	(49; 63)
	Vwo	B2	365	35.1	(30; 40)
Francese	Havo	A2	188	51.1	(43; 59)
	Vwo	B1	292	52.1	(46; 58)

Bibliografia

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Consiglio d'Europa (2002). *Quadro Comune Europeo di Riferimento per le Lingue: Apprendimento, Insegnamento, Valutazione*. Milano: La Nuova Italia-Oxford.

Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2008). *Over de drempels met taal en rekenen*. Enschede: SLO.

Fasoglio, D., Beeker, A., de Jong, K., Keuning, J. & van Til, A. (2014). *ERK-niveau schrijfvaardigheid Engels, Duits en Frans*. Enschede: SLO.

Feenstra, H. (2014). *Assessing writing ability in primary education*. Enschede: Universiteit Twente.

Feskens, R., Keuning, J., Til, van A. & Verheyen, R. (2014). *Prestatiestandaarden voor het ERK in het eindexamenjaar*. Arnhem: Cito.

Gelderen, A. van, Oostdam, R. & van Schooten, E. (2011). Does foreign language writing benefit from increased lexical fluency? Evidence from a classroom experiment. *Language Learning*, 61: 281-321.

Meijer, D. & Fasoglio, D. (2007). *Handreiking schoolexamens moderne vreemde talen havo/vwo - Engels, Frans, Duits*. Enschede: SLO.

Melse, L. & Kuhlemeier, H. (2000). *Beoordeling van schrijfproducten met en zonder ankers. Is er verschil?* Arnhem: Cito.

Sanders, P. & Verstralen, H. (2011). Het beoordelen van toetscores. In: P. Sanders (red.), *Toetsen op School*, Arnhem: Cito, pp. 146-157.

circa la metà degli allievi raggiunge nella formazione *havo* il livello A2 e *vwo* il livello B1. Per il tedesco: il 56% degli allievi *havo* raggiunge il livello B1, mentre nelle classi *vwo* solo il 35% raggiunge il B2.

Interessanti al riguardo sono le informazioni tratte dalle interviste con alcuni degli insegnanti che hanno partecipato allo studio: questi hanno identificato come fattori problematici il numero di lezioni a disposizione, la disponibilità di materiali, la modalità di valutazione e il numero di allievi in classe. Fra gli insegnanti c'è unanimità nel constatare che classi grandi e scarsità di tempo da dedicare all'abilità di scrittura influiscano negativamente sui risultati di apprendimento. Fra le attività di apprendimento aventi un impatto positivo sullo sviluppo dell'abilità scritta sono invece stati menzionati scambi internazionali, una didattica dell'azione, *peer-feedback* e l'elaborazione diretta del feedback.

Conclusione

Applicando il 75% come norma di riferimento, si osserva che tale risultato viene solo raggiunto per l'inglese nella forma-

zione *havo* (B1 77%). Per l'inglese *vwo*, francese *havo* e francese *vwo* l'obiettivo di apprendimento è raggiunto da meno del 75% degli allievi. Nonostante il test di francese *havo* fosse stato sviluppato per il livello A2+, il panel di esperti ha ritenuto più adeguato stabilire come standard di prestazione per francese *havo* il livello A2. Il 51% degli allievi raggiunge tale livello, per cui si può presumere che la percentuale che raggiunge A2+ sia ancora più bassa - in entrambi i casi comunque ben al di sotto del 75%.

Nella costruzione dei test per il tedesco si è presupposto che le prestazioni degli allievi fossero superiori agli obiettivi fissati (A2+ per *havo* e B1 per *vwo*). Questa ipotesi è stata confermata durante la prima fase della standardizzazione: il panel di esperti ha stabilito che gli elaborati scritti in tedesco si prestassero meglio ad una valutazione secondo i criteri dei livelli B1 (*havo*) e B2 (*vwo*). La percentuale degli allievi che raggiunge questi due livelli è tuttavia inferiore alla norma del 75%. È assai probabile che la norma sarebbe stata soddisfatta se gli standard di prestazione fossero stati A2+ e B1.

Testo tradotto e adattato da:

Beeker, A., Fasoglio, D., Jong, K. de, Keuning, J. & Til, A. van (2015). ERK-niveau schrijfvaardigheid Engels, Duits en Frans. Onderzoek naar het bereikte niveau aan het eind van *havo* en *vwo*. In: *Levende Talen Tijdschrift* 16/2: 16-25.

TESTVALIDITÄT IM UNTERRICHT: EINE CHECKLISTE FÜR REFLEKTIERTE(RE) BEURTEILUNGEN

Cosa valutare? E come? Per rispondere a queste domande in modo esaustivo l'autrice mette a disposizione una tabella pronta ad esser consultata per compiti in classe, ma anche e soprattutto per test con maggiore visibilità come esami d'ammissione o di maturità. Oltre a ciò, il presente contributo traccia brevemente un quadro dello sviluppo storico della discussione sulla 'validità' di un esame ed espone quale sia lo stato attuale della discussione sulla validità di un test – attenta alla situazione in cui si svolge, alle competenze delle persone testate e a criteri di qualità come la sua usabilità e la sua attendibilità.

Von den vielen Tests in meiner Schulkarriere ist mir eine Geschichtsprüfung in Erinnerung geblieben, die ich in der 6. Klasse abgelegt habe: Völlig unerwartet bestand ich die Prüfung nur knapp und war damit noch unter den besten meiner ansonsten recht erfolgsverwöhnten Klassenkameraden. Die Prüfung war von einer jungen Praktikantin konzipiert worden, die offensichtlich vorausgesetzt hatte, dass wir nicht nur im Unterricht aufgepasst, sondern auch die entsprechenden Seiten im Lehrbuch gelesen hatten. Damals hielten wir diese Erwartung für unfair und protestierten leidenschaftlich. Wir hatten, wie wohl die meisten Schülerinnen und Schüler, erwartet, dass in der Prüfung nur Inhalte beurteilt würden, die wir zuvor im Unterricht von der Lehrperson vermittelt bekommen hatten. Heute weiss ich, dass auch eine andere Lesart möglich ist: Für die junge Praktikantin war das selbstständige Nachlesen der Unterrichtsinhalte wohl schlicht ein integraler Bestandteil der Lernziele des Geschichtsunterrichts und konnte damit geprüft werden. Was auch immer zutrifft – die Anekdote ist vor allem deshalb interessant, weil

sie gut illustriert, warum Testvalidität relevant ist: Beurteilungen sollten gut überlegt, geplant und kommuniziert werden, um ihre Rolle in Unterricht, Schule und Gesellschaft erfüllen zu können. Der vorliegende Artikel soll dazu einen Beitrag leisten: Eine Checkliste soll dazu anregen, die Ziele und Eigenschaften von Beurteilungsinstrumenten und Tests systematisch(er) zu erfassen und deren Schwächen (besser) zu erkennen. Einige kurze Informationen zum Begriff der Testvalidität sollen dann einen Einblick in die Forschung geben und Ausgangspunkt für weitere Lektüren sein. Die Checkliste orientiert sich inhaltlich am *Assessment Use Argument* von Bachman & Palmer (2010). Die darin enthaltenen Kriterien wurden übersetzt und für den schulischen Kontext neu strukturiert. Fünf konkrete Fragen sollen dabei helfen, eine Beurteilung in ihrer Komplexität zu erfassen. So könnte man ausgehend von einem Beurteilungsziel überlegen, wie dieses am besten zu erreichen ist (z.B. eine Prüfung am Ende einer Unterrichtseinheit über Mark Twains *Tom Sawyer* oder eine Übersicht über die

Katharina Karges | ●
Fribourg



Katharina Karges ist Wissenschaftliche Mitarbeiterin am Kompetenzzentrum für Mehrsprachigkeit in Fribourg/Freiburg. Sie arbeitet an Forschungsprojekten u.a. zu den Themen kompetenzorientierte Beurteilung, computerbasiertes Testen, Entwicklung fremdsprachlicher Kompetenzen.

Die Autorin dankt Malgorzata Barras und Peter Lenz für ihre unendlichen Ideen, Patrick Karges und Katrin Berger für ihren gesunden Menschenverstand und Mathias Picenoni für das Auf-den-Punkt-Bringen.

Ein Sprachtest ist in dem Ausmass valide, wie es ihm gelingt, Sprachkompetenzen im Hinblick auf einen bestimmten Verwendungszweck zuverlässig, fair und inhaltlich aussagekräftig zu messen.

mündlichen Fertigkeiten der Lernenden zu Beginn eines Schuljahres). In diesem Fall dienen die ersten drei Fragen dazu, das Ziel klarer zu formulieren und zu begründen (geht es in der Prüfung über *Tom Sawyer* zum Beispiel darum, etwas über den Autor, das Buch oder die Figur zu wissen, die Erzählung zu interpretieren oder die Geschichte in der Fremdsprache nachzuerzählen). Die Fragen zu Form und Inhalt des Tests können dann dabei helfen, diese Ziele bei der konkreten Umsetzung der Beurteilung nicht aus den Augen zu verlieren. Umgekehrt sollen die Fragen aber auch dazu dienen, bestehende Beurteilungsinstrumente kritisch zu reflektieren, indem z.B. gefragt wird, ob ihr Einsatz in einer konkreten Lernsituation angemessen ist, welche Anpassungen nö-

tig wären oder ob es alternative Formen gibt, um die gewünschten Ziele zu erreichen. Dabei ist es nur schwer möglich, allgemeingültige Regeln aufzustellen, denen man nur noch folgen muss. Deshalb kann es in der Checkliste auch nicht darum gehen, möglichst viele Punkte anzukreuzen. Vielmehr muss der Zusammenhang zwischen der Beurteilung, ihren Zielen und ihren Konsequenzen für jede Anwendungssituation und jede Gruppe von Lernenden neu hergestellt werden, indem man sich über jene Punkte Gedanken macht, die für eine konkrete Beurteilungssituation relevant sind. Nur so kann dem komplexen Gefüge von Lehren, Lernen und Beurteilen Rechnung getragen werden, welches die Grundlage unseres Schulsystems bildet.

Checkliste zur Überprüfung der Validität eines Beurteilungsinstruments

<p>Welche positiven Konsequenzen soll die Beurteilung haben?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Den Lernprozess der Lernenden fördern. <input type="checkbox"/> Die Lernenden in das passende Unterrichtsniveau einordnen. <input type="checkbox"/> Die Qualität des Unterrichts verbessern. <input type="checkbox"/> Über den Lernstand der Lernenden informieren: <ul style="list-style-type: none"> <input type="checkbox"/> die Lernenden selber <input type="checkbox"/> die Lehrpersonen <input type="checkbox"/> Eltern <input type="checkbox"/> Schulbehörden <input type="checkbox"/> Arbeitgeber <input type="checkbox"/> ... 		Ziele der Beurteilung
<p>Wofür wollen Sie die Beurteilung konkret benutzen? Sie möchten Entscheidungen treffen über...</p>		
<p>den Erfolg der Lernenden:</p>	<ul style="list-style-type: none"> <input type="checkbox"/> zukünftige Unterrichtsinhalte. <input type="checkbox"/> die Formulierung von Lernempfehlungen. <input type="checkbox"/> die Einteilung der Testteilnehmenden in Gruppen: <ul style="list-style-type: none"> <input type="checkbox"/> für die Erteilung von Förderunterricht, <input type="checkbox"/> für die Zuteilung zu Unterrichtsniveaus, <input type="checkbox"/> für eine Gruppenarbeit oder ein Projekt. <input type="checkbox"/> die Vergabe einer Teilnote. <input type="checkbox"/> die Vergabe einer Note oder eines (Abschluss)Zertifikats. <input type="checkbox"/> ... 	
<p>den Erfolg des Unterrichts:</p>	<ul style="list-style-type: none"> <input type="checkbox"/> das Beibehalten, Ändern oder Abschaffen einer Unterrichtseinheit. <input type="checkbox"/> das Beibehalten, Ändern oder Abschaffen eines Unterrichtskonzepts. 	
<p>Welche Informationen benötigen Sie, um die Entscheidungen (s. oben) zu treffen? Die Lernenden müssen zeigen, dass ... Die Lernenden müssen zeigen, wie gut ...</p> <ul style="list-style-type: none"> <input type="checkbox"/> sie etwas <i>mitteilen</i> können. <input type="checkbox"/> sie etwas <i>verstehen</i>. <p style="text-align: right;">} savoir-faire</p> <ul style="list-style-type: none"> <input type="checkbox"/> sie über etwas <i>reflektieren</i> können. <input type="checkbox"/> sie ein Problem <i>lösen</i> können. <input type="checkbox"/> sie eine Strategie <i>anwenden</i> können. <p style="text-align: right;">} savoir apprendre / savoir être</p> <ul style="list-style-type: none"> <input type="checkbox"/> sie etwas <i>wissen</i>. <input type="checkbox"/> ... <p style="text-align: right;">} savoir</p>		

Wie können Sie am besten die Informationen erheben, die Sie für die Entscheidungen (s. oben) benötigen?

- Die Lernenden (regelmässig) im Unterricht beobachten.
- Im Unterricht konkrete Aufgaben stellen
 - und die Lernenden bei der Bearbeitung beobachten.
 - und das Produkt der Arbeit beurteilen.
- Ein Portfolio erstellen lassen und/oder zur Selbstbeurteilung anregen.
- Einen Test oder eine Prüfung durchführen.

(Wie) Haben Sie Folgendes sichergestellt?

- Die Schlussfolgerungen über die Fähigkeiten Ihrer Lernenden sind ...
 - relevant und ausreichend** für die zu fällende Entscheidung,
 - gerecht** für alle Testteilnehmenden,
 - evtl. **verallgemeinerbar** auf andere Aufgaben inner- und ausserhalb des Unterrichts.
- Die Beurteilungskriterien werden für die Lernenden und andere interessierte Personen transparent kommuniziert.
- Die Beurteilungskriterien werden bei allen Lernenden gleichermassen und vollständig angewendet.
- Die Aussagen über die Kompetenzen der Lernenden sind (für die Adressaten der Aussagen) verständlich formuliert.
- Niemand wird aufgrund von Faktoren benachteiligt, die nichts mit der Beurteilung zu tun haben (z.B. Lesekompetenzen in einem Test zur mündlichen Produktion).
- Niemandem entstehen wegen der Beurteilung unverhältnismässige Nachteile.

Inhalte und Form
der Beurteilung

„Testvalidität“ in der Bildungsforschung

Das Konzept der „Testvalidität“ gewann zu Beginn des 20. Jahrhunderts im Zusammenhang mit der Verbreitung von Intelligenztests an Bedeutung. Damals ging man davon aus, dass man Validität nachweisen kann, indem man die Resultate eines Messinstrumentes (also z.B. eines Intelligenztests) mit den Resultaten anderer Messungen des gleichen Sachverhalts (z.B. eines anderen Intelligenztests) vergleicht. Diese Vorstellung basierte allerdings auf der Hoffnung, man könne klar definierbare Kompetenzen eindeutig messen, was sich bis heute als problematisch erweist: Es besteht Uneinigkeit darüber, welche Fähigkeiten und Kenntnisse zu einer Kompetenz beitragen und wie sie dafür zusammenwirken. Im Laufe der Jahrzehnte wurde der Validitätsbegriff daher mehrmals neu konzeptualisiert (Newton & Shaw, 2014).

Das heute vorherrschende Verständnis von Validität entstand Mitte der 1950er Jahre in der Psychologie (Cronbach & Meehl, 1955) und wurde 1989 von Samuel Messick soweit ausgearbeitet, dass es seitdem in der Testforschung vorherrschend ist: Konstruktvalidität (Messick, 1989). Die darin enthaltene Idee eines zu testenden ‚Konstrukts‘ hat wissenschaftstheoretische Wurzeln und ist nicht immer leicht zu erfassen, sie hat aber zwei ganz konkrete Konsequenzen: a) sie setzt eine bewusste Formulierung

dessen voraus, was getestet werden soll (des Konstrukts); b) der Zusammenhang zwischen den Testinhalten, den Testaufgaben, den Resultaten und den daraus gezogenen Schlussfolgerungen muss hergestellt und nachgewiesen werden (Bachman & Palmer, 2010; Kane, 2006; Mislevy & Yin, 2012).

Moderne Testvalidität verlangt nicht mehr danach, einen Test in all seinen Facetten zu validieren. Stattdessen ist seine Verwendung in einem bestimmten Kontext und mit einer bestimmten Zielgruppe zu begründen. Dabei ist insbesondere zu zeigen, dass die Konsequenzen, die sich aus den Testresultaten ergeben, gerechtfertigt sind.

Die heutige Bildungsforschung verfolgt im Zusammenhang mit Testvalidität vor allem drei Fragestellungen:

- > Welche Funktion erfüllen konkrete Tests in welchen Anwendungssituationen (vgl. z.B. Bachman & Palmer, 2010)?
- > Spiegeln formale und inhaltliche Merkmale von Testaufgaben die Kompetenzen der Teilnehmenden wider (z.B. in diesem Heft: Barras, Lenz & Karges)?
- > Was ist eigentlich ‚Validität‘, wie ist sie nachzuweisen und in welchem Verhältnis steht sie zu anderen Qualitätsmerkmalen von Tests wie Reliabilität, Fairness, Usability u.Ä. (Kane, 2006; Markus & Borsboom, 2013; Newton & Shaw, 2014)?

Bibliografie

- Bachman, L. & Palmer, A. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford, New York: Oxford University Press.
- Cronbach, L. J. & Meehl, P. E. (1955). „Construct validity in psychological tests“. *Psychological bulletin*. 52 (4), 281-302.
- Kane, M. (2006). „Validation“. In: R. L. Brennan/ National Council on Measurement in Education/American Council on Education (Hrsg.), *Educational measurement*. 4. Aufl. Phoenix, AZ: Greenwood, pp. 17-64.
- Markus, K. A. & Borsboom, D. (2013). *Frontiers of test validity theory: Measurement, causation, and meaning*. New York: Routledge.
- Messick, S. (1989). „Validity“. In: R. L. Linn (Hrsg.), *Educational measurement*. 3. Aufl. New York: American Council on Education & Macmillan Publishers, pp. 13-103.
- Mislevy, R. J. & Y. Chengbin (2012). „Evidence-centered design in language testing“. In: G. Fulcher & F. Davidson (Hrsg.), *The Routledge handbook of language testing*. London, New York: Routledge, pp. 208-222.
- Newton, P. E. & Shaw, S. D. (2014). *Validity in educational and psychological assessment*. London, Cambridge: SAGE & Cambridge Assessment.

LASSOING IN LANGUAGES AT MANPOWER

In this contribution, Romain Hofer, Head of Marketing & Communication for Manpower, Switzerland, emphasizes the importance of foreign languages on the job and how future employees are screened for their language skills. He encourages young people to work abroad and return with both language and team-player skills in the spirit of "If you can make it there, you can make it anywhere"!

An interview with Romain Hofer | Manpower

Romain Hofer is the Head of Marketing and Communication for Manpower, Switzerland. He is from Lausanne, studied Business



Administration at the University of Lausanne and has, as well, spent many years in the German-speaking world as both a child and an adult in his work at UBS and Altran in Zurich. His daily work is in English, French and German: multilingualism is part of his "vie quotidienne". Même s'il peut des fois "s'arracher les cheveux" à cause de mauvaises traductions, il a toujours la tête pleine et ne risque pas de devenir chauve.

Humorvoll – Curious – Insightful – sont les mots qui me sont venus à l'esprit pendant mon entretien avec Romain Hofer, responsable Marketing et Communications chez Manpower Suisse.

Le but de cette interview était d'apprendre comment une entreprise comme Manpower évalue les compétences langagières de ses [futur-e-s] employé-e-s et de transférer les éléments valorisés par Manpower en principes pédagogiques. Même si Romain Hofer a clairement déclaré ne pas être pédagogue, ses remarques perspicaces peuvent nous faire réfléchir sur les liens entre travail et école.

Manpower, les jeunes, les langues...

Même si elle ne reçoit que peu de candidat-e-s sortant directement de formation, l'entreprise Manpower est proactive pour se faire connaître dans les gymnases, les écoles professionnelles et les écoles polytechniques, avec l'idée de pouvoir engager ces jeunes pour travailler avec ses clients dès qu'ils acquièrent une expérience minimale.

Manpower reçoit aussi les dossiers des candidats sans expérience professionnelle ou qui sortent de l'école et cherchent du travail temporaire. Ces personnes peuvent être engagées dans les domaines généralistes, par exemple dans les chaînes de productions ou les *life sciences*, mais moins dans des postes fixes et spécialisés. Selon les secteurs, Manpower administre des tests linguistiques (langues maternelles et étrangères) internes; ils demandent par exemple au candidat de corriger des lettres ou de suivre des instructions. Et puisque ce spécialiste des RH connaît les besoins des clients en ce qui concerne les langues, Manpower a ses propres mesures internes pour évaluer les compétences langagières des candidats.

Bulletins scolaires, ou..?

Dans le processus de recrutement, le recruteur est obligé de faire un scan rapide (il doit lire entre 100-200 CVs) et a rarement le temps de regarder les bulletins scolaires. Si la connaissance de la langue est un élément important demandé pour le poste, le fait d'avoir une codification améliorera les chances du candidat. Si le

candidat écrit simplement «allemand, je maîtrise plus ou moins l'écrit, et un petit peu l'oral», cela ne veut rien dire. S'il met «C1 en allemand», c'est déjà mieux pour le premier filtre, mais s'il a une attestation ou un examen du Goethe Institut ou TEOFL réalisé dans les 6 derniers mois, c'est encore mieux.

Séjours linguistiques, soft skills... Jawohl!

Manpower encourage les jeunes à faire soit un échange linguistique pendant l'école, soit un stage dans une autre région linguistique après l'école, avant d'entrer sur le marché du travail. Si le candidat peut dire qu'il a fait un séjour linguistique de 6 mois, c'est très bien reçu. L'apprentissage par immersion est bien plus valorisé qu'un certificat de langue, même si rien n'empêche de faire certifier le progrès qu'on y a fait par un examen.

Manpower préfère les stages comme forme d'échange parce qu'ils prouvent que les jeunes ont eu une première expérience de travail en équipe et dans une autre langue.

> Quand les jeunes entrent dans les entreprises, ils doivent d'abord montrer qu'ils peuvent s'intégrer dans l'équipe et suivre les étapes d'un projet (avec un début, une fin, des objectifs, des limites, un budget). S'ils peuvent le faire dans une langue étrangère, il y a de fortes chances qu'ils puissent le faire partout.

> La connaissance de la culture du pays est souvent importante pour nos clients. Il y des entreprises qui demandent que le candidat connaisse le pays de l'entreprise – il doit être capable de mener des conversations professionnelles (y compris à propos de la géographie, des structures politiques, etc).

> Dans le monde professionnel, si l'on passe 8 heures par jour dans une équipe, on parle d'autre chose que du travail: les candidats doivent donc être capables de mener des conversations sur tous les sujets!

Il faut donc davantage que la connaissance de la langue: il faut des connaissances culturelles et la capacité de s'intégrer rapidement. L'expérience en stage dans une autre langue en est une bonne indication.

L'école primaire/secondaire et le monde du travail..?

Aborder le thème du monde du travail dans l'enseignement des langues dès l'âge de 12-13 ans n'est pas trop tôt. Cette sensibilisation est importante: il est bien de parler de Heidi, mais il est plus pertinent de donner des exemples qui montrent *pourquoi* c'est un avantage de parler l'anglais, l'allemand et l'italien.

L'école peut simuler la situation post-scolaire et sensibiliser les enfants aux cultures et aux questions qu'ils peuvent poser à ce sujet, à la découverte des pays et de la cuisine locale.

L'école ne doit pas se focaliser sur la poésie, les choses trop abstraites pour les enfants. Il y a assez d'exemples concrets de personnes comme Stress, Roger Federer ou Noah Veraguth, qui maîtrisent les langues et évoluent dans des domaines sportifs ou culturels, et qui montrent pourquoi les langues sont importantes. De plus, si les jeunes sortant de l'école ont travaillé en groupe, ce n'est néanmoins pas la même chose que le travail en équipe dans un cadre professionnel: il faudra aussi faire davantage de projets en langues étrangères.

Manpower et les langues

Manpower est une entreprise spécialisée dans les secteurs de l'intérim et du recrutement. Au sein du groupe Manpower, elle est la seule ayant un site web en 4 langues (F/D/I/E); elle compte 60 antennes réparties dans 50 villes et 19 cantons, dans les 3 régions linguistiques de Suisse; chacun de ses 13 secteurs d'activités (service public, médical, etc...), de ses 6 métiers transversaux (administration, technique, ...) ont des besoins spécifiques concernant les langues: à Berne, c'est le français et l'allemand; à Zurich et Genève, c'est l'anglais qui domine. Manpower travaille avec des traducteurs, relecteurs et rédacteurs, mais «c'est un défi quotidien de pouvoir parler le même langage, d'avoir cette sensibilité. On a tous des mentalités différentes, et si on veut que le message passe bien, il ne faut pas seulement la même langue, mais le même langage; ce n'est pas une question de traduction.»

Il existe certainement des bases qu'on est obligés d'apprendre. Mais si on ne les contextualise pas, les mots ne vont pas être appris. Et pour contextualiser, il ne suffit pas de parler du temps qu'il fait. Il existe de nombreuses plateformes, pour tous les niveaux, qui traitent de l'actualité – on peut très bien parler des Grammy Awards ou des réfugiés dans les classes. Les mêmes articles qui ont paru dans les journaux se trouvent sur les sites en langage simplifié dont le vocabulaire est expliqué aux élèves – comme

L'apprentissage par immersion est bien plus valorisé qu'un diplôme de langue, même si rien n'empêche de faire certifier le progrès qu'on y a fait par un examen.

sur *BBC Learn English News Reports*. Il y a énormément d'outils. Il faut davantage penser «*out of the box*» et ne pas prendre les manuels comme guides. Ils sont trop «*trocken*», avec peu d'émotion, d'humour, et d'actualité – ce dont on a en revanche besoin dans les milieux professionnels. De plus, au travail on ne demande pas forcément la perfection – on a le droit de faire des fautes et de recevoir des explications. Les gens doivent se décomplexer: il faut oser parler, il ne faut pas avoir

peur. Ce n'est pas en se perfectionnant qu'on gravit les échelons dans le monde professionnel, c'est l'attitude qui compte. Dans la majorité des cas, on n'attend pas des bilingues parfaits pour un poste. Il est donc contradictoire de demander la perfection à l'école – que l'on enlève des points à cause d'accents oubliés. Dans le monde du travail, on est très content d'avoir des gens qui osent et qui peuvent communiquer. Si on est trop strict, on détruit la motivation.

Help wanted!

En règle générale, Manpower cherche des professionnels spécialisés – le niveau de langue est un filtre pour chercher les employés qui conviennent:

- Maîtrise du français médical indispensable
- Niveau d'allemand B1
- Maîtrise du français et de l'anglais indispensable, bilingues
- Anglais avancé
- Français: langue maternelle

Les clients demandent parfois que le candidat soit immédiatement opérationnel: par exemple, les grands hôpitaux ont souvent un besoin immédiat de spécialistes qui maîtrisent le français médical.

LANGUAGE

All L2 taught in primary and secondary schools

TARGET

Primary and secondary school

AIMS

- > Facilitate transitioning from primary to secondary school
- > Adapt learning to assessment and vice-versa
- > Select learning activities suitable for both primary and secondary levels

TRANSITION FROM PRIMARY TO SECONDARY LANGUAGE LEARNING: USING AN ASSESSMENT FOR LEARNING APPROACH TO UNPACK PUPILS' 'SUITCASES OF SKILLS AND ABILITIES'

Il passaggio dal livello primario al livello secondario può essere problematico dal punto di vista della valutazione e della sua comprensione da parte degli allievi. A Londra, l'utilizzo del metodo *Assessment for Learning*, usato nelle scuole del primario così come del livello secondario, permette una transizione facilitata: gli insegnanti del primario usano certi elementi del secondario, come per esempio certi esercizi del libro di testo del livello secondario oppure il fatto di concentrarsi maggiormente sulla produzione scritta. In questo contributo didattico, allievi, insegnanti e ricercatori propongono delle soluzioni valide per un apprendimento continuo.

Continuity and discontinuities

Research has indicated that transition from primary to secondary, an administrative arrangement, is generally problematic across all subjects causing 'discontinuity' of learning and other adjustment problems in many cultural contexts (Audin *et al.*, 2005; Evangelou *et al.*, 2008; Mayer, 2006). The breaks or transitions are not always easy for pupils to understand or manage and require preparation and support for the pupils from teachers in order to ensure a smooth transition in the learning journey. This is especially important in language learning given the need for continuity and consolidation as well as planned progression to follow the contours of the natural language learning curve. There are many transitions in a child's education that include 'home to starting school' and 'infants to junior' before the major 'primary to secondary' break. These are followed by other transitions from 'beginning' to 'middle' secondary education before another major transition to 'advanced stages of learning'. These breaks do not respect fully a child's holistic development and natural

learning curve. The large scale research of Werlen (2005) in primary schools in Germany highlighted the need for strategies that take a long term view of a child's school life in its entirety. Such strategies would involve both the careful management of transition points for the benefit and welfare of the pupil (not just satisfying administrative arrangements) as well as the planning of cohesive learning that reflects the child's world in terms of meaningful content and a pedagogy that promotes personalised, pupil-centred progress in learning.

'Assessment for Learning' to 'bridge the gap'

There is a need for both primary and secondary teachers to have a sustained dialogue about children, appropriate pedagogies and learning skills and to be able to articulate their role and expertise to each other. On the basis of such a dialogue and collaborative planning, then pupils' learning continuity and progression could be assured via a pedagogic learning bridge, the weakest link according to Galton *et al.* (1999). Indeed, the capabilities of what

Jane Jones | London ●

taught languages in secondary schools for 16 years and is a pioneer in introducing formative assessment into schools. She is currently Head of Modern Foreign Language Teacher Education at King's College London and a member of the King's College Language Assessment Group.



PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > [SCHEDE 6 E 7](#)



Pupils of all ages can show a remarkable capacity to discuss their learning in a considered and insightful way, although they might not always be able to articulate their ideas in the formal language of learning.

primary children can do and the richness of their primary language learning experiences are frequently ignored, as are the many strategic skills in which the primary children have become adept (Jones, 2010). Many primary teachers in particular are expert in the use of Assessment for Learning (AfL) strategies that focus on pupils having considerable ownership of their learning. AfL is the practical implementation of assessment that is formative and 'for which the first priority in its design and practice is to serve the purpose of promoting pupils' learning' (Black & Wiliam, 1998: 2). Key aspects of a language teaching approach that incorporate Assessment for Learning relevant for transition include:

- > Clarity about learning objectives and the success criteria- here the teacher shares a sense of orientation of the lesson and its aims, and pupils have a good idea about how they will know when they have achieved these objectives.
- > Peer assessment- pupils being able to engage with one another to undertake peer support, teaching and assessment.
- > Self- assessment- an awareness on the part of pupils of their own progress and being able to reflect and specify the next steps in their learning.
- > Feedback- the languages classroom provides ample opportunities for a constant flow of feedback from pupils to teacher, teacher to pupils and pupil to pupil.
- > Good questioning is essential in effective teaching and pupil talk would be central in the lessons enabling the pupils to learn interactively and the teacher to have evidence of the pupils' understanding.

AfL has been proven to be of immense benefit to the learners as well as improving teaching practices (Black *et al.*, 2003). As mentioned, many primary teachers make extensive use of such techniques as a natural part of their pedagogy and are often well trained in the techniques. Some secondary languages teachers have also embraced AfL practices enthusiastically and skilfully but less extensively than their primary colleagues. This means that the pupils are well practised in AfL techniques and capable of directing their own learning to a considerable degree. In research by Jones (2010), both primary and secondary pupils made it clear, with regard to their language learning and teaching experiences, that they understood AfL and felt comfortable

with, for example, peer and self- assessing as I will detail in the next section.

What pupils say about transition and AfL

It is important to ask for and listen to the views of pupils in all that affects them directly. Children are, it has been found, well able to give revealing and honest views about their learning as Flutter & Rudduck (2004: 7) found in their extensive research with children about teaching and learning: Our research with primary and secondary pupils across the UK has demonstrated that pupils of all ages can show a remarkable capacity to discuss their learning in a considered and insightful way, although they might not always be able to articulate their ideas in the formal language of learning.

Researching over the years with children, when primary pupils are asked a question about whether they thought their primary language learning would help them with languages at secondary school, the pupils show their own strategic awareness of how one stage of learning can underpin the next one, for example:

- > 'Yes I think it will help because we have learnt how to remember to pronounce words'
- > 'Maybe, yes because we have learnt the basics and got the flow, no because we will be learning more difficult words and have to write them down rather than say the words aloud.'

In a discussion about AfL in their learning, sample responses from primary pupils included:

- > 'We have lots of AfL in our lessons. Sometimes the teacher forgets the learning objectives so we have to remind the teacher or she sometimes asks one of us to do the learning objectives. Sometimes the learning objective is a question'
- > 'We do peer assessment every lesson. It used to worry me but now I have got used to it, I learn a lot from my friends and get ideas from them'
- > 'Our teacher asks us to traffic light our work [use green/amber/red according to level of understanding] so we can do this in our exercise books with coloured sticky dots and we traffic light other groups when we are doing group work plays'.

Secondary pupils mostly claimed to have enjoyed and profited from their early ex-

periences of language learning yet there was always widespread comment on apparent repetition by secondary teachers of work already covered in primary schools, as shown by this pupil's comment:

> 'I couldn't believe it when the teacher started teaching us 1 to 10 in French as we had done 1 -100 in primary school'.

New secondary pupils who were interviewed, being the same pupils interviewed when they were at primary school made reference to AfL at secondary school in these ways:

- > 'Some teachers do AfL all the time. The French teacher always starts with the learning objectives. Some teachers don't so we ask them'
- > 'We do peer marking in French but it's not always clear what we have to mark so we get worried'
- > 'The teacher tells us to do self- assessment on our homework before we hand it in'
- > 'Some teachers just tell us to write stuff down and don't tell us what we are doing but we can usually figure out what is going on'

These insightful comments from the pupils provide evidence of their understanding of the different ways in which they are taught and the indelible impact of AfL on their learning so that even in the absence of clarity about learning aims with some teachers, they have been able to develop strategies to 'figure it out' for themselves.

How teachers can respond constructively

To move forwards in learning rather than going over old ground, and in order to build and extend learning skills in a sustaining way, it is essential that primary and secondary teachers find a way to have a dialogue about learning during the transition period. There are still many misconceptions and a lack of trust that exist between primary and secondary teachers, for example, :

'A lot of secondary teachers seem to perceive what we primary teachers do as just playing, they don't see a lot of the grammar we put in and also what levels the children come out at, some children are really extremely gifted and we're really getting them to a good standard but the secondary teachers just wipe the slate clean.'

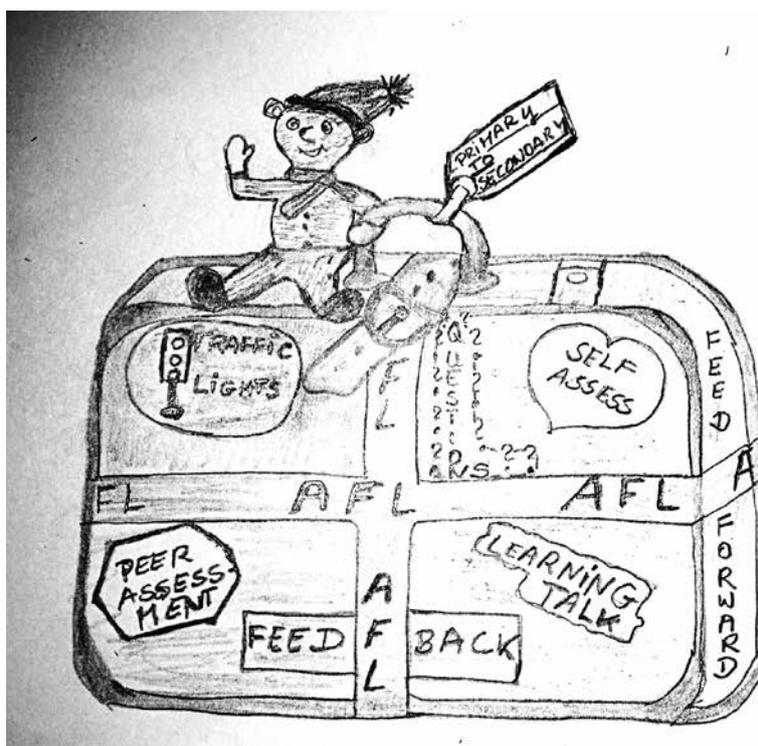
Conversely, a secondary teacher emphasised why she felt it necessary to 'wipe

the slate clean' in some respect: 'I often have to unteach incorrect French that has been taught incorrectly at primary school and start some grammar that hasn't been done at all'.

However, I want to draw attention to this devastating comment from a primary Headteacher who said:

'My pupils have been learning French since they started school. They have learnt a lot and learnt to be very independent learners. We have a whole school approach to AfL in this school. They take suitcases of skills and abilities with them to secondary school but these are simply ignored.'

This strongly- worded statement highlights the importance of primary teachers providing a sustainable language learning grounding and of secondary teachers validating and building on what children have learnt to ensure pedagogical continuity.



Drawing by Michaela Thomas

There was an acknowledgement by both primary and secondary teachers that dialogues about teaching and learning for successful transition across the phases were inadequate. A useful dialogue would establish the importance of collaboration and affirm a belief in the whole pupil language learning journey, through the drawing up of a kind of 'road map' (Black & Jones, 2007). The 'road map' would ideally include mutually useful visits to

There was an acknowledgement by both primary and secondary teachers that dialogues about teaching and learning for successful transition across the phases were inadequate.

each other's schools, sharing of topics, vocabulary, structures and skills taught and a discussion about assessment and AfL. Pupils would have a voice in this discussion. This would lead to some sharing of ideas on transitional learning activities that would draw on AfL some of which I discuss in the next section.

Transitional learning and AfL: ideas from the pupils and student teachers

These ideas are not all new but are examples of excellent creative transitional learning activities. The learners themselves had ideas about this, primary learners suggesting the following ideas for primary school:

- > Having a go at a lesson from a secondary text book, doing this in groups
- > Reading simple stories, poems or rhymes to establish phoneme-grapheme link
- > Doing spelling and grammar 'challenges' such as competitions and quizzes
- > Learning to write longer texts such as text messages to a friend or short letters.

It can be seen that such activities proposed by the learners respond to AfL principles through building on prior knowledge, and creating challenge for learning to move it forwards. Student teachers who are secondary specialists at King's College London recently shared their ideas for the first few weeks of learning in the secondary school:

- > The pupils can make posters of what they have done in their primary schools. They can do individual ones or do this in pairs or groups to start a learning dialogue and create rubrics for their work together
- > How about the pupils create a large learning tree for the wall and the pupils add on 'leaves' to show what they have learned and can remember? They can keep adding to this as they remember items
- > I saw them do little quizzes in the primary school so we could do lots of little quizzes with groups taking part to see what they already know

> My idea is a kind of Languages Show where the pupils present in groups to their peers what they feel confident about

> Pair up the pupils with learning buddies so that they can support each other

> Plan the teaching around a series of dialogues/conversations so that the pupils can use the vocabulary and phrases they have learnt

> Peer assessment – 2 stars and a wish-every opportunity as the pupils are good at this and they really learn from each other. It's a perfect opportunity for the teacher to see and hear what the pupils can do and what their misconceptions are'.

> The children can make passports containing as much information as they can about themselves. They could then use these in pair work to ask questions of each other and where the teacher sees how fluent and how many structures they can use

> Ask the pupils in groups to create plays using a list of items given by the teacher that relate to work previously covered and that include some new but guessable items.

> Get the classroom organized by labeling items and creating classroom rules

> Groups of pupils could prepare vocabulary/sentence mind maps based on topics that they covered in primary that will be useful in secondary

> Start the lessons with simple games (simple Maths games in the target language to practise numbers) so to make an explicit link to something the children will be working on in more complexity.

> Write a simple story with pictures. They could even contain a couple of simple activities such as 'Find the word for ...'. All the stories could then create a library and the children read each other's.

The student teachers stressed the need for revision and consolidation in early lessons in the secondary school to give confidence and to ascertain what the pupils already knew, essential to know how to plan the next stages of learning. Their

suggestions also emphasise collaborative learning as central with lots of pair and group work. Such activities would allow the teacher to become aware of where the children were in terms of their knowledge, and identify the pupils' strengths and the weak areas and allow for a dialogue between the teacher and the pupils in order to enable the pupils to create the goals themselves and take ownership of them. These activities would also reflect the extensive peer work that is embedded in the primary stage and would 'keep it warm' even in the more formal setting of the secondary school.

The suggestions above are based on common sense, practical conceptions of effective teaching. These practical ideas can be part of a repertoire that all teachers can adapt very easily on the basis of a shared understanding and implementation of an AFL approach to provide a strategic pedagogical bridge and cohesion in pupil learning as they transition from primary to secondary or from teacher to teacher.

Conclusion

All teachers on both sides of the transition bridge need to assess and value the 'suitcases of skills and capabilities' that the children prepare and take with them on their learning journeys. On the basis of the contents of those suitcases, teachers can develop their teaching in a way that provides a pedagogic bridge that is enabling and validating for the pupils, recognising previous learning and devising ways to build on this. It is a huge disappointment for learners to be told that their previous learning counts for nothing, a bit like losing a suitcase or other valued item on a journey that becomes lost forever. In the words of a Head teacher: *'If language learning is a learning journey and I happen to think that it is a very exciting journey, then we need a clearer route and closer links'*. AFL provides a strategy for teachers to frame their teaching to provide continuity and for the pupils to be able 'to figure' out, on a confident basis, the next steps in their language learning journey, re-packing their 'suitcases' with new skills and abilities.

References

- Audin, L., Luc, C. & Archimbaud, M.** (2005). *Enseigner l'anglais de l'école au collège: comment aborder les principaux obstacles à l'apprentissage*. Paris: Hatier.
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D.** (2003). *Assessment for Learning: Putting it into Practice*. Maidenhead: Open University Press.
- Black, P. & Jones, J.** (2007). Formative Assessment and the learning and teaching of MFL: sharing the language learning road map with the learners. *Language Learning Journal*, 34, 4-9.
- Black, P. & Wiliam, D.** (1998). *Inside the Black Box: Raising standards through classroom assessment*. School of Education, King's College London, United Kingdom.
- Evangelou, E., Taggart, B., Silva, K., Melhuish, E., Sammons, P. & Siraj-Blatchford, I.** (2008). *What makes a successful transition from Primary to Secondary School? Research Report DCSF-RR019*. London: DCSF.
- Flutter, J. & Rudduck, J.** (2004). *Consulting Pupils: What's in it for Schools?* London: RoutledgeFalmer.
- Galton, M., Gray, J. & Rudduck, J.** (1999). *The Impact of School Transitions and Transfers on Pupil Progress and Attainment*. DCSF Research Report No 131.
- Jones, J.** (2010). The role of Assessment for Learning in the management of primary to secondary transition: implications for language teachers. *Language Learning Journal*, 38, 2, 175-191.
- Jones, J. & Coffey, C.** (2013). *Modern Foreign Languages 5-11. A Guide for Teachers*. London: Routledge.
- Jones, J. & Wiliam, D.** (2008). *Modern Foreign Languages inside the Black Box*. London: GL Assessment.
- Mayer, N.** (2006). Fremdsprachenunterricht als Kontinuum- Der Übergang von der Grundschule in die Sekundarstufe. In: W. Gehring (Hrsg.), *Fremdsprachenunterricht heute*. Oldenburg: BIS-Verlag, pp. 215-33.
- Werlen, E.** (2005). *Zwei Jahre Fremdsprachenunterricht an Grundschulen in Baden-Württemberg: Erste Zwischenergebnisse der Wissenschaftlichen Begleitung der Pilotphase Fremdsprache in der Grundschule. Fremdsprachen praktisch: Beiträge und Mitteilungen aus dem Landesverband Baden-Württemberg des Fachverbandes Moderne Fremdsprachen*, 17: 63-76.

LANGUAGE

App available in 35 languages

TARGET

Ages 5-18

AIMS

- > Real-time assessment data collection
- > Easily manage reporting and communication with parents

MATERIALSClassDojo App, downloadable from www.classdojo.com

APP REVIEW: THE BENEFITS OF USING CLASSDOJO FOR ASSESSMENT DATA COLLECTION IN THE FOREIGN LANGUAGE CLASSROOM

ClassDojo is a free tool for teachers with numerous functions that encourage active involvement of learners and parents. Teachers can record notes on learners, keep in touch with parents, document daily activities and motivate learners. Caleb Howard describes how he is successfully using this tool to motivate and assess his learners.

● Caleb Howard | Vineland, NJ, USA

has been an elementary school Spanish teacher in New Jersey, USA since 2004. To keep his patience level high (and to remind him that language learning is hard), he tries to always be practicing a new foreign language. He grew up living in Ecuador and Costa Rica and completed undergraduate study at Houghton College in New York, USA. Every year he teaches over 600 students.



ClassDojo....My students love it and it's made my life, as a foreign language teacher, a lot easier. It allows me to...

- > ...save time,
- > motivate students and
- > record assessment data...

...without having to pay anything.

It's easy to do with a free app from Class-Dojo.com. It works by allowing teachers to award and deduct points based on student performance in the classroom.

Why do STUDENTS love it?

1. It's **attractive**. Great layout. Eye-catching design.
2. It's **personal**. Student names are next to their avatar/character. Students and their parents can review their progress anytime and anywhere.
3. It's **customizable**. Teachers can allow students to choose whatever avatar they would like.
4. It's **engaging**. Students can watch their progress and track their performance. Sound effects keep the students engaged. The random-student-picker keeps students paying attention.

Why does their TEACHER love it?

1. It's **versatile**. Since it's web-based you can access the program anywhere you can access the internet. Keep points private by only accessing it with your phone/tablet. Make it public by projecting points in front of the class to add a competitive dynamic. Students can compete to see whose avatar receives the most points. Some teachers offer awards for point accumulation.
2. **Quickly record assessment data**. Since positive and negative points are customizable, you can give a point for anything. For example, during (or immediately after) a student's oral presentation, you can use your device/computer to award pre-assigned, customizable points based on your assessment rubric. In a matter of seconds you can
 - > click the student's name/avatar.
 - > click 4 points, 3 points, 2 points or 1 point based on things like how intelligible/fluent the presentation was or based on how many target vocabulary words were used.

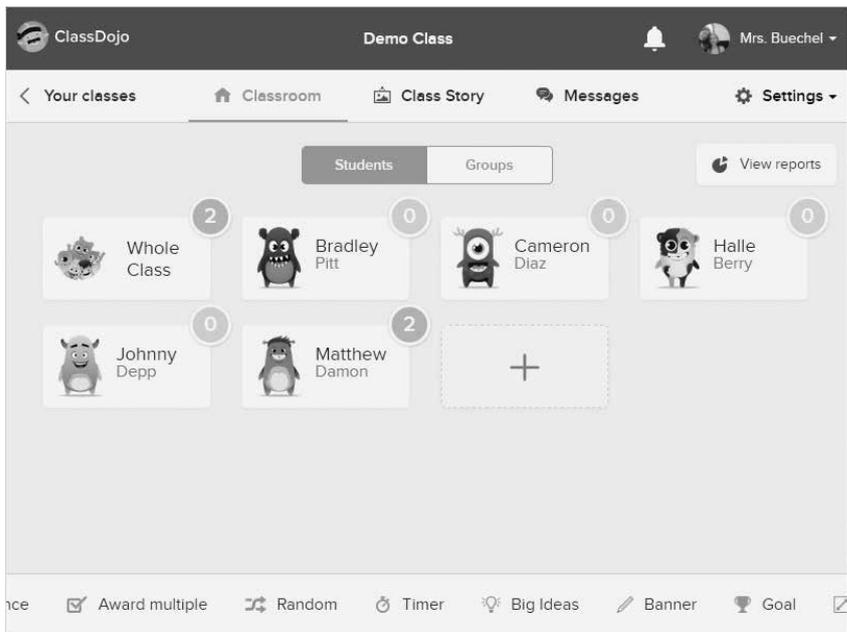
➤ PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > [SCHEDE 6 E 3](#)

If you wish, you can attach personalized comments or notes to any piece of assessment data you enter. The information is automatically stored, time-stamped and, by clicking 'Reports,' you can review the data whenever you're ready to formally report grades. (Note: I don't use ClassDojo for ALL assessment data collection. Generally I keep written assignments separate and save ClassDojo for oral performance tasks. Lesson Note is another noteworthy, free app for managing interpersonal mode performance assessments.)

3. **Randomly choose students to participate.** No one feels like the teacher is singling them out or playing favorites. It motivates students to be prepared to give an answer at any time. Use this function as much or as little as you'd like.

4. **Effectively redirect off-task behavior.** Even if you're staying in the target language with novice students, they will understand when they're doing something right or wrong because of ClassDojo.com's sound effects. Give students extra incentive to stay on task by connecting parents (in seconds) to the ClassDojo account.

5. **Consistently and manageably award exemplary behavior.** Generally, I give points to students every time they use the target language as a way to motivate them and track their participation. Since the points are customizable you could award points for anything including, pronunciation, fluency, ability to improvise, etc.



The interface of ClassDojo.

Need lesson ideas for what you can do with ClassDojo.com in your foreign language classroom?

Visit my 'Tips For Staying In The Target Language' blog (www.SenorHoward.com/blog) and type 'ClassDojo' in the search bar.

3 Questions to Caleb Howard about ICT and Teaching

Babylonia: Señor Howard, do kids bring their cell phones to class?

Caleb Howard: My students are too young for cell phones still. Also, I don't see that cell phone use would be necessary for ClassDojo. It would be something that students and their parents access outside of school hours to monitor progress/performance.

B: Would you say that the kids are using it as well for the board functions/posting homework/study tips and such?

CH: Regarding...board functions...I'm not familiar with that term. I don't think ClassDojo does anything with posting homework or study tips. I use Google Classroom for that. (That would be good content for another article/topic :) There are other apps out there, like Class123 which offer board functions, but then they are perhaps lacking a function ClassDojo has.

B: Do the kids know when you are using the randomizer (so do you use this function all the time or for specific activities -do you announce it)? Do you record whether the kids knew the answer/could respond? If so, can the kids see this and then in your experience, are they motivated not to let it happen again if they could not respond?

CH: Usually if I randomly choose a student and they don't know the answer...I stick with them and scaffold and give them cues until they get it. I never take away a point for students not knowing an answer. ...although I DO TAKE AWAY points all the time for students that aren't paying attention...or for students that make a peer feel embarrassed for trying/participating.



STANDARDS-BASED GRADING TOOLS

Alors que tous les enseignants suisses connaissent Lehreroffice et que de plus en plus d'entre eux se familiarisent avec ces fonctionnalités sur schulwebseite.com, il existe bon nombre d'outils similaires, axés sur l'évaluation standardisée. Cet article propose une brève description des plus communément utilisés.

● Laura Loder Büchel |
Zürich

The following tools can be used to help foreign language teachers collect data on learner performance in sophisticated ways – encouraging standards-based grading and monitoring learner progress. With the following tools, language teachers can directly enter the CEFR descriptors and record multiple instances of learner performance PER DESCRIPTOR.

Jump Rope
<https://www.jumpro.pe>

JumpRope allow teachers to collect data on what learners have learned (not just what they have done). Each score and set of items on tests and performance assessments have to be linked to standards so that standards-based reports can be generated. By working with this system, from the start teachers must think of what they are teaching and ultimately assessing.

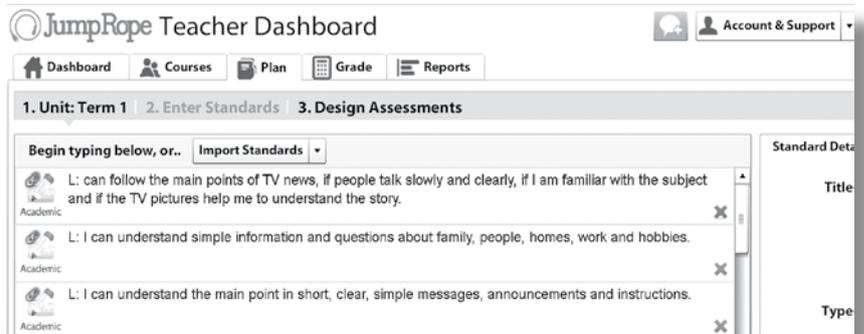
Active Grade
<http://activegrade.com/>

Active Grade gives the example of Mike (<http://activegrade.com/#features>):
"Take Mike ... Mike's second semester of Spanish class is going pretty well and he's

learned a lot of new vocabulary but his current grade is a C because he struggled with conjugating verbs on his midterm exam. Mike may understand that he needs to work on conjugating verbs, but even with a high score on the final exam he may not be able to raise his "C" to an "A." Or worse yet, his teacher may have moved on from simple verb conjugation to irregular verbs and Mike would have no opportunity to demonstrate his understanding of the initial concept. With the ActiveGrade approach, Mike's teacher would work with him to identify that it's the future tense causing Mike's difficulties, and after Mike clearly demonstrates an improvement in and understanding of this practice, the intelligent calculation done by ActiveGrade can reflect that achievement in Mike's final grade, despite Mike's initial struggle on the midterm. Mike's final grade becomes a reflection of the core skills and concepts he's learned in Spanish class, not simply an average of how he scored on the tests and assignments given."

Active Grade provides anyone interested with a very good overview of what standards-based grading is and a starter kit to help teachers set up a standards-based grading policy in their classroom.

➤ PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > [SCHEDE 6 E 3](#)



Traditional Math Gradebook

STUDENT	Homework 1	Homework 2	Quiz 1	Homework 3	Final Grd
Antonio P.	70%	78%	70%	75%	C
Arlene B.	76%	83%	76%	81%	B-
Benjamin P.	97%	95%	93%	98%	A
Bertha R.	70%	79%	70%	79%	C+

SBG Math Gradebook (ActiveGrade)

STUDENT	Add Like Fractions	Find Slope of Line	Find Area of Triangles
Antonio P.	3.5	3	3
Arlene B.	3.75	2	4
Benjamin P.	4	3.2	4
Bertha R.	3	2	2

Pearson Power School

<http://www.powerschool.com/powerschool-sis/>

PowerSchool includes a powerful gradebook and teacher tools that provide all the features teachers need in one place, making it easy-to-use and learn. All classes, rosters, student demographic information, grading periods, standards, rubrics and grades scales are automatically loaded into the gradebook. All data flows back to the central database in real time, providing all stakeholders, including parents and students, with instant visibility to assignments, scores, grades, comments and progress toward each standard.

Ken O'Conner writes about how Power School can be used for standards based grading here: <http://www.oconnorgrading.com/PowerTeacher.php>

Powerful Teacher Tools and Resources

PowerSchool includes a powerful gradebook and teacher tools that provides all the features teachers need in one place, making it easy-to-use and learn. All classes, rosters, student demographic information, grading periods, standards, rubrics and grades scales are automatically loaded into the gradebook. All data flows back to the central database in real time, providing all stakeholders, including parents and students, with instant visibility to assignments, scores, grades, comments and progress toward each standard.

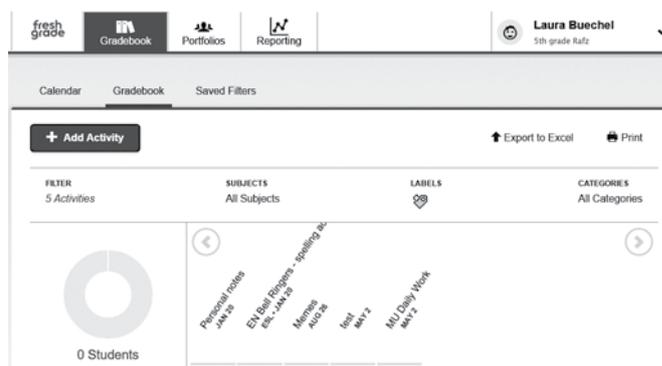
- Fast, easy attendance recording
- Automatic grade/standard calculation
- Schedules and student rosters automatically loaded and updated
- Link assignments to content for parents and students
- Supports co-teaching
- Supports district, state and Common Core Standards
- Email parents/guardians and students
- Create intervention groups and project teams and create "what-if" grade calculations
- Track each student's progress toward standards

Pearson Power School

Fresh Grade

<https://www.freshgrade.com>

Fresh Grade is a free tool offering a gradebook for standards-based grading and space for individual learner digital portfolios. On the website and Facebook site, they offer interesting discussions not only about the product, but also about assessment.



LANGUAGE

English as a foreign language

TARGET

Secondary Level I, Third Class

AIMS

- > Assess integrated performance of 5 skills
- > Construct the assessment criteria with the students
- > Increase critical thinking about assessment

MATERIALS

- > Voices 3, Unit 2, Task B
- > Voices 3, Unit 2, Worksheets 4,5,6,7,8



INTEGRATED PERFORMANCE ASSESSMENT IN ENGLISH: HOW TO ASSESS AND DEVELOP SEVERAL ASPECTS OF A DEBATE ON ANIMAL TESTING

L'article présente une expérience d'évaluation en anglais L2, basée sur le manuel *Voices 3*, qui a été menée avec succès par Esmeralda Meili, enseignante d'anglais au secondaire I dans le canton de Zurich. La tâche et l'évaluation présentées intègrent la compréhension orale et écrite, la production orale et écrite ainsi que l'interaction. En plus de la présentation du matériel et de la procédure d'évaluation, l'auteure résume également sa «politique» d'évaluation.

● Esmerelda Meili | Winterthur

Esmerelda Meili is a Secondary School English Teacher (Class 1- 3, level A-C) in Winterthur.



This assessment activity is based on *Voices 3 / Unit 2 Task B* (*Voices* book pp. 30-32 plus worksheets, see Figure 1). In order to explain this assessment I'll use the different language skills as described in *Lingualevel*, highlighting where each skill is focused on and needed for this task.

-  **Listening**
-  **Reading**
-  **Spoken interaction**
-  **Oral production**
-  **Writing**

Source of icons: www.lingualevel.ch

In this task the students had to read two articles giving background and current information on animal testing. We did the jigsaw reading in two big groups ( Reading), each with one of the articles and subsequently the learners

formed pairs. One person having read article A got together with someone who had read article B and they both informed each other about what they had read ( Spoken interaction). Once both partners had done this, they had to write down ( Writing) all the points for and against animal testing.

With the help of the language support (Figure 2) they had to create a debate, starting with an opening statement and one pupil arguing for, the other against animal testing (5 arguments was the minimum needed to be debated).

These debates were then recorded by the students on their mobile phones and sent via WhatsApp to the teacher. Before we structured the debates the pupils and I determined which aspects would be evaluated ( Spoken interaction), made sure that everybody understood what the criteria meant (examples were given for each category) and phrased the evaluation categories in simple language. We came up with the following assessment criteria for the debate with 1 as poor and 4 as excellent (See *Evaluation grid* on the next page).

Evaluation grid

Opening statement	1 2 3 4
The debate makes sense	1 2 3 4
Respond to other person	1 2 3 4
Use of English	1 2 3 4
Conclusion	1 2 3 4

Once the assessment criteria were determined, pairs were given time to rewrite their arguments, remodelling them to meet the criteria best. Subsequently they recorded their debates (🗣️ Oral production) and sent them to me.

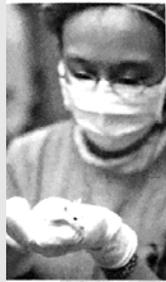
We discussed (🗣️🗣️ Spoken interaction) the sort of vocabulary that would be needed to write comments during the evaluation of the debates and students were aware that their spoken texts were being evaluated out of 20 and their assessment skills were being evaluated out of 10. The two marks together were entered for their final grade in this task.

After playing each recorded debate, we discussed and gave feedback on the different aspects (🗣️🗣️ Spoken interaction) before reaching a consensus between teacher and pupils on the final mark allocated to the pair out of 20. The focus was on speaking (🗣️ Oral production), listening comprehension (👂 Listening) and being able to assess (I call this skill critical thinking) what you hear through writing (✍️ Writing); this demands full concentration and involves all of the language skills, thus making the mark "Kompetenzübergreifend". The evaluation sheets were collected after we had heard all recordings and the written feedback evaluated by the teacher.

It takes a lot of time to prepare the students to be able to evaluate each other's work objectively. The more the students become aware of what is being evaluated and how, the more they understand that their opinions are important and taken seriously and the more they are involved with setting up the evaluation criteria, the easier it becomes for them to be a part of the process.

I love teaching with the *Voices* books and Task Based Learning. I simply follow the *Voices* Units and usually complete 6 to 7 units, doing 10 to 12 tasks per year plus other tests focusing on single language

- 3 Read this extract from an article in a consumer magazine about animal testing.
 - a Highlight six keywords or short phrases summarizing the main ideas.



Animal testing has been commonly used in scientific research for more than one hundred years. And for a long time, it has been a very controversial topic. Keeping animals in testing laboratories is often considered cruel. They may feel discomfort and stress, might suffer and ultimately even die. But it is not simply black or white, there are many different aspects that have to be taken into consideration. Pharmaceutical companies test their drugs on animals – drugs that might save thousands of lives. And cosmetics need to be tested to ensure that the products won't harm a person's eyes or skin. But we at Consumer Magazine still ask ourselves: Should animals be used to test beauty products? Should animals really die for our society's wish for eternal beauty?

- b Compare keywords with a classmate. If you have marked different words, decide on the six words or phrases that best summarize the text.
- 4 Divide your class into two groups, A and B.
 - Group A: Got to Worksheet 4.
 - Group B: Read the overview of animal testing carefully on your own and do Steps a–c.

** Animal testing: an overview

In the early 20th century, it became common to test the safety of cosmetics and drugs on animals. The reason was a tragedy which took place in 1937. A drug labelled "Elixir of Sulfanilamide" killed more than 100 people.

According to English law every new drug for medicine has to be tested on two different types of live mammals. The aim is to ensure that the drug does not cause harm to humans. But animals pay a high price for this.

There are laws which regulate how lab animals have to be treated during testing, so that they suffer as little as possible. Nevertheless, animals can get stressed and some animals might be treated badly by some scientists.

It can be debated if animal testing for medical reasons should be supported or not.

Supporters of the practice claim that many medical achievements in the 20th century were based on the use of animals in some way. Take the following examples:

- Through animal testing scientists could develop vaccines against illnesses like measles, mumps and rabies.



Figure 1

Task B (pp. 30–32)

6
b

- I (don't) agree with ...
- We should also take into consideration that ...
- I don't like the idea of ...
- It's horrible that ...
- I'm glad that ...
- You said in your opening statement that ... but I disagree with ...
- Another argument in favour of/against animal testing is ...
- You might be right but ...
- Without animal testing we would never have found/achieved/developed ...
- Animal testing ensures/does not ensure that ...

Figure 2

skills or grammar and vocabulary. My general grading policy involves the following per semester:

- > reading, writing, speaking and listening tests (about 12 per term, more writing(4) than other skills)
- > Tasks (about 5 per term like PowerPoint presentations, debates, designing a flyer, writing blogs ...)
- > A symbol system (o, +, ++) to indicate contributions to discussions, comments on others' work.

References

Stotz, D. & Rehm, S. (2011) *Voices 3, Course Book*. Zürich: Lehrmittelverlag Zürich.

Lenz, P. & Studer, T. *Lingualevel*. schulverlag plus AG. Available at: <http://lingualevel.ch/>

**LANGUAGE**

English as a foreign language

TARGET AND LEVEL

Primary School, Secondary School I and II, A1 - A2

AIMS

- > Think about how testing and teaching are seamless;
- > Apply Bloom's Taxonomy to test development.

OLDIES BUT GOODIES: BLOOM'S TAXONOMY AS A GUIDE FOR TEACHER-MADE TESTS

La taxonomie de Bloom peut être utilisée pour de multiples objectifs dans le cadre de la formation initiale des enseignants. Cette contribution propose un exercice d'utilisation de la taxonomie de Bloom pour créer un test, à travers lequel les futurs enseignants réfléchissent à l'adaptation de matériel pré-existant. Cet exercice permet non seulement de penser l'évaluation mais également de considérer l'organisation des savoirs à différents niveaux.

● Laura Loder Büchel |
Zürich

Introduction

Bloom's taxonomy (1956) is a helpful tool for teachers to ask different questions, to get away from the mere "comprehension" element of teaching language and to provide another angle of analyzing planning. There have also been many other models over the years such as the reworked Bloom's (Anderson & Krathwohl, 2001), Marzano's New Taxonomy (2000) the SOLO (structure of observed learning outcomes, Biggs & Collis, 1982), Socratic questioning and a good overview can be found in O'Neill & Murphy (2010). The challenge was given to the students at the Schaffhausen University of Teacher Education to apply Bloom's taxonomy to the creation of a test. This exercise was done within one 45-minute lecture and was meant as a thinking exercise in an attempt to get away from vocabulary translation tests, an extremely singular way of assessing language, to thinking about what it means to know words and what it means to test language skills meaningfully and encourage thought. At a later point in the course, students thought more about features of language

to be tested such as gist and inference for listening and reading, and mechanics, accuracy and range of writing (to name a few examples of subskills and constructs). The following examples of a pre-service teacher-developed test are based on the Anderson & Krathwohl (2001) revised version of Bloom's taxonomy for primary school learners using the First Choice series Senses (Littlejohn & Schofield, 2004) books towards the end of their first year of formal English language instruction.

Each example begins with the larger category of the respective cognitive process followed by the exact text from Anderson & Krathwohl (2001: Appendix 5.1). The concrete ideas listed below the category are first based on the primary level, but the tasks can certainly be used at the secondary or higher levels of language proficiency.

Practical ideas

Category 1: Remember – "Retrieve relevant knowledge from long-term memory"

Category 2: Understand – "Construct meaning from instructional messages, including oral, written, and graphic communication."

➤ PIÙ ARTICOLI SU QUESTO TEMA:
WWW.BABYLONIA.CH >
ARCHIVIO TEMATICO > [SCHEDE 6 E 19](#)

Bloom's lends itself to
 a relative expedient
 understanding of some
 categories of thinking
 skills and is a good support
 for beginning teachers.

Our thinking of the differences between remembering and understanding in the language classroom were that remembering was arbitrarily knowing words but not necessarily what they mean, and understanding would be using them in appropriate contexts. The other thought was that perhaps *remembering* is more receptive understanding and thus *understanding* would be related to production. With this in mind, for remembering, students came up with the following activities and discussions entailed about which cognitive domain they represented. Learners can simply match words from the unit to pictures or definitions. They can also answer true/false questions such as:

- > You touch with your hands.
- > You smell with your hands.

Learners can label a picture where people are doing things with their senses (smelling roses; petting a dog) or be prompted to say or what a specific person is doing. Students also came up with a quick dialogue (Figure 1) where learners can identify who likes what.

Another activity is to have learners describe various objects that were NOT in the book (Figure 2), but with the language that was taught in the book. Here only content and spelling are assessed, but the idea could be used for the creation of a riddle activity.

Figure 1: Dialogue

Listen to the two children talking together. Decide how they feel about the different objects. If they like it, put a smiley ☺. If they don't like it, put a frown ☹. Complete the chart. The first one is done for you.

Sarah (S): I really like oranges!
 John (J): Me, too, they taste really good!
 S: Do you like cats, John?
 J: No, they're horrible!
 S: I think they're nice! I like to smell them, but dogs, on the other hand, smell bad.
 J: I think they are nice. If you wash them, they smell good. Now listen to the music. What do you think about it?
 S: It's horrible.
 J: Oh, yes, that's true! Let's see, what's next: Coffee – what do you think?
 S: Ughh, it's too hot! What do you think?
 J: It's not too hot – and it smells good! What's last on our list – the flowers – oh, they smell great!
 S: Ahhchoo: They're horrible! They make me sneeze! Better get plastic ones next time!

	Object	Sarah	John
1.	oranges		
2.	cats		
3.	dogs		
4.	the music		
5.	coffee		
6.	flowers		

Figure 2: Describing novel objects

Object	What is it? It's a(n)...	Shape It's...	Color It's...	Texture It's...	Flavor It's...
	orange	round	orange	rough	tangy
	cereal box		colorful		
	cookie				

Reading comprehension with short answer questions for a text similar to one in the textbooks (Where does the light go? What does your brain do in the dark? Figure 3) is also a viable option.

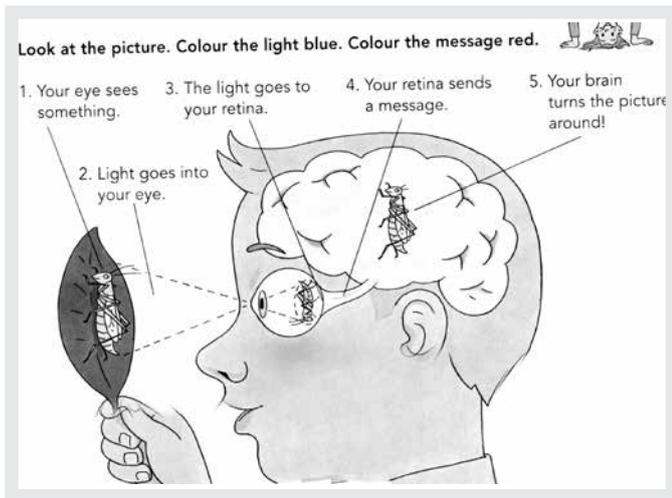


Figure 3. Senses Topic Book reading text p. 9.

Category 3: Apply - "Carry out or use a procedure in a given situation."

The simplest form of assessment with Senses that would show application would be to let the learners first do an experiment in class as part of the learning process (Figure 4). Then, give them the task of modelling this experiment orally with a partner for you at a later date. This would entail them learning / writing a role play such as:

A: Close your eyes and hold your nose! I am going to give you something to eat.

B: Okay! I'm ready.

A: Good, here we go. What is it?

B: I don't know! It's chewy. Is it gum?

A: Yes! But which flavor?

B: Oh, I don't know!

This would recycle language used previously and classroom language ("I don't know", "I'm ready") and allow the learners to take a context they are familiar with to apply their new language skills. Although normally with applying new knowledge, learners should be put in a novel situation, in Senses the learners are not required to say anything other than the name of the food. Like this, they are now producing the activity instead of passively doing it. Other ideas would be to let them do another experiment of their own choice that has to do with senses (such as coloring food and seeing if, for example, an orange tastes as good if it's been dyed black).

Finally, an open prompt such as "Look out your classroom window. What information do your senses give you?" would be an exercise in application.

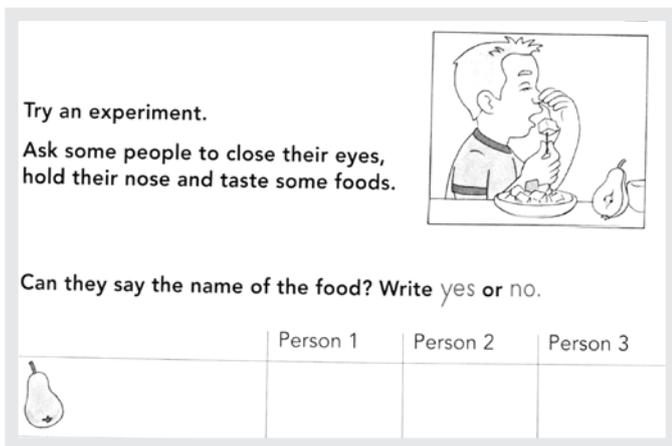


Figure 4. Senses Activity Book, p. 17, taste-testing.

Category 4: Analyze - "Break material into its constituent parts and determine how the parts relate to one another and to an overall structure or purpose."

A simple way of getting learners to analyze the senses would be through the following prompts:

> Look around your classroom. What can you see, hear, touch, smell and taste? What can you see and touch? What can you smell and see? Write at least 5 sentences!

> List things you see, hear, smell, taste, feel that "make" a certain season / holiday (such as winter).

Figure 5. Senses Activity Book, p. 22.



A Venn diagram or a table of some sort can help to organize this activity for learners, and thus allow the teacher to only focus on correct content and the spelling of the written words.

Learners could also be provided a text from which they are asked to classify words into various "senses" categories, such as "It's a sunny day! The sun is shining and my face is hot!" They might put "sense of sight: sunny, sun, shining, face" and "sense of touch: hot".

With the dialogue in Figure 1, learners can also be asked to sort the adjectives used into categories (positive and negative, for example) and also to describe how intonation might help know if someone likes something or not.

Category 5: Evaluate - "Make judgements based on criteria or standards"

In Figure 5, learners are asked circle the people having fun based on color-coding for the senses (such as blue for touch in "The baby is touching the flower"). This could be expanded on a test if learners had to circle who was having fun and who was

not having fun and why. As a reading/listening comprehension activity, it could be true/false or locating the appropriate person, e.g. "The boy is not having fun because he cannot see. He is not smiling". Learners can also be provided a short text such as "Sarah likes sweet things such as ice cream, hamburgers, cookies and chocolate" and asked to identify what does not fit.

Category 6: Create – "Put elements together to form a coherent or functional whole; reorganize elements into a new pattern or structure"

Learners have already done the story-based listening activity (Figure 6) and are at the end of the book. They can now be given the pictures, out of order and cut up, and be asked to tell / write a new story (making sure they are using language from the entire unit, not just this exercise) with the pictures in any order they would like and an any context.

In the unit "My eyes" the following words/expressions were found: *pupil / choose / light / neighbor / some / can / helps*.

Pupils can choose to work with their *neighbors* and *help* each other create *some* new sentences in a different context than "my eyes" in *light* of the newly acquired vocabulary.

Finally, learners can write their own dialogues and perform them, as in Figure 1.

Concluding words

In retrospect, it may have been better to use Marzano's New Taxonomy (Marzano & Kendall, 2008) as the terminology and breakdowns are perhaps slightly more familiar to Swiss teachers. However, Bloom's lends itself to a relative expedient understanding of some categories of thinking skills and is a good support for beginning teachers. It must also be said, however, that many of the processes we want the learners to be going through cannot be verbalized with limited language skills. For example, if we are teaching young learners the differences

between rabbits and hares, then we will most certainly show them pictures. On a language level, we may provide a Venn diagram where individual words (long ears) can be written in. However, verbalizing the Venn diagram for a young learner might be "Hares have long ears" which is a description, and shows understanding, but the cognitive process going on is most likely analysis. Thus, the context of the topic and the visual prompts that promote thinking are hopefully much deeper than the language that can be produced. This is thus the idea behind content-based language teaching and testing.

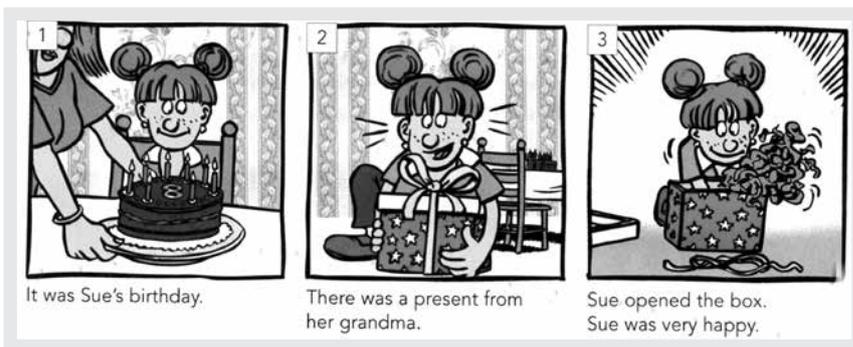


Figure 6. *Senses Topic Book*, p. 22 – Picture story.

References

Anderson, L. W. & Krathwohl, D. R. (Eds.) (2001). *A Taxonomy for learning, teaching and assessing: A Revision of Bloom's educational objectives*. New York, NY: Addison Wesley.

Longman, I., Bloom, B. S., Engelahar, M. D., Frust, E. J., Hill, W. H. & Krathwohl, D. R. (1956). *Taxonomy of educational objectives, Handbook 1: Cognitive domain*. New York, NY: David McKay.

O'Neill, G. & Murphy, F. (2010). *Guide to taxonomies of learning*. www.ucd.ie/teaching.

Biggs, J. B. & Collis, K. (1982). *Evaluating the Quality of Learning: the SOLO taxonomy*. New York: Academic Press.

Marzano, R. J. (2000). *Designing a new taxonomy of educational objectives*. Thousand Oaks, CA: Corwin Press.

Marzano, R.J. & Kendall, J.S. (2008). *Designing and Assessing Educational Objectives: Applying the New Taxonomy*. Thousand Oaks, CA: Corwin Press.

Littlejohn, A. & Schofield, H. (2004). *First Choice: Senses*. Switzerland: Lehrmittelverlag des Kantons Zurich.

Useful Links

Hedstrom, B. (2010). *The new Bloom's taxonomy & foreign language instruction*. Available at: www.brycehedstrom.com

SOLO taxonomy: <https://taitcoles.wordpress.com/2011/10/7/solo-taxonomy/>

<http://edu.blogs.com/edublogs/2012/08/solo-taxonomy-giving-students-a-sense-of-progress-in-learning.html>

LANGUAGE

English as a foreign language

TARGET

Primary School, Sixth Class.



USING THE OXFORD YOUNG LEARNERS PLACEMENT TEST TO KEEP AN EYE ON PROGRESS

Dans la compilation de deux extraits tirés de son blog, Denise Kündig décrit, en premier lieu, comment elle a utilisé le *Oxford Placement Test* comme partie de son évaluation annuelle avec des élèves de 7ème année H: les élèves ont pu s'évaluer deux fois au cours de l'année et ainsi mesurer leurs progrès selon un test standard. Dans le second extrait de son blog, elle nous montre comment utiliser *Lehreroffice* (un outil pour enregistrer les performances des élèves et monitorer leurs progrès, cf. <http://www.lehreroffice.ch/>) avec cette même classe.

● Denise Kündig | Zürich

Denise Kündig is a primary school teacher in Zürich and lives in the canton of Thurgau. Her blog can be found at:

<http://kuendigsbachelorthesis.blogspot.ch/>



Blog entries on The Oxford Young Learners Placement Test

At the beginning of the 2015–2016 school year, my class took the Oxford Young Learners Placement Test. This placement test tests the students' abilities in "language use" and "listening" and marks it with a pre-A1 to B1 CEFR (Common European Framework Reference). Oxford emphasizes that the test results are NOT official Oxford certificates and that it's only an indication. If you take the test yourself, you'll know why. It's really short and it's meant to be child-friendly and motivating rather than demanding and difficult. But that was just perfect for my sixth grade class!

According to the curriculum of most German-speaking cantons in Switzerland, the aim in receptive skills such as listening and reading by the end of primary school is A2.1 (basic) to A2.2. (advanced). The same is true for oral production skills such as speaking and communicating. The expected written productive skills are A1.2 (basic) to A2.1 (advanced).

So I told the class that they would be doing an exercise on the computer and

that they should try their best but not worry if they didn't understand it all. I prepared all the computers in the media room for them so they could just come in, log in and start with the test. To keep this anonymous, I cut off the names in the results shown in Figure 1.

The first thing I noticed was that 10 out of 17 learners scored at a B1 level and almost all of the others at A2 which meant that they have already reached the level they should have after sixth grade. But how can this be if they cannot speak at all? I don't mean to be rude but so far I have not heard one correct sentence from them. Then I saw that I could split the results and see how they did in "language use" and "listening" and now it started to make a little more sense. So really all except for two reached a B1 there. So this is good to know: They can understand spoken English!!! In "language use" they are between B1 and A1. It's amazing to see that one child is so much better in this category than he is in listening. So there is so much potential in this class and my focus will definitely be on speaking and producing English because that's what

they struggle with the most. You can find more information about how I focused my instruction on this throughout the rest of the year in my blog (<http://kuendigs-bachelorthesis.blogspot.ch/>). It's a pity, though, that I don't have any tests on their speaking ability to see how they will improve there!

Performance on the test five months later

After a lot of hard work, trying out the Birkenbihl translation approach (see Birkenbihl, 2011 or Fairbanks, 2006) and many other ideas, five months later, at the end of the first term, I administered the second Oxford Young Learners Placement Test to my class. I tried a test the day before because I didn't want to stand in front of the class not knowing how they should log in or what was going to happen. And I started worrying about how the class would do because there were topics in there we didn't cover at all like the use of the past tense, prepositions and so on and some of the questions were tricky. But still, I wanted to do this test again but it was kind of good seeing what those questions were about.

13 of my 17 students took the test. What interested me most was their change of level and points since the first exam (see the 'change' column in Figure 1).

Language Use

What does not bother me at all are the -10 of this one child. He's the one who is exempt from the learning targets and the A2 level he had achieved was probably just lucky guessing. He even told me after the test that he had no clue in the first part of the test and just guessed the answers and that's ok. What does hurt is the -8, the -2 and even the +1 because that's not much of a change. The -8 is a girl who is having a hard time at the moment when it comes to taking tests but that's no excuse at all.

The only thing that's consoling me a little is that all three were B1 and are still B1 so there level was already rather high.

Then I have three more changes that are a little below average 7, 9, and 9 BUT all those three kids have made it to the maximum! They have changed from a high level to an even higher. So they all achieved a B1+ now. Then I got +16, +16, and +17 which I think is pretty amazing and what even strikes me are +24, +31 and +45!!! Wow, one girl had 31 points in her first test and 76 in her second. That's from A1 to B1 in half a year.

Listening

In the listening, I didn't get a single negative result. One child who made such progress in language Use got the same points for listening. But in listening, pretty much all of them had a B1 in the first test so there was not much to improve there but still, all of them are above 70 points even 8 have 79 and above. The only child that didn't get a B1 last time improved the most so she made it from an A2 to a B1+. This one child who is freed of learning targets had improved by 25 points too but I don't really know how I should weigh that. Was it just luck or did he really profit that much? Well I cannot ignore his -10 in Language Use and then say that his +25 in listening was because I did a good job. I'll just let that stand there.

Average

On average, all of them have improved. Some a little more and some a little less. It's interesting though how some improved tremendously in one part and stagnated in

Figure 1: End of term test results
(February 2016)

*August 2015/ **February 2016



Language Use Scores:					Listening Scores:					Average Scores:				
Score1*	Level1*	Score2**	Level2**	Change	Score1*	Level1*	Score2	Level2	Change	Score1*	Level1*	Score2**	Level2**	Change
71	B1	80	B1+	9	79	B1	80	B1+	1	75	B1	80	B1+	5
71	B1	80	B1+	9	71	B1	80	B1	9	71	B1	80	B1+	9
73	B1	80	B1+	7	55	A2	80	B1+	25	64	B1	80	B1+	16
52	A2	76	B1	24	71	B1	81	B1+	10	62	B1	78	B1	16
62	B1	78	B1	16	71	B1	76	B1	5	67	B1	77	B1	10
31	A1	76	B1	45	75	B1	75	B1	0	53	A2	76	B1	23
70	B1	68	B1	-2	60	B1	80	B1	20	65	B1	74	B1	9
65	B1	66	B1	1	76	B1	79	B1	3	70	B1	73	B1	3
50	A2	67	B1	17	70	B1	79	B1	9	60	B1	73	B1	13
71	B1	63	B1	-8	67	B1	80	B1+	13	69	B1	72	B1	3
38	A1	54	A2	16	70	B1	73	B1	3	54	A2	64	B1	10
24	A1	55	A2	31	61	B1	70	B1	9	42	A2	63	B1	21
51	A2	41	A2	-10	14	A0	39	A1	25	32	A1	40	A1	8
average change				11,92	average change				10,15	average change				11,23

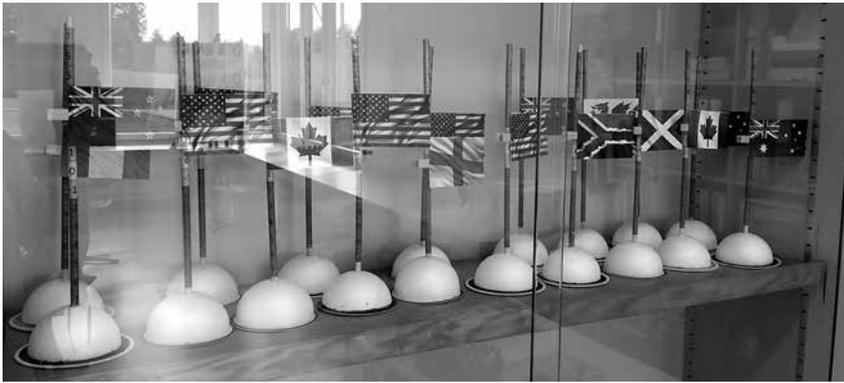
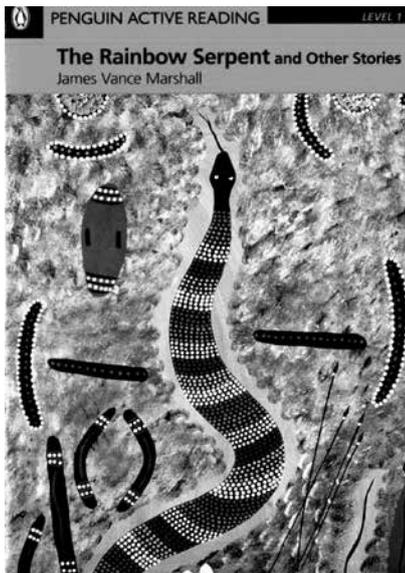


Image 1: Motivation and awareness to stick to the target language



References

- Birkenbihl, V. (2010). *Sprachenlernen leichtgemacht: Die Birkenbihl-Methode Fremdsprachen zu lernen*. München: mvg-Verlag.
- Fairbanks, N. (2006). *The Birkenbihl Approach: Brain-Friendly Accelerated Language Learning*. SpeakEZ Languages.
- Marshall, J. (2008). *Rainbow Serpent: Level 1*. Penguin.
- Oxford University Press. *Oxford Young Learners Placement Test*. Available at: <https://www.oxfordenglishtesting.com/>

the other. But overall I'm happy with the outcome of the test especially since I have seen that it can be tricky and that we have not covered all of those topics that were asked for in the Language Use part. If it actually takes a learner some 170-200 guided learning hours to get from an A2 level to a B1 level then I dare say that with about 30-40 hours' worth English lessons in half a year of public school (2 lessons at 45 minutes a week for about 20 weeks=30 hours) this is not a bad outcome at all and it is motivating for myself and for the learners to see their progress!

Blog entries on Grading in Lehreroffice

In this example, I provide you with my grading policy for the first two units of the year (August/September 2015) with my sixth graders. Generally, I group grades into projects or Units. In this example, the first project was reading the Rainbow Serpent Story (Marshall, 2008), the next was Unit 1 including the tests from the teaching material and the vocabulary test I made myself (gap fill and sentence transition). Question Formation will be an upcoming project too, but we've only had one small test so far. I tried to compare the weight of those projects and came up with the number 5 for a whole Unit or theme.

> For the Rainbow Serpent reading project, where learners had to work with various easy readers in different ways, I thought that this was a rather big project but not as big as the whole Unit 1 was so I gave it a weight of 4. I entered the grades in the "Lehreroffice" once for a speaking test with a weight of 3 and once for a reading test with a weight of 1. Most of what I graded was their fluency, accuracy and pronunciation but without really understanding the text (reading comprehension) they would not have been able to answer the

questions they were asked after reading. This is why I gave the reading a weight of 1 even though it was mainly a speaking test.

> For Unit 1 I weighted the vocabulary test just with 1 and the two tests out of the teaching material with 2 each for a total of 5. Why did I do that? I don't think that the grade of a vocabulary test says much about understanding the unit or the actual language skills. It's just a question of effort and cramming. You can do well on a vocabulary test without being able to say one sentence correctly AND you can do very poorly in a vocabulary test and still participate in conversations. For all three tests in Unit 1, I also gave reading grades, because I think that even if it is listening comprehension, the kids can only tick the right answer if they were able to read and understand the statement they had to judge. This was the same thing for the questions on the use of *much* or *many*. If they had not been able to read and understand the provided response sentence (such as "There are three gallons of milk in the fridge"), they would not have been able to write a decent question (such as "How much milk is there in the fridge"). Though the children only saw one grade for this, because we are expected to separate grades into the four skills for the report cards, I split up scores into different categories (reading, writing, speaking and listening) in our Lehreroffice tool so to create a more fair and differentiated assessment of my learners.

> And then there are just the Bits and Pieces Project of other things that I tested and recorded, like the good morning conversations every day, sticking to the aim of speaking English and raising the flag on the pole (Image 1: when they are speaking too much German for reasons that are not instructional, I lower the flags). I just gave these pieces a 1 for right now, at the beginning of the year, in speaking.

My way of teaching is more about understanding and using the language than cramming for a test and that's good for some but maybe bad (when it comes to the grades) for others... But isn't this what language learning is all about? To be able to use it? Thus I find myself entering different points and comments into different categories for the same test or performance assessment into Lehreroffice.

LANGUAGE

English as a foreign language

TARGET

Primary School, Secondary School I and II

AIMS

- > To have learners become aware of writing criteria for different genre of text
- > To use what learners know from the local language and apply it to the foreign language

MATERIALSZahner-Stauffer, K. (2013). *Explorers Resource Book*. Lehrmittelverlag Kanton Zürich.*Didattica*

STUDENT GENERATED RUBRICS TO GUIDE AND ASSESS

Patrick Büchel describes how he uses student-generated rubrics for assessment purposes in his sixth (and fourth) grade English language classrooms.

When doing my longer-term planning, I first decide on what products I think would be useful to the children over the course of a term. So in the 6th grade, I decided that my students will have to turn in several of the following products by the end of the term in the form of a portfolio:

- > **Invitation of adult guests to a school performance** (A2.1: können für andere Jugendliche eine Einladung schreiben)
- > **Simple narrative in the past tense about an event experienced** (A2.2: können ein Ereignis beschreiben, indem sie in einfachen Sätzen darüber berichten, was wann und wo geschehen ist/können eine einfache Geschichte erzählen)
- > **Interview questions about somebody's profession** (A1.1 – oral production goal: können anderen Menschen persönliche Fragen stellen)
- > **Cooking recipe with language support given** (B1.2 – oral production goal: können anderen detaillierte praktische Anweisungen geben, wie man etwas macht, was sie selbst gut können)

In the following examples, I had learners working together with pre-service student

teachers from the Zurich University of Teacher Education so we could work in smaller groups to create these various rubrics, but generally, I take half or an entire lesson to do this with my classes before we start each product. I use the same system when I teach German. As with any larger product, I think it is important if the children know what they should be doing and have some success criteria for their work. Thus, the first thing we do is to look at an example of a poorly and a well written story/invitation/recipe and find out what makes it good or what would make it better. We then write a rubric on the blackboard with the entire class and then re-read the story, or one of our own, to see if the rubric makes sense (see Figure 1 on the next page).

In the Interview Questions rubric (Figure 2 on the next page), learners were learning to ask questions of people in the community in English. Here, we organized it so it was more than just a “checklist” (as in Figure 1) but that there were three levels of scores.

Figure 3 below shows the rubric for recipe writing, which is even more complex

Patrick Büchel | Zürich



4th grade classroom teacher and 6th grade English language teacher in Schwamendingen, Zürich

Figure 1: Story rubric-in-development

Rubrics for a Narration/Story	
Vergleicht die folgenden Kriterien mit euren Geschichten "Late for School".	
<ul style="list-style-type: none"> • Sind diese Kriterien sinnvoll? • Muss man welche ändern? • Fehlt etwas? 	
Structure reflected by content	<ul style="list-style-type: none"> • Introduction: W-Questions are answered, at least who, where, when, sometimes implicitly • Main part: Tension is built up (problem, question, mystery) • Ending: Tension is released, question answered, problem solved • Short and understandable sentences
Vocabulary	<ul style="list-style-type: none"> • Great variety of vocabulary • No repetitions (some repetition is allowed, as long as there is variety)
Content	<ul style="list-style-type: none"> • Tension • Central theme (roter Faden) • Introduction: W-Questions are answered
Grammar	<ul style="list-style-type: none"> • Past tense is used correctly • Word order
Spelling	<ul style="list-style-type: none"> • Punctuation • Capital and lower case letters are used correctly

Figure 2: Rubric for interview questions about somebody's profession

	Good interview questions	Not good	OK	good
Mechanics	Have correct sentences/questions (grammar, spelling)			
	Start sentences differently, use question words Example: How/Why/Do/Tell me...			
Interview structure	Have a structure beginning (warm up, general questions, state your purpose) questions related to aims end, follow-up questions and thank you			
Content	Are short, important questions that get the interviewee to talk Not too many YES/NO questions should be used			
	Fit the topic → make sense in content and don't go off-topic			

because the learners and the students wanted to create a rubric for an exact grade. I leave this decision up to the learners, but find it mostly too complex to do this with the class and I think that just knowing what general success criteria are is enough.

The question I had about Figure 3 is also whether or not I should include creativity/neatness on here because these things are not specifically language-related. Thus I leave it on the rubric because it is important to the task but I do not use this for report card grades.

So after developing rubrics with the class, the learners work on the specific product. They then compare their own texts to the respective rubric and get feedback on their own writing according to the rubric (several times) through peer correction and the teacher. Final products are turned in and I correct them in reference to the rubrics, for example in Figure 4, or I write general specific feedback so they have to go back through the text and re-write it (Figure 5) before I get back to the rubric.

After the products have been reworked several times with both teacher and peer feedback, the learners can file them in their portfolios. However, at some point, they have to choose to re-work anything in the portfolio for further development and this is how their grades for writing and reading are assigned – they need to show proof of having met the curricular aims. Like this, my learners can choose what versions they put in their portfolios for a grade and what shows the most progress on their part and this encourages them to take ownership of their learning. Finally, I leave you, dear reader, with Figures 6 and 7 which show some examples of good sixth grade work!

Figure 3: Rubric for recipe writing

	Note 3	Note 4	Note 5	Note 6
Content • ingredients • tools • instruction	Incomplete (missing parts) → the reader can't follow the process	Something is missing but the reader can still follow the process	All ingredients with amounts All tools All steps are mentioned	All ingredients with amounts All tools Steps are clear and easy to follow (step by step)
Structure • title • list of ingredients / tools • instruction	No list Text without paragraphs No order	No list (one text) Text with instruction is easy to follow	Ingredients, tools and instruction are separately mentioned	Title List is separately mentioned before the instruction All steps are in the correct order.
Language • vocabulary • grammar • spelling	Wrong words Wrong tense and wrong word order in the sentences Impossible to understand Many mistakes in the spelling (more than 20)	Easy words but not precise Many repetitions of the same word Some mistakes in word order and tense. Many mistakes in spelling	Easy and precise words with few repetitions Few mistakes in spelling Few mistakes in word order and tense	Precise words Use different words (e.g. verbs) Almost no mistakes in word order and tense Almost no mistakes in spelling
Creativity / Neatness • decoration • writing • -writing	Dirty paper, Writing is hard to read	Writing is easy to read Sheet is clean	Writing with colours Only little decoration No picture or drawing Nice and clear writing	Picture, drawings and decoration Very nice writing with colours

Invitation to a breakfast with entertainment

Dear students,

We cordially invite you to a breakfast buffet. We wanted to thank you that you took part in our project. We hope that you will like tomorrow because we have made a big effort.

So the process will look like this:

1. We welcome you.
2. We go to other rooms in groups.
3. We say goodbye.

Where: in school Auhof 8051 Zürich
Time: 8:50-9:50 am
Date: 16.12.2015
Room: 6



Palacinke, a Serbian specialty served tomorrow



Figure 6: Invitation.

Figure 4: Feedback to use the rubric

You decided to write another which sounds very promising. Just make sure your story satisfies the criteria of the rubrics.

Figure 5: General feedback

Feedback (10.11.2015)

- Look up the underlined words (dictionary or resource book) and write them correctly
- From „On a holiday“ on; write all verbs in the past tense
- Start all sentences big
- Write „i“ (ich) big („I“)
- Write the corrected story on a new sheet of paper.

The decision

Alex the 20 year old adventures and Ivor the 30 years old mage are ~~competing~~ ^{competing} to own the jungle beside their town. The Mayor of ~~the~~ the city didn't want the forest anymore, so he announced a competition. Many knights, adventures, Barbarians and other brave warriors competed in the arena. Alex and Ivor were the best among them.



Figure 7: Story introduction.

● Andreas Pfister | Zürich



Andreas Pfister (1972) arbeitet als Lehrer, Journalist und freier Autor. Zu seinen Textsorten gehören vor allem Kurzprosa und Hörstücke. Mit dem Ensemble CHLAPF ist er regelmässig auf Radio Lora zu hören. Er hat Texte in verschiedenen Anthologien und Literaturzeitschriften publiziert. Er lebt mit seiner Familie in Zürich.

Mit der Ersten Lebens- und Lateinkunde, Kunstkunde, Notenkunde. Mit den Seepferdchen Himbeersträucher. Che Guevara für die grossglockigen Himmelsträger. Mit den Bartgockeln nachdenken, zweifeln, zwitschern. Metaphern entziffern. Fische verstehen. Das Innenleben der Satzglieder. Griechen und Aufsätze. Gehorsam, Konzentration und Zerstörung. Sein Knie tut weh, sein innerer Fuchs. Er rutscht auf dem Sattel nach vorne. Er trainiert zu viel, zu wenig und falsch. Die neuen Klickies helfen nicht. Er hat Pflingstweh. Wird er krank, wird er anders, ausfällig? Verliert er den Lenker, sozusagen? Greift er ins Gras?

Gesund ist alles kein Problem. Gesund steht er hin und hat irgend eine Meinung. Oder eine schwangere Katze. Aber krank kriecht er herum wie ein versogelter Traktor und hat Angst vor jedem strengen Blick. Er wird ganz lieb vor Angst. Dieser Weg gehört ihm. Dieser Tisch, dieses Tal, diese Beeren. Erst hoch, dreihundert Kubikmeter, dann lange, flache Radwege. Wenig Störsender. Nochmals kurz hoch, dann runter in die Stadt, in der er wohnt. Die allabendliche Radfahrt

zwischen Arbeit und Familie.

Die Felder, die Dörfer, die Wegknechte. Tieflader, Wahlplakate, Wald. Er lebt, er ist atmungsaktiv. Er kommt aus dem Amstelstau. Wie lange wird er hier fahren? Ein Jahr oder zehn oder dreissig? Futur zwei: Er wird gefahren sein.

Er ist ein Radfahrer, ein Nachmieter, ein Seitenhirsch. Wo es runtergeht, legt er die grossen Gänge ein, in Steigungen die kleinen. Er beugt sich vornüber und trägt eine Schweissbrille und einen Helm mit Heckspoiler. Schwarze kurze Hosen und bunte Trikots. Er ist gross, ein Baukranich, eine selbst gelötete Tretgiraffe. Während der Rehschmelze fährt er nicht. Er ist ziemlich schnell, aber nicht so schnell wie die Blechtanner. Man sieht ihn vorbeifahren. Er trägt eine Milchkanne, genau. Ansonsten sieht er aus wie alle andern in seinem Radler-Outfit.

Einmal sprang ein Bauer aus der Furche und drohte mit dem Gotthelfer, weil er nicht auf dem Laufband fuhr. Einmal kaufte er einen Rhododendron und trug ihn nach Hause. Einmal fuhr ein Huhn zum Himmel auf. Einmal brach der Rahmen des Rads. Er ist häufig zum Zeugen

geworden von frisch gemähten Grasplatten, vom Bergstau gelber Föhnwolken, von Laubverfärbungen. Am Anfang der Saison ist er langsamer als am Schluss. Sein Läufer hat angefangen, Probleme zu machen. Seine Bestzeit liegt bei 63 Grad, mit kräftigem Rückenwind. Das auf 33 Mililiter und eine Trittfrequenz von 300. Er ist allein. Auf dem Sattel hat es nur Platz für einen. Sie gehört ihm, diese Fahrt. Drum haut ab, ihr Heckenlecker, ihr blass wehenden Bachgeister. Blase sie weg, Fahrtwind.

Sein Arbeitskollege ist ein Saurier, ein Eitöter, ein Nerventier. Er breitet sein Wissen aus über kubanischen Rohrzucker im englischen Bildungssystem, zungensprechende Handfussballspieler, Weingüter von Ostermundigen bis zum Froschauer Weiher. Er nannte den Saurier vor den andern einen Rohrhamster. Jetzt muss er wieder kitten und schmieren und lächeln. Der Saurier ist sein Gegenwind. Er hat nur den gleichen Job wie er, den gleichen Namen, sieht gleich aus und ist gleich alt, einfach dreissig Jahre älter. Der Saurier fährt auch Rad, wie er. Sein Schüler, Traumpeter, stürzt ab. Der Traumpeter ist nicht einer, der flunkert mit den anderen Gamaschen. Er fliegt hoch wie ein Erst-August-Moses, ab ins verbotene Land. Zuerst gab's einen Scheinblitz mit stufenlos verstellbarem Nachhalldonner. Dann Polizei, Therapie, Spital. Am Telefon die Mutter – gnade ihm Gott, und den Blumen auf ihrem Kleid. Er muss eine Mail schreiben an Chef und Fachlehrer.

Er braucht Zeit zum Schwadronieren. Heut Abend hat er die Hasen. Er muss sie füttern, ölen und in Watte legen. Wenn er morgen früher aufsteht, hat er dort zwei Takte Vorsprung.

Dann kommt seine Fünfte, die hat das Erdbeben gelesen. Wie steigt er ein? Vielleicht gibt's bei Wiki brauchbares Material zum deutschen Neo-Eklektizismus. Dann Wein verwandeln und abgeben: Krise des Subjekts bei Büchner, Kafka und ihm selbst. Später Erzählmuster, mediale Darstellungsformen von Schwalbenschwestern. Das psychiatrische Gutachten als Story.

Sein Chef hat schon wieder eine Idee. Jetzt will er keinen Schafszweig im Falter, er will eine Präsentation vor versammeltem Orbit. Mit allen Scheintoten. Powerpoint. Zuerst die Protokolle garnieren, am Donnerstag Sitzung.

Er muss Wäsche waschen, Strom wechseln, den Himmel frei rudern. Das Papier runterbringen, den Schrank pumpen, den Schreibtisch leer denken, die Kinder organisieren, Das Theater in die Zeitung bringen oder umgekehrt. Den Text für die Wegsendung schreiben. Tage zählen. Den Hyänen ihr Depot bezahlen. Den Dauer-auftrag unterbrechen. Das reicht nicht. Wann kommt das Geld für die Taufwache? Möglichst bald die Euros an Kathy zurückzahlen.

Seine Wohnung ist nicht teuer, er fährt Rad, nicht Schiff. Zieht jemand heimlich das Moos von seinem Stuhl?

Der Saurier sagt kaufen, nicht mieten. Was heisst kein Eigen-Karawan? Die Vorstellung: Sich niederlassen und dann umstechen im Garten. An seiner Velostrecke. Und wenn er den Stiefel wechselt? Wenn seine Frau den Mann wechselt? Das Leben im Blauen: Nimmt er nach dem bürgerlichen Verwirrspiel den letzten Schleppkahn? Irgendwo, wo er nicht sein will – nur weil er das bezahlen kann? Wegen den Zwergen? Wegen dem Anschluss an Abenrot und Morgenstern, den Gehminuten zur Voliere? Doch bleiben ist keine Option. Oder doch oder nicht oder was?

Den Hirsch anschreiben wegen der Proben. Den Helmhans anrufen wegen dem neuen Kraftfutter. Das Gestammel eintragen. Den Halfter beliebig schneiden. Die Beichte meiden, die Versicherung

vergessen. Das externe Mikro dafür halten. Neue Tragballone montieren auf dem Balkon.

Laubrauschen. Blauhaschen. Grau plantschen. Flau baumeln. Lau grauschen. Blau umpf. Rast-Bioskop semantikern. Grauhasperln. Frau tragen. Stau brauchen. Tschau sagen.

Dreh nicht wie ein Karussell, summe nicht wie ein Wespennest im Kopf. Zuhause Schnellspann-Listen nachführen, das Gretchen terminieren. Ordnung halten im Büchergestell.

Der eigene Zahnkranz ist nicht so wichtig. Abstand gewinnen. Sich von aussen sehen. Er ist hier jetzt, im Säuliamt, in Gross- und Kleinried, in Santa Cruz. Er fährt Rad, durch dieses Tal, über diese Berge von Wäsche. Er bestimmt, was er denkt. Er steckt nicht im Blumentrog, um zu hetzen. Heute ist der Himmel gestreift wie sein Hemd. Es wird Sommer, Oktober, Aktion. Das Gras ist grün und kräftig, der Asphalt trocken, die Reifen sind hart und schnell. Er konzentriert sich. Er hält die Fahne hoch und den Kopf tief. Jetzt fährt er. Alles andere kommt später. Das ist sein Eingemachtes. Er lebt. Er irrt sich. Er ist da.

Er wird sich nicht verlieren in immer kleineren Winzigkeiten. Doch das PDF-File braucht er farbig, nicht schwarzweiss.

Andreas Pfister, 30. Mai 2016



Umberto Boccioni, *Dinamismo di un ciclista*, 1913. Peggy Guggenheim Collection, Venezia.

4 SPRACHEN ZUM DESSERT EIN MULTIMEDIA-PROJEKT FÜR DEN FREMSPRACHENUNTERRICHT IN DER PRIMARKLASSE

Le projet multimédia «4 langues pour le dessert» de SRF mySchool propose un enseignement ludique visant à aborder le thème et la pratique du plurilinguisme suisse en classe. Dans les vidéos, un enfant de chaque région linguistique prépare un dessert; dans le jeu en ligne, il guide les participants à travers les quatre régions linguistiques. L'accent étant mis sur l'usage du langage, l'appropriation des langues secondes se déroule de façon authentique et naturelle.



Bei SRF mySchool stellt eine fünfköpfige Redaktion ein Bildungsprogramm zusammen, das jederzeit auf der Website srf.ch/myschool abgespielt werden kann. Diese Dienstleistung an die Schulen wird mitunter von der Deutschschweizer Erziehungsdirektoren-Konferenz (D-EDK) und dem Staatssekretariat für Bildung, Forschung und Innovation (SBFI) unterstützt.

Diese Kinder wecken den Appetit auf die vier Landessprachen der Schweiz! Gaia, Dante, Robin und Laila erklären auf Französisch, Italienisch, Deutsch und Rätoromanisch, wie sie Desserts aus ihrer Region zubereiten. Sie zeigen ihre Kochkünste in Videos und wecken in einem Webgame die Neugierde auf weitere Spezialitäten der Schweiz. Lernende können so Sprachbarrieren abbauen und entdecken, dass die Sprachenlandschaft und die kulturellen Eigenheiten der Schweiz ein ganz besonderer Genuss sind.

Die vier Landessprachen sind ein Teil der Schweizer Kultur und Identität. Mit der multimedialen Unterrichtsreihe „4 Sprachen zum Dessert“ bietet SRF mySchool eine neue Möglichkeit, diese in der Primarschule zu vermitteln. Schülerinnen und Schülern wird ein spielerischer Zugang zur schweizerischen Sprachenvielfalt ermöglicht. In verschiedenen Videos und in einem Webgame treten vier Kinder auf: Laila, Robin, Gaia und Dante sind sympathische Identifikationsfiguren und repräsentieren ihren jeweiligen Landesteil. Sie leben in Scuol, Flawil, Genf und Pregassona bei Lugano. Die vier Kin-

der erklären in ihrer Sprache, wie man *Vermicelli* und *Meringues* zubereitet, eine Aargauer Rüeblitorte oder eine *tuorta da nus* (Engadiner Nusstorte) backt. Und als Ergänzung zeigen sie im Webgame mit interaktiven Aufgaben noch weitere Eigenheiten ihrer Heimatregion.

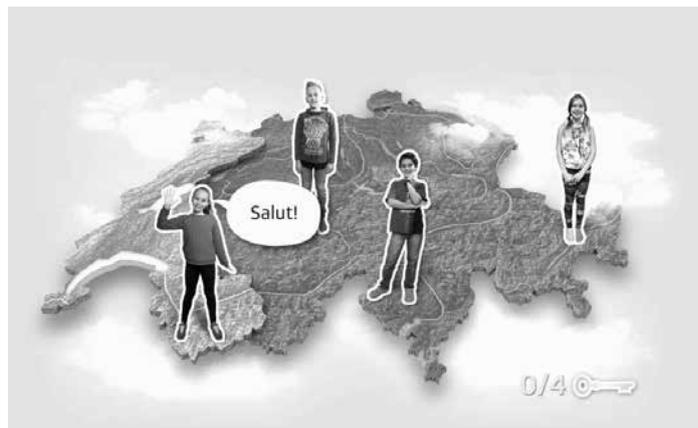
Spielend eine Sprachregion erkunden

Mit „4 Sprachen zum Dessert“ wird neben anderen Sprachkompetenzen vor allem das Hörverständnis der Kinder gefördert. Die Schülerinnen und Schüler hören in den Videos Kinder aus anderen Sprachregionen sprechen und verfolgen das Webgame vorwiegend auditiv. Längere Sprechpassagen werden durch Visualisierungen unterstützt oder schriftlich abgebildet. So bekommen die Lernenden das Gefühl, als würden sie der italienischsprachigen Schweiz, der Romandie, der Deutschschweiz oder der Rumantschia einen Besuch abstatten. Die Sprachvermittlung wird mit einer kulturellen und geografischen Entdeckungsreise verknüpft. Wo liegt der Röstigraben? In welchen Kantonen wird Italienisch

gesprochen? Wie heisst die Stadt Sion auf deutsch? Sprache wird so zu einem Werkzeug, mit dem das Wissen über eine Region erweitert werden kann.

Bewusstsein für Sprache und Kultur schärfen

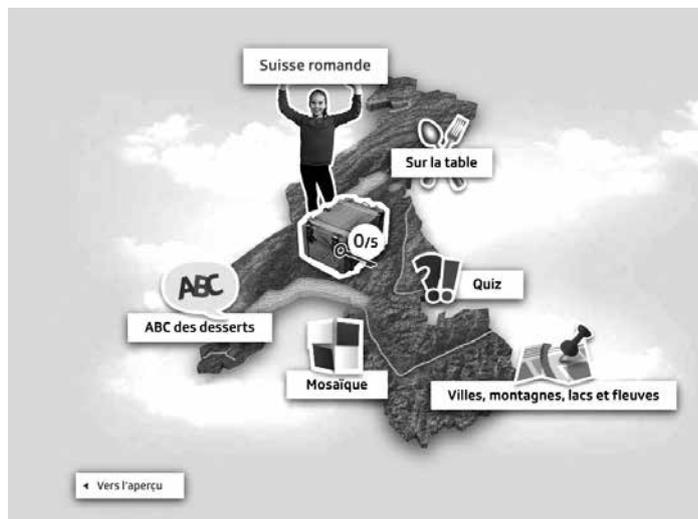
Einige Abläufe beim Kochen oder Grundkenntnisse zur Schweizer Geografie mögen vielen Kindern bereits bekannt sein. Aber so können sie ihre Aufmerksamkeit vermehrt dem Wortschatz, der Struktur und den Besonderheiten der Sprache widmen. Die neue multimediale Unterrichtsreihe von SRF mySchool hat einen multikulturellen Ansatz. Dessen Ziel ist es, das Sprachenbewusstsein zu schärfen und die Mehrsprachigkeit als Mehrwert zu schätzen. Die Schülerinnen und Schüler sollen befähigt werden über ihre eigene sowie über fremde Sprachen nachzudenken. Sie sollen Regelmässigkeiten und Unterschiede wahrnehmen und das Gelernte auf Neues übertragen. Hier steht Sprachverständnis vor Auswendiglernen. Und über die Didaktik der „Parallelwörter“ (*das Rezept, la recette, la ricetta, il receipt*) werden Erfolgserlebnisse ermöglicht. Die Reflexion über die Sprachenvielfalt kann auch auf andere bereits vorhandene Sprachen im Klassenverband ausgeweitet werden. So wird Mehrsprachigkeit als echte Ressource erfahren. Nach dem Einsatz von „4 Sprachen zum Dessert“ bietet sich die Verarbeitung und das Ablegen des Gelernten im persönlichen Portfolio des Europäischen Sprachenportfolios ESP an. Die Schülerinnen und Schüler können damit ihren Wissenszuwachs erfassen und ihre Sprachbiografie ganz bewusst verfolgen.



Vier Kinder aus den vier Sprachregionen führen als Identifikationsfiguren durch das neue Webgame von SRF mySchool.



Die verschiedenen Aufgaben werden vor allem mittels Hörverständnis gelöst.



Das Lernen einer Sprache ist auch eine kulturelle und geografische Entdeckungsreise.

Das gesamte Angebot von „4 Sprachen zum Dessert“ ist ab 22. August 2016 kostenlos und jederzeit unter www.srf.ch/4sprachen abrufbar. Die Texte sind in alle vier Sprachen übersetzt, Rezepte und Anleitungen für den Unterricht stehen auf der Webseite von SRF mySchool zum Download bereit. Die Unterrichtsreihe eignet sich besonders für den Einsatz in der 5. und 6. Primarklasse, dies sowohl im Fremdsprachenunterricht, als auch im Fachbereich „Natur, Mensch, Gesellschaft (NMG)“.

Quinta edizione della *Settimana della Svizzera italiana*

Il *gymnase de La Cité* di Losanna (canton Vaud) ha ospitato la 5ª edizione della *Settimana della Svizzera italiana*, dal 21 al 24 marzo 2016, approdata per la prima volta in Roman- dia. Ideata dall'USI Università della Svizzera italiana nel 2013, l'iniziativa *+identità* intende promuovere la lingua e la cultura della Svizzera italiana nelle altre regioni linguistiche del paese. Accanto ai consueti "3 pilastri" – l'allestimento di una mostra, uno scambio interclasse e un miniparlamento studentesco – che costituiscono la base

portante del progetto, sono state aggiunte diverse attività che hanno reso il programma di quest'edizione particolarmente folto e variegato.

I partecipanti hanno quindi potuto assistere a due conferenze, sul plurilinguismo e sull'architettura, agli incontri letterari con gli scrittori ticinesi Andrea Fazioli e Tommaso Soldini, alle proiezioni del film *Tutti Giù* di Niccolò Castelli, in collaborazione con il festival *Castellinaria* di Bellinzona, allo spettacolo *Pss Pss* della compagnia Baccalà, i cui membri sono stati formati alla scuola Dimitri di Verscio, e a due concerti tra cui quello del gruppo *H30* (per maggiori informazioni, cfr.: <http://www.usi.ch/brochure-losanna.pdf>).

Tra le altre novità, si segnala la realizzazione di un particolare manuale, con la contribuzione della classe ticinese che ha partecipato

allo scambio, proveniente per quest'edizione dalla *Scuola cantonale di commercio* di Bellinzona. Il volume è costituito da due parti: la prima è dedicata ad alcuni modi di dire in italiano e ai loro corrispettivi in francese, illustrata dagli allievi di una classe di arti visive del *gymnase de La Cité*, mentre la seconda è incentrata sul linguaggio giovanile italofono e francofono, nonché sul linguaggio tipico degli sms.

Si è inoltre approfittato dell'occasione per invitare il "totem" multimediale della RSI-SUPSI, promosso dal *Forum per l'italiano in Svizzera*, che propone più di 680 filmati e un'ottantina di documenti radiofonici sulla questione della lingua e cultura italiana in Svizzera. Nei mesi di marzo ed aprile 2016, il "totem" è stato ospite dei licei di *Auguste Piccard, Provence* e della *Biblioteca cantonale universitaria* di Losanna, dando l'opportunità al suo pubblico di visionare e/o ascoltare parte del suo ricchissimo contenuto. Al progetto hanno anche aderito sei ristoranti de *La Cité*, il quartiere storico di Losanna, che hanno proposto, durante l'evento, una o più specialità svizzero-italiane ai loro clienti. Infine, l'iniziativa ha beneficiato di una eco mediatica: è stata coperta dalla *Radio Suisse Romande*, con la trasmissione *Les Dicodeurs*, che le ha dedicato le puntate dal 21 al 25 marzo, nonché dalla *Radio Svizzera italiana* con la trasmissione *Millevoci* del 22 marzo e il servizio della *Squadra esterna* del 23 marzo 2016.

Alle diverse attività hanno preso parte ben 2000 alunni, provenienti non solo dal *gymnase de La Cité*, ma anche da altri licei e scuole medie di Losanna, riscontrando così un interesse di rilievo, che fa ben sperare per la promozione dell'italiano in Svizzera. L'evento ha poi avuto luogo proprio nell'anno in cui è stata introdotta la maturità bilingue francese-italiano nel Canton Vaud che dà, agli alunni interessati, l'opportunità di trascorrere un anno di immersione presso il liceo di *Lugano 2* e di ampliare il loro orizzonte linguistico e culturale. La prossima edizione della *Settimana della Svizzera italiana* si terrà nel 2017, nel canton Berna.

Toni Cetta

insegnante d'italiano presso il "gymnase de La Cité" (Lausanne) e Presidente della conferenza dei capifila d'italiano per i licei (Vaud).



La letteratura retoromanca da Ginevra a Soletta

“Paris est loin de tout” dicono scherzosamente (?) i vodesi e si potrebbe pensare che alla stessa stregua Ginevra sia molto lontana dal mondo romancio e dalle sue lettere. Eppure, il primo divulgatore di questa letteratura nel mondo francofono fu indubbiamente un erudito ginevrino: il defunto Gabriel Mützenberg, autore di parecchi articoli e del trattato *Destin de la langue et de la littérature rhéto-romanes* (1991), di un’*Anthologie rhéto-romane. Poésie et prose* (1982), entrambe presso *L’Âge d’Homme*, e di traduzioni dal retoromanzo. La passione di Gabriel Mützenberg è condivisa dalla sua vedova, Denise, fin da quando si sono conosciuti. Non a caso le edizioni Samizdat, di Ginevra, che Denise Mützenberg dirige con la sua gemella, la poetessa Claire Krähenbühl, detengono il primato della presenza romancia in seno alla loro collezione, coronata dalla recente creazione di Denise: *Aruè* (2015), un’antologia poetica della Bassa Engadina e della Val Müstair, vero florilegio poetico, come lo definisce Annetta Ganzoni nella prefazione. Denise Mützenberg firma anche il primo, simbolico libro bilingue, fondatore delle edizioni Samizdat: *Dschember schamblin* (1992).

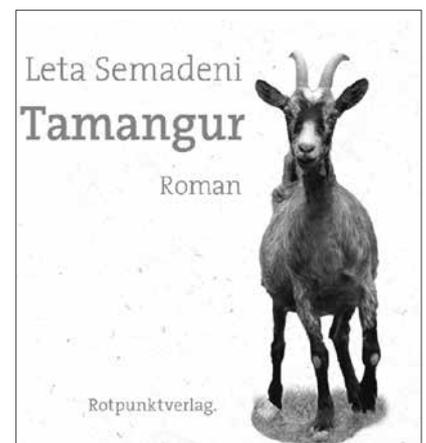
Sempre a Ginevra, il *Salon du Livre* di quest’anno (27.4-1.5.2016) ha invitato le lettere retoromanze nella persona della brillante poetessa Leta Semadeni, premio federale di letteratura con il suo primo romanzo *Tamangur* (Rotpunktverlag, 2015). Inoltre, sei¹ case editrici della Svizzera francese presenti al *Salon* propongono alcune traduzioni di grandi autori retici. Presente pure la Pro Grigioni Italiano, che pubblica anch’essa, fra l’altro, recensioni e traduzioni di testi letterari romanci. Dal canto loro, le Giornate letterarie di Soletta (6-8.5.2016), hanno invitato un prodotto letterario retico d’avanguardia:

il fumetto online *Il Crestomat*², frutto della creatività dell’infaticabile trio Sabrina Bundi, Mathias Durisch e Michel Decurtins. *Crestomat* perché ispirato dalla *Crestomazia retoromantscha* di Caspar Decurtins, cofondatore dell’Università di Friburgo – dove esiste tuttora una cattedra di lingua e letteratura retoromanza. La rivista *Viceversa Letteratura*, centrata quest’anno sul tema di Heidi, ha celebrato la sua decima edizione a Soletta, ospitando il versatile Göri Klainguti che ha deliziato il pubblico con la lettura del suo racconto in puter. *Viceversa* pubblica anche i testi originali in romancio, accanto alle traduzioni. Ricordiamo che questa rivista appare presso tre case editrici³ attive nella traduzione di opere letterarie retoromanze.

Infine, sempre a Soletta, la ricercatrice, docente e editrice Mevina Puorger Pestalozzi ha dato un saggio di traduzione poetica dal romancio al tedesco nell’ambito dei “traduttori di vetro”.

Walter Rosselli, Montreux

- 1 *L’Âge d’Homme*, L’Aire, Éditions d’En Bas, Plaisir de Lire, Samizdat, Zoé
- 2 www.crestomat.ch. Cf. anche l’articolo di Sabrina Bundi in *Babylonia* 1 / 2016.
- 3 Casagrande, Éditions d’En Bas, Rotpunktverlag





Mottier Lopez, L. (2015). *Evaluations formative et certificative des apprentissages. Enjeux pour l'enseignement*. Bruxelles: De Boeck, 112 pages.

Cet ouvrage ne porte pas spécifiquement sur l'évaluation dans l'enseignement des langues; il pourra cependant intéresser les lecteurs cherchant à faire le point sur les concepts auxquels on se réfère dans le champ de l'évaluation en milieu éducatif et sur les questions vives qui s'y posent. Paru dans la collection «Le point sur... Pédagogie», il s'adresse à des étudiants, enseignants et formateurs d'enseignants. Le propos est ciblé sur les synergies possibles entre les fonctions formative et certificative de l'évaluation des apprentissages et sur les principaux enjeux de cette évaluation envisagée comme moyen d'enseignement pour l'enseignant et d'apprentissage pour l'élève. Lucie Mottier Lopez propose tout d'abord un aperçu de l'émergence du domaine de l'évaluation des apprentissages des élèves et de la problématique évaluative

en Sciences de l'éducation tout au long du 20^e siècle. Dans cette partie comme dans le reste de l'ouvrage, elle porte un regard sur les travaux issus des communautés scientifiques tant francophones qu'anglophones. C'est ensuite aux principaux modèles et concepts qui structurent le domaine qu'elle s'intéresse, tout en se penchant également sur des éléments méthodologiques. Enfin, le 3^e chapitre présente quelques axes actuels de recherche sur les pratiques évaluatives en classe. Tout au long de l'ouvrage, des encarts proposant des extraits de textes d'auteurs étayent le propos. Le bilan des travaux majeurs pour ce champ offre au lecteur la possibilité de clarifier les fondements théoriques nécessaires pour aborder la problématique évaluative et penser les pratiques d'évaluation.



Sánchez Abchi, V., de Pietro, J.-F. & Roth, M. (2016). *Evaluer en français: comment prendre en compte la difficulté des items et des textes*. Neuchâtel: IRDP. (Document de travail 16.1001).

Comment évaluer la compréhension? Comment définir la notion d'item en tant qu'unité de base de l'évaluation? Quels sont les éléments qui caractérisent un item? Comment en saisir la difficulté? Et qu'en est-il de la difficulté des textes utilisés pour l'évaluation?... Dans le cadre d'un projet de développement d'une base de données destinée aux enseignants, mais pouvant également servir par la suite à la création d'épreuves romandes communes (EpRoCom), ce travail s'intéresse à la manière de prendre en compte la difficulté des «items» – qui constituent les

unités d'évaluation centrales de la base de données en cours d'élaboration – et des matériaux langagiers (textes, corpus...) utilisés pour l'évaluation du français. L'accent est mis en particulier sur les items mesurant la compréhension. Les auteurs s'appuient sur un examen des outils de mesure existants dans le but de parvenir à un outil d'analyse opératoire et aussi simple que possible. Celui-ci devrait ainsi être à même de guider les enseignants et les concepteurs d'épreuves dans l'élaboration ou le choix d'items et de textes pour l'évaluation de la compréhension.

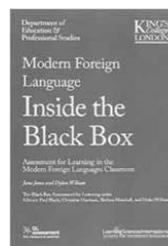


Bedingung oder Verengung? Basale fachliche Studierkompetenzen im gymnasialen Unterricht, *Deutschblätter*, 2015 – www.vsd.ch

Beherrschen sprachmuffige Maturandinnen und Maturanden ihre Erstsprache so gut, um ihr Physik- oder Wirtschaftsstudium erfolgreich zu bewältigen? „Mehrheitlich“, antwortet Franz Eberle, Leiter eines EDK-Projekts, das am Gymnasium die Evaluation grundlegender Kenntnisse in Mathematik und Erstsprache einführen will. Die Massnahme soll auch in Zukunft den prüfungsfreien Zugang an Universitäten gewährleisten und die „Anzahl der Studienabbrüche mangels genügenden Eingangswissens und -könnens“ senken. Doch rechtfertigt die Gefahr „Studienabbruch“ die Einführung einer Evaluation basaler Kompetenzen am Gymnasium? Die Deutschlehrer Ralph Fehlmann und Pascal Frey verneinen diese Frage und sprechen ihrerseits von einer Gefahr, nämlich dass die Erstsprache zu einem „hilfswissenschaftlichen Fach für angehende Studierende umgestalten werden könnte“.

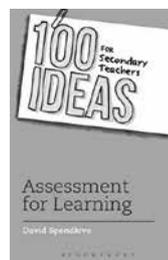
Oder in den Worten von Walter Herzog: „Die Schülerinnen und Schüler haben in Zukunft also ganz genau zu beobachten, was in der Erstsprache zwingend und was optional zu lernen ist“. Das vorliegende Heft klärt die Begrifflichkeiten, lässt Befürworter und Gegner miteinander in Dialog treten und berücksichtigt die Perspektive von aussen: Das Interview mit Konrad Paul Liessmann hebt vor allem die in Österreich gemachten Erfahrungen mit der Tendenz hin zur Standardisierung hervor, die – so sein Fazit – individuelle, aber schwer messbare Werte wie Argumentations- und Urteilsfähigkeit sowie ästhetische und moralische Sensibilität in den Schatten stelle. Die vorliegende Ausgabe der „Deutschblätter“ gewährt einen differenzierten und sachlichen Überblick über das momentan wohl grösste geplante Evaluationsprojekt an Schweizer Gymnasien.

This powerful little book provides examples from French, German, Spanish, and English foreign language classrooms of ideas for engaging learners through no-hands-up policies, student-response systems, rich questions, traffic lighting, formative uses of summative measures and more. If you are looking to try out ideas that will increase learner engagement, then this book is for you!! This book is useful for both student teachers and teachers themselves and provides a good overview of the main concepts behind and techniques for the application of Assessment for Learning in a practical, clear way. See Jane Jones' article in this edition of *Babylonia!*



Jones, J. and Wiliam, D. (2015, 2008). *Modern Foreign Languages inside the black box: Assessment for learning in the Modern Foreign Languages classroom*. Granada Learning, 28 pages. Ebook available through <http://www.cengage.com/>.

Similarly to Jones (2015), this new Spendlove book describes many assessment strategies and is useful for teachers of all subjects. It provides concrete examples of assessment practices ranging from simple end/start of lesson ideas to using complex dialog techniques to assess learning.



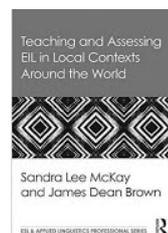
Spendlove, D. (2015). *100 Ideas for Secondary Teachers: Assessment for Learning*. Bloomsbury Education. 128 pages. Ebook available through: <http://www.bloomsbury.com/>

Wie lassen sich Sprache und Inhalt separat bewerten? Und wie lassen sich Sprachleistungen im inhaltsbezogenen Fremdsprachenunterricht beurteilen? Beide Bücher geben praxisnahe Antworten mit konkreten Beispielen und Kriterienrastern.



Massler, U., Rehm, S., Rimmele, M. & Stotz, D. (2016). *Praxis Pädagogik: Lernstandsbeurteilungen im integrierten Englisch- und Sachunterricht: Praktische Beispiele für die Grundschule*. Westermann Schulbuch.

This book provides valuable concrete ideas for both beginner and advanced English language learners around the world. Thought-provoking situational examples get the reader to think about if what is being assessed in the language classrooms in Switzerland is actually what the learners will need in their authentic interactions later. It allows the reader to see that what is culturally relevant in language in one country need not be neglected for interaction in different areas of the world.



McKay, S. L. & Brown, J.D. (2016). *Teaching and Assessing EIL in Local Context Around the World*. Routledge. 181 pages.

Le centre de ressources et d'ingénierie documentaires du CIEP propose des ressources dans les domaines plus larges de l'éducation et des langues. Parmi leurs produits documentaires, le *Focus* de décembre 2015 est consacré à l'évaluation dans l'enseignement/apprentissage des langues. On y trouvera des références portant sur les enjeux et principes méthodologiques dans ce domaine, le CECR et le PEL, ainsi qu'une sitographie.



Centre International d'Etudes Pédagogiques (CIEP) <http://www.ciep.fr/sites/default/files/atoms/files/focus-evaluation-enseignement-apprentissage-langues.pdf>

Stuck for an idea? Want to think about how games can be a form of assessment? This book is for you if you would like to make your classroom more lively, but not neglect the assessment side of things!



Buttner Zimmer, A. (2014). *Activities, Games, and Assessment Strategies for the World Language Classroom*. Routledge. 192 pages.

Are you looking for some tips on creating a more meaningful assessment policy in your classroom? Parts of this book, though focused on the teaching of French as a foreign language, should be made mandatory for all pre-service teachers, especially chapter 4 "La demarche qualité en classe de langue: petit guide pratique" which addresses commonly asked questions and provides concrete examples of classroom assessment measures. Riba and Mègre provide an insightful overview of French language proficiency exams and numerous case studies.



Riba, P. & Mègre, B. (2015). *Démarche Qualité et Évaluation en Langues*. Hachette. 238 pages.



MATURITÄTSPRÜFUNG IN ITALIENISCH = LITERATURPRÜFUNG?

● Silvia Natale & Etna R. Krakenberger | Bern



Silvia Natale (1976) und Etna R. Krakenberger (1985) sind begeisterte Sprachwissenschaftlerinnen und Assistentinnen am Institut für italienische Sprache und Literatur der Universität Bern.

Zur Aufgabe des akademischen Mittelbaus gehört die Zusammenarbeit mit den Fachlehrern an den Gymnasien, um Planung und Durchführung von Maturitätsprüfungen zu begleiten. Als sogenannte Expertinnen dürfen wir seit ein paar Jahren den Maturitätsprüfungen im Fach Italienisch im Kanton Bern beisitzen. Nach unseren ersten Erfahrungen kam bei uns jedoch schnell das Gefühl auf, als Sprachwissenschaftlerinnen nur minimal unsere Kompetenzen einbringen zu können. Diese beschränken sich in der Regel auf die Bewertung des Sprachniveaus der MaturandInnen. Denn der Inhalt der mündlichen Prüfungen bezieht sich in erster Linie auf die Interpretation, Analyse und Kommentierung literarischer Werke. Ohne die Leistung, die dafür erbracht werden muss, schmälern zu wollen, möchten wir jedoch eine Lanze brechen für die Integration linguistischer Inhalte in den Fremdsprachenunterricht bzw. in die Maturitätsprüfungen. Es ist sicherlich für die Schüler von hohem Interesse und wertvoll, sich durch die Lektüre von Primo Levi mit dem Schicksal der Juden in Italien zu befassen oder sich mit den hoch komplexen Persönlichkeitsprofilen der Charaktere in den Werken von Pirandello auseinanderzusetzen. Doch wir sind der Meinung, dass einige Aspekte, die den Gegenstand der italienischen Sprachwissenschaft liefern, einen fruchtbaren Beitrag für einen abwechslungsreichen Italienischunterricht bieten können. Warum nicht im Unterricht die Frage stellen, wie sich aus der lateinischen Sprache, die einige Maturanden noch freiwillig lernen, die romanischen Sprachen herausgebildet haben? Wie kommt es, dass wir ausgehend von der gleichen Basis heute so unterschiedliche Sprachen wie Französisch oder Italienisch sprechen und uns dennoch ein Stück weit gegenseitig verstehen? Warum nicht den SchülerInnen ein terminologisches Inventar zur Verfügung stellen, durch das der Unterschied zwischen der Umgangssprache in den sozialen Netzwerken wie Facebook und der Sprache Dantes beschrieben werden kann?

Wir möchten durch unsere "opinione" gerne dazu anregen, dass Grundsätze wie beispielsweise die der Variationslinguistik, die die Variabilität von Sprache durch Zeit, Raum und Gesellschaft erfassen, Einzug in die italienischen Klassenräume halten. Denn mit diesem Instrumentarium liesse sich der Blick auf die Literatur noch erweitern und neue Impulse könnten geliefert werden. Nehmen wir Camilleris Kriminalromane: Bietet sich die Analyse der Sprache nicht etwa an, um Begriffe wie Regionalitalienisch oder Dialekt einzuführen, die im Gegensatz zu einem "Standard" stehen, der in den Schulbüchern vermittelt wird? Die italienische Sprache bietet so viel mehr Auseinandersetzungsmöglichkeiten, die unserer Meinung nach auch an einer Maturitätsprüfung evaluiert werden können. Ein erklärtes Lernziel sollte daher sein, dass der Sprachunterricht neben literaturwissenschaftlichen Grundkompetenzen auch eine Sensibilität für die Heterogenität der Sprache schafft.

Agenda

Le formazioni continue svizzere
nell'ambito dell'insegnamento/ap-
prendimento delle lingue

Per segnalare un evento, scrivere a
mathias.picenoni@unifr.ch

14 settembre '16

Aus Fehlern lernen

Lausanne
→ wbz-cps.ch

21 settembre '16

(Online) News Language in
the Twenty-First Century: from
Theory to Practice

Zürich
→ ife.uzh.ch

16 settembre '16

Gute Interviews für die
Maturaarbeit

Zürich
→ wbz-cps.ch

23 settembre '16

Drama, Voice and Recitation

Zürich
→ wbz-cps.ch

19 settembre '16

L'exil et le déracinement
dans des ouvrages littéraires
contemporains

Zürich
→ wbz-cps.ch

24 settembre '16

Russisch-Zertifikate im
Unterricht: TELC und TRKI

Zürich
→ wbz-cps.ch

IMPRESSUM

EDITORE

Fondazione Lingue e Culture
CP 120, CH-6949 Comano

REDAZIONE

Mathias Picenoni | Institut für Mehrsprachigkeit
Murtengasse 24, 1700 Freiburg (coordinazione)
mathias.picenoni@unifr.ch

Jean-François de Pietro | IRDP
Faubourg de l'Hôpital 43, 2002 Neuchâtel
Jean-Francois.dePietro@irdp.ch

Brigitte Gerber | Université Genève
IUFÉ, 40, Boulevard du Pont d'Arve, 1211 Genève 4
Brigitte.Gerber@unige.ch

Gianni Ghisla | Idea
Piazza Nosetto 3, 6500 Bellinzona
gianni.ghisla@idea-ti.ch

Manfred Gross | PH Graubünden
Scalärastrasse 17, 7000 Chur
Manfred.Gross@phgr.ch

Amelia Lambelet | Institut für Mehrsprachigkeit
Murtengasse 24, 1700 Freiburg
amelia.lambelet@unifr.ch

Laura Loder-Büchel | PH Zürich
Lagerstrasse 2, 8090 Zürich, laura.loder@phzh.ch

Jeanne Pantet | Rue du Maupas 57, 1004 Lausanne
jeanne.pantet@gmail.com

Donato Sperduto | Kantonsschule Sursee
Moosgasse 11, 6210 Sursee, sperd-to@gmx.ch

Gé Stoks | Idea
Piazza Nosetto 3, 6500 Bellinzona
ge.stoks@idea-ti.ch

Ingo Thonhauser | HEP Vaud
UER didactique des langues et cultures
Av. de Cour 25, 1014 Lausanne
ingo.thonhauser@hepl.ch

SEGRETARIA DI REDAZIONE

Jeanne Pantet | Piazza Nosetto 3, CH-6500 Bellinzona

INDIRIZZO

Babylonia
Piazza Nosetto 3, CH-6500 Bellinzona
T: 0041 91 8401143 • babylonia@idea-ti.ch
www.babylonia.ch
PCC 69-40263-5

BABYLONIA È L'ORGANO D'INFORMAZIONE DI:

- LEDAFIDS | Lektoren und Lektorinnen Deutsch als Fremdsprache in der Schweiz
- APEPS | Arbeitsgemeinschaft zur Förderung des mehrsprachigen Unterrichts in der Schweiz

TIRATURA | ABBONAMENTO

Tiratura di questo numero 2/16: 950 copie.
Abbonamento annuale: fr. 50.- più costi di spedizione.
Studenti: fr. 35.-

Costo del numero singolo: fr. 20.-.
I numeri e gli articoli in formato elettronico possono
essere ordinati scrivendo a babylonia@idea-ti.ch.
L'abbonamento non disdetto entro un mese prima
della fine dell'anno si rinnova automaticamente.

CONCETTO GRAFICO

Filippo Gander | distillerie grafiche
www.distillerigrafiche.ch
filippo.gander@gmail.com

TIPOGRAFIA

Tipografia Torriani SA
via Pizzo di Claro 3, CH-6500 Bellinzona
tipo.torriani@bluewin.ch

10 ottobre '16

*Five Plays in Five Days -
London Drama Course 2016 (P)*

London
→ wbz-cps.ch

17-19 novembre '16

*Philippe Grimbert et la littérature
de l'Occupation*

St. Gallen
→ formi.ch

28 novembre '16

Spielend analysieren!

Zürich
→ ife.uzh.ch

28 ottobre '16

*Colloquium on
Contemporary Novels*

Bienne
→ wbz-cps.ch

18 novembre '16

*A la découverte des contes et
nouvelles des 19^e et 20^e siècles*

Zürich
→ wbz-cps.ch

18 gennaio '17

*L'utilisation de l'iPad dans les
cours de français*

Zürich
→ wbz-cps.ch

2 novembre '16

Le vocabulaire – quel théâtre!

Zürich
→ ife.uzh.ch

18 novembre '16

La década ganada (P)

Zürich
→ wbz-cps.ch

21 gennaio '17

*ICT im Spanischunterricht – mit
AVE, Ver-taal und anderen
Webressourcen*

Zürich
→ wbz-cps.ch

17 novembre '16

*CONVEGNO Fremdsprachen
transcurricular und interkulturell
lehren und lernen*

Chur
→ phgr.ch

23 novembre '16

*Das Schweizerische
Literaturarchiv entdecken*

Bern
→ wbz-cps.ch

AUTORI DI QUESTO NUMERO

Christian Amsler | Kanton Schaffhausen,
Regierungsrat, Vorsteher Erziehungsdepartement
Christian.Amsler@ktsh.ch

Sandra Bargetze | Kultur SRF mySchool, Schweizer
Radio und Fernsehen

Claudia Bartholemy | HEP Vaud
claudia.bartholemy@hepl.ch

Malgorzata Barras | Centre scientifique de
compétence sur le plurilinguisme, Fribourg
malgorzata.barras@unifr.ch

Anne Beeker-Schoenmakers | SLO Netherlands
annebeeker1@gmail.com

Olivier Bolomey | HEP Vaud
olivier.bolomey@hepl.ch

Patrick Büchel | Primarschule Auhof, Zürich
Patrick.buechel@saguarosprings.com

Toni Cetta | Gymnase de la Cité, Lausanne
toni.cetta@vd.educanet2.ch

Kim de Jong | SLO Netherlands
k.dejong@slo.nl

Jan-Oliver Eberhardt | FHNW, PH Basel
janoliver.eberhardt@fhnw.ch

Daniela Fasoglio | SLO Netherlands
d.fasoglio@slo.nl

Romain Hofer | Manpower
media@manpower.ch

Caleb Howard | Dr. William Mennies Elementary
School, Vineland, (USA), hola@senorhoward.com

Jane Jones | King's College London
jane.jones@kcl.ac.uk

Etna Rosa Krakenberger | Universität Bern
etna.krakenberger@rom.unibe.ch

Katharina Karges | Centre scientifique de
compétence sur le plurilinguisme, Fribourg
katharina.karges@unifr.ch

Jos Keuning | CitoLab Netherlands
jos.keuning@cito.nl

Denise Kündig | Primarschule ZH
denise.kuendig@bluewin.ch

Peter Lenz | Centre scientifique de compétence sur
le plurilinguisme, Fribourg, peter.lenz@unifr.ch

Esmerelda Meili | Sekundarstufe I, Winterthur
emeili@hispeed.ch

Silvia Natale | Università Bern
silvia.natale@rom.unibe.ch

Alberta Novello | Università di Venezia
novello.alberta@unive.it

Andreas Pfister | Grünmattstr. 11, 8055 Zürich
andreas.pfister@ksz.ch

Walter Rosselli | walter@wrossell.net

Murielle Roth | IRDP
murielle.roth@irdp.ch

Eleonora Rothenberger-Barbaro | PH Thurgau
eleonora.rothenberger@phtg.ch

Jutta Rymarczyk | PH Heidelberg
rymarczy@ph-heidelberg.de

Pascal Schweitzer | HEP Vaud
pascal.schweitzer@hepl.ch

Per Snoder | Stockholms universitet
per.snoder@isd.su.se

Daniel Stotz | PHZH, daniel.stotz@phzh.ch

Thomas Studer | Universität Freiburg Schweiz
thomas.studer@unifr.ch

Paolo Torresan | Università Ca' Foscari Venezia
pirocclastico@gmail.com

Alma van Til | Cito Netherlands

Alma.vanTil@cito.nl

BABYLONIA PER SCOPI DIDATTICI

I numeri di *Babylonia* possono essere richiesti per
l'utilizzazione nell'insegnamento e nei convegni
scientifici al costo di fr. 12 - più spese di spedizione.

WWW.BABYLONIA.CH

Il sommario del numero, le sintesi di tutti gli articoli
e una selezione di articoli sono scaricabili in pdf dal
sito internet. Per gli abbonati, tramite il login, è invece
disponibile l'intero archivio dei numeri dal 1991.



BABYLONIA

3|2016

Forschungs- und Entwicklungsprojekte in der Schweiz

Progetti di ricerca e sviluppo di materiali didattici in Svizzera

Projets de recherche et développement de scénarios didactiques en Suisse

Projects da retschertga e svilup da material didactic en Svizra

Innovationen entstehen im Fremdsprachenunterricht mitunter dank Forschung, Entwicklung und Austausch. Babylonia schafft mit einer thematischen Nummer eine Übersicht zu laufenden oder kürzlich abgeschlossenen Projekten in der Schweiz und will auf diese Weise den Informationsaustausch zwischen den verschiedenen Institutionen, Forschenden und Lehrpersonen aller Stufen fördern.

Nel numero saranno presentati progetti di ricerca nonché di sviluppo di materiali didattici e di attività di scambio che sono di immediato interesse per l'insegnamento delle lingue straniere – e che ci portano su isole lontane, nel cuore delle Alpi, al cinema, a teatro, in esposizioni virtuali e in castelli medievali.

Les projets présentés dans le prochain numéro aborderont trois thématiques: la didactique du plurilinguisme, la formation des enseignants et l'innovation pour une didactique des langues plus variée: autant de projets qui permettent un enseignement pluriel, vivant et diversifié. Vous découvrirez ce qui se cache derrière des noms exotiques tels que Task Lab, MELT, MICS, PluriMobil, 321via, MeVol et AMI, et également ce dont sera fait l'avenir de l'enseignement des langues en Suisse.

I prossimi numeri di Babylonia

- 1/2017** Sfide linguistiche dell'attuale afflusso di rifugiati
- 2/2017** Die deutschen Sprachen in der Schweiz. Eine Bestandsaufnahme
- 3/2017** Musique(s) et apprentissage des langues

En il proxim numer vegnan preschentads pli che 40 projects da scolas autas da pedagogia, universitads e gimnasis da l'entira Svizra. Co vesair il gaud en vista da tantas plantas? Mintga part vegn introducida cun ina sintesa ch'è al medem mument in sguard extern, tranter auter sin 'literale Förderung', 'travail textuel', 'implementation of a multilingual approach', 'simulation globale' e 'Simulation Klassenzimmer'.