Sprache, die bleibt

Fremdsprachenerwerb mit Film, Musik und Dialog — warum emotionale Verankerung im Gedächtnis besser hält

---

Dirk Streifler

1. Auflage · Berlin, 15. Oktober 2026

Sprache wird nicht durch Regel-Pauken erworben, sondern durch verständlichen, verankerten, gemeinschaftlich erlebten Input — der Film liefert ihn, das Lied verankert ihn, die Gemeinschaft macht ihn möglich; und was davon noch Wette ist, steht hier mit Namen und Messplan.

Zusammenfassung

Dieses Buch begründet das pädagogische Konzept von Pontivo, einer Plattform für Spracherwerb mit Film, Lied und Gemeinschaft. Sein Ausgangspunkt ist ein doppelter Befund: Trotz enormen Aufwands führt regelzentrierter Sprachunterricht verlässlich zu Wissen über die Sprache, aber zu selten zu Können in ihr — in deutschen Integrationskursen wie in italienischen Englischklassen. Die Spracherwerbsforschung benennt seit vier Jahrzehnten die Bedingungen, unter denen Menschen Sprachen tatsächlich erwerben: große Mengen verständlichen, bedeutsamen Inputs bei niedriger Angst.

Aus diesen Bedingungen entwickelt das Buch ein Prüfraster aus drei Fragen — Verständlich? Verankert? Verbunden? — und geht sie entlang der gebauten Plattform durch: den authentischen Film mit synchroner zweisprachiger Untertitelspur und Klick-Erklärung, das Lied als Gedächtnisform für Sprache samt der Bedingungen, unter denen es trägt und unter denen nicht, den Chor und das Klassenzimmer als angstsenkende, zugehörigkeitsstiftende Räume, in denen die Muttersprache Gerüst statt Störfaktor ist. Im übersetzten Heimatlied — dem Format, in dem eine Gruppe das mitgebrachte Lied eines Mitglieds gemeinsam in die neue Sprache überträgt und singt — laufen alle Stränge zusammen. Das Schlusskapitel benennt die Grenzen des Ansatzes und legt eine Messliste vor, an der sich das Konzept prüfen lassen will.

Das Buch unterscheidet durchgehend zwischen forschungsgestützten Prinzipien und den noch ungemessenen Design-Wetten der konkreten Bauform. Sein Maßstab ist der der Schriftenreihe: prüfbar statt behauptet.

Entstanden ist die Idee aus der Not der Geflüchteten; gebaut ist sie für alle, die eine neue Sprache brauchen, um dazuzugehören — die internationale Fachkraft und ihre Familie ebenso wie die vielsprachige Regelklasse. Und sie zielt weiter als auf Vokabeln: Wo eine Gruppe die Lieder ihrer Mitglieder gemeinsam in die neue Sprache übersetzt und singt, wird Integration zum ersten Mal in beide Richtungen hörbar. Was davon trägt, belegt dieses Buch; was Wette bleibt, benennt es — und es lädt Pädagogik, Medien und Kunst ein, die Probe aufs Exempel gemeinsam zu machen.

Inhaltsverzeichnis

Zusammenfassung 1
Teil I — Das Problem 4
Kapitel 1 · Sechshundert Stunden — und kein Wort 5
1.1 Zwei Szenen 5
1.2 Was die Zahlen sagen 5
1.3 Wissen über eine Sprache, Können in einer Sprache 6
1.4 Warum es nicht an den Lehrkräften liegt 7
1.5 Was dieses Buch behauptet — und was nicht 7
1.6 Der Weg durch dieses Buch 8
Kapitel 2 · Die Bedingungen des Erwerbs 9
2.1 Was ein Kind bekommt 9
2.2 Erste Bedingung: verständlicher Input 9
2.3 Die Probe aufs Exempel: die Bücherflut von Fidschi 10
2.4 Zweite Bedingung: ein niedriger Filter 11
2.5 Die dritte Frage: Was bleibt? 12
2.6 Der Sog: Warum Menschen wiederkommen 12
2.7 Vom Prinzip zum Bau 13
Teil II — Der Input: Sehen, hören, lesen 14
Kapitel 3 · Sehen, hören, lesen: der doppelt kodierte Film 15
3.1 Ein europäisches Zufallsexperiment 15
3.2 Zwei Kanäle sind mehr als einer 15
3.3 Die dritte Spur: Untertitel in der Zielsprache 16
3.4 Die Redundanz-Frage — ehrlich ausgetragen 17
3.5 Zwei Sprachen zugleich: die zweisprachige Wette 17
3.6 Was der Film noch nicht kann 18
Kapitel 4 · Das Wort im Moment des Bedürfnisses 19
4.1 Die Stelle, an der man hängen bleibt 19
4.2 Beiläufiges Lernen ist real — und es ist langsam 19
4.3 Was die Glosse ändert 20
4.4 Der Klick — und was an ihm Wette ist 20
4.5 Verstanden ist noch nicht behalten 21
Teil III — Die Melodie 22
Kapitel 5 · Warum das Lied im Kopf bleibt 23
5.1 Das Gedächtnis, das niemand bestellt hat 23
5.2 Was das Labor sagt: Die Melodie zerlegt den Strom 23
5.3 Ehrlich: Wann das Lied nicht hilft 24
5.4 Vom Labor zum Liederbuch 25
Teil IV — Die Gemeinschaft 26

Kapitel 6 · Der Chor: Zugehörigkeit senkt die Hürde 27
6.1 Neunzig Minuten 27
6.2 Was das gemeinsame Singen mit Menschen macht 27
6.3 Die geschützte Stimme 28
6.4 Für die, die ankommen müssen 29
6.5 Vom Befund zum Bau 30
Kapitel 7 · Das Klassenzimmer: geteilte Aufmerksamkeit, legitimierte Muttersprache 31
7.1 Das Verbot 31
7.2 Das Gerüst 31
7.3 Die Muttersprache als Gerüst 32
7.4 Das gebaute Klassenzimmer 33
7.5 Was dem Klassenzimmer noch fehlt 34
Teil V — Die Probe 35
Kapitel 8 · Coro One Sky: Das Lied aus der Heimat, gesungen in der neuen Sprache 36
8.1 Die Übung 36
8.2 Fünf Mechanismen 36
8.3 Die Erzählung: Warum das Ganze eine Geschichte braucht 38
8.4 Die Vorsichtsregel — und die eine große Wette 39
Teil VI — Die Szenarien & der offene Weg 41
Kapitel 9 · Drei Klassenzimmer und eine ehrliche Agenda 42
9.1 Drei Klassenzimmer 42
9.2 Amal, ein Jahr später — eine Hypothese 42
9.3 Was der Ansatz nicht kann 43
9.4 Das Rechte-Paradox und die Lieder-Spende 43
9.5 Die Messliste 44
9.6 Der Stammbaum 45
Quellenregister 47
Abkürzungen 51
Glossar 52
Zum Konzept & Team 54

Teil I — Das Problem

„… a gifted person ought to learn English (barring spelling and pronouncing), in 30 hours, French in 30 days, and German in 30 years."

— Mark Twain, A Tramp Abroad, Anhang D (1880)

Kapitel 1 · Sechshundert Stunden — und kein Wort

1.1 Zwei Szenen

Amal ist vierunddreißig, sie kam vor drei Jahren aus Syrien nach Deutschland. Sie hat getan, was man von ihr verlangt hat, und mehr als das: Sie hat den allgemeinen Integrationskurs besucht — sechshundert Unterrichtsstunden Sprachkurs, dazu hundert Stunden Orientierungskurs —, sie hat Vokabelhefte gefüllt, Arbeitsblätter bearbeitet, Modelltests geübt. Am Ende hat sie den „Deutsch-Test für Zuwanderer" bestanden, mit dem Ergebnis B1. Sie gehört damit zur besseren Hälfte ihres Jahrgangs. Das Zertifikat liegt zu Hause in einer Klarsichtfolie.

An der Supermarktkasse fragt die Kassiererin, schnell und beiläufig, über die Schulter: „Sammeln Se Punkte?" Amal versteht die Wörter nicht. Sie versteht nicht einmal, dass es eine Frage war. Sie spürt die Schlange hinter sich, das Warten, die Blicke. Sie lächelt, nickt, bezahlt und sagt — nichts. Auf dem Heimweg übersetzt sie sich den Satz, den sie nicht gesagt hat. Zu Hause könnte sie fehlerfrei erklären, wann der Dativ steht.

Die zweite Szene spielt tausend Kilometer südlich, in einer Oberstufenklasse in Bologna. Englisch steht hier seit der Grundschule auf dem Stundenplan; die meisten in diesem Raum haben mehr als zehn Jahre Unterricht hinter sich, Grammatikpensen, Vokabeltests, Klassenarbeiten. Die Lehrerin startet eine Dokumentation im englischen Original. Nach zwei Minuten hebt jemand die Hand und fragt, ob es die Fassung nicht auch auf Italienisch gebe. Es ist keine Provokation. Es ist eine ehrliche Auskunft: Der Ton erreicht sie nicht.

Beide Szenen sind erfunden. Die Zahlen, die hinter ihnen stehen, sind es nicht.

1.2 Was die Zahlen sagen

Der deutsche Integrationskurs ist kein Nebenangebot, sondern das zentrale sprachpolitische Instrument des Landes: siebenhundert Unterrichtsstunden, gesetzlich geregelt, mit dem erklärten Ziel, die Teilnehmenden im Abschlusstest auf das Niveau B1 zu führen.[1] Im Jahr 2024 erreichten von den Teilnehmenden des Abschlusstests aus allgemeinen Integrationskursen 58,6 Prozent dieses B1; 31,8 Prozent blieben bei A2, 9,6 Prozent darunter.[2] Man muss diese Zahl zweimal lesen. Sie zählt nur diejenigen, die überhaupt bis zum Test gekommen sind — und selbst von ihnen verfehlt vier von zehn das Kursziel.

Und die Zahl sagt noch etwas Zweites, Leiseres. Denn B1 ist kein hohes Niveau. Die europäische Referenzskala beschreibt es als Schwellenniveau: Wer B1 hat, kann sich über vertraute Dinge in einfachen, zusammenhängenden Sätzen äußern und versteht die Hauptpunkte, wenn klare

---

[1] § 10 Abs. 1 Satz 1 der Integrationskursverordnung (IntV): „Der Integrationskurs umfasst 700 Unterrichtsstunden" (gesetze-im-internet.de/intv/__10.html, abgerufen 2026-07-24); davon entfallen auf den Sprachkurs 600 Unterrichtsstunden, § 11 Abs. 1 Satz 1 IntV (gesetze-im-internet.de/intv/__11.html). Das Kursziel B1 formuliert das Bundesamt für Migration und Flüchtlinge (BAMF) in seiner Geschäftsstatistik ausdrücklich: Ziel sei, „dass die Teilnehmenden im Abschlusstest DTZ das Niveau B1 des Gemeinsamen Europäischen Referenzrahmens für Sprachen (GER) erreichen" (Bericht zur Integrationskursgeschäftsstatistik für das Jahr 2024, Anmerkungen zu den DTZ-Tabellen, S. 17 ff.).

[2] BAMF, *Bericht zur Integrationskursgeschäftsstatistik für das Jahr 2024*, S. 19 (Tabelle „Sprachniveau ‚Deutsch-Test für Zuwanderer' im Jahr 2024 nach Kursarten und Ergebnis"): Allgemeine Integrationskurse — 160.804 Teilnehmende (58,6 %) mit B1, 87.449 (31,8 %) mit A2, 26.314 (9,6 %) unter A2, von insgesamt 274.567 DTZ-Teilnehmenden (PDF, abgerufen 2026-07-24; Fundstellen am PDF-Volltext verifiziert). Zwei Lesehinweise: Gezählt wird je Person das höchste erreichte Ergebnis (Anm. 1 der Tabelle); und die Anteile beziehen sich allein auf Testteilnehmende — wer den Kurs abbricht oder nicht zum Test antritt, erscheint in dieser Quote nicht.

Standardsprache verwendet wird.[3] Die Kassiererin verwendet keine klare Standardsprache. Der Alltag tut es fast nie. Das heißt: Selbst das *Erreichen* des Kursziels garantiert nicht die Sprache, die man an der Kasse, beim Arzt, im Elterngespräch braucht. Amal ist kein Ausnahmefall des Systems. Sie ist sein Erfolgsfall.

Die italienische Szene hat ihre eigene Statistik. Italien testet seine Abschlussjahrgänge flächendeckend; im Jahr 2025 erreichten im Fach Englisch 53,3 Prozent der Schülerinnen und Schüler des letzten Jahrgangs das vorgesehene Niveau im Leseverstehen — im Hörverstehen waren es 43 Prozent, im Süden des Landes deutlich unter drei von zehn.[4] Mehr als ein Jahrzehnt Unterricht, und weniger als die Hälfte versteht die Sprache, wenn sie gesprochen wird. Wohlgemerkt: *versteht*. Vom Sprechen ist noch gar nicht die Rede.

Zwei Länder, zwei Systeme, zwei sehr verschiedene Lernergruppen — und dasselbe Muster: Der Aufwand ist enorm, der Ertrag an lebendiger Sprache ist es nicht. Wer nur eine der beiden Szenen kennte, könnte an nationale Besonderheiten glauben, an Kursqualität, an Motivation. Die Doppelung legt einen anderen Verdacht nahe: Es liegt nicht an den Ländern. Es liegt an der Art, wie wir glauben, dass Sprachen in den Kopf kommen.

1.3 Wissen über eine Sprache, Können in einer Sprache

Der Verdacht hat einen klassischen Namen. Stephen Krashen hat Anfang der 1980er Jahre die Unterscheidung, um die es hier geht, ins Zentrum der Spracherwerbsforschung gestellt: Erwachsene verfügen über „zwei getrennte und unabhängige Wege, Kompetenz in einer zweiten Sprache zu entwickeln".[5] Der eine Weg ist das *Lernen* — bewusstes Wissen über die Sprache: Regeln kennen, Regeln benennen, über Regeln sprechen können. Der andere ist der *Erwerb* — ein unbewusster Prozess, derselbe, mit dem Kinder ihre Muttersprache aufbauen: Man merkt nicht, dass er stattfindet, und sein Ergebnis ist kein Regelwissen, sondern ein Sprachgefühl. Ein Satz „klingt richtig" oder „klingt falsch", ohne dass man sagen könnte, warum.

Diese Unterscheidung erklärt beide Szenen mit einem Schlag. Amal hat gelernt, was das Kurssystem zu lehren versprach: Wissen über das Deutsche. Der Dativ, den sie erklären kann, ist echtes Wissen — aber es ist Wissen der falschen Sorte für den Moment an der Kasse. Dort ist keine Zeit, Regeln anzuwenden; dort müsste der Satz von selbst kommen, aus dem Gefühl. Und die Klasse in Bologna hat elf Jahre lang Englisch *gelernt* und dabei erstaunlich wenig Englisch *erworben* — weil der Unterricht ihr vor allem Sätze über die Sprache gezeigt hat und viel zu selten die Sprache selbst, in Mengen, in denen ein Gefühl hätte entstehen können.

---

[3] Europarat, *Gemeinsamer europäischer Referenzrahmen für Sprachen* (2001), Globalskala, Niveau B1 („Threshold"/Schwellenniveau); Formulierung im Text paraphrasiert nach der Globalskala und dem Selbstbeurteilungsraster (deutsche Fassung am Druck zu prüfen; Raster online: europass.europa.eu, abgerufen 2026-07-24).

[4] INVALSI, Ergebnisdarstellung „L'inglese nella Prova INVALSI 2025 tra successi e fragilità" (invalsi.it, abgerufen 2026-07-24): Im letzten Jahrgang der Sekundarstufe (grado 13) erreichen 53,3 % das vorgesehene Niveau im Reading und 43 % im Listening; im Süden („Sud" 28,6 %, „Isole" 27,3 %) liegt das Hörverstehen dramatisch unter dem Nordosten (über 60 %). Grundlage ist der Rapporto Nazionale INVALSI 2025 (am Volltext zu prüfen). Die Zielniveaus der Prüfung sind B2 (Gymnasien/Fachoberschulen) bzw. B1+ (berufsbildende Züge).

[5] Krashen, S., *Principles and Practice in Second Language Acquisition*, Pergamon 1982, hier zitiert nach der vom Autor freigegebenen Internet-Ausgabe (2009), S. 10: „adults have two distinct and independent ways of developing competence in a second language" (PDF auf sdkrashen.com, abgerufen 2026-07-24; Seitenangaben nach der Paginierung der Internet-Ausgabe, am Erstdruck zu prüfen). Dort auch die Charakterisierung des Erwerbs als unbewusster Prozess mit „Gefühl für Richtigkeit" (S. 10 f.).

Man muss dazu sagen: Die scharfe Trennung der beiden Wege ist in der Forschung umstritten geblieben — ob Gelerntes wirklich nie zu Erworbenem werden kann, wie Krashen es behauptete, bezweifeln viele.[6] Für dieses Buch ist der Streit weniger wichtig, als er scheint. Denn unstrittig ist das Phänomen, auf das die Unterscheidung zeigt: Regelwissen und flüssiges Können sind zweierlei; das eine stellt sich durch Erklären und Üben von Regeln ein, das andere nachweislich nicht in ausreichendem Maß. Vier Jahrzehnte Forschung haben an diesem Kernbefund wenig geändert.[7]

1.4 Warum es nicht an den Lehrkräften liegt

Es wäre bequem, an dieser Stelle ein Versagen zu diagnostizieren: schlechte Kurse, unmotivierte Lehrer, faule Schüler. Dieses Buch tut das nicht, und zwar nicht aus Höflichkeit. Die Lehrkräfte in Integrationskursen und Schulklassen arbeiten in einem System, das ihnen die Bedingungen diktiert: Es gibt ein Curriculum, das in Prüfungen mündet, und Prüfungen fragen bevorzugt ab, was sich abfragen lässt — Regelwissen, Wortlisten, geübte Formate. Es gibt Klassen von zwanzig und mehr Personen, in denen der einzelne Lernende pro Stunde auf wenige Sprechminuten kommt. Und es gibt die Angst: vor der Blamage beim Fehler, vor der Prüfung, vor dem Urteil der anderen — eine Angst, die gerade bei Erwachsenen, die im Zielland um ihren Platz kämpfen, jede Unterrichtsstunde begleitet.

Unter diesen Bedingungen ist regelzentrierter Unterricht keine Dummheit, sondern eine rationale Anpassung: Er ist planbar, prüfbar, abrechenbar. Nur erzeugt er eben das, was er prüft — und nicht das, was Amal an der Kasse gebraucht hätte. Wenn man das System nicht anklagen, sondern verstehen will, muss man die Frage anders stellen: Nicht „Wer macht es falsch?", sondern „Unter welchen Bedingungen entsteht Sprache tatsächlich — und wie stellt man diese Bedingungen her, auch außerhalb des Klassenzimmers, auch ohne dass jemand dafür scheitern muss?"

1.5 Was dieses Buch behauptet — und was nicht

Die These dieses Buches lässt sich in drei Sätzen sagen. Sprache wird erworben, nicht gepaukt: Sie entsteht aus *verständlichem* Input — aus Sprache, die man versteht, weil Bild, Zusammenhang und Vorwissen beim Verstehen helfen. Sie bleibt, wenn dieser Input *verankert* wird — emotional, rhythmisch, melodisch, so, wie ein Lied im Kopf bleibt, das man nie zu behalten versucht hat. Und sie entfaltet sich, wo Menschen *verbunden* sind — wo eine Gemeinschaft die Angst senkt und das Sprechen und Singen zu etwas macht, das man miteinander tut statt voreinander. Verständlich, verankert, verbunden: An diesen drei Bedingungen wird sich alles messen lassen, was die folgenden Kapitel vorschlagen.

---

[6] Die Kritik setzt an der Nicht-Falsifizierbarkeit der strikten Trennung an und an Krashens Zusatzthese, Gelerntes könne nie in Erworbenes übergehen: McLaughlin, B., *Theories of Second-Language Learning*, London 1987, Kap. 2 (am Druck zu prüfen); aus Sicht der Fertigkeitserwerbstheorie DeKeyser, R., Skill Acquisition Theory, in: VanPatten/Williams (Hg.), *Theories in Second Language Acquisition*, 2. Aufl., New York 2015, Kap. 5 (am Druck zu prüfen). Die moderatere, heute breit geteilte Fassung unterscheidet implizites und explizites Wissen und hält deren Verhältnis für begrenzt durchlässig („Interface-Debatte"): Ellis, R., *The Study of Second Language Acquisition*, 2. Aufl., Oxford 2008, Kap. 9 (am Druck zu prüfen).

[7] Forschungsüberblick: Lightbown, P./Spada, N., *How Languages are Learned*, 4. Aufl., Oxford 2013, insb. Kap. 6 zu den Grenzen expliziter Grammatikvermittlung (am Druck zu prüfen). Dass explizite Instruktion messbare, aber vor allem auf explizites Wissen beschränkte Effekte hat, zeigen auch die Meta-Analysen von Norris/Ortega 2000 (*Language Learning* 50) — mit der dort selbst benannten Einschränkung, dass die Messinstrumente überwiegend regelnahes Wissen abfragen (ebd., am Volltext zu prüfen).

Pontivo — die Plattform, deren pädagogisches Konzept dieses Buch begründet — ist der Versuch, diese drei Bedingungen in eine gebaute Form zu bringen: authentischer Film mit zweisprachigen, synchronen Untertiteln und einer Klick-Erklärung für die Stelle, die man gerade nicht verstanden hat; ein Liederbuch für das gemeinsame Singen in Gruppen; ein Klassenzimmer, in dem eine Lehrkraft führt und alle folgen können; eine Erzählung, die das Ganze zusammenhält.[8]

Ebenso wichtig ist, was dieses Buch *nicht* behauptet. Es behauptet nicht, dass Pontivo bewiesen hätte, was es baut. Die Prinzipien, auf denen die Plattform steht, sind durch Jahrzehnte von Forschung gestützt; das wird Kapitel für Kapitel mit Fundstellen belegt. Die konkrete Bauform aber — dieser Player, dieses Liederbuch, dieser Comic, in dieser Kombination — ist eine Reihe von Design-Wetten, die empirisch noch nicht gemessen sind. Das Buch wird beide Ebenen konsequent auseinanderhalten und im letzten Kapitel eine ehrliche Liste dessen aufstellen, was zu messen wäre. Der Maßstab, dem sich diese Schriftenreihe verpflichtet hat, gilt auch hier: prüfbar statt behauptet — nach außen wie an uns selbst.

1.6 Der Weg durch dieses Buch

Der Weg ist in sechs Teile gegliedert. Teil I benennt das Problem und legt in Kapitel 2 das Fundament: die Bedingungen, unter denen Menschen Sprachen tatsächlich erwerben — dort werden die drei Fragen „Verständlich? Verankert? Verbunden?" ihren wissenschaftlichen Boden bekommen. Teil II wendet sich dem Input zu: dem Film mit seinen zwei Untertitelspuren und der Frage, warum bewegtes Bild mit Text verständlicher Input von seltener Dichte ist — und wo die Grenzen dieser Verheißung liegen. Teil III gehört der Melodie: dem Befund, dass gesungene Sprache anders haftet als gesprochene, und den Bedingungen, unter denen das gilt. Teil IV betritt den sozialen Raum: den Chor, der nachweislich schneller verbindet als fast jede andere gemeinsame Tätigkeit, und das Klassenzimmer, in dem die Muttersprache vom Störfaktor zum Gerüst wird. Teil V führt die drei Stränge in einer einzigen Übung zusammen: dem Lied aus der Heimat, das eine Gruppe gemeinsam in die neue Sprache übersetzt und dann gemeinsam singt — dort wird sich zeigen, warum dieses unscheinbare Format die dichteste Stelle des ganzen Konzepts ist, und welche Rolle die Erzählung „Coro One Sky" dabei spielt. Teil VI schließlich kehrt zu Amal und zur Klasse in Bologna zurück: Es beschreibt die drei konkreten Lern-Szenarien, für die Pontivo gebaut ist, zieht die Grenzen des Ansatzes — und legt die Messlatte fest, an der er sich wird messen lassen müssen.

Am Anfang stand eine Frau, die schweigend bezahlt hat, obwohl sie ein Zertifikat besitzt. Am Ende dieses Buches soll nachvollziehbar sein, warum ihr sechshundert Stunden Unterricht dieses Schweigen nicht nehmen konnten — und was es stattdessen braucht. Das nächste Kapitel beginnt dort, wo jeder Spracherwerb beginnt: bei der Frage, was im Kopf geschieht, wenn ein Mensch eine Sprache nicht lernt, sondern bekommt.

[8] Funktions-Hinweis (Produktbeschreibung, kein Forschungsbefund): Die in diesem Buch beschriebenen Bauformen sind — soweit nicht ausdrücklich als Bauabsicht gekennzeichnet — auf der Plattform in Betrieb (pontivo.de, Stand 15. Oktober 2026). Wo ein Kapitel eine gebaute Funktion behandelt, verweist eine Fußnote dieser Art auf die laufende Umsetzung.

Kapitel 2 · Die Bedingungen des Erwerbs

2.1 Was ein Kind bekommt

Niemand hat je einem Dreijährigen den Dativ erklärt. Trotzdem steht er eines Tages in der Küche und sagt einen Satz mit Dativ — nicht, weil er eine Regel angewandt hätte, sondern weil der Satz für ihn *richtig klingt*. Das Kind hat die Sprache nicht gelernt wie ein Schulfach. Es hat sie bekommen: aus Tausenden von Stunden Sprache, die an es gerichtet war, eingebettet in Situationen, die es verstand — beim Essen, beim Anziehen, beim Vorlesen, im Streit um das letzte Gummibärchen. Verstehen kam zuerst; die Form kam hinterher, von selbst.

Man muss mit dieser Beobachtung vorsichtig sein. Erwachsene sind keine großen Kinder; sie haben weniger Zeit, andere Gehirne, ein fertiges Selbstbild und eine erste Sprache, die immer mitspricht.[9] Aber die Spracherwerbsforschung der letzten vier Jahrzehnte hat eine Kernaussage erstaunlich stabil gehalten: Der Mechanismus, der beim Kind die Sprache aufbaut, steht auch dem Erwachsenen noch zur Verfügung — wenn man ihm gibt, was das Kind bekommt. Die Frage dieses Kapitels ist deshalb nicht, ob Erwachsene Sprachen erwerben können, sondern unter welchen Bedingungen. Zwei davon lassen sich benennen, belegen — und bauen.

2.2 Erste Bedingung: verständlicher Input

Die erste Bedingung hat Stephen Krashen in eine Formel gefasst, die zum bekanntesten Kürzel der Disziplin geworden ist: *i + 1*. Wenn *i* der gegenwärtige Stand eines Lerners ist, dann geschieht Erwerb, wenn er Sprache *versteht*, die ein Stück darüber liegt. Krashens eigene Formulierung ist präziser, als ihre populäre Kurzform ahnen lässt: Eine „notwendige (aber nicht hinreichende) Bedingung" für den Schritt von *i* nach *i + 1* sei, „dass der Erwerbende Input versteht, der *i + 1* enthält — wobei ‚verstehen' heißt, dass er auf die Bedeutung und nicht auf die Form der Botschaft gerichtet ist".[10] Zwei Dinge stecken in diesem Satz. Erstens: Es genügt nicht, Sprache zu *hören*; sie muss verstanden werden. Zweitens: Das Verstehen gelingt gerade nicht durch Analyse der Form, sondern an ihr vorbei — mit Hilfe von allem, was außer der Sprache noch da ist: der Situation, dem Bild, dem Weltwissen, dem Gesicht des Sprechenden. Genau deshalb kann man Sprache verstehen, die man noch nicht *kann*. Und genau in diesem Moment, sagt die Hypothese, wächst das Können.

Daraus folgt ein dritter, oft übersehener Punkt, der für alles Weitere in diesem Buch wichtig ist: Man muss *i + 1* nicht dosieren. Krashen selbst hat betont, dass der Input nicht gezielt auf die „nächste Struktur" zugeschnitten werden muss — wenn genug verständlicher Input da ist, ist *i + 1* automatisch enthalten.[11] Der Lehrplan, der heute den Dativ und nächste Woche das Perfekt

---

[9] Der Vergleich mit dem Erstspracherwerb ist eine Analogie mit klaren Grenzen: Erwachsene Lerner unterscheiden sich in Zeitbudget, Vorwissen, Alphabetisierung und im Einfluss der Erstsprache; native-ähnliche Endniveaus sind im Erwachsenenerwerb die Ausnahme. Krashen selbst beansprucht nur, dass Erwachsene *denselben Erwerbsmechanismus* weiter nutzen können, nicht, dass sie ihn gleich weit tragen: Krashen, S., *Principles and Practice in Second Language Acquisition*, Pergamon 1982 (Internet-Ausgabe 2009), S. 10 (PDF, abgerufen 2026-07-24).

[10] Krashen 1982 (Fn. 1), S. 20 f.: „a necessary (but not sufficient) condition to move from stage i to stage i + 1 is that the acquirer understand input that contains i + 1, where ‚understand' means that the acquirer is focussed on the meaning and not the form of the message" (Übersetzung im Text von uns; Seitenangaben nach der Internet-Ausgabe).

[11] Krashen 1982 (Fn. 1), S. 20 f.: Input müsse *i + 1* nicht gezielt ansteuern; „if the acquirer understands the input, and there is enough of it, i + 1 will automatically be provided".

„dran nimmt", versucht etwas zu steuern, was sich bei ausreichender Menge von selbst einstellt. Die praktische Folge ist eine Umkehrung der üblichen Knappheitslogik: Nicht die *richtige Portion* Sprache ist das Ziel, sondern die *große Menge* verständlicher Sprache.

Man muss auch hier ehrlich bleiben: Dass verständlicher Input *notwendig* ist, bestreitet in der Forschung niemand; ob er *hinreichend* ist, sehr wohl. Die Interaktionsforschung hat gezeigt, dass das Aushandeln von Bedeutung im Gespräch den Input erst richtig wirksam macht, und die Output-Forschung, dass eigenes Sprechen und Schreiben eine eigene, nicht ersetzbare Funktion hat — wer nur versteht, muss nie Sätze *bauen*.[12] Dieses Buch übernimmt darum nicht Krashens Maximalthese, sondern den Konsenskern: Ohne massiven verständlichen Input entsteht keine Sprache; alles andere baut auf ihm auf. Was das für die Grenzen einer Plattform bedeutet, die vor allem Input liefert, wird Kapitel 9 offen bilanzieren.

2.3 Die Probe aufs Exempel: die Bücherflut von Fidschi

Hypothesen dieser Größe brauchen eine Probe außerhalb des Seminarraums. Eine der schönsten stammt aus den Jahren 1980/81, aus ländlichen Grundschulen auf Fidschi, wo Englisch Zweitsprache und der Unterricht streng strukturiert war. Warwick Elley und Francis Mangubhai statteten Klassen der Stufen 4 und 5 mit je 250 attraktiven englischen Geschichtenbüchern aus — eine „Bücherflut" — und ließen sie täglich zwanzig bis dreißig Minuten damit arbeiten: die eine Hälfte mit gemeinsamem, vorgelesenem Lesen, die andere mit stillem Selbstlesen. Die Kontrollklassen behielten ihr gewohntes strukturbasiertes Programm. Entscheidend: Die Buchklassen bekamen *keine zusätzliche Englischzeit* — dieselben Stunden, nur anders gefüllt.[13]

Nach acht Monaten hatten die Buchklassen der Stufe 4 ihr Leseverstehen um fünfzehn Monate verbessert, die Kontrollklassen um sechseinhalb; im Hörverstehen und bei den Strukturtests zeigte sich dasselbe Muster. Im zweiten Jahr griff der Vorsprung auf das Schreiben und andere Fächer über; über sieben verschiedene Sprachtests hinweg lagen die Effektstärken zwischen 0,76 und 1,18 — in der Bildungsforschung sind das große Werte. Kinder, die einfach nur verstehbare, interessante Geschichten in Menge bekamen, überholten Kinder, die systematisch unterrichtet wurden, auch und gerade in dem, was die anderen systematisch geübt hatten: Grammatik.[14]

Eine Vorsicht gehört dazu: Fidschi war eine Flut aus *Büchern*, bei Kindern, im Klassenzimmer. Dass sich der Befund auf Erwachsene vor einem Bildschirm überträgt, ist damit nicht bewiesen — was trägt, ist das Prinzip: Reichhaltiger, verständlicher, bedeutsamer Input schlägt formzentrierte Instruktion, ohne dass jemand mehr Zeit investieren müsste. Ob der Film diese Rolle

---

[12] Zur Interaktions-Hypothese Long, M., The role of the linguistic environment in second language acquisition, in: Ritchie/Bhatia (Hg.), *Handbook of Second Language Acquisition*, San Diego 1996, S. 413–468 (am Druck zu prüfen); zur Output-Hypothese Swain, M., Communicative competence: Some roles of comprehensible input and comprehensible output in its development, in: Gass/Madden (Hg.), *Input in Second Language Acquisition*, Rowley 1985, S. 235–253 (am Druck zu prüfen). Zur allgemeinen Kritik an Krashens Theoriearchitektur bereits Kapitel 1, Fn. 6.

[13] Elley, W./Mangubhai, F., The impact of reading on second language learning, *Reading Research Quarterly* 19 (1983) 53–67 (am Volltext zu prüfen); Design und Ergebnisse hier wiedergegeben nach der ausführlichen Darstellung der Autoren selbst: Elley, W./Cutting, B./Mangubhai, F./Hugo, C., *Lifting Literacy Levels with Story Books*, 1996, ERIC ED416441, S. 4 f. (Volltext, abgerufen 2026-07-24; dort ausdrücklich: „no extra time was spent on English by the Book Flood Groups").

[14] Elley/Cutting/Mangubhai/Hugo (Fn. 5), S. 5: Stufe 4 — 15 Monate Lesefortschritt (beide Buchgruppen) gegen 6,5 Monate (Kontrolle) nach acht Monaten; Stufe 5 — 15 bzw. 9 gegen unter 3 Monate; im zweiten Jahr „highly significant differences" auf allen sieben Sprachtests, Effektstärken 0,76–1,18. Zur Einordnung der Größenordnung: Effektstärken ab 0,8 gelten nach der üblichen Konvention (Cohen) als groß.

für Erwachsene spielen kann, und unter welchen Bedingungen, ist die Frage von Teil II — und sie hat ihre eigene Forschung.

2.4 Zweite Bedingung: ein niedriger Filter

Die zweite Bedingung betrifft nicht den Input, sondern den Menschen, der ihn empfängt. Krashen nannte sie den *affektiven Filter*: Angst, fehlendes Selbstvertrauen, fehlende Motivation wirken wie eine Blende zwischen dem verstandenen Input und dem Teil des Kopfes, der daraus Sprache baut. Ist der Filter hoch, kann jemand eine Botschaft vollständig verstehen — und es bleibt trotzdem nichts hängen. Krashens Fazit ist bemerkenswert praktisch: Die wirksame Lehrkraft sei, wer „Input liefert und ihn in einer Situation niedriger Angst verständlich macht".[15]

Das Bild vom Filter ist eine Metapher, kein nachgewiesener Mechanismus, und es ist dafür kritisiert worden.[16] Aber das Phänomen, auf das es zeigt, gehört zum am besten Belegten der ganzen Disziplin. Elaine Horwitz und ihre Kollegen haben der Fremdsprachenangst 1986 ihren Namen und ihr Messinstrument gegeben und sie als eigenständiges Syndrom beschrieben: „ein eigenes Geflecht von Selbstwahrnehmungen, Überzeugungen, Gefühlen und Verhaltensweisen", das aus der Einzigartigkeit des Sprachenlernens entsteht — denn kein anderes Schulfach greift Selbstbild und Selbstausdruck so an wie das Sprechen in einer Sprache, die man noch nicht beherrscht.[17] Vier Jahrzehnte Forschung mit dieser und verwandten Skalen zeigen durchweg dasselbe: Je höher die Sprachangst, desto schlechter die Leistung, und zwar quer durch Länder, Sprachen und Altersgruppen.[18] Jeder, der je in einer Schulklasse aufgerufen wurde und den eigenen Herzschlag lauter hörte als die Frage der Lehrerin, kennt den Mechanismus von innen.

Für den Bau von Lernumgebungen ist das keine Randnotiz, sondern eine Konstruktionsvorgabe erster Ordnung. Es genügt nicht, guten Input bereitzustellen; man muss die Situation so bauen, dass niemand Angst haben muss, ihn zu empfangen und zu erwidern. Das ist der tiefere Grund, warum dieses Buch der Gemeinschaft einen ganzen Teil widmet: Der Chor, in dem die einzelne Stimme in der Menge geschützt ist, und die geschlossene Gruppe, in der Fehler unter Freunden passieren, sind — das wird Teil IV zeigen — keine sozialen Beigaben, sondern Angstsenker, und damit im Wortsinn Erwerbsbedingungen.

---

[15] Krashen 1982 (Fn. 1), S. 30–32; die Zwei-Bedingungen-Zusammenfassung S. 32 („comprehensible … input containing i + 1 … and second, a low or weak affective filter to allow the input ‚in'"); das Lehrkraft-Zitat S. 32 („The effective language teacher is someone who can provide input and help make it comprehensible in a low anxiety situation", Übersetzung von uns).

[16] Krashen verortet den Filter selbst „outside the language acquisition device proper" (1982, Fn. 1, S. 31) — er ist ein Erklärungsbild für einen Zusammenhang, kein neurophysiologisch identifizierter Mechanismus; das Konzept geht auf Dulay/Burt (1977) zurück (ebd., S. 31). Kritisch zur Prüfbarkeit McLaughlin (Kapitel 1, Fn. 6).

[17] Horwitz, E. K./Horwitz, M. B./Cope, J., Foreign Language Classroom Anxiety, *The Modern Language Journal* 70 (1986) 125–132, hier S. 128: „we conceive foreign language anxiety as a distinct complex of self-perceptions, beliefs, feelings, and behaviors related to classroom language learning arising from the uniqueness of the language learning process" (Übersetzung im Text von uns); zur Sonderstellung des Fachs ebd., S. 128 („Probably no other field of study implicates self-concept and self-expression to the degree that language study does."), doi:10.1111/j.1540-4781.1986.tb05256.x (Fundstellen am Volltext verifiziert). Dort auch die Foreign Language Classroom Anxiety Scale (FLCAS), S. 129 f.

[18] Meta-analytisch zusammengefasst u. a. bei Teimouri, Y./Goetze, J./Plonsky, L., Second language anxiety and achievement: A meta-analysis, *Studies in Second Language Acquisition* 41 (2019) 363–387 (durchweg negative Korrelationen mittlerer Stärke; am Volltext zu prüfen); ähnlich Botes/Greiff/Dewaele 2020 zur FLCAS (am Volltext zu prüfen). Auf die Übernahme konkreter Koeffizienten wird verzichtet, bis die Primärstellen geprüft sind.

2.5 Die dritte Frage: Was bleibt?

Verständlich und angstfrei — reicht das? Noch nicht ganz. Zwischen dem Verstehen eines Satzes und seinem Besitz liegt eine dritte Hürde: das Behalten. Jeder kennt Sprache, die verstanden wurde und trotzdem spurlos verschwand — der Vortrag von letzter Woche, das Gespräch von gestern. Und jeder kennt das Gegenteil: den Werbejingle aus der Kindheit, das Lied, das nach dreißig Jahren noch wortwörtlich abrufbar ist, den Filmsatz, den man mit der Stimme des Schauspielers zitiert. Offenbar gibt es Formen von Input, die *haften* — weil sie mehrfach kodiert sind (Bild und Ton zugleich), weil sie emotional besetzt sind, weil Rhythmus und Melodie sie tragen, weil eine Geschichte ihnen einen Platz gibt.

Dieses Buch fasst diese dritte Bedingung unter dem Wort *verankert* zusammen. Anders als „Input" und „affektiver Filter" ist das kein eingeführter Fachbegriff, sondern eine Sammelklammer dieses Buches für mehrere getrennt erforschte Effekte — die doppelte Kodierung von Bild und Wort, das Gedächtnis für Gesungenes, die Kraft der Erzählung —, die in den Teilen II, III und V einzeln belegt werden.[19] Damit ist das Prüfraster komplett, das dieses Buch durchziehen wird. Es besteht aus drei Fragen, die man an jede Lernumgebung stellen kann, an ein Klassenzimmer ebenso wie an eine App:

- **Verständlich?** — Bekommt der Lernende Sprache in Menge, die er *verstehen* kann, weil Kontext, Bild und Vorwissen ihm helfen?
- **Verankert?** — Ist der Input so beschaffen, dass er *bleibt* — mehrfach kodiert, emotional, rhythmisch, erzählt?
- **Verbunden?** — Geschieht das Ganze in einem Raum, der die Angst *senkt*, statt sie zu erzeugen?

Der Regel-Unterricht aus Kapitel 1 beantwortet alle drei Fragen mit Nein: Sein Input ist knapp und formzentriert, sein Stoff ist auf Abfragbarkeit statt auf Haften gebaut, und seine Prüfungssituation ist eine Angstmaschine. Das ist, in drei Sätzen, die Diagnose — und zugleich der Bauplan für alles, was folgt.

2.6 Der Sog: Warum Menschen wiederkommen

Verständlich, verankert, verbunden — die drei Fragen beschreiben, was mit dem Input geschieht. Eine vierte Frage entscheidet darüber, ob es überhaupt Input gibt: Warum sollte jemand am Dienstag wiederkommen, der am Montag freiwillig gekommen ist? Die Motivationspsychologie hat darauf eine erstaunlich stabile Antwort. Nach der Selbstbestimmungstheorie von Deci und Ryan trägt sich menschliches Engagement dann selbst, wenn drei Grundbedürfnisse erfüllt sind: Autonomie (ich wähle), Kompetenzerleben (ich kann) und soziale Eingebundenheit (ich gehöre dazu).[20] Man lese die drei Begriffe neben den Bauplan dieses Buches: Die Auswahl des

---

[19] „Verankert" ist also Eigenbeitrag dieses Buches als *ordnender* Begriff, nicht als Befund: Die Bausteine — Dual-Coding-Theorie und multimediales Lernen (Teil II), Gedächtnis für gesungene Sprache (Teil III), narrative Struktur (Teil V) — sind fremde, je einzeln belegte Forschungslinien; neu ist nur ihre Zusammenfassung unter einer Bedingung neben „verständlich" und „verbunden". Wo dieses Buch integriert statt belegt, wird das — wie hier — ausgewiesen.

[20] Deci, E. L./Ryan, R. M., *Intrinsic Motivation and Self-Determination in Human Behavior*, New York 1985 (am Druck zu prüfen); zusammenfassend Ryan/Deci, Self-determination theory and the facilitation of intrinsic motivation, *American Psychologist* 55 (2000) 68–78 (am Volltext zu prüfen). Übertragungs-Hinweis: Die Selbstbestimmungstheorie ist eine allgemeine Motivationstheorie, keine sprachdidaktische; die Zuordnung ihrer drei Grundbedürfnisse zu Filmauswahl, Muttersprachzeile und Gruppe ist Deutung dieses Buches, kein Messbefund.

Films ist Autonomie; die Muttersprachzeile erzeugt Kompetenzerleben vom ersten Abend an — man *versteht*, wo man im Kurs scheiterte; und die Gruppe, von der Teil IV handelt, ist die Eingebundenheit. Der Lückentext einer Lern-App erfüllt keines der drei Bedürfnisse: Er wählt für den Lernenden, er dokumentiert vor allem Fehler, und er wird allein bearbeitet.

Wo die drei Bedürfnisse zusammenkommen, kann Beschäftigung in den Zustand kippen, den Csíkszentmihályi *Flow* genannt hat: Das Tun trägt sich selbst, die Anstrengung verschwindet aus dem Erleben.[21] Der Unterschied ist im Alltag mit bloßem Auge sichtbar: Die Hausaufgabe wird absolviert, weil das System sie abprüft; der Film wird zu Ende geschaut, weil man wissen will, wie es weitergeht. Für eine Methode, deren erste Bedingung *Menge* heißt (2.2), ist das keine psychologische Fußnote, sondern die Antwort auf die Mengenfrage selbst: Menge entsteht nicht aus Disziplin, sondern aus Sog.

2.7 Vom Prinzip zum Bau

Nimmt man die drei Fragen ernst, wird aus der Spracherwerbstheorie eine Konstruktionsaufgabe: Woher bekommt eine erwachsene Frau in Gelsenkirchen, ein Jugendlicher in Bologna *massenhaft* verständlichen Input in der Zielsprache — nicht fünfundvierzig Minuten pro Woche, sondern jeden Abend, freiwillig, gern? Wie muss dieser Input beschaffen sein, damit er nicht nur verstanden wird, sondern bleibt? Und wie baut man darum einen sozialen Raum, der die Angst senkt, statt sie zu vergrößern?

Die nächsten vier Teile dieses Buches gehen diese Fragen der Reihe nach durch. Die erste Antwort liegt näher, als der Sprachunterricht lange wahrhaben wollte: Sie flimmert seit hundert Jahren über Leinwände und Bildschirme, sie ist authentisch, emotional, massenhaft vorhanden und in fast jede Sprache der Welt untertitelt. Das nächste Kapitel handelt vom Film — und von der Frage, was aus einem Fernsehabend wird, wenn man ihn richtig baut.

[21] Csíkszentmihályi, M., *Flow: The Psychology of Optimal Experience*, New York 1990 (am Druck zu prüfen). Auch hier der Übertragungs-Hinweis: Die Flow-Forschung wurde an aktiven Tätigkeiten (Klettern, Musizieren, Schach) entwickelt; ob das verstehende Sehen mit Doppelspur die dort beschriebene Passung von Anforderung und Können erzeugt, ist plausibel, aber ungemessen — sie gehört sachlich zur Messliste von Kapitel 9 (Verbleibs- und Nutzungsmaße).

Teil II — Der Input: Sehen, hören, lesen

„maiorem habere vim ad discenda ista liberam curiositatem quam meticulosam necessitatem." — „… dass die freie Neugier mehr Kraft hat, uns lernen zu lassen, als der ängstliche Zwang."

— Augustinus, Confessiones I,14,23 (Übersetzung von uns)

Kapitel 3 · Sehen, hören, lesen: der doppelt kodierte Film

3.1 Ein europäisches Zufallsexperiment

Warum sprechen die Niederländer, die Skandinavier, die Portugiesen so auffällig gut Englisch — und die Deutschen, Spanier, Italiener so viel schlechter? Die übliche Antwort zeigt auf die Schulen. Die interessantere Antwort zeigt auf den Fernseher. Als in den 1920er und 30er Jahren der Tonfilm nach Europa kam, musste jedes Land entscheiden, was es mit fremdsprachigen Filmen tut. Große Sprachräume, die sich die teure Synchronisation leisten konnten — Deutschland, Italien, Spanien, Frankreich —, ließen die Stimmen neu einsprechen. Kleine Länder untertitelten. Die Entscheidung fiel aus ökonomischen und politischen Gründen, lange bevor jemand an Sprachdidaktik dachte — und sie blieb, wie Mediensysteme bleiben, für Generationen.

Damit hat die Geschichte ein Experiment angelegt, das kein Forscher je hätte genehmigen lassen: ganze Bevölkerungen, über Jahrzehnte, die eine Gruppe mit englischem Originalton und Untertiteln, die andere mit Synchronfassung. Eine ökonometrische Studie hat die Bilanz gezogen: Länder mit Untertitel-Tradition schneiden in den englischen Sprachtests markant besser ab — der geschätzte Untertitel-Effekt entspricht etwa 17 Prozent des durchschnittlichen Fertigkeitsniveaus, anderthalb Standardabweichungen im Ländervergleich, und er ist am stärksten genau dort, wo der Fernseher am meisten liefert: beim Hörverstehen.[22] Portugal untertitelt, Spanien synchronisiert; die portugiesischen TOEFL-Ergebnisse liegen deutlich über den spanischen — bei zwei Nachbarländern mit eng verwandten Sprachen.[23] Man muss diese Zahlen mit der gebotenen Vorsicht lesen: Es sind Ländermittelwerte, keine Messungen an einzelnen Lernern, und Länder unterscheiden sich in mehr als ihren Fernsehgewohnheiten.[24] Aber die Richtung ist schwer wegzudiskutieren — zumal sie exakt dorthin zeigt, wohin auch die Labor- und Klassenzimmerforschung zeigt, von der dieses Kapitel handelt.

Der Fernsehabend ist, mit anderen Worten, nie nur Unterhaltung gewesen. Er ist die größte Sprachlernmaschine, die je gebaut wurde — nur hat sie niemand als solche gebaut. Dieses Kapitel fragt, *warum* sie funktioniert, und das nächste, wie man sie vom Zufallsprodukt zum Instrument macht.

3.2 Zwei Kanäle sind mehr als einer

Der erste Grund liegt in der Architektur des Gedächtnisses. Allan Paivio hat in den 1970er und 80er Jahren die *Dual-Coding-Theorie* entwickelt: Menschen verarbeiten Information in zwei getrennten, aber verbundenen Systemen — einem sprachlichen für Wörter und einem bildlichen

---

[22] Rupérez Micola, A./Aparicio Fenoll, A./Banal-Estañol, A./Bris, A., TV or not TV? The impact of subtitling on English skills, *Journal of Economic Behavior & Organization* 158 (2019) 487–499, doi:10.1016/j.jebo.2018.12.019 (Fundstellen am Autoren-PDF verifiziert): Untertitel-Effekt „equivalent to 16.9% (one and a half standard deviations) of the average level of English skills"; am stärksten beim Hörverstehen. Zur historischen Wahl der Übersetzungsmodi in den 1920er/30er Jahren ebd., Abschnitt 2.

[23] Ebd., Abschnitt 2 (Beispiel Portugal/Spanien: „in Spain television is dubbed. And, as a result, Portugal's results in the TOEFL exams are much better than Spain's"); Subtitling-Länder liegen im internetbasierten TOEFL im Rohvergleich 3,4 Punkte vorn, am deutlichsten im Listening (ebd., deskriptive Ergebnisse).

[24] Übertragungs-Hinweis: Es handelt sich um Ländermittelwerte (TOEFL-Landesdurchschnitte) mit ökonometrischer Identifikation über die historische Wahl des Übersetzungsmodus — nicht um Lernverlaufsmessungen an Individuen. Was trägt: massenhafter untertitelter Originalton wirkt bevölkerungsweit auf Sprachfertigkeiten, besonders das Hörverstehen. Was nicht trägt: eine Aussage über ein bestimmtes Produkt, eine bestimmte Lernergruppe oder die Doppel-Untertitel-Form (dazu Fn. 13 f.). Selektionsfragen (wer nimmt am TOEFL teil?) diskutieren die Autoren selbst (ebd., Abschnitt 4).

für Anschauung. Was in beiden Systemen zugleich abgelegt wird, ein Wort *mit* seinem Bild, hat zwei Gedächtnisspuren statt einer und entsprechend mehr Wege, wiedergefunden zu werden.[25] Richard Mayer hat daraus eine empirisch breit geprüfte Theorie des multimedialen Lernens gebaut, deren Grundbefund inzwischen Lehrbuchwissen ist: Menschen lernen aus Wort *und* Bild besser als aus Wort allein.[26]

Für den Spracherwerb ist das mehr als eine Gedächtnistechnik. Man erinnere sich an Krashens Paradox aus Kapitel 2: Erwerb geschieht, wenn man Sprache versteht, die man eigentlich noch nicht kann — und das gelingt nur, weil außer der Sprache noch etwas da ist, das beim Verstehen hilft. Der Film ist genau dieses „außer der Sprache": Die Handlung zeigt, wovon die Rede ist; das Gesicht zeigt, wie es gemeint ist; die Szene liefert den Kontext, den der Anfänger braucht, um aus einem unverständlichen Lautstrom eine verstehbare Botschaft zu machen. Der Film ist *i + 1* mit eingebauter Verstehenshilfe — verständlicher Input in industrieller Menge.

3.3 Die dritte Spur: Untertitel in der Zielsprache

Doch Hören und Sehen allein lassen eine Lücke. Wer eine fremde Sprache nur hört, weiß oft nicht einmal, *wo ein Wort aufhört und das nächste beginnt* — der Lautstrom kennt keine Leerzeichen. Hier kommt die dritte Spur ins Spiel: der Untertitel in der Zielsprache, das Mitlesen dessen, was man hört. Robert Vanderplank hat schon in den 1980er Jahren, mit den Teletext-Untertiteln der BBC, systematisch beobachtet, dass Lerner aus untertiteltem Fernsehen Sprache ziehen — und dass der Untertitel nicht vom Hören ablenkt, sondern das Gehörte lesbar macht.[27] Seither ist daraus eine eigene Forschungsrichtung geworden. Ihre Meta-Analyse — achtzehn Studien zu Hörverstehen und Wortschatz — kommt zu einem klaren Ergebnis: Untertitel in der Zielsprache verbessern beides, das Verstehen und das Wörterlernen, mit großen Effekten.[28]

Warum, zeigt eine viel zitierte Studie aus Michigan besonders schön. Lerner des Arabischen, Chinesischen, Spanischen und Russischen sahen kurze Videos mit und ohne zielsprachliche Untertitel; mit Untertiteln schnitten sie in Wortschatz- und Verstehenstests besser ab. Die Erklärung der Autorinnen ist die entscheidende: Untertitel helfen, *Wortgrenzen zu identifizieren* — sie zerlegen den „ansonsten unverständlichen Sprachstrom" in Einheiten, an die sich Bedeutung heften kann, und verbinden das Gehörte mit dem Geschriebenen.[29] Und dass dieser Mechanis-

---

[25] Paivio, A., *Mental Representations: A Dual Coding Approach*, Oxford 1986 (am Druck zu prüfen); zusammenfassend ders., Dual coding theory: Retrospect and current status, *Canadian Journal of Psychology* 45 (1991) 255–287 (am Volltext zu prüfen).

[26] Mayer, R. E., *Multimedia Learning*, 3. Aufl., Cambridge 2021 (am Druck zu prüfen; „Multimedia-Prinzip"). Mayers Theorie verbindet Paivios Zwei-Kanal-Annahme mit begrenzter Verarbeitungskapazität und aktiver Verarbeitung; sie ist an Erklär-Lernstoff (nicht an Sprachlernstoff) entwickelt worden — zur Übertragungsgrenze siehe 3.4 und Fn. 10 f.

[27] Vanderplank, R., The value of teletext sub-titles in language learning, *ELT Journal* 42 (1988) 272–281 (am Druck zu prüfen); ders., *Captioned Media in Foreign Language Learning and Teaching*, London 2016 (am Druck zu prüfen).

[28] Montero Perez, M./Van Den Noortgate, W./Desmet, P., Captioned video for L2 listening and vocabulary learning: A meta-analysis, *System* 41 (2013) 720–739, doi:10.1016/j.system.2013.07.013: großer Effekt zielsprachlicher Untertitel auf Hörverstehen (15 Studien) wie Wortschatz (10 Studien). Der Volltext war elektronisch nicht frei zugänglich; die berichteten Effektstärken (in der Literatur mit g ≈ 0,99 bzw. 0,87 wiedergegeben) sind sekundär übernommen und **am Volltext zu prüfen**; übernommen wird hier nur die Kernaussage des Abstracts (großer Effekt auf beides).

[29] Winke, P./Gass, S./Sydorenko, T., The effects of captioning videos used for foreign language listening activities, *Language Learning & Technology* 14/1 (2010) 65–86, hier S. 65 (Volltext, Fundstellen am PDF verifiziert): „captioning was more effective than no captioning"; Untertitel helfen, Wortgrenzen zu identifizieren und den „ansonsten unverständlichen Sprachstrom" zu segmentieren (ebd., S. 65, Übersetzung von uns). Ehrlichkeitshalber: In den In-

mus nicht auf den Unterricht angewiesen ist, zeigt die Forschung an Kindern, die fremdsprachiges Fernsehen schlicht *geschaut* haben und dabei messbar Sprache aufnahmen — beiläufig, ohne Auftrag, ohne Test im Rücken.[30]

3.4 Die Redundanz-Frage — ehrlich ausgetragen

An dieser Stelle muss dieses Buch eine Spannung austragen, die man in Werbetexten vergeblich suchen würde. Denn dieselbe multimediale Lernforschung, auf die sich Abschnitt 3.2 stützt, kennt ein *Redundanzprinzip*, das gegen Bildschirmtext zu sprechen scheint: Wer eine Erklärung hört und denselben Text zusätzlich geschrieben mitliest, lernt oft *schlechter* als ohne den Text — die doppelte Darbietung desselben Wortlauts verstopft die Verarbeitung, statt sie zu stützen.[31] Wäre das uneingeschränkt wahr, wäre der untertitelte Film eine schlechte Idee.

Die Auflösung liegt im Lernziel, und sie verdient es, genau formuliert zu werden. Das Redundanzprinzip wurde an Muttersprachlern gewonnen, die aus gesprochenen Erklärungen *Inhalte* lernen sollten — Blitzentstehung, Bremssysteme. Für sie ist der mitlaufende Text tatsächlich redundant: Er enthält nichts, was das Ohr nicht mühelos liefert. Für den Sprachlerner ist die Lage umgekehrt. Sein Lernziel ist nicht der Inhalt hinter der Sprache, sondern die Sprache selbst — und sein Ohr liefert eben *nicht* mühelos: Es verliert Wortgrenzen, verschluckt Endungen, rät bei jedem dritten Wort. Der zielsprachliche Untertitel ist für ihn keine zweite Kopie des Verstandenen, sondern oft der einzige Zugang zum Nicht-Verstandenen. Die Redundanz-Literatur selbst benennt diese Randbedingungen inzwischen ausdrücklich: Bei Nicht-Muttersprachlern, bei schwer verständlichem Ton, bei kurzem, langsam dargebotenem Text kehrt sich das Prinzip um oder verschwindet.[32] Die Captioning-Befunde aus 3.3 sind also keine Anomalie, die man verschweigen müsste — sie sind der gut belegte Sonderfall, in dem aus Redundanz Stützung wird. Bleiben soll von dieser Debatte trotzdem ein Stachel: Text auf dem Bildschirm ist nicht gratis. Er kostet Aufmerksamkeit, und eine Lernumgebung, die Untertitel einsetzt, muss sich fragen lassen, ob sie diese Aufmerksamkeit gut investiert.

3.5 Zwei Sprachen zugleich: die zweisprachige Wette

Pontivo geht über den einsprachigen Untertitel hinaus: Sein Player zeigt *zwei* Spuren zugleich — die Zielsprache, synchron zum Ton, und darunter die Muttersprache des Lernenden. Die Logik ergibt sich aus allem Bisherigen: Die zielsprachliche Spur segmentiert den Lautstrom und verbindet Klang mit Schrift; die muttersprachliche Spur garantiert das Verstehen — sie hält die Bedingung „verständlich" auch dort aufrecht, wo Bild und Kontext allein nicht mehr reichen, und macht damit authentisches, ungestuftes Material für Anfänger überhaupt erst zugänglich. Sie ist,

---

terviews berichteten Lerner auch, Untertitel als „Krücke" zu benutzen (ebd., S. 65) — die Abhängigkeitsfrage ist real und gehört auf die Messliste (Kapitel 9).

[30] d'Ydewalle, G./Van de Poel, M., Incidental foreign-language acquisition by children watching subtitled television programs, *Journal of Psycholinguistic Research* 28 (1999) 227–244 (am Druck zu prüfen — Details des Designs, insbesondere welche Untertitel-Richtung welchen Zuwachs erzeugte, werden hier bewusst nicht referiert, bis die Primärstelle geprüft ist).

[31] Mayer 2021 (Fn. 5), Kapitel zum Redundanzprinzip (am Druck zu prüfen); grundlegend zur „redundancy" auch die Cognitive-Load-Forschung um Sweller/Kalyuga (am Druck zu prüfen).

[32] Randbedingungen des Redundanzprinzips — u. a. Nicht-Muttersprachler, schwer verständliche Audioquelle, kurzer/langsamer Text — bei Mayer 2021 (Fn. 5) ausdrücklich als „boundary conditions" geführt (am Druck zu prüfen). Die L2-Captioning-Meta-Analyse (Fn. 7) ist der empirische Gegenpol auf Seiten des Sprachenlernens.

um einen Begriff aus Kapitel 7 vorwegzunehmen, gelebtes *Translanguaging*: Die Muttersprache wird nicht ausgesperrt, sondern als Verstehensgerüst eingespannt.[33]

Wie steht die Forschung zu zweisprachigen Doppel-Untertiteln? Sie ist jung, klein — und ermutigend. Eine Blickbewegungsstudie mit chinesischen Englischlernern verglich alle vier Bedingungen (nur Muttersprache, nur Zielsprache, beide, keine): Die Doppelspur erzeugte *keine* messbare Mehrbelastung gegenüber einer einzelnen Spur, und sie half dem Verstehen gegenüber dem Sehen ohne Untertitel; die Blicke der Lerner verteilten die Arbeit selbst — stabil auf die Muttersprachzeile für die Bedeutung, flexibel auf die Zielsprachzeile für die Form.[34] Weitere Studien finden Wortschatzzuwächse aus doppelt untertiteltem Sehen.[35] Aber es wäre unehrlich, daraus mehr zu machen, als es ist: Das sind wenige Studien mit kleinen Stichproben und kurzen Clips. Dass Pontivos konkrete Bauform — die synchrone Doppelspur im Dauerbetrieb, über Monate, bei Geflüchteten und Schulklassen — hält, was die Prinzipien versprechen, ist eine *Design-Wette*: gut begründet, noch nicht gemessen. Sie steht auf der Messliste von Kapitel 9 weit oben.

3.6 Was der Film noch nicht kann

Ehrlich bleiben heißt auch: Der Film hat Grenzen, und zwei davon sind grundsätzlich. Die erste: Er liefert Verstehen, aber keine einzelne Worterklärung. Wer mitten im Satz an einem Wort hängen bleibt, dem hilft die schönste Doppelspur nicht weiter — er braucht eine Antwort auf *dieses eine Wort*, jetzt, ohne den Film zu verlassen. Wie aus diesem Moment des Bedürfnisses Wortschatz wird, ist die Frage des nächsten Kapitels. Die zweite Grenze ist tiefer: Der Film macht aus niemandem einen Sprecher. Er füllt das Ohr, nicht den Mund. Was es braucht, damit aus verstandener Sprache gesprochene wird — und warum dieser Weg über das Singen führen kann —, davon handeln die Teile III und IV.

---

[33] Funktions-Hinweis: Der Player mit synchroner Doppelspur ist in der Mediathek der Plattform in Betrieb (pontivo.de/mediathek, Stand 15. Oktober 2026).

[34] Liao, S./Kruger, J.-L./Doherty, S., The impact of monolingual and bilingual subtitles on visual attention, cognitive load, and comprehension, *The Journal of Specialised Translation* 33 (2020) 70–98, doi:10.26034/cm.jostrans.2020.549 (Open Access, Fundstellen am Volltext verifiziert): „Bilingual subtitles … did not appear to induce more cognitive load or produce more cognitive gain than monolingual subtitles"; gegenüber der Bedingung ohne Untertitel „more beneficial … linguistic support" (Abstract, S. 70); stabilere Blickzuwendung zur L1-Zeile ebd. Zwanzig Teilnehmende, vier Kurzvideos — eine kleine Studie, als solche zu gewichten.

[35] Wang, A./Pellicer-Sánchez, A., Incidental vocabulary learning from bilingual subtitled viewing: An eye-tracking study, *Language Learning* 72 (2022), doi:10.1111/lang.12495 (am Volltext zu prüfen); ferner die Folgeforschung zu Doppel-Untertiteln und Vokabelerwerb (u. a. *ReCALL*, *SSLA* 2022 ff., am Volltext zu prüfen). Die Gesamtlage: konsistent positiv, aber jung und schmal.

Kapitel 4 · Das Wort im Moment des Bedürfnisses

4.1 Die Stelle, an der man hängen bleibt

Man sitzt im Film, es läuft gut, man versteht — und dann kommt ein Wort, an dem alles hängt. *Der Schwarm verhält sich wie ein einziges Tier.* Schwarm? Die Doppelspur aus Kapitel 3 hilft ein Stück weit: Unten steht die Übersetzung des Satzes. Aber welches Wort im Satz war nun welches? Warum steht das Verb da, wo es steht? Und was genau heißt dieses eine Wort, um das es offenbar geht? In diesem Moment ist der Lernende in einem Zustand, den sich jeder Didaktiker nur wünschen kann: Er *will etwas wissen*. Nicht, weil es im Lehrplan steht, sondern weil es ihm gerade den Film versperrt. Die Frage dieses Kapitels ist, was man mit diesem Moment macht — und was aus Wörtern wird, wenn man ihn richtig nutzt.

4.2 Beiläufiges Lernen ist real — und es ist langsam

Die Vokabelforschung hat eine gute und eine unbequeme Nachricht. Die gute: Wortschatz wächst tatsächlich *beiläufig* — beim Lesen, Hören und Zuschauen, ohne Vokabelheft, einfach dadurch, dass Wörter in verständlichen Zusammenhängen wieder und wieder vorkommen. Der überwiegende Teil des Wortschatzes, den Muttersprachler besitzen, ist genau so entstanden; niemand hat ihnen zwanzigtausend Wörter beigebracht.[36] Auch vor dem Bildschirm funktioniert das: In der Leuvener Studie, die drei Untertitel-Varianten verglich, hatten die Zuschauer nach zweimaligem Sehen dreier französischer Clips messbar Wörter aufgenommen — die Untertitel-Gruppen erkannten die *Form* der neuen Wörter signifikant besser als die Kontrollgruppe und wussten noch, in welchem Clip sie ihnen begegnet waren.[37]

Die unbequeme Nachricht steckt in derselben Studie: Auf die Frage, was die Wörter *bedeuten*, war der Vorsprung schon dünner — nur die Gruppen, deren Untertitel die Schlüsselwörter hervorhoben, übertrafen die Kontrollgruppe beim Wiedererkennen der Bedeutung, und beim freien Erinnern der Bedeutung verschwand der Effekt ganz.[38] Das ist kein Ausreißer, sondern der Normalbefund des beiläufigen Lernens: Es baut zuerst ein Gefühl für die Form auf — *dieses Wort gibt es, so klingt es, da gehört es hin* — und erst über viele Begegnungen hinweg die sichere Bedeutung. Beiläufigkeit ist breit, aber flach; pro Begegnung bleibt wenig hängen, und seltene Wörter begegnen einem selten.[39] Wer nur schaut, lernt wirklich — aber langsam.

---

[36] Grundlegend zum beiläufigen (inzidentellen) Vokabellernen: Nation, I. S. P., *Learning Vocabulary in Another Language*, 2. Aufl., Cambridge 2013, insb. Kap. 4 (am Druck zu prüfen); Hulstijn, J., Incidental and intentional learning, in: Doughty/Long (Hg.), *The Handbook of Second Language Acquisition*, Oxford 2003, S. 349–381 (am Druck zu prüfen). Auf die Wiedergabe konkreter Schwellenwerte (wie viele Begegnungen ein Wort braucht, wie viele Wortfamilien welchen Textdeckungsgrad ergeben) wird verzichtet, bis die Primärstellen geprüft sind.

[37] Montero Perez, M./Peters, E./Clarebout, G./Desmet, P., Effects of captioning on video comprehension and incidental vocabulary learning, *Language Learning & Technology* 18/1 (2014) 118–141, hier S. 118 (Volltext, Fundstellen am PDF verifiziert): 133 flämische Studierende, drei französische Clips, zweimal gesehen; „the captioning groups scored equally well on form recognition and clip association and significantly outperformed the control group".

[38] Ebd., S. 118: „Only the keyword captioning and full captioning with highlighted keywords groups outperformed the control group on meaning recognition. Captioning did not affect comprehension nor meaning recall."

[39] Zur Charakteristik „breit, aber flach" des inzidentellen Lernens statt vieler: Nation 2013 (Fn. 1), Kap. 4; die kleinen Pro-Begegnung-Zuwächse sind der Grund, warum die Forschung Wiederholung und Frequenz so betont (am Druck zu prüfen).

4.3 Was die Glosse ändert

Es gibt ein altbewährtes Werkzeug, das genau an dieser Schwäche ansetzt: die *Glosse* — die kleine Worterklärung am Rand des Textes, im Buchdruck so alt wie das Buch selbst. Die Forschung hat sie gründlich vermessen. Eine Meta-Analyse über 42 Studien mit fast viertausend Lernern kommt zu einem bemerkenswert klaren Ergebnis: Lesen mit Glossen führte zu 45,3 Prozent gelernten Zielwörtern im sofortigen Test (gegenüber 26,6 Prozent ohne Glossen) und zu 33,4 Prozent im verzögerten Test (gegenüber 19,8) — die Glosse verdoppelt das beiläufige Lernen beinahe.[40] Zwei Details derselben Analyse sind für alles Weitere wichtig. Erstens: Glossen in der *Muttersprache* der Lernenden wirkten besser als Erklärungen in der Zielsprache — ein weiterer Beleg dafür, dass die Muttersprache im Lernprozess Gerüst ist, nicht Gegner.[41] Zweitens: Am stärksten wirkten Glossen, die den Lernenden kurz *arbeiten* ließen (etwa Mehrfachwahl-Glossen, bei denen man die passende Bedeutung auswählt), am schwächsten die, die man gar nicht erst aufsuchen muss.[42]

Der zweite Befund hat eine Theorie im Rücken, die man kennen sollte, weil sie den „Moment des Bedürfnisses" aus 4.1 direkt adressiert: die *Involvement-Load-Hypothese* von Laufer und Hulstijn. Sie besagt, grob gesprochen, dass ein Wort umso eher bleibt, je mehr von drei Zutaten im Lernmoment steckt: *Bedürfnis* (ich brauche dieses Wort jetzt), *Suche* (ich muss mir seine Bedeutung beschaffen) und *Bewertung* (ich muss prüfen, ob die Bedeutung in meinen Zusammenhang passt).[43] Das Vokabelheft am Sonntagabend hat von allen dreien wenig. Der Klick im Film hat alle drei: Das Bedürfnis kommt aus der Handlung, die Suche ist der Klick selbst, und die Bewertung geschieht im Moment, in dem die Erklärung auf die Szene trifft, die man gerade vor Augen hat.

4.4 Der Klick — und was an ihm Wette ist

Pontivos Klick-Erklärung ist die Glosse, in die Jahrhundertform übersetzt: Wer im Player an einer Stelle hängen bleibt, tippt sie an und bekommt eine Erklärung *dieser Stelle in diesem Zusammenhang* — in seiner Muttersprache, ohne den Film zu verlassen.[44] Das ist mehr als ein Wörterbuch-Aufruf, denn erklärt wird nicht das Wort im luftleeren Raum, sondern die Stelle: die Wendung, die Anspielung, die Grammatik, die den Satz so gebaut hat. Nach allem, was die Glossen-

---

[40] Yanagisawa, A./Webb, S./Uchihara, T., How do different forms of glossing contribute to L2 vocabulary learning from reading? A meta-regression analysis, *Studies in Second Language Acquisition* 42 (2020) 411–438, doi:10.1017/S0272263119000688 (Zahlen am Verlagsabstract verifiziert; Volltext am Druck zu prüfen): 359 Effektstärken aus 42 Studien, N = 3.802; gelernte Zielwörter mit Glossen 45,3 % (sofort) bzw. 33,4 % (verzögert) gegenüber 26,6 % bzw. 19,8 % ohne.

[41] Ebd. (Abstract): „L1 glosses yielded greater learning than L2 glosses" — ohne bedeutsame Interaktion mit dem Sprachstand; die Überlegenheit der Muttersprachglosse galt also auch für Fortgeschrittene.

[42] Ebd. (Abstract): Mehrfachwahl-Glossen am wirksamsten, In-Text-Glossen und Glossare am schwächsten. Die Deutung — wirksam ist, was kurze aktive Verarbeitung erzwingt — ist Interpretation der Autoren wie dieses Buches und deckt sich mit der Involvement-Load-Hypothese (Fn. 8).

[43] Laufer, B./Hulstijn, J., Incidental vocabulary acquisition in a second language: The construct of task-induced involvement, *Applied Linguistics* 22 (2001) 1–26 (am Volltext zu prüfen; die Drei-Komponenten-Darstellung — need, search, evaluation — folgt dem Aufsatztitelkonstrukt und der Standardrezeption, z. B. bei Nation 2013, Fn. 1).

[44] Funktions-Hinweis: Die Klick-Erklärung ist im Player der Mediathek in Betrieb (pontivo.de/mediathek); angeklickte Wörter wandern in eine persönliche Kartei mit Wiederholungs- und Prüfungsmodus (pontivo.de/review, pontivo.de/pruefung). Stand 15. Oktober 2026.

Forschung weiß, ist genau das die richtige Bauform: muttersprachlich, bedürfnisgetrieben, kontextgebunden.[45]

Aber hier muss dieses Buch zwei Wetten offen auf den Tisch legen. Die erste: Die Erklärungen erzeugt ein Sprachmodell — eine Maschine, die sehr gute und gelegentlich falsche Erklärungen produziert, und die Forschung, die Glossen vermessen hat, hat *redaktionell geprüfte* Glossen vermessen, keine maschinell erzeugten. Ob die maschinelle Glosse die Qualität hält, auf der ihre Belegbasis steht, ist offen und muss kontrolliert werden — das Schwesterbuch dieser Reihe hat dem Problem, dass maschinelle Auskünfte Belege brauchen, die tragen, ein ganzes Werk gewidmet.[46] Die zweite Wette betrifft das Maß: Die Glossen-Forschung weiß auch, dass Unterbrechung kostet. Wer alle fünf Sekunden klickt, schaut keinen Film mehr, sondern blättert ein Wörterbuch mit Bildern — und verliert genau das, was den Film als Input so wertvoll macht: den Fluss, die Geschichte, das Dranbleiben. Wo das gesunde Maß liegt, zwischen dem Klick, der ein Loch stopft, und dem Klicken, das den Abend zerlegt, ist empirisch nicht geklärt. Beides steht auf der Messliste von Kapitel 9.

4.5 Verstanden ist noch nicht behalten

Nehmen wir an, alles gelingt: Der Film liefert verständlichen Input in Menge, die Doppelspur hält das Verstehen hoch, der Klick vertieft die Stellen, an denen es klemmt. Dann hat der Lernende am Ende des Abends — verstanden. Aber zwischen dem Verstehen und dem Besitz einer Sprache liegt die Nacht, und die Nacht ist ein Sieb. Was am Morgen noch da ist, entscheidet sich nicht allein an der Zahl der Begegnungen, sondern daran, *wie* der Stoff im Kopf abgelegt wurde: ob er eine Form hatte, an der das Gedächtnis Halt findet. Die älteste solche Form der Menschheit ist nicht die Tabelle und nicht die Regel. Sie ist der Vers, der Rhythmus, die Melodie — das Lied, das man nicht zu behalten versucht und trotzdem nie wieder loswird. Warum das so ist, und was es für den Spracherwerb bedeutet, ist die Frage des nächsten Teils.

[45] Der Schluss von der Glossen-Evidenz (Fn. 5–7) auf die Klick-Erklärung ist eine Übertragung mit klaren Grenzen: Die Meta-Analyse betrifft *Lesen* mit vorbereiteten Glossen; der Klick betrifft *Video* mit erzeugten Erklärungen. Was trägt: muttersprachlich, bedürfnisnah, verarbeitungsaktiv. Was nicht trägt: eine Effektgröße für Pontivos konkrete Bauform — sie ist ungemessen (Kapitel 9).

[46] Zum Problem maschineller Auskünfte, deren Belege die Aussage tragen müssen: Streifler, D., *Kalibriertes Vertrauen*, ra.de-Verlag 2026 (Schwesterband dieser Reihe zu trust-on.net), dort insb. Kap. 1 zum *misgrounding*. Für Pontivos Klick-Erklärung folgt daraus eine Qualitätsschuld: Stichproben-Prüfung der erzeugten Erklärungen gegen fachliche Referenz, bevor der Klick als „Glosse im Sinne der Forschung" gelten darf.

Teil III — Die Melodie

„Wie wenig gehört zum Glücke! Der Ton eines Dudelsacks. — Ohne Musik wäre das Leben ein Irrthum."

— Friedrich Nietzsche, Götzen-Dämmerung, Sprüche und Pfeile 33 (1889)

Kapitel 5 · Warum das Lied im Kopf bleibt

5.1 Das Gedächtnis, das niemand bestellt hat

Fast jeder Mensch trägt Liedzeilen in Sprachen mit sich herum, die er nie gelernt hat. Wer in den 1980er Jahren jung war, kann bis heute italienische Refrains mitsingen, ohne Italienisch zu können; Kinder singen englische Popzeilen lautgetreu nach, Jahre bevor sie das erste englische Wort *verstehen*. Niemand hat diese Zeilen geübt, niemand wollte sie behalten — sie sind einfach geblieben. Der Sprachdidaktiker Tim Murphey hat für die harmloseste Form dieses Phänomens einen schönen Namen geprägt: das *song-stuck-in-my-head*-Phänomen, das Lied, das im Kopf weiterläuft, ob man will oder nicht.[47]

Für ein Buch über Spracherwerb ist das mehr als eine Kuriosität. Kapitel 4 endete mit der Frage, welche Form Input haben muss, damit er die Nacht übersteht. Hier liegt eine Antwort seit Jahrtausenden offen herum: Vor aller Schrift haben Menschen ihr Wissen in Versen, Rhythmen und Melodien überliefert — nicht aus ästhetischem Vergnügen, sondern weil gebundene Sprache die einzige war, die zuverlässig von Kopf zu Kopf kam. Die Frage ist, ob dieser alte Befund moderner Prüfung standhält: Bleibt gesungene Sprache wirklich besser haften als gesprochene — und wenn ja, unter welchen Bedingungen?

5.2 Was das Labor sagt: Die Melodie zerlegt den Strom

Die erste moderne Antwort kommt aus einem Labor in Marseille und knüpft unmittelbar an das Segmentierungsproblem aus Kapitel 3 an: Wo hört im fremden Lautstrom ein Wort auf, wo fängt das nächste an? Daniele Schön und Kollegen ließen Erwachsene einem künstlichen Sprachstrom zuhören — aneinandergereihte Silben ohne Pausen, aus denen sich „Wörter" nur über statistische Regelmäßigkeiten heraushören lassen. Wurde der Strom *gesprochen* dargeboten, blieben die Hörer nach sieben Minuten auf Zufallsniveau. Wurde derselbe Strom *gesungen* — jede Silbe konsequent an eine Tonhöhe gekoppelt —, erkannten sie die Wortgrenzen.[48] Das Fazit der Autoren: Gerade in der Anfangsphase, beim ersten Zerlegen einer neuen Sprache, wirken „die motivationalen und strukturierenden Eigenschaften der Musik" zusammen — die Melodie macht den Strom nicht nur angenehmer, sie *gliedert* ihn.[49]

Der zweite Baustein kommt aus Edinburgh und betrifft nicht das Hören, sondern das Sprechen. Karen Ludke, Fernanda Ferreira und Katie Overy ließen sechzig Erwachsene fünfzehn Minuten lang ungarische Alltagssätze lernen — nach dem Hör-und-Wiederhol-Verfahren, in drei Gruppen: sprechen, rhythmisch sprechen, singen. Die Sing-Gruppe schnitt insgesamt am besten ab, und ihr Vorsprung war statistisch signifikant genau bei den zwei Tests, die am schwersten wie-

---

[47] Murphey, T., The song stuck in my head phenomenon: A melodic din in the LAD?, *System* 18 (1990) 53–64 (am Druck zu prüfen); ders., *Music and Song* (Oxford 1992, am Druck zu prüfen). Murpheys Kernbeobachtung — Lieder laufen unwillkürlich im Kopf nach und liefern dabei Gratis-Wiederholung — ist hier nur als benanntes Phänomen übernommen, nicht als Wirkungsnachweis.

[48] Schön, D./Boyer, M./Moreno, S./Besson, M./Peretz, I./Kolinsky, R., Songs as an aid for language acquisition, *Cognition* 106 (2008) 975–983, doi:10.1016/j.cognition.2007.03.005 (Kernbefund am Abstract/PubMed verifiziert; Details der drei Experimente am Volltext zu prüfen): „a strong learning facilitation of song compared to speech". Die Sieben-Minuten-Fassung des Befunds (gesungene Silbenströme → Wortgrenzen erkannt, monotone Sprechfassung → Zufallsniveau) hier wiedergegeben nach der Darstellung bei Ludke et al. 2014 (Fn. 4, S. 41 f.).

[49] Schön et al. 2008 (Fn. 2), Abstract: der Anfangserwerb „may largely benefit of the motivational and structuring properties of music in song" (Übersetzung von uns).

gen: beim *Erinnern und Produzieren* der gesprochenen ungarischen Sätze.[50] Bemerkenswert ist, was den Unterschied *nicht* erklärte: weder Alter noch Stimmung, weder phonologisches Arbeitsgedächtnis noch musikalische Begabung oder Ausbildung.[51] Man musste kein Musiker sein — das Singen selbst tat die Arbeit. Die Autoren fassen ihre Deutung in eine Formel, die dieses Buch aufnimmt: Eine „Hör-und-Sing-Methode" kann das wortgetreue Gedächtnis für fremdsprachliche Sätze stützen.[52]

Warum das plausibel ist, lässt sich aus den vorigen Kapiteln zusammensetzen: Das Lied ist mehrfach kodierter Input im Sinne von Kapitel 3 — Text, Melodie, Rhythmus, oft noch Bewegung —, es erzwingt Wiederholung, ohne zu langweilen (kein Mensch liest einen Text zwanzigmal, aber jeder singt einen Refrain zwanzigmal), es verlangsamt und dehnt die Silben, macht also die Lautgestalt hörbarer als schnelles Sprechen, und es hängt Sprache an Emotion. Aber Plausibilität ist billig. Die Wahrheit ist komplizierter, und sie gehört in die Hauptlinie, nicht ins Kleingedruckte.

5.3 Ehrlich: Wann das Lied nicht hilft

Denn dieselbe Forschung kennt Befunde, die in die Gegenrichtung zeigen. Isabelle Peretz — eine der führenden Musikkognitionsforscherinnen — ließ mit Amélie Racette Erwachsene Liedzeilen lernen und fand: Beim Singen wurden *weniger* Wörter erinnert als beim Sprechen, und ob das Material gesungen oder gesprochen präsentiert wurde, machte für das Textgedächtnis keinen Unterschied.[53] Und Wanda Wallace hatte schon 1994 gezeigt, dass die Melodie den Text nur dann trägt, wenn sie über die Strophen hinweg *wiederkehrt* und leicht zu lernen ist; eine einmal gehörte oder ständig wechselnde Melodie hilft nicht — sie kostet.[54]

Widerlegt das den Teil III dieses Buches? Nein — es *präzisiert* ihn, und zwar genau dorthin, wo Pontivos Liederbuch gebaut ist. Liest man die Befunde nebeneinander, ergeben sie keine Kakophonie, sondern eine Bedingungsliste. Das Lied hilft, wenn die Melodie *einfach* ist und sich *wiederholt* — dann übernimmt sie die Gliederungsarbeit, die Wallace beschreibt, statt selbst Lernstoff zu sein; wo die Melodie schwer ist, wird sie zur Doppelaufgabe, und der Text verliert (das ist Racettes Befund: ihre Probanden mussten unbekannte Melodien *mit*lernen). Das Lied hilft ferner beim *wortgetreuen* Behalten von Wendungen und Sätzen — Ludkes Produktionstests —,

---

[50] Ludke, K. M./Ferreira, F./Overy, K., Singing can facilitate foreign language learning, *Memory & Cognition* 42 (2014) 41–52, hier S. 41, doi:10.3758/s13421-013-0342-5 (Fundstellen am Autoren-PDF verifiziert): 60 Erwachsene, Bedingungen Sprechen/rhythmisches Sprechen/Singen, 15 Minuten Ungarisch; „superior overall performance" der Sing-Gruppe, „statistically significant ($p < .05$) for the two tests that required participants to recall and produce spoken Hungarian phrases".

[51] Ebd., S. 41: „The differences in performance were not explained by potentially influencing factors such as age, gender, mood, phonological working memory ability, or musical ability and training."

[52] Ebd., S. 41: „a ‚listen-and-sing' learning method can facilitate verbatim memory for spoken foreign language phrases" (Übersetzung im Text von uns).

[53] Racette, A./Peretz, I., Learning lyrics: To sing or not to sing?, *Memory & Cognition* 35 (2007) 242–253 (Kernbefund am PubMed-Abstract verifiziert, Volltext am Druck zu prüfen): „Fewer words were recalled when singing than when speaking"; die Darbietungsform (gesungen/gesprochen) hatte keinen Einfluss auf das Textgedächtnis. Wichtig für die Einordnung: Gelernt wurden unvertraute Lieder — Text *und* Melodie waren neu, das Singen also eine Doppelaufgabe.

[54] Wallace, W. T., Memory for music: Effect of melody on recall of text, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20 (1994) 1471–1485 (Befund nach Abstract und Standardrezeption wiedergegeben, am Volltext zu prüfen): Melodie stützt Textabruf, wenn sie über Strophen wiederkehrt und leicht lernbar ist; bei einmaliger oder wechselnder Melodie kein Vorteil bzw. Nachteil. Wallaces Deutung — die Melodie „chunkt" den Text, markiert Zeilenlängen und Betonungsmuster — ist die strukturelle Erklärung des Schön-Befunds (Fn. 2) auf Liedebene.

nicht als Ersatz für das breite Verstehen, das der Film liefert. Und es hilft dem, der *mitsingt*, mehr als dem, der nur hört: Singen ist Aussprache-Übung mit eingebautem Metronom.[55] Kurz: Das Lied ist kein Wundermittel, sondern ein Spezialwerkzeug — das Verankerungswerkzeug für die Wendungen, die sitzen sollen.

5.4 Vom Labor zum Liederbuch

Genau als dieses Spezialwerkzeug ist das Liederbuch in Pontivo gebaut. Es setzt auf einfache, wiederholungsreiche Lieder — Volkslieder, Kinderlieder, Kanons, moderne Mitsinglieder —, also auf das Repertoire, das Wallaces Bedingung von selbst erfüllt: Strophenmelodien, die sich wiederholen, Refrains, die wiederkommen. Es zeigt Text und Akkorde zusammen (im ChordPro-Format, dem Standard der Mitsing-Praxis), damit eine Gitarre oder ein Klavier die Gruppe tragen kann. Und es ist — das ist der entscheidende Punkt, der in den nächsten Teil überleitet — kein Karaoke-Automat für Einzelne, sondern ein *Gruppenwerkzeug*: Der Lehrer- und Schüler-Modus lässt eine Person das Lied wählen, und alle singen dasselbe, zur selben Zeit, im selben Raum.[56]

Auch hier gilt die Ehrlichkeitsregel dieses Buches: Dass *dieses* Liederbuch, mit *diesen* Liedern, in *dieser* Einbettung den Laborvorsprung des Singens in messbaren Spracherwerb übersetzt, ist eine Design-Wette — die Laborstudien arbeiteten mit Minutenaufgaben und Einzelpersonen, nicht mit Monaten und Klassen.[57] Aber die Wette ist präzise platziert: Sie setzt auf den am besten belegten Fall (einfache, wiederholte Melodien; wortgetreues Behalten; aktives Mitsingen) und meidet den widerlegten (schwere Melodien, passives Hören, Lied als Allzweckersatz für Input).

Eines allerdings kann das Lied, was keine Laborstudie misst und kein Film ersetzt: Man kann es *gemeinsam* tun. Eine Klasse kann nicht gemeinsam einen Satz an der Supermarktkasse sagen — aber sie kann gemeinsam singen, vom ersten Tag an, auch die, die noch kein Wort sprechen. Was dabei mit einer Gruppe von Menschen geschieht — mit ihrer Angst, ihrer Bindung, ihrem Gefühl, dazuzugehören —, ist vielleicht der stärkste Einzelbefund dieses ganzen Buches. Er ist das Thema des nächsten Kapitels.

---

[55] Die Aussprache-Deutung (Singen dehnt Silben, erzwingt Artikulation im Takt) ist Interpretation dieses Buches auf Basis von Ludkes Produktionstest-Befund (Fn. 4) und der Prosodie-Literatur (statt vieler: Patel, A., *Music, Language, and the Brain*, Oxford 2008, am Druck zu prüfen); als eigenständiger Wirkungsnachweis für Aussprachetraining durch Singen wäre sie zu prüfen.

[56] Funktions-Hinweis: Das Liederbuch ist in Betrieb (pontivo.de/liederbuch, Stand 15. Oktober 2026): Texte mit Gitarren-Akkorden (ChordPro), Transponieren, Bühnen-Ansicht in großer Schrift, gemeinsames Singen in Gruppen — ohne Konto nutzbar.

[57] Übertragungs-Hinweis: Von 15-Minuten-Laborlernen (Fn. 4) und Kunstsprachströmen (Fn. 2) zu monatelangem Gruppen-Singen führt kein direkter Schluss; was trägt, sind Mechanismus (Segmentierung, Wiederholung, wortgetreues Behalten) und Bedingungsliste (einfach, wiederholt, aktiv). Die Feldwirkung des Liederbuchs gehört auf die Messliste (Kapitel 9).

Teil IV — Die Gemeinschaft

„Wo man singet, laß dich ruhig nieder, / Ohne Furcht, was man im Lande glaubt; / Wo man singet, wird kein Mensch beraubt; / Bösewichter haben keine Lieder."

— Johann Gottfried Seume, Die Gesänge (1804), erste Strophe

Kapitel 6 · Der Chor: Zugehörigkeit senkt die Hürde

6.1 Neunzig Minuten

Man stelle sich einen Dienstagabend in einem Gemeindesaal vor. Vierzehn Menschen, die sich nicht kennen, zwei davon erst seit Monaten im Land. Um sieben Uhr stehen sie verlegen zwischen Stuhlreihen. Um halb neun stehen sie im Halbkreis, haben dreimal denselben Kanon verpatzt und zweimal gelacht, und als jemand vorschlägt, nächste Woche zehn Minuten früher anzufangen, nicken alle. Irgendetwas ist in diesen neunzig Minuten passiert, das in Sprachkursen oft in Monaten nicht passiert.

Dieses Etwas ist messbar. Ein Forscherteam um den Evolutionspsychologen Robin Dunbar begleitete über sieben Monate neu gegründete Abendkurse für Erwachsene — Singklassen und, als Vergleich, Bastel- und Schreibkurse — und ließ die Teilnehmer zu drei Zeitpunkten einschätzen, wie nah sie sich ihrer Gruppe fühlten. Das Ergebnis: Nach dem ersten Monat zeigten die Sänger einen signifikant größeren Zuwachs an empfundener Nähe als alle Vergleichsgruppen; erst nach Monaten zogen die anderen nach.[58] Die Forscher tauften den Befund auf einen Namen, der hängen bleibt: den *Eisbrecher-Effekt* — Singen stellt schnelle soziale Bindung zwischen Fremden her, ohne dass es dazu langer persönlicher Gespräche bedürfte.[59] Ihre evolutionäre Deutung: Gesang ist vermutlich eines der ältesten Werkzeuge der Menschheit, um Gruppen zu verbinden, die zu groß sind, als dass sich jeder mit jedem einzeln anfreunden könnte.[60] Der Chor am Dienstagabend führt kein Sozialexperiment durch. Er benutzt eine sehr alte Maschine.

6.2 Was das gemeinsame Singen mit Menschen macht

Der Eisbrecher-Befund steht nicht allein; um ihn herum liegt ein breites Feld von Forschung zum Chorsingen und Wohlbefinden. Sänger eines Laienchors, vor und nach der Probe untersucht, zeigten nach dem *Singen* mehr positive Stimmung, weniger negative — und einen Anstieg des Immunglobulins A im Speichel, eines Markers der Immunabwehr; das bloße *Anhören* von Chormusik in derselben Gruppe bewirkte das nicht.[61] Befragt man Chormitglieder selbst, ist das Bild einhellig: In einer der frühen systematischen Erhebungen gaben 87 Prozent an, sozial vom Singen zu profitieren, 75 Prozent emotional.[62] Man muss diese Literatur nüchtern einordnen —

---

[58] Pearce, E./Launay, J./Dunbar, R. I. M., The ice-breaker effect: singing mediates fast social bonding, *Royal Society Open Science* 2 (2015) 150221, doi:10.1098/rsos.150221 (Abstract-Wortlaut am PubMed-Eintrag verifiziert): „we followed newly formed singing and non-singing (crafts or creative writing) adult education classes over seven months"; Messzeitpunkte Monat 1, 3 und 7; „Singers demonstrated a significantly greater increase in closeness at timepoint 1".

[59] Ebd.: „Although singers and non-singers felt equally connected by timepoint 3, singers experienced much faster bonding" — der Begriff „ice-breaker effect" stammt aus dem Titel der Studie. Ehrlichkeitshalber gehört dazu: Nach sieben Monaten waren die Gruppen gleichauf; der Chor-Vorteil ist ein *Geschwindigkeits*-, kein Endstandsvorteil. Für Lernkontexte, in denen die ersten Wochen über Abbruch oder Dranbleiben entscheiden, ist genau die Geschwindigkeit die relevante Größe (dazu 6.3).

[60] Zur evolutionären Rahmung (Gesang als Bindungsmechanismus für große Gruppen, „social bonding hypothesis") die Diskussion bei Pearce et al. (Fn. 1) sowie Dunbars Arbeiten zu Gruppengrößen und sozialer Fellpflege (statt vieler: Dunbar, *Grooming, Gossip, and the Evolution of Language*, Cambridge/MA 1996, am Druck zu prüfen). Die Evolutionsdeutung ist Hintergrundtheorie, kein tragender Beleg dieses Kapitels.

[61] Kreutz, G./Bongard, S./Rohrmann, S./Hodapp, V./Grebe, D., Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state, *Journal of Behavioral Medicine* 27 (2004) 623–635 (Abstract am PubMed-Eintrag verifiziert): Mitglieder eines Laienchors, zwei Proben im Wochenabstand, Bedingung Singen vs. Anhören von Chormusik; „Singing leads to increases in positive affect and S-IgA, while negative affect is reduced."

[62] Clift, S./Hancox, G., The perceived benefits of singing: findings from preliminary surveys of a university college choral society, *Journal of the Royal Society for the Promotion of Health* 121 (2001) 248–256 (Abstract am PubMed-

vieles beruht auf Selbstauskünften von Menschen, die freiwillig in Chören singen, die physiologischen Studien sind klein.[63] Aber die Konvergenz ist bemerkenswert: Quer durch Methoden und Länder findet sich dasselbe Muster aus gehobener Stimmung, gesenktem Stress und — vor allem — sozialer Verbundenheit.

Für die Frage dieses Buches ist ein dritter Befund der wichtigste. Die größte Untersuchung zum Singen von Kindern, die Evaluation des britischen Nationalprogramms „Sing Up" mit Daten von über sechstausend Kindern, fand einen klaren Zusammenhang: Je weiter die Singentwicklung eines Kindes, desto positiver sein Selbstbild und sein Gefühl, *sozial dazuzugehören* — unabhängig von Alter, Geschlecht und Herkunft.[64] Die Autoren mahnen selbst zur Vorsicht: Das ist eine Korrelation, keine nachgewiesene Wirkrichtung; wer sich zugehörig fühlt, singt vielleicht auch lieber.[65] Aber genau diese Studie hält fest, worauf es hier ankommt: Singen und Zugehörigkeit hängen aneinander, und zwar unabhängig davon, woher ein Kind kommt.

6.3 Die geschützte Stimme

Nun könnte man einwenden: Schön — aber was hat das mit *Spracherwerb* zu tun? Die Antwort führt zurück zu Kapitel 2, und sie hat zwei Stufen.

Die erste Stufe ist die Angst. Kein Moment des Sprachenlernens ist so angstbesetzt wie der, in dem man vor anderen den Mund aufmacht: Man klingt wie ein Kind, man wird bewertet, und das Selbstbild — erwachsen, kompetent, ernst zu nehmen — steht auf dem Spiel; die Fremdsprachenangst-Forschung hat gezeigt, dass gerade das Sprechen-vor-anderen ihr Zentrum ist und dass sie Leistung messbar drückt.[66] Der Chor ist die einzige soziale Form, die dieses Problem *konstruktiv* löst statt es zu verbieten: Die einzelne Stimme ist in der Menge unhörbar. Wer im Chor einen Fehler singt, blamiert sich nicht — es merkt buchstäblich niemand. Damit wird etwas möglich, das es sonst nirgends gibt: **laute, aktive Produktion in der Zielsprache vom ersten Tag an, mit einem Blamage-Risiko von null.** Eine Klasse kann nicht im Chor einen Behördenbrief formulieren; aber sie kann im Chor singen, auch die, die noch kein einziges Wort sprechen. Der Chor ist, in der Sprache von Kapitel 2, ein gebauter Filter-Senker: Er stellt die Bedingung her, unter der Input überhaupt ankommt und Produktion sich überhaupt traut.

Die Organisationspsychologie hat für diese Bedingung einen präzisen Namen. Amy Edmondson nannte sie *psychologische Sicherheit*: die geteilte Überzeugung einer Gruppe, dass zwischenmenschliches Risiko — fragen, irren, Nichtwissen zeigen — hier nicht bestraft wird; in ihrer

---

Eintrag verifiziert): n = 84 bzw. 91; 87 % soziale, 75 % emotionale Benefit-Zustimmung; sechs Nutzen-Dimensionen.

[63] Zur gebotenen Nüchternheit: Selbstselektion (wer Chöre mag, singt in Chören), Selbstauskunfts-Formate und kleine physiologische Stichproben begrenzen die Beweiskraft der Wohlbefindens-Literatur; die Bindungsbefunde (Fn. 1) stehen methodisch stärker da (Längsschnitt, Vergleichsgruppen). Dieses Buch stützt seine Kette (6.3) deshalb primär auf Bindung und Angst, nicht auf Gesundheitseffekte.

[64] Welch, G. F./Himonides, E./Saunders, J./Papageorgi, I./Sarazin, M., Singing and social inclusion, *Frontiers in Psychology* 5 (2014) 803, doi:10.3389/fpsyg.2014.00803 (Open Access, Fundstellen am PMC-Volltext verifiziert): 6.087 Kinder mit gepaarten Daten (aus 11.258 untersuchten): „The higher the normalized singing development rating, the more positive the child's self-concept and sense of being socially included, irrespective of singer age, sex and ethnicity."

[65] Ebd.: „Whilst caution is needed in the interpretation of the correlation and regression data, and in the pitfalls of ascription of causal effects between measures …" — die Autoren selbst lassen die Wirkrichtung offen; dieses Buch übernimmt den Befund entsprechend nur als Zusammenhang, nicht als Nachweis.

[66] Horwitz/Horwitz/Cope 1986 (Kapitel 2, Fn. 9), S. 125 ff.; zur Sonderstellung des Sprechens ebd. Die FLCAS-Items kreisen wesentlich um Sprechsituationen vor der Klasse.

Studie an einundfünfzig Arbeitsteams war genau diese Überzeugung die Voraussetzung dafür, dass Menschen lernen, statt sich abzusichern.[67] Der Chor muss diese Sicherheit nicht durch Appelle herstellen; er baut sie in seine Form ein: Die Menge verbirgt die Stimme, also existiert das Risiko, vor dem man geschützt werden müsste, gar nicht erst.

Die zweite Stufe ist das Dranbleiben. Spracherwerb braucht, das war die Lektion der Teile II und III, vor allem eines: Menge — Monate von Input, Hunderte von Begegnungen mit jedem Wort. Menge aber ist keine didaktische Größe, sondern eine motivationale: Sie entsteht nur, wo Menschen *wiederkommen*. Und Menschen kommen dorthin wieder, wo sie dazugehören. Der Eisbrecher-Effekt ist deshalb kein Wellness-Bonus am Rande der Methode, sondern ihr Fundament: Die Bindung der ersten Wochen entscheidet darüber, ob es die Monate überhaupt gibt, in denen die anderen Mechanismen wirken können.

Beides zusammen ist die Kette, die dieses Buch behauptet — und sie soll hier so offen wie möglich liegen: Die Chorforschung belegt Bindung, Stimmung und Zugehörigkeit; sie belegt *nicht* direkt, dass im Chor Sprache erworben wird. Der Schluss vom Chor auf den Spracherwerb läuft über zwei belegte Zwischenglieder — Angst hemmt Erwerb, Zugehörigkeit trägt Dranbleiben — und ist als Kette plausibel, aber als Ganzes ungemessen.[68] Kapitel 9 wird genau diese Kette auf die Messliste setzen.

6.4 Für die, die ankommen müssen

Für eine Gruppe ist diese Kette keine akademische Figur, sondern der Alltag: für Menschen, die neu in einem Land sind. Eine Geflüchtete wie Amal aus Kapitel 1 hat nicht nur ein Sprachproblem; sie hat ein Zugehörigkeitsproblem, und beide verstärken sich gegenseitig — wer nicht sprechen kann, findet schwer Anschluss, und wer keinen Anschluss hat, hat niemanden, mit dem er sprechen könnte. Der Sprachkurs unterbricht diesen Kreislauf nicht: Dort sitzt sie ausschließlich unter Menschen, die dasselbe Problem haben.

Gemeinsames Singen unterbricht ihn an der einzigen Stelle, die ohne Sprache erreichbar ist. Im Chor neben Einheimischen zu stehen setzt kein Sprachniveau voraus — die Melodie trägt, der Text steht da, die Stimme verschwindet in der Menge. Zugehörigkeit entsteht hier *vor* der Sprache, nicht als ihre Belohnung. Die Community-Music-Forschung, die musikalische Gruppen aus Geflüchteten und Einheimischen begleitet hat, beschreibt durchgängig dieses Muster: Wohlbefinden, Verständigung über Kulturgrenzen und ein Gefühl des Ankommens, lange bevor die gemeinsame Sprache trägt.[69] Es ist der seltene Fall einer Integrationsmaßnahme, bei der die Auf-

---

[67] Edmondson, A., Psychological Safety and Learning Behavior in Work Teams, *Administrative Science Quarterly* 44 (1999) 350–383: psychologische Sicherheit als „a shared belief held by members of a team that the team is safe for interpersonal risk taking" (Definition und Studiendesign — 51 Arbeitsteams eines Fertigungsunternehmens — am frei zugänglichen Volltext-Spiegel verifiziert; am Druck zu prüfen). Übertragungs-Hinweis: Die Befunde stammen aus Arbeitsteams, nicht aus Lerngruppen; übernommen wird das Konstrukt und sein Mechanismus (Sicherheit → Lernverhalten), nicht eine Effektgröße für Chöre.

[68] Übertragungs-Hinweis in aller Form: Belegt sind (a) schnelle Bindung durch Gruppensingen (Fn. 1 f.), (b) Zusammenhang Singen–Zugehörigkeit (Fn. 7), (c) leistungsdrückende Wirkung der Sprechangst (Fn. 9, Kapitel 2 Fn. 10) und (d) die Notwendigkeit großer Inputmengen (Teil II). NICHT direkt belegt ist der Gesamtschluss „Chor-Modul → besserer Spracherwerb"; er ist eine zweistufige Inferenz über Filter und Dranbleiben und wird in Kapitel 9 als Messaufgabe formuliert (Angst-, Zugehörigkeits- und Verbleibsmaße im Chor-Modus).

[69] Statt vieler: Dow, R./Vougioukalou, S., Wellbeing and Integration through Community Music: The Role of Improvisation in a Music Group of Refugees, Asylum Seekers and Local Community Members, *Contemporary Music Review* 38 (2019), doi:10.1080/07494467.2019.1684075 (am Volltext zu prüfen; Verlagszugriff gesperrt). Diese Literatur ist

nahmegesellschaft nicht Prüfer ist, sondern Mitsänger. Die Soziologie kennt für solche Räume einen eigenen Begriff: den *dritten Ort* — weder Zuhause noch Arbeit oder Behörde, sondern der neutrale Boden, auf dem Status draußen bleibt und Gespräch (hier: Gesang) die einzige Eintrittskarte ist.[70] Der Chor ist der dritte Ort, den eine Ankommende betreten kann, bevor sie die Sprache der ersten beiden beherrscht.

6.5 Vom Befund zum Bau

Pontivos Chor-Modul übersetzt diese Befunde in drei Bauentscheidungen. Erstens der *Lehrer-/Schüler-Modus*: Eine Person — Chorleiterin, Lehrer, Gruppengründer — wählt das Lied, und bei allen anderen erscheint dasselbe, im selben Moment, mit Text und Akkorden. Das ist technisch unscheinbar und sozial entscheidend: Es erzeugt den *einen gemeinsamen Gegenstand*, auf den alle zugleich schauen — die Grundform geteilter Aufmerksamkeit, von der das nächste Kapitel handelt. Zweitens die *geschlossene Gruppe*: Chöre, Klassen und Familien singen in privaten Räumen, die niemand von außen betritt. Was in der Gruppe passiert — der schiefe Ton, der Fehlversuch, das Lied, das jemandem etwas bedeutet —, bleibt in der Gruppe; der Schutzraum, der die Einzelstimme im Klang verbirgt, wird auf die soziale Ebene verlängert. Drittens das *Repertoire*: einfache, wiederholungsreiche Lieder mit Akkorden (Kapitel 5), darunter — und das wird sich in Teil V als der dichteste Baustein des ganzen Konzepts erweisen — Lieder, welche die Mitglieder der Gruppe selbst mitbringen.[71]

Ob diese gebaute Fassung die Wirkung der Gemeindesaal-Chöre erreicht, ist — Ehrlichkeitsregel — offen: Die zitierte Forschung hat physisch versammelte Gruppen untersucht, keine App-gestützten.[72] Aber die Bauentscheidungen sind keine Erfindungen, sondern Übersetzungen: Jede von ihnen hat einen benennbaren Befund hinter sich.

Bleibt die Frage, wer die Gruppe führt — und was aus der Muttersprache wird, wenn eine Lehrkraft eine Klasse durch Film und Lied steuert. Das ist das Klassenzimmer, und es hat seine eigene, hundert Jahre alte Theorie.

überwiegend qualitativ (Interviews, Beobachtung) und wird hier als konvergentes Muster, nicht als Effektnachweis geführt.

[70] Oldenburg, R., *The Great Good Place*, New York 1989 (am Druck zu prüfen): der „dritte Ort" als neutraler, statusnivellierender Gemeinschaftsraum neben Zuhause (erster) und Arbeit (zweiter Ort) — entwickelt an Cafés, Kneipen, Friseurläden. Übertragungs-Hinweis: Oldenburgs Kriterien (Neutralität, Niedrigschwelligkeit, Gespräch als Hauptaktivität, Stammpublikum) passen auf den Chor in bemerkenswertem Maß — mit der Abweichung, dass an die Stelle des Gesprächs der Gesang tritt; als soziologische Einordnung verwendet, nicht als Beleg.

[71] Funktions-Hinweis: Das Chor-Modul ist in Betrieb (pontivo.de/choere, Stand 15. Oktober 2026): geschlossene Gruppen, Lehrer-/Schüler-Modus, Mitlesen als Gast per Einladungscode ohne Konto.

[72] Zusätzlich offen und der Redlichkeit halber benannt: Ob synchron vermitteltes gemeinsames Singen über eine Plattform (gleicher Text, gleicher Moment, gegebenenfalls verteilte Räume) die Bindungseffekte physischer Chöre trägt, ist unerforscht; die Latenz-Physik verbietet verteiltes Echtzeit-Singen ohnehin — Pontivos Modul ist deshalb primär für den *gemeinsamen Raum* gebaut (Klasse, Chorprobe, Familie), die Plattform liefert Text, Akkorde und Gleichzeitigkeit, nicht den Raum.

Kapitel 7 · Das Klassenzimmer: geteilte Aufmerksamkeit, legitimierte Muttersprache

7.1 Das Verbot

In Amals Integrationskurs galt eine Regel, die so selbstverständlich wirkte, dass niemand sie begründete: *Hier wird nur Deutsch gesprochen.* Deutsch wurde auf Deutsch erklärt, an Menschen, die noch kein Deutsch konnten; das Arabische, das Amal mitbrachte — eine voll ausgebaute Sprache, in der sie denken, argumentieren, ironisch sein kann —, war im Kursraum nicht Werkzeug, sondern Störfall. Dieses Verbot hat eine Geschichte, und sie ist kürzer und profaner, als ihr Ansehen vermuten lässt: Die „direkte Methode" der Sprachschulen des frühen zwanzigsten Jahrhunderts machte die Einsprachigkeit zum Prinzip — nicht zuletzt, weil sie für Schulketten mit internationalem Lehrpersonal praktisch war: Wer nur in der Zielsprache unterrichtet, muss die Sprachen seiner Schüler nicht können.[73] Aus der Not wurde eine Doktrin, aus der Doktrin ein Reflex. Wissenschaftlich begründet war er nie besonders gut.

Dieses Kapitel handelt davon, was ein Klassenzimmer eigentlich ist, wenn man es von der Lernforschung her denkt — und warum die Muttersprache darin nicht der Feind der Zielsprache ist, sondern ihr wichtigstes Gerüst.

7.2 Das Gerüst

Die theoretische Grundlage stammt aus den 1930er Jahren, von Lew Vygotskij, und sie hat einen einfachen Kern: Lernen ist zuerst ein sozialer Vorgang und wird erst danach ein innerer. Was ein Kind heute nur an der Hand eines Könnenden schafft, schafft es morgen allein; die Spanne dazwischen — zwischen dem, was einer selbstständig kann, und dem, was er „unter Anleitung oder in Zusammenarbeit mit fähigeren Partnern" kann — nannte Vygotskij die *Zone der nächsten Entwicklung*.[74] Lernen findet in dieser Zone statt, und die Aufgabe des Lehrenden ist nicht, Wissen hinüberzureichen, sondern die Zone zu bewirtschaften: Aufgaben so zu stellen, dass sie einen Schritt über dem Alleingekonnten liegen, und genau so viel Halt zu geben, dass der Schritt gelingt. Für diesen Halt haben Wood, Bruner und Ross das Bild geprägt, das seither durch alle Didaktiken wandert: das *Gerüst* (scaffolding) — die Stützstruktur, die der Erfahrenere errichtet und Stück für Stück abbaut, während der Lernende trägt, was er schon tragen kann.[75]

Man beachte, wie genau sich das mit Krashens *i + 1* aus Kapitel 2 deckt, das aus einer ganz anderen Tradition stammt: Beide sagen, dass Lernen an der Kante des Könnens geschieht, einen

---

[73] Zur Geschichte der direkten Methode und des Einsprachigkeits-Prinzips (Berlitz-Schulen, Reformbewegung): Richards, J. C./Rodgers, T. S., *Approaches and Methods in Language Teaching*, 3. Aufl., Cambridge 2014, Kap. zur Direct Method (am Druck zu prüfen); kritisch zur wissenschaftlichen Basis des Dogmas Butzkamm/Caldwell (Fn. 4), Kap. 1. Die verkürzte Kausalgeschichte im Text (Personalpraxis → Prinzip) folgt Butzkamms Darstellung und ist als solche zu prüfen.

[74] Vygotskij, L. S., *Mind in Society: The Development of Higher Psychological Processes*, Cambridge/MA 1978, S. 86: die Zone der nächsten Entwicklung als „the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (Übersetzung im Text von uns; Seitenangabe nach der Standardausgabe, am Druck zu prüfen).

[75] Wood, D./Bruner, J. S./Ross, G., The role of tutoring in problem solving, *Journal of Child Psychology and Psychiatry* 17 (1976) 89–100 (am Druck zu prüfen; Verlagszugriff gesperrt). Der Begriff „scaffolding" für die abbaubare Stützfunktion des Tutors stammt aus diesem Aufsatz; die Wiedergabe folgt der Standardrezeption.

Schritt über dem Stand — und dass es *Hilfe beim Verstehen* ist, nicht Erklärung der Form, die über diese Kante trägt. Der zweisprachige Untertitel aus Kapitel 3 ist nichts anderes als ein Gerüst im Wortsinn: eine Stütze, die das Verstehen hält, wo die eigene Sprachkraft noch nicht reicht, und die man abbauen kann — Zeile für Zeile —, wenn sie nicht mehr gebraucht wird.

7.3 Die Muttersprache als Gerüst

Womit aber baut man Gerüste? Die stabilste Stütze, die ein erwachsener Lerner besitzt, ist die Sprache, die er schon kann. Wolfgang Butzkamm — der deutsche Anglist, der gegen das Einsprachigkeits-Dogma ein Forscherleben lang angeschrieben hat — hat es in die schlichteste Form gebracht: Die Muttersprache ist nicht der Störfaktor des Fremdsprachenlernens, sondern sein größtes Kapital; in ihr haben wir gelernt, was Sprechen überhaupt ist — dass Laute bedeuten, dass Sätze gebaut werden, dass man mit Worten handelt —, und dieses Wissen wandert in jede weitere Sprache mit.[76] Die neuere Mehrsprachigkeitsforschung hat denselben Gedanken unter dem Begriff *Translanguaging* entfaltet: Mehrsprachige besitzen nicht zwei getrennte Sprachfächer, sondern ein gemeinsames Repertoire, aus dem sie situationsgerecht schöpfen; ein Unterricht, der die Erstsprache aussperrt, sperrt einen Teil des Denkwerkzeugs seiner Schüler aus.[77]

Das ist keine bloße Theorieposition; die empirischen Spuren liegen quer durch dieses Buch verstreut, und es lohnt, sie einmal einzusammeln. Die Glossen-Meta-Analyse aus Kapitel 4: Worterklärungen in der *Muttersprache* schlagen Erklärungen in der Zielsprache, auch bei Fortgeschrittenen.[78] Die Blickbewegungsstudie aus Kapitel 3: Die muttersprachliche Untertitelzeile ist der stabile Anker, von dem aus die Lerner die Zielsprachzeile erkunden — die Bedeutung ruht, die Form wird erobert.[79] Und die Bücherflut aus Kapitel 2 wirkte in Schulen, deren Kinder zu Hause Fidschianisch oder Hindi sprachen — niemand hat ihnen das ausgetrieben, das Englische kam *dazu*.[80] Wo immer man die Muttersprache als Werkzeug zulässt statt als Leck behandelt, arbeitet sie mit.

Damit kein Missverständnis entsteht: Nichts davon spricht gegen möglichst viel Zielsprache im Raum — die Menge verständlichen Inputs bleibt die erste Bedingung von allen. Es spricht gegen das *Verbot* der Erstsprache: gegen den Kurs, der Verstehenshilfe verweigert, die einen Halbsatz kosten würde, und gegen die Beschämung derer, die in ihrer eigenen Sprache denken. Das Gerüst steht nicht im Weg des Gebäudes. Es macht es möglich.

Wer das für eine Außenseiterposition hält, unterschätzt, wie weit die Sprachenpolitik hier schon ist: Der Europarat hat die Mehrsprachigkeit im Begleitband zum Referenzrahmen (2020) aus-

---

[76] Butzkamm, W./Caldwell, J. A. W., *The Bilingual Reform: A Paradigm Shift in Foreign Language Teaching*, Tübingen 2009 (am Druck zu prüfen); Kernthese: „The mother tongue is the greatest asset people bring to the task of foreign language learning" (Standardformulierung der Autoren, Fundstelle am Druck zu prüfen). Deutschsprachig: Butzkamm, *Psycholinguistik des Fremdsprachenunterrichts*, 3. Aufl., Tübingen 2002 (am Druck zu prüfen).

[77] García, O., *Bilingual Education in the 21st Century: A Global Perspective*, Malden 2009 (am Druck zu prüfen); García, O./Wei, L., *Translanguaging: Language, Bilingualism and Education*, Basingstoke 2014 (am Druck zu prüfen). Übertragungs-Hinweis: Translanguaging ist als Beschreibung mehrsprachiger Praxis und als pädagogische Haltung gut ausgearbeitet, aber kein experimentell quantifiziertes Wirkprinzip; die quantitativen Stützen dieses Kapitels sind die in Fn. 6–8 versammelten Einzelbefunde.

[78] Yanagisawa/Webb/Uchihara 2020 (Kapitel 4, Fn. 5 f.): „L1 glosses yielded greater learning than L2 glosses", ohne Interaktion mit dem Sprachstand.

[79] Liao/Kruger/Doherty 2020 (Kapitel 3, Fn. 13): stabilere Blickzuwendung zur L1-Zeile, flexible Exploration der L2-Zeile, keine Mehrbelastung durch die Doppelspur.

[80] Elley/Mangubhai (Kapitel 2, Fn. 5 f.); zum Übergreifen der Effekte auf die Erstsprache (Fidschianisch) im zweiten Jahr die dortige Autorendarstellung (Fiji Intermediate Examination, Mittelwert 58 % auch im Fach Fijian Language).

drücklich zur pädagogischen Leitlinie erhoben — plurilinguale Kompetenz gilt dort nicht als Summe getrennter Einsprachigkeiten, sondern als ein zusammenhängendes Repertoire, das der Unterricht nutzen soll statt zu unterdrücken.[81] Die Erstsprache einzubinden ist damit kein didaktisches Wagnis mehr, sondern die Richtung, in die das europäische Rahmenwerk selbst zeigt.

7.4 Das gebaute Klassenzimmer

Pontivos Klassenzimmer setzt diese beiden Ideen — geteilte Aufmerksamkeit unter Führung, Muttersprache als Gerüst — technisch um, und an einer Stelle tut es etwas, das ein physisches Klassenzimmer nicht kann.

Der *Lehrer-/Schüler-Modus* ist die gebaute Form der geteilten Aufmerksamkeit: Die Lehrkraft wählt — den Film, die Szene, das Lied —, und bei allen erscheint dasselbe im selben Moment. Das klingt banal und ist der Kern jeder Unterrichtsform seit der Dorfschule: ein Gegenstand, viele Blicke, eine führende Hand. Die Plattform nimmt der Lehrkraft dabei nicht die Führung ab, sondern die Reibung — das Verteilen, das „Seite 34, zweiter Absatz", das „bei mir läuft es nicht".[82]

Das Besondere liegt in der zweiten Ebene: Derselbe Film läuft für jeden Schüler *mit seiner eigenen Muttersprachzeile*. Eine Klasse, in der Arabisch, Ukrainisch, Tigrinya und Polnisch zusammensitzen — der Normalfall jeder Integrationsklasse —, schaut denselben deutschen Film, und jedes Kind hat sein eigenes Gerüst darunter. Im physischen Klassenzimmer ist das unmöglich; kein Lehrer kann in vier Sprachen zugleich flüstern. Hier ist es eine Zeile Konfiguration: Die Engine trennt Zielsprache, Muttersprache und Bedienoberfläche als unabhängige Achsen.[83] Das ist Translanguaging als Infrastruktur: Die Klasse teilt den Gegenstand und die Zielsprache — und jeder behält sein eigenes Werkzeug.

Und die Lehrkraft? Sie wird nicht ersetzt, sie wird *entlastet* — von genau dem Teil ihrer Arbeit, den die Maschine kann (Verstehenssicherung in vier Sprachen gleichzeitig), für den Teil, den nur sie kann: auswählen, deuten, ermutigen, das Gespräch nach dem Film führen, die Aufgabe stellen, die aus dem Gesehenen Sprache macht. Die Bildungsforschung gibt dieser Arbeitsteilung recht: In Hatties großer Forschungssynthese gehört die Qualität der Beziehung zwischen Lehrkraft und Lernenden zu den stärksten bekannten Einflussfaktoren auf den Lernerfolg — Zeit, die eine Maschine der Beziehungsarbeit zurückgibt, ist also in die wirksamste Größe des Unterrichts investiert.[84] Ein Werkzeug, das Lehrer überflüssig machen soll, wäre auch pädagogisch falsch

---

[81] Council of Europe, *Common European Framework of Reference for Languages: Learning, teaching, assessment — Companion volume*, Straßburg 2020: plurilinguale/plurikulturelle Kompetenz als Schlüsselkonzept und Leitlinie der Sprachenbildung (Kernaussage an den offiziellen Europarats-Seiten verifiziert; Pinpoints am PDF-Volltext zu prüfen — der Direktzugriff auf rm.coe.int war elektronisch gesperrt). Bewusst als „Leitlinie" bezeichnet: Der Referenzrahmen ist ein Rahmenwerk, keine Rechtsnorm.

[82] Funktions-Hinweis: Der Klassenraum ist in Betrieb (pontivo.de/fuer-lehrer, Stand 15. Oktober 2026): Klasse mit Einladungscode, Hausaufgaben, Klassenchat, gemeinsames Sync-Kino; Anzeige auf Beamer oder TV über pontivo.de/tv.

[83] Die Achsen-Architektur (Zielsprache / Muttersprache / Oberflächensprache je Nutzer unabhängig) ist Produktbeschreibung, kein Forschungsbefund; sie wird hier referiert, weil sie die technische Bedingung der Mehrsprachen-Klasse ist. Ob die Vielsprachen-Klasse mit individueller Gerüstzeile den erwarteten Vorteil gegenüber einer Einheitszeile hat, ist ungemessen (Kapitel 9).

[84] Hattie, J., *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*, London 2009 (am Druck zu prüfen): Lehrkraft-Lernenden-Beziehung mit d ≈ 0,72 unter den stärksten Einflussfaktoren — ausdrücklich nicht „der stärkste": In Hatties späteren Rangreihen liegen andere Faktoren darüber (etwa kollektive Lehrer-Wirksamkeit, d ≈ 1,57; Werte nach den vom Autor betriebenen Übersichten, am Druck zu prüfen). Zudem ist Hatties

gebaut: Das Gerüst-Prinzip *braucht* den Könnenden, der weiß, wann er stützt und wann er loslässt.[85]

Zur Redlichkeit gehört an dieser Stelle auch eine Absichtserklärung, die als solche gekennzeichnet ist: Auf Dauer soll die Praxis die Plattform mitentwickeln — Lehrkräfte, die eine ungenaue Klick-Erklärung melden oder eine gelungene Liedübersetzung in den gemeinsamen Bestand geben, nach dem Vorbild partizipativer Aktionsforschung.[86] Gebaut ist diese Rückkopplungsschleife noch nicht; sie steht hier als Bauplan, nicht als Behauptung.

7.5 Was dem Klassenzimmer noch fehlt

Ein Klassenzimmer, das gemeinsam schaut und gemeinsam singt, hat viel — aber es hat noch keinen Stoff, der *aus der Klasse selbst* kommt. Bisher fließt alles Material in eine Richtung: vom Katalog zur Gruppe. Der dichteste Baustein des ganzen Konzepts dreht diese Richtung um: Er macht das, was die Lernenden mitbringen — die Lieder ihrer Herkunft —, zum Unterrichtsmaterial, und die Lernenden selbst zu denen, die es erklären. Was dabei geschieht, sprachlich, sozial und menschlich, ist das nächste Kapitel — und mit ihm ist dieses Buch bei seiner Pointe.

Methodik (Aggregation heterogener Meta-Analysen) substanziell kritisiert worden — die Größenordnung „Beziehung wirkt stark" gilt gleichwohl als robust; mehr wird hier nicht behauptet.

[85] Zur Aufgabenorientierung (die Aufgabe nach dem Input, die aus Verstehen Sprache macht): Ellis, R., *Task-based Language Learning and Teaching*, Oxford 2003; Long, M., *Second Language Acquisition and Task-Based Language Teaching*, Malden 2015 (beide am Druck zu prüfen). Pontivo liefert bewusst den Input- und Gemeinschaftsteil und überlässt die Aufgabenebene der Lehrkraft — eine Arbeitsteilung, keine Vollständigkeitsbehauptung.

[86] Zum methodischen Vorbild: Kemmis, S./McTaggart, R./Nixon, R., *The Action Research Planner: Doing Critical Participatory Action Research*, Singapur 2014 (am Druck zu prüfen). Der Status ist im Text ausgewiesen: erklärte Bauabsicht, kein vorhandenes Feature — die Rückkopplungsschleife (Glossen-Korrektur, Strophen-Einspeisung) existiert in der Plattform derzeit nicht.

Teil V — Die Probe

„Ihre Gesänge sind das Archiv des Volks, der Schaz ihrer Wissenschaft und Religion … Abdruck ihres Herzens, Bild ihres häuslichen Lebens in Freude und Leid, beym Brautbett und Grabe."

— Johann Gottfried Herder, Deutsches Museum 1777, S. 433

Kapitel 8 · Coro One Sky: Das Lied aus der Heimat, gesungen in der neuen Sprache

8.1 Die Übung

Die Übung, von der dieses Kapitel handelt, ist in einem Satz beschrieben: Ein Mitglied des Chors bringt ein Lied aus seiner Heimat mit; die Gruppe übersetzt es gemeinsam in die neue Sprache; dann singen es alle — auf Italienisch, auf Deutsch — auf die alte Melodie.

Man muss sich den Vorgang konkret vorstellen, um zu sehen, was in ihm steckt. Ein syrischer Teenager summt die Melodie, die seine Mutter gesungen hat; er schreibt die arabischen Zeilen auf und erklärt der Gruppe, Zeile für Zeile, was sie bedeuten — was sie *wirklich* bedeuten, denn die Wörterbuchbedeutung ist ihm zu wenig. Die Gruppe schlägt italienische Fassungen vor, verwirft, feilt: Diese Zeile ist zu lang für die Melodie, dieses Wort trifft den Ton nicht, hier muss die Betonung auf die zweite Silbe. Der Lehrer hilft mit Grammatik, wo es klemmt. Nach zwei Abenden steht eine singbare italienische Strophe. Dann steht der Chor auf und singt sie — und der Junge, dessen Lied es ist, hört zwanzig Menschen in der neuen Sprache singen, was seine Mutter in der alten sang.

In dieser unscheinbaren Übung laufen sämtliche Stränge dieses Buches in einem einzigen Punkt zusammen. Es lohnt sich, die Mechanik auseinanderzulegen — es sind fünf Mechanismen, und für jeden einzelnen liegt der Beleg in einem früheren Kapitel bereit.

8.2 Fünf Mechanismen

Erstens: Die Melodie ist schon da. Kapitel 5 endete mit einer Bedingungsliste: Das Lied trägt den Text, wenn die Melodie einfach ist, sich wiederholt und nicht selbst gelernt werden muss — Racettes Gegenbefund entstand ja gerade daran, dass ihre Probanden Text *und* unbekannte Melodie zugleich bewältigen mussten.[87] Das Heimatlied ist der Grenzfall, in dem diese Bedingung perfekt erfüllt ist: Die Melodie sitzt seit der Kindheit, samt Rhythmus, Zeilenlängen und Betonungsmuster — das ganze Gerüst, das nach Wallace den Text „chunkt", steht fertig im Gedächtnis.[88] Die neuen italienischen Wörter werden in eine alte, tief verankerte und emotional besetzte Struktur eingehängt. Einen günstigeren Speicherplatz für fremde Sprache kennt die Gedächtnisforschung nicht.

Zweitens: Übersetzen ist die tiefste Form der Wortarbeit. Kapitel 4 hat die drei Zutaten benannt, die ein Wort haften lassen: Bedürfnis, Suche, Bewertung.[89] Die gemeinsame Liedübersetzung treibt alle drei auf ihr Maximum. Das Bedürfnis ist nicht didaktisch simuliert, sondern echt — es ist *mein* Lied, und die Fassung muss stimmen; die Suche läuft durch Wörterbuch, Lehrer und Gruppendiskussion; und die Bewertung ist die strengste denkbare: Das Wort muss nicht nur

---

[87] Racette/Peretz 2007 (Kapitel 5, Fn. 7): Gelernt wurden dort unvertraute Lieder — die Melodie war Mitlernstoff, das Singen eine Doppelaufgabe. Das Heimatlied-Format eliminiert genau diese Doppelaufgabe für den, der das Lied mitbringt; für die übrigen Gruppenmitglieder ist die Melodie nach wenigen Durchgängen gelernt (einfache, strophische Lieder vorausgesetzt, Kapitel 5, Fn. 8 und 11).

[88] Wallace 1994 (Kapitel 5, Fn. 8): Melodie stützt Textabruf, wenn sie wiederkehrt und leicht lernbar ist; ihre Gliederungsleistung (Zeilenlängen, Betonungsmuster, Chunking) überträgt sich auf jeden Text, der in sie eingepasst wird — also auch auf die Übersetzung. Dass die Übertragung der Gedächtnisstütze von der Original- auf die übersetzte Textfassung verlustfrei gelingt, ist plausible Erwartung, kein geprüfter Befund.

[89] Laufer/Hulstijn 2001 (Kapitel 4, Fn. 8): need, search, evaluation.

richtig sein, sondern *singbar* — Silbenzahl, Betonung, Klang. Jede Zeile ist eine Mehrfachwahl-Glosse in Gruppenform: Fassungen erzeugen, vergleichen, verwerfen.[90] Und was am Ende dasteht, hat die Gruppe *selbst erzeugt* — der Generierungseffekt, einer der robustesten Befunde der Gedächtnispsychologie, besagt genau dies: Selbst Hervorgebrachtes wird besser behalten als bloß Gelesenes, über alle Testformen hinweg.[91]

Drittens: Die Rollen kehren sich um. Im Sprachkurs ist der Geflüchtete strukturell der Defizitäre — der, dem etwas fehlt. In dieser Übung ist er für die Dauer des Liedes der einzige Experte im Raum: Er kennt das Original, er entscheidet, ob die Übersetzung trifft, Lehrer und Gruppe sind auf ihn angewiesen. Die Fremdsprachenangst, deren leistungsdrückende Wirkung Kapitel 2 belegt hat, hängt am bedrohten Selbstbild — am Erwachsenen, der sich als Kind klingen hört.[92] Hier passiert das Gegenteil: Kompetenz wird sichtbar, bevor die neue Sprache sie ausdrücken kann. Die Sing-Up-Studie aus Kapitel 6 gab dem eine empirische Spur: Gelingendes Singen hängt mit Selbstbild und Zugehörigkeitsgefühl zusammen — die Autoren vermuten als Mechanismus gerade die erlebte eigene Wirksamkeit.[93]

Viertens: Integration läuft in beide Richtungen. Assimilation sagt: Lass deine Lieder zurück und lerne unsere. Diese Übung sagt das Gegenteil: *Dein Lied ist es wert, in der neuen Sprache zu existieren.* Die Herkunft wird nicht abgelegt, sondern hinübergetragen; die neue Sprache wird zum Träger der eigenen Geschichte statt zu ihrer Konkurrenz. Das ist Translanguaging nicht als Unterrichtstechnik, sondern im Wortsinn: zwei Sprachen, die an derselben Strophe arbeiten.[94] Und die Aufnahmegesellschaft ist nicht Prüfer, sondern Mitübersetzer und Mitsänger — sie lernt das Lied, die Geschichte dahinter, ein Stück der Menschen, mit denen sie singt. Von allen Formaten, die dieses Buch beschreibt, ist dies das einzige, in dem die Einheimischen etwas von den Ankommenden *lernen* — Integration hört auf, eine Einbahnstraße zu sein.

Dass genau diese Konstellation Vorurteile abbaut, ist dabei keine fromme Hoffnung, sondern einer der bestgeprüften Befunde der Sozialpsychologie. Gordon Allports Kontakt-Hypothese benennt die Bedingungen, unter denen Begegnung zwischen Gruppen wirkt: gleicher Status in der Situation und ein gemeinsames Ziel — und die große Meta-Analyse von Pettigrew und Tropp über mehr als fünfhundert Studien bestätigt: Kontakt senkt Vorurteile typischerweise ohnehin, unter Allports Bedingungen aber am stärksten.[95] Man lese die Bedingungen neben die Übung:

---

[90] Zur Überlegenheit verarbeitungsaktiver Glossenformen Yanagisawa/Webb/Uchihara 2020 (Kapitel 4, Fn. 5 und 7); die Analogie Liedzeile ↔ Mehrfachwahl-Glosse ist Deutung dieses Buches.

[91] Slamecka, N. J./Graf, P., The generation effect: Delineation of a phenomenon, *Journal of Experimental Psychology: Human Learning and Memory* 4 (1978) 592–604 (Abstract-Kernaussage verifiziert: Überlegenheit der Generate- über die Read-Bedingung in allen fünf Experimenten, über freie/gestützte Erinnerung, Wiedererkennen und Konfidenzurteile hinweg; am Volltext zu prüfen). Übertragungs-Hinweis: Der klassische Generierungseffekt betrifft Einzelwörter/Wortpaare im Labor; seine Ausdehnung auf gemeinschaftlich erzeugte Liedzeilen ist Analogie.

[92] Horwitz/Horwitz/Cope 1986 (Kapitel 2, Fn. 9): Fremdsprachenangst als Bedrohung von Selbstbild und Selbstausdruck; Kapitel 2, Fn. 10 zur Leistungswirkung.

[93] Welch et al. 2014 (Kapitel 6, Fn. 7 f.): Zusammenhang Singentwicklung–Selbstbild–Zugehörigkeit; als Mechanismus vermuten die Autoren „perceived competence and singing self-efficacy" — ausdrücklich als Inferenz, nicht als Nachweis.

[94] García/Wei 2014 (Kapitel 7, Fn. 5). Die Liedübersetzung ist zugleich der Punkt, an dem die in Kapitel 7 zitierte L1-Glossen-Überlegenheit (Kapitel 4, Fn. 6) praktisch wird: Die Bedeutungsklärung läuft konstitutiv über die Erstsprache des Liedgebers.

[95] Allport, G. W., *The Nature of Prejudice*, Reading/MA 1954, insb. Kap. 16 (am Druck zu prüfen); meta-analytisch Pettigrew, T. F./Tropp, L. R., A meta-analytic test of intergroup contact theory, *Journal of Personality and Social Psychology* 90 (2006) 751–783 (Kernbefunde am Abstract verifiziert): 713 unabhängige Stichproben aus 515 Studien; Kontakt reduziert Vorurteile typischerweise auch ohne Allports Optimalbedingungen, mit ihnen stärker — die Be-

Die Rollenumkehr stellt den gleichen Status her, die singbare Strophe ist das gemeinsame Ziel. Was die Musik darüber hinaus beiträgt — geteilter Rhythmus, geteilter Atem, und im übersetzten Lied die hörbare Erfahrung, dass das Lied aus Asmara von denselben Hoffnungen erzählt wie das einheimische —, wird in der Musikkognition als sozialer Bindungsmechanismus diskutiert; als gesicherten neurobiologischen Vorurteils-Löscher darf man es nicht ausgeben, als plausiblen Verstärker der Kontaktbedingungen sehr wohl.[96]

Fünftens: Der Vollzug ist angstfrei. Am Ende wird gesungen — die einzige Form lauter Produktion in der neuen Sprache, die vom ersten Tag an ohne Blamage-Risiko möglich ist, mit der geschützten Einzelstimme in der Menge (Kapitel 6). Und das übersetzte Lied verschwindet nicht: Es geht ins Repertoire der Gruppe ein und wird wieder und wieder gesungen — die Wiederholung, die das beiläufige Lernen braucht (Kapitel 4), ist eingebaut, ohne dass sie je „Übung" heißt.

Man sieht: Kein einziger dieser fünf Mechanismen ist eine Erfindung dieses Buches — Melodie-Gerüst, Verarbeitungstiefe, Generierungseffekt, Selbstwirksamkeit, Angstsenkung sind je für sich belegte, teils jahrzehntealte Befunde. Die Übung ist ihre Zusammenführung in einem einzigen Format. Genau das ist ihre Stärke — und, ehrlich gesagt, auch ihr Beweisproblem, auf das der letzte Abschnitt zurückkommt.

8.3 Die Erzählung: Warum das Ganze eine Geschichte braucht

Pontivo erzählt diese Übung — und die Methode dahinter — als Geschichte: *Coro One Sky*, ein Comic über einen Chor, in dem Kinder aus verschiedenen Ländern unter demselben Himmel zusammensingen. Man könnte das für Verpackung halten. Die Kognitionsforschung legt nahe, dass es mehr ist.[97]

Jerome Bruner — derselbe, von dem das Gerüst-Bild aus Kapitel 7 stammt — hat argumentiert, dass das narrative Denken ein eigener, nicht ersetzbarer Modus des Verstehens ist: Menschen ordnen soziale Wirklichkeit in Geschichten, nicht in Sätzen mit Wahrheitswert.[98] Mar und Oatley haben daraus ein Forschungsprogramm gemacht: Fiktion funktioniert als *Simulation* der sozialen Welt — sie abstrahiert, verdichtet und lässt den Leser soziale Erfahrung durchleben, die er selbst nicht gemacht hat.[99] Für die Sprachdidaktik kommt ein handfesterer Befund hinzu, der in Kapitel 2 schon auftrat: Die Bücherflut wirkte mit *Geschichten*büchern, und die Extensive-Reading-Tradition hat seither immer wieder gezeigt, dass erzählendes, selbstgewähltes, leichtes

---

dingungen sind förderlich, nicht zwingend notwendig. Genau deshalb wird hier nicht behauptet, die Übung sei die einzige wirksame Form — nur, dass sie die förderlichen Bedingungen ungewöhnlich vollständig herstellt.

[96] Zur Diskussion von Musik als sozialem Bindungsmechanismus statt vieler Koelsch, S., *Brain and Music*, Chichester 2012 (am Druck zu prüfen), sowie die Bindungsbefunde aus Kapitel 6 (Fn. 1 f., 4). Ausdrücklich zurückgewiesen wird die stärkere Lesart, gemeinsames Musizieren löse „die neurobiologischen Grenzen zwischen In-Group und Out-Group auf" — dafür gibt die Befundlage keine Deckung. Wer den Moment lieber soziologisch fasst: Hartmut Rosas Begriff der *Resonanz* (Rosa, *Resonanz. Eine Soziologie der Weltbeziehung*, Berlin 2016, am Druck zu prüfen) beschreibt ihn als geglückten Weltbezug — als Deutungsrahmen zitiert, nicht als Beleg.

[97] Funktions-Hinweis: Der Comic ist auf der Plattform lesbar (pontivo.de/comic, Stand 15. Oktober 2026).

[98] Bruner, J., *Actual Minds, Possible Worlds*, Cambridge/MA 1986, Kap. 2 (paradigmatischer vs. narrativer Modus; am Druck zu prüfen); ders., The narrative construction of reality, *Critical Inquiry* 18 (1991) 1–21 (am Volltext zu prüfen).

[99] Mar, R. A./Oatley, K., The function of fiction is the abstraction and simulation of social experience, *Perspectives on Psychological Science* 3 (2008) 173–192 (Abstract am PubMed-Eintrag verifiziert): Fiktion bietet „models or simulations of the social world via abstraction, simplification, and compression" und erzeugt „a deep and immersive simulative experience of social interactions".

Lesen — ausdrücklich auch Comics — Sprache trägt, weil die Geschichte den Leser zieht, wo die Übung ihn schieben müsste.[100]

Im Konzept übernimmt die Erzählung deshalb drei Aufgaben. Sie gibt der Methode *Figuren und Reihenfolge*: Wer den Comic liest, sieht Kinder, die genau das durchleben, was die Übung verlangt — das Fremdsein, das erste Mitsingen, das eigene Lied —, und weiß dadurch, was kommt; Erwartbarkeit ist Verstehenshilfe, für Kinder wie für die Lehrerin, die das Format zum ersten Mal wagt. Sie liefert *leichten Lesestoff* in der Zielsprache, doppelt kodiert im Wortsinn — Bild und Text —, als niederschwelligste Input-Form der Plattform. Und sie macht die Methode *erzählbar*: Eine Schule, ein Elternabend, ein Chorleiter versteht „ein Chor unter einem Himmel" schneller als jede Hypothesenliste.

Auch hier die Ehrlichkeitsgrenze: Die narrative Evidenz stammt aus der Lese- und Fiktionsforschung; dass ein *Plattform-Comic* diese Funktionen erfüllt, ist Übertragung, nicht Nachweis.[101] Der Comic ist das am wenigsten messbare Stück des Konzepts — und zugleich das billigste: Er kostet keine Unterrichtszeit und konkurriert mit nichts. Wetten mit kleinem Einsatz und plausiblem Gewinn darf man eingehen, wenn man sie als Wetten führt.

8.4 Die Vorsichtsregel — und die eine große Wette

Zwei Dinge müssen am Ende dieses Kapitels stehen, gerade weil es das begeistertste des Buches ist.

Das erste ist eine Schutzregel. Heimatlieder sind nicht neutral: An ihnen hängen Kindheit, Familie, manchmal Verlust und Schlimmeres. Ein Format, das Menschen auffordert, ihr Innerstes zum Unterrichtsmaterial zu machen, braucht Freiwilligkeit als hartes Prinzip — *wessen* Lied gesungen wird, entscheidet allein, wem es gehört; wer (noch) nicht will, singt die Lieder der anderen mit, und das ist kein Defizit. Und niemand wird zum Repräsentanten „seiner Kultur" ernannt: Der Junge bringt *sein* Lied mit, nicht „die syrische Musik". Die Grenze zwischen Würdigung und Zurschaustellung ist schmal, und sie verläuft dort, wo die Entscheidung liegt: beim Sänger, nicht beim Format.[102]

Das zweite ist die Bilanz der Beweislage, ohne Schutzbehauptung: Die fünf Mechanismen sind einzeln belegt; ihre Kombination in diesem Format ist es nicht. Nach Kenntnis dieses Buches ist das übersetzte Heimatlied als integriertes Lern- und Integrationsformat empirisch nicht untersucht — es ist die zentrale Design-Wette des ganzen Konzepts, mit dem besten Begründungsfundament und der dünnsten Direktevidenz.[103] Ein Konzept, das sich wissenschaftlich nennt, hat

---

[100] Elley/Mangubhai (Kapitel 2, Fn. 5 f.); zur Extensive-Reading-Tradition Day, R./Bamford, J., *Extensive Reading in the Second Language Classroom*, Cambridge 1998 (am Druck zu prüfen); zu Comics und leichtem Lesen Krashen, S., *The Power of Reading*, 2. Aufl., Westport 2004, Kap. 3 (am Druck zu prüfen); Cary, S., *Going Graphic: Comics at Work in the Multilingual Classroom*, Portsmouth 2004 (am Druck zu prüfen).

[101] Übertragungs-Hinweis in aller Form: Belegt sind Simulations- und Verstehensfunktionen von Erzählungen (Fn. 12 f.) und die Sprachwirkung erzählenden Lesestoffs (Fn. 14). Nicht belegt ist die Wirkung eines bestimmten Comics als narratives Dach einer Lernplattform; messbar wäre am ehesten seine Nutzung (wird er gelesen? von wem? vor oder nach den Filmen?) — Kapitel 9.

[102] Die Schutzregel ist pädagogische Wertung dieses Buches, kein Forschungsreferat; sie folgt aus der Trauma-Sensibilität der Arbeit mit Geflüchteten (vgl. die qualitative Community-Music-Literatur, Kapitel 6, Fn. 12, die Freiwilligkeit und Eigentum an den Beiträgen durchgängig betont).

[103] Verwandte, aber nicht deckungsgleiche Forschungslinien existieren — Singen im L2-Unterricht (Teil III), Community Music mit Geflüchteten (Kapitel 6, Fn. 12), Übersetzen im Sprachunterricht (Translation in Language Teaching, statt vieler Cook, G., *Translation in Language Teaching*, Oxford 2010, am Druck zu prüfen) —; die Schnittmenge aller

dafür genau einen redlichen Umgang: die Wette benennen, den Einsatz klein halten, die Messung planen. Was genau zu messen wäre — und was das Konzept insgesamt kann und nicht kann —, ist das letzte Kapitel.

drei ist nach hiesiger Recherche unbeforscht. Sollte sich das bei vertiefter Literatursuche ändern, gehört der Befund in die nächste Revision dieses Buches. — Nummerierung dieses Apparats in rev. 3 angepasst (Fn. 9 f. neu; 9–14 alt → 11–16); in rev. 4 erneut: Funktions-Fußnote 11 neu, 11–16 alt → 12–17.

Teil VI — Die Szenarien & der offene Weg

„Nicht die Wahrheit, in deren Besitz irgend ein Mensch ist, oder zu sein vermeinet, sondern die aufrichtige Mühe, die er angewandt hat, hinter die Wahrheit zu kommen, macht den Werth des Menschen."

— Gotthold Ephraim Lessing, Eine Duplik (1778), Erster Abschnitt

Kapitel 9 · Drei Klassenzimmer und eine ehrliche Agenda

9.1 Drei Klassenzimmer

Alles, was dieses Buch entwickelt hat, läuft in drei konkreten Szenarien zusammen, für die Pontivo gebaut ist — und es ist dieselbe Maschine, die alle drei trägt.

Das erste Klassenzimmer steht in Deutschland: Geflüchtete lernen Deutsch. Der Abendfilm — eine Dokumentation im deutschen Original — läuft mit deutscher Zielspur und der jeweils eigenen Muttersprachzeile; der Klick erklärt die Stelle, an der es hakt; einmal die Woche singt die Gruppe, und in unregelmäßigen Abständen wandert ein Lied aus Aleppo, Cherson oder Asmara durch die Übersetzungswerkstatt aus Kapitel 8. Das zweite Klassenzimmer steht in Italien und unterscheidet sich nur in einer Achsenstellung: Geflüchtete lernen Italienisch — der Chor heißt hier *One Sky*, und er ist das namensgebende Szenario des Comics. Das dritte Klassenzimmer ist die Klasse aus Bologna, mit der dieses Buch begann: Italienische Jugendliche lernen Englisch — dieselbe Doppelspur, nur dass unten jetzt Italienisch steht und im Player die Dokumentation läuft, die im Unterricht bisher an der Synchronfassung scheiterte.

Dass drei so verschiedene Lagen — Integration unter Druck, Integration im Ankunftsland, Schulfremdsprache im Wohlstand — mit einer Engine bedienbar sind, ist keine Vermarktungspointe, sondern eine Konsequenz der Theorie: Wenn Spracherwerb überall denselben Bedingungen folgt — verständlicher Input, Verankerung, niedrige Angst —, dann unterscheiden sich die Szenarien nur in der Stellung der Achsen: welche Sprache Ziel ist, welche Gerüst, wer die Gruppe führt.[104]

Die drei Klassenzimmer sind darum auch keine abschließende Liste. Die Verwaltung sortiert Zuwanderung in getrennte Sphären — hier „Expats", dort „Geflüchtete" —, und in klassischen Lernumgebungen begegnen die Gruppen einander nie. Kognitiv aber teilen die irische Ingenieurin im Konzernbüro und der syrische Junge aus Kapitel 8 dieselbe Erfahrung: den Verlust sprachlicher Souveränität, die Erschöpfung des ständigen Nicht-Verstehens, die Aufgabe, sich neu zu verorten. Für die Bedingungen dieses Buches ist der Aufenthaltstitel keine Variable; dieselbe Achsenstellung, die den Integrationskurs trägt, trägt die internationale Fachkraft und ihre Familie. Und sie trägt die Regelschule, deren Klassen längst so vielsprachig sind wie jede Integrationsklasse: Gemeinsames Sehen und Singen in heterogenen Klassen verlangt keine kognitive Eingangshürde — es lässt Kinder als Gruppe handeln, bevor die Notenskala sie sortiert. Was in der Not der Geflüchteten entstanden ist, ist an keiner Stelle nur für sie gebaut.

9.2 Amal, ein Jahr später — eine Hypothese

Es ist verführerisch, an dieser Stelle die Eröffnungsszene aufzulösen: Amal, ein Jahr mit Abendfilmen und Chorproben, an derselben Supermarktkasse — und diesmal antwortet sie. Dieses Buch wird der Verführung nicht ganz nachgeben, denn genau diese Szene ist keine Tatsache, sondern die *Hypothese*, um deren ehrliche Prüfung es hier geht. Was sich sagen lässt, ist bescheidener und fester: Nach einem solchen Jahr hätte Amal einige hundert Stunden echtes, verstandenes Deutsch gehört — mehr gesprochene Alltagssprache, als sechshundert Kursstunden je

[104] Die Achsen-Architektur ist Produktbeschreibung (Kapitel 7, Fn. 11). Der Satz „gleiche Bedingungen, andere Achsenstellung" ist die Konsequenz der in Teil I entwickelten Erwerbstheorie und teilt deren Geltungsgrenzen — insbesondere die Alphabetisierungsgrenze (9.3).

enthalten können; sie hätte Dutzende Wendungen im Wortlaut verfügbar, weil sie sie gesungen hat; und sie stünde nicht mehr allein: Menschen, neben denen man ein Jahr gesungen hat, sind keine Fremden mehr, vor denen man verstummt. Ob das reicht, damit sie „Sammeln Se Punkte?" versteht und lachend verneint, weiß niemand — *noch* niemand. Der Unterschied zum Kurssystem aus Kapitel 1 ist nicht das Versprechen. Der Unterschied ist, dass die Bedingungen stimmen — und dass man das Ergebnis messen kann und messen will.

9.3 Was der Ansatz nicht kann

Zuerst aber die Grenzen, und zwar die grundsätzlichen, nicht die kosmetischen.

Der Ansatz macht keine Sprecher. Die Forschung, die diesem Buch sein Fundament gibt, hat eine offene Flanke, und sie wurde in Kapitel 2 benannt: Verstehen ist notwendig, aber vermutlich nicht hinreichend. Eigenes, frei formuliertes Sprechen — das Aushandeln im Gespräch, das Ringen um den eigenen Satz — hat nach allem, was die Output- und Interaktionsforschung zeigt, eine eigene Funktion, die kein Input ersetzt.[105] Pontivo deckt von der Produktion nur das Singen ab — wortgetreue, geschützte Produktion, wertvoll und begrenzt. Das freie Gespräch findet außerhalb der Plattform statt oder gar nicht; die Aufgabe, es herbeizuführen, bleibt bei Lehrkräften, Tandems, Begegnungen. Wer den Ansatz als Vollersatz des Sprachunterrichts verkauft, hat dieses Buch nicht gelesen.

Der Ansatz setzt Lesen voraus. Die Doppelspur, das Herzstück von Teil II, ist Text. Wer nicht oder kaum lesen kann — und in Integrationskursen sitzen Menschen, die nie alphabetisiert wurden oder eine andere Schrift mitbringen —, für den trägt genau der Mechanismus nicht, der hier am besten belegt ist.[106] Für diese Lernenden bleiben Ton, Bild und Lied; ob das genügt, ist offen. Es wäre wohlfeil, diese Grenze zu verschweigen, weil sie die verletzlichste Gruppe betrifft.

Der Ansatz hängt an den Rechten. Authentischer Film ist fremdes Eigentum; was gezeigt, untertitelt und angeklickt werden darf, entscheiden Lizenzen, nicht Didaktiker. Die Quellenlage begrenzt den Katalog, und ein Konzept, das auf Menge baut, muss diese Abhängigkeit als Betriebsrisiko führen — der folgende Abschnitt entfaltet diese Grenze und die Wege hinaus.

Der Klick kann falsch erklären. Die maschinell erzeugte Glosse ist eine Qualitätsschuld (Kapitel 4): Ohne Stichprobenprüfung gegen fachliche Referenz bleibt jede einzelne Erklärung ein kleines Versprechen ohne Beleg.

Und die Lehrkraft bleibt unersetzt. Aufgaben, Gespräch, Ermutigung, die Entscheidung, wann ein Gerüst fällt — das alles kann die Plattform rahmen, aber nicht leisten (Kapitel 7). Wo sie als Sparmodell für Personal missverstanden würde, wäre sie pädagogisch falsch eingesetzt.

9.4 Das Rechte-Paradox und die Lieder-Spende

Eine der Grenzen aus 9.3 verdient einen eigenen Abschnitt, weil sie nicht kognitiver, sondern rechtlicher Natur ist — und weil sie sich gestalten lässt. Das Urheberrecht kennt zwar eine

[105] Swain 1985; Long 1996 (beide Kapitel 2, Fn. 4, am Druck zu prüfen). Die Grenzziehung folgt dem dort markierten Konsenskern: Input notwendig, Output/Interaktion mit eigener, nicht ersetzbarer Funktion.

[106] Zu Schriftferne und Alphabetisierung in Integrationskursen führt das BAMF eigene Alphabetisierungskurse (vgl. den in Kapitel 1, Fn. 2 zitierten Bericht, Kursarten-Tabellen); die Captioning-Evidenz (Kapitel 3) ist durchgängig an lesefähigen Stichproben gewonnen. Für nicht alphabetisierte Lernende ist der Untertitel-Mechanismus nicht belegt und plausibel unwirksam — eine Konzeptgrenze, keine Randnotiz.

Schranke für Unterricht und Lehre: § 60a UrhG erlaubt Bildungseinrichtungen, zu nicht kommerziellen Zwecken bis zu 15 Prozent eines veröffentlichten Werkes zu nutzen, Werke geringen Umfangs — und damit typischerweise auch einzelne Liedtexte — sogar vollständig. Aber die Schranke hat zwei Haken, und beide treffen dieses Konzept ins Zentrum. Erstens privilegiert sie die Bildungseinrichtungen selbst — Schulen, Hochschulen, Einrichtungen der Berufs- und Weiterbildung —, nicht eine Plattform: Für ihren Katalog kann sich Pontivo auf § 60a nicht berufen; privilegiert ist nur der Einsatz *durch* eine Bildungseinrichtung in deren Unterricht. Zweitens nimmt das Gesetz ausgerechnet die Noten aus: Grafische Aufzeichnungen von Werken der Musik dürfen auch im Unterricht nicht ohne Lizenz vervielfältigt werden.[107] Eine Methode aber, die von der Immersion in ganze Werke lebt — der ganze Film, das ganze Lied —, kommt mit fünfzehn Prozent nirgends hin. Das ist das Rechte-Paradox der Sprachbildung: Ausgerechnet das Material, das sprachlich und emotional am stärksten trägt, ist im Lernraum am schwersten rechtssicher zu nutzen.

Der Weg führt deshalb nicht über Schranken, sondern über Lizenzen — und damit durch das internationale Geflecht der Verwertungsgesellschaften: in Deutschland die GEMA und, für Noten und Liedtexte, die VG Musikedition, in Österreich die AKM, in der Schweiz die SUISA, daneben SACEM, PRS, ASCAP und ihre Schwestern. Deren Tarifwerke sind historisch auf Rundfunk, Aufführung und Schulbuch zugeschnitten, nicht auf geschützte digitale Lernräume, die ganze Werke brauchen.[108] Zwei Auswege zeichnen sich ab. Öffentlich-rechtliche und öffentlich finanzierte Sender können Bildungsfreigaben für geschlossene Lernumgebungen erteilen — ihr Auftrag trägt das, und die in Kapitel 3 versammelte Untertitel-Evidenz liefert die Begründung gleich mit. Und für die Lieder etabliert das Konzept ein Gegenseitigkeitsmodell, die *Lieder-Spende*: Musikerinnen, Musiker und Chöre nutzen das Liederbuch mit seinen ChordPro-Sätzen als Werkzeug ihrer eigenen Proben- und Bühnenpraxis; im Gegenzug können Rechteinhaber einzelne eigene Werke per Direktlizenz für den didaktischen Raum freigeben. So wird aus dem Rechtehindernis ein Tauschverhältnis — freiwillig, dokumentiert, kündbar. Ob es trägt, entscheidet sich nicht am Konzept, sondern an den Verträgen; auch das gehört zur ehrlichen Bilanz dieses Kapitels.

9.5 Die Messliste

Ein Konzept, das sich wissenschaftlich nennt, sagt selbst, woran es scheitern könnte. Die folgende Liste ist nach Tragweite geordnet: oben die Wetten, an denen das Konzept hängt.

1. **Die Doppelspur.** Lernen Anfänger mit synchroner Ziel- und Muttersprachzeile messbar mehr als mit einsprachigen Untertiteln — über Monate, nicht Minuten? (Design: Vergleichsgruppen im Feld; Vokabel- und Hörverstehenstests; die Kurzclip-Befunde aus Kapitel 3 sind Startpunkt, nicht Antwort.) 2. **Der Chor-Effekt auf den Verbleib.** Senkt das Chor-Modul Sprachangst und

[107] § 60a Abs. 1 UrhG (15 %-Grenze, nicht kommerzielle Zwecke, privilegierter Nutzerkreis), Abs. 2 (vollständige Nutzung u. a. von „Werken geringen Umfangs" und vergriffenen Werken), Abs. 3 Satz 1 (u. a. keine „Vervielfältigung von grafischen Aufzeichnungen von Werken der Musik" = Notenausschluss), Abs. 4 (erfasste Einrichtungen) — Wortlaut am amtlichen Volltext verifiziert (gesetze-im-internet.de/urhg/__60a.html, abgerufen 2026-07-25); eingeführt durch das Urheberrechts-Wissensgesellschafts-Gesetz (UrhWissG, 2017). Ob ein konkreter Liedtext „Werk geringen Umfangs" ist, entscheidet sich im Einzelfall (Rahmenverträge der Länder mit den Verwertungsgesellschaften, am Wortlaut zu prüfen); die Vergütungspflicht nach § 60h UrhG bleibt unberührt. Der Abschnitt ersetzt keine Rechtsberatung im Einzelfall.

[108] Zur Zuständigkeitsverteilung (GEMA: Aufführungs-/Wiedergaberechte; VG Musikedition: u. a. grafische Rechte an Noten und Texten) und zu den internationalen Schwestergesellschaften: Selbstdarstellungen der Gesellschaften (am jeweiligen Tarif-/Vertragswerk zu prüfen). Die Aussage über den Zuschnitt der Tarifwerke ist eine Struktur-Beobachtung, kein Tarifgutachten.

erhöht es Zugehörigkeit und Dranbleiben? (Instrumente liegen vor: die Angstskala FLCAS, Zugehörigkeitsmaße, schlicht die Abbruchquote — die zweistufige Kette aus Kapitel 6 wird damit prüfbar.[109]) 3. **Das Heimatlied-Format.** Bleiben Wörter aus selbst übersetzten Liedern besser als gleich häufige Wörter aus dem Filmkorpus? Steigt die erlebte Selbstwirksamkeit der Liedgeber? (Kleines Vorher/Nachher-Design pro Chorgruppe; die zentrale Wette aus Kapitel 8.) 4. **Der gesungene Wortlaut.** Sind gesungene Wendungen nach Wochen noch abrufbar, wo gehörte verblasst sind? (Die Laborbefunde aus Kapitel 5, in den Feldmaßstab übersetzt.) 5. **Die Klick-Glosse.** Stichproben-Audit der maschinellen Erklärungen gegen fachliche Referenz (Fehlerquote, Fehlertypen); dazu: Behalten geklickter gegenüber nicht geklickter Wörter, und die Flow-Frage — ab welcher Klickdichte kippt der Filmabend in Wörterbucharbeit? 6. **Vokabelzuwachs pro Sehstunde** als laufende Basisgröße (Stichproben-Vokabeltests entlang des tatsächlich gesehenen Materials), um die Beiläufigkeits-Erwartung aus Kapitel 4 gegen die Realität zu halten. 7. **Der Comic:** Wird er gelesen, von wem, und geht seinem Lesen Nutzung der übrigen Module voraus oder nach? (Reine Nutzungsmessung — mehr gibt die Evidenzlage nicht her, Kapitel 8.)

Nichts an dieser Liste ist exotisch; jede Zeile ist mit Bordmitteln einer Plattform und den Standardinstrumenten der Forschung machbar. Was sie verlangt, ist nur eines: den Willen, sich messen zu lassen — auch auf das Risiko hin, dass eine Wette verloren geht. Verlorene Wetten sind im Maßstab dieser Schriftenreihe kein Makel; unbenannte sind einer.

9.6 Der Stammbaum

Zum Schluss die Inventur, die jedes Buch dieser Reihe sich selbst schuldet: Was daran ist geliehen, was ist eigen?

Geliehen ist fast alles, und zwar mit Absicht. Die Erwerbs- statt Lern-Perspektive und die beiden Bedingungen stammen von Krashen und der auf ihn folgenden Forschung; die Zwei-Kanal-Architektur des Gedächtnisses von Paivio und Mayer; die Untertitel-Evidenz von Vanderplank über die Leuvener Meta-Analyse bis zu den Doppelspur-Studien; die Glossen- und Involvement-Befunde von Nation, Laufer, Hulstijn und der Meta-Analyse von Yanagisawa; das Gesangsgedächtnis von Wallace, Schön und Ludke samt der Ernüchterung durch Racette und Peretz; der Eisbrecher-Effekt von Pearce, Launay und Dunbar; Gerüst und Zone der nächsten Entwicklung von Vygotskij, Wood, Bruner und Ross; die Rehabilitierung der Muttersprache von Butzkamm und der Translanguaging-Forschung; die Erzähltheorie von Bruner, Mar und Oatley; Motivation und Sicherheit von Deci, Ryan, Csíkszentmihályi und Edmondson; die Kontakt-Hypothese von Allport, Pettigrew und Tropp.[110]

Eigen sind drei Dinge. Erstens die *Integration*: die Zusammenführung dieser Stränge in eine Architektur — Film als Input, Lied als Anker, Gruppe als Filter-Senker — und in eine Engine, die alle drei Szenarien mit denselben Achsen bedient. Zweitens das *Heimatlied-Format* als Übung, in der die Stränge sich kreuzen; die Bausteine sind fremd, die Kombination ist, soweit ersichtlich,

[109] FLCAS: Horwitz/Horwitz/Cope 1986 (Kapitel 2, Fn. 9), Skala S. 129 f.; Zugehörigkeits-/Nähemaße nach dem Vorbild von Pearce et al. 2015 (Kapitel 6, Fn. 1). Dass Verbleib („Retention") als Integrationsmaß taugt, ist Setzung dieses Buches: Wer bleibt, bekommt Menge — und Menge ist nach Teil II die erste Bedingung.

[110] Sämtliche Fundstellen in den Kapiteln 2 bis 8; das Quellenregister im Anhang führt sie gesammelt, mit Prüfstatus (verifiziert / am Druck zu prüfen) je Fundstelle — der Verifikationsstand ist Teil der Publikation, nicht internes Material: Auch das gehört zu „prüfbar statt behauptet".

neu. Drittens der *Merk-Anker* als Prüfraster: die drei Fragen, mit denen sich jede Lernumgebung — auch jede künftige Version von Pontivo selbst — verhören lässt.

Und damit zur letzten Anwendung dieses Rasters, auf das Buch selbst. **Verständlich?** Es hat seine Hauptlinie für Leser ohne Fachstudium gebaut und die Tiefe in die Fußnoten gelegt; ob das gelungen ist, entscheiden die Chorleiterinnen und Lehrer, die es lesen. **Verankert?** Es hat, wo es konnte, an Szenen erzählt statt an Konstrukten — an einer Frau an einer Kasse, an einem Jungen, dessen Lied zwanzig Stimmen singen. **Verbunden?** Es erscheint offen, für jeden lesbar, in einem Verlag, der Wissen nicht hinter Schranken legt — und es lädt ausdrücklich zur Widerlegung ein: Jede Zahl trägt ihre Fundstelle, jede Wette ihren Namen.

Am Anfang stand eine Frau, die schweigend bezahlt hat, obwohl sie ein Zertifikat besitzt. Dieses Buch hat nicht versprochen, dass sie sprechen wird. Es hat gezeigt, unter welchen Bedingungen Menschen zu sprechen beginnen — und ein Instrument beschrieben, das diese Bedingungen herstellt, samt der Liste, an der es sich prüfen lassen muss. Der Rest ist keine Theorie mehr. Der Rest ist ein Dienstagabend in einem Gemeindesaal, ein Film mit zwei Zeilen Text, und ein Lied, das jemand mitgebracht hat.

Quellenregister

Alphabetisch. Status: **[V]** = an der Primärstelle verifiziert (Volltext oder Verlagsabstract, Prüfweg im Verifikations-Log) · **[D]** = am Druck/Volltext zu prüfen (so auch in der jeweiligen Fußnote ausgewiesen). Kapitel-Fundstellen in Klammern.

- Allport, G. W., *The Nature of Prejudice*, Reading/MA 1954 — [D] (Kap. 8)
- Augustinus, *Confessiones*, Buch I, Kap. 14 (I,14,23) — [V, lateinischer Volltext] (Motto Teil II)
- BAMF, *Bericht zur Integrationskursgeschäftsstatistik für das Jahr 2024*, S. 17 ff., 19 — [V, PDF-Volltext] (Kap. 1, 9)
- Bruner, J., *Actual Minds, Possible Worlds*, Cambridge/MA 1986 — [D] (Kap. 8)
- Bruner, J., The narrative construction of reality, *Critical Inquiry* 18 (1991) 1–21 — [D] (Kap. 8)
- Butzkamm, W., *Psycholinguistik des Fremdsprachenunterrichts*, 3. Aufl. 2002 — [D] (Kap. 7)
- Butzkamm, W./Caldwell, J. A. W., *The Bilingual Reform*, Tübingen 2009 — [D] (Kap. 7)
- Cary, S., *Going Graphic: Comics at Work in the Multilingual Classroom*, Portsmouth 2004 — [D] (Kap. 8)
- Clift, S./Hancox, G., The perceived benefits of singing, *J Royal Society for the Promotion of Health* 121 (2001) 248–256 — [V, Abstract] (Kap. 6)
- Cook, G., *Translation in Language Teaching*, Oxford 2010 — [D] (Kap. 8)
- Council of Europe, *CEFR Companion Volume*, Straßburg 2020 — [V, Kernaussage an coe.int; Pinpoints am PDF zu prüfen] (Kap. 7)
- Csíkszentmihályi, M., *Flow: The Psychology of Optimal Experience*, New York 1990 — [D] (Kap. 2)
- Day, R./Bamford, J., *Extensive Reading in the Second Language Classroom*, Cambridge 1998 — [D] (Kap. 8)
- Deci, E. L./Ryan, R. M., *Intrinsic Motivation and Self-Determination in Human Behavior*, New York 1985 — [D] (Kap. 2)
- DeKeyser, R., Skill Acquisition Theory, in: VanPatten/Williams (Hg.), *Theories in SLA*, 2. Aufl. 2015 — [D] (Kap. 1)
- Dow, R./Vougioukalou, S., Wellbeing and Integration through Community Music, *Contemporary Music Review* 38 (2019) — [D, Verlagszugriff gesperrt] (Kap. 6)
- Dunbar, R., *Grooming, Gossip, and the Evolution of Language*, 1996 — [D] (Kap. 6)
- d'Ydewalle, G./Van de Poel, M., Incidental foreign-language acquisition by children watching subtitled TV, *J Psycholinguistic Research* 28 (1999) 227–244 — [D] (Kap. 3)
- Edmondson, A., Psychological Safety and Learning Behavior in Work Teams, *Administrative Science Quarterly* 44 (1999) 350–383 — [V, Definition am Volltext-Spiegel; am Druck zu prüfen] (Kap. 6)
- Elley, W./Cutting, B./Mangubhai, F./Hugo, C., *Lifting Literacy Levels with Story Books*, ERIC ED416441 (1996), S. 4 f. — [V, Volltext] (Kap. 2, 7, 8)

- Elley, W./Mangubhai, F., The impact of reading on second language learning, *Reading Research Quarterly* 19 (1983) 53–67 — [D; Befunde nach der Autoren-Darstellung ED416441 referiert] (Kap. 2)
- Ellis, R., *Task-based Language Learning and Teaching*, Oxford 2003 — [D] (Kap. 7)
- Ellis, R., *The Study of Second Language Acquisition*, 2. Aufl., Oxford 2008 — [D] (Kap. 1)
- Europarat, *Gemeinsamer europäischer Referenzrahmen für Sprachen* (2001), Globalskala — [D; Raster online geprüft] (Kap. 1)
- García, O., *Bilingual Education in the 21st Century*, Malden 2009 — [D] (Kap. 7)
- García, O./Wei, L., *Translanguaging*, Basingstoke 2014 — [D] (Kap. 7, 8)
- Hattie, J., *Visible Learning*, London 2009 — [D; Effektstärken tw. sekundär, am Druck zu prüfen] (Kap. 7)
- Herder, J. G., Von Ähnlichkeit der mittlern englischen und deutschen Dichtkunst, *Deutsches Museum* 1777, S. 433 — [V, Digitalisat] (Motto Teil V)
- Horwitz, E. K./Horwitz, M. B./Cope, J., Foreign Language Classroom Anxiety, *Modern Language Journal* 70 (1986) 125–132 — [V, Volltext; Definition S. 128, FLCAS S. 129 f.] (Kap. 2, 6, 8, 9)
- Hulstijn, J., Incidental and intentional learning, in: Doughty/Long (Hg.), *Handbook of SLA*, Oxford 2003, S. 349–381 — [D] (Kap. 4)
- INVALSI, Ergebnisdarstellung Prova INVALSI 2025, Inglese grado 13 (invalsi.it) — [V, Institutsseite; Rapporto Nazionale am Volltext zu prüfen] (Kap. 1)
- IntV (Integrationskursverordnung), § 10 Abs. 1, § 11 Abs. 1 — [V, gesetze-im-internet.de] (Kap. 1)
- Kemmis, S./McTaggart, R./Nixon, R., *The Action Research Planner*, Singapur 2014 — [D] (Kap. 7)
- Koelsch, S., *Brain and Music*, Chichester 2012 — [D] (Kap. 8)
- Krashen, S., *Principles and Practice in Second Language Acquisition*, Pergamon 1982 (Internet-Ausgabe 2009), S. 10 f., 20 f., 30–32 — [V, Volltext] (Kap. 1, 2)
- Krashen, S., *The Input Hypothesis*, London 1985 — [D] (Kap. 2)
- Krashen, S., *The Power of Reading*, 2. Aufl., Westport 2004 — [D] (Kap. 8)
- Kreutz, G. et al., Effects of choir singing or listening on secretory immunoglobulin A, cortisol, and emotional state, *J Behavioral Medicine* 27 (2004) 623–635 — [V, Abstract] (Kap. 6)
- Laufer, B./Hulstijn, J., Incidental vocabulary acquisition: The construct of task-induced involvement, *Applied Linguistics* 22 (2001) 1–26 — [D] (Kap. 4, 8)
- Lessing, G. E., *Eine Duplik* (1778), Erster Abschnitt — [V, Wortlaut an geprüften Zitatquellen; am Druck zu prüfen] (Motto Teil VI)
- Liao, S./Kruger, J.-L./Doherty, S., The impact of monolingual and bilingual subtitles, *J Specialised Translation* 33 (2020) 70–98 — [V, Open-Access-Volltext] (Kap. 3, 7)
- Lightbown, P./Spada, N., *How Languages are Learned*, 4. Aufl., Oxford 2013 — [D] (Kap. 1)
- Long, M., The role of the linguistic environment in SLA, in: Ritchie/Bhatia (Hg.), *Handbook of SLA*, 1996, S. 413–468 — [D] (Kap. 2, 9)

- Long, M., *SLA and Task-Based Language Teaching*, Malden 2015 — [D] (Kap. 7)
- Ludke, K. M./Ferreira, F./Overy, K., Singing can facilitate foreign language learning, *Memory & Cognition* 42 (2014) 41–52 — [V, Volltext] (Kap. 5, 8)
- Mar, R. A./Oatley, K., The function of fiction is the abstraction and simulation of social experience, *Perspectives on Psychological Science* 3 (2008) 173–192 — [V, Abstract] (Kap. 8)
- Mayer, R. E., *Multimedia Learning*, 3. Aufl., Cambridge 2021 — [D] (Kap. 3)
- McLaughlin, B., *Theories of Second-Language Learning*, London 1987 — [D] (Kap. 1, 2)
- Montero Perez, M./Peters, E./Clarebout, G./Desmet, P., Effects of captioning on video comprehension and incidental vocabulary learning, *Language Learning & Technology* 18/1 (2014) 118–141 — [V, Volltext] (Kap. 4)
- Montero Perez, M./Van Den Noortgate, W./Desmet, P., Captioned video for L2 listening and vocabulary learning: A meta-analysis, *System* 41 (2013) 720–739 — [V, Abstract-Kernaussage; g-Werte am Volltext zu prüfen] (Kap. 3)
- Murphey, T., The song stuck in my head phenomenon, *System* 18 (1990) 53–64 — [D] (Kap. 5)
- Murphey, T., *Music and Song*, Oxford 1992 — [D] (Kap. 5)
- Nation, I. S. P., *Learning Vocabulary in Another Language*, 2. Aufl., Cambridge 2013 — [D] (Kap. 4)
- Nietzsche, F., *Götzen-Dämmerung*, Sprüche und Pfeile 33 (1889) — [V, Gutenberg-Volltext] (Motto Teil III)
- Norris, J./Ortega, L., Effectiveness of L2 instruction, *Language Learning* 50 (2000) — [D] (Kap. 1)
- Oldenburg, R., *The Great Good Place*, New York 1989 — [D] (Kap. 6)
- Paivio, A., *Mental Representations: A Dual Coding Approach*, Oxford 1986 — [D] (Kap. 3)
- Paivio, A., Dual coding theory: Retrospect and current status, *Canadian J Psychology* 45 (1991) 255–287 — [D] (Kap. 3)
- Patel, A., *Music, Language, and the Brain*, Oxford 2008 — [D] (Kap. 5)
- Pearce, E./Launay, J./Dunbar, R. I. M., The ice-breaker effect, *Royal Society Open Science* 2 (2015) 150221 — [V, Abstract wörtlich] (Kap. 6, 9)
- Pettigrew, T. F./Tropp, L. R., A meta-analytic test of intergroup contact theory, *J Personality and Social Psychology* 90 (2006) 751–783 — [V, Abstract] (Kap. 8)
- Racette, A./Peretz, I., Learning lyrics: to sing or not to sing?, *Memory & Cognition* 35 (2007) 242–253 — [V, Abstract] (Kap. 5, 8)
- Richards, J. C./Rodgers, T. S., *Approaches and Methods in Language Teaching*, 3. Aufl., Cambridge 2014 — [D] (Kap. 7)
- Rosa, H., *Resonanz. Eine Soziologie der Weltbeziehung*, Berlin 2016 — [D] (Kap. 8)
- Ryan, R. M./Deci, E. L., Self-determination theory and the facilitation of intrinsic motivation, *American Psychologist* 55 (2000) 68–78 — [D] (Kap. 2)
- Rupérez Micola, A./Aparicio Fenoll, A./Banal-Estañol, A./Bris, A., TV or not TV? The impact of subtitling on English skills, *J Economic Behavior & Organization* 158 (2019) 487–499 — [V, Volltext] (Kap. 3)

- Schön, D. et al., Songs as an aid for language acquisition, *Cognition* 106 (2008) 975–983 — [V, Abstract; Experimentdetails am Volltext zu prüfen] (Kap. 5)
- Seume, J. G., *Die Gesänge* (1804) — [V, Digitalisate] (Motto Teil IV)
- Slamecka, N. J./Graf, P., The generation effect, *JEP: Human Learning and Memory* 4 (1978) 592–604 — [V, Abstract] (Kap. 8)
- Streifler, D., *Kalibriertes Vertrauen*, ra.de-Verlag 2026 — [V, Schwesterband der Reihe] (Kap. 4)
- Swain, M., Communicative competence, in: Gass/Madden (Hg.), *Input in SLA*, Rowley 1985, S. 235–253 — [D] (Kap. 2, 9)
- Teimouri, Y./Goetze, J./Plonsky, L., Second language anxiety and achievement: A meta-analysis, *SSLA* 41 (2019) 363–387 — [D] (Kap. 2)
- Twain, M., *A Tramp Abroad*, Anhang D: The Awful German Language (1880) — [V, Wikisource-Transkription] (Motto Teil I)
- UrhG (Urheberrechtsgesetz), § 60a (mit § 60h) — [V, gesetze-im-internet.de, abgerufen 2026-07-25] (Kap. 9)
- Vanderplank, R., The value of teletext sub-titles in language learning, *ELT Journal* 42 (1988) 272–281 — [D] (Kap. 3)
- Vanderplank, R., *Captioned Media in Foreign Language Learning and Teaching*, London 2016 — [D] (Kap. 3)
- Vygotskij, L. S., *Mind in Society*, Cambridge/MA 1978, S. 86 — [V, Standardrezeption mehrfach; am Druck zu prüfen] (Kap. 7)
- Wallace, W. T., Memory for music: Effect of melody on recall of text, *JEP: LMC* 20 (1994) 1471–1485 — [V, Abstract/Rezeption; am Volltext zu prüfen] (Kap. 5, 8)
- Wang, A./Pellicer-Sánchez, A., Incidental vocabulary learning from bilingual subtitled viewing, *Language Learning* 72 (2022) — [D] (Kap. 3)
- Welch, G. F. et al., Singing and social inclusion, *Frontiers in Psychology* 5 (2014) 803 — [V, PMC-Volltext] (Kap. 6, 8)
- Winke, P./Gass, S./Sydorenko, T., The effects of captioning videos, *Language Learning & Technology* 14/1 (2010) 65–86 — [V, Volltext] (Kap. 3)
- Wood, D./Bruner, J. S./Ross, G., The role of tutoring in problem solving, *J Child Psychology and Psychiatry* 17 (1976) 89–100 — [D, Verlagszugriff gesperrt] (Kap. 7)
- Yanagisawa, A./Webb, S./Uchihara, T., How do different forms of glossing contribute to L2 vocabulary learning from reading?, *SSLA* 42 (2020) 411–438 — [V, Verlagsabstract] (Kap. 4, 7, 8)

Abkürzungen

- **BAMF** — Bundesamt für Migration und Flüchtlinge
- **DTZ** — Deutsch-Test für Zuwanderer
- **ES** — Effektstärke
- **FLCAS** — Foreign Language Classroom Anxiety Scale
- **GER** — Gemeinsamer europäischer Referenzrahmen für Sprachen
- **IntV** — Integrationskursverordnung
- **INVALSI** — Istituto nazionale per la valutazione del sistema educativo di istruzione e di formazione (italienisches Institut für Bildungsevaluation)
- **JEP** — Journal of Experimental Psychology
- **L1 / L2** — Erstsprache (Muttersprache) / Zweit- bzw. Fremdsprache
- **LLT** — Language Learning & Technology (Zeitschrift)
- **SLA** — Second Language Acquisition (Zweitspracherwerb)
- **SSLA** — Studies in Second Language Acquisition (Zeitschrift)
- **TBLT** — Task-Based Language Teaching (aufgabenbasierter Sprachunterricht)
- **TOEFL** — Test of English as a Foreign Language
- **UE** — Unterrichtseinheit(en)
- **ZPD** — Zone of Proximal Development (Zone der nächsten Entwicklung)

Glossar

- **Affektiver Filter** — Krashens Bild für die Blockade zwischen verstandenem Input und Spracherwerb durch Angst, fehlendes Selbstvertrauen oder fehlende Motivation; hoher Filter = wenig kommt an (Kap. 2).
- **Captions / zielsprachliche Untertitel** — Untertitel in derselben Sprache wie der Ton; segmentieren den Lautstrom und verbinden Klang mit Schrift (Kap. 3).
- **ChordPro** — Textformat für Liednoten, das Akkordsymbole in den Liedtext einbettet; Standard der Mitsing-Praxis, Grundlage des Liederbuch-Moduls (Kap. 5).
- **Doppel-Untertitel (bilinguale Untertitel)** — gleichzeitige Anzeige von Ziel- und Muttersprachzeile; Pontivos Player-Bauform (Kap. 3).
- **Dritter Ort** — Oldenburgs Begriff für neutrale, statusnivellierende Gemeinschaftsräume neben Zuhause und Arbeit; in diesem Buch: der Chor (Kap. 6).
- **Dual Coding / doppelte Kodierung** — Paivios Theorie zweier verbundener Gedächtnissysteme (sprachlich und bildlich); doppelt kodierte Inhalte haben mehr Abrufwege (Kap. 3).
- **DTZ** — „Deutsch-Test für Zuwanderer", skalierter Abschlusstest des Integrationskurses (B1/A2) (Kap. 1).
- **FLCAS** — Foreign Language Classroom Anxiety Scale; Standardskala zur Messung von Fremdsprachenangst (Horwitz et al. 1986) (Kap. 2, 9).
- **Flow** — Csíkszentmihályis Zustand selbsttragenden Tuns, in dem die Anstrengung aus dem Erleben verschwindet (Kap. 2).
- **Generierungseffekt** — Selbst erzeugte Inhalte werden besser behalten als bloß gelesene (Slamecka & Graf 1978) (Kap. 8).
- **GER** — Gemeinsamer europäischer Referenzrahmen für Sprachen; Niveaustufen A1–C2, B1 = „Schwellenniveau" (Kap. 1).
- **Glosse** — kurze Worterklärung am Rand eines Textes; digital: die Klick-Erklärung (Kap. 4).
- **i + 1** — Krashens Formel für Input knapp über dem aktuellen Stand *i*; Erwerb geschieht beim Verstehen solchen Inputs (Kap. 2).
- **Inzidentelles (beiläufiges) Lernen** — Lernen als Nebenprodukt verstehender Beschäftigung, ohne Lernabsicht (Kap. 4).
- **Involvement Load** — Laufer/Hulstijns Drei-Faktoren-Modell der Verarbeitungstiefe beim Wortlernen: Bedürfnis, Suche, Bewertung (Kap. 4).
- **Kontakt-Hypothese** — Allports Bedingungen, unter denen Begegnung zwischen Gruppen Vorurteile abbaut (Statusgleichheit, gemeinsames Ziel); meta-analytisch bestätigt durch Pettigrew/Tropp (Kap. 8).
- **Lieder-Spende** — Gegenseitigkeitsmodell dieses Konzepts: Bühnen- und Probennutzung des Liederbuchs gegen Direktlizenz einzelner Werke für den Lernraum (Kap. 9).
- **Merk-Anker** — das Prüfraster dieses Buches: „Verständlich? Verankert? Verbunden?" (Kap. 2).
- **Misgrounding** — ein Beleg, der existiert, aber die Aussage nicht trägt (Begriff aus dem Schwesterband *Kalibriertes Vertrauen*) (Kap. 4).

- **Psychologische Sicherheit** — Edmondsons geteilte Überzeugung einer Gruppe, dass zwischenmenschliches Risiko (fragen, irren) nicht bestraft wird; Vorbedingung des Lernens (Kap. 6).
- **Rechte-Paradox** — Befund dieses Buches: Das sprachlich und emotional wirksamste Material ist im Lernraum am schwersten rechtssicher nutzbar (Kap. 9.4).
- **Scaffolding (Gerüst)** — abbaubare Stützstruktur, die ein Könnender für einen Lernenden errichtet (Wood/Bruner/Ross 1976) (Kap. 7).
- **Selbstbestimmungstheorie (SDT)** — Deci/Ryans Motivationstheorie: Autonomie, Kompetenzerleben und soziale Eingebundenheit als Grundbedürfnisse, aus denen sich Engagement selbst trägt (Kap. 2).
- **Translanguaging** — Verständnis der Mehrsprachigkeit als ein gemeinsames Repertoire; pädagogisch: die Erstsprache als legitimes Werkzeug im Unterricht (Kap. 7).
- **Zone der nächsten Entwicklung (ZPD)** — Vygotskijs Spanne zwischen dem, was ein Lernender allein kann, und dem, was er mit Hilfe kann; der Ort des Lernens (Kap. 7).

Zum Konzept & Team

Pontivo (pontivo.de) ist eine Plattform für Spracherwerb mit Film, Lied und Gemeinschaft: authentisches Video mit synchronem, zweisprachigem Untertitel-Player und Klick-Erklärung; ein Liederbuch mit Akkorden für gemeinsames Singen in Gruppen mit Lehrer-/Schüler-Modus; ein Klassenzimmer für Lern-Szenarien über dieselbe mehrsprachige Engine; und der Comic „Coro One Sky" als erzählerisches Dach. Die Plattform wird von einem kleinen Team um die ra.de Verlags-KG entwickelt.

Dieses Buch ist zu Pontivo, was der Schwesterband *Kalibriertes Vertrauen* zu trust-on.net ist: die offene wissenschaftliche Begründung des Ansatzes — einschließlich der Liste dessen, was noch nicht bewiesen ist. Es erscheint als Band 1 der Reihe **„Schriften zur Sprachbildung mit Film und Musik"** im ra.de-Verlag (Diamond Open Access: frei lesbar, ohne Gebühren für Leser wie Autoren).

Prüfbarkeit. Jede Fundstelle dieses Buches trägt einen Prüfstatus (an der Primärstelle verifiziert oder ausdrücklich „am Druck zu prüfen"); der Verifikations-Log ist Teil der Publikationsunterlagen. Korrekturen und Widerlegungen sind willkommen und werden in Revisionen mit Änderungsprotokoll eingearbeitet — der Maßstab der Reihe gilt auch für dieses Buch: prüfbar statt behauptet.