

Uwe Neugebauer/Michael Becker-Mrotzek/Petra Stanat

Ermittlung von Sprachförderbedarf bei Kindern im Elementarbereich aus pädagogisch-psychologischer Sicht

1 Einleitung

Am 12.2.2007 war in Spiegel Online zu lesen: „Nordrhein-Westfalen: Sprachtest für Vierjährige wird Pflicht. Als erstes Bundesland setzt Nordrhein-Westfalen auf Sprachtests für so kleine Knirpse. Sie müssen künftig beim Brettspiel ‚Besuch im Zoo‘ zeigen, was sie schon können.“¹ Dies war nur eine von vielen Neuerungen, die in den letzten 15 Jahren im Bereich der Sprachstandsfeststellung und Sprachförderung in Deutschland umgesetzt worden sind. Den Ausgangspunkt der Maßnahmen bildete die Erfahrung, dass Kinder aus zugewanderten und sozial schwachen Familien oft schon bei Eintritt in die Schule über vergleichsweise ungünstige Lernvoraussetzungen verfügen, weil ihr sprachlicher Entwicklungsstand nicht altersgemäß ist. Gleichzeitig hat sich wiederholt gezeigt, dass Entwicklungsrückstände umso leichter und wirksamer reduziert werden können, je früher die Förderung einsetzt.² Aus diesem Grund ist es sowohl aus Sicht der individuellen Kinder als auch aus gesellschaftlicher Sicht sinnvoll, sprachlichen Förderbedarf möglichst früh zu erkennen und anschließend geeignete Fördermaßnahmen umzusetzen. Dies ist jedoch nicht nur aus juristischer, sondern auch aus pädagogisch-psychologischer Perspektive mit erheblichen Herausforderungen verbunden, die konzeptuelle, methodische und praktische Aspekte betreffen und im folgenden Beitrag beschrieben werden sollen. Weiterhin wird der Stand der Entwicklungen skizziert, der im Bereich der Sprachstandsdiagnostik in Deutschland bislang erreicht worden ist.

2 Notwendigkeit und Möglichkeiten früher Sprachdiagnostik

In Diskussionen über die Bildungsbenachteiligung von Kindern und Jugendlichen mit Zuwanderungshintergrund und Heranwachsenden aus bildungsfernen Familien wurde seit den 1990er Jahren zunehmend der Elementarbereich in den Blick genommen und dabei die zentrale Rolle der frühen sprachlichen Entwicklung betont³. Auf der Grundlage von Forschungsbefunden zu Effekten kompensatorischer Förderung⁴ wird argumentiert, dass sprachliche Defizite so früh wie mög-

1 Erst kürzlich, am 19.11.2013, titelte der WDR „Sprachtest Delfin 4 wird abgeschafft“ und schrieb „Die Sprachfähigkeit der Kinder soll künftig nicht mehr von Außenstehenden geprüft werden, sondern im Alltag von Erziehern“. Damit wurde das 2007 in Nordrhein-Westfalen eingeführte Verfahren der Sprachstandsfeststellung sechs Jahre später wieder aufgegeben.

2 Heckman, J. J., The economics of inequality: The value of early childhood education, *American Educator*, 2011, S. 31–47.

3 Reich, H. H., Tests und Sprachstandsmessungen bei Schülern und Schülerinnen, die Deutsch nicht als Muttersprache haben, in: Bredel, U. et al. (Hrsg.), *Didaktik der deutschen Sprache*, Paderborn 2006, S. 915.

4 Schneider, H./Becker-Mrotzek, M./Sturm, A./Jambor-Fahlen, S./Neugebauer, U./Efing, C./Kernen, N., Wirksamkeit von Sprachförderung, Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache (Hrsg.), Köln 2013; Stanat, P./Felbrich, A., Sprachförderung als Voraussetzung für die Sicherung von Mindeststandards im Bildungssystem: Ansatzpunkte und Herausforderungen, in: Deißner, D. (Hrsg.) *Chancen bilden. Wege zu einer gerechteren Bildung – ein internationaler Erfahrungsaustausch*, Wiesbaden 2013, S. 79–100; Paetsch, J./Wolf, K. M./Stanat, P./Darsow, A., Sprachförderung von Kindern und Jugendlichen aus Zuwandererfamilien, *Zeitschrift für Erziehungswissenschaft*, Son-

lich erkannt werden sollten, um Fördermaßnahmen im Elementarbereich zu initiieren und allen Kindern gute Startchancen zu Beginn der Grundschule zu ermöglichen.⁵

Diese Entwicklung hat sich durch den sogenannten „PISA-Schock“ weiter intensiviert. Die 2001 veröffentlichten Ergebnisse der ersten PISA-Studie⁶ wiesen darauf hin, dass Schülerinnen und Schüler aus zugewanderten und sozial schwachen Familien im schulischen Kompetenzerwerb in Deutschland besonders stark benachteiligt waren. Ausgehend von der Annahme, dass Sprache eine wichtige Voraussetzung für erfolgreiche Bildungsprozesse ist und möglichst früh gefördert werden sollte, identifizierte die Kultusministerkonferenz als Konsequenz aus PISA 2000 insgesamt sieben Handlungsfelder, von denen sich die erste auf „Maßnahmen zur Verbesserung der Sprachkompetenz bereits im vorschulischen Bereich“ bezog. Weiterhin verpflichteten sich die Bundesländer im Nationalen Integrationsplan von 2007, vor der Einschulung Verfahren zur Feststellung und Beobachtung des Sprachstandes einzusetzen und je nach Bedarf eine Sprachförderung zu ermöglichen.⁷ Da jedoch die Aktivitäten in diesem Bereich kaum länderübergreifend koordiniert wurden, ist das Vorgehen bei der Sprachstandsdiagnostik und Sprachförderung in Kindertagesstätten sehr unterschiedlich, so dass sich eine „[...] große Heterogenität zwischen den verschiedenen Ländern bei der Umsetzung dieser Vorhaben“⁸ zeigt. Dies betrifft zum einen, ob Sprachstandserhebungen verpflichtend für alle Kinder durchgeführt werden und in welchem Alter die Erhebungen stattfinden. Zum anderen unterscheiden sich die Länder darin, welche Verfahren eingesetzt werden und welche Konsequenzen mit den Ergebnissen verknüpft sind. Die konkreten Fördermaßnahmen, die bei einem identifizierten Förderbedarf folgen, variieren zwischen den Ländern ebenfalls erheblich.

In den Ländern kommen vier Arten von Verfahren der Sprachstandsfeststellung zum Einsatz: Beobachtungsbögen, Einschätzverfahren, Testverfahren und Screenings. Bei den *Beobachtungsbögen* werden möglichst konkrete, verhaltensnahe Beschreibungen vorgelegt (z. B. „Das Kind kann Oberbegriffe finden und verwenden, wie z. B. Kleidung, Obst ...“), die zumeist durch Erzieherinnen auf der Grundlage von Beobachtungen des Kindes im Alltag zu beurteilen sind (z. B. anhand solcher Kategorien wie „trifft häufig zu“ oder „trifft selten/nie zu“). *Einschätzverfahren* verwenden dagegen eher allgemein gehaltene Beschreibungen, wie z. B. „Das Kind kann Gegenstände des täglichen Lebens in der deutschen Sprache benennen“, die etwa als „ausreichend“ oder „nicht ausreichend“ einzuschätzen sind.⁹ Bei *Testverfahren* werden dem Kind dagegen gezielt Aufgaben gestellt und dessen Reaktionen darauf dokumentiert. Dabei können verschiedene Aspekte sprachlicher Kompetenz erfasst und vielfältige Aufgabenstellungen eingesetzt werden, wie etwa die Bestimmung des Wortschatzumfangs eines Kindes durch die Benennung von Objekten oder die Beschreibung von Handlungen auf Bildkarten. Insbesondere mit Testbatterien, die mehrere Dimensionen der sprachlichen Entwicklung messen, lassen sich zudem spezifische Stär-

derband Herkunft und Bildungserfolg von der Vorschule bis zur Universität: Forschungsstand und Interventionsmöglichkeiten aus interdisziplinärer Perspektive (im Druck).

5 Kniffka, G./Siebert-Ott, G., Deutsch als Zweitsprache in der Schule, Paderborn 2007, S. 122.

6 Baumert, J./Klieme, E./Neubrand, M./Prenzel, M./Schiefele, U./Schneider, W./Stanat, P./Tillmann, K.-J./Weiss, M. (Hrsg.), PISA 2000: Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich, Opladen 2003.

7 Die Bundesregierung, Der Nationale Integrationsplan. Neue Wege – neue Chancen, Berlin 2007, S. 52 ff.

8 Lisker, A., Sprachstandsfeststellung und Sprachförderung im Kindergarten und beim Übergang in die Schule, München 2010, S. 3, sowie Neugebauer, U./Becker-Mrotzek, M., Die Qualität von Sprachstandsverfahren im Elementarbereich. Eine Analyse und Bewertung, Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache (Hrsg.), Köln 2013, Schneider, W./Baumert, J./Becker-Mrotzek, M./Hasselhorn, M./Kammermeyer, G./Rauschenbach, T./Roßbach, H.G./Roth, H.-J./Rothweiler, M./Stanat, P., Expertise „Bildung durch Sprache und Schrift (BISS)“, erhältlich unter http://www.bmbf.de/pubRD/BISS_Expertise.pdf (26.04.2014).

9 Beispiel entnommen aus dem Verfahren „Früh Deutsch lernen“ des Ministeriums für Bildung, Saarland.

ken- und Schwächenprofile bestimmen, die für die Gestaltung der Förderung hilfreich sein können. Bei einem *Screening* schließlich handelt es sich ebenfalls um ein Testverfahren, das jedoch zunächst nur allgemein identifizieren soll, ob ein Kind Sprachförderbedarf hat oder nicht. Eine genauere Diagnose, auf dessen Grundlage z. B. eine spezifische Förderung geplant werden kann, ist anhand eines Screenings nicht möglich.

3 Erfassung von Sprachkompetenz

3.1 Der kindliche Spracherwerb

Der Spracherwerb von Kindern gehört zu den faszinierendsten Entwicklungen im Laufe des Lebens eines Menschen: Scheinbar beiläufig und ohne besondere Anstrengungen eignen sich Kinder die Sprache ihrer Eltern bzw. Umgebung an und nutzen diese beim täglichen Spielen und Kommunizieren mit anderen. Ab einem Wortschatzumfang von 150 Wörtern (mit ca. 1,0–2,5 Jahren) beginnt die Entdeckung der Syntax, d. h. es werden erste Zwei- und Dreiwortäußerungen gebildet und grammatische Regeln erprobt.¹⁰ Zwischen Kindern gleichen Alters bestehen schon sehr früh erhebliche Unterschiede in der Sprachentwicklung. Diese beziehen sich primär auf den Zeitpunkt und die Geschwindigkeit des Erwerbs neuer Fähigkeiten (siehe hierzu z. B. den Elternfragebogen ELFRA 2 zu sprachlichen Fähigkeiten von zweijährigen Kindern¹¹). Bereits bei einjährigen Kindern variiert die Anzahl der verwendeten Wörter (aktiver Wortschatz) zwischen Null und 199. Mit 2,6 Jahren hat sich die Spannweite auf 218–565 Wörter, mit vier Jahren auf 1.500–4.000 Wörter erhöht¹². Diese Variabilität findet sich auch in den anderen sprachlichen Bereichen, wie z. B. dem Grammatikerwerb.

Im Alter von ca. 3,5–4,0 Jahren erreicht die Sprachfähigkeit des Kindes ein Niveau, welches differenzierte sprachliche Äußerungen in Bezug auf Wortschatz, Syntax und sprachliches Handeln ermöglicht. Kinder in diesem Alter können sich in der Regel mit ihrer sozialen Umwelt sicher verständigen. Im Alter von 4 bis 6 Jahren nimmt der Wortschatzumfang weiter zu, neue Wort- und Satzformen sowie sprachliche Handlungen werden angeeignet. Ab einem Alter von ca. 9 bis 12 Jahren schließlich ist der allgemeine Spracherwerb weitgehend abgeschlossen und es sind keine ausgeprägten Leistungsunterschiede mehr zu beobachten, die systematisch mit dem Alter variieren. Es bestehen jedoch weiterhin individuelle Leistungsunterschiede zwischen Kindern im gleichen Alter, die erheblich sein können. Diese Varianz in den individuellen Erwerbsverläufen wirft die Frage auf, welche Entwicklungen als problematisch zu erachten sind und einer besonderen Förderung bedürfen.

3.2 Definition von Sprachkompetenz

Aus pädagogisch-psychologischer Sicht erfordert die Erfassung des Konstrukts „Sprachkompetenz“ bzw. „Sprachförderbedarf“ zunächst, dass dieses konzeptuell definiert wird. Eine wissenschaftlich fundierte, interdisziplinär akzeptierte Definition existiert bislang jedoch nicht. Ein Indiz hierfür sind etwa die unterschiedlichen Fachtermini wie „Sprachdefizit“, „verzögerte Sprachent-

10 Szagun, G., Sprachentwicklung beim Kind. Ein Lehrbuch, Weinheim und Basel 2011, S. 126 ff.

11 Grimm, H./Doil, H., Elternfragebogen für die Früherkennung von Risikokindern, ELFRA 2: Elternfragebogen für zweijährige Kinder: Sprache und Kommunikation, Göttingen 2000.

12 Vgl. hierzu Szagun (Anm. 10), Kapitel 8: Individuelle Unterschiede beim Spracherwerb, S. 206–240.

wicklung“ oder „unausgeglichene Sprachentwicklung“, die in der Literatur teils synonym, teils spezifisch verwendet werden. Implizit wird unter einem „Sprachförderbedarf“ verstanden, dass die sprachlichen Kompetenzen eines Kindes so unzureichend entwickelt sind, dass es ohne besondere Förderung Schwierigkeiten haben wird, einen erfolgreichen Einstieg in schulisches Lernen zu finden. Dabei geht es ausdrücklich nicht um Sprachentwicklungsstörungen im engen Sinne, die u. a. durch gravierende Defizite und atypische Entwicklungsverläufe gekennzeichnet sind. Solche Störungen erfordern spezifische Verfahren der Diagnostik und Maßnahmen der Förderung, die bereits existieren und etabliert sind.¹³

Eine mögliche Systematisierung des Konstrukts „Sprachliche Kompetenz“ haben Ehlich et al. vorgelegt¹⁴, die den Prozess der Sprachaneignung als Interaktion von sieben sog. Basisqualifikationen beschreiben. Trotz Einschränkungen ist diese Kategorisierung hilfreich, um zu bestimmen, welche sprachlichen Bereiche mit den in Deutschland eingesetzten Sprachstandsverfahren bzw. Aufgabentypen erfasst werden und inwieweit die Konstrukte und ihre Operationalisierung durch empirische Forschung fundiert sind. Als basal gilt dabei die *phonische Basisqualifikation*, die als die Fähigkeit definiert ist, Laute der Zielsprache zu rezipieren und zu produzieren. Die *semantische Basisqualifikation* umfasst sowohl den aktiven und passiven Wortschatz als auch die Begriffsbildung, während die *morphologisch-syntaktische Basisqualifikation* (Bildung von Wortformen durch Flexion und Stellung der Wörter im Satz) den Grammatikerwerb beinhaltet. Zu diesen sprachlichen Basisqualifikationen im engeren Sinne kommen die handlungsbezogenen-pragmatischen und diskursiven Fähigkeiten (etwa auf Fragen zu antworten, Anweisungen auszuführen oder Selbsterlebtes zu erzählen), deren Entwicklung jedoch bislang kaum untersucht worden ist. Auch ist nur teilweise bekannt, in welchem Verhältnis die Basisqualifikationen zueinander stehen, d. h. in welchem Grade sie untereinander korrelieren und miteinander interagieren. Zudem ist nur für einen Teil dieser Fähigkeiten hinreichend geklärt, ob sie besonders prädiktiv für einen erfolgreichen Einstieg in das schulische Lernen sind und daher bei Sprachstandserhebungen in jedem Fall erfasst werden sollten. Die Herausforderung, den Nachweis prädiktiven Wertes zu führen, ist für Erhebungsverfahren besonders groß, die sich an der sog. natürlichen Sprachproduktion (v. a. diskursive und pragmatische Basisqualifikation) orientieren.

3.3 Aufgabentypen zur Erfassung sprachlicher Teilfähigkeiten

Eine weitere Herausforderung besteht in der „Kluft“ zwischen alltäglicher Sprachproduktion und Testaufgaben¹⁵, die der aktiven Sprachproduktion zugrundeliegende Fähigkeiten erfassen sollen. Wenn ein vierjähriges Kind Nonsens-Sätze wie z. B. „*Wenn die Insel lacht, wandert sie in den Tag*“ oder Kunstwörter wie z. B. „*Knegiba*“ nachsprechen soll, ist nicht unmittelbar ersichtlich, welchen Zweck das haben soll bzw. welche dahinter stehenden sprachlichen Fähigkeiten überprüft werden – weshalb sowohl Erzieherinnen als auch Fachwissenschaftlerinnen und -wissenschaftler an solchen Aufgaben Kritik üben, die jedoch im Falle von nachweislich validen Verfahren unberechtigt ist (s. u.). Wenn ein Kind Nonsens-Sätze, die grammatisch korrekt sind, nachsprechen kann, so ist dies ein Indikator für eine hohe morphologisch-syntaktische Kompetenz. Dies kann

13 *Kannengieser, S.*, Sprachentwicklungsstörungen. Grundlagen, Diagnostik, Therapie, München 2012.

14 *Ehlich, K. et al.*, Anforderungen an Verfahren der regelmäßigen Sprachstandsfeststellung als Grundlage für die frühe und individuelle Sprachförderung von Kindern mit und ohne Migrationshintergrund, BMBF (Hrsg.), Bildungsforschung 11, Berlin 2007.

15 Die im Folgenden verwendeten Beispiele entstammen den Verfahren SETK 3-5 (*Grimm/Aktas/Frevert*, 2010) Delfin 4 (*Fried/Briedigkeit*, 2007) und „Fit in Deutsch“ (*Niedersächsisches Kultusministerium*, 2007).

kontrastiert werden mit syntaktisch falschen Sätzen („Wenn die Insel lacht, Tag sie in den wandert“), die Kindern mit bereits ausgebauten grammatischen Fähigkeiten erheblich schwerer fallen als Kindern mit geringeren Kompetenzen in diesem Bereich, für die beide Sätze gleich schwer sind. Kann ein Kind Wörter wie „Knegiba“ nachsprechen, ist es offensichtlich in der Lage, phonische Lautketten zu identifizieren und zu reproduzieren, was eine unabdingbare Voraussetzung für einen ungestörten Spracherwerb darstellt. Kunstwörter haben gegenüber tatsächlich existierenden Wörtern den Vorteil, dass die Kinder sie nicht kennen, die erfolgreiche Aufgabebearbeitung daher nicht vom Vorwissen abhängt.¹⁶

Außer der phonischen und der morphologisch-syntaktischen Basisqualifikation wird häufig die semantische Fähigkeit geprüft, also der passive und aktive Wortschatz sowie die Begriffsbildung. Dazu wird das Kind aufgefordert, Bilder in unterschiedlicher Weise zu benennen oder zu beschreiben: „Wo ist der Müllmann?“, „Was macht das Kind?“, „Welches Kind lacht gerade?“. Die Begriffsbildung wird über Fragen erfasst wie „Was kann man mit einem Löffel / einem Stift / einem Ball alles machen?“; „Welche Sachen sind alle rot?“ oder auch „Erzähle mir mal, was du zum Anziehen hast / was es für Vögel / Autos gibt“.

Die Erfassung von pragmatischen und diskursiven Fähigkeiten ist dagegen weitaus schwieriger. Am einfachsten ist es zu beobachten, ob ein Kind auf Handlungsanweisungen angemessen reagiert (z. B. „Bringe mir mal den Ball, nachdem Du Deinen Becher weggebracht hast“) oder das Angebot zu einem Sprecherwechsel nutzt (z. B. „Wer möchte jetzt etwas erzählen?“, „Hat jemand Lust, einen Vorschlag zu machen?“). Hierbei ist es besonders herausfordernd, die Standardisierung der Beobachtungssituation sicherzustellen und ob es an Kompetenzen liegt, wenn das Kind nicht angemessen reagiert: Hat es die Aufforderung nicht verstanden oder ist es z. B. zu schüchtern?

4 Güte der Erfassung sprachlicher Fähigkeiten und sprachlichen Förderbedarfs

Sprachliche Kompetenz umfasst also eine Vielzahl von Facetten, die für den Bildungserfolg von Kindern und Jugendlichen bedeutsam sind. Bei Merkmalen wie „Wortschatz“ oder „Erzählfähigkeit“ handelt es sich um psychologische Konstrukte, die sich – anders als etwa die Körpergröße einer Person – nicht direkt messen lassen. Vielmehr ist es erforderlich, Indikatoren für diese Konstrukte zu identifizieren, die beobachtbar sind, und Instrumente zu entwickeln, die zur Messung der Indikatoren eingesetzt werden können. Für die Entwicklung solcher Instrumente gelten Standards, die in der psychologischen Diagnostik allgemein anerkannt sind.¹⁷ Die Hauptgütekriterien sind dabei die Objektivität, die Reliabilität und die Validität, die im Folgenden kurz beschrieben werden sollen, um zu verdeutlichen, wie anspruchsvoll es ist, solide Instrumente zur Erfassung von Sprachförderbedarf bereitzustellen.

Die *Objektivität* eines Messverfahrens betrifft die Frage, ob die Messungen unabhängig davon sind, wer sie durchgeführt, ausgewertet und interpretiert hat. Bei der Sprachstandsdiagnostik geht

16 Hasselhorn, M./Seidler-Brandler, U./Körner, K., Ist das „Nachsprechen von Kunstwörtern“ für die Entwicklungsdiagnostik des phonologischen Arbeitsgedächtnisses geeignet?, in: Hasselhorn, M., Schneider, W./Marx, H. (Hrsg.), Diagnostik von Lese-Rechtschreibschwierigkeiten, Tests und Trends. Jahrbuch der pädagogisch-psychologischen Diagnostik, Göttingen 2000, S. 119–133.

17 Vgl. z. B. Brunner, M./Stanat, P./Pant, H. A., Diagnostik und Evaluation, in: Seidel, T./Krapp, A. (Hrsg.), Pädagogische Psychologie, Weinheim (im Druck).

es also darum, ob bei einem bestimmten Kind dasselbe Urteil über seine Sprachkompetenz bzw. über seinen Förderbedarf resultieren würde, wenn unterschiedliche Personen das Messinstrument anwenden. Je standardisierter die Durchführung, Auswertung und Interpretation bei einem Verfahren verläuft, desto höher ist die Objektivität der Messungen. Dies ist bei Sprachstandsverfahren im Elementarbereich eine erhebliche Herausforderung, da Kinder in diesem Alter nur bedingt in der Lage sind, Instruktionen zu folgen. Eng umrissene Aspekte des Sprachstands, wie zum Beispiel der Wortschatz, lassen sich dabei noch recht gut standardisiert erfassen, etwa indem Kindern Bildkarten mit klar erkennbaren Objekten oder Tätigkeiten vorgelegt werden, die sie benennen sollen¹⁸. Anhand einer Liste von möglichen Antworten wird dann für jede Bildkarte beurteilt, ob das Kind eine korrekte Bezeichnung genannt hat oder nicht. Deutlich schwieriger ist dagegen die objektive Erfassung sprachlicher Kompetenz mit sehr offenen Aufgabenstellungen, wie etwa beim Hamburger Verfahren zur Analyse des Sprachstands Fünfjähriger (HAVAS-5), bei dem das Kind aufgefordert wird, die in einer Folge von sechs Bildern dargestellte Geschichte zu erzählen¹⁹. Die Aufzeichnung der Erzählung wird anschließend anhand eines Auswertungsbogens in Bezug auf die Bewältigung der Aufgabe, das kommunikative Handeln, die Grammatik und den Wortschatz analysiert. Sowohl die Durchführung der Erhebung als auch die anschließende Auswertung der Äußerungen lassen beim HAVAS-5 deutlich mehr Spielraum zu als beim stärker standardisierten Wortschatztest, und es ist daher erheblich schwieriger, objektive Urteile zu gewährleisten. Der Grad der Objektivität eines Messverfahrens wird in der Regel anhand der Übereinstimmung von Urteilen, zu denen zwei oder mehr Anwender des Instruments unabhängig voneinander bei denselben Kindern gekommen sind, bestimmt. Eine ausreichend hohe Objektivität kann bei Sprachstandsfeststellungen im Allgemeinen nur dann erreicht werden, wenn die Personen, die das jeweilige Verfahren anwenden sollen, darin geschult wurden. Je offener das Verfahren angelegt ist, wie etwa im Falle des HAVAS-5, desto größer ist der Schulungsbedarf.

Bei der *Reliabilität* (Messgenauigkeit) geht es um die Frage, wie genau die Messwerte das zu erfassende Konstrukt abbilden, und zwar unabhängig davon, ob das Instrument tatsächlich das Konstrukt misst, das es messen soll (letzteres betrifft die Validität, s. u.). Zur Bestimmung der Reliabilität stehen mehrere Methoden zur Verfügung, die auf Zusammenhängen zwischen mehrfachen Messungen mit dem Instrument basieren (z. B. wiederholter Einsatz des Instruments zu verschiedenen Zeitpunkten, zeitgleicher Einsatz von parallelen Fassungen des Instruments). Weisen Kinder, die bei der einen Messung hohe Werte erzielt haben, auch bei der anderen Messung hohe Werte auf, so ist dies ein Hinweis darauf, dass das Instrument präzise ist. Fällt der Zusammenhang hingegen gering aus, wäre der Messfehler sehr groß und das Instrument als unreliabel zu bewerten.

Die *Validität* (gelegentlich auch als „Gültigkeit“ bezeichnet) schließlich betrifft die Frage, ob das Instrument das misst, was es messen soll. Dies umfasst bei Sprachstandsverfahren, die als Grundlage für Förderentscheidungen herangezogen werden, drei Aspekte. Erstens sollte der Nachweis geführt werden, dass tatsächlich das angezielte Konstrukt, wie etwa der Wortschatz oder die Erzählfähigkeit, erfasst wird und nicht z. B. die allgemeine Sprechfreude, die Schüchternheit oder der Grad der Extraversion eines Kindes. Zweitens wäre zu prüfen, ob das erfasste Konstrukt für den Schulerfolg relevant ist. Dieser Aspekt ist besonders wichtig, da Sprachstandserhebungen im Elementarbereich nicht als Selbstzweck durchgeführt werden, sondern zur Identifikation von Kindern dienen sollen, die in ihrer sprachlichen Entwicklung gezielt unterstützt werden müssen, um

18 Wie z. B. beim „Aktiven Wortschatztest für 3- bis 5-jährige Kinder“: Kiese-Himmel, C., AWST-R., Göttingen 2005.

19 Reich, H. H./Roth H.-J., HAVAS-5. Hamburger Verfahren zur Analyse des Sprachstandes bei 5-Jährigen, Landau/ Hamburg 2003. Unveröffentlichtes Manuskript.

ihnen einen erfolgreichen Einstieg in schulisches Lernen zu ermöglichen. Vor diesem Hintergrund müsste vor dem Einsatz eines Instruments gezeigt werden, dass das damit erfasste Konstrukt tatsächlich für die schulische Leistungsentwicklung prädiktiv ist. Dies erfordert jedoch relativ aufwändige Längsschnittstudien und wurde bislang lediglich für drei der derzeit 21 Verfahren, die in Deutschland zu Einsatz kommen, im Ansatz überprüft (DESK 3-6, SETK 3-5, Vorstellungsverfahren Viereinhalbjähriger) sowie für ein Verfahren (Heidelberger Auditives Screening in der Einschulungsdiagnostik, HASE) umfassender, über einen Zeitraum von fünf Jahren.²⁰ Dabei wurde u. a. der Zusammenhang zwischen der Fähigkeit, mit vier Jahren Sätze nachzusprechen, mit dem in der dritten Klasse erreichten Wortschatz ermittelt und gezeigt, dass eine Korrelation von 0.54 besteht. Demnach lässt sich mit diesem Aufgabentyp knapp 30% der Varianz des späteren Wortschatzes aufklären.²¹

Ein dritter Aspekt der Validität betrifft den Schwellenwert, der als Grundlage für die Entscheidung herangezogen wird, ob ein Kind gezielte Förderung erhalten sollte oder nicht. Im Idealfall läge dieser Entscheidung eine kriteriale Bezugsnorm zugrunde, die angibt, ab welchem Messwert Kinder aufgrund von sprachlichen Schwierigkeiten mit hoher Wahrscheinlichkeit Probleme beim Übergang in die Schule haben werden und daher einer besonderen Förderung bedürfen. Dieser Schwellenwert sollte anhand von theoretisch und empirisch fundierten Erwägungen (z. B. „Welchen Wortschatzumfang benötigt ein Kind mindestens, um in der Schulanfangsphase erfolgreich lernen zu können?“) festgelegt und mit Längsschnittstudien validiert werden. In der Regel wird bei Entscheidungen über den Förderbedarf jedoch die soziale Bezugsnorm angewendet, d. h. es werden diejenigen Kinder als förderbedürftig klassifiziert, die im jeweiligen Verfahren die geringsten Leistungen erzielen (z. B. die leistungsschwächsten 16 Prozent). Auch ein solcher sozialnormorientierter Schwellenwert sollte validiert werden, was ebenfalls nur bei den wenigsten Sprachstandsverfahren, die in Deutschland eingesetzt werden, erfolgt ist (s. u.).

Die drei Hauptgütekriterien der psychologischen Diagnostik stehen insofern in einem hierarchischen Verhältnis zueinander, als ein Verfahren, das nicht objektiv ist, auch nicht reliabel sein kann und für Validität Messungen vorausgesetzt werden, die objektiv und reliabel sind. Gleichzeitig wird in der Pädagogik aber häufig ein Spannungsverhältnis zwischen Objektivität und Validität gesehen, das sich auf die Akzeptanz von Sprachstandsverfahren auswirkt. Wie oben beschrieben, erfordert eine objektive Erfassung von Aspekten sprachlicher Kompetenz ein gewisses Maß an Standardisierung der Erhebung. Die Künstlichkeit solcher Erhebungssituationen wird jedoch vielfach als Einschränkung der Validität der Messungen empfunden.²² Daher ist bei Pädagoginnen und Pädagogen eine Tendenz zu beobachten, Verfahren zu präferieren, die auf aufwändigen Beobachtungen von Kindern in natürlichen Interaktionssituationen basieren, deren Objektivität – und damit auch deren Reliabilität und Validität – jedoch unklar ist.²³ Stark standardisierte Testungen, wie etwa das Nachsprechen von Sätzen, das sich als objektiver, reliabler und valider Indikator für Schulschwierigkeiten erwiesen hat²⁴, werden dagegen nicht nur von der pädagogischen Praxis,

20 Vgl. Neugebauer, U./Becker-Mrotzek, M. (Anm. 8).

21 Schöler, H., Prognose schriftsprachlicher Leistungen und Risiken im Vorschulalter am Beispiel des Heidelberger Auditiven Screenings in der Einschulungsuntersuchung (HASE), in: Hasselhorn, M./Schneider, W. (Hrsg.), Frühprognose schulischer Kompetenzen, Tests und Trends, 2011. Jahrbuch der pädagogisch-psychologischen Diagnostik, Göttingen 2011.

22 Mischo, C., Über den Umgang mit grundlegenden Problemen von Beobachtungs- und Diagnoseprozessen, Frühe Bildung, 1(2), 2012, S. 106–108.

23 Kliche, T./Wittenborn, C./Koch, U., Was leisten Entwicklungsbeobachtungen in Kitas? Eigenschaften und Verbreitung verfügbarer Instrumente, Praxis der Kinderpsychologie und Kinderpsychiatrie 58, 2009, S. 419–433.

24 Roos, J./Schöler, H./Trettlein, A., Zur prognostischen Validität des Heidelberger Auditiven Screenings in der Einschul-

sondern teilweise auch von Wissenschaftlerinnen und Wissenschaftlern in der Pädagogik und Linguistik abgelehnt²⁵. Die zugrundeliegende Gleichsetzung „hohe Standardisierung = geringe Validität“ ist jedoch unangemessen.

Kontrovers wird auch die Frage diskutiert, inwieweit die jeweiligen Verfahren geeignet sind, den Sprachstand von Kindern mit unterschiedlichen Herkunftssprachen zu erfassen. Dies ist eine Frage der differenziellen Validität: Häufig wird angenommen, dass Verfahren, die dazu geeignet sind, bestimmte Aspekte der sprachlichen Kompetenz einsprachiger Kinder zu bestimmen, nicht zwangsläufig auch für mehrsprachige Kinder valide Messungen liefern. Auf diesen Aspekt wird im Folgenden noch genauer eingegangen.

5 Überblick über Verfahren zur Bestimmung des Förderbedarfs

Eine exakte Bestandsaufnahme der derzeit in Deutschland eingesetzten Verfahren der Sprachstandsfeststellung im Elementarbereich ist schwierig. Im Bildungsbericht 2012²⁶ werden 17 Verfahren genannt, die in 14 Bundesländern eingesetzt werden, während das Deutsche Jugendinstitut in seiner aktualisierten Expertise mit Stand vom Dezember 2012 für 15 Bundesländer 18 unterschiedliche Verfahren aufführt, darunter zwei Verfahren, die in keiner Liste auftauchen²⁷. Nach eigenen Recherchen in den Bundesländern mit Stand vom September/Oktober 2013 werden derzeit 21 unterschiedliche Verfahren eingesetzt, um zu einer Förderentscheidung zu gelangen²⁸. Auch diese Bestandsaufnahme wird jedoch bald wieder angepasst werden müssen – so wurde z. B. das Verfahren Delfin 4 in NRW auf Grundlage des Koalitionsvertrages kürzlich abgeschafft (s. o.).

In mindestens fünf Bundesländern werden für Kita-Kinder und Kinder, die keine Kita besuchen, zudem unterschiedliche Verfahren eingesetzt, so z. B. in Schleswig-Holstein die Beobachtungsbögen Sismik oder Seldak bei Kita-Kindern und das Verfahren HAVAS 5 bei Kindern, die nicht in die Kita gehen. Das gleiche Prinzip findet sich auch in Berlin – hier kommt bei Kita-Kindern das Verfahren QuaSta („Qualifizierte Stuserhebung“) zum Einsatz, während Kinder, die keine Kita besuchen, mit dem Screening-Verfahren Deutsch Plus 4 untersucht werden. Dies wird inhaltlich-methodisch damit begründet, dass sich die Informationsbasis für die beiden Gruppen unterscheidet: Ein Kita-Kind kann auf breiter Basis über einen längeren Zeitraum beobachtet werden, während Kinder, die nicht institutionell betreut werden, nur punktuell beobachtet werden können und für sie Screenings oder Testverfahren daher besser geeignet sind. Insgesamt werden in den Bundesländern drei unterschiedliche Beobachtungsbögen, zwei Einschätzverfahren, sieben Screenings, vier Testverfahren sowie fünf gemischte Verfahren (die sowohl Beobachtungsskalen als auch Durchführungsaufgaben umfassen) eingesetzt.

Beobachtungsbögen erfahren dabei eine vergleichsweise breite Akzeptanz bei den Erzieherinnen. Ihr Hauptnutzen besteht darin, die pädagogischen Fachkräfte auf die wesentlichen Sprachsituationen aufmerksam zu machen, und ihre Beobachtungen stärker zu objektivieren und zu doku-

lungsuntersuchung HASE. Abschlussbericht des Projektes EVER, Heidelberg 2007.

25 Valtin, R., Phonologische Bewusstheit: Ein kritischer Blick auf ein modisches Konstrukt, *Frühe Bildung*, 1 (4), 2012, S. 223–225.

26 *Autorengruppe Bildungsberichterstattung*, *Bildung in Deutschland 2012: Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*, Bielefeld 2012.

27 Lisker, A., Sprachstandsfeststellung und Sprachförderung vor der Einschulung. Aktualisierte Expertise im Auftrag des DJI, Deutsches Jugendinstitut e.V., München 2013.

28 Vgl. Neugebauer, U./Becker-Mrotzek, M. (Anm. 8).

mentieren. Für die Verfahren Sismik und Seldak liegen Hinweise darauf vor, dass dies tatsächlich gelingt.²⁹ So berichten Erzieherinnen, dass durch den Einsatz der beiden Verfahren ihre Wahrnehmung der sprachlichen Entwicklung bei den Kindern sehr viel genauer geworden sei. Die Hauptschwäche solcher Beobachtungsbögen liegt allerdings darin, dass ihre Objektivität gering sein dürfte. Zu den wenigsten Verfahren liegen jedoch zuverlässige Angaben zur Objektivität vor, die z. B. anhand von Beobachter-Übereinstimmungen relativ leicht zu bestimmen wären. Zumeist wird darauf vertraut, dass die Beschreibungen des Verhaltens („*Benutzt Verben angemessen*“; „*Vorwiegend sachlich angemessene Nomen*“) hinreichend genau sind, um zu gewährleisten, dass alle Erzieherinnen das gleiche Verhalten darunter verstehen, was jedoch fraglich erscheint.

Screening-Verfahren, und vor allem Testverfahren, werden von Erzieherinnen in der Regel kritischer betrachtet.³⁰ Auf inhaltlicher Ebene wird moniert, dass die Testaufgaben nicht der Erfahrungswelt der Kinder entsprechen und sie daher etwas anderes messen als die sprachlichen Fähigkeiten der Kinder, die diese in Alltagssituationen zeigen – es wird also die Validität der Verfahren in Frage gestellt. Auch sind Einsatz und Auswertung dieser Art von Verfahren aufwändig und nur schwer in den Kita-Alltag zu integrieren. Baden-Württemberg hat daher ein zweistufiges Verfahren implementiert: Im ersten Schritt wird das relativ einfach durchzuführende und auszuwertende Screening HASE eingesetzt, bei dem die vier Aufgabentypen (a) Nachsprechen von Sätzen, (b) Nachsprechen von Kunstwörtern, (c) Wiederholen von Zahlenfolgen und (d) Erkennen von Wortfamilien unter Nutzung einer Audio-CD zum Einsatz kommen. Dies dauert ca. 10 Minuten. Kinder mit auffälligem Befund werden anschließend von geschulten Personen des Gesundheitsamtes mit dem aufwändigeren Verfahren SETK 3-5 untersucht, das fünf Aufgabentypen umfasst und mit knapp 30 Minuten deutlich mehr Zeit beansprucht. Ein zweistufiges Vorgehen kommt auch in Brandenburg zur Anwendung. Bei der Durchführung von Screenings und Testverfahren werden die Personen, die diese Verfahren einsetzen, teilweise vorher dafür geschult, so z. B. in Baden-Württemberg und Mecklenburg-Vorpommern.

Eine weitere Herausforderung für Sprachkompetenzerfassungen, die bereits erwähnt wurde, ist der Umgang mit Mehrsprachigkeit: Wie sollen Kinder getestet werden, die (a) kürzlich nach Deutschland zugewandert sind und gerade erst Deutsch lernen oder (b) in einer Familie mit Migrationsgeschichte mehrsprachig aufwachsen? In mindestens zwei Bundesländern werden bei einsprachig aufwachsenden Kindern andere Verfahren eingesetzt als bei Kindern, die mehrsprachig bzw. mit Deutsch als Zweitsprache (DaZ) aufwachsen. Dieses kontrovers diskutierte Vorgehen basiert auf der fachwissenschaftlichen Annahme, dass der Erwerbsprozess mehrsprachiger Kinder Besonderheiten aufweist, was relativ unstrittig ist. Strittig ist dagegen die Frage, ob und in welcher Weise diese Besonderheiten zu berücksichtigen sind, um die Validität von Sprachstandsfeststellungen zu gewährleisten. So dürfte die Problemlage bei einem vierjährigen Kind, das seine Muttersprache altersgemäß beherrscht, das aber Defizite hinsichtlich Wortschatz und Grammatik in der deutschen Sprache aufweist, anders gelagert sein als bei einem einsprachig Deutsch aufwachsenden Kind, das die gleichen Defizite aufweist. Wie dem mit vertretbarem Aufwand in der Sprachstandsdiagnostik und in der Sprachförderung in angemessener Weise Rechnung getragen werden könnte, ist jedoch unklar.

29 Z. B. Ulrich, M./Mayr, T., Und wie steht es um die sprachliche Bildung deutscher Kinder? Der Beobachtungsbogen SELDAK, Kindergarten heute12, 2006, S. 22–27.

30 Vgl. z. B. die Stellungnahme der LIGA der Wohlfahrtsverbände Mecklenburg-Vorpommern vom 18.11.2010; erhältlich unter http://www.kita-portal-mv.de/documents/liga_stellungnahme_fruehbi-vo.pdf (26.04.2014).

6 Ab wann besteht ein „Förderbedarf“?

Wenn ein vierjähriges Kind einen aktiven Wortschatz von 800 Wörtern aufweist, ist dies (sehr) gering, ein aktiver Wortschatz von ca. 3.000 Wörtern entspricht hingegen dem Durchschnitt. Ab welchem Wortschatzumfang kann aber gesagt werden, dass dieser „förderbedürftig“ ist? Die Kriterien, die bei der Bestimmung von Sprachförderbedarf herangezogen werden, variieren zwischen den Verfahren, und damit auch zwischen den Ländern, erheblich.

Bei den meisten Verfahren sind die Entscheidungskriterien sozial verankert, d. h. es wird den am schwächsten abschneidenden Kindern ein Förderbedarf attestiert. Dies umfasst zumeist 10% (z. B. Sismik in Bayern 2013), 16% (z. B. HASE) oder 25% (Delfin 4, SETK 3-5) der Kinder. Teilweise, insbesondere bei Beobachtungsbögen, werden die Kriterien ausschließlich normativ formuliert („für das Alter angemessener Wortschatz“). Dies sind Beispiele für bildungspolitische (maximal 10% oder 20% können eine Förderung erhalten) oder subjektive („angemessen“) Setzungen, die inhaltlich nicht fundiert sind. Problematisch ist dabei auch, dass die Stichproben, auf denen die Normwerte basieren, oft nicht repräsentativ und/oder veraltet sind.

Verfahren, bei denen Förderentscheidungen auf einer validierten kriterialen Norm basieren, kommen in der Praxis der Sprachstandsfeststellungen in Deutschland bislang nicht zum Einsatz. Hierzu wäre es erforderlich, klar zu definieren und empirisch zu prüfen, was ein Kind können muss, um am Schuleintritt erfolgreich teilhaben zu können. Dies erfordert weitere Grundlagenforschung wie etwa Untersuchungen zur Frage, welche bildungssprachlichen Anforderungen Kinder in der Grundschule bewältigen müssen³¹ und woran man bei 4- bis 5-jährigen Kindern erkennen kann, ob sie diese Fähigkeiten im Laufe der nächsten 12–24 Monate ohne gezielte Förderung entwickeln werden oder nicht. Um die sozial definierten Kriterien, die bislang verbreitet sind, abzulösen, sind also empirische Hinweise darauf erforderlich, wie ein Kind, das aufgrund sprachlicher Schwächen Probleme in seiner Schullaufbahn aufweist, im Alter von 4–5 Jahren bei welchen Aufgaben wie abgeschnitten hat, um zukünftig förderbedürftige Kinder zuverlässig identifizieren und gezielt fördern zu können.

7 Fazit und Ausblick

Ein erheblicher Anteil der Kinder in Deutschland weist bei Schuleintritt deutliche sprachliche Schwächen auf, so dass sie in ihren schulischen Entwicklungschancen von Beginn an benachteiligt sein dürften. Davon sind insbesondere das Lesen und Schreiben betroffen, die für die weitere Lernentwicklung zentral sind. Aber auch in anderen Fächern, wie etwa Mathematik und Sachunterricht, sind Kinder, die die Unterrichtssprache nur eingeschränkt beherrschen, in ihrer Lernentwicklung beeinträchtigt. Auf diese Herausforderung haben die Länder in Deutschland sehr unterschiedlich reagiert, was u. a. damit zusammenhängt, dass zum Zeitpunkt der Entscheidung, Sprachförderung in Kitas zu intensivieren, kaum Verfahren zur Verfügung standen, die zur Erhebung des Sprach-

31 *Heppt, B./Dragon, N./Berendes, K./Stanat, P./Weinert, S.*, Beherrschung von Bildungssprache bei Kindern im Grundschulalter, *Diskurs Kindheits- und Jugendforschung* 3, 2012, S. 349–356, *Haag, N./Heppt, B./Stanat, P./Kuhl, P./Pant, H. A.*, Second language learners' performance in mathematics: Disentangling the effects of academic language features, *Learning and Instruction* 28, 2013, S. 24–34.

stands unmittelbar hätten eingesetzt werden können.³² Bei den vorhandenen Verfahren lagen die Rechte zumeist bei Verlagen, die für den Einsatz erhebliche Nutzungsgebühren verlangen, so dass insbesondere in großen Bundesländern ein flächendeckender und regelmäßiger Einsatz der Instrumente kaum bezahlbar wäre. Insofern bestand nur die Möglichkeit, neue Verfahren zu entwickeln oder vorhandene anzupassen. Obwohl die Herausforderungen in vielen Regionen sehr ähnlich sind, haben die Länder darauf nicht gemeinsam reagiert, sondern jeweils eigene Instrumente entwickelt. Dies hat zur Folge, dass heute sehr viele verschiedene Verfahren eingesetzt werden, die jeweils zu unterschiedlichen Förderquoten kommen. Zudem ist die Qualität vieler der eingesetzten Verfahren unklar. Bei den 21 derzeit eingesetzten Verfahren liegen für zehn Instrumente keine Angaben zur Objektivität vor, für fünf Verfahren ist die Reliabilität unbekannt und bei neun Verfahren fehlen empirische Belege für die angenommene Faktorenstruktur (als Indikator für die Konstruktvalidität).³³ Die besonders wichtige prädiktive Validität, die eine Überprüfung des Zusammenhangs zwischen Messwerten, die von Kindern im jeweiligen Sprachstandsverfahren erreicht wurden, mit später erhobenen Indikatoren für den schulischen Erfolg dieser Kinder erfordert, wurde bislang nur für eines der eingesetzten Verfahren gründlicher überprüft.

Vor diesem Hintergrund erscheint es sinnvoll darüber nachzudenken, ob es nicht zweckmäßig wäre, ausgewählte Verfahren mit unterschiedlichen Zielsetzungen (Screening-Verfahren, differenzierte Beobachtungsbögen, genauere Erfassung des Sprachstands bei mehrsprachigen Kindern, Tests für wissenschaftliche Untersuchungen, Verfahren für die Ableitung von spezifischen Förderempfehlungen etc.) arbeitsteilig weiter zu entwickeln. Die gemeinsame Entwicklung würde nicht nur Ressourcen schonen, sondern auch die Qualität der Verfahren verbessern und zudem zu vergleichbaren Förderempfehlungen in den Ländern führen. Flankierend dazu ist gezielte Forschung erforderlich, die untersucht, welche Indikatoren und Kriterien für die Bestimmung von Förderbedarf am besten geeignet sind, welche Art der Förderung effektiv ist und wie dabei die zunehmende Mehrsprachigkeit von Kindern zu berücksichtigen ist. Viele dieser Fragen lassen sich nur in soliden Längsschnittstudien beantworten, von denen bislang nur sehr wenige durchgeführt worden sind.

Verf.: Dr. Uwe Neugebauer/Prof. Dr. Michael Becker-Mrotzek, Universität zu Köln, Mercator-Institut für Sprachförderung und Deutsch als Zweitsprache, Albertus-Magnus-Platz, 50923 Köln, E-Mail: Uwe.Neugebauer@uni-koeln.de/becker.mrotzek@uni-koeln.de.

Prof. Dr. Petra Stanat, Humboldt-Universität zu Berlin, Institut zur Qualitätsentwicklung im Bildungswesen (IQB), Unter den Linden 6, 10099 Berlin, E-Mail: petra.stanat@iqb.hu-berlin.de.

32 Fried, L., Expertise zu Sprachstandserhebungen für Kindergartenkinder und Schulanfänger. Eine kritische Betrachtung, Expertise im Auftrag des DJI, Deutsches Jugendinstitut e. V., 2004; verfügbar unter: http://dji.de/bibs/271_2232_ExpertiseFried.pdf (26.04.2014).

33 Vgl. Neugebauer, U./Becker-Mrotzek, M. (Anm. 8).