

Zur Problematik der Evaluation von Sprachfördermaßnahmen

1. Einleitung

Sprachförderung wird seit der Jahrtausendwende von Wissenschaft, Politik und Öffentlichkeit als Medium der Integration von Kindern mit Migrationshintergrund erkannt. In dieser Zeit ist eine Vielzahl an Fördermaterialien und –konzepten entstanden, die unter den verschiedensten Rahmenbedingungen eingesetzt werden. Die Beurteilung der Wirksamkeit der unterschiedlichen Maßnahmen ist von großer gesellschaftspolitischer Relevanz, da sie mit Kosten verbunden sind und den Kindern möglicherweise keine optimalen Bedingungen für ihre sprachliche und, damit verbunden, ihre schulische Entwicklung bieten.

Die Evaluation von Sprachfördermaßnahmen stellt die Wissenschaft dabei vor erhebliche, schwer lösbare Probleme. Das hat zum einen mit der Menge an internen und externen Faktoren zu tun, die den Zweitspracherwerb in der Kindheit beeinflussen und es erschweren, in allen relevanten Aspekten vergleichbare Untersuchungs- und Kontrollgruppen zu bilden. Zum anderen stellen die eingesetzten Messinstrumente einen Problembereich bei der Beurteilung von Maßnahmen dar. Die Ergebnisse vorliegender Evaluationsstudien sind vor diesem Hintergrund als äußerst fragwürdig zu betrachten.

Im Folgenden wird eine Studie kritisch reflektiert, die Sprachfördermaßnahmen im Rhein-Neckar-Raum betrifft: die von J. Roos und H. Schöler im Auftrag der Landesstiftung Baden-Württemberg durchgeführte EVAS-Studie (Polotzek; Roos; Schöler 2009), bei der die Ergebnisse von Sprachförderung nach drei Konzepten, darunter einer Vorläuferversion von „Deutsch für den Schulstart“¹, mit denen einer „unspezifischen“ Förderung im Kindergartenalltag verglichen wurden. In dieser Studie wurden, ebenso wie in einer zeitgleich durchgeführten Studie der PH Weingarten zu anderen Fördermaßnahmen, keine positiven Effekte der gezielten Sprachförderung gegenüber einer „situativen“, in den Alltag integrierten sprachlichen Förderung im Kindergarten festgestellt. Es stellt sich die Frage, ob alle oder einige der einbezogenen, gezielten Fördermaßnahmen tatsächlich keinen „Mehrwert“ erbringen, oder ob dieser in den Evaluationsstudien nur nicht erfasst wird. Diese

¹ Das Förderkonzept wurde im Rahmen eines Projekts am Seminar für Deutsch als Fremdsprachenphilologie der Universität Heidelberg unter Leitung der Verfasserin dieses Artikels entwickelt. Wir danken der Günter-Reimann-Dubbers-Stiftung (Heidelberg) und der Dürr-Stiftung (Hamburg) für die finanzielle Unterstützung, die die Projektarbeit möglich macht. Joanna Ermonies-Jargiello, Gunde Kurtz, Giulio Pagonis und Tanja Vasylyeva danke ich für Unterstützung und Anregungen beim Verfassen dieses Artikels.

Frage wird im Folgenden insbesondere für das Förderkonzept „Deutsch für den Schulstart“ in Bezugnahme auf die EVAS-Studie, aber auch auf eine eigene, projektinterne Überprüfung von Förderergebnissen reflektiert. Im Zentrum stehen dabei die eingesetzten Messinstrumente. Für diese erscheinen die folgenden Anforderungen wesentlich:

- Sie sollten die zentralen Förderbereiche berücksichtigen, d.h. auf die Inhalte der Förderung bezogen sein.
- Sie sollten den Stand und die Entwicklung der Kinder in diesen Bereichen angemessen erfassen können, d.h. entwicklungssensitiv sein.

Für die Bewertung der Ergebnisse sind Richtwerte erforderlich, welche Entwicklungsschritte bei den für die jeweiligen Kinder gegebenen Erwerbs- und Förderbedingungen in welchem Zeitrahmen möglich sind. Diese liegen beim derzeitigen Forschungsstand nicht vor, so dass die Interpretation der Daten und damit die Bewertung der Maßnahmen problematisch ist.

2. Das Förderkonzept „Deutsch für den Schulstart“

Das Sprachförderkonzept „Deutsch für den Schulstart“ hat die Zielsetzung, insbesondere mehrsprachige Kinder auf die sprachlichen Anforderungen der Schule vorzubereiten (Kaltenbacher; Klages 2007; 2008). Es umfasst Fördermaterialien, die in vier aufeinander aufbauende Phasen eingeteilt sind, einen Förderzeitraum von zwei bis drei Jahren abdecken und im Altersbereich von vier bis acht Jahren, d.h. in den letzten beiden Kita- und den ersten beiden Grundschuljahren, einsetzbar sind.² Im Sinne einer diagnosebasierten Förderung wurden Sprachstandserhebungsverfahren entwickelt, mit denen der Einstieg in die Förderung festgelegt und der Fortschritt der Kinder dokumentiert werden kann. Die Förderkräfte werden im Rahmen von zweitägigen Kurzfortbildungen in die Arbeit nach dem Förderkonzept eingeführt.³

Die Förderung dient einerseits der Unterstützung der Kinder beim Aufbau grundlegender alltagssprachlicher Fähigkeiten, insbesondere aber auch der Hinführung zur schulischen Bildungssprache. Sie umfasst die folgenden Bereiche:

- Wortschatz (verschiedene Lebensbereiche der Kinder; Wortbildungsmuster; relationale Ausdrücke; etc.);
- Grammatik (Syntax des einfachen Satzes; einfache Nebensatztypen; nominale und verbale Formenbildung);

² Derzeit arbeitet die Projektgruppe an einer Adaptation des ursprünglich für den Vorschulbereich entwickelten Materials an die Bedingungen einer Grundschulförderung (1. und 2. Klasse).

³ Zum Förderkonzept gehören außerdem eine einjährige E-Learning-Fortbildung, die der vertieften Auseinandersetzung mit dem Förderkonzept und der Ausbildung von Multiplikatorinnen dient, sowie Materialien und Anregungen für die Elternarbeit.

- Text (Geschichtenverständnis; Textsorten Erzählen, Beschreiben und Anweisungen);
- Mathematische Vorläuferfertigkeiten (Kategorisierung; Zählen und Mengenerfassung);
- Phonologische Bewusstheit (Verse & Reime; Betonungsmuster von Wörtern; Anlaute).

Die EVAS-Studie bezieht sich auf die Vorläuferversion des Jahres 2005/06 von „Deutsch für den Schulstart“, in der bereits alle genannten Förderbereiche enthalten waren.⁴ Die Komponenten des Konzepts wurden jedoch seit Projektbeginn 2004 im Rahmen einer internen Begleitforschung und unter Einbezug der Praxiserfahrungen vieler mit uns kooperierender Förderkräfte kontinuierlich erprobt und überarbeitet (Kaltenbacher; Klages; Pagonis 2009). So sind die Materialien gegenüber der evaluierten Version inzwischen deutlich erweitert und bezüglich verschiedener Aspekte, u.a. der didaktischen Ausgestaltung, modifiziert.⁵

3. Die EVAS-Studie

Die an der PH Heidelberg unter der Leitung von J. Roos und H. Schöler durchgeführte „Evaluationsstudie zur Sprachförderung von Vorschulkindern“ (EVAS) betrifft das Förderjahr 2005/06 und erfasst sprachliche, kognitive und ausgewählte schulische Leistungen der Kinder im Rahmen eines Prätest/Posttest- und Follow-up-Designs. Die Datenerhebung erfolgte vor Beginn der Förderung im letzten Kita-Jahr der Kinder, am Ende des Förderjahres sowie nach dem 1. und 2. Schuljahr (Polotzek et al. 2009: 8f). Verglichen werden über alle vier Zeitpunkte hinweg:

- drei Untersuchungsgruppen, die eine „spezifische“ Förderung erhalten haben, und zwar nach der Konzeption von Zvi Penner (89 Kinder), Rosemarie Tracy (39 Kinder) und Kaltenbacher & Klages, der Vorläuferversion von „Deutsch für den Schulstart“ (58 Kinder); die Kinder nahmen dabei durchschnittlich an 88 Stunden Förderung in Gruppen von sechs bis acht Kindern teil;
- eine Vergleichsgruppe von Kindern mit Sprachförderbedarf, die nur im üblichen, „unspezifischen“ Rahmen innerhalb der Kita bzw. des Kindergartens sprachlich gefördert wurde (95 Kinder);
- eine Vergleichsgruppe von Kindern ohne Förderbedarf (199 überwiegend einsprachige Kinder).

⁴ Der Einfachheit halber beziehen wir uns auch auf diese Vorläuferversion mit der Bezeichnung „Deutsch für den Schulstart“.

⁵ Die Didaktik des Förderkonzepts folgt dem Ziel, die Kinder im Rahmen einer systematischen Förderung in ihrem natürlichen Erwerbsverlauf zu unterstützen, indem jeweils eine gezielte Förderung im Bereich des nächsten Entwicklungsschritts erfolgt. Dabei wird ein impliziter, beiläufiger Erwerb durch die Kinder angestrebt.

Die Untersuchungsgruppen waren beim 2. Follow-up um 38 % reduziert, so dass keine getrennten Berechnungen für die drei verschiedenen „gezielten“ Förderkonzeptionen mehr durchgeführt werden konnten.

Als Messinstrumente für den Entwicklungsstand in der gesprochenen Sprache wurden vier Subtests des Heidelberger Sprachentwicklungstests (HSET), der in den 70er Jahren zur Diagnose von Sprachstörungen bei einsprachigen Kindern entwickelt wurde, eingesetzt (Grimm; Schöler 1978). Erfasst wurden Fähigkeiten im Wortschatz (Subtest Wortfindung), in der Flexionsmorphologie (Subtest Plural-Singular-Bildung) und in der Syntax (Subtests Verstehen von Strukturformen und Imitation von Strukturformen). Schriftsprachliche Fähigkeiten wurden beim Follow-up mit – ebenfalls für einsprachige Kinder entwickelten – Testverfahren für die Lese- und Rechtschreibfähigkeit erfasst.

Die Verfasser von EVAS halten als Ergebnis fest, dass sich am Ende des Evaluationszeitraums, d.h. am Ende des 2. Schuljahres, „die Gruppe der nicht in spezielle Sprachfördermaßnahmen involvierten Kinder mit Förderbedarf in allen untersuchten Aspekten nicht von der mit viel Aufwand zusätzlich geförderten Gruppe unterscheidet“ (Polotzek et al. 2009: 68). Auch eine Annäherung an die Vergleichsgruppe der Kinder ohne Förderbedarf konnte nicht festgestellt werden.

Die Autoren weisen darauf hin, „dass nicht alle vorstellbaren und bedeutsamen Einflussgrößen (...) in die vorliegenden Analysen einfließen oder kontrolliert und variiert werden“ konnten, die „das Ausbleiben eines substanziellen Fördereffekts“ mit verursacht haben könnten (S. 69). Diese Aussage gesteht zwar Einschränkungen ein, lässt diese jedoch als für die Schlussfolgerung nicht wesentlich erscheinen. Tatsächlich weist die Evaluationsstudie jedoch methodische und inhaltliche Probleme auf, die ihre Aussagekraft ganz erheblich einschränken.

Kritikpunkte an der EVAS-Studie lassen sich in ganz unterschiedlichen Bereichen finden, was in Anbetracht ihrer Komplexität nicht überrascht. Neben grundsätzlichen Mängeln in der Methodik, die u.a. die fehlende psycholinguistische Hypothesenbildung und, damit verbunden, die Eignung der Testverfahren betrifft (s.u.), liegt eine ganze Reihe weiterer Probleme vor. So ist die Vergleichbarkeit der Kinder der Untersuchungsgruppen mit denen der Vergleichsgruppe mit Förderbedarf nicht gewährleistet, da für die Letztere bessere individuelle bzw. soziale Rahmenbedingungen angenommen werden können (Kaltenbacher; von Stutterheim 2009). Auch wurde im Bereich der kognitiven Fähigkeiten zwar die – nonverbale – Intelligenz, nicht jedoch die für den Spracherwerb grundlegende Aufmerksamkeitsspanne erfasst. Bezüglich

der Datenerhebung wurde von Seiten der beteiligten Kindertagsstätten die Eignung der eingesetzten Testleiter/innen für diese Aufgabe in Frage gestellt.

Die Methodik der EVAS-Studie wurde von einem Experten (Rogge 2009) kritisch analysiert. Hauptkritikpunkte betreffen:

- die fehlende theoretische (sprachwissenschaftliche und psycholinguistische) Basis, auf deren Grundlage klare Hypothesen zu erwartbaren Lernfortschritten hätten formuliert werden können. Bei diesem methodischen Defizit sind die Ergebnisse der Studie nicht zu interpretieren;

- die fragliche Eignung der eingesetzten statistischen Auswertungsverfahren. Sie erlauben es nicht, die Entwicklung von einzelnen Kindern und differenzierten Wirkungszusammenhängen in den Blick zu nehmen;

- unzulässige Verallgemeinerungen. So schränkt die Anlage der Untersuchung (das quasi-experimentelle Design) die Verallgemeinerbarkeit der Ergebnisse stark ein. Die fehlende Thematisierung dieser Einschränkung in den Berichten stellt eine Verletzung der wissenschaftlichen Sorgfaltspflicht dar.

Insgesamt lässt es die Studie demnach nicht zu, Schlussfolgerungen für künftige Fördermaßnahmen abzuleiten, da eine differenzierte Erfassung und Bewertung von Erwerbsverläufen unterbleibt.

Im Folgenden soll die Eignung der eingesetzten Testverfahren zur Überprüfung des Fördererfolgs genauer diskutiert werden. Diese Frage stellt einen Kernpunkt bei der Kritik der EVAS-Studie aus der Sicht der Spracherwerbsforschung dar.⁶

Bezüglich der Auswahl der erfassten Sprachbereiche ist EVAS dadurch gekennzeichnet, dass die Textkompetenz der Kinder, d.h. ihre Fähigkeit beim Verstehen und Gebrauch größerer sprachlicher Einheiten (wie Geschichtenverständnis und Erzählen) nicht berücksichtigt wird. Damit wird ein wesentlicher Inhalt der Sprachförderung nach „Deutsch für den Schulstart“ vollständig ausgeklammert. Für die Überprüfung des Wortschatzes wurde neben einer einfachen Benennungsaufgabe, die aufgrund des Deckeneffekts keine interpretierbaren Daten erbrachte, der Subtest „Wortfindung“ des HSET eingesetzt. Ebenso wie für den überprüften grammatischen (morphologischen und syntaktischen) Bereich wird damit ein Instrument verwendet, das nicht zum Zweck der Evaluation, sondern zu dem der Differentialdiagnose von Sprachstörungen entwickelt wurde. Ein solches Vorgehen wird von Wottawa & Thierau (1990: 125f) in ihrem Lehrbuch zur Evaluation kritisch reflektiert. Die Verfasser stellen fest, dass sich Testverfahren, die für angewandt-diagnostische Zwecke

⁶ Für andere Evaluationen, etwa die ebenfalls von der Landesstiftung Baden-Württemberg in Auftrag gegebene der PH Weingarten, trifft das ebenso zu.

entwickelt wurden, für Evaluationen wenig eignen. So erfordert die diagnostische Zielsetzung eine geringe Änderungssensitivität, während bei Evaluationen gerade - auch kleinere - Veränderungen von Interesse sind. Die Autoren stellen die verbreitete Bevorzugung von standardisierten Verfahren in Frage (S. 126):

„Das Bestreben vieler Evaluatoren, objektive Testverfahren einzusetzen, ist verständlich. Man muss aber auch die Grenzen der Möglichkeiten dieser Instrumente zumindest bei ihrem derzeitigen Entwicklungsstand sehen, und es kann sinnvoller sein, mit einer „weichen“, weniger exakten und elaborierten Methode ein interessantes Kriterium zu messen als mit hoher Präzision etwas, was inhaltlich nicht zu den eigentlichen Evaluationszielen passt.“ Diese Überlegungen treffen für die bei EVAS eingesetzten Testverfahren in hohem Maße zu, wie im Folgenden am Beispiel der beiden Subtests „Verstehen von Strukturformen“ (VS) und „Imitation von Strukturformen“ (IS), mit denen die syntaktische Kompetenz der Kinder überprüft wurde, gezeigt werden soll. Die beiden Subtests bestehen aus Testitems, mit denen späte Entwicklungen in der Syntax erfasst werden können – solche, die sich im Erstspracherwerb überwiegend erst im Übergang zum Schulalter herausbilden. Im Einzelnen handelt es sich dabei um Passivsätze (VS, IS), Relativsätze (VS, IS), Temporalsätze (VS, IS), Expletivstrukturen (IS) und performative Sätze (VS), die den Kindern kontextlos vorgegeben werden und mit Spielfiguren ausagiert (Verstehenstest) bzw. nachgesprochen (Imitationstest) werden sollen. Die damit verbundenen sprachlichen Anforderungen werden zunächst an einigen Beispielen für den Verstehens-, danach für den Imitationstest erläutert.

Für den Verstehenstest sollen dazu die Testsätze mit Passivstrukturen und mit temporalen Nebensätzen herangezogen werden. Passivstrukturen sind in den Testitems Nr. 2, 4 und 8 enthalten:

VS Testitem 2: *Die Katze wird von dem Jungen gefangen* (irreversibel)

VS Testitem 4: *Das Mädchen wird von dem Jungen gewaschen* (reversibel)

VS Testitem 8: *Die Mutter wird von dem kleinen Kind gewaschen* (kontra-intuitiv)

Die Testsätze enthalten einen zunehmenden Schwierigkeitsgrad: Der irreversible, d.h. nur in einer Handlungsrichtung (mit dem Jungen als Agens) plausible Passivsatz 2 kann ohne Beachtung der Passivmorphologie aufgrund des Weltwissens, d.h. auf der Basis einer Semantikstrategie, korrekt ausagiert werden. Die von Bever für den Erstspracherwerb postulierte N-V-N-Strategie (vgl. Grimm; Schöler; Wintermantel 1975: 62f), nach der das erste Nomen im Satz als Agens interpretiert wird, muss dabei allerdings außer Kraft gesetzt werden. Bei dem reversiblen Passivsatz 4, bei dem beide Handlungsrichtungen („Mädchen wäscht Junge“ und „Junge wäscht Mädchen“) gleich plausibel sind, hilft weder eine

Semantikstrategie noch die N-V-N-Strategie weiter. Hier müssen die strukturellen Merkmale des Passivs genutzt werden, um zu einem korrekten Verständnis zu kommen. Bei Testsatz 8 müssen diese so gut beherrscht werden, dass sie den kontra-intuitiven Inhalt des Satzes „aushebeln“.

Eine entsprechende Schwierigkeitsabfolge ergibt sich für die Testitems mit temporalen Nebensätzen:

VS Testitem 6: *Die Ente wackelt weg, bevor das Schaf umfällt*

VS Testitem 7: *Nachdem das Pferd rennt, springt der Elefant*

VS Testitem 12: *Bevor der Hund rennt, springt das Pferd*

Satz 6 kann auf der Basis einer „order of mention“ – Strategie (Grimm et al. 1975: 134f) korrekt verstanden werden: Die Nennung der Ereignisse im Satz entspricht hier der chronologischen Abfolge. Dasselbe trifft auf Satz 7 zu, bei dem allerdings die den Literaturbefunden zufolge schwierigere Konjunktion *nachdem*, d.h. die Nachzeitigkeit, betroffen ist. Bei Satz 12 ist die „order of mention“ – Strategie irreführend, hier ist ein präzises Verständnis der Konjunktion erforderlich.

Für alle Testsätze gilt, dass aufgrund ihrer isolierten Darbietung keine kontextbezogenen Verstehensstrategien eingesetzt werden können. In Anbetracht der hohen Anforderungen an die grammatische Kompetenz erstaunt es daher nicht, dass die Untersuchungsgruppen im Posttest durchschnittlich nur sechs Aufgaben korrekt ausagierten und damit deutlich hinter der Vergleichsgruppe ohne Förderbedarf zurückblieben, die auf durchschnittlich 11 korrekte Lösungen kam. Über den Fördererfolg sagt dieses Ergebnis jedoch wenig aus.

Zur Illustration der sprachlichen Anforderungen beim Imitationstest sollen zwei – für den Subtest durchaus typische - Testsätze herangezogen werden:

IS Testitem 1: *Der Teppich wird von dem Vater ausgeklopft*

IS Testitem 5: *Es sitzt der kleine Vogel im Gebüsch*

Die Imitation von Sätzen dieser Länge setzt eine Restrukturierung und damit sprachliches Wissen auf verschiedenen Ebenen voraus. Die Testsätze beinhalten dabei für die Förderkinder problematische Wörter (*ausklopfen, Gebüsch*) und Sachverhalte („Teppich klopfen“). Ihre präzise Wiedergabe setzt neben der Wortschatzkenntnis differenziertes Wissen im formal-grammatischen Bereich voraus: in Bezug auf die Artikelwahl in der Genus- und Kasusmorphologie, im verbalen Bereich in der Passivmorphologie, beim expletiven „es“ die Beachtung eines rein formalen, inhaltsleeren Elements. Forschungsergebnisse zum frühen Zweitspracherwerb zeigen, dass entsprechende Bereiche in einem langwierigen Prozess erworben werden (Kaltenbacher; Klages 2007; 2008; Montanari 2010; Wegener 1995; 1998).

Da für die Bewertung mit der vollen Punktzahl (zwei Punkte je Testitem) eine weitestgehend präzise Wiedergabe erforderlich ist und nur einige eher umgangssprachliche Abweichungen mit einem Punkt honoriert werden, erstaunt das schlechte Ergebnis der Förderkinder nicht: Sie erreichen im Posttest durchschnittlich sechs, die Kinder ohne Förderbedarf dagegen 16 Punkte. Die exemplarische Analyse verdeutlicht die Komplexität der sprachlichen Aufgaben, die die Kinder bei den beiden Subtests VS und IS bewältigen müssen. Die überprüften grammatischen Strukturen sind dabei nicht Gegenstand der gezielten Sprachförderung, wie sie nach „Deutsch für den Schulstart“ im Vorschulalter angeboten wird. Sie können für diese Altersgruppe auch kein sinnvolles Förderziel sein. Wie die Ausführungen in Abschnitt 4 zeigen, geht es bei diesen Kindern um den Aufbau grundlegender Strukturen, auf deren Basis komplexe Sätze erst verstanden und korrekt produziert werden können. Die HSET-Subtests gehen damit an den Zielen der Sprachförderung vorbei und setzen eine unangemessen hohe Messlatte an.

Die Verfahren sind auch ungeeignet dafür, tatsächliche Entwicklungen in der Sprachkompetenz der Kinder abzubilden, d.h. sie sind nicht entwicklungs sensitiv. So orientieren sich die Subtests VS und IS an den Verarbeitungsstrategien und Entwicklungssequenzen, die für den Erstspracherwerb ermittelt wurden (Grimm; Schöler; Wintermantel 1975). Inwiefern diese auch für den Zweitspracherwerb in der Kindheit Gültigkeit haben, ist kaum untersucht. Forschungsergebnisse der letzten Jahre machen jedoch deutlich, dass sich bereits ab einem Einstiegsalter⁷ von ca. drei Jahren Unterschiede im grammatischen Bereich gegenüber dem Erstspracherwerb (und Übereinstimmungen mit dem Zweitspracherwerb bei Erwachsenen) zeigen, die sich mit zunehmendem Alter bei Kontaktbeginn verstärken (u.a. Habertzettl 2005; Kaltenbacher; Klages 2007; Meisel 2007). Generell ist davon auszugehen, dass der Zweitspracherwerb durch eine stärkere Konzentration auf die inhaltliche Seite der Sprache, also auf die Semantik und Pragmatik, gekennzeichnet ist, die sprachliche Form von Ausdrücken dagegen weniger Beachtung findet. Das bedeutet für das Sprachverstehen von Kindern im Zweitspracherwerb, dass kontextbezogene Strategien vermutlich einen höheren Stellenwert haben als beim Erstspracherwerb gleichaltriger Kinder. Solche Strategien werden bei VS nicht berücksichtigt, durch den intensiven Einbezug von Geschichten in die Förderung nach „Deutsch für den Schulstart“ sind hier jedoch Fördereffekte anzunehmen. Die Verständnisleistung der überprüften Kinder wird damit nicht adäquat erfasst.

⁷ Damit ist der Zeitpunkt gemeint, ab dem das Kind in aktiven Kontakt mit der Zweitsprache kommt.

Bei IS ist fraglich, ob sich die zweisprachigen Kinder im selben Maße wie Kinder mit der Erstsprache Deutsch am Wortlaut der Sätze orientieren, oder sich verstärkt auf die Wiedergabe des Inhalts der Sätze konzentrieren, was sich negativ auf die Bewertung auswirkt. Entwicklungsspezifische „Fehler“, d.h. solche Abweichungen von der Sprachnorm, die charakteristische Zwischenschritte des frühen Zweitspracherwerbs darstellen, werden bei der Auswertung – die ja auch für eine andere Zielgruppe und Zielsetzung entwickelt wurde – ebenfalls nicht berücksichtigt. Bei einem entwicklungs sensitiven Verfahren sollten sie bei der Auswertung positiv bewertet werden.⁸

Für die – ebenfalls für einsprachige Kinder entwickelten – Verfahren zur Erfassung der Rechtschreibleistung und des Leseverstehens gilt Ähnliches wie für die HSET-Untertests. Auch hier zeigt sich das gravierende linguistische Defizit der EVAS-Studie: Die Tests setzen Sprachleistungen der mehrsprachigen Kinder voraus, die bei den gegebenen Erwerbs- und Förderbedingungen unrealistisch sind. Prüft man eine Fähigkeit zu zwei Zeitpunkten, die zum zweiten Zeitpunkt nicht erreicht sein kann, so lassen sich auf Grund der Testergebnisse keine Schlussfolgerungen zu Lernfortschritten ziehen.

4. Ergebnisse der internen Begleitforschung

Seit Beginn des Projekts „Deutsch für den Schulstart“ im Jahr 2004 ist die Überprüfung des Lernerfolgs der Kinder wesentlicher Teil der Projektarbeit. Im Rahmen einer Begleitforschung wurden Verfahren der Sprachstandserhebung entwickelt, mit denen verschiedene Bereiche der grammatischen Kompetenz, aber auch die Erzählfähigkeit und das Geschichtenverständnis überprüft werden können. Diese Verfahren sind eng auf die Förderinhalte bezogen und werden vor Beginn der Förderung zur Festlegung des Fördereinstiegs und zur Bildung der Fördergruppen, während der Förderung gegebenenfalls zur Modifikation des Vorgehens und an ihrem Ende zur Erfassung der Lernfortschritte eingesetzt. Sie können als Teil einer „formativen Evaluation“ (Wottawa; Thierauf 1990: 30) betrachtet werden, bei der die Überprüfung der Fördereffekte zu einer Verbesserung der Förderkonzeption genutzt wird, sind allerdings nicht standardisiert. Neben den verschiedenen Testverfahren wurden Rückmeldungen von Förderkräften zur Umsetzbarkeit und Wirksamkeit der Förderspiele in diesen Evaluierungs- und Optimierungsprozess einbezogen.⁹

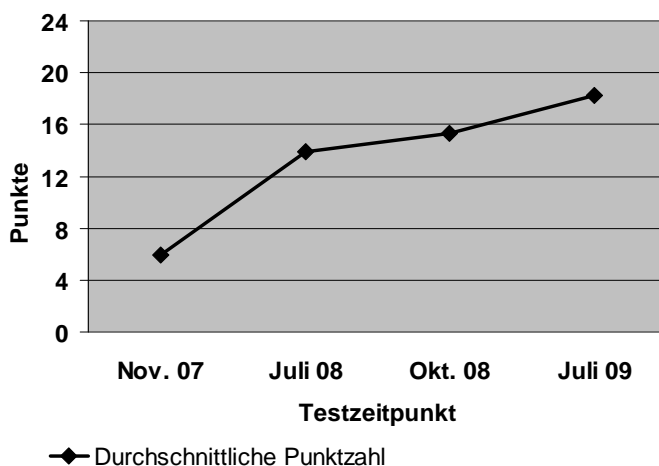
⁸ Das betrifft u.a. die Genusform der Artikel, bei der die Kinder im Zweitspracherwerb zunächst ein zweigliedriges System aus maskulinen und femininen Formen aufbauen und Neutra mit der maskulinen Form verwenden, aber auch Entwicklungsstufen beim Passiverwerb, wo „ist“ als Auxiliar übergeneralisiert wird.

⁹ Wir danken an dieser Stelle den zahlreichen Personen, die auf diese Weise durch ihr Engagement zur stetigen Verbesserung der Förderkonzeption beigetragen haben.

Im Folgenden werden Förderergebnisse, die im Rahmen der internen Begleitforschung ermittelt wurden, kurz vorgestellt. Sie betreffen eine Gruppe von 51 Kindern, die in Heidelberger Tagesstätten über zwei Jahre hinweg (2007/08 und 2008/09), in ihrem vorletzten und letzten Kita-Jahr, mit dem Förderkonzept „Deutsch für den Schulstart“ gefördert wurden.¹⁰ Aus Platzgründen beschränken wir uns auf die Entwicklung im grammatischen Bereich. Eine vollständigere Analyse ist in Kaltenbacher; Klages; Vasylyeva (im Druck) enthalten.

Die hier vorgestellten Ergebnisse beziehen sich auf einen Einzeltest, bei dem den Kindern nacheinander 15 Bildkarten vorgelegt werden mit der Bitte, die abgebildeten Sachverhalte zu beschreiben.¹¹ Die mit diesem Verfahren erhobenen Aussagen der Kinder werden in einem ersten Auswertungsschritt in Bezug auf Syntax und Wortschatz, in weiteren Schritten in Bezug auf die Genus- und Kasuskompetenz analysiert. Der erste Auswertungsschritt erfasst die Fähigkeit des Kindes, den abgebildeten Sachverhalt in Bezug auf Wortschatz und Satzstruktur angemessen sprachlich wiederzugeben. Bei der Satzstruktur werden dabei die syntaktische Vollständigkeit sowie verschiedene Stellungsmuster des Verbs (S-V-O; S-Aux-O-V; S-V-O-Part¹²) erfasst. Abbildung 1 gibt die Entwicklung der Kinder über die beiden Förderjahre hinweg wieder.¹³

Abb. 1: Entwicklung der Kinder bei Syntax und Wortschatz in zwei Förderjahren



¹⁰ In den beiden Jahren zusammen nahmen die Kinder durchschnittlich ca. 180 Stunden an der Sprachförderung teil.

¹¹ Die Abbildungen haben teilweise irrealen Charakter, was nicht nur den Unterhaltungswert steigert, sondern auch die Überprüfung der produktiven Kompetenz der Kinder begünstigt. Gleichzeitig können so Sätze besser elizitiert werden, bei denen das Subjekt nicht auf menschliche Wesen, sondern auf unbelebte Handlungsträger Bezug nimmt, was für die Ermittlung der Genuskompetenz von Interesse ist.

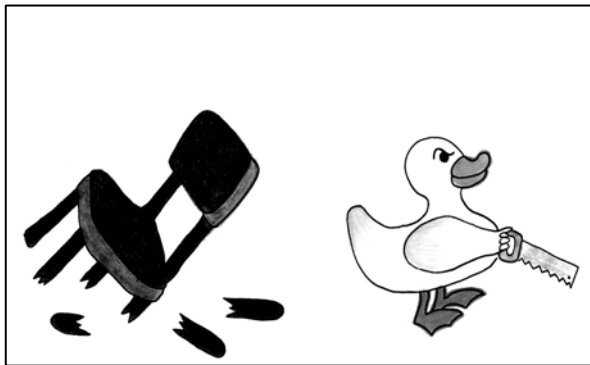
¹² Die Abkürzungen stehen für: S = Subjekt; V = Verb; O = Objekt; Aux = Auxiliar; Part = trennbare Verbpartikel.

¹³ Zwischen der Endtestung des 1. Jahres (Juli 2008) und der Anfangstestung des 2. Jahres (November 2008) fand keine Förderung statt.

Abbildung 1 zeigt einen Anstieg der Leistungen von durchschnittlich sechs auf 18 von 24 möglichen Punkten für die Gesamtgruppe der Kinder.

Die Abbildungen 2 und 3 geben am Beispiel von zwei Bildkarten typische Äußerungen der Kinder zu Beginn, in der Mitte und am Ende des Förderzeitraums wieder.

Abb. 2
 „Die Ente hat den Stuhl kaputt gemacht“



Anfang 1. Förderjahr
 Die Vogel schneid was
 Die hat kaputt gemacht
 Hier ist alle kaputt

Ende 1. Förderjahr
 Der Ente hat Stuhl geschneidet
 Der macht den Stuhl kaputt
 Ein Vögel hat ein Stuhl gesägt

Ende 2. Förderjahr
 Die Ente hat den Stuhl mit Säge geschnitten
 Ein Entchen hat den Stuhl gesägt
 Die Ente hat den Stuhl kaputt gemacht,
 weil sie böse ist

Abb. 3
 „Das Haus gießt den Baum“



Anfang 1. Förderjahr
 Die Haus, das ist eine Wasser
 Der macht Wasser Blume
 Die pflückt Baum

Ende 1. Förderjahr
 Die Haus mag den Baum pflanzen
 Die Haus macht die Baum nass
 Die duscht den Baum

Ende 2. Förderjahr
 Der Haus gießt den Baum
 Ein Haus pflanzt den Baum
 Der Haus gibt den Baum Wasser
 Das Haus gießt den Baum

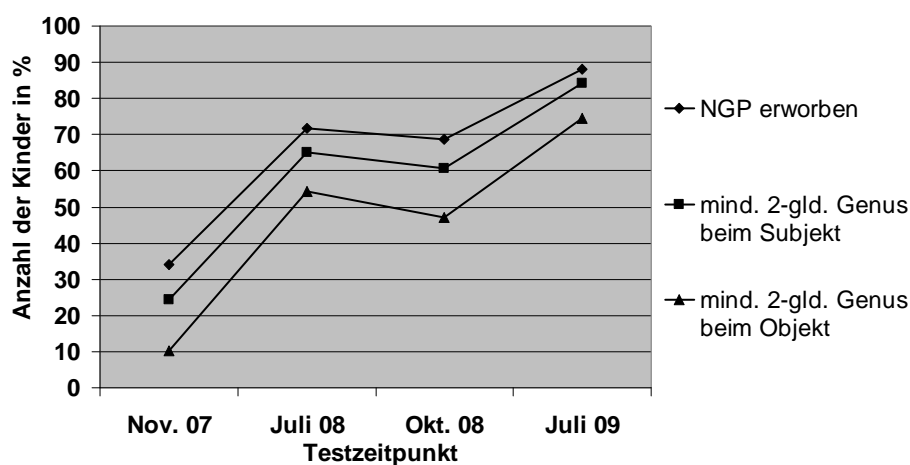
Die Äußerungen verdeutlichen den schrittweisen Aufbau der Fähigkeit, konkrete Sachverhalte syntaktisch weitgehend korrekt und mit adäquatem Wortschatz auszudrücken. Dabei sind Überdehnungen und Neuschöpfungen von Wörtern (z.B. *duschen*, *pflanzen* und *pflücken* für *gießen*), der Gebrauch von Wörtern mit sehr allgemeiner Bedeutung (*machen*) sowie der Gebrauch von Pronomen anstelle voller Nominalphrasen (*die hat kaputt gemacht*; *die duscht den Baum*) als systematische Vereinfachungen festzustellen.

Ein entsprechendes Bild ergibt sich für den Bereich der Nominalflexion (Genus und Kasus), der in weiteren Auswertungsschritten anhand der elizitierten Äußerungen erfolgt. Wie Abbildung 4 zeigt, ist auch in diesem Bereich ein kontinuierlicher Anstieg der

Sprachleistungen festzustellen. Erfasst wird dabei der Erwerb des Natürlichen-Geschlechts-Prinzips (NGP) sowie des grammatischen Genus beim Subjekt und direkten Objekt. Beim grammatischen Genus wird dabei ermittelt, ob ein Kind beim Subjekt zumindest systematisch zwischen den Formen *der* und *die*, beim direkten Objekt zwischen *den* und *die* unterscheidet. Ein solches zweigliedriges System, bei dem das Neutrum unberücksichtigt bleibt, stellt einen Zwischenschritt beim frühen Zweitspracherwerb dar (Dieser 2008; Kaltenbacher; Klages 2007; Kaltenbacher et al. im Druck).

Die Abbildung zeigt auch, dass die Kinder die drei verschiedenen Manifestationsbereiche des Genus, das semantisch basierte NGP und das eher phonologisch bestimmte grammatische Genus beim Subjekt und Objekt¹⁴, nicht gleichzeitig, sondern etwas zeitversetzt erwerben. Diese Beobachtung gilt für Kinder ohne gezielte Förderung ebenso wie für solche, die mit „Deutsch für den Schulstart“ gefördert wurden, wo diese Abfolge als Progression im Fördermaterial vorgegeben ist.

Abb. 4: Ergebnisse beim NGP, grammatischen Genus beim Subjekt und direkten Objekt in zwei Förderjahren



Insgesamt zeichnen die Ergebnisse ein Bild, bei dem die geförderten Bereiche schrittweise im Rahmen von Entwicklungssequenzen erworben werden. Sie verdeutlichen auch, dass komplexe Sätze wie die bei den HSET-Subtest VS und IS überprüften weit jenseits des Kompetenzniveaus der geförderten Kinder liegen.

5. Fazit

Die Verfasser von EVAS vertreten den Standpunkt, dass eine den wissenschaftlichen Anforderungen genügende Evaluationsstudie nur auf der Basis standardisierter Testverfahren

¹⁴ Das Deutsche weist eine lautliche Tendenz bei der Genuszuweisung auf: Wörter, die auf unbetontes „e“ enden, sind in der überwiegenden Mehrheit feminin; einsilbige Wörter sind mehrheitlich maskulin.

möglich ist. Wie in Abschnitt 3 gezeigt wurde, nehmen sie dabei eine mangelnde Validität durch den Einsatz von Verfahren in Kauf, die die Entwicklungsprozesse der untersuchten Kinder im frühen Zweitspracherwerb nicht erfassen können. Dieser Mangel, der auch in einer aktuellen Studie zur Sprachförderung und Sprachdiagnostik (Redder et al. 2010) äußerst kritisch diskutiert wird, wird in der EVAS-Studie allerdings nicht thematisiert. Mögliche Gründe für den fehlenden Nachweis eines Fördereffekts werden in den Rahmenbedingungen der Förderung und dem didaktischen Vorgehen gesehen, nicht jedoch im Design der Studie selbst. Mit dieser fehlenden Einsicht in die Schwachpunkte der Studie geht eine offensive Verbreitung der Ergebnisse einher, die die wissenschaftliche Sorgfaltspflicht vermissen lässt. Die im Rahmen des Projekts „Deutsch für den Schulstart“ entwickelten, „weichen“ Testverfahren weisen kontinuierliche Entwicklungen der Kinder nach, für die jedoch eine Vergleichsgruppe von anders geförderten Kindern fehlt. Es lässt sich damit nicht feststellen, in wieweit die Fortschritte der Kinder auf die gezielte Förderung zurückzuführen sind. Die Ergebnisse von Montanari (2010) zum Genuserwerb weisen jedoch darauf hin, dass sich kontinuierliche Fortschritte in diesem Bereich ohne gezielte Fördermaßnahmen keineswegs selbstverständlich einstellen.

Die Ergebnisse von EVAS werfen auch die Frage auf, wie viel in einem Förderjahr realistisch erreicht werden kann. Die Studie suggeriert, dass mit einer einjährigen Maßnahme grundsätzliche Rückstände aufgeholt werden können. Die Erfahrungen der letzten Jahre zeigen jedoch, dass Sprachförderung längerfristig angelegt sein muss, um den Kindern im notwendigen Maße weiterzuhelfen. Welche Maßnahmen dabei für bestimmte Altersgruppen die besten Ergebnisse erbringen, ist weit von einer wissenschaftlichen Klärung entfernt. Die EVAS-Studie trägt bedauerlicherweise nicht zur Klärung dieser Frage bei.

Literaturnachweis

Dieser, Elena (2008): Strategien beim Erwerb des russischen und deutschen Genus (am Beispiel mehrsprachiger Kinder). In: *Zeitschrift für Slavistik* 53, 394 – 402.

Grimm, Hannelore; Schöler, Hermann; Wintermantel, Margret (Hrsg.) (1975). *Zur Entwicklung sprachlicher Strukturformen bei Kindern*. Weinheim u.a.: Beltz.

Grimm, Hannelore; Schöler, Hermann (1978): *Heidelberger Sprachentwicklungstest*. Göttingen: Hogrefe.

Haberzettl, Stefanie (2005): *Der Erwerb der Verbstellungsregeln in der Zweitsprache Deutsch durch Kinder mit russischer und türkischer Muttersprache*. Tübingen: Niemeyer.

Kaltenbacher, Erika; Klages, Hana (2007): Sprachprofil und Sprachförderung bei Vorschulkindern mit Migrationshintergrund. In: Ahrenholz, Bernt (Hrsg.): *Kinder mit Migrationshintergrund. Spracherwerb und Fördermöglichkeiten*. 2.Aufl. Freiburg: Fillibach, 80 – 97.

Kaltenbacher, Erika; Klages, Hana (2008): Deutsch für den Schulstart: Zielsetzungen und Aufbau eines Förderprogramms. In: Ahrenholz, Bernt (Hrsg.): *Deutsch als Zweitsprache – Voraussetzungen und Konzepte für die Förderung von Kindern und Jugendlichen mit Migrationshintergrund*. 2., überarb. u. erg. Aufl. Freiburg: Fillibach, 135 – 154.

Kaltenbacher, Erika; Klages, Hana; Pagonis, Giulio (2009): Projekt *Deutsch für den Schulstart* – Arbeitsbericht April 2009. <http://www.deutsch-fuer-den-schulstart.de/upload/arbeitsbericht.pdf>

Kaltenbacher, Erika; von Stutterheim, Christiane (2009): Stellungnahme zur EVAS-Studie. <http://www.deutsch-fuer-den-schulstart.de/upload/stellungnahme1.pdf>

Kaltenbacher, Erika; Klages, Hana; Vasylyeva, Tetyana (im Druck): Sprachförderung als Vorbereitung auf die sprachlichen Anforderungen der Schule: Zielsetzungen und ihre Realisierbarkeit. In: Mehlem, Ulrich; Sahel, Said (Hrsg.): *Erwerb schriftsprachlicher Kompetenzen im DaZ-Kontext: Diagnose und Förderung*. Freiburg: Fillibach.

Meisel, Jürgen (2007): Mehrsprachigkeit in der frühen Kindheit: Zur Rolle des Alters bei Erwerbsbeginn. In: Anstatt, Tanja (Hrsg.): *Mehrsprachigkeit bei Kindern und Erwachsenen*. Tübingen: Attempto, 93 – 113.

Nieding, Gerhild (2006): *Wie verstehen Kinder Texte? Die Entwicklung kognitiver Repräsentationen*. Lengerich: Pabst.

Redder, Angelika; Schwippert, Knut; Hasselhorn, Marcus; Forschner, Sabine; Fickermann, Detlef; Ehlich, Konrad (2010): Grundzüge eines nationalen Forschungsprogramms zu Sprachdiagnostik und Sprachförderung. In: *ZUSE-Diskussionspapier Nr. 1*. Hamburg: Zentrum zur Unterstützung der wissenschaftlichen Begleitung und Erforschung schulischer Entwicklungsprozesse.

Rogge, Klaus-Eckart (2009): Stellungnahme zur Methodik der Studie: „Evaluationsstudie zur Sprachförderung von Vorschulkindern (EVAS)“. <http://www.deutsch-fuer-den-schulstart.de/upload/stellungnahme2.pdf>

Polotzek, Silvana; Roos, Jeanette; Schöler, Hermann: *Evaluationsstudie zur Sprachförderung von Vorschulkindern. Vorläufiger Abschlussbericht*, 16.03.2009.

Wegener, Heide (1995): Das Genus im DaZ-Erwerb. Beobachtungen an Kindern aus Polen, Russland und der Türkei. In: Handwerker, Brigitte (Hrsg.): *Fremde Sprache Deutsch*. Tübingen: Narr, 1 – 24.

Wegener, Heide (1998): Das Passiv im DaZ-Erwerb von Grundschulkindern. In: Wegener, Heide (Hrsg.) *Eine zweite Sprache lernen. Empirische Untersuchungen zum Zweitspracherwerb*. Tübingen: Narr, 143-172.

Wottawa, Heinrich; Thierau, Heike (1990): *Lehrbuch Evaluation*. Bern u.a.: Huber.