



NEPS WORKING PAPERS

Kristin Hecker, Anna Südkamp, Carina Leser und
Sabine Weinert

ENTWICKLUNG EINES TESTS ZUR ERFASSUNG VON HÖRVERSTEHEN AUF TEXTEBENE BEI SCHÜLERIN- NEN UND SCHÜLERN DER KLASSEN- STUFE 9

NEPS Working Paper No. 53
Bamberg, März 2015

Working Papers of the German National Educational Panel Study (NEPS)

at the Leibniz Institute for Educational Trajectories (LifBi) at the University of Bamberg

The NEPS Working Papers publish articles, expertises, and findings related to the German National Educational Panel Study (NEPS).

The NEPS Working Papers are edited by a board of researchers representing the wide range of disciplines covered by NEPS. The series started in 2011.

Papers appear in this series as work in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character.

Any opinions expressed in this series are those of the author(s) and not those of the NEPS Consortium.

The NEPS Working Papers are available at

<https://www.neps-data.de/projektübersicht/publikationen/nepsworkingpapers>

Editorial Board:

Jutta Allmendinger, WZB Berlin

Cordula Artelt, University of Bamberg

Jürgen Baumert, MPIB Berlin

Hans-Peter Blossfeld, EUI Florence

Wilfried Bos, University of Dortmund

Claus H. Carstensen, University of Bamberg

Henriette Engelhardt-Wölfler, University of Bamberg

Frank Kalter, University of Mannheim

Corinna Kleinert, IAB Nürnberg

Eckhard Klieme, DIPF Frankfurt

Cornelia Kristen, University of Bamberg

Wolfgang Ludwig-Mayerhofer, University of Siegen

Thomas Martens, DIPF Frankfurt

Manfred Prenzel, TU Munich

Susanne Rässler, University of Bamberg

Marc Rittberger, DIPF Frankfurt

Hans-Günther Roßbach, LifBi

Hildegard Schaeper, DZHW Hannover

Thorsten Schneider, University of Leipzig

Heike Solga, WZB Berlin

Petra Stanat, IQB Berlin

Volker Stocké, University of Kassel

Olaf Struck, University of Bamberg

Ulrich Trautwein, University of Tübingen

Jutta von Maurice, LifBi

Sabine Weinert, University of Bamberg

Contact: German National Educational Panel Study (NEPS) – Leibniz Institute for Educational Trajectories – Wilhelmsplatz 3 – 96047 Bamberg – Germany – contact@lifbi.de

Entwicklung eines Tests zur Erfassung von Hörverstehen auf Textebene bei Schülerinnen und Schülern der Klassenstufe 9

*Kristin Hecker, Leibniz-Institut für Bildungsverläufe
Anna Südkamp, Technische Universität Dortmund
Carina Leser, Leibniz-Institut für Bildungsverläufe
Sabine Weinert, Otto-Friedrich-Universität Bamberg
Unter spezifischer Mitarbeit von Sarah Gunkel, Leibniz-Institut für
Bildungsverläufe*

E-Mail-Adresse der Erstautorin:

kristin.hecker@lifbi.de

Bibliographische Angaben bei deutschsprachigen Papers:

Hecker, K., Südkamp, A., Leser, C. & Weinert, S. (2015). Entwicklung eines Tests zur Erfassung von Hörverstehen auf Textebene bei Schülerinnen und Schülern der Klassenstufe 9 (NEPS Working Paper No. 53). Bamberg: Leibniz-Institut für Bildungsverläufe, Nationales Bildungspanel.

Danksagung

Diese Arbeit entstand mit Unterstützung der Professorinnen und Professoren sowie Mitarbeiterinnen und Mitarbeitern des Arbeitsbereichs „Kompetenzentwicklung im Lebenslauf“ des NEPS. An dieser Stelle möchten wir uns insbesondere bei Frau Prof. Dr. Cordula Artelt und Herrn Prof. Dr. Claus H. Carstensen sowie bei Anika Langmann und den wissenschaftlichen Hilfskräften Miriam Hildenbrand, Sarah Seiler und Max Schnell für die unermüdliche Arbeit und die zahlreichen konstruktiven Gespräche bedanken, die einen erheblichen Beitrag zur Testentwicklung beigetragen haben.

Entwicklung eines Tests zur Erfassung von Hörverstehen auf Textebene bei Schülerinnen und Schülern der Klassenstufe 9

Abstract

Im Nationalen Bildungspanel werden verschiedene Indikatoren von Sprachkompetenz erfasst. Während einige dieser Indikatoren kohärent über die Lebensspanne gemessen werden (z. B. der rezeptive Wortschatz, die Lesekompetenz), werden andere Indikatoren nur zu ausgewählten Zeitpunkten im Bildungsverlauf erfasst. Dazu gehört auch das textbezogene Hörverstehen im Deutschen, das in Klassenstufe 9 gemessen wird.

In diesem Beitrag wird die Entwicklung eines Tests zur Erfassung von Hörverstehen auf Textebene beschrieben. Zunächst wird auf die theoretischen und empirischen Grundlagen der Testentwicklung eingegangen. Dabei finden die Konzeption und Messung von Hörverstehen in anderen nationalen und internationalen Large Scale Assessments, Modelle von Sprachkompetenz und Theorien zur kognitiven Verarbeitung von gesprochener Sprache Berücksichtigung. Außerdem wird der Stellenwert von Hörverstehen im Deutschen innerhalb der Konzeption zur Erfassung von Sprachkompetenz im Nationalen Bildungspanel deutlich gemacht. Aus dem theoretischen und empirischen Hintergrund wird daraufhin ein Konzept zur Messung von Hörverstehen im Deutschen abgeleitet. Im Anschluss daran wird die konkrete Umsetzung eines Tests zur Erfassung von Hörverstehen in der neunten Klassenstufe beschrieben. Dabei werden die Kriterien für die Auswahl von Hörtexten als Testmaterial veranschaulicht. Außerdem wird dargestellt, wie Überlegungen zur Aufgabenform (Fragen zur den Hörtexten) und zum Antwortformat umgesetzt wurden. Schließlich wird ein Ausblick auf den Einsatz des Testverfahrens im Nationalen Bildungspanel gegeben.

Schlagworte

Hörverstehen, Textverarbeitung, Kompetenzmessung, Testentwicklung

Abstract

The National Educational Panel Study (NEPS) includes the assessment of different language indicators. While some language indicators are measured coherently across the lifespan (e. g., receptive vocabulary, reading competence), others are measured at critical points in individuals' educational careers. Amongst them is the indicator of listening comprehension in German language, which is measured at text-level in Grade 9.

In this paper, we describe the development of a test of listening comprehension at text-level. First, the theoretical and empirical background of test development is presented. It acknowledges concepts on the assessments of listening comprehension in other national and international large scale assessments, draws on different models of language competence, and deals with theories on the cognitive processing of spoken language. Furthermore, the role of the assessment of listening comprehension within the framework for the assessment of language competencies within the NEPS is highlighted. On the basis of this theoretical and empirical background, we build up a framework for the assessment of

listening comprehension. Then, the practical realization of a test for the assessment of listening comprehension in Grade 9 is presented. Here, we give a rationale for selecting certain episodes of spoken language as test material. We also explain how the items and the response format were implemented. Finally, we give an outline on the implementation of the test within the NEPS.

Keywords

Listening comprehension, text comprehension, competence assessment, test development

Inhalt

1. Einleitung.....	5
2. Die Erfassung von Sprachindikatoren in nationalen und internationalen Large Scale Assessments	6
2.1 Internationale Bildungsstudien mit deutscher Beteiligung	6
2.2 Nationale Bildungsstudien in den USA.....	7
2.3 Nationale Bildungsstudien in Deutschland	8
3. Hörverstehen als Teilfähigkeit sprachlicher Kompetenz	11
3.1 Modelle von Sprachkompetenz: Komponenten von Sprachwissen und funktional-integrative Sicht.....	11
3.2 Lese- und Hörverstehen als kognitiver Verarbeitungsprozess: Gemeinsamkeiten und Unterschiede	12
3.3 Merkmale gesprochener Sprache	15
4. Einordnung des Hörverstehens in das Konzept zur Erfassung von Sprachkompetenz im Nationalen Bildungspanel	16
4.1 Rahmenkonzept zur Erfassung der Lesekonzept im NEPS.....	17
4.2 Weitere Sprachindikatoren im NEPS.....	17
5. Konzept zur Erfassung von Hörverstehen auf Textebene im Nationalen Bildungspanel	19
6. Entwicklung eines Tests zur Erfassung von Hörverstehen.....	21
6.1 Auswahl von Hörtexten	21
6.2 Itementwicklung und Antwortformat.....	22
6.3 Merkmale zur Beschreibung der Testschwierigkeit.....	25
7. Fazit und Ausblick.....	27
8. Literatur.....	30

1. Einleitung

Das Nationale Bildungspanel (National Educational Panel Study – NEPS) ist eine deutschlandweite Large Scale Studie, deren Ziel es ist, Bildungsverläufe abzubilden und die Bedingungen, Moderatoren und Folgen von Bildungsprozessen zu identifizieren. Im NEPS wird ein Multikohortensequenzdesign umgesetzt, das die regelmäßige Befragung und Testung repräsentativer Stichproben zu bedeutsamen Zeitpunkten im Bildungsverlauf vorsieht. Daten über Kompetenzen und deren Entwicklung über die Lebensspanne sind dabei ein zentraler Anknüpfungspunkt für alle anderen Bereiche der Studie (Artelt, Weinert & Carstensen, 2013; Weinert, Artelt, Prenzel, Senkbeil, Ehmke & Carstensen, 2011).

Sprachkompetenzen, wie z. B. die Lesekompetenz in der Mehrheitssprache, gehören zweifelsohne zu den zentralen, bildungsrelevanten kognitiven Kompetenzen (siehe z. B. Weinert, 2006). Im NEPS werden deshalb verschiedene, gezielt ausgewählte Sprachindikatoren erfasst. Die Auswahl der Sprachindikatoren und die Erarbeitung einer Rahmenkonzeption zur Erfassung dieser Indikatoren basieren jeweils auf umfangreichen Vorüberlegungen. Das NEPS Konzept zur Erfassung von Sprachkompetenz sieht dabei auch vor, Hörverstehen im Deutschen¹ als rezeptive Sprachkompetenz zu erfassen.

Im vorliegenden Beitrag geht es um die Konzeption und Entwicklung eines Tests zur Erfassung von Hörverstehen, der sich in die Gesamtkonzeption im NEPS eingliedert und theoretische sowie praktische Aspekte berücksichtigt, die spezifisch für das Hörverstehen sind. Aufgrund ihrer großen bildungsbezogenen und gesellschaftlichen Relevanz, werden Sprachkompetenzen seit langem auch in anderen Large Scale Assessments erfasst. Deshalb wird zunächst ein kurzer Überblick darüber gegeben, welche domänenspezifischen und insbesondere welche sprachlichen Kompetenzen in anderen nationalen und internationalen Large Scale Assessments erfasst werden. Außerdem wird jeweils darauf eingegangen, ob und wenn ja, wie das Hörverstehen auf Textebene erfasst wird (Abschnitt 2). Für die Verortung von Hörverstehen innerhalb der Sprachkompetenz wird zudem auf verschiedene Modelle von Sprachkompetenz eingegangen (Abschnitt 3.1). In die Konzeption fließen darüber hinaus Überlegungen zu Gemeinsamkeiten und Unterschieden in den kognitiven Verarbeitungsprozessen beim Lese- und Hörverstehen ein (3.3) sowie Merkmale gesprochener Sprache, an die Hörverstehen als rezeptive Fähigkeit gebunden ist (3.3). Im Anschluss an die Beschreibung unserer Vorüberlegungen wird das Hörverstehen in das Konzept zur Erfassung der Sprachkompetenz im NEPS eingeordnet (4). Basierend auf diesen theoretischen Überlegungen haben wir die Konzeption zur Erfassung des Hörverstehens im NEPS spezifiziert (5), bevor schließlich die Schritte der Testentwicklung und der neu entwickelte Test konkret beschrieben werden (6).

¹ Hier geht es um die Erfassung von Hörverstehen in der deutschen Sprache als gesellschaftliche Mehrheits- und Unterrichtssprache in Deutschland (Weinert, 2010), unabhängig davon, ob Deutsch als Erst- oder Zweitsprache erworben wurde. Bei Personen mit türkischem und russischem Sprach- bzw. Migrationshintergrund wird in NEPS zudem das Hörverstehen in diesen Sprachen erfasst.

2. Die Erfassung von Sprachindikatoren in nationalen und internationalen Large Scale Assessments

In Vorbereitung auf die Entwicklung eines Tests zur Erfassung von Hörverstehen auf Textebene stellen sich die Fragen, welche sprachlichen Kompetenzen in anderen Bildungsstudien erfasst werden, ob Hörverstehen auf Textebene als eigenständige sprachliche Kompetenz berücksichtigt wird, und welche Definitionen und Operationalisierungen zugrunde gelegt werden. Diese Fragen sollen zunächst unter Berücksichtigung verschiedener internationaler Schulleistungsuntersuchungen, an denen sich Deutschland beteiligt, beantwortet werden. Im Anschluss daran werden nationale Bildungsstudien innerhalb der USA und Deutschlands beschrieben.

2.1 Internationale Bildungsstudien mit deutscher Beteiligung

Das deutsche Bildungswesen nimmt regelmäßig an internationalen Schulleistungsuntersuchungen teil, um die Leistungsfähigkeit der deutschen Schülerinnen und Schüler im internationalen Vergleich festzustellen und gegebenenfalls Maßnahmen für das Bildungswesen ableiten zu können. Derzeit nimmt Deutschland am „Programme for International Student Assessment“ (PISA), an der „Internationalen Grundschul-Lese-Untersuchung“ (IGLU bzw. PIRLS²) und an den „Trends in International Mathematics and Science Studies“ (TIMSS) teil (Autorengruppe Bildungsberichterstattung, 2012).

PISA erfasst Kenntnisse und Fähigkeiten von 15-jährigen Schülerinnen und Schülern in den Bereichen Textverständnis/Lesen, Mathematik und Naturwissenschaften. Die Studie wurde von der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) initiiert und wird seit dem Jahr 2000 alle drei Jahre durchgeführt (Artelt, Drechsel, Bos & Stubbe, 2008; Helmke, 2007; Baumert et al. 2001). Pro Erhebungswelle wird jeweils eine der drei Leistungen schwerpunktmäßig erfasst. In den Jahren 2000 und 2009 lag der Schwerpunkt auf der Lesekompetenz, 2003 und 2012 auf Mathematik und 2006 wurde die naturwissenschaftliche Grundbildung schwerpunktmäßig erfasst (Klieme et al., 2010; OECD, 2010). Im Jahr 2012 wurden darüber hinaus die Bereiche Problemlösen und „financial literacy“ mit in die Erhebung aufgenommen (OECD, 2013). In PISA wird die Lesekompetenz über das Textverständnis erfasst und ist somit an einem funktional-integrativen Modell zur Messung von Sprachkompetenz orientiert (hierzu siehe auch Abschnitt 4). Zudem liegt dem Konzept der Literacy-Gedanke, also die Erfassung von Kompetenzen in möglichst authentischen Situationen, zu Grunde. Bei der Messung von Lesekompetenz werden drei Dimensionen berücksichtigt: Art des Lesestoffs (Textformat), Art der Leseaufgabe sowie Art des Gebrauchs, für den der Text geschrieben wurde (Situationen bzw. Kontext des Textes; Deutsches PISA - Konsortium, 2000; OECD, 2010). In der PISA-Studie werden zwei Textformate unterschieden: kontinuierliche Texte bzw. Prosatexte (z. B. Erzählungen, Argumentationen) und diskontinuierlichen Texte wie z. B. Formulare oder Grafiken (Artelt et al., 2008; Deutsches PISA – Konsortium, 2000; OECD, 2010). Ziel der PISA-Studie für den Bereich Lesekompetenz ist es, insbesondere das Textverständnis und die Fähigkeit, Texte einzuordnen, zu erfassen. Die Auswahl der Lesetexte erfolgt anhand unterschiedlicher Lesesituationen, indem Texte aus unterschiedlichen Umfeldern privater (z. B. Romane),

² PIRLS ist die internationale Abkürzung (Progress in International Reading Literacy Study). Im Folgenden wird ausschließlich die deutsche Abkürzung IGLU verwendet.

öffentlicher (z. B. Anschreiben), berufsbezogener (z. B. Manuale) und bildungsbezogener Art (z. B. Lehrbücher) eingesetzt werden. Weiterhin werden vier Dimensionen kognitiver Anforderungen unterschieden: „Informationen aus einem Text extrahieren“, „eine textbasierte Interpretation entwickeln“, „über den Inhalt des Textes reflektieren“ und „über die Form des Textes reflektieren“ (Artelt et al., 2008; Deutsches PISA - Konsortium, 2000; Klieme et al., 2010; OECD, 2010).

Die IGLU-Studie ist ein Large Scale Assessment, das im Bereich der Kompetenzmessung spezifisch auf die Erfassung von Sprachkompetenz ausgerichtet ist. Auch hier wird auf die Erfassung rezeptiv-schriftlicher Sprachkompetenz (Lesekompetenz) fokussiert. Im Rahmen der IGLU-Studie wird die Lesekompetenz von Schülerinnen und Schülern am Ende der 4. Jahrgangsstufe getestet. Die Studie wird seit 2001 in einem fünfjährigen Rhythmus durchgeführt. Initiator und Verantwortlicher auf internationaler Ebene ist die International Association for the Evaluation of Educational Achievement (IEA; Campbell, Kelly, Mullis, Martin & Sainsbury, 2001). Deutschland nimmt daran teil, um für die Primarstufe einen internationalen Vergleich gewährleisten zu können und einen Überblick über das Leseverständnis der Schülerinnen und Schüler, ihre Einstellung zum Lesen und ihre Lesegewohnheiten am Ende der Grundschulzeit zu bekommen (Bundesministerium für Bildung und Forschung, 2012). Es werden drei Dimensionen der Lesekompetenz spezifiziert: Leseintention, Prozesse des Leseverständnisses sowie Lesemotivation, Leseselbstkonzept und Leseverhalten (Bos et al., 2007). Wie bei PISA werden Texte unterschiedlicher Art und Aufgaben mit verschiedenen kognitiven Anforderungen berücksichtigt (Bos et al., 2005).

2.2 Nationale Bildungsstudien in den USA

Auch in anderen Ländern werden regelmäßig Large Scale Assessments durchgeführt, um die Kompetenzen von Schülerinnen und Schülern und deren Entwicklung zu erheben. In der Early Childhood Longitudinal Study (ECLS) werden drei altersunterschiedliche Kohorten längsschnittlich über mehrere Jahre begleitet. Gegenstand der Studie ist die Erfassung der emotionalen, kognitiven, sozialen und körperlichen Entwicklung der Kinder sowie ihre Fähigkeiten und Fertigkeiten in diesen Bereichen während wichtiger Schnittstellen in ihrer Entwicklung (z. B. Übergang zur Schule; Mulligan, Hasted & McCarrol, 2012). Außerdem wird die Beziehung zwischen verschiedenen familiären, schulischen, individuellen und gesellschaftlichen Variablen und der Entwicklung des Kindes, seinem Lernverhalten sowie seinen Leistungen in der Schule ermittelt (Pollack, Atkins-Burnett, Rock & Weiss, 2005). Dafür wird eine breite Palette an Befragungsinstrumenten für Kinder, Eltern, Lehrer und andere Interaktionspartner eingesetzt (Njarian, Snow, Lennon & Kinsey, 2010; Pollack et al., 2005). Aus dem Bereich der Sprachkompetenz werden Basisfähigkeiten (z. B. Erkennung von Buchstaben), Wortschatz und das Leseverstehen getestet (Pollack et al., 2005). Dabei sind unmittelbares Verstehen von Texten, die Erstellung von Zusammenhängen und Interpretation von Textteilen sowie persönliche Reflektion der Inhalte und die Entwicklung einer kritischen Einstellung von Relevanz (Pollack et al., 2005).

Das National Assessment of Educational Progress (NAEP) ist die wohl größte nationale Längsschnittstudie in Amerika. Sie wird seit 1969 alle vier Jahre bei Personen im Alter von 9, 13 und 17 Jahren durchgeführt und soll ein Bild über die schulische Entwicklung von Schülerinnen und Schülern geben (U.S. Department of Education, 1999; U.S. Department of Education, 2003). Die Untersuchungen umfassen zwölf Bereiche, u. a. Mathematik,

Naturwissenschaften, Lesen und Schreiben, aber auch Wirtschaft, Geographie, Kunst oder Geschichte. Das Konstrukt zur Lesekompetenz basiert auf der Annahme, dass unterschiedliche Texte mit differenten Absichten gelesen werden und unterscheidet somit zwischen informierenden (Expositionen, argumentative und überzeugende Texte) sowie literarischen Texten (fiktionalen, literarischen nichtfiktionalen und poetischen Texten). Die Testfragen zielen auf die kognitiven Anforderungen *Lokalisieren und Wiedergeben*, *Integrieren und Interpretieren* und *Reflektieren und Bewerten* (U.S. Department of Education, 1999; National Center for Education Statistics, 2011).

Das National Assessment of Adult Literacy (NAAL) ist ein amerikanisches Large Scale Assessment aus dem Jahr 2003, das auf die Erfassung von Sprachkompetenzen fokussiert. Hier wurden englischsprachige Lese- und Schreibkompetenzen von 16-Jährigen erfasst (Baldi et al., 2009; Kutner, Greenberg & Baer, 2005; White & Dillow, 2005). NAAL unterscheidet drei Arten von Lese- und Schreibfähigkeit, die auf unterschiedlichen Skalen gemessen werden: Prosa mittels kontinuierlicher Texte, Dokumente mittels diskontinuierlicher Texte und „quantitative Literatur“. Mit „quantitativer Literatur“ sind Texte gemeint (Prosa oder Dokumente), die Zahlen beinhalten (z. B. ein Checkbuch ausgleichen oder Berechnungen durchführen) und den Umgang mit Zahlen innerhalb eines gedruckten Textes erfordern (Baldi et al., 2009; Kutner, Greenberg & Baer, 2005; White & Dillow, 2005).

2.3 Nationale Bildungsstudien in Deutschland

Auf nationaler Ebene verfolgt Deutschland ebenfalls ein umfassendes Bildungsmonitoring. Damit soll eine regelmäßige Überprüfung des deutschen Bildungswesens sichergestellt und Maßnahmen der Unterrichts- und Qualitätsentwicklung eingeleitet werden.

In den Jahren 2001 bis 2006 wurde unter Leitung des Deutschen Instituts für Internationale Pädagogische Forschung (DIPF) die repräsentative Studie Deutsch-Englisch-Schülerleistungen International (DESI) durchgeführt (Beck & Klieme, 2007). Zu Beginn und am Ende des 9. Schuljahres wurden die sprachlichen Kompetenzen von Neuntklässlern umfassend getestet. Im Mittelpunkt der Studie standen rezeptive (Leseverstehen und Hörverstehen) und produktive (schriftliche Textproduktion und Sprechen) Sprachkompetenzen im Deutschen und Englischen. Darüber hinaus wurden u. a. Bedingungen von Kompetenzerwerb und -entwicklung, Sprachbewusstheit (Einstellung zur Sprache) und interkulturelle Kompetenzen untersucht (Klieme, 2006). Das Konzept der Sprachkompetenz im Fach Deutsch umfasst in DESI Wortschatz, Lesen, Argumentation, Bewusstheit, Rechtschreibung und Textproduktion. Dabei wurde auf die schriftsprachlichen Kompetenzen Lesen und Schreiben fokussiert. Hörverstehen im Deutschen wurde in der DESI-Studie nicht erfasst. Es kam aber ein englischsprachiger Hörverstehenstest zum Einsatz (Jude & Klieme, 2007; Klieme et al., 2006).

An die Stelle von Erweiterungen der IGLU- und PISA-Studie (IGLU-E, PISA-E) zum nationalen Schulleistungsvergleich sind seit 2009 Ländervergleiche gerückt, deren Grundlage die Bildungsstandards der Kultusministerkonferenz KMK bilden (Böhme & Bremerich-Vos, 2012). Die Bildungsstandards der KMK „legen fest, welche Kompetenzen die Schülerinnen und Schüler bis zu einer bestimmten Jahrgangsstufe an wesentlichen Inhalten erworben haben sollen“ (KMK, 2005a, S. 9). Die Ländervergleiche werden vom Institut zur Qualitätsentwicklung im Bildungsbereich (IQB) durchgeführt und überprüfen auf Basis von Länderstichproben, ob die Bildungsstandards erreicht werden (KMK, 2006). Sie finden

zeitgleich mit PISA und damit alle fünf Jahre im Primarbereich und alle sechs Jahre im Sekundarbereich statt (KMK, 2006). Die Ländervergleiche finden in den Bereichen Deutsch, erste Fremdsprache, Mathematik und Naturwissenschaften in allen allgemeinbildenden Schulen aller Schularten des Regelschulsystems und für alle Bundesländer statt. Es werden jeweils Kompetenztests sowie Schüler-, Eltern- und Lehrerfragebögen eingesetzt (Böhme & Bremerich-Vos, 2012).

In den Bildungsstandards (KMK, 2004; KMK, 2005b) sind für das Fach Deutsch verschiedene Kompetenzbereiche festgelegt: Sprache und Sprachgebrauch, Sprechen und Zuhören, Schreiben und Lesen (mit Texten und Medien umgehen). Der Kompetenzbereich Sprechen und Zuhören umfasst wiederum fünf Teilkompetenzen: (1) zu anderen sprechen, (2) verstehend zuhören, (3) Gespräche führen, (4) szenisch spielen und (5) über Lernen sprechen. Damit werden sowohl produktive als auch rezeptive Sprachkompetenzen berücksichtigt. Der Bereich „verstehend zuhören“ meint die Fähigkeit, Gesprächsbeiträge anderer zu verfolgen und aufzunehmen, wesentliche Informationen zu verstehen, wiederzugeben und kritisch zu hinterfragen und Aufmerksamkeit für verbale und nonverbale Äußerungen zu entwickeln. Für die Erhebung des Bereiches Sprechen und Zuhören schlagen die Bildungsstandards zum Beispiel Gruppendiskussionen, Rollenspiele oder Präsentationen vor (KMK, 2004; KMK, 2005b; Böhme & Bremerich-Vos, 2012).

Im Jahr 2009 wurden in der Sekundarstufe für Deutsch und die erste Fremdsprache die Teilkompetenzen Leseverstehen und Hörverstehen erhoben. Das Hörverstehen wurde anhand eines Tests erfasst, bei dem die Schülerinnen und Schüler verschiedene Hörtexte von CD vorgespielt bekommen und anschließend Fragen zu den Texten beantworten sollen. Die Bildungsstandards unterscheiden literarische und informierende Textsorten (Behrens & Eriksson, 2009). In Anlehnung an Tests zur Erfassung von Lesekompetenz in PISA wurden bei der Erfassung des Hörverstehens zur Überprüfung der Bildungsstandards Aufgaben mit unterschiedlichen kognitiven Anforderungen entwickelt. Dabei wird zwischen den Anforderungen „Informationen ermitteln“, „textbezogen Interpretieren“ und „Reflektieren und Bewerten“ unterschieden. Es gab sowohl Aufgaben im Multiple-Choice-Format als auch kurze offene Fragen (Böhme & Bremerich-Vos, 2012). Der Test wurde in zwei Varianten durchgeführt. In einer Variante wurde zunächst der Hörtext präsentiert. Dabei hatten die Schülerinnen und Schüler keine Kenntnis von den Fragen zum Text, die sie später in ihrem Aufgabenheft schriftlich beantworten sollten. In der anderen Variante wurden den Schülerinnen und Schülern zunächst die Aufgaben vorgelegt und sie wurden angeleitet, die Aufgaben sorgfältig zu lesen. Erst danach wurde der Hörtext präsentiert und die Schülerinnen und Schüler sollten die Fragen im Aufgabenheft bearbeiten. Bei dieser Variante bestand also die Möglichkeit, den Hörtext gezielt im Hinblick auf relevante Informationen anzuhören.

Auch beim Ländervergleich 2011 im Primarbereich wurde ein Hörverstehentest eingesetzt, der ähnlich konzipiert war wie der Test im Sekundarbereich. Der Test aus dem Kompetenzbereich Zuhören bestand aus fünf Hörtexten mit insgesamt 51 schriftlich vorgegebenen Items. Die Art der Hörtexte reichte von klassischen Hörspielen bis zu Kinderradiosendungen. Die Hörtexte wurden den Kindern einmal von CD vorgespielt. Anschließend mussten sie die dazugehörigen Fragen beantworten. Insgesamt gab es 30 Items mit einem geschlossenen Format, 15 Items mit einem halboffenen Format, das eine Kurzantwort verlangte, und sechs offene Items, bei denen mehrere Sätze mit Begründungen

gefragt waren. Die Hörtexte hatten eine Länge von zwei bis viereinhalb Minuten, nur ein Hörtext war neun Minuten lang, um Schüler im oberen Kompetenzbereich differenzieren zu können. Die Gesamttestzeit im Fach Deutsch betrug an Regelschulen 80 Minuten (Böhme & Bremerich-Vos, 2012).

Neben den Ländervergleichen werden seit 2003 jährlich Vergleichsarbeiten (VERA-3 und VERA-8) durchgeführt, die der Standortbestimmung und Überprüfung der Leistungsfähigkeit von Schulen und Klassen der dritten und achten Jahrgangsstufe dienen (KMK, 2006; Helmke, 2007). Sie sind für alle allgemeinbildenden Schulen in Deutschland verpflichtend (mit Ausnahme von Baden-Württemberg). Die Aufgaben von VERA orientieren sich an den KMK-Bildungsstandards (KMK, 2012). Die konkrete Auswahl der Testaufgaben sowie die Durchführung des Tests, die Auswertung und Rückmeldung der Ergebnisse übernehmen die jeweiligen Schulen (Helmke, 2007). Bei VERA werden die Fächer Deutsch und Mathematik getestet, wobei mindestens eines der Fächer obligatorisch untersucht werden muss. Wird das Fach Deutsch getestet, so muss mindestens der Kompetenzbereich *Lesen* untersucht werden. Weitere, optionale Kompetenzbereiche sind *Sprechen und Zuhören*, *Schreiben* sowie *Sprache und Sprachgebrauch untersuchen*. Die Domäne Zuhören wird somit optional getestet. Bei VERA-8 wird zusätzlich die erste Fremdsprache erfasst. Dabei soll mindestens der Bereich Lese- oder (fremdsprachliches) Hörverstehen getestet werden (KMK, 2012). Da die KMK-Bildungsstandards die Grundlage für die Vergleichsarbeiten bilden, liefern sie auch die theoretischen Annahmen über die Kompetenzbereiche. Das Konstrukt Hörverstehen wird hier also genauso verstanden wie unter dem Punkt „Ländervergleiche“ bereits beschrieben (KMK, 2005b). Die Aufgaben zum Zuhören werden den Schülern in gedruckter Form vorgelegt, jedoch dürfen sie diese erst nach dem Hören des Textes anschauen. Während dem Hören des Textes ist es den Schülerinnen und Schülern erlaubt, sich Notizen zu machen. Die Fragen zu den Hörtexten haben unterschiedliche Formate. So gibt es z. B. einfache und komplexe Multiple Choice Fragen und offene Fragen (IQB, 2013).

Zusammenfassend lässt sich festhalten, dass in vielen Large Scale Assessments auf die Erfassung der Lesekompetenz als einem rezeptiven Sprachindikator fokussiert wird. Hörverstehen auf Textebene wird dagegen nur selten im Rahmen von Large Scale Assessments erfasst. Eine Ausnahme bilden in Deutschland die Ländervergleiche und Vergleichsarbeiten, die Instrumente zur Erfassung des Hörverstehens auf Textebene in bestimmten Klassenstufen implementieren.

Gründe für den Fokus auf die Lesekompetenz liegen unter anderem in den Beschränkungen in Large Scale Assessments, die beispielsweise die Möglichkeiten der Testdurchführung oder Zeitbeschränkungen betreffen.

Auch das NEPS setzt bei der Erfassung der Sprachkompetenz einen Schwerpunkt darauf, die Lesekompetenz kohärent über die Lebensspanne zu erfassen. Lesekompetenz ist einerseits von großer Bedeutung für Bildung, Bildungsverläufe und gesellschaftliche Teilhabe. Andererseits ermöglichen die Erfassung von Lesekompetenz im NEPS vergleichende Analysen mit anderen nationalen und internationalen Studien. Mit Blick auf die Messung kann auf bewährte Modelle und Rahmenkonzeptionen zurückgegriffen und aufgebaut werden.

Dennoch erscheint es aus praktischer und theoretischer Sicht wichtig, zu bestimmten Zeitpunkten im Bildungsverlauf über die Lesekompetenz hinaus weitere Sprachindikatoren zu erfassen. Einer dieser Indikatoren ist das Hörverstehen auf Textebene.

3. Hörverstehen als Teilfähigkeit sprachlicher Kompetenz

Ganz allgemein kann Hörverstehen zunächst als die Fähigkeit definiert werden, Aspekte gesprochener oder auditiv vermittelter Sprache zu verstehen (Porsch & Tesch, 2010). Hörverstehen ist eine sehr früh verfügbare Fähigkeit, die dem Erwerb anderer sprachlicher Fähigkeiten – insbesondere dem Schriftspracherwerb, also dem Lesen und Schreiben – zunächst vorausgeht (Weinert & Grimm, 2012). Der Erwerb sprachlicher Fähigkeiten ist einerseits genetisch verankert, wird aber andererseits durch soziale und bildungsbedingte Faktoren beeinflusst (z. B. Weinert & Ebert, 2013). Außerdem wird dem Spracherwerb eine grundlegende Funktion für den Erwerb weiterer domänenspezifischer Kompetenzen und metakognitiver Fähigkeiten zugeschrieben (Weinert et al., 2011; Weinert, 2007). So ermöglicht Hörverstehen, wie später auch das Lesen, über die gesamte Lebensspanne hinweg die Aneignung von neuem Wissen in den unterschiedlichsten Bereichen (Jude & Klieme, 2007). Ebenso wie das Sprechen, das Lesen und das Schreiben ist die Fähigkeit des Hörverstehens eine Schlüsselkompetenz für die Bewältigung von Alltagsanforderungen und wichtig für gesellschaftliche Teilhabe (Feyten, 1991). Hörverstehen steht dabei mit anderen sprachlichen Fähigkeiten und Fertigkeiten in Zusammenhang. Für die Erfassung von Hörverstehen ist es daher notwendig, diese Fähigkeit zunächst im Bereich der Sprachkompetenz zu verorten und definitorisch abzugrenzen. In der Forschung haben sich zwei Ansätze zur Konzeptualisierung von Sprachkompetenz bewährt: funktionale Kompetenzmodelle und Komponenten-orientierte Modelle.

3.1 Modelle von Sprachkompetenz: Komponenten von Sprachwissen und funktional-integrative Sicht

Aus pädagogisch-bildungsbezogener Perspektive rücken funktionale Kompetenzmodelle in den Vordergrund, die zwischen produktiven und rezeptiven Kompetenzen der gesprochenen und geschriebenen Sprache unterscheiden (Berendes, Weinert, Zimmermann & Artelt, 2013; Weinert, 2010; Jude & Klieme, 2007). Sprachkompetenz wird vor diesem Hintergrund des Sprachgebrauchs (Verstehen und Produzieren von Sprache) als Komplex von Fähigkeiten betrachtet, die sich in vier Bereiche differenzieren lassen: interaktives Sprechen als produktiv-auditive, Schreiben als produktiv-schriftsprachliche, Lesekompetenz als rezeptiv-schriftsprachliche und Hörverstehen als rezeptiv-auditive Teilfähigkeit (Weinert & Grimm, 2012; Berninger & Abbott, 2010; Leucht, Retelsdorf, Möller & Köller, 2010; Jude & Klieme, 2007). Diese funktional-integrative Perspektive auf Sprachkompetenz berücksichtigt in besonderer Weise reale Situationen, in denen Kommunikation stattfindet und in denen entsprechende Teilfähigkeiten für einen kontextangemessenen Sprachgebrauch interagieren (Berendes et al., 2013).

Komponenten-orientierte Modelle werden häufig in der psychologischen und sprachwissenschaftlichen Forschung angewendet und konzeptualisieren Sprachkompetenz primär aus der Perspektive von Sprachkomponenten bzw. sprachlichen Wissenssystemen (Berendes et al., 2013; Weinert, 2010). Für den erfolgreichen Umgang mit Sprache ist das Zusammenwirken der verschiedenen Wissenssysteme bzw. Sprachkomponenten notwendig, die sich nach Weinert und Grimm (2012) in prosodisches, linguistisches und pragmatisches

Wissen unterteilen lassen. Prosodisches Wissen bezieht sich auf Sprachmelodie, Sprachrhythmus und Betonungen. Diese erleichtern z. B. die Verarbeitung, Speicherung und den Erwerb von Sprache. Linguistisches Wissen ist die Voraussetzung für die Produktion und das Verstehen von Sprache. Dazu müssen phonologische Regularitäten (Lautstruktur der Sprache), Syntax (Regeln der Satzbildung), Morphologie (Regeln der Wortbildung) sowie lexikalisch-semantische Regeln (Bedeutungsstruktur des Wortschatzes, Satzbedeutungen) verfügbar sein. Für eine erfolgreiche Kommunikation ist es zudem notwendig, Sprache kontextangemessen und handlungsorientiert anzuwenden. In den Verstehensprozess fließen unter anderem Wissen über soziale Beziehungen, Rollenverständnis sowie Interpretationen und Reflexionen ein. Der Gebrauch von Sprache wird als pragmatische Komponente bezeichnet. Sozialinteraktive Beziehungen und Diskurswissen werden sowohl beim Verstehensprozess als auch beim produktiven Sprachgebrauch einbezogen. Die sprachkompetente Person kann Sprechhandlungen vollziehen, Konversation steuern und am Diskurs teilhaben sowie Zusammenhänge erschließen (Weinert & Grimm, 2012). Hierbei wirken alle sprachlichen Wissenssysteme zusammen.

Bei der Erfassung sprachlicher Kompetenzen können demnach ganz unterschiedliche Komponenten von Sprachwissen berücksichtigt werden. Damit ist nicht gemeint, dass einzelne Komponenten nur für die eine oder andere der funktionalen sprachlichen Teilfähigkeiten benötigt werden. Vielmehr sind für einen kontextangemessenen und erfolgreichen Umgang mit Sprache ausreichend Kenntnisse in allen Wissenssystemen notwendig. Dennoch lässt sich vermuten, dass je nach Verarbeitungsmodalität (Sprechen, Schreiben, Leseverstehen oder Hörverstehen) die Komponenten von Sprachwissen eine unterschiedliche Relevanz haben können. So ist beispielsweise pragmatisches Wissen besonders für das Sprechen in dem Sinne bedeutsam, dass die Sprechenden in einer Kommunikationssituation eine bestimmte Rolle einnehmen und ihre Sprache daran anpassen (Anrede, Wortwahl etc.). In der Zuhörsituation können insbesondere Wissen über Sprachmelodie, Rhythmus und Betonung Aufmerksamkeit lenken und Hinweise auf Bedeutungen geben (Buck, 2004).

Gemeinsamkeiten und Unterschiede zwischen den Teilfähigkeiten sprachlicher Kompetenz, insbesondere zwischen dem Lese- und Hörverstehen, lassen sich nicht nur an der Relevanz der Komponenten des Sprachwissens, sondern auch an theoretischen Modellen der zugrundeliegenden kognitiven Verarbeitungsprozesse sowie an entsprechenden empirischen Befunden festmachen.

3.2 Lese- und Hörverstehen als kognitiver Verarbeitungsprozess: Gemeinsamkeiten und Unterschiede

Theoretische Überlegungen zum Prozess des Hörverstehens werden in der kognitionspsychologischen Literatur häufig in Zusammenhang mit Theorien des Leseverstehens³ gestellt. Dementsprechend haben Kürschner und Schnotz (2008) ein integriertes Modell des Hör- und Leseverstehens entwickelt, das u. a. auf hierarchischen Modellen des Textverstehens beim Lesen (van Dijk & Kintsch, 1983; Kintsch & van Dijk, 1978) und auf der Mehrspeichertheorie des Gedächtnisses (Baddeley, 1986, 2000; Atkinson &

³ Wir verwenden den Begriff des Leseverstehens im Sinne des Verstehens von geschriebenen Texten.

Shiffrin, 1968) basiert. In dem Modell werden verschiedene Stufen auditiver und visueller Informationsverarbeitung unterschieden. Zunächst liegen die gesprochenen bzw. geschriebenen Texte als externe Repräsentation vor. Die externe Repräsentation ist kurzfristig im sensorischen Register verfügbar, bevor die Informationen unter Aufwendung von Aufmerksamkeit ins Arbeitsgedächtnis überführt werden. Hier entsteht eine mentale Repräsentation des Gehörten bzw. Gelesenen. In Abhängigkeit von der Verarbeitungstiefe entstehen daraufhin unterschiedlich elaborierte und stabile Gedächtnisspuren im Langzeitgedächtnis. Kürschner und Schnotz (2008) nehmen an, dass die ersten Stufen des Verarbeitungsprozesses (externe Repräsentation, sensorisches Register) eher modalitätsspezifisch, also spezifisch für das Zuhören bzw. Lesen, sind. Für die Erstellung einer mentalen Repräsentation und die Überführung der Informationen ins Langzeitgedächtnis nehmen die Autoren modalitätsunabhängige Prozesse an.

Kürschner und Schnotz (2008) integrieren in ihrem Modell des Hör- und Leseverstehens zwei grundlegende theoretische Ansätze zur Beschreibung der Informationsverarbeitungsprozesse, die dem Hör- und Leseverstehen zugrunde liegen: die monistische und die dualistische Position (Rost & Buch, 2010). Der dualistische Erklärungsansatz nimmt für das Lese- und das Hörverstehen jeweils modalitätsspezifische, also differente Verarbeitungsprozesse an, wonach Unterschiede zwischen dem Schriftsprach- und Hörverstehen auf Fähigkeiten und/oder Einschränkungen in unterschiedlichen zugrunde liegenden kognitiven Prozessen zurückgehen. Vertreter der monistischen Sichtweise hingegen nehmen Modalitätsunabhängigkeit an und erachten den Verstehensprozess beim Lesen und Zuhören als im Grundsatz identisch. Im Gegensatz zur dualistischen Position werden Leistungsunterschiede zwischen Lese- und Hörverstehen in der monistischen Erklärung nicht auf unterscheidbare Verständnisprozesse, sondern auf lesespezifische Faktoren, z. B. mangelnde Dekodierfähigkeit, zurückgeführt. Sind diese Defizite behoben, wird sich das Leseverstehen dem Hörverstehen angleichen, so die zugrunde liegende Annahme (Rost & Buch, 2010). Sowohl für die monistische als auch für die dualistische Position lassen sich empirische Belege finden.

Studien zum „simple view of reading“-Ansatz liegt eine monistische Sichtweise zugrunde. Dieser Ansatz geht davon aus, dass das Leseverstehen einerseits durch Worterkennen (Rekodier- und Dekodierfähigkeit) und andererseits durch Hörverstehen beeinflusst wird (Hoover & Gough, 1990). Marx und Jungmann (2000) untersuchten in einer Querschnittsstudie deutschsprachige Kinder der ersten bis vierten Jahrgangsstufe. Die Rekodier- und Dekodierfähigkeit, sowie das Hör- und Leseverstehen wurden mithilfe von „Knuspels Leseaufgaben“ (Marx, 1998) erfasst. Insgesamt zeigten die Schülerinnen und Schüler in den höheren Klassenstufen bessere Leistungen im Hör- und Leseverstehen als Schülerinnen und Schüler niedrigerer Klassenstufen. Erwartungsgemäß zeigten sich insbesondere bei den Leseanfängern bessere Leistungen im Hörverstehen als im Leseverstehen. Während die Korrelationen zwischen den Leistungen im Hör- und Leseverstehen in der ersten Klassenstufe gering ausfielen, war der Zusammenhang zwischen beiden Kompetenzen am Ende der vierten Klassenstufe hoch, was die Autoren dahingehend interpretieren, dass sich die Fähigkeiten des Hör- und Leseverstehens mit zunehmender Lesefähigkeit angleichen. Verhoeven und van Leeuwe (2008) untersuchten in einer Längsschnittstudie eine repräsentative Stichprobe niederländischer Grundschulinnen und Grundschüler. Auch hier zeigte sich im Sinne des „simple view of reading“-Ansatzes, dass Dekodierfähigkeit und Hörverstehen das Leseverstehen signifikant vorhersagen.

Insbesondere differente Gedächtnisbeanspruchung oder Einsatz metakognitiver Strategien können als unterschiedliche Einflussfaktoren beim Hörverstehen gegenüber dem Leseverstehen aufgrund der Darbietungsmodalität angesehen werden (Kürschner & Schnotz, 2008; Rost & Buch, 2010). Richter und Christmann (2002) verweisen darauf, dass hohe Korrelationen zwischen dem Hör- und Leseverstehen nicht einfach auf identische Verarbeitungsprozesse zurückgeführt werden können. Die Lesesituation stellt beispielsweise bei globalen Kohärenzbildungsprozessen deutlich höhere Anforderungen an den Leser, weil der Prozess sozial und physisch vom Kontext losgelöst ist (Richter & Christmann, 2002; Danks, 1980). Demgegenüber ergeben sich beim Hörverstehen Vorteile durch Betonungen und Sprechpausen, die zusätzliche Verständnishinweise bieten (Rost & Buch, 2010). Einen interessanten Befund liefern auch Böhme und Wehrich (2012), die beim IQB-Ländervergleich im Jahr 2011 Bundesländer identifizieren konnten, in denen die Kompetenzen der Schülerinnen und Schüler im Leseverstehen deutlich vom Hörverstehen abwichen. Im Bereich der Lesekompetenz erreichten die Schülerinnen und Schüler überdurchschnittliche Werte während beim Hörverstehen Kompetenzwerte erreicht wurden, die signifikant unter dem deutschen Mittelwert lagen. Die Autoren vermuten, dass (in den entsprechenden Bundesländern) im Bereich Lesen andere Lerngelegenheiten verfügbar sind, die die Lesekompetenzentwicklung befördern, aber nicht die Kompetenz im Hörverstehen (Böhme & Wehrich, 2012). Eine Analyse zur Struktur der sprachlichen Kompetenzen der Bildungsstandards zeigte, dass sich insbesondere das Hörverstehen von den anderen Kompetenzen im Fach Deutsch abgrenzen lässt und damit teilweise unabhängige sprachliche Teilfähigkeiten angenommen werden können (Böhme & Bremerich-Vos, 2012). Berninger und Abbott (2010) untersuchten in einer Studie nicht nur die Zusammenhänge zwischen Hör- und Leseverstehen, sondern zusätzlich Zusammenhänge mit Sprechen und Schreiben. Ziel war es zu überprüfen, ob die vier Teilbereiche von Sprache auf identischen kognitiven Verarbeitungsprozessen beruhen oder ob spezifische kognitive Verarbeitungsprozesse beteiligt sind. Zur Untersuchung dieser Hypothese wurden das Hörverstehen, das Leseverstehen, der mündliche und schriftliche Ausdruck bei Schülerinnen und Schülern der ersten, dritten, fünften und siebten Klasse in einem 2-Kohorten-Längsschnittdesign mithilfe standardisierter Testverfahren erfasst. Modellvergleiche im Rahmen von konfirmatorischen Faktorenanalysen zeigten, dass ein vier-faktorielles Modell sprachlicher Kompetenz eine bessere Modellgüte aufwies als ein ein-faktorielles Modell, was für die theoretische Annahme der Existenz von Teilfähigkeiten sprachlicher Kompetenz spricht. Das Hörverstehen klärte außerdem über alle Altersgruppen und Kohorten hinweg bei der Vorhersage der Lesekompetenz durch Hörverstehen, schriftlichen und mündlichen Ausdruck einen unigen Varianzanteil auf. Ähnlich wie das oben beschriebene Modell des Hör- und Leseverstehens von Kürschner und Schnotz (2008) spricht die Studie von Berninger und Abbott (2010) sowohl für modalitätsspezifische als auch für geteilte Informationsverarbeitungsprozesse beim Lesen, Zuhören, Sprechen und Schreiben, die je nach Verarbeitungsmodalität (Hören oder Lesen) zu verschiedenen mentalen Repräsentationen führen und sich in unterschiedlichen Lernleistungen niederschlagen können (Kürschner & Schnotz, 2008). Bei der Erfassung und Untersuchung sprachlicher Kompetenz ist es folglich wichtig, den Modus der Informationsverarbeitung zu berücksichtigen.

Als rezeptiv-auditive Fähigkeit ist Hörverstehen an gesprochene Sprache gebunden. Gesprochene Sprache weist im Vergleich zur Schriftsprache einige Besonderheiten auf, auf die im folgenden Abschnitt eingegangen wird.

3.3 Merkmale gesprochener Sprache

Gesprochene Sprache wird in der persönlichen menschlichen Interaktion und Kommunikation realisiert, aber auch über Medien wie das Radio, Podcasts oder Hörbücher vermittelt. Im Vergleich zu schriftlichen Texten, müssen Zuhörer gesprochene Sprache in der Regel in Echtzeit verarbeiten. Dies stellt hohe Anforderungen an das Arbeitsgedächtnis. Wie die gesprochene Sprache linguistisch gestaltet ist, ist u. a. davon abhängig, ob sie spontan im Alltag angewendet wird oder ob sie im Vorfeld einer sprachlichen Situation gezielt vorbereitet wird. Vor diesem Hintergrund ist ein wichtiges Unterscheidungsmerkmal gesprochener Sprache ihre Mündlichkeit bzw. Schriftlichkeit (Buck, 2004; Dirven, 1984; Rubin, Hafer & Arata, 2000). Nach Dirven (1984) lässt sich gesprochene Sprache auf einem Kontinuum der Spontanität mit den Polen „spontan“ und „nicht spontan“ abtragen (vgl. Abbildung 1). Im Alltag ist Sprache von einer hohen Spontanität geprägt, d. h. Sprache wird ohne gezielte Vorbereitung produziert. Spontan gesprochene Sprache ist z. B. durch Auslassungen und Satzabbrüche, Wiederholungen und Korrekturen, Pausen und Häsitationen sowie eine geringe semantische Variation gekennzeichnet. Diese Art von Sprache entspricht dem Konzept der Mündlichkeit (oral-based discourse). Auf der anderen Seite des Kontinuums bildet sich nicht-spontan gesprochene Sprache ab. Diese Art von Sprache ist häufig vorbereitet und z. B. vorab schriftlich verfasst worden. Im Moment der Darbietung wird sie z. B. abgelesen oder stark am schriftlichen Text orientiert vorgebracht. Sprache von vorbereiteten, schriftlichen Texten ist größtenteils fehlerfrei, enthält keine Auslassungen oder Häsitationen und ist dialektfrei. Diese Beschreibung gesprochener Sprache entspricht dem Konzept der Schriftlichkeit (literate-based discourse). Der Zuhörer muss sich der jeweiligen Hörsituation anpassen, wenn er das Gesagte verstehen will. Rubin et al. (2000) gehen davon aus, dass konzeptionell eher mündliche Texte besser verstanden werden, wenn sie mündlich präsentiert und konzeptionell eher schriftliche Texte besser verstanden werden, wenn diese schriftlich präsentiert werden. In ihrer Studie, in der die Teilnehmerinnen und Teilnehmer konzeptionell eher mündliche und schriftliche Texte sowohl lesen als auch hören mussten, konnten sie diese Hypothesen jedoch nicht bestätigen.



Abbildung 1: Kontinuum gesprochener Sprache (vgl. hierzu Dirven, 1984)

Aus bildungswissenschaftlicher Perspektive ist die Unterscheidung zwischen konzeptionell eher mündlichen und konzeptionell eher schriftlichen Texten aus einem weiteren Grund bedeutsam. Nach Cummins (2008) unterscheiden sich Sprachfähigkeiten zur Verständigung in Alltagssituationen (basic interpersonal communicative skills = BICS) von kognitiven bildungsbezogenen Sprachfähigkeiten (cognitive academic language proficiency = CALP), die dem Autor folgend insbesondere für Bildungserfolge ausschlaggebend sind (Cummins, 2008). Bildungssprache – also Fähigkeiten auf der Ebene von CALP – unterscheidet sich von Alltagssprache in lexikalischer, morphologischer und syntaktischer Hinsicht (Gogolin &

Lange, 2010; Heppt, Dragon, Berendes, Stanat & Weinert, 2012; Berendes, Dragon, Weinert, Heppt & Stanat, 2013). Sie ist durch komplexere grammatikalische Satzstrukturen und teilweise durch einen spezifisch verwendeten Wortschatz gekennzeichnet. Bildungssprache ist „ein bestimmter Ausschnitt sprachlicher Kompetenz, [die sich] auch dann, wenn sie im Mündlichen vorkommt, an den Regeln des Schriftsprachgebrauchs orientiert“ (Gogolin & Lange, 2010, S. 111). Hinzu kommt, dass Bildungssprache häufig losgelöst von sozialen und situativen Kontexten verwendet wird und damit andere Fähigkeiten (nämlich CALP) in der pragmatischen Sprachkomponente erfordert (Berendes et al. 2013; Cummins, 2008). Solche kontext-reduzierten Interaktionen finden sich häufig in formalen Bildungskontexten, wie es beispielsweise in typischen Unterrichtssituationen der Fall ist (Bailey & Butler, 2003; Heppt et al., 2012).

Zusammenfassend lässt sich festhalten, dass Hörverstehen aus funktionaler Perspektive eine Teilfähigkeit von Sprachkompetenz ist, neben den Fähigkeiten Lesen, Schreiben und Sprechen, die separat erfasst werden sollte. Es wird angenommen, dass beim Hörverstehen zumindest teilweise differente Verarbeitungsprozesse erfolgen sowie unterschiedliche Komponenten von Sprachwissen benötigt werden. Der auditive Inhalt kann je nach Zuhörsituation eher alltagssprachlich oder bildungssprachlich geprägt sein. Während Alltagssprache in soziale und situative Kontexte eingebettet ist und sich an den Regeln der Mündlichkeit orientiert, sind die eher kontext-reduzierten und kognitiv anspruchsvollen Inhalte von Bildungssprache an den schriftlichen Regeln des Sprachgebrauchs angelehnt. Bei der Erfassung von Hörverstehen sollten diese besonderen Charakteristika berücksichtigt werden.

4. Einordnung des Hörverstehens in das Konzept zur Erfassung von Sprachkompetenz im Nationalen Bildungspanel

Grundsätzlich liegt den Kompetenzmessungen im Nationalen Bildungspanel (NEPS) ein funktionales Literacy-Konzept zugrunde, nach dem funktionale, durch Bildung beeinflussbare Kompetenzen mit (quasi)authentischen Aufgaben erfasst werden. So sind die konstruierten Aufgaben an Situationen orientiert, wie sie im Alltag vorkommen können und haben damit einen konkreten praktischen Bezug zur Lebensbewältigung. Das NEPS Design sieht mehrere parallele Längsschnittuntersuchungen vor, in denen bildungsrelevante Kompetenzen (weitgehend) kohärent in ihrer Entwicklung über die Lebensspanne erfasst werden sollen. Dies gilt im Grundsatz auch für die im NEPS erfassten Sprachkompetenzen. So basiert die Erfassung von Lesekompetenz auf einem funktional-integrativen Ansatz und folgt einer Literacy-Konzeption (vgl. hierzu Gehrler, Zimmermann, Artelt & Weinert, 2013). Lesekompetenz ist unbestritten eine der wichtigsten Kompetenzen für die Bewältigung von Alltagssituationen, für Bildungs- und Berufserfolge. Aus bildungswissenschaftlicher Sicht sind sowohl rezeptive als auch produktive Teilfähigkeiten von Sprachkompetenz (funktional-integrativer Ansatz) wichtig, da – wie in Abschnitt 3 bereits erläutert – die Teilfähigkeiten miteinander interagieren. In Large Scale Studien wird, und das trifft auch für das NEPS zu, häufig auf die rezeptiven Fähigkeiten fokussiert, die mit standardisierten Testverfahren reliabel und objektiv erfasst werden können und dabei vergleichsweise einfach zu administrieren sind (Jude & Klieme, 2007). Als alleiniger Indikator für die Erfassung von Sprachkompetenz ist Lesekompetenz in einer Large Scale Studie wie NEPS jedoch nicht ausreichend (Berendes et al., 2013). Daher wird Hörverstehen als weiteres rezeptives Sprachmaß erfasst.

Im Folgenden wird ein kurzer Überblick über die Konzeption der Erfassung von Lesekompetenz im NEPS gegeben, um anschließend auf Hörverstehen als weiteren Indikator für Sprachkompetenz im NEPS einzugehen.

4.1 Rahmenkonzept zur Erfassung der Lesekompetenz im NEPS

Zur Erfassung der Lesekompetenz werden im NEPS Lesetests mit quasiauthentischen Texten eingesetzt, die systematisch fünf verschiedene Textsorten bedienen. Textsorten folgen hinsichtlich ihres Aufbaus und ihrer Struktur bestimmten Regeln. Es wird angenommen, dass Wissen über Struktur und Aufbau eines Textes die Rezeption vereinfacht, indem bereits vorhandenes Wissen aktiviert wird (Artelt et al., 2005). Anders formuliert wird die Kommunikationsabsicht eines Textes mit bestimmten, konventionell geltenden und verbindlich festgelegten Mitteln ausgedrückt, z. B. indem Sachtexte vorwiegend Informationen vermitteln oder literarische Texte eher unterhalten sollen (Brinker & Ausborn-Brinker, 2010). Diese kommunikative Intention (des Autors) wird mit der Textsorte transportiert und beeinflusst den Prozess der Bildung einer mentalen Repräsentation (Schnotz & Dutke, 2004). Textsortenwissen und Kenntnisse über die Intention des Autors erzeugen beim Leser Erwartungen, die das Textverständnis erhöhen (Artelt et al., 2005). Um die Lesekompetenz breit und kohärent über die Lebensspanne zu messen, spezifiziert die Rahmenkonzeption zur Erfassung der Lesekompetenz im NEPS fünf Textsorten: (1) Texte mit kommentierender Funktion, (2) Sachtexte, (3) Anleitungs- und Gebrauchstexte, (4) literarisch-ästhetische Texte und (5) Anzeige- und Werbetexte. Je nach Alter der Testpersonen variieren die Texte hinsichtlich Schwierigkeit, Länge und Thema während die Textsorten identisch bleiben. Außerdem unterscheidet die Rahmenkonzeption zur Erfassung der Lesekompetenz im NEPS – in Anlehnung an PISA – drei differente Verstehensanforderungen, um verschiedenen Lesesituationen Rechnung zu tragen. Die Lesetests enthalten Aufgaben, die (a) nach Informationen im Text fragen, (b) textbezogenes Schlussfolgern erfordern und (c) Reflektieren und Bewerten einschließen. Die kohärente Messung von Lesekompetenz beginnt im NEPS in Klassenstufe 4 (wobei erste Fähigkeiten des Lesens bereits in der Grundschule erfasst werden) und wird im zweijährigen Rhythmus wiederholt getestet. Das Rahmenkonzept zur Erfassung der Lesekompetenz ist im Artikel von Gehrler, Zimmermann, Artelt und Weinert (2013) detailliert beschrieben. Zusätzlich zur Lesekompetenz wird ein Lesegeschwindigkeitstest mindestens einmal in jeder Kohorte eingesetzt.

4.2 Weitere Sprachindikatoren im NEPS

Die Berücksichtigung verschiedener Teilfähigkeiten der Sprachkompetenz in der Mehrheitssprache, speziell die getrennte Erfassung von Lese- und Hörverstehen, ist für das NEPS auch deshalb bedeutsam, weil sich immer wieder Leistungsunterschiede in den sprachlichen Fähigkeiten, insbesondere beim Leseverstehen, z. B. zwischen Kindern mit und ohne Migrationshintergrund zeigen (z. B. Klieme et al., 2010). Diese Leistungsunterschiede gelten als besonders relevant für soziale Disparitäten in Bildungskarrieren. Die Untersuchung des Erwerbs der Mehrheitssprache – sowohl bezogen auf Alltagssprachliches als auch auf schul- bzw. bildungssprachliches Hörverstehens sowie mit Blick auf Lesekompetenz – ist damit für eine Bildungsstudie besonders wichtig. Obwohl die Unterschiede geringer werden, scheint es Kindern mit Migrationshintergrund jedoch nicht immer zu gelingen, die Unterschiede vollständig aufzuholen.

Marx und Roick (2012) konnten die Unterschiede in der Sprachkompetenz im Deutschen zwischen Kindern mit und ohne Migrationshintergrund speziell für das Hörverstehen belegen. In ihrer Studie wiesen die Jugendlichen mit Migrationshintergrund schwächere Leistungen im Hörverstehen, im Wortschatz und in den morphosyntaktischen Kenntnissen auf als Jugendliche mit deutscher Herkunftssprache. Allerdings verschwanden die Unterschiede in den morphosyntaktischen Kenntnissen zwischen den untersuchten Sprachgruppen wenn der sozioökonomische Status berücksichtigt wurde, anders formuliert wiesen die Jugendlichen mit gleichem sozialen Hintergrund – unabhängig von ihrer Herkunftssprache – ähnliche morphosyntaktische Kenntnisse auf. Bei allen Jugendlichen mit Schwächen im Wortschatz und auf grammatikalischer Ebene reichten die Kenntnisse nicht aus, um den Text auf der Ebene der Bedeutungsgenerierung (Parsing) angemessen zu erschließen (Marx & Roick, 2012).

Geringere sprachliche Fähigkeiten weisen auch Schülerinnen und Schüler mit besonderem Förderbedarf im Bereich Lernen auf. Über die Fähigkeit dieser Schülergruppen, gesprochene Sprache zu verstehen, ist bisher wenig bekannt. Dabei gilt das Hörverstehen als wichtige Voraussetzung für den Erwerb der Schriftsprache (Rost & Buch, 2010). Im Hinblick auf Personen mit geringen Lesefähigkeiten wäre es interessant, mögliche Unterschiede zwischen dem Lese- und Hörverstehen zu untersuchen.

Zuletzt bleibt noch zu erwähnen, dass es Personengruppen gibt, die über (noch) keine oder sehr eingeschränkte Lesefähigkeiten verfügen. Das sind einerseits Kinder im Vorschulalter und andererseits erwachsene Personen (z. B. mit unzureichender Alphabetisierung), deren Sprachkompetenz über andere Indikatoren erfasst werden muss.

Im NEPS werden daher verschiedene Tests zum Hörverstehen auf Wort- und Satzebene eingesetzt, die – basierend auf komponentenorientierten Modellen – zur Erfassung einzelner Wissenskomponenten von Sprachkompetenz abzielen. Der rezeptive Wortschatz gilt als guter Indikator für Sprachkompetenz und als Prädiktor für die spätere Lesekompetenz (Weinert, 2010). Die Durchführung von Tests zum Hörverstehen auf Wortebene ist relativ einfach, in verschiedenen Altersstufen möglich und erfordert keine Lesefähigkeiten. Im NEPS werden in jeder Kohorte Varianten des Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1981, 2004) eingesetzt, die eine gute Anschlussfähigkeit an andere nationale und internationale Large Scale Assessments ermöglichen. Zusätzlich zum Hörverstehen auf Wortebene werden rezeptive Grammatikkenntnisse (Hörverstehen auf Satzebene) zu zwei Messzeitpunkten – im Kindergarten und in der ersten Klasse der Grundschule⁴ – als potenziell wichtige Prädiktoren späterer Lesekompetenz erfasst.

Darüber hinaus werden im NEPS weitere Sprachindikatoren (z. B. Aufgaben zur phonologischen Bewusstheit; Orthographietests) zu bestimmten Zeitpunkten gemessen, zu denen sie besonders praktisch und theoretisch relevant für den weiteren Bildungsverlauf sind (vgl. hierzu detailliert Berendes et al, 2013).

⁴ Das NEPS ist mit insgesamt 6 Kohorten gestartet, deren Zielpersonen zum Startzeitpunkt unterschiedlichen Alters waren – Neugeborene, Kindergartenkinder, Schulkinder der Klassen 5 und 9, Studierende sowie Erwachsene. Hörverstehen auf Satzebene wurde bisher zweimal in Startkohorte 2 eingesetzt.

Was derzeit noch aussteht, ist ein Test zur Erfassung des Hörverstehens auf Textebene, um das Hörverstehen in der Mehrheitssprache im Sinne des Literacy-Ansatzes und auf Basis von funktionalen Modellen zu erfassen und gleichzeitig eine Vergleichbarkeit zu im NEPS erfassten Lesekompetenzen zu ermöglichen. Abbildung 2 gibt einen Überblick zu den erfassten Indikatoren zum Hörverstehen.

	Hörverstehen		
Ebene	auf Wortebene	auf Satzebene	auf Textebene
Sprachkompetenzmodell	komponenten-orientiertes Sprachmaß		funktionales Sprachmaß
Messzeitpunkt	mindestens einmal je Kohorte, tlw. mit Wiederholungsmessung	zweimalig, im Kindergarten und Klassenstufe 1	vorgesehen für Klassenstufe 9

Abbildung 2: Erfassung von Hörverstehen im NEPS

In Klassenstufe 9 sieht das Design für die Kompetenztestung im NEPS die Testung zahlreicher Kompetenzen vor, da es sich hier um einen besonders wichtigen Zeitpunkt im Bildungsverlauf handelt. Zum einen ist es der letzte Zeitpunkt, zu dem (fast) alle Schülerinnen und Schüler noch in den Schulen erreichbar sind, und zum anderen können hier Kompetenzen kurz vor dem Übergang in den Ausbildungs- und Berufsmarkt sowie vor weiteren Schullaufbahnentscheidungen erhoben werden. Aus diesen Gründen soll in Klassenstufe 9 neben anderen Sprachindikatoren auch das Hörverstehen auf Textebene erfasst werden. Ziel ist es, in Klassenstufe 9 ein weiteres funktionales Sprachmaß einzusetzen. Im folgenden Abschnitt wird das Konzept zur Erfassung von Hörverstehen auf Textebene im NEPS beschrieben.

5. Konzept zur Erfassung von Hörverstehen auf Textebene im Nationalen Bildungspanel

Wie bereits gezeigt, erfolgt die Untersuchung von sprachlichen Kompetenzen in zahlreichen Large-Scale Studien mit einem Fokus auf Lesekompetenz, sodass zu diesem Bereich auf umfangreiche Vorüberlegungen und Erfahrungen zurückgegriffen werden kann. Neben der Lesekompetenz stellt jedoch auch das Hörverstehen einen zentralen Indikator von Sprachkompetenz dar, der im NEPS erfasst wird. Während Hörverstehen aus komponenten-orientierter Perspektive bereits häufig in Large-Scale Studien untersucht wird (z. B. Einsatz von Wortschatztests im NEPS), trifft dies für eine Messung auf Textebene und damit basierend auf einem funktional-integrativen Ansatz eher seltener zu. Eine Ausnahme bilden die Ländervergleiche zur Erfassung der Bildungsstandards in Deutschland, die Instrumente zur Messung von Hörverstehen auf Textebene implementierten (Böhme & Bremerich-Vos, 2012). In den Testverfahren, die im Rahmen der Ländervergleiche eingesetzt wurden, wurden die Fragen zu den Hörtexten schriftlich präsentiert. Zur erfolgreichen Bearbeitung des Hörverstehenstests war also nicht nur Hörverstehen sondern auch Lesekompetenz erforderlich.

Im NEPS ist es uns ein zentrales Anliegen, Hörverstehen auf Textebene unabhängig von Lesekompetenz zu messen, um spezifische Subgruppen wie beispielsweise Personen mit Migrationshintergrund oder Personen mit unzureichender Alphabetisierung aufgrund geringerer Lesekompetenz nicht zu benachteiligen und ein valides Instrument zur Erfassung des Hörverstehens zu implementieren. Da für dieses Vorhaben keine vorhandene Konzeption aus anderen Large-Scale Studien zur Verfügung steht, auf die zurückgegriffen werden kann, stellt die kohärente Messung des Hörverstehens auf Textebene über die gesamte Lebensspanne hinweg eine Herausforderung dar. Aus diesem Grund erfolgt die Entwicklung eines Tests für die Erfassung von Hörverstehen auf Textebene für die Klassenstufe 9, die eine zentrale Schnittstelle im Übergang zum Arbeitsmarkt bzw. zur höheren Sekundarbildung darstellt, so dass hier das Hörverstehen auf Textebene als ein zentraler Prädiktor für den beruflichen bzw. weiteren schulischen Erfolg untersucht werden soll.

Für die Entwicklung einer eigenen Konzeption unter Berücksichtigung der spezifischen Anforderungen wird Hörverstehen in Anlehnung an Buck (2004) als die Fähigkeit definiert, kürzere und längere Abschnitte realistisch gesprochener Sprache in Echtzeit zu verarbeiten, dem Hörtext die sprachlichen Informationen zu entnehmen, die darin eindeutig enthalten sind und Schlussfolgerungen zu ziehen, die eindeutig durch den Inhalt des Hörtextes impliziert werden. Anschließend an diese Auffassung und an die Literacy-Perspektive – analog zur Erfassung der Lesekompetenz im NEPS – wird Hörverstehen in alltäglichen Kontexten und damit möglichst authentisch untersucht. Im Unterschied zur Lesekompetenz, deren Texte in schriftlicher Form vorliegen, sollen mit der Erfassung von Hörverstehen speziell auch alltagssprachliche Fähigkeiten abgedeckt werden, sodass das Konzept der Mündlichkeit und konzeptuellen Schriftlichkeit von Texten gezielt berücksichtigt wird. Während Alltagssprache in soziale Kontexte eingebettet ist und sich eher an den Regeln der Mündlichkeit orientiert, findet sich die anspruchsvollere Bildungssprache oftmals eher in schriftlichen Hörtexten. Unsere Konzeption unterscheidet daher die verschiedenen Hörtexte auf dem Kontinuum der Spontanität gesprochener Sprache in konzeptionell eher mündliche (z. B. Diskussion) und konzeptionell eher schriftliche Hörtexte (z. B. Rede), wie in Abbildung 3 verdeutlicht.

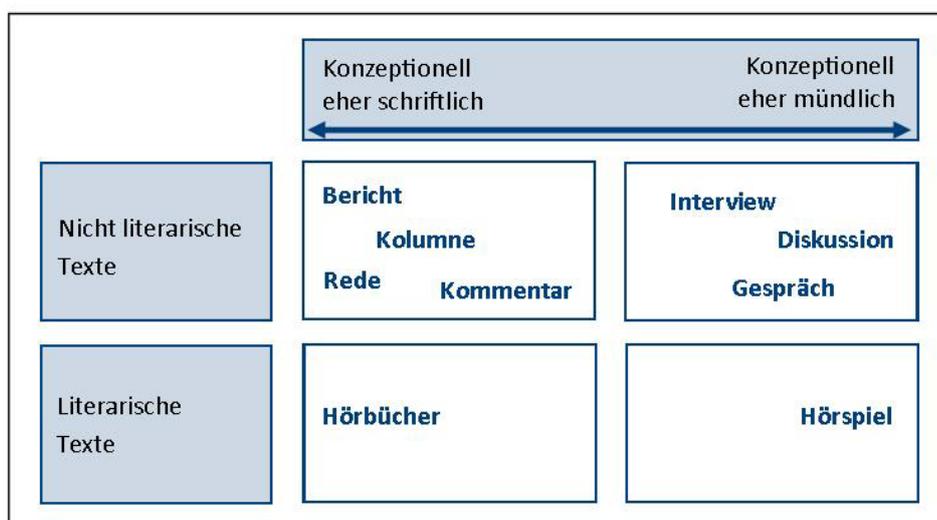


Abbildung 3: Konstrukt für die Auswahl von Hörtexten

Neben dieser Unterscheidung differenziert die Konzeption weiterhin zwischen verschiedenen Textsorten, wie beispielsweise zwischen Hörspielen (literarisch) und Interviews (nicht literarisch bzw. informierend) und deckt somit ein breites Spektrum an verschiedenen Hörtexten ab. Die funktionsbasierte Textkategorisierung – in Anlehnung an das Konzept zur Erfassung von Lesekompetenz im NEPS (Gehrer et al., 2013) – und die Unterscheidung zwischen spontaner und konzeptuell schriftlicher Sprache (Dirven, 1984) führt so zu einem Konstrukt, welches die Berücksichtigung der Vielfältigkeit gesprochener Sprache in Hörtexten breit gewährleisten kann.

Basierend auf der erläuterten theoretischen Rahmenkonzeption in Abbildung 3 soll im nachfolgenden Kapitel die Konstruktion eines Tests zur Erfassung von Hörverstehen auf Textebene auf Basis einer funktional-integrativen Sichtweise vorgestellt werden.

6. Entwicklung eines Tests zur Erfassung von Hörverstehen

Wie in anderen großen Bildungsstudien auch, ist der Einsatz von Testverfahren im NEPS an gewisse Rahmenbedingungen geknüpft, die bei der Konstruktion des Hörverstehentests berücksichtigt werden müssen. Dazu gehören die Festlegung der Testzeit je Domäne/Kompetenz auf maximal 30 Minuten und die Testung im Schulkontext in Gruppensettings von maximal 25 Schülerinnen und Schülern. Ziel war es, vor dem Hintergrund der theoretischen Grundlagen einen Test zur Erfassung von Hörverstehen im Deutschen auf Textebene für die Klassenstufe 9 zu entwickeln, der unter den gegebenen Rahmenbedingungen durchführbar ist. Im Folgenden wird auf die Testentwicklung, d. h. die Auswahl der Hörtexte, die Itementwicklung und das Antwortformat, näher eingegangen.

6.1 Auswahl von Hörtexten

Ziel war es, kürzere und längere Abschnitte gesprochener Sprache auf der Basis authentischer Hörtexte auszuwählen. Diese Hörtexte wurden aus im Internet zugänglichen Hörtexten ausgewählt (z. B. Ausschnitte aus Radiosendungen, Podcasts, Hörspielen und Hörbüchern). Die Auswahl der Hörtexte orientiert sich zum einen am Konzept der Mündlichkeit und Schriftlichkeit von Texten sowie an der Unterscheidung verschiedener Textsorten. Diese Konzeption folgt der Annahme, dass verschiedenartige Texte unterschiedliche Anforderungen an die Zuhörer stellen und gewährleistet eine Berücksichtigung der Vielfältigkeit gesprochener Sprache im Alltag (vgl. hierzu Abschnitte 3.3 und 4.1).

Die Auswahl der Hörtexte orientierte sich darüber hinaus am Thema und Inhalt der Texte, an ihrer Länge und Verständlichkeit. Um die 14- und 15-jährigen Schülerinnen und Schüler zum Zuhören zu motivieren, wurden thematisch abwechslungsreiche und für die Altersgruppe interessante bzw. relevante Hörtexte ausgewählt. Dies betrifft beispielsweise Inhalte wie Schule, Beruf oder zwischenmenschliche Beziehungen. Demgegenüber wurde auf Themen wie Religion, Politik, Wirtschaft oder Finanzen verzichtet, die oftmals eher mäßiges Interesse bei Jugendlichen hervorrufen (Shell-Jugendstudie, 2010). Speziell wurde darauf geachtet, dass die Themen und Fragen nicht einfach auf bereits verfügbares Vorwissen Bezug nehmen. Ein wichtiges Kriterium für die Auswahl stellte daher der Ausschluss von Hörtexten dar, die Allgemeinwissen vermitteln, da vermutet werden kann, dass Fragen zu dieser Art von Hörtexten von einem Teil der Schüler unabhängig vom Verständnis des Hörtextes mit entsprechendem Vorwissen beantwortet werden können.

Ein weiteres Auswahlkriterium für Hörtexte bezog sich auf ihre Länge. In der vorgegebenen 30-minütigen Testzeit sollten verschiedene Arten von Hörtexten vorgegeben werden können. Gleichzeitig sollten die Schülerinnen und Schüler den Texten gut zuhören können und ihre Aufmerksamkeitsspanne nicht überschritten werden. Die Dauer der Hörtexte wurde deshalb auf minimal eineinhalb und maximal sechs Minuten festgelegt. Um die grundsätzliche Verständlichkeit der Hörtexte für alle Schülerinnen und Schüler sicherzustellen, wurden nur hochdeutsche Hörtexte ohne fremdsprachliche oder dialektale Einfärbungen ausgewählt. Auch wurden Texte mit musikhinterlegten Passagen oder anderen lauten und dauerhaften Hintergrundgeräuschen ausgeschlossen. Schließlich war die Unabhängigkeit der Hörtexte vom Tagesgeschehen ein Kriterium für die Auswahl. So wurden z. B. Nachrichtentexte ausgeschlossen, die über aktuelle politische Geschehnisse berichten.

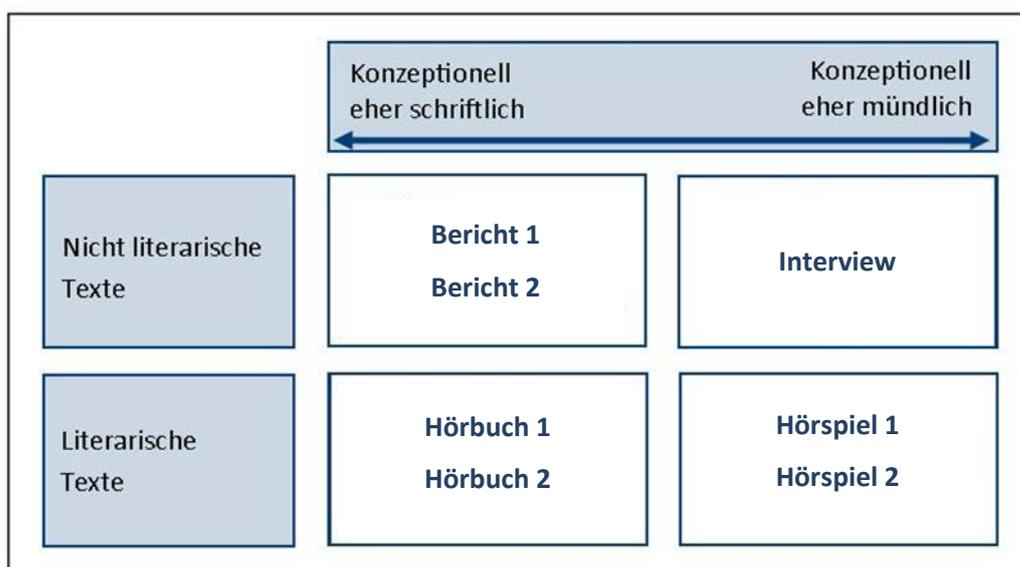


Abbildung 4: Spektrum ausgewählter Hörtexte

Für die Erstellung einer ersten Testversion wurden insgesamt sieben Hörtexte ausgewählt, die sich den Kategorien konzeptionell eher schriftlich und konzeptionell eher mündlich zuordnen lassen sowie literarischer bzw. nicht-literarischer (informierender) Art sind. Die folgende Abbildung stellt die sieben Texte entsprechend ihrer Kategorienzugehörigkeit dar (vgl. Abbildung 4).

Der kürzeste Text hat eine Dauer von 1:59 Minuten und der längste Text von 5:39 Minuten. Im literarischen Bereich wurde z. B. eine Geschichte ausgewählt, in der es um den Sinn des Lebens geht, und eine spannende Kriminalgeschichte. Zu den nicht-literarischen Hörtexten gehören z. B. ein Bericht über einen ungewöhnlichen Rekord und ein Interview über die Aufgaben in einem spannenden Beruf.

6.2 Itementwicklung und Antwortformat

Zu allen ausgewählten Hörtexten wurden Fragen entwickelt, die das Verständnis gesprochener Sprache erfassen sollen. In Anlehnung an die Literatur zu Prozessen des Textverstehens werden drei unterschiedliche kognitive Anforderungsniveaus unterschieden (vgl. hierzu Richter & Christmann, 2002; Buck, 2004). Fragen des Typus 1 erfordern das *Wörtliche Verstehen expliziter Äußerungen*. Diese Fragen zielen auf das Verstehen kürzerer

Äußerungen auf einer wörtlichen Ebene ab. Bei geschlossenen Antwortformaten kann die Textinformation in den Antwortoptionen paraphrasiert wiedergegeben sein. Fragen des Typus 2 verlangen *das Herstellen von Inferenzen*. Sie fordern den Zuhörer dazu auf, auf Basis des Hörtextes Schlussfolgerungen zu ziehen. Dazu gehören auch Fragen, die sich auf die Authentizität von Texten beziehen oder Stimmungen und Haltungen aus dem Text (zum Beispiel aus dem Tonfall) erfragen. Fragen vom Typus 3 zielen auf *das Verstehen impliziter Bedeutungen und Aussagen* ab. Die Beantwortung ist nur möglich, wenn Informationen im Text über verschiedene Textstellen hinweg kombiniert und der Textinhalt reflektiert und bewertet werden kann. Hierunter fallen beispielsweise Fragen, die auf die Wirkung des Textes auf den Zuhörer oder auf die Verfasserintention abzielen. Fragen aller drei Typen wurden grundsätzlich so konstruiert, dass sie sich auf wesentliche Aussagen des Hörtextes beziehen. Außerdem werden die Informationen aus dem Hörtext in der Reihenfolge abgefragt, in der sie im Hörtext genannt werden. Textübergreifende Fragen schließen sich an Fragen mit direktem Textstellenbezug an. Eine bereits abgefragte Information wird kein zweites Mal abgefragt. Für eine erste Testversion wurden zu jedem ausgewählten Hörtext sechs bis zwölf Fragen nach diesen Prinzipien entwickelt, davon jeweils mindestens zwei von jedem Fragetypus (vgl. Tabelle 1).

Alle Aufgaben der vorliegenden Testversion enthalten ein geschlossenes Antwortformat. Die Aufgabe der Schülerinnen und Schüler besteht darin, Inhalte des Hörtextes korrekt (wieder) zu erkennen. Auf ein offenes Antwortformat (freie Produktion) wurde verzichtet, weil offene Antworten in Gruppensettings die Fähigkeit zum schriftlichen Ausdruck erfordern. Ziel war es jedoch, Hörverstehen möglichst unabhängig von schriftsprachlichen Fähigkeiten zu messen. Des Weiteren haben wir uns für den Einsatz komplexer Multiple Choice-Aufgaben entschieden. Eine Aufgabe besteht aus einer Frage zum Hörtext und vier Optionen. Für jede Antwortoption sollen die Schülerinnen und Schüler entscheiden, ob die Antwort stimmt oder nicht stimmt. Die Antwortmöglichkeiten wurden so entwickelt, dass bei mindestens einer und bei maximal drei Antwortoptionen „stimmt“ zutrifft. Für die Verwendung komplexer Multiple-Choice-Aufgaben haben wir uns aus zwei Gründen entschieden. Zum einen sollten sowohl die Hörtexte als auch die dazugehörigen Aufgaben auditiv präsentiert werden.

Tabelle 1: Merkmale der ausgewählten Hörtexte

Text	Dauer [min]	Textsorte	Textgestaltung	Anzahl Fragen	Typ 1	Typ 2	Typ 3
1	03:12	Bericht	konzeptionell schriftlich	6	2	2	2
2	01:59	Hörspiel	konzeptionell mündlich	12	5	4	3
3	03:33	Hörbuch	konzeptionell schriftlich	9	4	2	3
4	04:37	Hörbuch	konzeptionell schriftlich	11	5	3	3
5	03:57	Hörspiel	konzeptionell mündlich	9	3	3	2
6	03:02	Bericht	konzeptionell schriftlich	8	2	3	3
7	05:39	Interview	konzeptionell mündlich	9	4	3	2

In Abbildung 5 ist dargestellt, welche Aspekte der Aufgaben den Schülerinnen und Schülern rein auditiv (links) und welche Aspekte im Antwortheft präsentiert wurden (rechts). Wie zu erkennen ist, wurden nur wenige Informationen schriftlich im Testheft abgedruckt. Um die Anforderungen an die Aufmerksamkeit der Schülerinnen und Schüler zu minimieren, wurde deshalb nur ein Antwortformat verwendet. Das komplexe Multiple-Choice-Format hat außerdem den Vorteil, dass nach dem Hören einer Antwortoption direkt eine Antwort im Testheft gegeben werden kann und nicht mehrere Möglichkeiten über einen längeren Zeitraum erinnert werden müssen.

Zum anderen hat sich das einfache Multiple-Choice-Format im Kontext von Gruppentestungen in Pretests als problematisch herausgestellt, da es hier dazu kommen kann, dass mehrere Personen parallel und für alle sichtbar eine bestimmte Antwort wählen. Zu Beginn der Testdurchführung erhalten die Teilnehmerinnen und Teilnehmer eine Instruktion, die das Antwortformat und das korrekte Ausfüllen verdeutlicht. Es schließt sich eine Übungsaufgabe für die Schülerinnen und Schüler mit dem Ziel an, den Testablauf zu proben und die Testpersonen an das Format zu gewöhnen. Für die Testdurchführung ist vorgesehen, dass die Testpersonen einen Hörtext einmalig präsentiert bekommen. Im Anschluss wird die erste Aufgabe zum Hörtext (Frage plus drei bis vier Antwortoptionen) auditiv präsentiert. Parallel dazu können die Teilnehmerinnen und Teilnehmer ihre Antworten im Testheft vergeben, indem sie „stimmt“ oder „stimmt nicht“ ankreuzen. Dann werden Frage und Antwortalternativen ein zweites Mal vorgegeben und die Teilnehmerinnen und Teilnehmer können ihre Antworten kontrollieren und ggf. korrigieren. Daraufhin folgen die weiteren Fragen zum Hörtext nach demselben Schema.



Auditiv präsentierte Frage



Antwortformat im Testheft

Wie ist Jans erste Reaktion, als er den Anruf entgegennimmt?

- a) Er gibt an, keine Zeit zu haben.
- b) Er sagt, der Anrufer ist falsch verbunden.
- c) Er gibt an, Nele Lös nicht zu kennen.
- d) Er sagt, er kann gerade nicht telefonieren.

	Stimmt	Stimmt nicht
a)	<input type="checkbox"/>	<input type="checkbox"/>
b)	<input type="checkbox"/>	<input type="checkbox"/>
c)	<input type="checkbox"/>	<input type="checkbox"/>
d)	<input type="checkbox"/>	<input type="checkbox"/>

Abbildung 5: Auditiv präsentierte Item und dazugehöriges abgedrucktes Antwortformat.

Um die kognitiven Prozesse und Strategien abschätzen zu können, die während der Testung angewandt werden und um Hinweise auf das Verständnis der Fragen zu erhalten, wurden

vorab kognitive Interviews durchgeführt. Hierzu wurden die Aufgaben vier Schülern⁵ einer neunten Klasse vorgelegt, die aus dem NEPS-Expertenpool rekrutiert wurden und sich zur Durchführung des Tests mit anschließender Bewertung der Aufgaben bereit erklärt hatten. Der Ablauf im Pretest erfolgte analog der vorgesehenen Testung. Einer der beiden Testleiterinnen begrüßte den Schüler und gab eine kurze Einführung. Anschließend wurde die vorbereitete CD abgespielt, auf der zunächst ein Hörtext und daran anschließend die dazugehörigen Fragen präsentiert wurden. Entsprechend des vorgesehenen Testablaufs bearbeitete der Schüler die Aufgaben im Testheft. Eine zweite Testleiterin protokollierte während der Testdurchführung den Ablauf der Testsitzung und beobachtete Auffälligkeiten im Verhalten des Schülers (z. B. langes Zögern bei bestimmten Fragen). Im Anschluss an die Testdurchführung fand eine Phase der Rückmeldung durch den Schüler statt, bei dem erste Mängel oder Schwierigkeiten gemeldet werden konnten, die auf dem Protokollbogen vermerkt wurden. Außerdem wurden die korrekten Lösungen der Aufgaben mit dem jeweiligen Schüler besprochen. Diese Phase mündete direkt in das kognitive Interview, welches mittels eines strukturierten Leitfadens durchgeführt wurde. Dieser sah den Einsatz verschiedener Fragetechniken mithilfe des Probings (Nachfragen) vor, wie etwa Fragen zum Verständnis, zur Wahl von Antwortkategorien, zum Besitz relevanter Informationen sowie unspezifische Nachfragen (vgl. hierzu Prüfer & Rexroth, 2005).

Weiterhin wurden die Schüler ganz allgemein um Einschätzung zur Testdurchführung gebeten. Im Hinblick auf den Testablauf zeigte sich, dass das Antwortformat aufgrund der Abweichung von bekannten Hörverstehenstests aus der Schule (z. B. im Fach Englisch) durch die rein auditive Präsentation zunächst gewöhnungsbedürftig erschien, aber durch die Instruktion mit dem Übungsbeispiel gut möglich war. Das zweimalige Vorspielen der Fragen hat sich insbesondere bei schwierigen Items als notwendig herausgestellt. Die Schüler gaben an, dass sie Antwortoptionen nach dem ersten Hören nicht mit Sicherheit beantworten könnten. Eine Antwort im Nachhinein zu korrigieren, erweist sich nur dann als möglich, wenn die Antwortoptionen ein zweites Mal präsentiert werden. Die Orientierung im Testheft erwies sich als problemlos. Insgesamt hat sich die Durchführbarkeit bestätigt. Keiner der Teilnehmer hat den Test, der insgesamt zwei Stunden Zeit in Anspruch nahm, als belastend empfunden, vielmehr wurde hier direkt auf den Vorteil der zweimaligen Antwortpräsentation hingewiesen. Die kognitive Beanspruchung scheint demnach angemessen zu sein.

6.3 Merkmale zur Beschreibung der Schwierigkeit eines Hörverstehenstests

Die Schwierigkeit eines Tests zum Hörverstehen ergibt sich einerseits aus der Schwierigkeit des Hörtextes. Auf inhaltlicher Ebene spielen beispielsweise die Anzahl der im Text enthaltenen Themen, die Informationsdichte und die Spezifität der Themen eine Rolle dafür, dass der Text eher einfach oder eher schwierig zu verstehen ist. Auf textlicher Ebene haben Merkmale wie Satzstruktur, Satzlänge und verwendeter Wortschatz aber auch Gliederungsmerkmale und Erzählreihenfolge einen Einfluss auf die Textschwierigkeit. Neben den inhaltlichen und textlichen Kriterien beeinflussen akustische Merkmale wie Hintergrundgeräusche, Artikulation und Intonation den Verstehensprozess (Grotjahn, 2001; vgl. auch Buck, 2004). Werden authentische Hörtexte wie in unserem Test verwendet, lässt sich die

⁵ Für die kognitiven Interviews wurden Testpersonen aus dem NEPS Expertenpool rekrutiert. Es haben sich lediglich Jungen für die Durchführung bereit erklärt.

Schwierigkeit der Texte anhand dieser Kriterien bestimmen, jedoch nicht steuern, sofern keine Veränderungen an den Hörtexten vorgenommen werden. Der Einsatz von Schwierigkeitsindizes – wie bei Lesetests oft üblich – ist nicht sinnvoll, da mit diesen Indizes die Lesbarkeit eines Textes ermittelt wird und keine Aussagen zum Verstehen eines Hörtextes getroffen werden können. Im vorliegenden Test zur Erfassung des Hörverstehens sprechen mehrere Gründe dafür, die Schwierigkeit des Hörtextes nicht zu verändern. Wie bereits erwähnt werden authentische Texte eingesetzt, deren Authentizität nicht durch Veränderungen zugunsten eines höheren oder niedrigeren Schwierigkeitsgrades aufgegeben werden sollte. Ein zentrales Anliegen bei der Testkonstruktion war zudem, dass jede Testperson grundsätzlich in der Lage sein soll, den Hörtext akustisch zu verstehen. Akustische Merkmale scheiden vor diesem Hintergrund gänzlich für die Schwierigkeitssteuerung aus. Auch die vermutlich hohe Beanspruchung des Arbeitsgedächtnisses beim Hörverstehen sollte in unserem Test nicht zusätzlich durch eine hohe Themenvielfalt innerhalb eines Textes belastet werden.

Die Schwierigkeit eines Hörverstehenstest ergibt sich andererseits aus der Schwierigkeit der Aufgaben. Grotjahn (2005; 2001) unterscheidet hier Variablen zu Iteminhalten und Variablen zur Itemformulierung. Wie in den Abschnitten 4.1 und 6.2 bereits herausgestellt, kann die Konstruktion von Variablen zu Iteminhalten auf unterschiedliche kognitive Anforderungen abzielen (hier: Wörtliches Verstehen expliziter Informationen, textbezogenes Schlussfolgern, Verstehen impliziter Bedeutungen und Äußerungen). Die Schwierigkeit einzelner Items ist jedoch nicht nur von den Iteminhalten, sondern auch von der Verknüpfung des Items mit dem Hörtext abhängig. Es ist denkbar, dass detailabfragende Items in einem Hörverstehenstest aufgrund der Gedächtnisbeanspruchung vergleichsweise schwieriger zu beantworten sind, da sich der Zuhörer an Details möglicherweise weniger gut erinnert. So ist auch denkbar, dass Fragen zum Gesamtzusammenhang in einem Hörtext gut beantwortbar sind und das Schlussfolgern erleichtern. Daher ist es aus unserer Sicht notwendig, Items zu allen kognitiven Anforderungen zu berücksichtigen, um Hörverstehen adäquat zu erfassen. Wir gehen jedoch nicht von einer Schwierigkeitshierarchie in den kognitiven Anforderungen aus.

Bei der Itemformulierung können sprachliche Merkmale wie Satzlänge und -bau, verwendeter Wortschatz, Konnexionen sowie Passiv- und Konjunktivformulierungen die Schwierigkeit von Items beeinflussen (vgl. Drosdowski, 1998; Hoffmann, 2010; Kessel & Reimann, 2010). So weisen beispielsweise Satzkonstruktionen aus alleinstehenden Hauptsätzen eine geringere Komplexität als Satzgefüge auf. Die Verwendung von häufig verwendeten und gängigen Verknüpfungen und Konnektoren (z. B. und, weil, oder) wirken verständnis erleichternd, während ungewöhnliche Konnexionen (z. B. dessen ungeachtet) die Komplexität steigern (Pasch, Brauße, Breindl & Waßner, 2003). Durch ihren weniger häufigen Gebrauch in der Alltagskommunikation gelten Konjunktivformulierungen (z. B. hätte, wäre, könnte) oder Passivkonstruktionen (z. B. wird gebildet, ist geschlossen) als verständniserschwerend (IDS, 2013; vgl. Helbig & Kempter, 1997). Die Formulierung der Aufgaben variierte entsprechend mit Blick auf die verschiedenen sprachlichen Mittel, um sowohl bildungs- als auch alltagssprachliche Formulierungen zu verwenden. Pro Aufgabe wurde auf die einheitliche Verwendung der sprachlichen Mittel geachtet. So sollten beispielsweise bei allen Antwortalternativen innerhalb einer Aufgabe ähnliche Satzkonstruktionen verwendet (z. B. Satzreihe, Satzgefüge, Satzstellung), und Formulierungen

auf ähnlicher sprachlicher Ebene vorgenommen werden (z. B. gebräuchliche vs. ungebräuchliche Konnektoren, Aktiv- vs. Passivformulierungen).

7. Fazit und Ausblick

Anhand von sprachwissenschaftlichen und sprachpsychologischen Grundlagen wurde ein theoretisches Rahmenkonzept zur Erfassung von Hörverstehen als wichtige Teilkomponente von Sprachkompetenz entwickelt. Das Konstrukt basiert auf der Literacy-Konzeption und einer funktional-integrativen Sichtweise von Sprachkompetenz und greift dabei auf das integrierte Modell des Hör- und Leseverstehens von Kürschner und Schnotz (2008) zurück, nach dem differente Verarbeitungsprozesse bei den rezeptiven Teilfähigkeiten von Sprachkompetenz (also zwischen Hör- und Leseverstehen) angenommen werden.

Demnach erscheint es wichtig, bei der Erfassung von Sprachkompetenz den Modus der Informationsvermittlung zu berücksichtigen und sowohl Lese- als auch Hörverstehen zur Messung von Sprachkompetenz zu erfassen. Die funktionale Sicht ermöglicht es, Sprachkompetenz einerseits als Ganzes zu verstehen und gleichzeitig einzelne Teilfähigkeiten zu erfassen, um Unterschiede in der Sprachkompetenz bei besonderen Gruppen, wie beispielsweise Personen mit Migrationshintergrund oder Personen mit geringen Lesefähigkeiten, aufzudecken und die Prüfung und Analyse von theoretischen Erklärungsansätzen zu ermöglichen. Zudem war es ein zentrales Anliegen, Hörverstehen möglichst ohne Einfluss anderer Teilfähigkeiten von Sprachkompetenz zu erfassen. Für das Konstrukt zur Erfassung von Hörverstehen ist es – wie in anderen Bildungsstudien auch – sinnvoll und notwendig, Hörtexte anhand ihrer Funktion, die der Text erfüllen soll, sowie anhand der verwendeten (alltags- oder bildungssprachlich orientierten) Sprache zu unterscheiden.

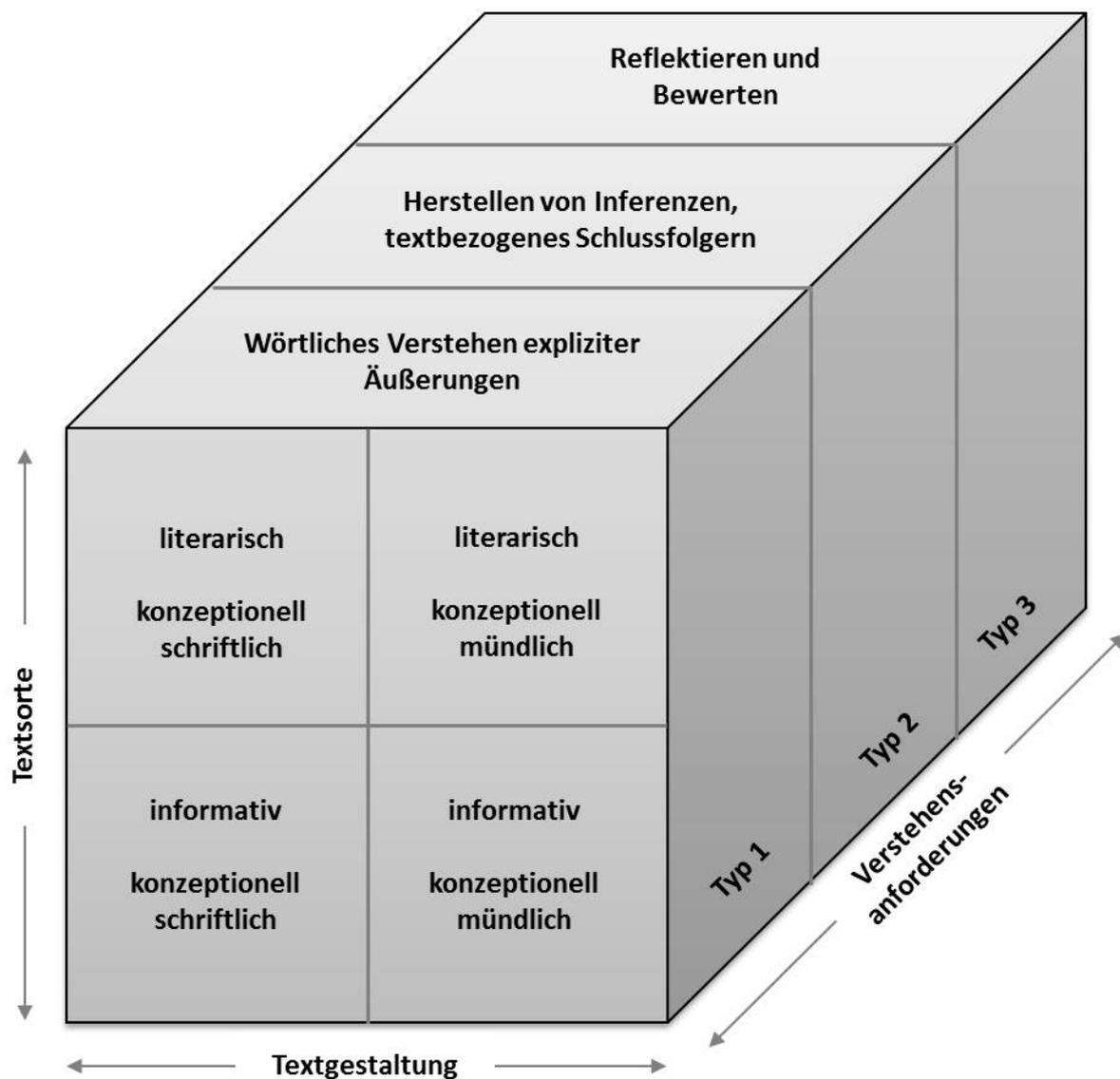


Abbildung 6: Rahmenkonzept zur Erfassung von Hörverstehen auf Textebene

So ist es möglich, eine Vielzahl an unterschiedlichen Hörtexten in einen Test zu integrieren und die Vielfalt gesprochener Sprache adäquat zu berücksichtigen. Weiterhin gehen in unser Konzept verschiedene kognitive Verstehensanforderungen ein, um unterschiedliche Anforderungen an verschiedene Hörsituationen abzudecken. Das folgende Modell fasst das Rahmenkonzept nochmal anschaulich zusammen (vgl. Abbildung 6).

Auf Basis dieses Rahmenkonzepts wurde eine erste Version eines Hörverstehenstests für Schülerinnen und Schüler der Klassenstufe 9 entwickelt. Die neunte Jahrgangsstufe gilt als wichtiger Übergang für weitere Bildungs- und Berufsentscheidungen, weshalb zu diesem Zeitpunkt eine Vielzahl von Kompetenzen im NEPS erhoben wird. Es ist das Ziel, den Hörverstehenstest erstmalig in einer NEPS-Haupterhebung in der neunten Klasse einzusetzen, der für die erste Jahreshälfte 2015 geplant ist (vgl. Abbildung 7). Die Phase der Testentwicklung erfolgte zwischen 2011 und 2012. Im Herbst 2012 wurde der Test erstmals in einer Vorstudie eingesetzt und daraufhin weitere Anpassungen vorgenommen. Im Frühjahr 2014 fand eine weitere Vorstudie statt, deren Ergebnisse derzeit analysiert werden.

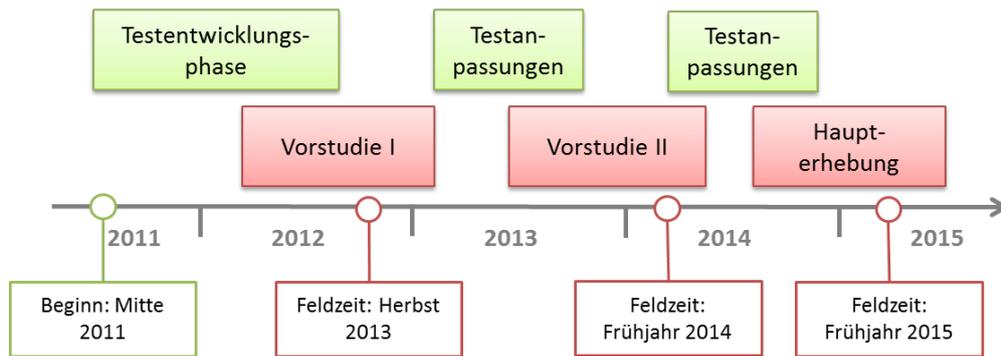


Abbildung 7: Zeitplan für die Testentwicklung und Studiendurchführung

Mit dem vorliegenden Test liegt ein weiteres Sprachmaß vor, welches die vorhandenen komponentenbezogenen Hörverstehenstests um einen funktionalen Indikator ergänzt; das zugrunde liegende Konzept eröffnet die Möglichkeit, Hörverstehen funktional auf Textebene über die Lebensspanne zu erfassen. Hierzu sind jedoch weitere Testentwicklungen für noch festzulegende Altersgruppen notwendig.

8. Literatur

- Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., Schneider, W., Stanat, P., Ostermeier, C., Schiefele, U., Valtin, R., Ring, K. & Saalbach, H. (2005). *Expertise – Förderung von Lesekompetenz (Bildungsreform Nr. 17)*. Berlin: Bundesministerium für Forschung und Bildung.
- Artelt, C., Drechsel, B., Bos, W. & Stubbe, T.C. (2008). Lesekompetenz in PISA und PIRLS/IGLU - Ein Vergleich. *Zeitschrift für Erziehungswissenschaft, 10*, 35-52.
- Artelt, C., Weinert, S. & Carstensen, C. (2013). Assessing competencies across the lifespan within the German National Educational Panel Study (NEPS) – Editorial. *Journal for Educational Research Online, 5*, 5-14.
- Atkinson, R.C. & Shiffrin, R.M. (1968). Human memory: A proposed system and its control processes. In: K.W. Spence & J.T. Spence (Hrsg.), *The psychology of learning and motivation: Advances in research and theory*. New York: Academy Press.
- Autorengruppe Bildungsberichterstattung (2012) (Hrsg.). *Bildung in Deutschland 2012. Ein indikatorengestützter Bericht mit einer Analyse zur kulturellen Bildung im Lebenslauf*. Bielefeld: Bertelsmann Verlag.
- Baddeley, A.D. (1986). *Working memory*. Oxford: Clarendon Press.
- Baddeley, A.D. (2000). *The episodic buffer: A new component of working memory? Trends in Cognitive Science, 4* (11), 417-423.
- Bailey, A.L. & Butler, F.A. (2003). An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document. CSE-Report 611. Los Angeles.
- Baldi, S., Kutner, M, Greenberg, E., Jin, Y., Baer, J., Moore, E., Dunleavy, E., Berlin, M., Mohadjer, L., Binzer, G., Krenzke, T., Hogan, J., Amsbary, M., Forsyth, B., Clark, L., Annis, T., Bernstein, J. & White, S. (2009). Technical Report and Data File User's Manual for the 2003 National Assessment of Adult Literacy (NCES 2009-476). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office. Verfügbar unter: <http://nces.ed.gov/pubs2009/2009476.pdf> [07.01.2014].
- Baumert, J., Klieme, E., Neubrand M., Prenzel, M., Schiefele, U. Schneider, W., Stanat. P., Tillmann, K-J. & Weiß, M. (2001). *PISA 2000, Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich*. Opladen: Leske und Budrich.
- Beck, B. & Klieme, E. (2007). (Hrsg.). *Sprachliche Kompetenzen - Konzepte und Messung*. DESI-Studie (Deutsch Englisch Schülerleistungen International). Weinheim: Beltz.
- Behrens, U. & Eriksson, B. (2009). Sprechen und Zuhören. In A. Bremerich-Vos, D. Granzer, U. Behrens & O. Köller (Hrsg.), *Bildungsstandards für die Grundschule: Deutsch konkret: Aufgabenbeispiele - Unterrichtsanregungen – Fortbildungsideen*. Berlin: Cornelsen Scriptor.

- Berendes, K., Weinert, S., Zimmermann, S. & Artelt, C. (2013). Assessing language indicators across the lifespan within the German National Educational Panel Study (NEPS). *Journal for Educational Research Online*, 2, 15-49.
- Berendes, K., Dragon, N., Weinert, S., Heppt, B. & Stanat, P. (2013). Hürde Bildungssprache? Eine Annäherung an das Konzept „Bildungssprache“ unter Einbezug aktueller empirischer Forschungsergebnisse. In A. Redder & S. Weinert (Hrsg.), *Sprachförderung und Sprachdiagnostik – interdisziplinäre Perspektiven*. Münster: Waxmann.
- Berninger, V.W. & Abbott, R.D. (2010). Listening comprehension, oral expression, reading comprehension, and written expression: Related yet unique language systems in grades 1, 3, 5, and 7. *Journal of Educational Psychology*, 102, 635-651.
- Böhme, K. & Bremerich-Vos, A. (2012). Beschreibung der im Fach Deutsch untersuchten Kompetenzen. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik, Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Böhme, K. & Wehrich, S. (2012). Der Ländervergleich im Fach Deutsch. In P. Stanat, H. A. Pant, K. Böhme & D. Richter (Hrsg.), *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik, Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K. & Valtin, R. (2007). *IGLU 2006, Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.M., Prenzel, M., Schwippert, K., Valtin, R., & Walther, G. (2004). *IGLU. Einige Länder der Bundesrepublik Deutschland im nationalen und internationalen Vergleich*. Münster: Waxmann.
- Bos, W., Lankes, E.-M., Prenzel, M., Schwippert, K., Valtin, R. & Walther, G. (2005). *IGLU, Vertiefende Analysen zu Leseverständnis, Rahmenbedingungen und Zusatzstudien*. Münster: Waxmann.
- Brinker, K. & Ausborn-Brinker, S. (2010). *Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden*. Berlin: Schmidt.
- Buck, G. (2004). *Assessing listening*. New York: Cambridge University Press.
- Bundesministerium für Bildung und Forschung (2012). Deutschlands Grundschülerinnen und Grundschüler im Lesen, in Mathematik und in Naturwissenschaften im internationalen Vergleich weiterhin im oberen Drittel. Verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_12_11-PM_TIMSS_IGLU.pdf [07.01.2014].
- Campbell, J.R., Kelly, D.L., Mullis, I.V.S., Martin, M.O. & Sainsbury, M. (2001). *Framework and specifications for PIRLS assessment 2001* (2nd ed.). Chestnut Hill, MA: Boston College.
- Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N.H. Hornberger (Hrsg.): *Encyclopedia of Language and Education, 2nd Edition, Volume 2: Literacy*. New York: Springer Science + Business Media LLC.

- Danks, J.H. (1980). Comprehension in listening and reading: Same or different? In J. Danks & K. Pezdek (Hrsg.), *Reading and understanding*. Newark, Delaware: International Reading Association.
- Deutsches PISA - Konsortium (2000). *Schülerleistungen im internationalen Vergleich. Eine neue Rahmenkonzeption für die Erfassung von Wissen und Fähigkeiten*. Berlin: Max-Planck-Institut für Bildungsforschung.
- Dirven, R. (1984). Was ist Hörverstehen? Synopse vorhandener Theorien und Modelle. In A. Schumann, K. Vogel & B. Voss (Hrsg.), *Hörverstehen. Grundlagen, Modelle, Materialien zur Schulung des Hörverstehens im Fremdsprachenunterricht der Hochschule*. Tübingen: Narr.
- Drosdowski, G. (Hrsg.) (1998). *Duden. Grammatik der deutschen Gegenwartssprache. 6., völlig neu bearbeitete und erweiterte Auflage*. Mannheim, Wien, Zürich: Dudenverlag.
- Dubowy, M., Ebert, S., von Maurice, J. & Weinert, S. (2008). *Sprachlich-kognitive Kompetenzen beim Eintritt in den Kindergarten. Ein Vergleich von Kindern mit und ohne Migrationshintergrund*. Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 40, 124-134.
- Dunn, L. M., & Dunn, D. M. (1981). *Peabody Picture Vocabulary Test – Revised (PPVT-R)*. Circle Pines, NM: American Guidance Service.
- Dunn, L. M., & Dunn, D. M. (2004). *Peabody Picture Vocabulary Test (PPVT) (deutsche Version)*. Göttingen, Germany: Hogrefe.
- Feyten, C. M. (1991). The power of listening ability: An overlooked dimension in language acquisition. *The Modern Language Journal*, 75, 173-80.
- Gehrer, K., Zimmermann, S., Artelt, C., & Weinert, S. (2013). NEPS-Framework for assessing reading competence and results from an adult pilot study. *Journal of Educational Research Online*, 5, 2, 50-79.
- Gogolin, I. & Lange, I. (2010). Bildungssprache und Durchgängige Sprachbildung. In Fürstenau, S. & Gomolla, M. (Hrsg.): *Migration und schulischer Wandel: Mehrsprachigkeit*. Wiesbaden: VS-Verlag.
- Grotjahn, R. (2001). Determinants of the Difficulty of Foreign Language Reading and Listening Comprehension Tasks: Predicting Task Difficulty in Language Tests. In: H. Pürschel & U. Raatz (Hrsg.), *Tests and translation. Papers in memory of Christine Klein-Braley*. Bochum: AKS 2001.
- Grotjahn, R. (2005). Testen und Bewerten des Hörverstehens. In M. Ó Dúill, R. Zahn & K.D.C. Höppner (Hrsg.), *Zusammenarbeiten: Eine Festschrift für Bernd Voss*. Bochum: AKS-Verlag.
- Helbig, G. & Kempter, F. (1997). *Das Passiv*. Leipzig u. a.: Langenscheidt Verlag Enzyklopädie.
- Helmke, A. (2007). *Unterrichtsqualität erfassen, bewerten, verbessern*. Seelze: Klett.
- Heppt, B., Dragon, N., Berendes, K., Stanat, P. & Weinert, S. (2012). Beherrschung von Bildungssprache bei Kindern im Grundschulalter. *Diskurs Kindheits- und Jugendforschung* (3), 349-356.

- Hoffmann, M. (2010). *Deutsch fürs Studium. Grammatik und Rechtschreibung. 2., aktualisierte und überarbeitete Auflage*. Paderborn: Schöningh.
- Hoover, W. A. & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160.
- IDS (2013). Verfügbar unter <http://www.ids-mannheim.de/gra/konnektoren/anfrage.html> [26.06.2013].
- IQB (2011). *Kompetenzstufenmodell zu den Bildungsstandards im Fach Deutsch und im Kompetenzbereich „Sprechen und Zuhören“ für den Primarbereich*. Entwurf in der Version vom 22.03.2011. Verfügbar unter: https://www.iqb.hu-berlin.de/bista/ksm/KSM_GS_Deutsch_Z.pdf [20.01.2014].
- IQB (2013). Deutschaufgaben der Domäne Zuhören für die 8. Klasse. Verfügbar unter <http://www.iqb.hu-berlin.de/vera/aufgaben> [26.06.2013].
- Jude, N. & Klieme, E. (2007). Sprachliche Kompetenz aus Sicht der pädagogisch-psychologischen Diagnostik. In: E. Klieme & B. Beck (Hrsg.), *Sprachliche Kompetenzen. Konzepte und Messung*. DESI-Studie (Deutsch Englisch Schülerleistungen International), (S. 9-22). Weinheim u.a. : Beltz 2007.
- Kessel, K. & Reimann, S. (2010). *Basiswissen Deutsche Gegenwartssprache. Dritte, überarbeitete Auflage*. Tübingen, Basel: A. Francke.
- Kintsch, W. & van Dijk, T.A. (1978). *Towards a model of text comprehension and production*. *Psychological Review*, 85 (5), 363-394.
- Klieme, E. (2006). *Zusammenfassung zentraler Ergebnisse der DESI-Studie*. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.
- Klieme, E., Eichler, W., Helmke, A., Lehmann, R.H., Nold, G., Rolff, H.-G., Schröder, K., Thomé, G. & Willenberg, H. (2006). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Zentrale Befunde der Studie Deutsch-Englisch-Schülerleistungen-International (DESI)*. Frankfurt am Main: Deutsches Institut für Internationale Pädagogische Forschung.
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M., Schneider, W., Stanat, P. (2010). *PISA 2009. Bilanz nach einem Jahrzehnt*. Waxmann: Münster.
- KMK (2004). *Bildungsstandards im Fach Deutsch für den Mittleren Schulabschluss. Beschluss vom 04.12.2003*. München: Luchterhand.
- KMK (2005a). *Bildungsstandards der Kultusministerkonferenz. Erläuterungen zur Konzeption und Entwicklung*. München: Luchterhand.
- KMK (2005b). *Bildungsstandards im Fach Deutsch für den Primarbereich. Beschluss vom 15.10.2004*. München: Luchterhand.
- KMK (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring. Beschluss vom 02.06.2006*. München: Luchterhand.
- KMK (2012). *Vereinbarung zur Weiterentwicklung von VERA. Beschluss der Kultusministerkonferenz vom 08.03.2012*. Berlin.
- Kürschner, C. & Schnotz, W. (2008). Das Verhältnis gesprochener und geschriebener Sprache bei der Konstruktion mentaler Repräsentationen. *Psychologische Rundschau*, 59, 139-149.

- Kutner, M., Greenberg, E. & Baer, J. (2005). *A first Look at the Literature of American's Adults in the 21st century (NCES 2006-470)*. U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Leucht, M., Retelsdorf, J., Möller, J. & Köller, O. (2010). *Zur Dimensionalität rezeptiver englischsprachiger Kompetenzen*. *Zeitschrift für Pädagogische Psychologie*, 24, 123-138.
- Marx, A. & Roick, T. (2012). *Prädiktoren des Hörverstehens bei Jugendlichen deutscher und Jugendlichen nichtdeutscher Herkunftssprache*. *Zeitschrift für Pädagogische Psychologie*, 26 (2), 121-134.
- Marx, H. (1998). *Knuspels Leseaufgaben: Gruppenlesetest für Kinder Ende des ersten bis vierten Schuljahres*. Göttingen: Hogrefe.
- Marx, H. & Jungmann, T. (2000). *Abhängigkeit der Entwicklung des Leseverstehens von Hörverstehen und grundlegenden Lesefertigkeiten im Grundschulalter. Eine Prüfung des <Simple View of Reading>-Ansatzes*. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 32, 81-93.
- Mulligan, G.M., Hastedt, S. & McCarroll, J.C. (2012). *First-Time Kindergartners in 2010-11: First Findings From the Kindergarten Rounds of the Early Childhood Longitudinal Study, Kindergarten Class of 2010-11 (ECLS-K:2011)*. (NCES 2012-049). Washington, DC: National Center for Education Statistics.
- Najarian, M., Snow, K., Lennon, J. & Kinsey, S. (2010). *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B) Preschool-Kindergarten 2007 Psychometric Report*. (NCES 2010-009). Washington, DC: National Center for Education Statistics.
- National Center for Education Statistics (2011). *The Nation's Report Card: Reading 2011 (NCES 2012-457. 2011)*. Institute of Educational Sciences, U.S. Department of Education, Washington, D.C.
- Niklas, F., Segerer, R., Schmiedeler, S. & Schneider, W. (2012). *Findet sich ein „Matthäus-Effekt“ in der Kompetenzentwicklung von jungen Kindern mit oder ohne Migrationshintergrund?* *Frühe Bildung*, 1 (1), 26-33.
- OECD (2010). *PISA 2009 Ergebnisse: Was Schülerinnen und Schüler wissen und können Schülerleistungen in Lesekompetenz, Mathematik und Naturwissenschaften (Band I)*.
- OECD (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. OECD Publishing.
- Pollack, J., Atkins-Burnett, S., Rock, D. & Weiss, M. (2005). *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K), Psychometric Report for the Third Grade (NCES 2005-062)*. U.S. Department of Education. Washington, DC: National Center for Education Statistics.
- Pasch, R., Brauße, U., Breindl, E. & Waßner, U.H. (2003): *Handbuch der deutschen Konnektoren. Linguistische Grundlagen der Beschreibung und syntaktische Merkmale der deutschen Satzverknüpfers (Konjunktionen, Satzadverbien und Partikeln)*. Berlin, New York: Walter de Gruyter.

- Porsch, R. & Tesch, B. (2010). Messung der fremdsprachlichen Schreibkompetenz für Französisch als erste Fremdsprache. In R. Porsch, B. Tesch & O. Köller (Hrsg.), *Standardbasierte Testentwicklung und Leistungsmessung: Französisch in der Sekundarstufe*. Münster: Waxmann.
- Prüfer, P. & Rexroth, M. (2005). Kognitive Interviews. *ZUMA How-to-Reihe 15*.
- Richter, T. & Christmann, U. (2002). Lesekompetenz: Prozessebenen und interindividuelle Unterschiede. In: N. Groeben & B. Hurrelmann (Hrsg.), *Lesekompetenz: Bedingungen, Dimensionen, Funktionen*. Weinheim: Juventa.
- Rost, D.H. & Buch, S. (2010). Leseverständnis. In D.H. Rost (Hrsg.), *Handwörterbuch Pädagogische Psychologie*. Weinheim: Beltz.
- Rubin, D. L., Hafer, T. & Arata, K. (2000). Reading and listening to oral-based versus literate-based discourse. *Communication Education, 49*, 121-133.
- Schnotz, W. & Dutke, S. (2004). Kognitionspsychologische Grundlagen der Lesekompetenz: Mehrebenenverarbeitung anhand multipler Informationsquellen. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz*. Vertiefende Analysen im Rahmen von PISA 2000. VS: Wiesbaden.
- Shell-Jugendstudie (2010). Eine pragmatische Generation behauptet sich. Verfügbar unter: <http://www.shell.de/aboutshell/our-commitment/shell-youth-study/2010.html> [07.01.2014].
- U.S. Department of Education, Office of Educational Research and Improvement (1999). National Center for Education Statistics. The NAEP 1996 Technical Report, NCES 1999–452, by Allen, N.L., Carlson, J.E., & Zelenak, C.A. (1999). Washington, DC: National Center for Education Statistics.
- U.S. Department of Education, National Center for Education Statistics (2003). NAEP Validity Studies: An Agenda for NAEP Validity Research, NCES 2003–07, by Frances B. Stancavage, et al. Project Officer: Patricia Dabbs. Washington, DC: 2003.
- Van Dijk, T.A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Verhoeven, L. & Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: A longitudinal study. *Applied Cognitive Psychology, 22*, 407-423.
- Weinert, S. (2006). Sprachentwicklung. In W. Schneider & B. Sodian (Hrsg.), *Kognitive Entwicklung (Enzyklopädie der Psychologie, Serie 5 Entwicklungspsychologie, Vol. 2, S. 609-719)*. Göttingen: Hogrefe.
- Weinert, S. (2007). Wie Sprache das Wissen und Denken beeinflusst. In A. A. Bucher, K. Laueremann & A. M. Kalcher (Hrsg.), *Sprache leben. Kommunizieren & Verstehen*. Wien: G&G Verlag.
- Weinert, S. (2010). Erfassung sprachlicher Fähigkeiten. In E. Walther, F. Preckel & S. Mecklenbräuer (Hrsg.), *Befragung von Kindern und Jugendlichen*. Göttingen: Hogrefe.
- Weinert, S., Artelt, C., Prenzel, M., Senkbeil, M., Ehmke, T., & Carstensen, C. H. (2011). *Development of competencies across the lifespan*. *Zeitschrift für Erziehungswissenschaft, 14*, 67-86.

- Weinert, S. & Grimm, H. (2012). Sprachentwicklung. In W. Schneider & U. Lindenberger (Hrsg.), *Entwicklungspsychologie*. Weinheim: Beltz.
- Weinert, S. & Ebert, S. (2013). Spracherwerb im Vorschulalter: Soziale Disparitäten und Einflussvariablen auf den Grammatikerwerb. *Zeitschrift für Erziehungswissenschaft*, *16*, 303-332.
- White, S. & Dillow, S. (2005). *Key Concepts and Features of the 2003 National Assessment of Adult Literacy (NCES 2006-471)*. U.S. Department of Education. Washington, DC: National Center for Education Statistics.