

**Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin**

Hören und Verstehen:
Struktur und Determination des Hörverstehens bei Schülerinnen und Schülern
mit Deutsch als Erst- bzw. Zweitsprache

Dissertation
zur Erlangung des akademischen Grades
Doktorin der Philosophie (Dr. phil.)

vorgelegt von
Dipl.-Psych. Alexandra Marx

Berlin, 2015

Erstgutachter: Prof. Dr. Martin Brunner

Zweitgutachterin: Prof. Dr. Petra Stanat

Betreuer: Dr. Thorsten Roick

Tag der Disputation: 8. Februar 2016

Gib mir Gehör und ich werde dir Stimme geben.

Khalil Gibran

Diese Dissertation ist im Rahmen des Projektes „Determinanten der Lesekompetenz“ (DELKO) unter Leitung von Prof. Dr. Petra Stanat entstanden. Das in Kooperation mit der Julius-Maximilians-Universität Würzburg durchgeführte DELKO-Projekt wurde ermöglicht durch Sachbeihilfen der Deutschen Forschungsgemeinschaft (STA 626/6-1 und SCHN 315/40-1).

Danksagung

Prof. Dr. Petra Stanat möchte ich herzlich für ihre Begleitung und ihr Interesse während der Promotionsphase danken. Den Austausch mit ihr und ihren fachlichen Rat habe ich stets als große Bereicherung empfunden.

Bei Prof. Martin Brunner bedanke ich mich herzlich für die Bereitschaft zur Übernahme der Erstbegutachtung meiner Arbeit.

Mein besonderer Dank gilt Dr. Thorsten Roick, der mir eine wichtige Unterstützung während der Promotionsphase war. Seine konstruktiven Rückmeldungen und Anmerkungen haben wesentlich zum Gelingen dieser Arbeit beigetragen.

Den Kolleginnen und Kollegen am IQB möchte ich darüber hinaus für die kollegiale und herzliche Atmosphäre bedanken, innerhalb derer ich meine Dissertation anfertigen konnte.

Inhaltsverzeichnis

1. Einleitung.....	1
1.1 Ausgangslage: Bildungsbeteiligung und Kompetenzerwerb von Schülerinnen und Schülern mit Zuwanderungshintergrund in Deutschland	3
1.2 Disparitäten im Erwerb schulischer Kompetenzen zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund: Effekte des sozioökonomischen und sprachlichen Hintergrunds	5
1.3 Theorien zum Zweitspracherwerb	8
2. Hörverstehen in einer Erst- und einer Zweitsprache	12
2.1 Konzeptionen des Hörverstehens	12
2.2 Klassifikationen von Prädiktoren des Hörverstehens	14
2.3 Kognitive Prozessmodelle des Hörverstehens.....	16
2.4 Interaktionistische Modelle des Textverstehens.....	18
2.5 Hörverstehen in einer Zweitsprache	22
3. Linguistische Komponenten des Hörverstehens.....	24
3.1 Arbeitsgedächtnis	25
3.2 Phonologische Bewusstheit	29
3.3 Wortschatz.....	32
3.4 Morpho-syntaktische Kenntnisse	34
4. Die Rolle des sprachlichen Kontextes: Bildungssprachliches und Alltagssprachliches Hörverstehen.....	37
5. Die Bedeutung des Hörverstehens für Leseverstehen in einer Erst- und Zweitsprache.....	43
6. Anliegen dieser Arbeit.....	47
7. Ziele, Methoden und Ergebnisse der Einzelstudien	48
7.1 Methodisches Vorgehen der Einzelstudien	48
7.1.1 Stichprobe.....	48

7.1.2	Teilstudie 1: Prädiktoren des Hörverstehens bei Jugendlichen deutscher und Jugendlichen nichtdeutscher Herkunftssprache (Marx & Roick, 2012).....	51
7.1.3	Teilstudie 2: Prädiktoren der Lesekompetenz in einer Erst- bzw. Zweitsprache (Marx, Stanat, Roick, Segerer, Marx & Schneider, 2015).....	54
7.1.4	Teilstudie 3: Bildungssprachliches und alltagssprachliches Hörverstehen in einer Erst- bzw. Zweitsprache (Marx, Heppt & Henschel; in Revision).....	57
7.2	Analysestrategie	59
7.3	Umgang mit fehlenden Werten	62
7.4	Ergebnisse.....	65
7.4.1	Teilstudie 1: Prädiktoren des Hörverstehens bei Jugendlichen deutscher und Jugendlichen nichtdeutscher Herkunftssprache (Marx & Roick, 2012).....	65
7.4.2	Teilstudie 2: Prädiktoren der Lesekompetenz in einer Erst- bzw. Zweitsprache (Marx, Stanat, Roick, Segerer, Marx & Schneider, 2015).....	67
7.4.3	Teilstudie 3: Bildungssprachliches und alltagssprachliches Hörverstehen in einer Erst- bzw. Zweitsprache (Marx, Heppt & Henschel; in Revision).....	71
8.	Zusatzanalyse: Die Bedeutung alltagssprachlicher und bildungssprachlicher Wortschatzkenntnisse für das Hörverstehen in einer Erst- bzw. Zweitsprache	73
8.1	Forschungsfragen	80
8.2	Methode.....	82
8.3	Ergebnisse.....	83
9.	Abschließende Diskussion und Ausblick	87
9.1	Hörverstehen in einer Erst- bzw. Zweitsprache: Linguistische Prädiktoren und Bedeutung für das Leseverstehen	87
9.2	Alltagssprachliches und bildungssprachliches Hörverstehen in einer Erst- bzw. Zweitsprache.....	100
9.3	Grenzen der vorliegenden Arbeit	105
9.4	Ausblick.....	107
9.4.1	Implikationen für die pädagogische Praxis: Förderungen des Hörverstehens ..	107
9.4.2	Implikationen für die zukünftige Forschung.....	111
10.	Literatur	114

Anhang A: Prädiktoren des Hörverstehens bei Jugendlichen deutscher und Jugendlichen nichtdeutscher Herkunftssprache (Teilstudie 1)	137
Anhang B: Prädiktoren der Lesekompetenz in einer Erst- bzw. Zweitsprache (Teilstudie 2).....	165
Anhang C: Bildungssprachliches und Alltagssprachliches Hörverstehen in einer Erst- bzw. Zweitsprache (Teilstudie 3)	197
Erklärung	230

Zusammenfassung

Im Rahmen der an die PISA 2009-Erhebung angegliederten Studie *Determinanten der Lesekompetenz* (DELKO) wurden drei empirische Studien zur Untersuchung linguistischer Prädiktoren des Hörverstehens in einer Erst- (L1) bzw. Zweitsprache (L2) sowie dem Zusammenhang zwischen Hör- und Leseverstehen durchgeführt. Ziel der Studien war eine Klärung der besonderen Herausforderungen, denen Schülerinnen und Schülern beim Hörverstehen in einer Zweitsprache begegnen und die ihre Hörverstehensleistungen beeinträchtigen können. Weiterhin wurden die Abgrenzbarkeit und die Ausprägung bildungssprachlicher und alltags-sprachlicher Hörverstehenskompetenzen bei Jugendlichen mit unterschiedlicher Herkunftssprache vergleichend untersucht. In allen drei Studien wurden Schülerinnen und Schüler mit deutschem, türkischem sowie anderem nicht-deutschen Sprachhintergrund einbezogen.

Die Ergebnisse der ersten Studie erbrachten Hinweise darauf, dass die Hörverstehensleistung in einer Erst- bzw. Zweitsprache im Wesentlichen von denselben Prädiktoren bestimmt wird, L2-Jugendliche allerdings im Mittel schwächere Hörverstehenskompetenzen aufweisen als L1-Schülerinnen und -Schüler. Dabei zeigte sich, dass mit dem Wortschatz und den morpho-syntaktischen Kenntnissen diejenigen Prädiktoren einen Effekt auf die Hörverstehensleistung haben, die für die Bildung einer mentalen Repräsentation des Bedeutungsinhalts des Hörtextes entscheidend sind. Insbesondere L2-Jugendliche mit türkischer Herkunftssprache weisen bei diesen Prädiktoren gegenüber L1-Schülerinnen und -Schülern schwächer ausgeprägte Kompetenzen auf. Wie Teilstudie 2 zeigt, scheinen Nachteile in sprachlichen Kompetenzen bei L2-Jugendlichen mit einer anderen Herkunftssprache als Türkisch vornehmlich auf den niedrigen sozioökonomischen Status dieser Gruppe zurückführbar zu sein. Für L2-Jugendliche mit türkischer Herkunftssprache finden sich hingegen auch unter Berücksichtigung des sozialen Hintergrunds persistierende Schwächen im Hörverstehen sowie dessen Prädiktoren. Da in allen Sprachgruppen das Hörverstehen eng mit dem Leseverstehen verbunden ist, ist denkbar, dass diese Schwächen auch Disparitäten im Leseverstehen zu einem erheblichen Teil mit bedingen.

Teilstudie 3 zeigt auf, dass L2-Jugendliche im Mittel schwächere bildungssprachliche und alltags-sprachliche Hörverstehenskompetenzen aufweisen als L1-Jugendliche, wobei die Differenzen im Bereich des alltags-sprachlichen Hörverstehens ausgeprägter sind. Auf Basis dieser Befunde wurde in einer Zusatzstudie die Rolle des bildungssprachlichen und des alltags-sprachlichen Wortschatzes für das Hörverstehen näher untersucht. Es zeigt sich, dass Unterschiede im alltags-sprachlichen Wortschatz ausgeprägter sind als Unterschiede im bildungssprachlichen Wortschatz. Eine mögliche Ursache für diesen Befund liegt im Ausmaß des

Kontakts zu den unterschiedlichen Wortschatzfacetten in der Zweitsprache. Die durchgeführten Studien legen nahe, dass dieselben linguistischen Prädiktoren bedeutsam für das Hörverstehen in einer Erst- und Zweitsprache sind. Allerdings weisen L2-Jugendliche mit türkischer Herkunftssprache auch unter Berücksichtigung des sozialen Hintergrunds im Vergleich zu monolingualen Jugendlichen häufig schwächer ausgeprägte Kompetenzen im Bereich dieser Prädiktoren auf. Disparitäten im Hörverstehen betreffen dabei nicht nur bildungssprachliche Kompetenzen, sondern auch das alltagssprachliche Hörverstehen, dessen Beherrschung von Lehrkräften in den höheren Klassenstufen häufig vorausgesetzt wird. Implikationen dieser Befunde für die pädagogische Praxis, insbesondere in Hinblick auf die Förderung von Hörverstehenskompetenzen, sowie für zukünftige Forschungsvorhaben werden diskutiert.

Schlüsselwörter: Hörverstehen, Zweitsprache, Bildungssprache, Sozioökonomischer Status

Abstract

Within the research project “Determinanten der Lesekompetenz (DELKO)” (*Determinants of Reading Comprehension*), three studies were carried out to investigate linguistic determinants of listening comprehension as well as the effects of listening comprehension on reading competencies in first-language (*L1*) and second-language (*L2*) students. The studies aimed at analyzing the specific challenges students encounter in second-language listening. Furthermore, the separability and peculiarity of academic and everyday listening competencies in *L1* and *L2* students was tested empirically.

Results from the first study indicate that listening comprehension in *L1* and *L2* adolescents is influenced by the same set of linguistic determinants, but that *L2* students show weaker listening comprehension skills than their *L1* peers. For all language groups, vocabulary and morpho-syntactic knowledge, i.e. determinants that are needed to generate a mental representation of a listening texts’ gist, had a strong impact on students’ listening skills. At the same time, especially *L2* students with Turkish as a first language exhibited poorer performance than *L1* students on measures of these determinants. Study 2 shows that performance differences between *L1* students and *L2* students who speak another language than Turkish as a first language mainly result from the lower socioeconomic background of the latter group. For *L2* students with Turkish as a first language, however, performance disadvantages persist even after controlling for students’ socioeconomic status. As listening comprehension is strongly associated with reading comprehension in all language groups, these performance disadvantages may also impact *L2* students’ reading skills.

Study 3 shows that *L2* adolescents exhibit weaker academic listening skills as well as weaker everyday listening skills compared to their *L1* peers. Contrary to what was expected, the performance differences were more pronounced for students’ everyday listening comprehension than for academic listening. Based on these results, study four explores the impact of academic and everyday vocabulary on listening comprehension. Findings indicate that *L1* and *L2* students differ more strongly in terms of everyday vocabulary than in academic vocabulary. These results may possibly explained by the differing exposure *L2* students have to the respective facets of *L2* vocabulary.

The studies point out that the same set of linguistic determinants has effects on *L1* students’ and *L2* students’ listening comprehension while especially *L2* students with Turkish as a first language exhibit persistent weaknesses in terms of these determinants. The poorer performance of *L2* students not only pertains to their academic listening skills, but also to their everyday listening skills which teachers often assume to be well developed. The implications

of these results for educational practice, especially in terms of fostering students' listening comprehension skills, as well as for future research are discussed.

Keywords: Listening Comprehension, Second-Language Students, Academic Language.
Socioeconomic Status

1. Einleitung

Hörverstehenskompetenzen können als wichtige Voraussetzung des Erwerbs weiterer sprachlicher und schulischer Fähigkeiten betrachtet werden, da das inhaltliche und formale Verständnis gesprochener Sprache in nahezu allen Unterrichtssituationen und -fächern eine wichtige Rolle spielt (Imhof, 2010). Gut ausgebildete Hörverstehensfähigkeiten bilden somit nicht nur eine zentrale Grundlage der kommunikativen Kompetenz im Alltag, sondern beeinflussen auch den schulischen Kompetenzerwerb und den Bildungserfolg von Schülerinnen und Schülern in erheblichem Maße. Entsprechend wird in den von der Kultusministerkonferenz formulierten Bildungsstandards die Bedeutung des Hörverstehens für die unterrichtliche Teilhabe und den Kompetenzerwerb betont (z. B. KMK, 2005). Insbesondere in den sprachlichen Fächern wird verstehendes Zuhören demnach als ein wesentliches Lernziel des Unterrichts begriffen. Aber auch in den übrigen Fächern sind Hörverstehensfähigkeiten wichtig, um Äußerungen von Anderen zu fachspezifischen Inhalten verstehen und überprüfen zu können. Hörverstehen bildet somit eine Schlüsselqualifikation, die den Erwerb weiterer schulischer Kompetenzen entscheidend mitbedingt.

Aufgrund dieser Bedeutsamkeit ist das Hörverstehen in jüngster Zeit verstärkt in den Fokus der pädagogisch-psychologischen sowie fachdidaktischen Forschung gerückt. Im schulischen Unterricht werden gut ausgebildete Hörverstehensfähigkeiten der Schülerinnen und Schüler allerdings bislang häufig als gegeben vorausgesetzt. In der Folge wird der unterrichtliche Fokus oft eher auf die Vermittlung fachspezifischer Kompetenzen gelegt, während Maßnahmen zur Vermittlung oder Überprüfung des Hörverstehens nur selten Gegenstand des Unterrichts sind. Für Schülerinnen und Schüler mit Zuwanderungshintergrund kann dies problematisch sein: Für viele dieser Heranwachsenden ist Deutsch die Zweitsprache, so dass sie teilweise nicht über dasselbe sprachliche Vorwissen verfügen wie ihre Peers ohne Zuwanderungshintergrund. Obwohl das Aufwachsen mit mehreren Sprachen mit einer Reihe von Vorteilen hinsichtlich kognitiver oder metasprachlicher Fähigkeiten einhergehen kann, können in Folge des geringeren sprachlichen Vorwissens geringere Hörverstehensfähigkeiten im Vergleich mit Schülerinnen und Schülern ohne Zuwanderungshintergrund resultieren.

Für Lernende in einer Zweitsprache ist eine gut ausgebildete Hörverstehenskompetenz auch deshalb von zentraler Bedeutung für die weitere schulische Laufbahn, weil sie die Grundlage des Erwerbs weiterer sprachlicher und schulischer Fähigkeiten in der Zweitsprache bildet (Rost, 1994). Ergebnisse nationaler und internationaler Erhebungen wie dem Ländervergleich des IQB oder den PISA-Studien zeigen, dass Schülerinnen und Schüler mit Zuwanderungshintergrund häufig schwächere Leistungen in Verfahren zur Überprüfung ihrer schuli-

schen Kompetenzen erzielen als Kinder und Jugendliche ohne Zuwanderungs-hintergrund. Im IQB-Ländervergleich 2012 fanden sich beispielsweise deutliche geringere mathematische und naturwissenschaftliche Kompetenzen bei Neuntklässlern mit Zuwanderungshintergrund gegenüber ihren Peers ohne Zuwanderungshintergrund (Pöhlmann, Haag & Stanat, 2013). Ähnliche Leistungs Nachteile, auch in Bezug auf die Lesekompetenz, wurden in den PISA-Studien berichtet (OECD, 2002; 2010; 2013). Diese Befunde deuten darauf hin, dass es Jugendlichen mit Zuwanderungshintergrund in Deutschland bis zum Ende der Schulpflicht zu einem großen Teil nicht gelingt, Kompetenzen zu entwickeln, die denen von Schülerinnen und Schülern ohne Zuwanderungshintergrund entsprechen.

Über die Ursachen dieser Disparitäten zwischen Jugendlichen mit bzw. ohne Zuwanderungshintergrund herrscht bislang keine Klarheit. Angesichts der Bedeutsamkeit des Hörverstehens für den Erwerb weiterer sprachlicher und schulischer Fähigkeiten ist jedoch denkbar, dass diese Kompetenzunterschiede mittelbar auch durch Schwächen im Hörverstehen der Jugendlichen mit Zuwanderungshintergrund bedingt sein könnten. Allerdings wurden in den meisten der bisherigen Schulleistungsstudien die Hörverstehenskompetenzen von Schülerinnen und Schülern nicht berücksichtigt, so dass unklar ist, inwiefern diese Annahme empirische Unterstützung erfährt. Lediglich in den Ländervergleichsstudien 2009 und 2011 wurden auch Hörverstehenskompetenzen erfasst, wobei sich jeweils Leistungsunterschiede zwischen Kinder mit bzw. ohne Zuwanderungs-hintergrund zeigten (Böhme, Tiffin-Richards, Schipowski & Leucht, 2010; Haag, Böhme & Stanat, 2012; vgl. Kapitel 1.1). Allerdings fokussieren die Ländervergleichsstudien auf den Grundschulbereich. Über die Ausprägung des Hörverstehens und dessen Komponenten sowie dem Zusammenhang zu weiteren schulischen Kompetenzen bei älteren Jugendlichen, die gegenüber Grundschülerinnen und -schülern in der Regel über eine längere Geschichte des Kontakts zur Zweitsprache verfügen, ist wenig bekannt.

Vor diesem Hintergrund beschäftigt sich die vorliegende Arbeit mit dem Hörverstehen und dessen linguistischen Prädiktoren bei Jugendlichen mit und ohne Zuwanderungshintergrund, die Deutsch als Erst- oder als Zweitsprache erworben haben. Im Fokus der Analysen stehen dabei neben der Ausprägung einzelner Komponenten des Hörverstehens auch deren prädiktive Validität zur Vorhersage der Hörverstehenskompetenz bei Schülerinnen und Schülern mit Deutsch als Erst- bzw. Zweitsprache. Darüber hinaus wird die Rolle des Hörverstehens für die Erklärung schwacher Leseleistungen bei Jugendlichen unterschiedlicher Herkunftssprache betrachtet. Abschließend wird analysiert, inwiefern Schwächen von Jugendlichen nicht-deutscher Herkunftssprache im Bereich des Hörverstehens das alltagssprachliche

bzw. das bildungssprachliche Hörverstehen betreffen. Im Folgenden wird zunächst auf die Situation von Schülerinnen und Schülern mit Zuwanderungshintergrund in Deutschland in Bezug auf die Bildungsbeteiligung und den schulischen Kompetenzerwerb eingegangen. Darauf aufbauend folgt eine genauere Darstellung der theoretischen Grundlagen sowie der Erkenntnisziele der vorliegenden Arbeit. Anschließend werden die Methoden und Ergebnisse der durchgeführten Studien beschrieben und die zentralen Ergebnisse zusammenfassend diskutiert.

1.1 Ausgangslage: Bildungsbeteiligung und Kompetenzerwerb von Schülerinnen und Schülern mit Zuwanderungshintergrund in Deutschland

Sowohl der berufliche Erfolg als auch die gesellschaftliche Teilhabe werden maßgeblich durch das erreichte Ausbildungsniveau beeinflusst. Eine hohe Bildungsbeteiligungsquote kann somit als wesentliche Grundlage für die Integration von Menschen mit Zuwanderungshintergrund angesehen werden. Daten des aktuellen Bildungsberichts 2014 (Autorengruppe Bildungsberichterstattung, 2014) legen allerdings nahe, dass Unterschiede im Bildungserfolg von Personen mit bzw. ohne Zuwanderungshintergrund teilweise erheblich sind. So bestehen in der Gruppe der 16- bis unter 30-jährigen Menschen mit bzw. ohne Zuwanderungshintergrund erhebliche Unterschiede hinsichtlich der Art der besuchten Bildungseinrichtung (Autorengruppe Bildungsberichterstattung, 2014). Disparitäten in den erzielten Bildungsabschlüssen finden sich dabei sowohl hinsichtlich höherwertiger Abschlüsse, bei denen Personen mit Zuwanderungshintergrund unterrepräsentiert sind, als auch hinsichtlich niedrigqualifizierter oder fehlender Bildungsabschlüsse, wo Personen mit Zuwanderungshintergrund überrepräsentiert sind. Unklar bleibt allerdings, inwiefern diese Disparitäten Unterschiede im Erwerb schulischer Kompetenzen reflektieren, die wesentliche Prädiktoren des erreichten Bildungsniveaus darstellen. Da in der Berichterstattung für diese Variablen nicht kontrolliert wurde, ist beispielsweise denkbar, dass geringerwertige Bildungsabschlüsse von Personen mit Zuwanderungshintergrund eine geringere Ausprägung schulischer Kompetenzen abbilden.

Ergebnisse nationaler und internationaler Studien sind im Einklang mit dieser Annahme und zeigen, dass sich Nachteile von Personen mit Zuwanderungshintergrund gegenüber nicht zugewanderten Personen nicht nur in der Bildungsbeteiligung, sondern auch auf Ebene des Kompetenzerwerbs finden. So bestehen bereits in der Grundschule in Abhängigkeit vom Zuwanderungshintergrund zum Teil erhebliche Unterschiede in den erzielten schulischen Kompetenzen: Kinder aus Familien mit Zuwanderungshintergrund erreichen sowohl im Lesen als auch im mathematischen und naturwissenschaftlichen Bereich durchgängig geringere Kompe-

tenzniveaus als Kinder ohne Zuwanderungshintergrund (Bonsen, Kummer & Bos, 2008; Schwippert, Bos & Lankes, 2003). In der IGLU-Studie 2006 beispielsweise lag die durchschnittlich erzielte Leseleistung der Kinder aus zugewanderten Familien 48 Leistungspunkte und somit rund eine halbe Standardabweichung unter der Leseleistung von Kindern aus nicht zugewanderten Familien. Dieser Rückstand im Kompetenzerwerb entspricht in etwa dem Lernzuwachs, der in der Regel innerhalb eines Schuljahres erzielt wird (Schwippert, Hornberg, Freiberg & Stubbe, 2007). Ähnliche Disparitäten zeigten sich im IQB-Ländervergleich 2011 auch für den Kompetenzbereich „Zuhören“ im Fach Deutsch. Hier fanden sich bei Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe besonders stark ausgeprägte Mittelwertsunterschiede in der Hörverstehenskompetenz zwischen Kindern mit und ohne Zuwanderungshintergrund (Haag, Böhme & Stanat, 2012). Verglichen mit dem erwarteten mittleren Leistungszuwachs von der dritten bis zur vierten Jahrgangsstufe entsprachen die Differenzen im Zuhören dem Lerngewinn, der in der Regel in etwas mehr als einem Schuljahr erzielt wird. Insbesondere Schülerinnen und Schüler, deren Eltern aus der Türkei zugewandert waren, wiesen schwache Hörverstehenskompetenzen auf.

Auch für ältere Schülerinnen und Schüler am Ende der Schullaufbahn erbrachten die Ergebnisse der bisherigen PISA-Studien Hinweise darauf, dass Jugendliche mit Zuwanderungshintergrund in allen betrachteten Kompetenzbereichen durchgängig schwächere Leistungen aufweisen als ihre Peers ohne Zuwanderungshintergrund. Allerdings ergab sich seit der ersten PISA-Erhebung im Jahr 2000 eine positive Entwicklung: Während für Schülerinnen und Schüler ohne Zuwanderungshintergrund in der PISA 2009-Erhebung im Mittel keine Verbesserung der Lesekompetenz seit der ersten Erhebung im Jahr 2000 zu konstatieren war, fand sich für Jugendliche mit Zuwanderungshintergrund ein positiver Trend (Stanat, Rauch & Segeritz, 2010). Dennoch zeigten sich auch in der PISA 2009-Erhebung noch Kompetenznachteile im Bereich des Lesens für Schülerinnen und Schüler mit Zuwanderungshintergrund gegenüber ihren Peers ohne Zuwanderungshintergrund, die mit 44 Leistungspunkten in etwa einer halben Standardabweichung auf der PISA-Lesekompetenz-skala entsprachen. Trotz einer positiven Entwicklung seit der PISA 2000 Erhebung erreicht somit immer noch rund ein Drittel der Schülerinnen und Schüler mit Zuwanderungshintergrund nicht ein als Voraussetzung für die weitere gesellschaftliche Teilhabe betrachtetes Niveau der Lesekompetenz (Naumann, Artelt, Schneider & Stanat, 2010).

1.2 Disparitäten im Erwerb schulischer Kompetenzen zwischen Schülerinnen und Schülern mit und ohne Zuwanderungs-hintergrund: Effekte des sozioökonomischen und sprachlichen Hintergrunds

Zur Erklärung der Unterschiede im schulischen Erfolg zwischen Kindern mit und ohne Zuwanderungshintergrund sind eine Reihe von Faktoren diskutiert worden (vgl. Diefenbach, 2002 für einen Überblick). Der sozioökonomische und kulturelle Hintergrund der Familie wird dabei im Sinne der Humankapitaltheorie (Becker, 1993) als eine wesentliche Determinante des Bildungserfolgs betrachtet. Dieser Theorie zufolge wird der schulische Kompetenzerwerb durch die Verfügbarkeit bildungsrelevanter Ressourcen innerhalb der Familien von Schülerinnen und Schülern sowie durch Kosten-Nutzen-Abwägungen (z. B. entgangener Lohn, der während der Ausbildungsphase hätte erzielt werden können oder Aussicht auf ein höheres Einkommen aufgrund eines höherwertigen Bildungsabschlusses) bestimmt. Andere Ansätze zur Erklärung des Effekts des sozioökonomischen Hintergrunds auf den Kompetenzerwerb von Schülerinnen und Schülern betonen eher die Rolle prozessualer Faktoren, wie etwa die innerfamiliäre Sprachpraxis. Angenommen wird dabei, dass mit Unterschieden im sozioökonomischen Status auch Unterschiede in Quantität und Qualität des sprachlichen Inputs innerhalb der Familie einhergehen, die den Zusammenhang zwischen dem sozioökonomischen Hintergrund von Kindern und ihren sprachlichen Fähigkeiten moderieren (Hart & Risley, 1995; Hoff, 2003).

In vielen OECD-Ländern bestehen Differenzen im sozioökonomischen Status von Familien mit und Familien ohne Zuwanderungshintergrund, die die Disparitäten im schulischen Erfolg erklären könnten. In Deutschland sind Unterschiede im sozio-ökonomischen Status dabei verhältnismäßig stark ausgeprägt, unabhängig davon, wie dieser erfasst wird (Autorengruppe Bildungsberichterstattung, 2014; OECD, 2013). In den PISA-Studien wird der sozioökonomische Status anhand der sozialen, kulturellen und ökonomischen Bedingungen innerhalb der Familien von Schülerinnen und Schülern operationalisiert. In den *Index of Economic, Social and Cultural Status* (ESCS) gehen dabei Angaben der Schülerinnen und Schüler zum ausgeübten Beruf und den höchsten Bildungsabschlüssen der Eltern sowie Angaben zum kulturellen, lernbezogenen und computerbezogenen Besitz ein (vgl. Ehmke & Siegle, 2005; OECD, 2006 zur Konstruktion des ESCS). Der ESCS wird bei der Konstruktion so normiert, dass ein Wert von 0 dem sozioökonomischen, sozialen und kulturellen Status eines durchschnittlichen OECD-Jugendlichen entspricht. Die Standardabweichung ist auf 1 festgesetzt. Ergebnisse der PISA 2012-Erhebung zeigen für Deutschland, dass Jugendliche aus Familien ohne Zuwanderungshintergrund einen ESCS von 0.32 aufweisen, Jugendliche

mit Zuwanderungshintergrund hingegen einen ESCS von -0.41 (OECD, 2013). Diese Befunde belegen eine enorme Diskrepanz von mehr als 0.7 Standardabweichungen im sozioökonomischen Hintergrund von Jugendlichen mit bzw. ohne Zuwanderungshintergrund. Kinder und Jugendliche aus zugewanderten Familien verfügen somit häufig über geringere finanzielle, kulturelle und ökologische Ressourcen als ihre Peers ohne Zuwanderungshintergrund.

Entsprechend der Annahme, dass Unterschiede im sozioökonomischen Status die Disparitäten im Erwerb schulischer Kompetenzen zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund zumindest teilweise erklären könnten, reduzieren sich die Leistungsdifferenzen in der Regel, wenn für den sozioökonomischen Hintergrund der Familien kontrolliert wird. Im IQB-Ländervergleich 2011 beispielsweise verringern sich im Kompetenzbereich „Zuhören“ im Fach Deutsch für alle Kinder mit Zuwanderungshintergrund die Leistungsnachteile gegenüber Kindern ohne Zuwanderungshintergrund deutlich, wenn der sozioökonomische Status der Familie berücksichtigt wird. Für Schülerinnen und Schüler, deren Eltern beide aus der Türkei zugewandert sind, reduzieren sich Disparitäten im Zuhören gegenüber Kindern ohne Zuwanderungshintergrund beispielsweise um 35 Leistungspunkte, wenn der sozioökonomische Hintergrund berücksichtigt wird. Diese Verringerung entspricht in etwa der Hälfte des Leistungszuwachses zwischen der dritten und vierten Jahrgangsstufe, der im Kompetenzbereich Zuhören bei ca. 60 Punkten liegt (Haag et al., 2012). Ähnliche Verringerungen in den Leistungsdisparitäten zwischen Kindern mit bzw. ohne Zuwanderungshintergrund ergeben sich im Ländervergleich auch für den Bereich der Lesekompetenz, wenn für Merkmale des sozialen Hintergrunds der Familien kontrolliert wird (Stanat et al., 2012). Auch in der IGLU-Studie sowie der PISA 2009-Erhebung reduzieren sich Leistungsdifferenzen im Lesen zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund unter Kontrolle von Variablen des sozialen Hintergrunds (Schwippert et al., 2007; Stanat, Rauch & Segeritz, 2010).

Trotz eines substanziellen Zusammenhangs des sozioökonomischen und kulturellen Hintergrunds mit dem Kompetenzerwerb von Schülerinnen und Schülern mit Zuwanderungshintergrund (vgl. Baumert & Schümer, 2001) bleiben allerdings in der Regel auch nach Kontrolle dieses Strukturmerkmals familiärer Lebensverhältnisse Disparitäten in den schulischen Leistungen zwischen Schülerinnen und Schülern mit bzw. ohne Zuwanderungshintergrund bestehen (vgl. Schwippert et al., 2007; Stanat, Rauch & Segeritz, 2010). Diese können zumindest teilweise darauf zurückzuführen sein, dass für einen Teil der Schülerinnen und Schüler mit Zuwanderungshintergrund Deutsch die Zweitsprache darstellt und sie daher über weniger Kontakt zur Instruktionssprache verfügen als Peers ohne Zuwanderungshintergrund. Mit dem

Gebrauch einer anderen Sprache als Deutsch innerhalb der Familie könnte ein Mangel an Lerngelegenheiten zum Erwerb der Instruktionssprache im Sinne einer *time-on-task* verbunden sein, der in der Folge die weitere sprachliche und schulische Entwicklung von Schülerinnen und Schülern mit Zuwanderungshintergrund negativ beeinflusst (vgl. August & Shanahan, 2006).

In Deutschland zeigt sich, dass viele der Schülerinnen und Schüler mit Zuwanderungshintergrund Deutsch als Zweitsprache erworben haben. Ergebnisse der Studie „Deutsch-Englisch Schülerleistungen International“ (*DESI*) beispielsweise zeigen, dass 95% der Kinder, die Deutsch als Zweitsprache erwarben einen Zuwanderungshintergrund hatten (Hesse, Göbel & Hartig, 2008). Auch hinsichtlich der innerhalb der Familie gesprochenen Sprache finden sich Unterschiede zwischen Schülerinnen und Schülern mit bzw. ohne Zuwanderungshintergrund. So zeigte sich in der PISA 2012-Studie, dass ein großer Anteil der Jugendlichen mit Zuwanderungshintergrund innerhalb der Familie eine andere Sprache als Deutsch verwendet und somit über weniger Kontakt zur deutschen Sprache verfügt als Jugendliche mit deutscher Familiensprache. Allerdings finden sich dabei in Abhängigkeit vom Generationsstatus (ein Elternteil im Ausland geboren, erste oder zweite Zuwanderergeneration) sowie dem jeweiligen Herkunftsland teilweise große Unterschiede in der Angabe der Familiensprache (Stanat, Pant, Böhme & Richter, 2012). In der PISA 2012-Erhebung beispielsweise gaben nur 36% der Jugendlichen, die der ersten Zuwanderergeneration angehören, also selbst nicht in Deutschland geboren wurden, an, innerhalb der Familie Deutsch zu sprechen. Unter Jugendlichen mit türkischem Zuwanderungshintergrund findet sich hingegen ein besonders hoher Anteil an Jugendlichen, die innerhalb der Familie nicht oder nur wenig Deutsch sprechen: Für Schülerinnen und Schüler, deren Eltern beide zugewandert sind, liegt der Anteil an Jugendlichen, die zuhause vorwiegend oder ausschließlich Türkisch sprechen bei ca. 53% (Stanat, Pant, Böhme & Richter, 2012).

Im Einklang mit der Annahme, dass der Gebrauch einer anderen Sprache als Deutsch innerhalb der Familie mit einem Mangel an Lerngelegenheiten zum Erwerb der deutschen Sprache einhergehen könnte, zeigte sich in nationalen und internationalen Leistungsvergleichen wiederholt, dass die zu Hause gesprochene Sprache einen bedeutsamen Prädiktor für weitere schulische Fähigkeiten, wie etwa die Lesekompetenz, darstellt. So erreichten in der PISA 2009-Erhebung Jugendliche, die in ihrer Familie vorwiegend eine andere Sprache als Deutsch sprachen, unter Kontrolle des ESCS im Schnitt etwa 24 Punkte weniger auf der Lesekompetenzskala als Jugendliche, deren Familiensprache Deutsch war (Stanat, Rauch & Segeritz, 2010). Auch im IQB-Ländervergleich 2011 reduzierten sich die auch unter Kontrolle

des sozioökonomischen Hintergrunds bestehenden Disparitäten den Kompetenzbereich *Zuhören* im Fach Deutsch zwischen Schülerinnen und Schülern der 4. Jahrgangsstufe mit bzw. ohne Zuwanderungshintergrund deutlich, wenn zusätzlich für die innerhalb der Familien gesprochene Sprache kontrolliert wurde (Haag et al., 2012). Ohne Berücksichtigung des sozioökonomischen Hintergrunds ergaben sich beispielsweise für Kinder, deren Eltern beide aus der Türkei zugewandert waren, für das Zuhören Kompetenznachteile von 110 Punkten gegenüber Kindern ohne Zuwanderungshintergrund. Unter Berücksichtigung des HISEI und des elterlichen Bildungsniveaus verringerten sich die Disparitäten auf 75 Punkte. Wurde zusätzlich die Familiensprache der Schülerinnen und Schüler mit Zuwanderungshintergrund kontrolliert, reduzierten sich Leistungsunterschiede zwischen Kindern, deren Eltern beide aus der Türkei zugewandert sind und solchen ohne Zuwanderungshintergrund auf 65 Punkte. Damit zeigen Kinder aus zugewanderten türkischstämmigen Familien gegenüber Kindern ohne Zuwanderungshintergrund allerdings immer noch einen Rückstand im erreichten Kompetenzniveau, der dem erwarteten Leistungszuwachs eines Schuljahres entspricht (Haag et al., 2012).

Diese Befunde deuten darauf hin, dass der sprachliche Hintergrund Disparitäten im Erwerb schulischer Kompetenzen zwischen Schülerinnen und Schülern mit und ohne Zuwanderungshintergrund wesentlich mitbedingt. Obwohl das Aufwachsen mit mehreren Sprachen per se keinen Risikofaktor für den schulischen Kompetenzerwerb darstellt (vgl. z. B. Lindholm-Leary & Borsato, 2006), scheint es vielen Schülerinnen und Schülern, die Deutsch als Zweitsprache erwerben, auch unter Berücksichtigung differenzieller sozioökonomischer Bedingungen nicht zu gelingen, schulische Leistungen zu erzielen, die denen ihrer monolingualen Peers entsprechen. Es kann daher vermutet werden, dass mit dem Erwerb des Deutschen als Zweitsprache, der für viele der Kinder mit Zuwanderungshintergrund der Regelfall ist, spezifische Schwierigkeiten verbunden sind. Im folgenden Abschnitt werden zunächst theoretische Annahmen und Modelle zum Zweitspracherwerb vorgestellt, bevor anschließend das Hörverstehen in einer Erst- und Zweitsprache diskutiert wird.

1.3 Theorien zum Zweitspracherwerb

Umgangssprachlich, aber teilweise auch in der wissenschaftlichen Literatur werden Begriffe wie *Zweitsprache*, *Fremdsprache*, *Bilingualismus* und *Mehrsprachigkeit* zur Kennzeichnung von Personen, die mehrere Sprachen sprechen, häufig synonym und nicht trennscharf verwendet. Eine mögliche Ursache dafür ist, dass die zugrundeliegenden Definitionen einige Ähnlichkeiten und Überlappungen aufweisen, die eine klare Begriffsbestimmung erschweren (Ahrenholz, 2008). Unter *Erstsprache (L1)* wird in der Regel jene Sprache verstanden, die ein

Kind im Zuge seiner sprachlichen Entwicklung als erste erwirbt. Synonym dazu werden auch die Begriffe *Muttersprache* oder *Herkunftssprache* genutzt (Oksaar, 2003). Demgegenüber kennzeichnet die *Zweitsprache (L2)* jene Sprache, die Kinder nach dem Erwerb der Erstsprache erlernen. Abzugrenzen ist der Erwerb einer Zweitsprache dabei vom Fremdspracherwerb. Wesentliches Unterscheidungskriterium für die Differenzierung von Fremd- und Zweitsprache ist der Erwerbskontext: Der Erwerb einer Zweitsprache findet meist als ungesteuerter Spracherwerb innerhalb der Aufnahmegesellschaft statt, in der die Zweitsprache die Verkehrssprache darstellt. Die Aneignung einer Fremdsprache hingegen erfolgt in der Regel in der Herkunftsgesellschaft und wird zum Beispiel durch schulische Instruktionen gesteuert, wie etwa im Rahmen des Englischunterrichts an Schulen (Kniffka & Siebert-Ott, 2009). Kinder aus zugewanderten Familien in Deutschland können demnach als L2-Lernende betrachtet werden, wenn sie die deutsche Sprache im Zuge eines ungesteuerten Spracherwerbs und der jeweiligen Erstsprache nachgeordnet erlernen.

Zur Klassifizierung des Zweitspracherwerbs wird weiter differenziert, ob der Spracherwerb parallel erfolgt oder ob Erst- und Zweitsprache sukzessiv erworben werden. Im Falle eines parallelen, bilingualen Spracherwerbs wird davon ausgegangen, dass ein Kind von frühester Kindheit an beide Sprachen simultan erlernt. Erfolgt der Erwerb einer zweiten Sprache demgegenüber erst nach dem Erwerb der Erstsprache, wird in der Regel nicht von Bilingualität, sondern von dem Erwerb einer Zweitsprache gesprochen. Als Altersgrenze für die Unterscheidung von Bilingualismus bzw. Zweitspracherwerb wird der Beginn des dritten Lebensjahres diskutiert, allerdings besteht in Hinblick auf diese Festlegung kein Konsens (Kniffka & Siebert-Ott, 2009). Neben der auf den Erwerbszeitpunkt bezogenen Unterscheidung zwischen Bilingualität und Zweitspracherwerb wird häufig auch der Aspekt des in den jeweiligen Sprachen erzielten Kompetenzniveaus berücksichtigt. Teilweise wird postuliert, dass Bilingualität nur dann vorliegt, wenn in beiden Sprachen dieselbe Kompetenz erreicht wird, die jener eines muttersprachlichen Lernenden entspricht (vgl. Oksaar, 2003). Darüber hinaus existieren auch Ansätze, Bilingualität nicht unter Rückgriff auf Erwerbszeitpunkt und den erreichten Kompetenzstand, sondern anhand des Gebrauchs der involvierten Sprachen zu definieren. Dieser Ansatz definiert eine Person als bilingual, wenn sie zwei oder mehr Sprachen in ihrem täglichen Leben nutzt (Grosjean, 1998). In der Regel werden die involvierten Sprachen dabei in unterschiedlichen Kontexten genutzt, so dass bilinguale Menschen nur selten beide Sprachen in einem ähnlichen Maße beherrschen. Grosjean (1998) postuliert vor diesem Hintergrund ein Kontinuum zwischen Monolingualität und Bilingualität, auf dem sich mehrsprachige Personen bewegen.

Da keine allgemeingültige Definition für Bilingualität vorliegt, wird in der wissenschaftlichen Forschung wie auch in der Didaktik häufiger auf die Klassifizierung von Lernenden als Erst- oder Zweitsprachlernenden zurückgegriffen. Zum Erwerb einer zweiten Sprache existiert eine Reihe von Theorien, die sich darin unterscheiden, inwiefern sie Interferenzen bzw. Transfereffekte zwischen der Erst- und der Zweitsprache erwarten. Der *Interferenz- oder Kontrastivhypothese* (Fries, 1945; Lado, 1957) zufolge orientiert sich der Erwerb der Zweitsprache an Regularitäten der Erstsprache. Der Sprachaneignungsprozess in der Zweitsprache verläuft demnach aufgrund eines positiven Transfers für jene Strukturen und Regularitäten der Zweitsprache vereinfacht, die sich auch in der Erstsprache finden lassen. Negativer Transfer bzw. Interferenzprozesse entstehen demgegenüber, wenn sich die Strukturen der Erst- und Zweitsprache nicht entsprechen und eine Anwendung der aus der Erstsprache abgeleiteten Regularitäten auf die Zweitsprache zu Fehlern führt.

Empirische Ergebnisse konnten die zentralen Annahmen der Kontrastivhypothese jedoch nicht stützen, sondern sprechen eher dafür, dass Kontrastivfehler nur eine geringe Rolle im Zweitspracherwerb spielen (z. B. Dulay & Burt 1974; Richards, 1971). Aufbauend auf diesen Befunden wurde von Dulay und Burt (1974) die sogenannte *Identitätshypothese* des Zweitspracherwerbs postuliert. Dieser Hypothese zufolge baut der Zweitspracherwerb nicht auf der Erstsprache auf, vielmehr folgen beide Spracherwerbsprozesse demselben Entwicklungsmuster. Zentrale Annahme dieser Hypothese ist die Existenz einer angeborenen, universellen Sprachgrammatik (vgl. Chomsky, 1986), die die Voraussetzung sowohl zum Erwerb der Erst- als auch zum Erwerb der Zweitsprache bildet. Allerdings kann auch die Identitätshypothese als empirisch nicht ausreichend gestützt gelten, da entsprechende Studien vorwiegend auf Analysen einzelner Lerneräußerungen zu Aspekten von Syntax und Morphologie beruhen (vgl. Müller, 2000).

Eine Mittlerposition zwischen den Annahmen der Kontrastiv- und der Identitätshypothese nimmt die *Interlanguage-Hypothese* (Selinker, 1992) ein. Postuliert wird, dass sich im Zuge des Zweitspracherwerbs eine sogenannte Interlanguage bzw. Interimssprache entwickelt. Diese bildet ein sprachliches Übergangssystem, das aus Elementen der Erst- und der Zweitsprache besteht, aber auch eigene Regularitäten aufweist, wie beispielsweise die Auslassung von Präpositionen oder Übergeneralisierungen (Kniffka & Siebert-Ott, 2009). Im Laufe des Zweitspracherwerbs nähert sich diese Interlanguage sukzessive der zu erlernenden Zweitsprache an. Einzelne sprachliche Regularitäten der Interlanguage können jedoch auch bei fortgeschrittenen Zweitsprachlernenden fossilisiert vorliegen, d. h. überdauernd bestehen bleiben und die sprachliche Kompetenz in der Zweitsprache negativ beeinflussen.

Eine weitere einflussreiche Theorie des Zweitspracherwerbs ist die *Interdependenztheorie* bzw. die daraus abgeleitete *Schwellenhypothese* (Cummins, 1979). Zentrale Annahme der Interdependenztheorie ist, dass das erzielte Kompetenzniveau in der Zweitsprache vom Kompetenzniveau in der Erstsprache abhängt, über das ein Lernender zu Beginn des Zweitspracherwerbs verfügt. Sind zum Zeitpunkt des Erwerbsbeginns die erstsprachlichen Kompetenzen nur schwach entwickelt, so ist der Erwerb der Zweitsprache gegenüber L2-Lernenden, die über gute erstsprachliche Kompetenzen verfügen, erschwert (Cummins & Swain, 1986). Die Schwellenhypothese spezifiziert diesen Zusammenhang weiter und postuliert ein Schwellenniveau der erstsprachlichen Fähigkeiten, das erreicht werden muss, um die Zweitsprache erfolgreich erwerben zu können. Differenziert wird dabei zwischen den kommunikativen Grundfähigkeiten, die für die soziale Interaktion im Alltag benötigt werden (*Basic Interpersonal Communication Skills, BICS*) und komplexeren kognitiv-akademischen Sprachfähigkeiten (*Cognitive Academic Language Proficiency, CALP*), die im Zuge der formellen Bildung, z. B. in der Schule erworben werden. Aufgrund der erhöhten Komplexität der CALP gegenüber der BICS geht Cummins (2000) davon aus, dass CALP schwerer zu erwerben ist als BICS es sind, bzw. dass der Erwerb von CALP mehr Zeit benötigt als der Erwerb von BICS. Im Zuge des Zweitspracherwerbs können Cummins (Cummins & Swain, 1986) zufolge nur dann gute Leistungen erzielt werden, wenn ein Kind in der Erstsprache bereits eine Schwelle in Richtung CALP übertreten hat. Der Beherrschung der CALP kommt demnach eine Schlüsselrolle beim Erwerb der Zweitsprache zu. Die Schwellenhypothese hat im wissenschaftlichen Diskurs und in der Forschung zum Zweitspracherwerb eine breite Aufmerksamkeit erhalten. Obwohl die Unterscheidung von BICS und CALP in der Vergangenheit kontrovers diskutiert wurde (vgl. Cummins, 2008 für eine Übersicht einzelner Kritikpunkte), sind die zentralen Annahmen der Hypothese konform mit internationalen Befunden, die zeigen, dass viele Kinder mit Zuwanderungshintergrund in ihrer Zweitsprache über gute alltagssprachliche Kompetenzen verfügen, bei der Beherrschung schulsprachlicher Anforderungen jedoch häufig Schwächen aufweisen (Beck, McKeown & Kucan, 2002; Scarcella, 2003).

Dennoch kann nicht davon ausgegangen werden, dass die Schwellenhypothese im Vergleich der vier betrachteten Modelle den Zweitspracherwerb am treffendsten charakterisiert. So beruhen die betrachteten Hypothesen zum Erwerb zweitsprachlicher Kompetenzen auf unterschiedlichen Perspektiven und sind somit nur eingeschränkt miteinander vergleichbar (Kniffka & Siebert-Ott, 2009). Zudem hängt die Beurteilung der einzelnen Theorien auch von den jeweiligen Erkenntniszielen und Fragestellungen ab (Oksaar, 2003), so dass eine einheitliche Theorie, die die mit dem Erwerb einer Zweitsprache einhergehenden Prozesse umfas-

send beschreibt, derzeit nicht vorliegt. Für die vorliegende Arbeit ist insbesondere die Schwellenhypothese von Bedeutung, da diese Theorie einen Erklärungsansatz dafür bietet, dass Jugendliche mit Zuwanderungshintergrund häufig über gute kommunikative Fähigkeiten im Alltag verfügen, gleichzeitig aber Schwächen in Hinblick auf ihre schriftsprachlichen Leistungen im schulischen Kontext aufweisen (Beck, McKeown & Kucan, 2002; Scarcella, 2003).

2. Hörverstehen in einer Erst- und einer Zweitsprache

Nachdem im vorangegangenen Kapitel theoretische Annahmen zum Zweitspracherwerb vorgestellt wurden, beschreibt der folgende Abschnitt theoretische Modelle und empirische Befunde zur Struktur von Hörverstehenskompetenzen. Hörverstehensfähigkeiten können dabei als eine wesentliche Grundlage des Erwerbs weiterer sprachlicher und schulischer Fähigkeiten betrachtet werden. Rost (1994) verdeutlicht die Notwendigkeit gut ausgebildeter Hörverstehensfähigkeiten für den weiteren Wissenserwerb wie folgt: „Without understandable input at the right level, any learning simply cannot begin“ (S. 141). Dies wird umso deutlicher, betrachtet man den Anteil des Hörverstehens an der Kommunikation im schulischen Alltag: Im Mittel wird in 50% bis 70% der Unterrichtszeit von Schülerinnen und Schülern Zuhören erwartet (Imhof, 2004). Für Lernende einer Zweitsprache ist es daher von zentraler Bedeutung, zunächst gute Hörverstehensfähigkeiten in der Zweitsprache zu entwickeln, um darauf aufbauend weitere Kompetenzen erwerben zu können, indem z. B. neue Wörter aus Hörtexten erlernt werden (Richards, 2005; Vidal, 2003).

Im Folgenden wird zunächst das Konstrukt Hörverstehen definiert sowie theoretische Modelle zum Erwerb von Hörverstehenskompetenzen in einer Erst- und einer Zweitsprache dargestellt. Dabei werden zunächst generelle Definitionen und Konzeptionen von Hörverstehen beschrieben, bevor Klassifikationen von für das Hörverstehen notwendigen Fähigkeiten sowie spezifische Hörverstehensmodelle expliziert werden. Da bislang keines der Klassifikationsschemata oder die Annahmen spezifischer Hörverstehensmodelle empirisch validiert wurden, werden abschließend empirisch abgesicherte Modelle des Textverstehens dargestellt, die sowohl auf das Hörverstehen als auch auf das Leseverstehen anwendbar sind.

2.1 Konzeptionen des Hörverstehens

Für eine Konzeption des Konstrukts Hörverstehen ist es zunächst notwendig, Hörverstehen bzw. Zuhören vom reinen Hören abzugrenzen. Umgangssprachlich werden die Termini *hören* und *zuhören* häufig synonym verwendet. Im wissenschaftlichen Sinn bezeichnet *hören* jedoch

lediglich den physiologischen Prozess der Verarbeitung von Geräuschen, d. h. die Wahrnehmung von Schallwellen sowie deren Verarbeitung im sensorischen Sprachzentrum des Gehirns. Die Begriffe *zuhören* bzw. *Hörverstehen* bezeichnen demgegenüber einen willentlich gesteuerten, psychologischen Prozess der Bedeutungskonstruktion, bei dem die eingehenden Informationen selektiv verarbeitet werden (Rost, 2011). Während die Fähigkeit zum Hören bei den meisten Individuen von Geburt an vorhanden ist, entwickelt sich die Fähigkeit zum Zuhören erst im Lauf der sprachlichen Entwicklung eines Kindes: “Humans are born with an innate ability to hear. However, hearing is not listening, and many factors need to be considered in the development of effective listening skills” (Flowerdew & Miller, 2005, S. 28).

In der Vergangenheit wurde Hörverstehen meist als einfache Signalübertragung einer auditiven Nachricht von einem Sender zu einem Empfänger beschrieben. Innerhalb solcher sogenannter *information processing*-Modelle besteht die Rolle des Hörers lediglich darin, die vom Sender übermittelte Nachricht zu empfangen (vgl. Abbildung 1). Die Bedeutung eines Hörtextes liegt nur als seitens des Senders sprachlich kodierte Information vor, so dass Hörverstehen als passive Fähigkeit angesehen wird, bei der der Hörer keine aktive Verarbeitung der übermittelten Informationen vornimmt (Rost, 1990; Solmecke, 1993).

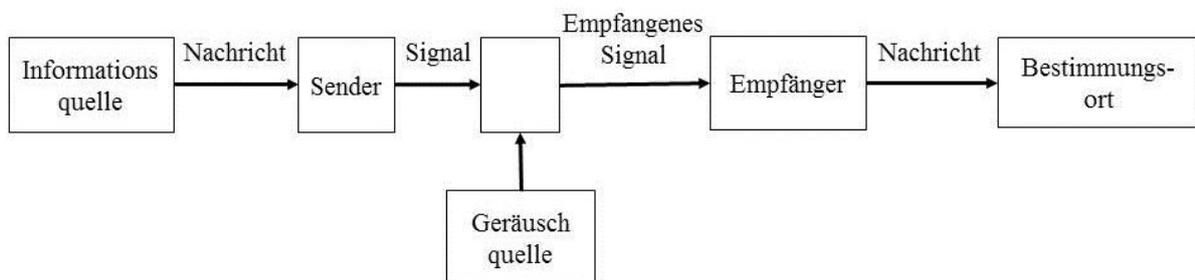


Abbildung 1. Ein einfaches *information processing*-Modell des Hörverstehens (Rost, 1990)

Seit den 1970er Jahren entwickelten neuere Verstehensmodelle hingegen definieren sprachliches Verstehen als einen interaktiven und komplexen Prozess, der neurologische, linguistische, semantische und pragmatische Verarbeitungsmechanismen seitens des Hörers integriert (Rost, 2011). Die *International Listening Association* beispielsweise definiert Hörverstehen als „the process of receiving, constructing meaning from, and responding to spoken and/or nonverbal messages“ (Purdy, 1997, S. 6). Auch andere der zahlreichen Definitionen des Hörverstehens betonen die Notwendigkeit einer aktiven Verarbeitung des Hörinhalts durch den Hörenden. Hörverstehen wird dabei zum Beispiel definiert als

- „the process of receiving, attending to, and assigning meaning to aural and visual stimuli“ (Wolvin & Coakley, 1996, S. 69)
- „hearing, understanding, remembering, interpreting, evaluating, and responding“ (Brownell, 2002, S. 18)
- „an active and conscious process in which the listener constructs meaning by using cues from contextual information and from existing knowledge, while relying upon multiple strategic resources to fulfill the task requirement“ (O'Malley, Chamot & Küpper, 1989, S. 434).

Diesen Definitionen ist gemeinsam, dass dem Hörenden eine aktive Rolle zugeschrieben wird, innerhalb derer er versucht, unter Verarbeitung linguistischen und kontextuellen Wissens den Bedeutungsgehalt einer eingehenden Sprachnachricht zu extrahieren. Die Fähigkeit, neu eingehende Informationen mit bereits bestehendem Wissen abzugleichen und effektiv integrieren zu können, ist dabei aufgrund des flüchtigen Charakters einer Sprachnachricht von zentraler Bedeutung für das Gelingen des Hörverstehensprozesses (Rost, 2005). Im Idealfall läuft die Informationsintegration hoch automatisiert und mit geringer bewusster Aufmerksamkeit des Hörenden ab, so dass ausreichend Gedächtniskapazität zur Verfügung steht, um den weiterhin eingehenden Sprachinput verarbeiten zu können (Chiu, McBride-Chang & Lin, 2012; Hogan, Adlof & Alonzo, 2014).

Obgleich diese zentralen Annahmen zu Hörverstehensprozessen in der Forschung häufig vertreten werden, liegt eine allgemeingültige Konzeptualisierung des Hörverstehens derzeit nicht vor. Es sind jedoch zahlreiche theoretische Modelle des Hörverstehens entwickelt worden, die die Prozesse der Informationsintegration im Zuge des auditiven Textverstehens beschreiben. Darüber hinaus liegt eine Reihe von Klassifikationsschemata vor, die jene Fähigkeiten auflisten, die hörer- bzw. textseitig das Gelingen des Hörverstehensprozesses bedingen. Im Folgenden werden zunächst entsprechende Klassifikationen von Prädiktoren des Hörverstehens und anschließend Prozessmodelle des Hörverstehens vorgestellt.

2.2 Klassifikationen von Prädiktoren des Hörverstehens

Bereits 1971 veröffentlichte Lundsteen eine Klassifikation von dem Hörverstehen zugrundeliegenden Faktoren, in der zunächst zentrale Voraussetzungen für gelingendes Hörverstehen aufgeführt werden, wie beispielsweise eine adäquate Gedächtnisspanne oder die Fähigkeit, Pausen, Rhythmus und Ton innerhalb einer Nachricht zu identifizieren. Die eigentlichen Hörverstehensfähigkeiten sind bei Lundsteen (1971) hierarchisch angeordnet und lassen sich in generelle Hörverstehensfähigkeiten (z. B. sequentielle Ordnung der Informationen eines Hör-

textes erinnern) und sogenannte kritische Hörverstehensfähigkeiten (z. B. die Stimmung des Sprechers identifizieren) unterteilen. In ähnlicher Weise definieren Wolvin und Coakley (1979) eine hierarchisch angelegte Klassifikation von spezifischen Fähigkeiten, die für unterschiedliche Hörverstehensarten benötigt werden, zum Beispiel für *diskriminierendes Zuhören*, bei dem im Hörtext enthaltene Fakten von subjektiven Meinungen unterschieden werden müssen oder für *verstehendes Zuhören*, bei dem die im Hörtext enthaltene Nachricht verstanden werden muss. Ähnlich wie Lundsteen (1971) beschreiben somit auch Wolvin und Coakley (1979) eine Reihe von Fähigkeiten, die dem Hörverstehen zugrunde liegen. Der Fokus beider Klassifikationen liegt allerdings auf der Definition unterschiedlicher Arten des Zuhörens bzw. verschiedener Hörziele und die zur Erreichung dieser Ziele notwendigen Strategien und Prozesse. Merkmale des Hörers, die den Hörverstehensprozess beeinflussen, spielen in beiden Taxonomien hingegen eine weniger bedeutsame Rolle.

Eine weitere Klassifikation von Hörverstehensfähigkeiten, die dem interaktiven Charakter des Hörverstehensprozesses Rechnung trägt, indem neben Hörermerkmalen auch Aspekte des Hörtextes, des Senders, des Kontextes sowie weiterer Faktoren berücksichtigt werden, stammt von Samuels (1984). Im Rahmen dieses Klassifikationsschemas wird zwischen sogenannten *inside-the-head*- und *outside-the-head*-Faktoren unterschieden. Unter *inside-the-head*-Faktoren versteht Samuels (1984) jene Prädiktoren, die den Verstehensprozess hörerseitig beeinflussen, wie etwa Intelligenz, sprachliche Fähigkeiten, die Verfügbarkeit von Hintergrundwissen oder die Verfügbarkeit von metakognitiven Hörstrategien. *Outside-the-head*-Faktoren beschreiben demgegenüber äußere und vom Hörenden nicht beeinflussbare Rahmenbedingungen, wie etwa das Thema eines Hörtextes oder die Klarheit des sprachlichen Inputs. Obwohl Samuels (1984) explizit den interaktiven Charakter des Hörverstehens betont, werden keine Angaben dazu gemacht, in welcher Weise Faktoren *inside-the-head* und Faktoren *outside-the-head* interagieren.

Neben den beschriebenen Klassifikationen findet sich eine Vielzahl weiterer Auflistungen von das Hörverstehen beeinflussenden Faktoren (vgl. auch Grotjahn, 2012 für die Angabe weiterer Literatur). Wichtige Prädiktoren des Hörverstehens, die in den meisten dieser Klassifikationen genannt werden, sind beispielsweise Akzent und Dialekt des Sprechers, Prosodie, Sprechgeschwindigkeit, Zahl der Sprecher, Pausen im Sprachfluss, Häufigkeit der Textdarbietung oder die konzeptuelle Schwierigkeit des Textinhalts (Grotjahn, 2012). Häufig wird dabei, ähnlich wie bei Samuels (1984), zwischen hörer- und textseitigen Prädiktoren des Hörverstehens unterschieden, aber nur selten finden sich explizite Annahmen dazu, in welcher Weise die einzelnen Faktoren zum Aufbau des Textverstehens beitragen und wie sie miteinander

interagieren. Zudem ist die überwiegende Zahl der Klassifikationen nicht empirisch überprüft, so dass ihre Gültigkeit unklar bleibt. Deskriptive Auflistungen von Faktoren, die den Hörverstehensprozess beeinflussen, liefern somit zwar Ansatzpunkte für die Identifikation von Merkmalen, die das Hörverstehen unterstützen bzw. hemmen können. Sie bilden jedoch keine überprüfbare Theorie des Hörverstehens und sind nicht ausreichend, um den komplexen interaktiven Prozess des Hörverstehens zu beschreiben.

2.3 Kognitive Prozessmodelle des Hörverstehens

Neben deskriptiv ausgerichteten Klassifikationen der für das Hörverstehen bedeutsamen Faktoren liegt eine Reihe von Modellen vor, die den Prozess des Hörverstehens beschreiben. Die meisten dieser Modelle nehmen dabei einen mehrstufigen Prozess an, bei dem verschiedene Phasen der Informationsverarbeitung unterschieden werden können. So postuliert etwa Dunkel (1986), dass beim Hörverstehen Strategien angewandt werden, die sich in die drei Phasen *Vorbereitung des Zuhörens*, *Zuhören* und *Nachbereitung des Zuhörens* unterteilen lassen. Auch Imhof (1998) geht anhand einer Analyse von Hörprotokollen davon aus, dass der Hörverstehensprozess durch die Anwendung spezifischer Strategien vor, während und nach dem Zuhören gekennzeichnet ist. Empirische Befunde, die diese Annahme unterstützen, stammen von Stein (1999). In ihrer Studie wurden Hörende aufgefordert, ihre Gedanken und Gefühle während des Zuhörens zu beschreiben. Anhand einer Analyse der angefertigten Hörprotokolle konnte Stein (1999) die affektiven und kognitiven Prozesse erfolgreichen Zuhörens beschreiben. Wie bereits von Dunkel (1986) und Imhof (1998) postuliert, lässt sich erfolgreiches Hörverstehen demnach in die drei Phasen *Vorbereitung* (z. B. Hörziele definieren), *Zuhören* (z. B. den Verstehensprozess kognitiv überwachen, Notizen anfertigen) und *Nachbereitung* (z. B. Nachfragen, retrospektive Evaluation des Hörinhalts) unterteilen. Allerdings bleibt in den genannten Modellen unklar, welche kognitiven Prozesse an den einzelnen Phasen des Hörverstehensprozesses beteiligt sind.

Neben Modellen, die den Prozess des Hörverstehens anhand der in den jeweiligen Prozessphasen verwendeten Strategien beschreiben, existieren auch Hörverstehensmodelle, die die am Zuhören beteiligten kognitiven Prozesse beschreiben. Einfache kognitive Modelle sehen das Hörverstehen dabei als zweistufigen Prozess, der hierarchieniedrige (*lower-level*) und hierarchiehöhere (*higher-level*) Verarbeitungsprozesse umfasst (vgl. Buck, 1991). Field (2008) beispielsweise unterscheidet zwischen Dekodier- und Bedeutungsgewinnungsprozessen, die sich auf unterschiedlichen Ebenen des Hörverstehens abspielen. *Dekodierprozesse* laufen dabei auf der Ebene von Phonemen, Silben oder der Syntax ab und betreffen

beispielsweise die Identifizierung von Konsonanten oder Vokalen oder die Ermittlung von Wort- und Satzgrenzen. Vorrangige Aufgabe des Hörers auf dieser unteren Verarbeitungsebene ist die Übersetzung der eingehenden Sprachsignale in Phoneme, Wörter und Sätze, um ein wörtliches Verständnis des Texts zu entwickeln (Field, 2008). Anschließend laufen auf einer hierarchiehöheren Stufe eine Vielzahl von Prozessen ab, die der *Bedeutungsgewinnung* dienen. Beispiele hierfür sind der Einbezug von Hintergrundwissen, die Bildung von Inferenzen über den Textinhalt, das Erkennen von redundanten Informationen oder die Überprüfung des bisher Gesagten auf Konsistenz.

Neben diesen einfachen Modellen existieren komplexere Modelle, in denen Hörverstehen als mehrstufiger Prozess mit einer ganzen Reihe von Einflussfaktoren gesehen wird. Ein solches Modell wurde von Brownell (1994) postuliert. Dem sogenannten *HURIER-Modell* zufolge besteht Hörverstehen aus sechs Komponenten, die Brownell (1994) mit *Hearing* (Konzentration auf die Nachricht), *Understanding* (Verständnis der wörtlichen Bedeutung einer Nachricht), *Remembering* (Erinnern der Nachricht), *Interpreting* (Beachtung nonverbaler und kontextueller Aspekte der Nachricht), *Evaluating* (Bestimmung der Wertigkeit der Nachricht) und *Responding* (Auswahl einer angemessenen Antwort auf die Nachricht) beschreibt. Im Zuge des Hörverstehens laufen diese Prozesse nicht linear oder chronologisch ab, sondern interagieren während der Verarbeitung einer eingehenden Sprachnachricht miteinander und werden dabei von einer Reihe von Hintergrundfaktoren, wie etwa früheren Erfahrungen oder den Werten und Haltungen des Hörers, beeinflusst.

Ein weiteres hörverstehensspezifisches Modell des Textverstehens stammt von Imhof (2004). Zuhören wird in diesem sogenannten *SOI-Modell* als ein mehrstufiger Prozess definiert, der die intentionale *Selektion*, *Organisation* und *Integration* verbaler und nonverbaler akustischer Informationen erfordert. Es wird angenommen, dass als Ausgangspunkt für das Zuhören zunächst eine Zuhörabsicht ausgebildet werden muss, die bestimmt, welche der zahlreichen eingehenden akustischen Informationen potentiell relevant und welche vernachlässigbar sind. Im Zuge des Zuhörens werden diese relevanten Informationen zunächst ausgewählt und dann organisiert, d. h. sortiert und interpretiert. Abschließend wird der Hörinhalt im Langzeitgedächtnis abgespeichert und mit relevantem Vorwissen abgeglichen. Imhof (2004) zufolge zeichnet sich ein kompetenter Zuhörer dadurch aus, dass er in der Lage ist, die zum Hörverstehen notwendigen Prozesse eigenständig und effektiv zu steuern.

Weitere kognitive Modelle, die den Prozess des Hörverstehens beschreiben, stammen beispielsweise von Nagle und Sanders (1986), Steil, Barker und Watson (1983) oder Solmecke (1992). Bislang bleibt allerdings unklar, inwiefern die theoretisch postulierten Prozesse

der Verarbeitung gesprochener Sprache sich tatsächlich voneinander abgrenzen lassen. Zudem machen diese Modelle des Hörverstehens keine Angaben dazu, inwiefern Prozesse der unterschiedlichen Verarbeitungsebenen miteinander interagieren. Dies führt zu einer gewissen „bottom-up-primacy“ von Hörverstehensmodellen (vgl. Wilson, 2003), bei der die Verarbeitung eines Hörtextes als von den grundlegenden Verarbeitungsmechanismen (z. B. Identifikation einzelner Laute) aus zu höherwertigen Verarbeitungsmechanismen verlaufend betrachtet wird. Prozesse, bei denen Schritte der hierarchiehöheren Verarbeitungsebenen die darunterliegenden beeinflussen, etwa wenn der Kontext die Erkennung eines Wortes erleichtert (vgl. Tyler & Marslen-Wilson, 1977), werden in diesen Hörverstehensmodellen demgegenüber weniger stark betont. Hörverstehen ist jedoch ein dynamischer Prozess, in dessen Zuge sowohl Prozesse hierarchieniedriger Verarbeitungsebenen als auch Prozesse höherer Verarbeitungsebenen miteinander interagieren (Buck, 2001; Field, 2008; Flowerdew & Miller, 2005b). Vor diesen Hintergrund erscheint es angebracht, Hörverstehen anhand interaktionistischer Modelle des Textverstehens zu beschreiben, die empirisch bereits einige Unterstützung erfahren haben.

2.4 Interaktionistische Modelle des Textverstehens

Hör- und Leseverstehen sind eng miteinander verknüpft und beruhen im Wesentlichen auf ähnlichen Verarbeitungsmechanismen. Entsprechend liegen interaktionistische Modelle der Sprachverarbeitung vor, die das Textverstehen unabhängig von der Modalität beschreiben und sowohl auf das Hör- wie auch auf das Leseverstehen anwendbar sind. Diese Modelle gehen davon aus, dass verschiedene Ebenen des Textverstehens miteinander interagieren, um die Bildung einer mentalen Repräsentation der Inhalte von auditiv oder visuell dargebotenen Texten zu ermöglichen. Ein solches interaktionistisches Modell der Sprachverarbeitung stammt von Anderson (1995). Er differenziert sprachliches Verstehen in drei miteinander verbundene und rekursive Prozesse: das *perceptual processing*, die Wahrnehmung der sprachlichen Informationen, das *parsing*, eine semantische und syntaktische Analyse des Inhalts sowie die *utilization*, die Verwendung dieser Informationen, z. B. zur Bildung von Inferenzen über zentrale Aussagen des Hörinhalts (vgl. Abbildung 2). Während des Hörverstehensprozesses können diese Prozesse ineinander übergehen oder anhand der Analysen aus vorhergehenden bzw. folgenden Verarbeitungsschritten modifiziert werden (Kurita, 2012). So können beispielsweise hierarchieniedrige Prozesse durch hierarchiehöhere Prozesse beeinflusst werden, etwa wenn aufgrund des Satzkontextes die Bedeutung eines einzelnen Wortes erschlossen wird. Umgekehrt können die Ergebnisse der hierarchieniedrigen Prozesse

die weitere Verarbeitung des Hörinhalts determinieren, etwa, wenn es aufgrund eines fälschlich identifizierten Wortes zu einer Fehlinterpretation eines Satzes kommt.

Obwohl die zentrale Annahme dreier unterscheidbarer Stufen des Textverstehens empirische Unterstützung erfahren hat (vgl. Christmann & Groeben, 1999; Goh, 2000), ist das von Anderson (1995) postulierte Modell in zweierlei Hinsicht unzureichend: Erstens bezieht sich das Modell nur auf das Verstehen einzelner Sätze. Es ist daher auf das Verstehen ganzer Texte nicht unbedingt anwendbar, da es keine Aussagen zur Bildung einer globalen Kohärenz eines Textes erlaubt.

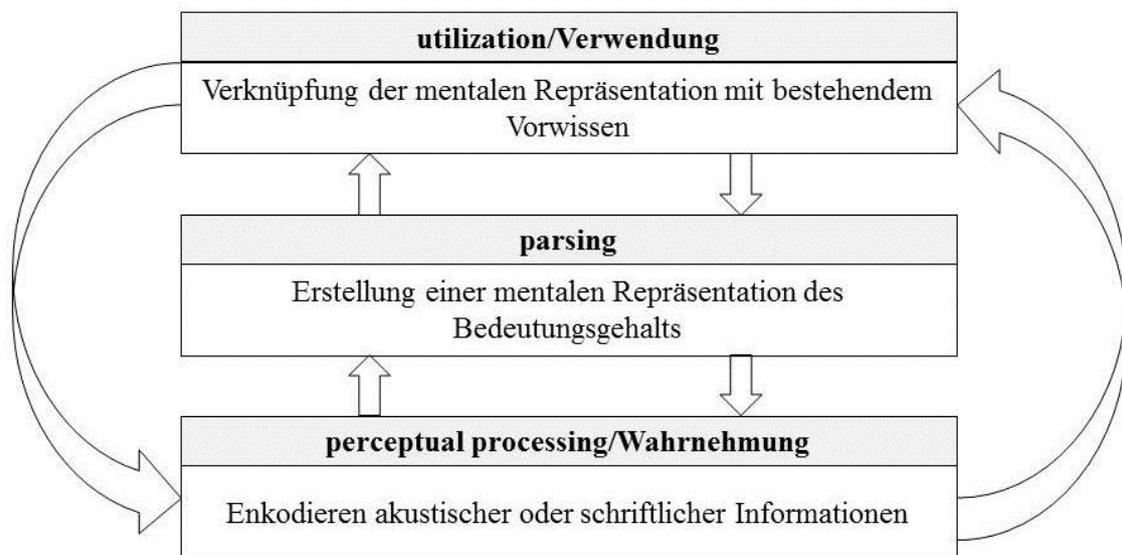


Abbildung 2. Das interaktionistische Modell des Satzverstehens nach Anderson (1995)

Zweitens bezieht sich die von Anderson (1995) beschriebene Satzrepräsentation stark auf die im Text enthaltenen Informationen, aus denen dann eine Repräsentation der Inhalte erstellt wird. Zwar geht auch Anderson (1995) davon aus, dass zum Verstehen eines Satzes Inferenzen nötig sind, die über das in einem Text Gesagte hinausgehen, d. h. Elaborationen gebildet werden müssen. Allerdings macht er keine Angaben dazu, inwiefern diese dazu genutzt werden, eine bedeutungsbasierte Repräsentation des Satzinhalts zu erstellen. Tatsächlich scheinen jedoch nicht nur die im Zuge des Parsings erschlossenen syntaktischen und semantischen Inhalte einer Sprachbotschaft im Gedächtnis abgespeichert zu werden, sondern vor allen Dingen die im Text beschriebenen Situationen, d. h. ein abstrahiertes Modell der im Text beschriebenen Sachverhalte (Bransford & Franks, 1971). Mentale Modelle von Situationen, sogenannten Situationsmodelle, bilden demnach einen wesentlichen Bestandteil des Textverstehens (Johnson-Laird, 1983).

Ein Modell des Textverstehens, das die Rolle und das Zustandekommen von Situationsmodellen expliziert, stammt von Kintsch und van Dijk (1978). Wie bei Anderson (1995) werden auch hier bei der Erstellung eines mentalen Textmodells drei Ebenen der Bedeutungsrepräsentation unterschieden (vgl. Abbildung 3). Auf der Ebene der Oberflächenrepräsentation werden zunächst der genaue Wortlaut und der Satzbau eines Textes gespeichert. Hier erfolgt die Enkodierung sprachlicher Informationen, d. h. die Segmentierung einzelner Phoneme und Wörter aus dem kontinuierlichen Sprachstrom.

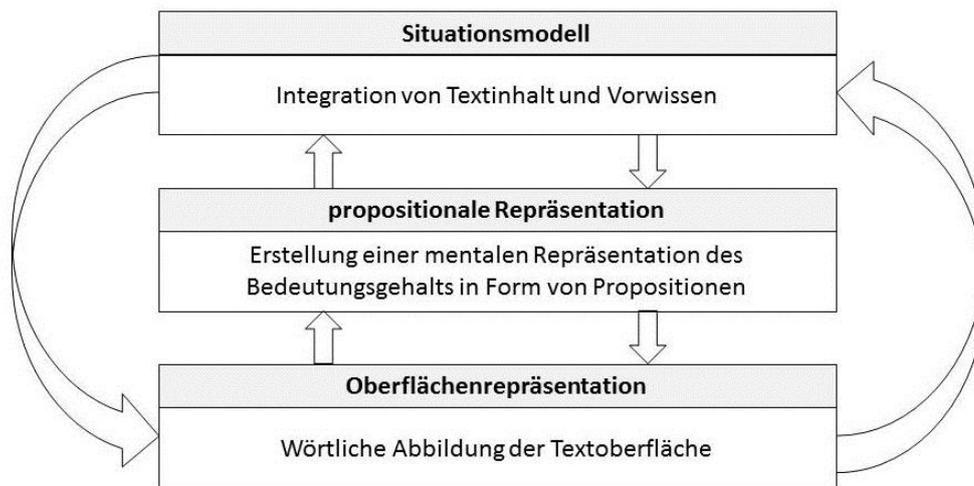


Abbildung 3. Das interaktionistische Modell des Textverstehens nach Kintsch & van Dijk (1978)

Die anschließenden beiden Verarbeitungsphasen werden im sogenannten *Construction-Integration-Modell* (Kintsch, 1988; 1998) weiter spezifiziert. Zentrale Annahme dieses Modells ist, dass beim Textverstehen zunächst der Bedeutungsgehalt des Textes in propositionaler Form extrahiert wird. Propositionen stellen dabei Repräsentationen des Bedeutungsgehalts einzelner Sätze dar und bestehen aus einem Prädikat, das die grundlegende Struktur der Argumente organisiert und einem oder mehreren Argumenten, die wiedergeben, worüber etwas gesagt wird. In der vom Construction-Integration-Modell postulierten Konstruktionsphase wird zunächst ein kohärentes propositionales Netzwerk des Textinhalts aufgebaut, in dem die einzelnen Propositionen miteinander verbunden sind. Der genaue Wortlaut eines Textes ist auf dieser Ebene nicht länger von Bedeutung. Entsprechend ist es für Hörende in dieser Phase wichtig, sich nicht länger auf einzelne Wörter zu konzentrieren, sondern die Aufmerksamkeit auf Phrasen oder ganze Sätze zu richten, wobei z. B. Intonation oder Pausen im Sprachfluss zur Orientierung innerhalb des Sprachflusses genutzt werden können (Kurita, 2012; O'Malley et al., 1989). Auf dieser Ebene werden zunächst sogenannte Mikrostrukturen

ausgebildet, die aus jenen Propositionen erstellt werden, die aus dem Text extrahiert werden können sowie zugehörigen Assoziationen aus dem Hintergrundwissen.

In der anschließenden Integrationsphase erfolgt schließlich die Erstellung einer mentalen Repräsentation der im Text dargestellten Umstände und Ereignisse, die durch relevantes Hintergrundwissen angereichert ist. In der Folge entsteht ein mentales Situationsmodell des Textes, in dem nicht länger die propositionale Bedeutung des Textes selbst, sondern der wesentliche Inhalt bzw. Sachverhalt repräsentiert ist. Zur Erstellung dieses Modells werden vom Leser oder Hörer aufgrund bereits bestehenden Wissens Assoziationen zu den Textinhalten gebildet. Dabei müssen passende Konzepte und Schemata aktiviert und weniger passende Konzepte deaktiviert werden, bis eine schlüssige Textrepräsentation entsteht, die keine widersprüchlichen Informationen oder unnötigen Propositionen mehr enthält. Diese Textrepräsentation wird anschließend ins Langzeitgedächtnis überführt (Zwaan & Singer, 2003).

Erfolgreiches Textverstehen ist im Rahmen der Modelle von Kintsch und van Dijk (1978) und Anderson (1995) somit ein Resultat von Verarbeitungsprozessen, die sowohl *bottom-up*, d. h. ausgehend von der Perzeption von Lauten hin zu größeren sprachlichen Einheiten wie Silben, Lexemen oder Sätzen, als auch *top-down*, also von höheren zu basaleren Verarbeitungsebenen, verlaufen (Chiu et al., 2012; Kurita, 2012). Bottom-up-Prozesse werden dabei von linguistischen Merkmalen einer Äußerung, wie Phonologie, Lexik, Syntax und Semantik ausgelöst und ermöglichen durch eine Analyse des semantischen und syntaktischen Gehalts einer Nachricht den Aufbau einer propositionalen bzw. situativen Repräsentation des Textinhalts. Top-down-Prozesse finden demgegenüber statt, wenn die Erstellung der Oberflächenrepräsentation oder der propositionalen Repräsentation durch die Resultate der Verarbeitung auf höheren Ebenen des Textverstehens beeinflusst wird. Beispielsweise wird bereits auf der untersten Verarbeitungsebene, d. h. auf der Ebene der Oberflächenrepräsentation, die Wahrnehmung des Hörinhalts durch Einflüsse aus höheren Verarbeitungsebenen beeinflusst, etwa wenn Lücken im Sprachfluss unter Zuhilfenahme eines Situationsmodells gefüllt werden (Warren & Warren, 1970). Hörverstehen kann somit als Resultat einer Wechselwirkung zwischen bottom-up und top-down verlaufenden Verarbeitungsprozessen angesehen werden, bei dem einerseits situative Elemente bzw. Textelemente gespeichertes Vorwissen aktivieren, dieses Vorwissen aber seinerseits die Einordnung und Verknüpfung von Textinformationen steuert (vgl. Kurita, 2012; Paschke, 2000). Nur wenn es dem Hörer gelingt, diese Prozesse zu integrieren, ist es möglich, ein Situationsmodell des Hörtextes aufzubauen und einen Hörtext zu verstehen.

Zahlreiche empirische Studien konnten die zentralen Annahmen zum Zusammenwirken von bottom-up- und top-down gesteuerten Verarbeitungsstrategien bei der Erstellung eines Situationsmodelles stützen (z. B. Goh, 2000; O'Malley et al., 1989; van Dijk, & Kintsch, 1983; Zwaan, 1993). Interaktionistische Modelle des Textverstehens scheinen somit gut geeignet, um den komplexen Prozess der Bildung mentaler Repräsentationen im Zuge des Hörverstehens zu beschreiben.

2.5 Hörverstehen in einer Zweitsprache

Wie andere Textverstehensprozesse kann auch das Hörverstehen als mentale Konstruktion des Bedeutungsgehalts einer Nachricht verstanden werden. Schülerinnen und Schüler in einer Zweitsprache begegnen im Zuge dieses Prozesses jedoch häufig spezifischen Herausforderungen, die ihr Hörverstehen beeinträchtigen können. Entsprechend erzielten L2-Lernende auch unter Kontrolle möglicher konfundierender Variablen, wie etwa dem sozioökonomischen Status, häufig geringere Hörverstehensleistungen als ihre monolingualen Peers (z. B. Eckhardt, 2008; Heppt, Stanat, Dragon, Berendes & Weinert, 2014; Stanat, Pant, Böhme & Richter, 2012). Hörverstehen in einer Zweitsprache scheint demnach für L2-Hörerinnen und -Hörer mit besonderen Schwierigkeiten zu verbunden sein.

Grundsätzlich unterscheidet sich der Prozess des Hörverstehens in einer Zweitsprache nicht vom erstsprachlichen Hörverstehen. Modelle des Textverstehens, die für Lernende einer Erstsprache entwickelt wurden, scheinen prinzipiell auch auf Lernende einer Zweitsprache anwendbar (Færch & Kasper, 1986). Empirische Hinweise auf die Anwendbarkeit des Textverarbeitungsmodells von Anderson (1995) in einer Zweitsprache liefern beispielsweise qualitative Studien von Goh (2000) und O'Malley et al. (1989). In beiden Studien wurden L2-Lernende während des Hörens eines Textes in regelmäßigen Abständen unterbrochen und nach ihren aktuellen Gedanken gefragt. Die Ergebnisse zeigen, dass sich auch beim L2-Hörverstehen die drei Prozess- bzw. Verarbeitungsebenen nach Anderson (1995) bzw. Kintsch und van Dijk (1978) finden lassen, der Hörverstehensprozess in einer Zweitsprache also offenbar ähnlich dem in einer Erstsprache abläuft. Buck (2001, S. 51) konstatiert entsprechend: "I believe that the difference between first- and second-language listening is not that the processes are fundamentally different in any way, but only that the knowledge necessary to understand is often grossly inadequate for the second-language listener."

Seit den 1980er Jahren wurde eine Vielzahl an Faktoren identifiziert, die das Hörverstehen in einer Zweitsprache beeinflussen und ggf. erschweren können. Darauf aufbauend wurde eine Reihe an Klassifikationen erstellt, die die für das Hörverstehen in einer Zweit-

sprache potentiell bedeutsamen Faktoren beschreiben, wie etwa Aspekte des Textes (z. B. Pausen, Sprecherwechsel) oder hörerseitige Faktoren (z. B. die Anwendung geeigneter Verstehensstrategien) (vgl. z. B. Lund, 1990; Rubin, 1994). Ein zusammenfassender Überblick über diese Klassifikationen von für das L2-Hörverstehen benötigten Faktoren findet sich bei Bloomfield et al. (2010). Allerdings kann angenommen werden, dass, wie von Buck (2001) postuliert, ein Großteil dieser Faktoren nicht spezifisch verstehensförderlich bzw. -hemmend für das L2-Hörverstehen wirkt, sondern auch für die erstsprachliche Hörverstehensleistung von Bedeutung ist. Entsprechend ähneln dann auch die für Lernende einer Zweitsprache aufgestellten Klassifikationen größtenteils den für das erstsprachliche Hörverstehen postulierten. Es kann demnach vermutet werden, dass sich Unterschiede in der Hörverstehensleistung von L1- und L2-Schülerinnen und -Schülern nicht primär aus unterschiedlichen Prädiktoren des Hörverstehens ergeben, sondern vielmehr aus der differenziellen Ausprägung oder unterschiedlichen Bedeutsamkeit dieser Prädiktoren.

Vielfach ist vermutet worden, dass L2-Lernende aufgrund ihrer schwächeren sprachlichen Fähigkeiten dem linguistischen Gehalt einer Hörmacht verstärkt Aufmerksamkeit zukommen lassen müssen (Anderson, & Lynch, 1988; Færch & Kasper, 1986; Voss, 1984). Einige Studien weisen darauf hin, dass L2-Lernende mit geringen zweitsprachlichen Kompetenzen bei der Verarbeitung von Hörtexten auf semantische und lexikalische Eigenschaften des Hörtextes fokussieren und beispielsweise versuchen, einen Text Wort für Wort zu übersetzen (Lynch, 1998; Rubin, 1994, Vandergrift, 2003). Dieser Fokus auf die bottom-up-Verarbeitung kann allerdings das Verständnis des Hörtextes erschweren, da weniger Aufmerksamkeit für die Verarbeitung des Inhalts zur Verfügung steht. Rost (2011, S. 133) gibt an: „When a learner attends to the form of the message, this attention to form competes for the processing capacity in short-term memory that is available to attend to content.“

Andere Studien hingegen zeigen, dass L2-Lernende auch vermehrt Hintergrundwissen, d. h. top-down-Strategien, nutzen, um für Lücken im Verständnis des Sprachinputs, d. h. der bottom-up-Verarbeitung zu kompensieren (Goh, 2000; Tyler, 2001; Voss, 1984). Die Nutzung von top-down-Strategien zur Füllung von Verständnislücken auf den unteren Hierarchieebenen des Hörverstehensprozesses findet sich zwar auch bei L1-Lernenden (Goh, 1998). Allerdings wird angenommen, dass L2-Lernende diese Strategien in vermehrtem Maße nutzen, da sie weniger vertraut mit der zweitsprachlichen Phonologie, Syntax und dem Wortschatz sind, die für eine bottom-up-Verarbeitung des Hörinhalts nötig wären (Voss, 1984). Dieser Rückgriff auf top-down Strategien kann problematisch sein, da bei L2-Lernenden damit gerechnet werden muss, dass sie über weniger kulturspezifisches Welt- sowie Diskurswis-

sen verfügen, so dass eine hauptsächlich top-down geleitete Verarbeitung häufig wenig erfolgreich ist (Field, 2008; Paschke, 2000).

Eine Reihe von Studien liefert empirische Evidenz, die die Annahme einer unterschiedlichen Nutzung von top-down bzw. bottom-up verlaufenden Hörstrategien in Abhängigkeit von der Zweitsprachkompetenz (z. B. Goh, 2000; O'Malley, Chamot & Küpper, 1989; Shohamy & Inbar, 1991; Tsui & Fullilove, 1998). Allerdings findet sich in keiner der genannten Studien eine L1-Vergleichsgruppe und nur O'Malley et al. (1989) machen Angaben dazu, wie die zweitsprachlichen Kompetenzen der Probanden operationalisiert wurden, greifen hierzu jedoch auf Angaben des Schulbezirks zurück, die unscharf definiert sind. Insgesamt liefern diese Studien somit erste Hinweise darauf, dass das Ausmaß des Rückgriffs auf top-down bzw. bottom-up geleitete Strategien von der zweitsprachlichen Kompetenz beeinflusst wird, es sind allerdings weitere, methodisch belastbare Studien nötig, die diese Ergebnisse replizieren.

Insgesamt kann festgehalten werden, dass es sich beim Hörverstehen sowohl in einer Erst- als auch in einer Zweitsprache um einen komplexen Textverarbeitungsprozess handelt, bei dem auf Basis des sprachlichen Inputs sowie relevanter Hintergrundinformationen ein mentales Modell der im Hörtext beschriebenen Situation erstellt wird. Obwohl dieser Prozess in einer Zweitsprache grundsätzlich ähnlich verläuft wie in einer Erstsprache, ist anzunehmen, dass der Hörverstehensprozess vom Kompetenzniveau in der Zweitsprache moderiert wird. L2-Lernende mit geringen zweitsprachlichen Kompetenzen scheinen verstärkt entweder auf top-down *oder* bottom-up geleitete Verarbeitungsstrategien zurückzugreifen. L2-Lernende mit fortgeschrittenen Zweitsprachkenntnissen sind demgegenüber in der Lage, beide Verarbeitungsstrategien zu nutzen: Sie können das eingehende Sprachsignal adäquat bottom-up verarbeiten und gleichzeitig Hintergrundwissen im Sinne einer top-down verlaufenden Strategie integrieren (Field, 2004). Gelingt eine solche integrierte Verarbeitung eines Hörtextes nicht, entstehen mit hoher Wahrscheinlichkeit Verständnislücken.

3. Linguistische Komponenten des Hörverstehens

Für Hörende in einer Zweitsprache stellt die automatisierte bottom-up-Verarbeitung des Hörinputs auf der linguistischen Ebene aufgrund der häufig schwächer ausgeprägten sprachlichen Kompetenzen eine besondere Herausforderung dar. So konstatiert Solmecke (1993, S. 34):

„Ist [...] das Dekodieren ein mühsamer Prozess, der viel Aufmerksamkeit und Arbeitsgedächtniskapazität erfordert, dann bleibt wenig Raum für die Anwendung von Kenntnissen und Fähigkeiten auf höheren Ebenen, auch wenn sie

prinzipiell verfügbar sind. Den daraus resultierenden Verstehensschwierigkeiten begegnen Lernende auf unterschiedliche Art und Weise: Entweder sie erschöpfen sich auf der Ebene des Dekodierens, ohne zu den höheren Verstehensebenen vorzudringen, oder sie geben ihre Bemühungen auf der untersten Ebene vorschnell auf und versuchen durch Einsatz ihres Vorwissens zu einem Resultat zu kommen. Das Ergebnis ist in beiden Fällen mit großer Wahrscheinlichkeit Nicht- oder Falschverstehen.“

Bei einer Untersuchung des zweitsprachlichen Hörverstehens ist es demnach angebracht, zunächst auf linguistische Prädiktoren der bottom-up Verarbeitung von Hörtexten zu rekurrieren, da diese die Basis für den weiteren Verstehensprozess bilden. Im Folgenden werden Befunde zu linguistischen Komponenten des Hörverstehens und deren Ausprägung in einer Erst- bzw. Zweitsprache referiert.

3.1 Arbeitsgedächtnis

Unter dem Arbeitsgedächtnis versteht man ein System zur kurzzeitigen simultanen Speicherung und Verarbeitung von Informationen. Zur Beschreibung des Arbeitsgedächtnisses sind mehrere Konzeptionen entwickelt worden, von denen die Arbeitsgedächtniskonzeption nach Baddeley (1986) das am weitesten verbreitete Modell darstellt, das zudem als gut bestätigt gelten kann. Diesem Ansatz zufolge besteht das Arbeitsgedächtnis aus einer übergeordneten und modalitätsübergreifenden *zentralen Exekutive*, die zwei untergeordnete modalitätsspezifische Subsysteme steuert und überwacht: Während der *visuell-räumliche Notizblock* für die Aufrechterhaltung und Verarbeitung visuell-räumlicher Informationen verantwortlich ist, dient die *phonologische Schleife* der Speicherung und Verarbeitung sprach-klanglicher Informationen. Akustische bzw. verbale Informationen werden hier zunächst für einige Sekunden in einem temporären phonologischen Speicher gehalten, ehe sie verblasen. Durch subartikulatorische Rehearsalprozesse, d. h. aktive innere Wiederholung der Informationen, kann der Prozess des Verblasens herausgezögert werden, so dass die Informationen im Subsystem der phonologischen Schleife verbleiben (Baddeley, 1996). In späteren Fassungen wurde das System des Arbeitsgedächtnisses nach Baddeley (1996) um die Annahme eines sogenannten episodischen Puffers ergänzt, der verantwortlich für die Integration des verbalen Kurzzeitgedächtnisses mit Inhalten des Langzeitgedächtnisses, d. h. relevantem Vorwissen ist (Baddeley, 2000). Aufgrund des Fokus der vorliegenden Studien auf bottom-up Prädiktoren des Hörverstehens wird im Folgenden jedoch nur auf die Komponenten des Arbeitsgedächtnismodells

eingegangen, die für die Text-verarbeitung auf den unteren Verarbeitungsebenen von Bedeutung sind.

Für das Hörverstehen sind insbesondere zentral-exekutive Prozesse sowie die phonologische Schleife bedeutsam. Sie dienen der unmittelbaren und fortlaufenden Interpretation des sprachlichen Inputs und ermöglichen die Speicherung der gewonnenen Informationen sowie den Abgleich mit bereits bestehenden Gedächtnisinhalten (Call, 1985; Shohamy & Inbar, 1991). Entsprechend kann eine geringe Kapazität der phonologischen Schleife oder der zentralen Exekutive mit einer geringeren Fähigkeit einhergehen, die Vielzahl der beim Hören eingehenden Informationen zu analysieren und zu speichern (Daneman & Carpenter, 1980). Nimmt die laufende Verarbeitung des sprachlichen Inputs viel Arbeitsgedächtniskapazität in Anspruch, stehen zudem weniger Ressourcen für die Speicherung des Gehörten zur Verfügung und weniger Hörinhalte können erinnert werden. Aufgrund der in einer Zweitsprache häufig geringeren Wortschatz- und Syntaxkenntnisse muss der sprachliche Input bei Lernenden einer Zweitsprache häufig aktiv verarbeitet werden, so dass die automatisierte Verarbeitung eines Hörtextes nicht möglich ist. Es kommt daher bei L2-Lernenden häufiger als bei muttersprachlichen Hörenden zu einer Überlastung des Arbeitsgedächtnisses, die in der Folge zu einer schwächeren Hörverstehensleistung führt (Service, Simola, Metsänheimo & Maury, 2002; Vandergrift, 2007).

Für die generelle Speicher- und Verarbeitungskapazität sind aus theoretischer Sicht allerdings keine Unterschiede zwischen Lernenden einer Erst- bzw. Zweitsprache zu erwarten, da davon ausgegangen werden kann, dass das Arbeitsgedächtnis ein sprachunabhängiges System bildet (Baddeley, 1986). Die Befundlage zu Unterschieden in der Arbeitsgedächtniskapazität zwischen muttersprachlichen Lernenden und Lernenden einer Zweitsprache ist jedoch nicht eindeutig. So ließen sich in einer Reihe von Studien tatsächlich keine Unterschiede zwischen L1-Lernenden und L2-Lernenden in Hinblick auf die Kapazität der zentralen Exekutive und der phonologischen Schleife feststellen (da Fontoura & Siegel, 1995; Engel de Abreu, Baldassi, Puglisi & Befi-Lopes, 2013), während andere Studien auf eine geringere Arbeitsgedächtniskapazität bei L2-Lernenden im Vergleich zu ihren L1-Peers hindeuten (Call, 1985; Grotjahn, 2005; Lesaux, Lipka & Siegel, 2006).

Die divergierenden Befunde lassen sich dadurch erklären, dass einige der in den Studien verwendeten Maße zur Erfassung der Arbeitsgedächtniskapazität mit den sprachlichen Fähigkeiten der Probanden konfundiert sind. Typischerweise wird die Kapazität der phonologischen Schleife durch Maße erfasst, die das Erinnern und die Wiedergabe von visuell oder auditiv dargebotenem Material erfordern, wie etwa Ziffern- oder Wortspannen. Zur Erfassung

der zentralen Exekutive werden häufig komplexere Maße verwendet, die das Behalten von Informationen bei gleichzeitiger Verarbeitung eingehender Informationen erfordern. Beispiele hierfür sind Satzspannenmaße, bei denen mehrere Sätze gelesen und das jeweils letzte Wort jedes Satzes wiedergeben werden soll oder komplexe Ziffernspannen, bei denen die akustisch oder visuell dargebotenen Ziffernfolgen in umgekehrter Reihenfolge wiedergegeben werden müssen. Satz- und Wortspannenmaße könnten allerdings L2-Lernende benachteiligen, da diese weniger vertraut mit einzelnen Wörtern oder den phonologischen Strukturen der Zweitsprache sind als muttersprachliche Lernende (Olsthoorn, Andringa & Hulstijn, 2012). Auch bei Kunstwörtern kann es zu einem ähnlichen Effekt kommen, etwa wenn die phonologischen Eigenschaften der zu erinnernden Kunstwörter denen einer bestimmten Sprache ähneln (Lanfranchi & Swanson, 2005).

Entsprechend dieser Annahme wiesen in einigen Studien L2-Probanden geringere Arbeitsgedächtniskapazitäten auf als L1-Probanden, wenn Arbeitsgedächtnismaße genutzt wurden, die verbales Material verwendeten (z. B. Ennemoser, Pepouna & Hartung, 2012; Lesaux, Lipka & Siegel, 2006; Olsthoorn, Andringa & Hulstijn, 2012). Erfolgt die Erfassung der Arbeitsgedächtniskapazität hingegen anhand numerischer Spannen, so zeigten sich keine differenziellen Arbeitsgedächtnisleistungen bei Lernenden unterschiedlicher sprachlicher Herkunft (Engel de Abreu, Baldassi, Puglisi & Befi-Lopes, 2013; Olsthoorn et al., 2012). Arbeitsgedächtnismaße, die verbales Material verwenden, scheinen demnach zur Erfassung der Arbeitsgedächtniskapazität in einer Zweitsprache nur eingeschränkt anwendbar zu sein. Obwohl Ziffern häufig unter den ersten in einer Zweitsprache erlernten Wörtern sind und somit als gut eingeübtes und stark automatisiertes Stimulusmaterial gelten können, kann allerdings auch die erzielte Ziffernspannenlänge in einer Erstsprache höher als in einer Zweitsprache ausfallen (vgl. Olsthoorn et al., 2012). Eine mögliche Ursache hierfür liegt in der unterschiedlichen Länge der Ziffernwörter in den jeweiligen Sprachen, die Personen beim Erinnern der Ziffernspannen subartikulatorisch wiederholen. Im Deutschen erfordert die vorgegebene Ziffernspanne 9-3-8 beispielsweise das Erinnern dreier einsilbiger Zahlwörter. Von den entsprechenden türkischen Zahlwörtern sind hingegen zwei zweisilbig (*sekiz* für 8 und *dokuz* für 9), so dass insgesamt fünf Silben erinnert und wiedergegeben werden müssen. Wird die Ziffernspanne von L2-Probanden subvokal in der Zweitsprache wiederholt, stellt die vorgegebene Ziffernspanne demnach im Türkischen höhere Anforderungen an die Arbeitsgedächtniskapazität als im Deutschen, was differenzielle Leistungen türkisch-sprachiger resp. deutschsprachiger Personen erklären könnte.

Allerdings scheint dieser Effekt der unterschiedlichen Zahlwortlänge nur bei auditiver Darbietung der Ziffernfolgen aufzutreten, bei visueller Darbietung der Ziffernfolgen hingegen nicht oder nur in sehr geringem Maße (Olsthoorn et al., 2012). Eine mögliche Ursache für diesen Befund liegt darin, dass bei akustischer Darbietung die wiederzugebenden Ziffern direkt in Form der entsprechenden Zahlwörter präsentiert werden, während bei visueller Darbietung die Ziffern zunächst in die Zahlwörter rekodiert werden müssen, um sie subvokal wiederholen zu können. Olsthoorn et al. (2012) berichten, dass dies für L1-Probanden zu verbesserten Leistungen bei akustischer Präsentation von Ziffernspannen gegenüber einer visuellen Präsentation führen kann. Bei L2-Probanden tritt ein entsprechender Effekt hingegen nicht auf, da diese auch bei akustischer Präsentation von Ziffern die Zahlenspannen in einer Zweitsprache verarbeiten müssen. In der Folge könnten sich differenzielle Leistungen zwischen L1- und L2-Lernenden bei der akustischen Darbietung von Ziffernspannen ergeben, die für die L1-Lernenden, nicht aber für die L2-Lernenden, gegenüber der visuellen Darbietung einfacher zu verarbeiten ist.

Zusammenfassend kann den berichteten Befunden zufolge davon ausgegangen werden, dass sich die Arbeitsgedächtniskapazität in einer Erst- bzw. Zweitsprache nicht voneinander unterscheidet, sofern zu deren Erfassung Maße verwendet werden, die für L2-Lernende kulturfair sind, wie etwa visuell dargebotene Ziffernspannen. Allerdings ist anzunehmen, dass vor allem die zentral-exekutive Komponente des Arbeitsgedächtnisses im Zuge des zweitsprachlichen Hörverstehens häufig stärker belastet ist als in einer Erstsprache, da der eingehende linguistische Input aufgrund mangelnder Sprachkenntnisse häufig nicht automatisiert verarbeitet werden kann. In der Folge stehen weniger Arbeitsgedächtniskapazitäten für die Verarbeitung weiterer Informationen zur Verfügung, was schwächere Hörverstehensleistungen in einer Zweitsprache nach sich ziehen könnte. Bisherige Studien zu diesem Thema haben jedoch vielfach Material zur Erfassung der Arbeitsgedächtniskapazität verwendet, das für L2-Schülerinnen und -Schüler differenziell valide ist (z. B. verbale Spannenmaße), so dass der Zusammenhang zwischen Arbeitsgedächtnis und Hörverstehen in einer Zweitsprache unklar bleibt. Vermutet werden kann, dass das Arbeitsgedächtnis und das Hörverstehen in einer Zweitsprache besonders eng zusammenhängen, da eine automatisierte Verarbeitung des Hörtextes aufgrund geringen sprachlichen Wissens häufig nicht möglich ist und eine arbeitsgedächtnisintensive aktive Analyse des Hörinhalts nötig wird (Vandergrift, 2007).

3.2 Phonologische Bewusstheit

Als phonologische Bewusstheit wird jene Fähigkeit bezeichnet, die es erlaubt, die Lautstruktur der gesprochenen Sprache zu erkennen sowie darüber zu reflektieren und lautliche Manipulationen vorzunehmen (Kurita, 2012). Unterschieden werden kann dabei die *phonologische Bewusstheit im weiteren Sinne*, d. h. die Bewusstheit für größere sprachliche Einheiten wie Silben oder Reime von der *phonologischen Bewusstheit im engeren Sinne*, d. h. der Bewusstheit für kleinere sprachliche Einheiten, die sogenannten Phoneme (Laute) (Duzy, Gold, Schneider & Elmar Souvignier, 2013). Für das Hörverstehen relevant ist vor allem die phonologische Bewusstheit im engeren Sinne, die benötigt wird, um den eingehenden Sprachstrom in einzelne Wörter zu segmentieren (Cheung, 2007).

Grundsätzlich erfolgt der Erwerb der phonologischen Bewusstheit in einer Erst- und einer Zweitsprache weitgehend ähnlich (Tilstra, McMaster, Van den Broek, Kendeou & Rapp, 2009). Allerdings wird davon ausgegangen, dass die phonologische Bewusstheit als ein Aspekt metasprachlicher Bewusstheit durch den Kontakt zu verschiedenen sprachlichen bzw. lautlichen Systemen gefördert werden könnte (Bialystok, 2009; Bruck & Genesee, 1995). So ist davon auszugehen, dass L2-Lernende dem eingehenden Sprachstrom besondere Aufmerksamkeit widmen müssen, um zwischen der Erst- und Zweitsprache diskriminieren zu können. In der Folge könnten L2-Lernende gegenüber L1-Kindern eine ausgeprägtere Bewusstheit für die lautliche Struktur der von ihnen gesprochenen Sprachen erwerben.

Tatsächlich zeigen einige Studien, dass L2-Kinder gegenüber Kindern, die mit nur einer Sprache aufwachsen, häufig über eine erhöhte phonologische Bewusstheit verfügen (z. B. Bruck & Genesee, 1995; Campbell & Sais, 1995; Rubin & Turner, 1989). Neben diesen Befunden existieren allerdings auch Studien, die für diesen Bereich keine Vorteile oder gar Nachteile mehrsprachiger Kinder gegenüber einsprachigen Kindern belegen (z. B. Linck, Osthus, Koeth & Bunting, 2014; Wade-Woolley & Siegel, 1997). Die entsprechenden Befunde könnten jedoch mit dem sozioökonomischen Hintergrund der Kinder konfundiert sein, der eng mit dem Erwerb der phonologischen Bewusstheit verbunden ist (vgl. Duncan & Seymour, 2000) und in den Analysen zumeist nicht kontrolliert wurde. Eine weitere mögliche Erklärung für die divergierenden Befunde bisheriger Studien zur phonologischen Bewusstheit in einer Zweitsprache liegt darin, dass der Effekt der Zweisprachigkeit auf die Ausprägung der phonologischen Bewusstheit von den jeweils involvierten Sprachen abzuhängen scheint (Bruck & Genesee, 1995; Kang, 2012). Der in Abhängigkeit von der gesprochenen Sprache variierende phonologische Input trägt dabei zu Unterschieden in der Ausprägung einzelner Aspekte der phonologischen Bewusstheit bei. Beispielsweise weisen italienischsprachige Kinder oftmals

eine hohe phonologische Bewusstheit im engeren Sinne auf, die daraus resultiert, dass im Italienischen eine hohe Phonem-Graphem-Korrespondenz existiert, die die Ausbildung einer Bewusstheit für die phonemischen Regularitäten der Sprache erleichtert (Cossu, Shankweiler, Liberman, Katz & Tola, 1988). Stärken und Schwächen in Hinblick auf einzelne Aspekte der phonologischen Bewusstheit scheinen somit die die Häufigkeit des Vorkommens bestimmter phonologischer Regularitäten in den jeweiligen Sprachen widerzuspiegeln.

Vor dem Hintergrund der Annahme, dass die Ausbildung der phonologischen Bewusstheit von den phonologischen Regularitäten der involvierten Sprachen abhängt, wäre zu erwarten, dass türkisch-deutschsprachige Kinder gegenüber rein deutschsprachigen Kindern über eine erhöhte phonologische Bewusstheit verfügen sollten, die sich aus der klar abgrenzbaren Silbenstruktur und der einfachen Identifizierbarkeit einzelner Phoneme des Türkischen ergeben sollte. Die lautliche Transparenz des Türkischen sollte die Entwicklung sowohl der Bewusstheit für Silben als auch die der phonemischen Bewusstheit erleichtern. Sowohl in einer Studie von Limbird und Stanat (2006) als auch in einer Studie von Duzy und Kollegen (2013) fand sich jedoch kaum empirische Unterstützung für diese Annahme.

Zu beachten ist allerdings auch hier, dass in den referierten Studien der sozioökonomische Status der Schülerinnen und Schüler als potenziell konfundierte Variable häufig nicht hinreichend erfasst wurde. So wurde bei Duzy et al. (2013) das kulturelle Kapital innerhalb der Familien anhand der Anzahl der Bücher im elterlichen Haushalt operationalisiert. Kinder aus türkisch-deutschsprachigen Familien wiesen dabei ein geringeres kulturelles Kapital auf als Kinder aus rein deutschsprachigen Familien. Allerdings wurde in den Analysen nicht für diesen differenziell ausgeprägten kulturellen Hintergrund kontrolliert. Limbird und Stanat (2006) berichten, dass die an der Studie beteiligten Schulen sich hinsichtlich der sozioökonomischen Situation der Bevölkerung im jeweiligen Schulumfeld nicht unterschieden. Der sozioökonomische Status auf Individualebene wurde jedoch nicht erhoben. Da türkischstämmige Schülerinnen und Schüler eher als Kinder ohne Zuwanderungshintergrund aus sozioökonomisch benachteiligten Familien stammen (vgl. z. B. Autorengruppe Bildungsberichterstattung, 2014; OECD, 2013), wäre auch für diese Studie denkbar, dass die Befunde nicht frei von Effekten eines differenziell ausgeprägten sozioökonomischen Status der Kinder sind. Dieser scheint mit dem Erwerb der phonologischen Bewusstheit verbunden zu sein (vgl. z. B. Duncan & Seymour, 2000; Lonigan, Burgess, Anthony & Barker, 1998). Als mögliche Ursache der differenziellen Entwicklung phonologischer Fähigkeiten bei Kindern mit niedrigem bzw. hohem sozioökonomischen Status werden dabei Unterschiede in spracherwerbsrelevanten familiären Variablen, wie etwa der Verfügbarkeit von Büchern oder

gemeinsamen Vorleseaktivitäten diskutiert, die Kindern unterschiedliche Gelegenheiten bieten, sprachliche Fähigkeiten zu erwerben (Lonigan et al., 1998). Wird in Studien zu Unterschieden in der phonologischen Entwicklung von Erst- und Zweitsprachlernenden nicht für den sozioökonomischen Hintergrund kontrolliert, bleibt demnach offen, ob die Ergebnisse auf den Status der Kinder als Erst- bzw. Zweitsprachlernende oder auf ihren differenziellen sozioökonomischen Status zurückführbar sind.

Über Beziehungen zwischen der phonologischen Bewusstheit und dem Hörverstehen bei L2-Lernenden im Vergleich zu L1-Lernenden lässt sich derzeit kaum eine gesicherte Aussage treffen, da bislang nur wenige Studien zu diesem Aspekt vorliegen (Kieffer, 2014). Aus theoretischer Sicht ist davon auszugehen, dass die Identifizierung von Wortgrenzen im Sprachfluss L2-Hörende, vor allem in den frühen Phasen des Zweitspracherwerbs, vor große Herausforderungen stellt und ihre Hörverstehensleistung in erheblichem Umfang beeinflusst (Field, 2003). So argumentiert Stæhr (2009) beispielsweise: „spoken language is characterized by assimilation as well as unclear articulation, and lexical units are not necessarily as marked as in written texts; this lack of clarity of spoken language makes word segmentation an extremely difficult task for L2 listeners“ (S. 582). In einer quantitativen Studie mit chinesischesprachigen Studierenden, die Englisch als Fremdsprache erwarben, gab entsprechend ein Großteil der Probanden an, beim zweitsprachlichen Hören Schwierigkeiten bei der Identifizierung von Phonemen und dem Erkennen von Wortgrenzen zu haben (Goh, 2000). Ergebnisse einer querschnittlich angelegten Studie von Li und Kollegen (2012) liefern weitere Hinweise auf einen besonders ausgeprägten Effekt der phonologischen Bewusstheit auf das zweitsprachliche englische Hörverstehen bei chinesisch-englischsprachigen Grundschulkindern. Der entsprechende Effekt fiel unter Kontrolle der erstsprachlichen Kompetenzen mit $\beta = .32$ in der vierten Klassenstufe stärker aus als bei jüngeren Schüler-innen und Schülern in der zweiten Klassenstufe ($\beta = .18$). Die Autoren schlussfolgern, dass die älteren Kinder im Zuge des Schriftspracherwerbs in der Zweitsprache gelernt hatten, Lauten bzw. Phonemen eine verstärkte Aufmerksamkeit zukommen zu lassen und so eine bessere phonologische Bewusstheit entwickelt hatten, die sich in der Folge positiv auf das Hörverstehen der Schülerinnen und Schüler auswirkte. Allerdings gab es in der Studie von Li et al. (2012) keine monolinguale Kontrollgruppe, so dass unklar bleibt, ob die phonologische Bewusstheit und das Hörverstehen in einer Zweitsprache gleichermaßen stark miteinander assoziiert sind wie in einer Erstsprache oder ob differenzielle Zusammenhangsmuster vorliegen.

Zusammenfassend lässt sich festhalten, dass phonologische Kompetenzen bei L2-Schülerinnen und -Schülern häufig ähnlich gut ausgebildet sind wie bei ihren L1-Peers. Aller-

dings lassen sich für diesen Bereich kaum allgemeingültige Aussagen treffen, da die Ausprägung der phonologischen Bewusstheit in einer Zweitsprache von einer Reihe von Faktoren, wie etwa den phonologischen Charakteristika der Erst- und der Zweitsprache sowie dem sozioökonomischen Status abzuhängen scheint. Obwohl insgesamt nur wenig empirische Befunde hierzu vorliegen, liefern bisherige Studien Hinweise darauf, dass phonologische Fähigkeiten mit dem Hörverstehen in einer Zweitsprache besonders eng verbunden sein könnten. Es fehlt bislang allerdings an Studien, die die Ausprägung der phonologischen Bewusstheit sowie deren Effekte auf die Hörverstehensleistung in einer Erst- bzw. Zweitsprache unter Berücksichtigung des sozioökonomischen Hintergrunds vergleichend untersuchen.

3.3 Wortschatz

Für die Bildung einer bedeutungsbasierten Repräsentation eines Hörtextes sind lexikalische Kenntnisse von zentraler Bedeutung, so dass Worterkennung beim Hörverstehen als wichtigster semantischer Verarbeitungsprozess gilt (sog. *lexis first principle*, vgl. Rost, 1994). In einer Zweitsprache sind gut ausgebildete Wortschatzkenntnisse für das Hörverstehen von besonderer Bedeutung, da der Sprachinput häufig durch Assimilationen oder undeutliche Artikulation geprägt ist, so dass lexikalische Einheiten nicht so deutlich markiert sind wie beispielsweise in geschriebener Sprache (Stæhr, 2009). Die Segmentierung des Gehörten in einzelne sprachliche Einheiten ist daher für L2-Lernende besonders herausfordernd. Vor allem L2-Schülerinnen und -Schüler mit geringen zweitsprachlichen Kompetenzen versuchen dabei häufig, top-down geleitete Verarbeitungsstrategien anzuwenden, d. h. den eingehenden Sprachstrom kontinuierlich zu überwachen und mit bekannten Wörtern abzugleichen (Field, 2003). Dies kann problematisch sein, wenn Informationen aus der bottom-up Verarbeitung eines Hörtextes missachtet werden, etwa wenn in der gehörten Phrase „went to assist a passenger“ das bekannte Wort „sister“ identifiziert wird und in der Folge die Erstellung eines kohärenten Situationsmodells erschwert wird (Field, 2003; vgl. auch Voss, 1984 für weitere Beispiele). Geringere Wortschatzkenntnisse erschweren somit die Fähigkeit, die Bedeutung eines Hörtextes extrahieren zu können. Meara (1996) argumentiert entsprechend, dass der Umfang des Wortschatzes die Basis sprachlicher Kompetenz bildet und erst ein breiter Wortschatz Lernende dazu befähigt, kompetent mit Sprache umzugehen. Für den Wortschatzumfang wird dabei davon ausgegangen, dass mindestens 80% bis 98% der im Hörtext vorhandenen Wörter erkannt werden müssen, um einen Text verstehen zu können (Bonk, 2000; Nation, 2006).

Grundsätzlich verläuft der Wortschatzerwerb in einer Zweitsprache ähnlich des erstsprachlichen Wortschatzerwerbs und erfolgt auch in einer vergleichbaren Geschwindigkeit

(Jean & Geva, 2009; Kieffer & Vukovic, 2013). In der Mehrzahl vorliegender empirischer Studien mit Lernenden aus unterschiedlichen Sprachräumen und Altersgruppen fand sich allerdings wiederholt eine geringere Wortschatzbreite von L2-Schülerinnen und -Schülern gegenüber ihren monolingualen Peers, auch wenn die Gruppen in Bezug auf ihren sozialen Hintergrund vergleichbar waren (z. B. Kieffer & Vukovic, 2013; Limbird, Maluch, Rjosk, Stanat & Merkens, 2014; Umbel & Oller, 1994; Umbel, Pearson, Fernández & Oller, 1992; Verhoeven, 1990). Im Rahmen einer längsschnittlich angelegten Studie zeigten beispielsweise Kieffer und Vukovic (2013), dass L2-Kinder im Verlauf der ersten vier Grundschuljahre auch bei vergleichbarem sozioökonomischem Hintergrund konsistent einen geringeren Wortschatz aufwiesen als L1-Kinder. Obwohl sich für die L2-Kinder in dieser Studie eine Zunahme des Wortschatzes über den Verlauf der Grundschulzeit fand, war diese Entwicklung offenbar nicht ausreichend, um Wortschatzkenntnisse zu erwerben, die denen ihrer monolingualen Peers entsprachen.

Zu beachten ist dabei, dass der Wortschatz bilingualer Kinder in der Regel über beide Sprachen verteilt ist, so dass die Kinder einige Worte ausschließlich in ihrer Erst- und andere Worte ausschließlich in ihrer Zweitsprache kennen (Oller, 2005). Es ist daher denkbar, dass das kombinierte erst- und zweitsprachliche Lexikon von L2-Lernenden den Umfang des Lexikons einsprachiger Personen übersteigt (Oller, 2005; Oller, Pearson & Cobo-Lewis, 2007). Die Befundlage zu dieser Annahme ist jedoch inkonsistent: Während eine Reihe von Studien die Annahme einer Verteilung des Wortschatzes von L2-Lernenden über die gesprochenen Sprachen hinweg stützen konnte (vgl. Oller et al., 2007), wird die Annahme, der kombinierte Wortschatz bilingualer Personen sei größer als der Wortschatz monolingualer Personen weniger gut gestützt. Engel de Abreu und Kollegen (2013) beispielsweise fanden für portugiesischsprachige L2-Lernende in Luxemburg einen geringeren Wortschatz in den jeweiligen Sprachen gegenüber portugiesischsprachigen monolingualen Kindern wie auch luxemburgischen monolingualen Kindern, obwohl die Gruppen in Bezug auf zentrale Hintergrundvariablen wie dem sozioökonomischen Status vergleichbar waren. Darüber hinaus war auch der kombinierte Wortschatz der L2-Lernenden in dieser Studie geringer als der jeweilige Wortschatz der beiden monolingualen Gruppen.

Ausgehend von Studien, die darauf hindeuten, dass Zweitsprachlernende im Vergleich zu L1-Lernenden häufig weniger breite Wortschatzkenntnisse aufweisen, ist vermutet worden, dass diese Schwächen im Bereich lexikalischer Kenntnisse für die im Vergleich zu L1-Lernenden oftmals geringeren Textverstehenskompetenzen von L2-Kindern verantwortlich sind (vgl. Kieffer & Vukovic, 2012). Tatsächlich scheint der Umfang des Wortschatzes in

einer Zweitsprache eng mit dem zweitsprachlichen Hörverstehen verbunden zu sein. So zeigte sich in einer Studie mit spanischsprachigen Grundschülerinnen und Grundschülern, die Englisch als Zweitsprache erlernten, in einer strukturanalytischen Modellierung ein bedeutsamer Effekt von $\beta = .85$ des zweitsprachlichen Wortschatzes auf das Hörverstehen in der Zweitsprache (Proctor et al., 2005). Auch in weiteren Studien mit L2-Lernenden, die Englisch als Fremdsprache erwarben, wurden mit $r = .45$ bzw. $r = .70$ mittlere bis starke Zusammenhänge zwischen dem Umfang des englischen Wortschatzes und dem englischen Hörverstehen berichtet (Mecarty, 2000; Stæhr, 2009).

Bei der Interpretation dieser Ergebnisse ist zu beachten, dass in keiner der bislang genannten Studien zum Zusammenhang zwischen Wortschatz und Hörverstehen eine monolinguale Vergleichsgruppe eingesetzt wurde. Es kann daher lediglich festgehalten werden, dass der Wortschatz und das Hörverstehen in einer Zweitsprache eng miteinander verbunden sind; inwiefern diese Assoziation stärker ausfällt als in einer Erstsprache kann hingegen nicht abgeschätzt werden. Befunde einer längsschnittlich angelegten Studie aus den Niederlanden stützen allerdings die auf theoretischer Ebene postulierte Annahme eines besonders engen Zusammenhangs zwischen dem Wortschatz und dem Hörverstehen bei L2-Lernenden. Droop und Verhoeven (2003) untersuchten verschiedene Aspekte des Sprachverstehens und deren Effekte auf das Leseverstehen bei Grundschülerinnen und Grundschülern mit Niederländisch als Erst- bzw. Zweitsprache. Die Befunde zeigten, dass für die monolingual niederländischen Kinder das auditive Textverstehen maßgeblich von ihren morpho-syntaktischen Fähigkeiten beeinflusst wurde. Für die L2-Kinder gab es zusätzlich einen Effekt der Wortschatzkenntnisse auf das niederländische Hörverstehen. Diese Befunde liefern erste Hinweise darauf, dass das auditive Textverstehen in einer Zweitsprache stärker vom Wortschatz abhängig sein könnte als in einer Erstsprache. Allerdings sind weitere Studien nötig, die diese Befunde replizieren und die theoretische Annahme eines besonders ausgeprägten Effekts des Wortschatzes auf das auditive Textverstehen in einer Zweitsprache empirisch stützen.

3.4 Morpho-syntaktische Kenntnisse

Auf der Ebene des Parsing bzw. der Erstellung einer propositionalen Repräsentation eines Hörtextes werden grammatikalische Fähigkeiten im Bereich der Morphologie und der Syntax benötigt, um durch eine Analyse der im Hörtext enthaltenen morpho-syntaktischen Informationen die Bedeutungen einzelner Satzteile extrahieren zu können. Der eingehende akustische Input muss zu diesem Zweck zunächst in einzelne grammatische Kategorien, d. h. einzelne Worte oder Satzglieder unterteilt werden, um anschließend Bezüge zwischen diesen Elemen-

ten herzustellen, die es erlauben, die propositionale Struktur des Hörinhalts mental abzubilden. Auf der Wortebene ist im Zuge dieses Prozesses vor allem eine gut ausgebildete morphologische Bewusstheit wichtig, die es erlaubt, die morphologische Struktur und somit die grammatikalische Funktion eines Wortes innerhalb eines Satzes zu erkennen. Unter Morphologie wird dabei die Lehre vom Aufbau von Wörtern aus kleineren, bedeutungs- oder funktionstragenden Einheiten, den sogenannten *Morphemen* verstanden (Tilstra et al., 2009). Auf Satzebene ist für die Bildung einer propositionalen Abbildung eines Hörtextes eine gut ausgebildete syntaktische Bewusstheit nötig, d. h. die Fähigkeit, die grammatikalische Struktur von Sprache zu erkennen und grammatikalisch korrekte von grammatikalisch inkorrekten Sätzen zu unterscheiden sowie syntaktische Fehler korrigieren zu können (Linck et al., 2014; Tilstra et al., 2009).

Hinsichtlich der Ausprägung der morpho-syntaktischen Bewusstheit in einer Zweitsprache finden sich für jene L2-Lernende, bei denen die Zweitsprache die dominante Sprache darstellt sowie bei balancierter Zweisprachigkeit, d. h. in dem Fall, dass sowohl die Erst- als auch die Zweitsprache gleichermaßen gut beherrscht werden, in der Regel morpho-syntaktische Fähigkeiten, die denen von monolingualen Personen entsprechen (Gutierrez-Clellen & Simon-Cerejido, 2007; Paradis, 2010). Für L2-Lernende, bei denen die Zweitsprache die der jeweiligen Erstsprache nachgeordnete Sprache darstellt, wurde in empirischen Studien jedoch wiederholt eine gegenüber L1-Schülerinnen und -Schülern geringere morpho-syntaktische Bewusstheit gezeigt (Chiappe, Siegel & Wade-Woolley, 2002; Da Fontoura & Siegel, 1995; Lesaux & Siegel, 2003; Paradis, 2010; Verhoeven, 1990), die sich auch dann findet, wenn sich die Gruppen der L1- und der L2-Lernenden in Hinblick auf ihren sozioökonomischen Hintergrund ähneln. Verhoeven (1990) beispielsweise fand in einer längsschnittlich angelegten Studie schwächere morpho-syntaktische Leistungen bei türkisch-sprachigen L2-Grundschülerinnen und -Grundschülern im Vergleich zu einer in Bezug auf den sozioökonomischen Hintergrund gematchten Stichprobe von niederländischen L1-Kindern. Weitere Studien deuten darauf hin, dass Schwierigkeiten von L2-Kindern im Bereich morpho-syntaktischer Kenntnisse auch im Verlauf der Grundschulzeit persistieren. So zeigen L2-Kinder in empirischen Studien auch gegen Ende der Grundschulzeit häufig eine geringe morpho-syntaktische Bewusstheit als monolingual deutschsprachige Kinder, die sich beispielsweise im Umgang mit Relativsätzen, Passivkonstruktionen oder Präpositionalgruppen manifestiert (Grießhaber, 1999; Reich, 2009). Dabei scheinen morpho-syntaktische Regularitäten, die sich nur aus der Erfahrung mit einer Sprache ableiten lassen (z. B. irreguläre Pluralformen) eine größere Hürde für L2-Schülerinnen und -Schüler darzustellen, als morpho-

syntaktische Regularitäten, die sich aus einfachen Regeln ableiten lassen, wie etwa der Gebrauch regulärer Pluralformen (Gottardo & Mueller, 2009).

Wie in einer Erstsprache sollte die morpho-syntaktische Bewusstheit auch für die Erstellung einer propositionalen Abbildung des Hörinhalts und in der Folge für das Hörverstehen in einer Zweitsprache eine bedeutsame Rolle spielen. Aus theoretischer Sicht wird von einigen Autoren angenommen, dass die häufig geringeren morpho-syntaktischen Fähigkeiten in einer Zweitsprache das Erkennen der sprachlichen Struktur eines Textes erschweren und somit eine Hauptursache für das Zustandekommen schwächerer Verstehensleistungen bilden (vgl. Droop & Verhoeven, 2003; Paschke, 2000). Dieser Sichtweise nach sollten morpho-syntaktische Fähigkeiten und Hörverstehen in einer Zweitsprache besonders eng miteinander verbunden sein. Empirisch lässt sich diese Annahme bislang jedoch nicht bestätigen. Felser und Clahsen (2009) beispielsweise zeigten, dass auch fortgeschrittene L2-Lernende beim Hörverstehen weniger stark auf morpho-syntaktische Parsingprozesse zurückgriffen als ihre monolingualen Peers. Ähnliche Befunde erbrachte die bereits genannte Studie von Droop und Verhoeven (2003), in der die Bedeutsamkeit verschiedener sprachlicher Prädiktoren auf das Hörverstehen von niederländischen und türkisch-niederländischen Kindern der dritten und vierten Klassenstufe untersucht wurde. Der Effekt der morpho-syntaktischen Bewusstheit auf das Hörverstehen in der Zweitsprache war dabei auch unter Kontrolle des sozioökonomischen Status der Kinder geringer ausgeprägt als in der Erstsprache. Für den deutschen Sprachraum erbrachte eine Studie mit Grundschülerinnen und -schülern, die Deutsch als Zweitsprache erlernten, ebenfalls keine Hinweise auf Nachteile im Hörverstehen von L2-Schülerinnen und -Schülern gegenüber ihren L1-Peers, die sich auf die syntaktische Komplexität der Hörtexte zurückführen ließen (Eckhardt, 2008). Felser und Clahsen (2009) geben an, dass L2-Lernende offenbar eher auf lexikalisch-semantische Informationen bei der Erstellung einer Textrepräsentation zurückgreifen als auf morpho-syntaktische Informationen. Zu einem ähnlichen Schluss kommen auch Droop und Verhoeven (2003).

Insgesamt zeigt sich, dass die morpho-syntaktische Bewusstheit in einer Zweitsprache häufig geringer ausgeprägt ist, sofern Erst- und Zweitsprache nicht im Kontext einer balancierten Bilingualität erworben werden (vgl. Gutierrez-Clellen & Simon-Cereijido, 2007; Paradis, 2010). Für L2-Lernende, die erst spät mit dem Erwerb der Zweitsprache begonnen haben bzw. geringen Kontakt zu Zweitsprache haben besteht somit die Gefahr der Ausbildung einer schwachen morpho-syntaktischen Bewusstheit (vgl. z. B. Chiappe, Siegel & Wade-Woolley, 2002; Paradis, 2010). In Hinblick auf die Bedeutsamkeit der morpho-syntaktischen Kenntnisse für das Hörverstehen von L2-Schülerinnen und -Schülern deuten bisherige Stu-

dien eher darauf hin, dass die morpho-syntaktische Bewusstheit in einer Zweitsprache eine geringere Rolle für das Hörverstehen spielen könnte als in einer Erstsprache (z. B. Droop & Verhoeven, 2003; Felser & Clahsen, 2009). Allerdings fokussierte die Mehrheit der bislang vorliegenden Studien auf den Grundschulbereich oder auf Lernende in den ersten Phasen des L2-Erwerbs. Weitgehend unklar bleibt bislang, inwiefern eventuelle Schwächen im Bereich morpho-syntaktischer Kenntnisse das Textverstehen bei L2-Lernenden beeinflussen, die in ihrer Zweitsprache beschult wurden und bereits über fortgeschrittene L2-Kenntnisse verfügen.

4. Die Rolle des sprachlichen Kontextes: Bildungssprachliches und alltagssprachliches Hörverstehen

Neben den vorgestellten linguistischen Prädiktoren wird das Hörverstehen auch von Kontextfaktoren beeinflusst. Von besonderer Bedeutung sind in diesem Zusammenhang die in Abhängigkeit von der sprachlichen Umgebung variierenden Anforderungen an die Zuhörenden: Sprachliches Verstehen findet in unterschiedlichen Kontexten und Situationen statt und stellt unterschiedliche Ansprüche an die Rezipienten. Eine wichtige Unterscheidung sprachlicher Anforderungsmomente betrifft die Abgrenzung der im Alltag verwendeten Sprache von der Sprache, die typischerweise in akademischen Kontexten wie etwa im Schulunterricht oder in einer universitären Vorlesung verwendet wird. Die in diesen Bereichen genutzte Sprache hat zwar viele Merkmale mit anderen Typen gesprochener Sprache, wie etwa einem alltäglichen Gespräch, gemeinsam. Da sie in der Regel jedoch einen sehr formalen, geplanten und informationsvermittelnden Charakter aufweist, ähnelt sie in vielerlei Hinsicht der Schriftsprache (Flowerdew & Miller, 1997). Wie genau bildungssprachliches Hörverstehen zu definieren ist und in welchem Maß man es beherrschen muss, um in Bildungskontexten erfolgreich zu sein, ist allerdings bislang unklar (Song, 2011).

Zentrales Kennzeichen bildungssprachlichen Hörverstehens ist verschiedenen Ansätzen zufolge ein hoher Grad von Schriftsprachlichkeit (z. B. Csomay, 2006; Koch & Oesterreicher, 1985; Rubin, 1994). Mündliche Kommunikation in Bildungskontexten ist demnach häufig durch sprachliche Merkmale gekennzeichnet, die typisch für die schriftsprachliche Kommunikation sind, etwa eine hohe lexikalische Dichte, eine geringe Kontextualisierung der Inhalte und der Gebrauch von Textstrukturen, wie sie in der Schriftsprache üblich sind (Flowerdew & Miller, 1997). Koch und Oesterreicher (1985) unterscheiden entsprechend zwischen konzeptionell schriftlicher und konzeptionell mündlicher Sprache, die die beiden Endpunkte eines sprachlichen Kontinuums markieren. Konzeptionell mündliche Sprache ist dabei als sogenannte *Sprache der Nähe* zu verstehen, die sowohl in der geschriebenen als auch in der münd-

lichen Modalität durch spezifische Parameter, wie etwa die Vertrautheit der Kommunikationspartner, Dialogizität, Spontaneität, referenzielle Nähe und eine Situations- und Handlungsbindung gekennzeichnet ist (Koch & Oesterreicher, 1985; 2007). Demgegenüber steht auf der konzeptuell schriftsprachlichen Seite die sogenannten *Sprache der Distanz*, die sich durch Eigenschaften wie die Fremdheit der Kommunikationspartner, Monologizität, Reflektiertheit, referenzielle Distanz und eine Situations- und Handlungsentbindung auszeichnet.

Auch andere Ansätze betonen die konzeptionelle Unterscheidung von Mündlichkeit und Schriftlichkeit. Ochs (1979) beispielsweise regt an, zwischen geplantem Diskurs, der eher schriftsprachtypische Merkmale aufweist, und ungeplantem Diskurs, der eher mit Mündlichkeit assoziiert ist, zu unterscheiden. Geplante schriftliche Kommunikation und spontane mündliche Äußerungen wären demnach als Extrempunkte eines Kontinuums zu verstehen, auf dem sich sprachliche Äußerungen verorten lassen. Diese Idee wird in der Annahme eines sogenannten *oral-literate continuums* (Tannen, 1980) aufgegriffen. Diesem Ansatz zufolge besteht ein wesentliches Unterscheidungsmerkmal kommunikativer Praktiken nicht in Hinblick auf den Modus der Mündlichkeit bzw. Schriftlichkeit per se, sondern ergibt sich aus den unterschiedlichen Strategien, die mit mündlichen bzw. schriftlichen Erzähltraditionen assoziiert sind und in jedem sprachlichen Modus verwendet werden können. Als *oral* wird dabei eine Sprache bezeichnet, in der Sender und Empfänger einer Nachricht über geteiltes Wissen oder eine Beziehung zueinander verfügen. *Literale* Kommunikation hingegen findet statt, wenn keine oder nur wenig Interaktion zwischen Sender und Empfänger stattfindet und der Inhalt dekontextualisiert ist. Sprachliche Äußerungen können diese Merkmale in unterschiedlichem Ausmaß beinhalten und sich entsprechend an unterschiedlichen Stellen des oral-literate-Kontinuums verorten lassen (Tannen, 1980). Mündliche Kommunikation in Bildungskontexten kann an der Schnittstelle zwischen Mündlichkeit und Schriftlichkeit angesiedelt werden: In der Regel enthält gesprochene Bildungssprache dabei weniger mit konzeptioneller Schriftlichkeit bzw. Literalität assoziierte Eigenschaften als geschriebene Bildungssprache (vgl. Nagy & Townsend, 2012). Dennoch beinhaltet auch gesprochene Sprache in akademischen Kontexten häufig Merkmale, die typisch für den Schriftsprachbereich sind, wie etwa einen Fokus auf die Vermittlung von Informationen und einen abstrakten Stil (Csomay, 2006). Entsprechend nimmt Tannen (1982) an, dass sich Hörtexte auf einem Kontinuum ansiedeln lassen, bei dem natürliche gesprochene Sprache das eine Extrem bildet und das andere Extrem durch jene Sprache gebildet wird, die Merkmale enthält, die eher mit geschriebenen Texten in Verbindung gebracht werden (z. B. hohe Informationsdichte, wenige Pausen, komplexe Sätze).

Häufig wird davon ausgegangen, dass die Verarbeitung bildungssprachlicher Texte gegenüber der Verarbeitung alltagssprachlicher Texte besondere Anforderungen an die sprachlichen Kompetenzen von Lernenden stellt (Feilke, 2012; Scarcella & Rumberger, 2000; Snow & Uccelli, 2009; Nagy & Townsend, 2012). Entsprechend wurde in den vergangenen Jahrzehnten eine Vielzahl von Beschreibungen des Hörverstehens in Bildungskontexten und den damit verbundenen Schwierigkeiten erstellt (vgl. Ferris & Tagg, 1996; Taylor & Geranpayeh, 2011). Darüber hinaus enthalten auch einige der generellen für das Hörverstehen entwickelten Klassifikationen (vgl. Kapitel 2.2) spezifische Angaben zu den für Hörverstehen in Bildungskontexten notwendigen Fähigkeiten (z. B. Buck, 2001; Flowerdew, 1995; Richards, 1983; Rubin, 1994). Flowerdew (1995) zufolge zeichnet sich bildungssprachliches Hörverstehen durch eine Reihe von Charakteristika aus, die spezielle Anforderungen an Zuhörer und Zuhörerinnen stellen. Dazu gehören beispielsweise die Verfügbarkeit des notwendigen Hintergrundwissens, die Fähigkeit, zwischen relevanten und irrelevanten Informationen unterscheiden zu können und die Verfügbarkeit nützlicher Strategien, wie z. B. das Anfertigen von Notizen während eines Vortrags. Unterschieden werden kann dabei zwischen authentischen Vorlesungen, die vom Sprechenden frei gehalten werden und schriftlich ausgearbeiteten Vorlesungen (Flowerdew & Miller, 1997). Erstere enthalten häufig inkomplette Sätze sowie Brückenwörter, die die Struktur eines Textes verdeutlichen (z. B. *und*, *aber*, *nun*). In authentischen Vorlesungen finden sich zudem mehr Pausen, Wiederholungen und Redundanzen als bei schriftlich ausgearbeiteten Vorlesungen.

Wie Field (2008) betont, sind solche Taxonomien der für das bildungssprachliche Hörverstehen benötigten Kompetenzen durchaus von Relevanz, da sie Lehrenden einen hilfreichen Überblick über jene Fähigkeiten bieten, die von den Lernenden erworben werden müssen, um Hörtexte im schulischen Kontext verstehen zu können. Allerdings werden in keiner der genannten Klassifikationen Angaben dazu gemacht, wie genau die beschriebenen Fähigkeiten operationalisiert werden können. Zudem bleibt unklar, inwiefern diese Fähigkeiten spezifisch für das bildungssprachliche Hörverstehen sind. Für eine empirische Überprüfung bildungssprachlicher Hörverstehenskompetenzen sind diese Klassifikationen daher von geringerer Bedeutung, da bislang keine empirische Validierung der postulierten zugrundeliegenden Fähigkeiten vorliegt.

Eine allgemeine und nicht nur auf das Hörverstehen bezogene Definition der in Bildungskontexten benötigten sprachlichen Kompetenzen stammt von Cummins (1980, 2008). Wie in Kapitel 1.3 dargestellt, unterscheidet Cummins (1980, 2000) dabei zwischen „*Basic Interpersonal Communication Skills (BICS)*“, d. h. der minimalen kommunikativen Kompe-

tenz, die in sozialen Interaktionen benötigt wird, und „*Cognitive Academic Language Proficiency (CALP)*“, die im Zuge der formellen Bildung erworben wird. Ausgehend von dieser Differenzierung hat das *National Center for Research on Evaluation, Standards, and Student Testing (CREST)* in den USA eine Konzeption von Bildungssprache entwickelt, die die Definition von CALP aufgreift und in einigen Punkten erweitert (Bailey & Butler, 2003). Bildungssprache wird dabei definiert als diejenige Sprache, die in der Schule oder anderen akademischen Kontexten genutzt wird, um Wissen zu vermitteln und zu erwerben (Anstrom et al., 2010): Es ist die Sprache “that is used by teachers and students for the purpose of acquiring new knowledge and skills, imparting new information, describing abstract ideas, and developing students’ conceptual understanding” (O’Malley et al., 1989, S.40). Inhaltlich ist das Register der Bildungssprache durch kognitiv anspruchsvolle und schul- bzw. bildungsbezogene Inhalte gekennzeichnet, wie man sie beispielsweise bei Vorträgen zu abstrakten Ideen und Konzepten findet (Bailey & Butler, 2003). Bildungssprache ist dabei weitgehend dekontextualisiert, d. h. sie beinhaltet im Gegensatz zur im Alltag genutzten Sprache keine oder nur wenige kontextuelle Hinweisreize, wie etwa Gesten, die das Verständnis unterstützen könnten (Bailey & Butler, 2003). Auf der sprachlichen Ebene ist Bildungssprache durch den Gebrauch bestimmter lexikalischer und syntaktischer Merkmale gekennzeichnet, die in der Alltagssprache vergleichsweise selten vorkommen. Im Bereich des Wortschatzes handelt es sich dabei z. B. vorwiegend um den Gebrauch eines domänenübergreifenden generellen bildungssprachlichen Wortschatzes (z. B. *erklären, beschreiben*) sowie eines domänenspezifischen speziellen bildungssprachlichen Wortschatzes (z. B. *Adjektiv, Nenner*) (vgl. Heppt, Haag, Böhme & Stanat, 2015). Grammatikalische Charakteristika, die Bildungssprache kennzeichnen, sind u.a. der vermehrte Gebrauch von Konnektoren, Nominalphrasen, Relativsätzen und Passivkonstruktionen (Bailey, Butler, Stevens & Lord, 2007; Butler, Bailey, Stevens & Huang, 2004). Bildungssprache und alltägliche Sprache sind durch die graduelle An- bzw. Abwesenheit der genannten Merkmale gekennzeichnet und lassen sich auf einem Kontinuum verorten, bei dem die beiden sprachlichen Register die Endpunkte markieren (Berendes, Dragon, Weinert, Heppt & Stanat, 2013; Snow, 2010). Die für den englischen Sprachraum diskutierten Charakteristika von Bildungssprache werden dabei auch als zentrale Eigenschaften von Bildungssprache im Deutschen angenommen (Eckhardt, 2008; Heppt et al., 2015).

Bislang existieren nur wenige Studien, die Bildungssprache und alltägliche Sprache auch empirisch voneinander abzugrenzen versuchen. Im deutschen Sprachraum liegen derzeit kaum empirische Überprüfungen der angenommenen Merkmale von Bildungssprache vor.

Für das Englische konnte Schleppegrell (2004) in einer Analyse der sprachlichen Charakteristika von Transkripten alltäglicher Konversationen einerseits und schulischer Texte (z. B. Textbücher, Aufgabenstellungen) andererseits die postulierten Unterschiede zwischen den in alltäglichen Situationen und den in schulischen Kontexten verwendeten Registern aufzeigen. Eine Reihe weiterer Studien sowohl aus dem englischsprachigen als auch aus dem deutschsprachigen Raum stützt die Annahme, dass die linguistische Komplexität von Testaufgaben, z. B. im Rahmen von Verfahren zur Erfassung mathematischer Kompetenzen, die Aufgabenschwierigkeit beeinflusst (Haag, Heppt, Stanat, Kuhl & Pant, 2013; Martiniello, 2008; Shaftel, Belton-Kocher, Glasnapp & Poggio, 2006). Aufgaben, die vermehrt bildungssprachliche Merkmale enthalten, sind demnach für Schülerinnen und Schüler im Durchschnitt schwieriger zu lösen als Aufgaben, die vorrangig Alltagssprachliche Merkmale enthalten.

Häufig wird angenommen, dass die Beherrschung der komplexen linguistischen Anforderungen bildungssprachlicher Texte Lernende in einer Zweitsprache vor besonders große Herausforderungen stellt, da sie im Mittel über weniger Lerngelegenheiten zum Erwerb dieses spezifischen linguistischen Wissens bzw. geringer ausgeprägte linguistische Fähigkeiten verfügen als ihre Peers (Aarts, Demir & Vallen, 2011; Gogolin & Lange, 2011; Haag et al., 2013). Empirisch ergibt sich allerdings ein inkonsistentes Bild hinsichtlich der Annahme besonders ausgeprägter negativer Effekte bildungssprachlicher Merkmale auf das Textverstehen bzw. die schulischen Leistungen von L2-Schülerinnen und -Schülern. Einige Studien lieferten Hinweise darauf, dass die linguistische Komplexität, d. h. das Ausmaß in dem bildungssprachlicher Merkmale innerhalb von Testitems in mathematischen oder naturwissenschaftlichen Tests vorkommen, die Itemschwierigkeiten für L1- und L2-Schülerinnen und -Schüler unterschiedlich stark beeinflusst (z. B. Haag et al., 2013; Martiniello, 2008).

Studien, die sich mit den Effekten bildungssprachlicher Merkmale auf die Hörverstehensleistung von L1- und L2-Lernenden beschäftigten, erbrachten allerdings keine Hinweise auf differenzielle Effekte. In einer quasi-experimentellen Studie mit Grundschülerinnen und Grundschülern variierte Eckhardt (2008) beispielsweise bildungssprachliche Merkmale in Hörverstehentexten, zu deren Inhalt die Schülerinnen und Schüler nach dem Hören einige Multiple Choice-Fragen beantworten sollten. Sowohl für die L1- als auch für die L2-Kinder erwiesen sich dabei Texte, die komplexe grammatikalische Merkmale und bildungssprachliches Vokabular enthielten, als schwieriger als eher Alltagssprachlich geprägte Texte. In einer zweiten Studie konnte Eckhardt (2008) zudem zeigen, dass Hörtexte mit schulrelevanten Themen für alle Schülerinnen und Schüler schwieriger zu verarbeiten waren als Texte, die sich inhaltlich mit alltäglichen Themen beschäftigten. Auch waren dekontextualisierte Texte

(Nacherzählungen von Filmszenen) für die Kinder schwieriger zu verstehen als Texte, die in einen visuellen Kontext eingebettet waren (Filmszenen). Die von Eckhardt (2008) angenommenen differenziellen Effekte der bildungssprachlichen Merkmale auf die Hörverstehensleistung von L1- und L2-Kindern fanden sich jedoch nicht. Allerdings wurden in der Studie von Eckhardt (2008) die Merkmale bildungssprachlichen Hörverstehens getrennt voneinander variiert, so dass eine simultane Auswertung aller bildungssprachlichen Merkmale nicht möglich war. Da Bildungssprache den vorliegenden Konzeptionen zufolge durch eine ganze Reihe von Merkmalen gekennzeichnet ist, ist denkbar, dass die getrennte Operationalisierung zentraler Eigenschaften bildungssprachlicher Hörtexte nicht geeignet ist, um bildungssprachliche Hörverstehenskompetenzen zu erfassen.

Ausgehend von diesen Befunden führten Heppt, Stanat, Dragon, Berendes und Weinert (2014) eine Studie durch, bei der Alltagssprachliches und Bildungssprachliches Hörverstehen bei Grundschülerinnen und Grundschulern der zweiten und dritten Klassenstufe überprüft wurde. Die Stichprobe umfasste sowohl monolinguale deutsche Kinder, als auch Kinder, die Deutsch als Zweitsprache erwarben. Die Mehrheit dieser Kinder sprach Türkisch als Herkunftssprache. Anders als in den Studien von Eckhardt (2008) wurden in der Studie von Heppt et al. (2014) mehrere Merkmale bildungssprachlichen Hörverstehens simultan variiert. Bildungssprachliche Hörtexte zeichneten sich im Vergleich zu Alltagssprachlichen Hörverstehenstexten dabei durch eine höhere Anzahl mehrsilbiger Wörter, mehr Nominalisierungen und den Gebrauch von fachspezifischem und allgemeinem bildungssprachlichem Wortschatz (z. B. „Temperatur“, „bezeichnen“) aus. Auf Ebene der Grammatik enthielten bildungssprachliche Hörverstehenstexte mehr Wörter pro Satz, eine höhere Anzahl komplexer Nebensatzgefüge sowie mehr Passivkonstruktionen als Alltagssprachliche Hörverstehenstexte. Wie Eckhardt (2008) nahmen auch Heppt und Kolleginnen (2014) eine Interaktion zwischen dem Sprachhintergrund der Kinder und der Textart an, d. h. sie vermuteten, dass L2-Schülerinnen und -Schüler gegenüber L1-Kindern generell schwächere Hörverstehenskompetenzen aufweisen sollten, bildungssprachliche Hörverstehenstexte allerdings noch einmal disproportional schwieriger für die L2-Schülerinnen und -Schüler sein sollten. Diese Annahme überproportionaler Leistungsnachteile konnte jedoch empirisch nicht gestützt werden. Zwar erzielten die L2-Kinder durchgehend geringere Hörverstehensleistungen als die L1-Kinder und bildungssprachliche Hörverstehenstexte waren für alle Schülerinnen und Schüler schwieriger zu verarbeiten als Alltagssprachliche. Ein Interaktionseffekt zwischen dem Sprachhintergrund und den bildungssprachlichen Anforderungen der Texte fand sich jedoch nicht. Eine mögliche Erklärung für das Ausbleiben von differenziellen Effekten bildungs-

sprachlicher Merkmale auf die Hörverstehensleistung bei Kindern mit unterschiedlichem Sprachhintergrund liegt im Alter der Kinder. Sowohl Heppt et al. (2014) als auch Eckhardt (2008) vermuten, dass sich deutlichere Unterschiede in der Verarbeitung von Alltags- und Bildungssprache erst in höheren Klassenstufen finden, wenn auch die im Unterricht vermittelten Inhalte anspruchsvoller werden. Es ist daher möglich, dass sich die vermuteten besonderen Leistungsnachteile der L2-Lernenden bei bildungssprachlichen Texten erst bei älteren Schülerinnen und Schülern zeigen (Chudaske, 2012). Für Kinder mit einem niedrigen sozioökonomischen Status gehen Chall et al. (1990) davon aus, dass sich dies ungefähr ab dem Ende der Grundschulzeit, d. h. der vierten Klassenstufe manifestieren könnte, da dies dem Zeitpunkt entspricht, zu dem die in der Schule vermittelten Inhalte zunehmend weniger kontextualisiert sind.

Insgesamt liefern die bislang vorliegenden Studien zu differenziellen Auswirkungen bildungssprachlicher und alltagssprachlicher Merkmale auf die Textverstehenskompetenzen in einer Erst- bzw. Zweitsprache somit ein uneindeutiges Bild. Allerdings existieren für den deutschen Sprachraum bislang nur wenige Studien, die zudem auf den Grundschulbereich fokussieren, in dem die sprachlichen Anforderungen noch vorwiegend alltagssprachlich geprägt sein könnten. Es besteht daher die Notwendigkeit, auch in den höheren Klassen bildungssprachliches und alltagssprachliches Hörverstehen empirisch zu untersuchen. Von besonderem Interesse ist dabei die Frage danach, ob sich bildungssprachliches und alltagssprachliches Hörverstehen nicht nur in Hinblick auf die sprachlichen Merkmale, sondern auch in Bezug auf die Kompetenzen von Schülerinnen und Schülern unterscheiden lassen; bildungssprachliches und alltagssprachliches Hörverstehen also distinkte Facetten der Hörverstehensfähigkeiten bilden. Derzeit existieren noch keine Studien, die die Abgrenzbarkeit der entsprechenden Kompetenzen von Schülerinnen und Schülern empirisch überprüfen. Es bleibt somit bislang unklar, inwiefern die Annahme, Bildungssprache stelle - vor allem für Lernende einer Zweitsprache - besondere Anforderungen an die sprachlichen Kompetenzen von Schülerinnen und Schülern, tatsächlich zutreffend ist.

5. Die Bedeutung des Hörverstehens für Leseverstehen in einer Erst- bzw. Zweitsprache

Aus den bisher referierten Befunden wird deutlich, dass Schülerinnen und Schüler in einer Zweitsprache zum Teil nicht über ein vergleichbares Wissen im Bereich linguistischer Prädiktoren des Hörverstehens verfügen wie ihre L1-Peers. Es ist denkbar, dass diese Schwächen, vor allem im Wortschatz und den morpho-syntaktischen Kenntnissen, die Hör-

verstehensleistung von L2-Jugendlichen negativ beeinflussen, indem sie die Bildung einer mentalen Repräsentation von Hörtexten erschweren. In der Folge könnte auch der Erwerb weiterer schulischer Kompetenzen für L2-Schülerinnen und -Schüler gegenüber L1-Schülerinnen und Schülern erschwert sein.

Insbesondere für den Erwerb der Lesekompetenz ist ein bedeutsamer Effekt des Hörverstehens zu erwarten, der sich aus der Ähnlichkeit der beiden Textverarbeitungsprozesse ergibt, denen im Wesentlichen dieselben Verarbeitungsmechanismen zugrunde liegen (Marx & Jungmann, 2000). So müssen sowohl im Zuge des Hör- als auch während des Leseverstehensprozesses einzelne Gedanken und Konzepte aus dem Text extrahiert und in Bezug zu Vorinformationen und Hintergrundwissen gebracht werden, um in der Folge ein mentales Modell des Textinhalts erstellen zu können (Field, 2008). Unterschiede in der Verarbeitung von Texten in den beiden Modalitäten ergeben sich vor allem aus der Tatsache, dass Hörtexte aufgrund von sprachlichen Phänomenen wie Elisionen oder Assimilationen von Lauten sowie aufgrund von Hintergrundgeräuschen häufig weniger eindeutig sind als geschriebene Texte. Darüber hinaus bündelt die Verarbeitung von Hörtexten, die in Echtzeit abläuft, mehr Aufmerksamkeit als die Verarbeitung räumlich und zeitlich stabiler Texte (Buck, 2001).

Angesichts der Ähnlichkeit der grundlegenden Verarbeitungsprozesse erscheint es wenig verwunderlich, dass Hör- und Leseverstehen in starkem Ausmaß korrelieren, sobald grundlegende Lesefähigkeiten erworben wurden (Marx & Jungmann, 2000; Rost & Hartmann, 1992). Dieser enge Zusammenhang wird in dem *Simple View of Reading*-Ansatz (Hoover & Gough, 1990) aufgegriffen, demzufolge die Lesekompetenz durch zwei multiplikativ miteinander verbundene Faktoren, dem Dekodieren und dem linguistischen Verstehen, determiniert wird. Dekodieren wird dabei als effiziente Worterkennung definiert, während linguistisches Verstehen die Fähigkeit bezeichnet, einen Satz syntaktisch zu analysieren, dessen Bedeutung zu erfassen und ihn mit den umgebenden Sätzen in Verbindung zu setzen. Im *Simple View of Reading* wird diese Fähigkeit über die Hörverstehenskompetenz operationalisiert, die im Sinne eines nicht in der schriftsprachlichen Modalität stattfindenden linguistischen Verstehens verstanden werden kann. Während das Dekodieren vor allem in den frühen Phasen des Leseerwerbs den bedeutsameren Prädiktor des Leseverstehens darstellt, bildet das als Hörverstehen operationalisierte linguistische Verstehen in den späteren Phasen des Leseerwerbs, wenn das Dekodieren bereits automatisiert abläuft, die wichtigste Determinante der Lesekompetenz. Die zentralen Annahmen des *Simple View of Reading* hinsichtlich des Effekts des Hörverstehens auf das Leseverstehen können als empirisch gut belegt gelten (z. B. Droop & Verhoeven, 2003; Gottardo & Mueller, 2009; Hoover & Gough, 1990).

Schwächen im Bereich des Hörverstehens können demnach auch Schwächen im Bereich des Leseverstehens nach sich ziehen. Für L2-Lernende ist dies besonders problematisch, da vermutet werden kann, dass Hörverstehenskompetenzen und Leseverstehenskompetenzen in einer Zweitsprache besonders eng zusammenhängen. Diese Annahme basiert auf der *time-on-task*-Hypothese (vgl. z. B. Gathercole, 2002), der zufolge die Ausbildung sprachlicher Fähigkeiten wie dem Hörverstehen vom Sprachkontakt abhängt: Je weniger Kontakt Kinder zu ihrer Zweitsprache haben, desto geringer die Wahrscheinlichkeit, dass sprachliche Kompetenzen in der Zweitsprache in ausreichendem Maß erworben werden. Dies betrifft insbesondere den Wortschatz und die morpho-syntaktischen Kenntnisse und in der Folge auch das Hörverstehen, deren Ausprägung vom Ausmaß und der Dauer des Sprachkontaktes abzuhängen scheint (z. B. Cummins, 1981). Auch für die phonologische Bewusstheit ergeben sich Effekte des Sprachkontakts, allerdings zeigen sich hier in der Regel auch bei ansonsten geringen Zweitsprachkenntnissen keine Nachteile von L2-Schülerinnen und -Schülern gegenüber ihren L1-Peers (z. B. Bruck & Genesee, 1995; Campbell & Sais, 1995; Rubin & Turner, 1989). Da die phonologische Bewusstheit für das Erlernen von Buchstaben-Laut-Zuordnungen und somit der effizienten Worterkennung eher förderlich ist (Ehri, 1995), ist davon auszugehen, dass L2-Schülerinnen und -Schüler im Bereich des Dekodierens in der Regel Leistungen zeigen, die denen von L1-Schülerinnen und -Schülern entsprechen oder jene sogar übertreffen (Verhoeven & van Leeuwe, 2012). Gleichzeitig zeigen sich gegenüber monolingualen Schülerinnen und Schülern häufig Schwächen im Bereich der lexikalischen und morpho-syntaktischen Kenntnisse sowie im Textverstehen von L2-Lernenden. Da L2-Lernende somit meist über gut ausgeprägte Fähigkeiten im Dekodieren verfügen, gleichzeitig jedoch Schwächen im Hörverstehen aufweisen, wird häufig eine besonders enge Verbindung zwischen dem Hör- und dem Leseverstehen in einer Zweitsprache vermutet, das häufig ebenfalls schwächer ausgeprägt ist als in einer Erstsprache (vgl. August & Shanahan, 2006).

Empirisch fanden sich in einer Reihe von Studien Belege dafür, dass Hörverstehen in einer Zweitsprache ein bedeutsamerer Prädiktor für das Leseverstehen sein könnte als in einer Erstsprache. Für den Grundschulbereich stammen entsprechende Befunde beispielsweise aus Studien von Droop und Verhoeven (2003), Gottardo und Müller (2009) oder Limbird (2007). Auch in einer Studie von Geva und Farnia (2012) konnte gezeigt werden, dass neben dem Wortschatz und der phonologischen Bewusstheit auch das Hörverstehen bei L2-Schülerinnen und -Schülern der oberen Grundschulklassen das Leseverstehen vorhersagte, während bei den L2-Lernenden kein Effekt des Hörverstehens auf das Leseverstehen zu beobachten war. Für den deutschen Sprachraum berichtet Limbird (2007) stärkere Zusammenhänge ($r = .32$ bzw.

$r = .33$) zwischen dem Hör- und dem Leseverstehen für deutsch-türkisch bilinguale Kinder der Klassenstufen 2 und 3 als für monolingual deutschsprachige Kinder ($r = .19$ bzw. $r = .14$). Während die Ergebnisse dieser Studien somit für die theoretisch postulierten stärkeren Zusammenhänge zwischen Hör- und Leseverstehen in einer Zweitsprache sprechen, berichtet Babayiğit (2014) gegenteilige Befunde aus einer Untersuchung von L1- und L2-Grundschulern in England. In dieser Studie fanden sich keine differenziellen Zusammenhänge zwischen dem Hör- und dem Leseverstehen bei L1- und L2-Kindern, die die Annahme einer stärkeren Rolle von Hörverstehensfähigkeiten für das Leseverstehen in einer Zweitsprache stützen würden.

Bei der Interpretation dieser Ergebnisse ist zu beachten, dass in den meisten der genannten Studien der sozioökonomische Hintergrund der Kinder nicht kontrolliert wurde. Lediglich in der Studie von Droop und Verhoeven (2003) wurde der sozioökonomische Status der beteiligten Schülerinnen und Schüler berücksichtigt. Limbird (2007) gibt an, dass alle Kinder aus Schulbezirken mit einem ähnlich niedrigen sozioökonomischen Status kamen und Babayiğit (2014) berichtet auf Basis des Anteils von Schülerinnen und Schülern, die am *free lunch*-Programm der jeweiligen Schule teilnahmen, von einem vermutlich geringeren sozioökonomischen Hintergrund der L2-Kinder, der in den Analysen jedoch nicht kontrolliert wird. Die Ergebnisse dieser Studien könnten somit mit dem oftmals unterschiedlichen sozioökonomischen Status der L1- und L2-Kinder konfundiert sein, der ebenfalls einen bedeutsamen Prädiktor des Leseverstehens darstellt.

Obwohl die Mehrzahl bisheriger Studien insgesamt Unterstützung für die Annahme einer besonders engen Verbindung zwischen Hör- und Leseverstehen in einer Zweitsprache liefert, besteht daher die Notwendigkeit der Durchführung weiterer Studien, die potenziell konfundierte Variablen wie den sozioökonomischen Hintergrund der Schülerinnen und Schüler berücksichtigen. Zudem besteht weiterer Forschungsbedarf in Hinblick auf Zusammenhänge zwischen Hör- und Leseverstehen bei älteren L1- und L2-Lernenden nach der Grundschulzeit. Da der Zusammenhang zwischen dem Hör- und dem Leseverstehen nach dem Erwerb automatisierter Dekodierfähigkeiten zunehmend enger wird (vgl. Kapitel 5), kann vermutet werden, dass sich differenzielle Zusammenhänge zwischen Hör- und Leseverstehen von L1- bzw. L2-Schülerinnen und -Schülern erst in den höheren Klassenstufen finden lassen, wenn die Dekodierfähigkeiten im Allgemeinen bereits automatisiert sind und Hörverstehen der wesentliche Prädiktor für die Lesefähigkeiten wird.

6. Anliegen dieser Arbeit

Trotz der zentralen Bedeutung des Hörverstehens für die weitere sprachliche und schulische Entwicklung von Schülerinnen und Schülern mangelt es derzeit an empirischen Untersuchungen zur Struktur und Determination von Hörverstehensfähigkeiten in einer Zweitsprache (vgl. Bonk, 2000; Eckhardt, 2008). Es kann daher zum derzeitigen Zeitpunkt nicht abgeschätzt werden, inwiefern Schwächen im Hörverstehen bei Schülerinnen und Schülern in einer Zweitsprache ein Resultat von geringer ausgeprägten Prädiktoren des Hörverstehens sind bzw. ob die einzelnen Prädiktoren bei Schülerinnen und Schülern unterschiedlicher Herkunftssprache möglicherweise differenziell bedeutsam für die Hörverstehensleistung sind. Auch die Bedeutung des Hörverstehens für das zweitsprachliche Leseverstehen ist in der Vergangenheit vornehmlich für Leseanfänger oder L2-Lernende in den frühen Phasen des Zweitspracherwerbs untersucht worden. Derzeit ist wenig darüber bekannt, welchen Effekt das Hörverstehen auf die Lesekompetenz bei jugendlichen L2-Schülerinnen und -Schülern hat, die ihre gesamte Schullaufbahn im zweitsprachlichen Kontext durchlaufen haben. Zudem mangelt es bislang an Studien, die die Struktur zweitsprachlichen Hörverstehens bei älteren Schülerinnen und Schülern in Hinblick auf die Abgrenzbarkeit von alltags- und bildungssprachlichem Hörverstehen untersuchen. In der Folge bleibt bislang unklar, ob die in Large-Scale Assessments wie PISA oder dem IQB-Ländervergleich berichteten Schwächen von L2-Schülerinnen und -Schülern im Bereich des Hörverstehens nur das bildungssprachliche Hörverstehen oder darüber hinaus auch ihre alltagssprachlichen Hörverstehenskompetenzen betreffen.

Vor diesem Hintergrund werden in der vorliegenden Arbeit die Struktur und die Ausprägung der Hörverstehensfähigkeiten von Jugendlichen mit Deutsch als Erst- bzw. Zweitsprache sowie Effekte der Hörverstehensfähigkeiten auf die Lesekompetenz vergleichend untersucht. Durch die Berücksichtigung des sozioökonomischen Status der Schülerinnen und Schüler in den Analysen, ist es dabei auch möglich, Effekte des sprachlichen Hintergrunds der L2-Schülerinnen und -Schülern von jenen zu isolieren, die aus ihrem gegenüber L1-Jugendlichen häufig geringeren sozioökonomischen Status resultieren. Aufgrund der besonders benachteiligten Situation von Schülerinnen und Schülern mit türkischer Herkunftssprache in Bezug auf die Ausprägung schulischer Kompetenzen (vgl. Kapitel 1.1) liegt das Hauptaugenmerk der Analysen auf dieser Gruppe.

7. Ziele, Methoden und Ergebnisse der Einzelstudien

Für die vorliegende Arbeit wurden vier empirische Teilstudien durchgeführt. In Teilstudie 1 wird untersucht, ob verschiedene linguistische Komponenten des Hörverstehens in einer Erst- bzw. Zweitsprache unterschiedlich prädiktiv für die Hörverstehenskompetenzen sind. In diesem Zusammenhang werden auch Unterschiede in der Ausprägung von Hörverstehenskompetenzen und dessen Prädiktoren zwischen Jugendlichen mit Deutsch als Erstsprache (*L1*) bzw. Deutsch als Zweitsprache (*L2*) untersucht. Teilstudie 2 geht der Frage nach möglicherweise differenziellen Effekten des Hörverstehens für die Lesekompetenz in einer Erst- bzw. Zweitsprache nach. Berücksichtigt werden dabei auch Effekte des sozioökonomischen und kulturellen Hintergrunds auf die Hörverstehensleistung. In Teilstudie 3 wird überprüft, ob sich alltagssprachliche Hörverstehensleistungen von bildungssprachlichen Hörverstehensleistungen abgrenzen lassen und ob bildungssprachliches Hörverstehen für Schülerinnen und Schüler mit Deutsch als Zweitsprache eine besondere Hürde darstellt. Aufbauend auf dieser Studie wird in einer Zusatzanalyse (Teilstudie 4) der Frage nachgegangen, ob sich bildungssprachlicher und alltagssprachlicher Wortschatz voneinander abgrenzen lassen und inwiefern beide Wortschatzfacetten differenziell bedeutsam für alltags- bzw. bildungssprachliches Hörverstehen in einer Erst- und einer Zweitsprache sind.

7.1 Methodisches Vorgehen der Einzelstudien

Die Daten der Teilstudien wurden im Rahmen der von der Deutschen Forschungsgemeinschaft (DFG) geförderten Studie „Determinanten der Lesekompetenz“ (DELKO; DFG-Geschäftszeichen STA 626/1-6 und SCHN 315/40-1) erhoben, in der Bedingungsfaktoren schwacher Lesekompetenz bei Schülerinnen und Schülern mit Deutsch als Erstsprache bzw. Deutsch als Zweitsprache vertiefend untersucht wurden. Die DELKO-Studie wurde im Mai/Juni 2009 als nationale Ergänzungsstudie zur PISA 2009-Erhebung durchgeführt. Im Folgenden wird zunächst die allen Studien zugrundeliegende Stichprobe beschreiben. Anschließend werden die Ziele und die eingesetzten Testverfahren der Einzelstudien beschreiben, ehe die Analysestrategie und der Umgang mit fehlenden Werten gemeinsam für alle Teilstudien abgehandelt werden. Abschließend werden die Ergebnisse der Einzelstudien zusammengefasst.

7.1.1 Stichprobe

Insgesamt nahmen an der DELKO-Studie 479 Jugendliche (44.9% weiblich) der 9. Klassenstufe aus 20 Schulen in verschiedenen Bundesländern teil. Das mittlere Alter lag bei 16;1 Jah-

ren ($s = 0.77$ Jahre). Aufgrund des Fokus der DELKO-Studie auf Schülerinnen und Schüler mit schwach ausgeprägten Lesekompetenzen wurden für die Teilnahme staatliche Haupt-, Real- oder Gesamtschulen in Ballungsgebieten ausgewählt, bei denen ein hoher Anteil an Jugendlichen mit schwachen Lesekompetenzen zu erwarten war. Zudem wurden vorrangig Schulen mit einem hohen Anteil ($> 40\%$) von Schülerinnen und Schülern mit Zuwanderungshintergrund sowie einem Anteil von mind. 15% an Schülerinnen und Schülern mit türkischer Herkunftssprache in die Stichprobe aufgenommen. Um abschätzen zu können, inwiefern eventuelle Unterschiede in der Ausprägung und dem Zusammenspiel von Prädiktoren des Hörverstehens zwischen türkisch-deutschsprachigen und deutschsprachigen Schülerinnen und Schülern auf die spezifische Herkunftssprache zurückzuführen sind, wurde zusätzlich zur Gruppe von L2-Jugendlichen mit türkischer Herkunftssprache ($L2_{Türk}$) eine weitere L2-Gruppe mit einer anderen Herkunftssprache als Türkisch untersucht ($L2_{Andere}$). Die Herkunftssprache wurde dabei anhand eines Items im Schülerfragebogen erfasst („*Welche Sprache hast du in deiner Familie zuerst gelernt?*“). Auf Basis dieses Vorgehens wurden 179 (37,4%) der Jugendlichen als L1-Schülerinnen und -Schüler identifiziert, 186 (38,8%) Schülerinnen und Schüler als L2-Jugendliche mit türkischer Herkunftssprache ($L2_{Türk}$) und 114 (23,8%) der Schülerinnen und Schüler als L2-Jugendliche mit einer anderen Herkunftssprache als Türkisch ($L2_{Andere}$). Innerhalb der $L2_{Andere}$ -Gruppe waren die häufigsten von den Jugendlichen genannten Erstsprachen Italienisch (11,4%), Arabisch (11,4%) und Polnisch (9,6%).

In beiden L2-Gruppen gab die Mehrheit der Schülerinnen und Schüler an, in Deutschland geboren worden zu sein, aber im Ausland geborene Eltern zu haben (86% für $L2_{Türk}$ und 61% für $L2_{Andere}$) und somit zur zweiten Zuwanderergeneration zu gehören (vgl. OECD, 2010). Der entsprechende Prozentsatz war in der Gruppe der $L2_{Andere}$ -Jugendlichen geringer ausgeprägt als in der Gruppe der $L2_{Türk}$ -Jugendlichen ($\chi^2_{(1)} = 20.73$, $p = <.01$; $\phi = .29$), d. h. bei den $L2_{Andere}$ -Jugendlichen gaben mehr Schülerinnen und Schüler an, selbst nach Deutschland zugewandert zu sein und somit der ersten Zuwanderergeneration anzugehören als in der Gruppe der $L2_{Türk}$ -Jugendlichen. Der Anteil von Schülerinnen und Schülern, die die deutsche Sprache vor dem Schuleintritt (mit 6 Jahren) erworben hatten, unterschied sich zwischen den beiden L2-Gruppen allerdings nicht: Sowohl $L2_{Türk}$ -Jugendliche als auch $L2_{Andere}$ -Jugendliche gaben zum überwiegenden Teil an, bereits vor der Einschulung die deutsche Sprache erlernt zu haben (93% bzw. 87 %; $\chi^2_{(1)} = 3.39$, $p = .34$, $\phi = .14$). Hinsichtlich der Unterstützung der L2-Schülerinnen und -Schüler beim Erwerb schulischer Kompetenzen berichteten 51% der teilnehmenden Schulen, Förderunterricht für alle L1- und L2-Schülerinnen und -Schüler mit entsprechendem Bedarf anzubieten. 41% der Schulen boten Vorbereitungsklassen für L2-

Schülerinnen und -Schüler an, aber nur eine der zwanzig teilnehmenden Schulen bot muttersprachlichen Unterricht an. Insgesamt 35% der Schulen berichteten, L2-Schülerinnen und -Schüler in Klassen mit einer geringeren Klassenstärke zu beschulen als L1-Schülerinnen und Schüler.

Tabelle 1.

Mittleres Alter und mittlere kognitive Grundfähigkeiten für die in der DELKO-Studie betrachteten Sprachgruppen

	L1 _{Deutsch} (<i>n</i> =179) <i>M</i> (<i>SD</i>)	L2 _{Türk} (<i>n</i> =186) <i>M</i> (<i>SD</i>)	L2 _{Andere} (<i>n</i> =114) <i>M</i> (<i>SD</i>)	<i>F</i>	<i>p</i>	η^2
Alter in Jahren	16.01(0.75)	16.04 (0.77)	16.14 (0.81)	0.98	.38	.00
kognitive Grundfähigkeiten (z-Werte)	0.11 (0.98)	-0.13 (1.01)	0.07 (0.99)	2.49	.09	.01

Die drei in den vorliegenden Studien betrachteten Schülergruppen unterschieden sich in ihrer Geschlechterzusammensetzung nicht voneinander ($\chi^2 = 1.13$; $p = .57$). Auch in Bezug auf das mittlere Alter fanden sich keine Sprachgruppenunterschiede (vgl. Tabelle 1). Zur Erfassung der kognitiven Grundfähigkeiten der Schülerinnen und Schüler wurde in der DELKO-Studie der Subtest N2 der revidierten Fassung des Kognitiven Fähigkeitstest für 4. bis 12. Klassen (KFT 4-12+R; Heller & Perleth, 2000) eingesetzt, der das Erkennen figuraler Analogien erfordert. Für die kognitiven Grundfähigkeiten ergab sich zwar ein kleiner Effekt des Sprachhintergrunds, allerdings ist die entsprechende Differenz in den kognitiven Grundfähigkeiten der Sprachgruppen nicht signifikant (vgl. Tabelle 1).

Der sozioökonomische Hintergrund der Schülerinnen und Schüler wurde mittels des auch in den PISA-Studien verwendeten *Index des ökonomischen, sozialen und kulturellen Status* (ESCS; vgl. OECD, 2006) erfasst. Für die DELKO-Studie wurde der ESCS anhand des Mittelwerts der OECD-Population zentriert und an deren Standardabweichung standardisiert. Es zeigte sich, dass alle Schülerinnen und Schüler der Stichprobe aus Familien mit einem geringen sozioökonomischen Status stammten. Der Mittelwert des ESCS lag mit $M = -0.63$ ($s = 0.92$) mehr als eine halbe Standardabweichung unter dem deutschen Durchschnitt in der PISA 2009-Studie (OECD, 2010). Zwischen den Sprachgruppen fanden sich dabei signifikante Unterschiede in der Ausprägung des ESCS ($F_{(2; 363)} = 12.09$; $p < .01$; $\eta^2 = .06$): L2_{Türk}-Schülerinnen und -Schüler ($M = -0.86$; $s = 0.92$) kamen aus Familien mit einem geringeren sozioökonomischen Status als L1-Schülerinnen und -Schüler ($M = -0.43$; $s = 0.79$) und

L2_{Andere}-Jugendliche ($M = -0.62$; $s = 0.85$). Der ESCS unterschied sich zwischen L1- und L2_{Andere}-Jugendlichen nicht ($F_{(1; 218)} = 3.27$; $p = .07$; $\eta^2 = .01$.)

7.1.2 Teilstudie 1: Prädiktoren des Hörverstehens bei Jugendlichen deutscher und Jugendlichen nichtdeutscher Herkunftssprache (Marx & Roick, 2012)

Das Ziel von Teilstudie 1 bestand in einer Überprüfung der Effekte sprachlicher Prädiktoren auf die Hörverstehensleistung bei Jugendlichen unterschiedlicher Herkunftssprache. Ausgehend von der Annahme, dass Hörende in einer Zweitsprache im Vergleich zu L1-Hörenden in verstärktem Maße auf *bottom-up* basierte Verarbeitungsstrategien zurückgreifen, sollten Effekte basaler sprachlicher Prädiktoren auf das Hörverstehen in einer Erst- bzw. Zweitsprache vertiefend untersucht werden. In das Modell des Hörverstehens aufgenommen wurden dabei einerseits auf Ebene der wahrnehmungsbasierten Verarbeitung die phonologische Bewusstheit, sowie andererseits die morpho-syntaktischen Kenntnisse auf der Ebene des Parsings. Weiterhin gingen das Arbeitsgedächtnis und der Wortschatz ein, die den Hörverstehensprozess sowohl auf der wahrnehmungsbasierten Ebene als auch auf der Ebene des Parsing beeinflussen (vgl. Abbildung 4).

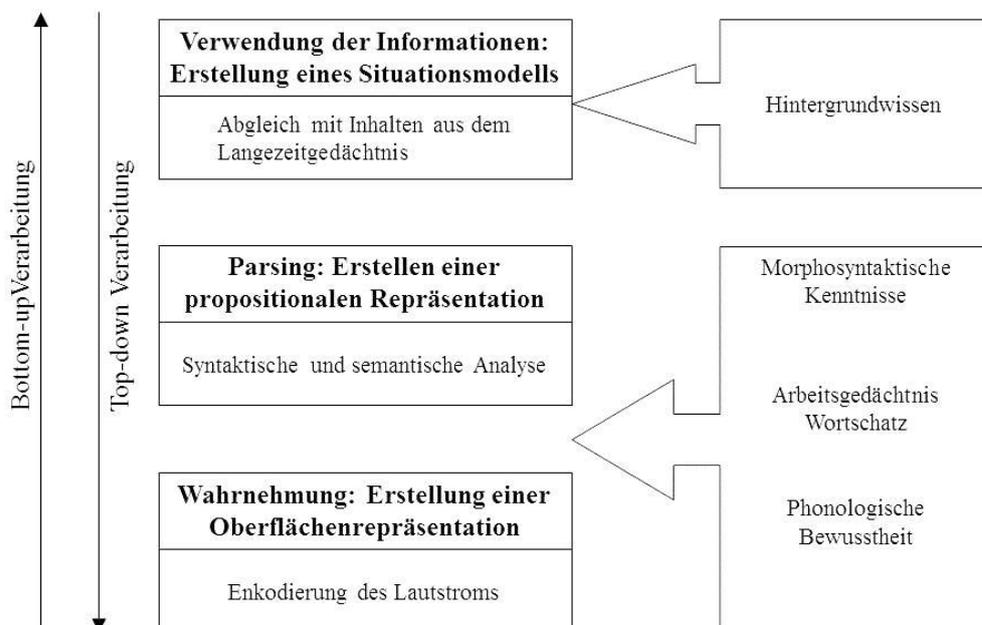


Abbildung 4. Modell des Hörverstehens und seiner Prädiktoren, angelehnt an Anderson (1995) bzw. Kintsch und van Dijk (1978)

In Teilstudie 1 wurden folgende Forschungsfragen untersucht:

- 1.1 Unterscheidet sich die Ausprägung linguistischer Prädiktoren des Hörverstehens zwischen L1- und L2-Schülerinnen und -Schülern?

Ausgehend von der *time-on-task*-Hypothese ist zu vermuten, dass L2-Schülerinnen und -Schüler aufgrund des häufig geringeren Sprachkontakts zur Zweitsprache über geringere sprachliche Fähigkeiten verfügen als L1-Jugendliche (vgl. Gathercole, 2002). Auf Basis dieser Hypothese sowie bisheriger empirischer Befunde aus Studien mit jüngeren Schülerinnen und Schülern (z. B. Bruck & Genesee, 1995; Haag et al., 2012; Kieffer & Vukovic, 2013; Verhoeven, 1990) wurde angenommen, dass L2-Jugendliche auch gegen Ende ihrer schulischen Ausbildung schwächere Leistungen im Hörverstehen, dem Wortschatz sowie den morpho-syntaktischen Fähigkeiten aufweisen als L1-Jugendliche. Für die phonologische Bewusstheit wurden keine Gruppenunterschiede angenommen, da diese als Aspekt metasprachlicher Bewusstheit durch den Kontakt zu mehreren Sprachen nicht beeinträchtigt werden sollte (Bialystok, 2009; Bruck & Genesee, 1995). Ausgehend von Befunden, die darauf hindeuten, dass die Arbeitsgedächtniskapazität in einer Erst- bzw. Zweitsprache ähnlich ausgeprägt ist, sofern sprachfreies Material zur Erfassung der Arbeitsgedächtniskapazität verwendet wird (vgl. z. B. Olsthoorn et al., 2012), wurden für diesen Prädiktor des Hörverstehens ebenfalls keine differenziellen Leistungen der L1- bzw. L2-Schülerinnen und -Schüler erwartet.

- 1.2 Unterscheiden sich die Effekte linguistischer Prädiktoren des Hörverstehens auf die Hörverstehensleistung in einer Erst- bzw. Zweitsprache?

Da die Identifizierung von einzelnen Wörtern im Sprachfluss für L2-Lernende eine größere Hürde darstellen sollte als in einer Erstsprache (vgl. z. B. Stæhr, 2009), wurden besonders starke Effekte des Wortschatzes und der phonologischen Bewusstheit auf das Hörverstehen in einer Zweitsprache erwartet. Für das Arbeitsgedächtnis wurde für die L2-Jugendlichen ebenfalls ein im Vergleich zu den L1-Jugendlichen stärkerer Effekt auf die Hörverstehensleistung angenommen, der sich daraus ergibt, dass L2-Lernende aufgrund ihrer schwächeren sprachlichen Fähigkeiten dem linguistischen Gehalt einer Hörbotschaft verstärkt Aufmerksamkeit widmen müssen (Anderson, & Lynch, 1988; Færch & Kasper, 1986; Rost, 2011; Voss, 1984). Aus theoretischer Sicht besteht keine Veranlassung zur Annahme differenzieller Effekte der morpho-syntaktischen Kenntnisse auf das Hörverstehen in einer Erst- bzw. Zweitsprache. Ausgehend von Befunden, die für Grundschülerinnen und -schüler bzw. Lernende in den frühen Phasen des L2-Erwerbs darauf hindeuten, dass den morpho-syntaktischen Kenntnissen in einer Zweitsprache

keine besonders ausgeprägte Rolle für die Hörverstehensleistung zukommt (z. B. Droop & Verhoeven, 2003; Felser & Clahsen, 2009), wird allerdings angenommen, dass der entsprechende Zusammenhang bei L2-Jugendlichen geringer ausfällt als bei L1-Jugendlichen.

An Teilstudie 1 nahmen alle 479 Schülerinnen und Schüler der DELKO-Stichprobe teil (vgl. Kapitel 7.1.1). Die Überprüfung der Hörverstehenskompetenzen erfolgte mittels eines computerbasierteren Tests zur Erfassung basaler Hörverstehensfähigkeiten bei Jugendlichen, der im Rahmen der DELKO-Studie entwickelt worden war ($\alpha = .78$). Die phonologische Bewusstheit der Schülerinnen und Schüler wurde anhand computerbasierter Skalen zur Erfassung der Fähigkeit zur *Lautkategorisierung* (11 Items, $\alpha = .71$) bzw. der *Vokallängenbestimmung* (15 Items, $\alpha = .77$) operationalisiert. Für die Überprüfung der morphosyntaktischen Kenntnisse wurde ein computergestütztes Verfahren mit Unterskalen zur Erfassung der *syntaktischen Bewusstheit* (22 Items, $\alpha = .68$) sowie der *morphologischen Bewusstheit* (25 Items; $\alpha = .83$) entwickelt. Die Arbeitsgedächtniskapazität wurde anhand von visuellen Ziffernspannen überprüft, die als sprachfreies Maß kulturell fair sind (Olsthoorn, Andringa, & Hulstijn, 2012). Die Leistung der phonologischen Schleife wurde mittels *Ziffernspannen vorwärts* erfasst, zur Erhebung der zentral-exekutiven Funktion des Arbeitsgedächtnisses wurden *Ziffernspannen rückwärts* eingesetzt. Die interne Konsistenz der Ziffernspannen lag bei $\alpha = .66$ für die Vorwärtsspannen und $\alpha = .84$ für die Rückwärts-spannen. Die beiden Spannenmaße korrelierten mit $r = .44$ ($p < .01$). Die Überprüfung der Wortschatzkenntnisse erfolgte anhand der Wortschatzerweiterung des Grundintelligenztests Skala 2 (CFT 20 R-WS; Weiß, 2006). Der anhand des ESCS operationalisierte sozio-ökonomische Status der Schülerinnen und Schüler ging als manifeste Kovariate in die Analysen ein.

Die Überprüfung der Fragestellungen erfolgte anhand eines strukturanalytischen Mehrgruppenmodells, in dem die Variablen auf latenter Ebene modelliert wurden. Zur Schätzung der latenten Variablen wurden je Konstrukt zwei Indikatoren über eine odd-even Item-Parcelung gebildet und auf Gleichheit der Residualvarianz restringiert. Für die Schätzung der Modellparameter wurde ein *Maximum Likelihood-Schätzer* verwendet.

7.1.3 Teilstudie 2: Prädiktoren der Lesekompetenz in einer Erst- bzw. Zweitsprache (Marx, Stanat, Roick, Segerer, Marx & Schneider, 2015)

Ziel von Teilstudie 2 war die Überprüfung der Effekte von Prädiktoren der Lesekompetenz, insbesondere des Hörverstehens, auf das Leseverstehen bei leseschwachen Jugendlichen in einer Erst- bzw. Zweitsprache. Zentrale Grundlage dieser Untersuchung war der *Simple View of Reading*-Ansatz (Hoover & Gough, 1990), demzufolge Hörverstehen einen starken Effekt auf die Lesekompetenz bei Schülerinnen und Schülern haben sollte, die bereits über automatisierte Fähigkeiten zur Worterkennung verfügen. In Teilstudie 2 wurde überprüft, inwiefern Unterschiede in der Ausprägung und im Zusammenhang zwischen Hörverstehen und Lesekompetenz bestehen. Berücksichtigt wurde dabei auch der sozioökonomische Hintergrund von L1- und L2-Schülerinnen und -Schülern, der zur Erklärung differenziell ausgeprägter sprachlicher Kompetenzen beitragen könnte und daher als Kontrollvariable in die Analysen einging. In Teilstudie 2 wurden somit die folgenden Forschungsfragen überprüft:

2.1 Finden sich differenziell ausgeprägte Leistungen im Hör- und Leseverstehen von L1- und L2-Jugendlichen?

Auf Grundlage theoretischer Annahmen und bisheriger empirischer Befunde aus dem Grundschulbereich wurde vermutet, dass L2-Schülerinnen und -Schüler auch am Ende der schulpflichtigen Zeit schwächere Leistungen im Hör- sowie im Leseverstehen aufweisen als ihre L1-Peers (vgl. z. B. Eckhardt, 2008; Heppt et al., 2012). Basierend auf den Ergebnissen bisheriger Schulleistungsstudien, in denen sich wiederholt besonders schwache schulische Leistungen bei L2-Schülerinnen und -Schülern mit türkischer Herkunftssprache fanden (vgl. Kapitel 1.1), wurde angenommen, dass diese Leistungsunterschiede für L2_{Türk}-Jugendliche stärker ausgeprägt sind als für L2_{Andere}-Jugendliche. Für die Lesegeschwindigkeit, die insbesondere bei älteren Schülerinnen und Schülern als Maß für das Dekodieren betrachtet werden kann (vgl. Savage, 2006; Wolf & Katzir-Cohen, 2001), wurden keine Gruppenunterschiede erwartet, da das Dekodieren eng mit der phonologischen Bewusstheit zusammenhängt, die durch den Kontakt zu mehreren sprachlichen Systemen nicht beeinträchtigt werden sollte (vgl. z. B. Bialystok, Majumder & Martin, 2003; Koda, 1998). Für übrigen Prädiktoren des Hörverstehens bzw. der Lesegeschwindigkeit (Wortschatz, morpho-syntaktische Kenntnisse, Arbeitsgedächtnisleistung und phonologische Bewusstheit) entsprachen die Annahmen den in Teilstudie 1 unter Angabe der relevanten theoretischen und empirischen Grundlagen genannten Hypothesen. Da der sozioökonomische Hintergrund eng mit dem Erwerb lin-

guistischer Fähigkeiten verbunden ist und L2-Schülerinnen und -Schüler in Deutschland häufig aus Familien mit einem geringeren sozioökonomischen Hintergrund stammen als ihre L1-Peers, wurde angenommen, dass Unterschiede in den sprachlichen Kompetenzen zwischen L1- und L2-Jugendlichen unter Kontrolle des sozio-ökonomischen Status geringer ausfallen als ohne eine entsprechende Kontrolle.

2.2 Lassen sich Unterschiede im Zusammenhang zwischen Hörverstehens- und Leseverstehenskompetenzen bei L1- und L2-Jugendlichen finden?

Obwohl sich in der Forschungsliteratur die Vermutung findet, Hörverstehenskompetenzen seien in einer Zweitsprache besonders eng mit dem Leseverstehen verbunden (vgl. August & Shanahan, 2006), existiert bislang keine theoretische Annahme, die ein solches differenzielles Zusammenhangsmuster zwischen Hör- und Lesekompetenzen für L1- bzw. L2-Lernende erklären könnte. Vor diesem Hintergrund wurde angenommen, dass der Zusammenhang zwischen Hör- und Leseverstehen in einer Erst- und in einer Zweitsprache ähnlich stark ausgeprägt sein sollte.

In Teilstudie 2 wurden alle 479 Schülerinnen und Schüler der DELKO-Stichprobe einbezogen (vgl. Kapitel 7.1.1). Zur Überprüfung der Forschungsfragen wurde ein latentes Strukturgleichungsmodell berechnet, in das neben dem Hörverstehen und der Lesegeschwindigkeit - als Maß für die Dekodierfähigkeit bei Jugendlichen - auch weitere linguistische Prädiktoren der Lesekompetenz eingingen (vgl. Abbildung 5). Das Leseverstehen wurde mittels des in der PISA 2009-Studie verwendeten Lesetests erfasst (vgl. OECD, 2006). Dabei bearbeiteten die Jugendlichen in der DELKO-Studie ein sogenanntes „easy set“ von Items, das auf der generellen PISA-Lesekompetenzskala verankert werden kann, aber im unteren Leistungsbereich genauer differenziert. Die Antworten der Schülerinnen und Schüler wurden analog zu dem in PISA verwendeten Vorgehen zunächst mittels *Item Response Techniques* skaliert, anschließend wurden fünf *Plausible Values* für jeden Jugendlichen gebildet und gingen als Schätzer der Lesekompetenz in die Analysen ein (vgl. OECD, 2006).

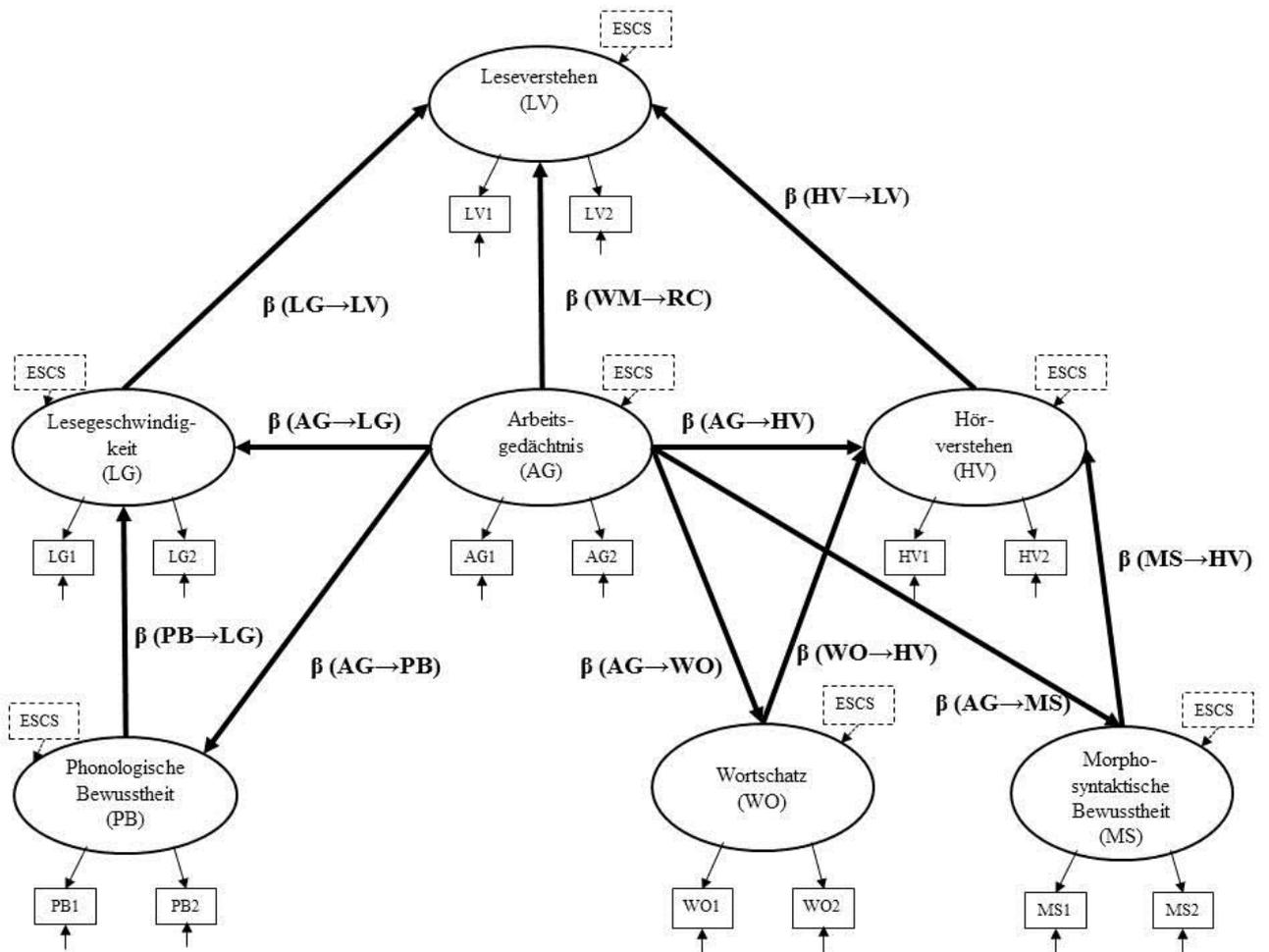


Abbildung 5. Theoretisches Modell des Leseverstehens

Zur Erfassung der Lesegeschwindigkeit als Aspekt des Dekodierens für ältere Schülerinnen und Schüler (vgl. Savage, 2006) wurde der *Lesegeschwindigkeits- und Verständnistest* (LGVT 6-12; Schneider, Schlagmüller & Ennemoser, 2007) eingesetzt. Da in der vorliegenden Studie nur die Lesegeschwindigkeit, nicht aber die Leseverstehensleistung durch den LGVT erfasst werden sollte, wurden nur für die Lesegeschwindigkeit Scores gebildet, die anhand der Anzahl korrekt bearbeiteter Verstehensaufgaben verifiziert wurden. Die Reliabilität des LGVT lag in der DELKO-Studie bei $\alpha = .72$. Die Operationalisierung der Hörverstehenskompetenzen, der morpho-syntaktischen Kenntnisse, des Wortschatzes, der phonologischen Bewusstheit, des Arbeitsgedächtnisses sowie des sozioökonomischen Hintergrunds erfolgte analog zu Teilstudie 1.

Die Fragestellungen von Teilstudie 2 wurden anhand latenter Mehrgruppenmodelle überprüft, wobei für die latenten Konstrukte je zwei Parcels über eine odd-even Item Parcelung gebildet wurden. Ausnahmen bildeten dabei das Leseverstehen, für das die fünf Plausible Values als Parcels eingingen, sowie die Lesegeschwindigkeit, für die die gewichteten Scores

der nach 4 Minuten und nach acht Minuten gelesenen Wörter als Parcels eingingen. Um die Rolle des sozioökonomischen Hintergrunds auf die Ausprägung von bzw. die Zusammenhänge zwischen Hör- und Leseverstehen und deren Prädiktoren abschätzen zu können, wurde zunächst ein Strukturgleichungsmodell ohne Kovariate berechnet, bevor anschließend ein zweites Modell berechnet wurde, in das der ECSS als manifeste Kovariate einging. Da die untersuchten linguistischen Prädiktoren des Leseverstehens sowie Hör- und Leseverstehen selbst eng mit der Kapazität der phonologischen Schleife und der zentralen Exekutive assoziiert sind (vgl. z. B. Baddeley, Gathercole & Papagno, 1998; Call, 1985; Gathercole & Baddeley, 1989), wurde in allen Analysen die Arbeitsgedächtniskapazität der Schülerinnen und Schüler kontrolliert.

7.1.4 Teilstudie 3: Bildungssprachliches und alltagssprachliches Hörverstehen in einer Erst- bzw. Zweitsprache (Marx, Heppt & Henschel; in Revision)

Das Ziel von Teilstudie 3 bestand in der Untersuchung alltagssprachlicher und bildungssprachlicher Hörverstehenskompetenzen in einer Erst- bzw. Zweitsprache. Bisherige Studien zu Aspekten bildungssprachlicher Kompetenzen fokussierten vorwiegend auf Grundschülerinnen und -schüler, so dass bislang wenig über die Ausprägung bildungssprachlicher Kompetenzen bei Jugendlichen in den höheren Klassenstufen bekannt ist, in denen die bildungssprachlichen Anforderungen ansteigen. Zudem liegen derzeit keine Studien vor, die untersuchen, ob sich die im Alltag verwendete Sprache und Bildungssprache auch auf der Ebene sprachlicher Kompetenzen unterscheiden lassen und voneinander abgrenzbare Facetten des Sprachverstehens bilden. Vor diesem Hintergrund wurden in Teilstudie 3 die Hörverstehenskompetenzen in einer Erst- bzw. Zweitsprache strukturanalytisch untersucht. Dabei wurden folgende Forschungsfragen überprüft:

3.1 Lassen sich alltagssprachliche und bildungssprachliche Hörverstehenskompetenzen voneinander abgrenzen?

Auf Basis theoretischer Annahmen zum Konstrukt der Bildungssprache wurde angenommen, dass ein zweifaktorielles Modell des Hörverstehens, das alltags- und bildungssprachliche Hörverstehenskompetenzen als getrennte Facetten des Hörverstehens differenziert, eine bessere Anpassung an die Daten aufweist als ein einfaktorielles Modell mit einem globalen Hörverstehensfaktor. Da bildungssprachliches Hörverstehen höhere Anforderungen an die Wortschatzkenntnisse und morpho-syntaktischen Fähigkeiten von Schülerinnen und Schülern stellt, wurde zusätzlich angenommen, dass diese Kompeten-

zen enger mit dem bildungssprachlichen Hörverstehen zusammenhängen sollten als mit dem alltagssprachlichen Hörverstehen.

3.2 Existieren Leistungsunterschiede im alltagssprachlichen bzw. bildungssprachlichen Hörverstehen zwischen L1- und L2-Schülerinnen und -Schülern?

Für L2-Schülerinnen und -Schüler wurden generell schwächere Hörverstehensleistungen als für ihre L1-Peers erwartet. Zudem wurde für beide Sprachgruppen angenommen, dass die bildungssprachlichen Hörverstehensleistung aufgrund der Komplexität bildungssprachlicher Anforderungen geringer ausgeprägt sein sollten als das alltagssprachliche Hörverstehen. Für die L2-Schülerinnen und -Schüler wurden für das bildungssprachliche Hörverstehen besonders ausgeprägte Leistungsnachteile vermutet, die sich daraus ergeben sollten, dass diese Jugendlichen das entsprechende Register in ihrer Zweitsprache erwerben müssen.

In Teilstudie 3 wurden alle 479 Schülerinnen und Schüler der DELKO-Stichprobe einbezogen. Für 20 Jugendliche lagen auf keiner der betrachteten Variablen Daten vor, so dass diese Schülerinnen und Schüler von den weiteren Analysen ausgeschlossen wurden. Die Zuweisung zu den Untersuchungsgruppen erfolgte analog zu Studie 1 und 2 anhand der von den Jugendlichen angegebenen Herkunftssprache. Im Gegensatz zu Teilstudie 1 und 2 liegen keine Annahmen bezüglich der Effekte konkreter Herkunftssprachen auf die Ausprägung alltagssprachlicher bzw. bildungssprachlicher Kompetenzen vor. Daher wurde bei den Auswertungen neben einer L1-Sprachgruppe nur eine globale L2-Sprachgruppe berücksichtigt.

Zur Erfassung des Wortschatzes und der morpho-syntaktischen Fähigkeiten wurden die für Teilstudie 1 beschriebenen Instrumente eingesetzt. Die Operationalisierung alltagssprachlicher und bildungssprachlicher Hörverstehensfähigkeiten erfolgte ebenfalls mittels des in Teilstudie 1 beschriebenen Hörverstehenstests, der für Teilstudie 3 in Hinblick auf die Bildungssprachlichkeit der Texte weiter differenziert wurde. Bildungssprachliche Units des Hörverstehenstests waren dabei eher in formellen, akademischen Settings (z. B. Bewerbungsgespräch) angesiedelt, während alltagssprachliche Units alltägliche Situationen thematisierten (z. B. Einkauf). Zudem wiesen die bildungssprachlichen Units weniger kontextuelle Bezüge auf als alltagssprachliche Units, d. h. die Texte enthielten weniger Informationen über die Protagonisten oder das Setting der thematisierten Handlungen. Auf syntaktischer Ebene enthielten bildungssprachliche Units mehr Nominalphrasen, Relativsätze und Passivkonstruktionen als alltagssprachliche Units. Lexikalisch waren bildungssprachliche Texte durch den gegenüber alltagssprachlichen Texten vermehrten Gebrauch von Komposita, Nominalisierungen

gen, mehrsilbigen Wörtern mit mehr als drei Silben sowie einen domänenspezifischen und -übergreifenden bildungssprachlichen Wortschatz gekennzeichnet.

Die Überprüfung der Forschungsfragen von Teilstudie 3 erfolgte mittels konfirmatorischer Faktorenanalysen, in die der sozioökonomische Status der Schülerinnen und Schüler als Kovariate einging. Die Operationalisierung des sozioökonomischen Hintergrunds erfolgte analog zu Teilstudie 1 und 2 anhand des Index des ökonomischen, sozialen und kulturellen Status (*ESCS*; OECD; 2006).

7.2 Analysestrategie

Die Prüfung der Fragstellungen in Teilstudie 1 und 2 erfolgte anhand von latenten Mehrgruppenmodellen, die unter Verwendung des Programms Mplus Version 7.11 (Muthén, & Muthén, 1998-2012) geschätzt wurden. Um sicherzustellen, dass die multivariaten Verteilungsparameter der latenten Konstrukte in den drei betrachteten Sprachgruppen invariant sind, wurden die Strukturgleichungsmodelle zunächst auf das Vorliegen von Messinvarianz überprüft. Hierzu werden auf Basis eines unrestringierten Modells sukzessive Parameter für die Gruppen auf Gleichheit restringiert (vgl. Steenkamp & Baumgartner, 1998). Im ersten Schritt wird dabei lediglich die Spezifikation der Faktormodelle auf Invarianz restringiert (*konfigurale Invarianz*), anschließend werden auch die Faktorladungen (*metrische Invarianz*) und in einem letzten Schritt zusätzlich die Intercepts für metrische Variablen bzw. Schwellenwerte beim Vorliegen von Ordinaldaten (*skalare Invarianz*) fixiert. Mittels eines χ^2 -Differenzentests wird die Anpassungsgüte eines restriktiveren mit dem jeweils weniger restriktiven Modell verglichen. Zeigt das weniger restriktive Modell keine bessere Anpassung auf die Daten als das restriktivere Modell, wird letzteres beibehalten. Wesentlich für die folgenden Analysen ist das Vorliegen von skalarer Invarianz, da unter dieser Bedingung sowohl die Ausprägungen der latenten Variablen als auch deren Beziehungen untereinander verzerrungsfrei zwischen den Sprachgruppen verglichen werden können (vgl. Baumgartner & Steenkamp, 1998). Die zusätzliche Annahme von Messfehlerinvarianz, bei der alle Elemente der Varianz-Kovarianz-Matrix des Messfehlers zwischen den Gruppen gleich gesetzt sind, wurde im Rahmen der Analysen nicht überprüft, da davon ausgegangen werden kann, dass Fehlervarianzen und -kovarianzen stichprobenspezifisch sind und ein Test auf Gleichheit der Messfehler somit als zu streng angesehen werden kann (vgl. Byrne, Shavelson & Muthén, 1989).

Die Anpassungsgüte der messinvarianten Modelle in Teilstudie 1 und 2 wurde zunächst anhand des χ^2 -Anpassungsmaßes beurteilt, das einen Wahrscheinlichkeitswert für die Hypothese liefert, dass die empirische Kovarianzmatrix der modelltheoretischen Kovarianzmatrix

entspricht (Backhaus, Erichson, Plinke & Weiber, 2006). Allerdings ist dieses Maß sensitiv dem Stichprobenumfang gegenüber, d. h. bei ausreichend großem Stichprobenumfang wird nahezu jedes Modell abgelehnt (vgl. Jöreskog & Sörbom, 1982). Aufgrund der Abhängigkeit des Signifikanzwertes von der Stichprobengröße wird daher in der Praxis oftmals lediglich das χ^2/df -Verhältnis beurteilt, das unter 2,5 liegen sollte (vgl. Backhaus, Erichson, Plinke & Weiber, 2006). Zusätzlich zu dem χ^2 -Anpassungsmaß wurde in Bezug auf die Anpassungsgüte der geschätzten Modelle der *Root Mean Square Error of Approximation* (RMSEA) betrachtet. Der RMSEA gibt die Diskrepanz zwischen Populationskovarianzmatrix und der am besten angepassten Matrix zur Populationskovarianzmatrix an und kann als Maß dafür betrachtet werden, wie schlecht ein gegebenes Modell auf die vorliegenden Daten passt. RMSEA-Werte kleiner als .06 weisen auf eine gute Modellanpassung hin (Hu und Bentler, 1999). Alle Modelle wurden weiterhin anhand des *Standardized Root Mean Square Residual* (SRMR) beurteilt, der die standardisierte Differenz zwischen beobachteten und vorher-gesagten Korrelationen im Modell abbildet. SRMR-Werte unter .08 können als Hinweis auf eine gute Modellanpassung gewertet werden (Hu & Bentler, 1999). Zusätzlich zu diesen absoluten Fitindizes, die angeben, wie gut bzw. schlecht das a-priori Modell die Daten repliziert, wurden mit dem *Tucker Lewis Index* (TLI) und dem *Comparative Fit Index* (CFI) zwei inkrementelle Fitindizes betrachtet. Für diese Indizes wird der Fit des aktuellen Modells mit dem eines zuvor gewählten und geschätzten Vergleichsmodells, dem Nullmodell, verglichen, das gegenüber dem aktuellen Modell restriktiver ist. Je besser die Anpassung des geschätzten Modells gegenüber dem Nullmodell ausfällt, desto höhere Werte resultieren für CFI und TLI (Hu & Bentler, 1999). Für das Nullmodell des CFI wird angenommen, dass alle latenten Variablen unkorreliert sind, anschließend wird dieses Modell mit der Kovarianzmatrix des relevanten Modells verglichen. Für die Berechnung des TLI werden die χ^2 -Werte des relevanten Modells zu den χ^2 -Werten des Basismodells in Bezug gesetzt. Im Gegensatz zu den absoluten Fitindizes sind die inkrementellen Fitmaße unabhängig von der Stichprobengröße. Werte unter 0.90 für den TLI bzw. CFI weisen auf eine unbefriedigende und Werte ≥ 0.90 auf eine gute Modellanpassung hin (Hu & Bentler, 1999).

Sowohl in Teilstudie 1 als auch in Teilstudie 2 erfolgte die inferenzstatistische Überprüfung von Gruppenunterschieden auf latenter Ebene anhand von Modellvergleichen mittels Wald- χ^2 -Tests, bei denen die interessierenden Parameter entweder über die Gruppen frei geschätzt oder gleich gesetzt wurden. Unterschiede in den Pfadkoeffizienten der Mehrgruppenmodelle wurden ebenfalls anhand von Wald- χ^2 -Tests auf ihre Signifikanz überprüft.

In Teilstudie 3 wurden zunächst für die betrachteten Variablen *differential item functioning* (DIF)-Analysen für den Sprachhintergrund der Schülerinnen und Schüler durchgeführt. DIF tritt auf, wenn ein Item in verschiedenen Subpopulationen unter Kontrolle der Fähigkeitsausprägungen unterschiedliche Lösungswahrscheinlichkeiten aufweist (Clauser & Mazor, 1998). Liegt DIF vor, variiert die Lösungswahrscheinlichkeit für ein bestimmtes Item demnach auch bei gleicher Personenfähigkeit zwischen den Gruppen, d. h. das Item kann nicht als messäquivalent betrachtet werden. Eine wesentliche Grundannahme bei der Analyse von DIF ist die Multidimensionalität des gemessenen Konstrukts (Shealy & Stout, 1993). Dabei stellt das eigentliche durch den Test gemessene Konstrukt die primäre Dimension dar; es existiert darüber hinaus aber noch mindestens eine weitere, sekundäre Dimension, die die Itemantwort innerhalb des Tests bestimmt. Vernachlässigbarer (*benign*) DIF tritt auf, wenn DIF auf eine sekundäre Dimension zurückführbar ist, die als dem Konstrukt inhärent gelten kann. Demgegenüber bezeichnet sogenannter *adverser* DIF das Vorliegen von DIF aufgrund von Störfaktoren, die nicht als relevant für das eigentlich zu messende Konstrukt angesehen werden können. Items, die adversen DIF aufweisen, sollten daher von weiteren Analysen ausgeschlossen werden (Douglas, Roussos & Stout, 1996). In Teilstudie 3 erfolgte die Überprüfung auf das Vorliegen von DIF auf Basis eines Vergleichs der Differenzen der Item-Logits von L1- und L2-Schülerinnen und -Schülern in einem zweidimensionalen Raschmodell mittels des Computerprogramms ConQuest 2.0 (Wu, Adams, Wilson, & Haldane, 2007). Items, die in dieser Betrachtung signifikante Logitdifferenzen von mehr als 0.64 Logits aufwiesen, wurden als substantieller adverser DIF betrachtet und aus den weiteren Analysen ausgeschlossen (vgl. Douglas et al., 1996).

Die Überprüfung der Dimensionalität des Hörverstehens in einer Erst- und Zweitsprache erfolgte in Teilstudie 3 über eine konfirmatorische Faktorenanalyse unter Verwendung von Mplus 7.11 (Muthén & Muthén, 1998-2012). Da in die Analyse kategoriale Daten eingingen, wurde als Schätzer der *weighted least square estimator* (WLSMV) verwendet. Der Fit des ein- bzw. zweidimensionalen Modells des Hörverstehens wurde analog zu Teilstudie 1 und 2 anhand der Parameter RMSEA, CFI und TLI sowie eines χ^2 -Differenzentests verglichen. Zusätzlich wurden Korrelationen zu den Außenkriterien des Wortschatzes sowie der morpho-syntaktischen Kenntnisse berechnet, die als latente Variablen in die Analyse eingingen. Unterschiede in den Zusammenhängen der Außenkriterien zu alltags- bzw. bildungssprachlichem Hörverstehen wurden mittels Wald- χ^2 -Tests auf statistische Bedeutsamkeit geprüft.

Die Überprüfung auf Sprachgruppenunterschiede in der Ausprägung alltagssprachlicher bzw. bildungssprachlicher Hörverstehenskompetenzen erfolgt analog zu dem für Teilstudie 1 und 2 beschriebenen Vorgehen über eine Prüfung der Messinvarianz des geschätzten Modells sowie eine anschließende Überprüfung auf Differenzen in den Intercepts zwischen L1- und L2-Jugendlichen mittels Wald- χ^2 -Tests.

7.3 Umgang mit fehlenden Werten

Aufgrund von Nichtteilnahme an der Fragebogenerhebung oder der Verweigerung einzelner Antworten bzw. Flüchtigkeitsfehlern innerhalb des Schülerfragebogens ergaben sich für die Gruppierungsvariable *Herkunftssprache* in der DELKO-Studie 34.4% fehlende Werte. Mittels des in der Statistiksoftware SPSS 18.0 implementierten *Missing Values Analyses*-Moduls wurde zunächst die Annahme überprüft, dass die fehlenden Daten *missing at random* (MAR) waren, d. h. zufällig und unsystematisch fehlten und somit als ignorierbare Nichtantwort betrachtet werden können (vgl. Allison, 2001). Im Rahmen des *Missing Values Analyses Tool* wird hierfür der Little'sche MCAR-Test genutzt, der die Annahme testet, dass das Auftreten fehlender Werte auf einer Variablen nicht von Ausprägungen anderer beobachteter Variablen innerhalb des Datensatzes abhängt (Little, 1988). Die Klassifikation des Ausfallprozesses als MAR erfolgt demnach relativ zu dem betrachteten Datensatz und ist somit abhängig von den erhobenen Daten. Lüdtke und Kollegen (2007) empfehlen vor diesem Hintergrund die Aufnahme zusätzlicher Variablen in den Datensatz, die mit den Ausfallprozessen auf den eigentlich interessierenden Variablen zusammenhängen könnten. Für die Analyse der Gruppierungsvariable *Herkunftssprache* wurden daher zusätzliche Variablen als Hintergrundvariablen in den Datensatz aufgenommen, die mittlere bis hohe Korrelationen zur Herkunftssprache aufwiesen¹. Die anschließende Analyse mittels des *Missing Values Analyses*-Tools zeigte, dass die Bedingung MAR für die fehlenden Daten zur Herkunftssprache angenommen werden konnte. Für die Gruppierungsvariable wurde daher eine multiple Imputation mittels des Programms SPSS 18.0 durchgeführt, bei der für jeden fehlenden Wert unter Einbezug der im Datensatz vorliegenden Informationen mehrere Ersetzungen vorgenommen und so mehrere vollständige Datensätze erstellt werden (vgl. Lüdtke et al. 2007 für eine Beschreibung der Methode). Die Parameterschätzung erfolgte über das *Markov-Chain-Monte-Carlo* (MCMC)-Verfahren, generiert wurden $m = 5$ Datensätze.

¹ In die Analyse eingehende Variablen waren Angaben der Lehrkräfte zum Migrationsstatus sowie zur Herkunftssprache der Schülerinnen und Schüler, Schülerangaben zu den Geburtsländern von Vater und Mutter, den Herkunftsländern der Großeltern, der Herkunftssprache beider Elternteile, der zuhause gesprochenen Sprache sowie dem Sprachgebrauch im sozialen Umfeld und Elterngaben bezüglich des Gebrauchs der Zweitsprache bei kommunikativen Aktivitäten.

Im Anschluss an die multiple Imputation wurde unter Bezug auf Überlegungen von Little und Rubin (2002) eine Simulation durchgeführt, bei der aus einem in Bezug auf die Angabe der Herkunftssprache vollständigen Datensatz zunächst ein Datensatz mit fehlenden Daten generiert und anschließend anhand der beschriebenen multiplen Imputation aufgefüllt wurde. Die geschätzten Daten für die Herkunftssprache der Schülerinnen und Schüler wurden dann mit den tatsächlichen Schülerangaben abgeglichen. Für die fünf Imputationsdurchgänge fanden sich dabei Prozentsätze korrekter Klassifikationen zwischen lediglich 53% und 60%. Auch die Verteilung der Schülerinnen und Schüler auf die drei Sprachgruppen entsprach nicht der des Ausgangsdatensatzes (vgl. Tabelle 2).

Tabelle 2.

Simulation zur Ersetzung der Gruppierungsvariable Herkunftssprache: Verteilung auf die drei Sprachgruppen

	kompletter Datensatz	Simulation: multiple Imputation ^a	Simulation: Ersetzen durch Lehrerangabe
L1 _{Deutsch}	42%	36%	46%
L2 _{Türk}	39%	56%	39%
L2 _{Andere}	18%	7%	15%

^agepoolte Werte für die fünf Imputationsdurchgänge

Aufgrund der unbefriedigenden Ergebnisse wurde eine erneute Simulation vorgenommen, bei der fehlende Werte auf der Gruppierungsvariable *Herkunftssprache* durch die Angaben der Lehrkräfte zur Herkunftssprache der Schülerinnen und Schüler ersetzt wurden, die vollständig vorlagen. Es zeigte sich, dass durch dieses Verfahren 86% der fehlenden Daten korrekt repliziert werden konnten. Zudem zeigte sich eine gute Anpassung an die Originaldaten im Hinblick auf die Verteilung auf die drei Sprachgruppen (vgl. Tabelle 2). Für die folgenden Auswertungen wurden fehlende Werte auf der Gruppierungsvariablen *Herkunftssprache* daher durch die entsprechenden Angaben der Lehrkräfte ergänzt.

Fehlende Werte auf den übrigen in die Analysen eingehenden Skalen wurden für Studie 1 und 2 mittels SPSS 18.0 zunächst auf das Vorliegen von *missingness at random* überprüft. Da alle fehlenden Werte diese Bedingung erfüllten, wurden Missings anschließend unter Einbezug aller betrachteten Variablen unter Verwendung des Programm SPSS 18.0 multipel imputiert (vgl. Lüdtke et al., 2007). Ähnlich wie bei dem Vorgehen zur Überprüfung des

Imputationsmodells für die Gruppierungsvariable wurde auch für die Imputierung der übrigen Variablen eine Simulation durchgeführt. Die Ergebnisse zeigten, dass sich die imputierten Datensätze in Bezug auf Minimal- und Maximalwerte sowie die arithmetischen Mittel und Standardabweichungen nur minimal vom Originaldatensatz unterschieden. Entsprechend gingen die fünf imputierten Datensätze in die latenten Modellierungen in Teilstudie 1 und 2 ein.

Für Teilstudie 3 gingen mit den dichotom kodierten Items kategoriale Daten in die Analysen ein. Die Parameterschätzung erfolgte daher mittels des *weighted least square estimator (WLSMV)*, der einen mittelwert- und varianzjustierten robusten Schätzer für kategoriale Daten bildet. Für diesen Schätzer werden Vergleiche genesteter Modelle im für die Analysen genutzten Programm Mplus 7.11 (Muthén & Muthén, 1998-2012) mittels der Option DIFFTEST durchgeführt, die nicht für imputierte Datensätze implementiert ist. Berücksichtigt werden bei WLSMV-Schätzung alle vorhandenen Daten im Rahmen einer *pairwise present* Analyse (vgl. Asparouhov & Muthén, 2010). Das entsprechende Vorgehen kann auch bei Vorliegen fehlender Werte als robust gelten, sofern der Anteil an fehlenden Werten nicht substantiell ist (Asparouhov & Muthén, 2010).

7.4 Ergebnisse

7.4.1 Teilstudie 1: Prädiktoren des Hörverstehens bei Jugendlichen deutscher und Jugendlichen nichtdeutscher Herkunftssprache (Marx & Roick, 2012)

Zur Prüfung auf Unterschiede in der Ausprägung sprachlicher Prädiktoren des Hörverstehens in einer Erst- bzw. Zweitsprache wurde zunächst die Messinvarianz des geschätzten Modells überprüft. Die Ergebnisse der Überprüfung sowie die Fitwerte der Modelle werden in Anhang A dargestellt. Zur Überprüfung von Forschungsfrage 1.1 nach Unterschieden in der Ausprägung sprachlicher Prädiktoren des Hörverstehens in einer Erst- bzw. Zweitsprache wurden anschließend die latenten Mittelwerte im skalaren Modell verglichen. Tabelle 3 liefert eine Übersicht der standardisierten Mittelwerte für die betrachteten latenten Variablen unter Kontrolle des sozioökonomischen Hintergrunds der Schülerinnen und Schüler. Die Gruppe der deutschsprachigen Jugendlichen diente dabei als Referenzgruppe, für die die Mittelwerte der latenten Variablen auf Null fixiert wurden. Für den Vergleich der beiden L2-Sprachgruppen untereinander wurde die Gruppe der L2_{Türk}-Jugendlichen als Referenzgruppe gewählt. Die Differenzen der standardisierten Mittelwerte zu den übrigen Sprachgruppen können im Rahmen dieser Modellierung analog zur Effektstärke d interpretiert werden. Dabei kennzeichnen Werte von $d \geq .20$ kleine, $d \geq .50$ mittlere und $d \geq .80$ starke Effekte (Cohen, 1988).

Tabelle 3.

Standardisierte latente Mittelwerte der Sprachgruppen im messinvarianten Mehrgruppenmodell für das Hörverstehen in Teilstudie 1 (unter Kontrolle des ESCS)

	L1 _{Deutsch}	L2 _{Türk}	L2 _{Andere}	Gruppenvergleiche der latenten Mittelwerte		
				L1 _{Deutsch} – L2 _{Türk}	L1 _{Deutsch} – L2 _{Andere}	L2 _{Türk} – L2 _{Andere}
α (HV)	0.00	-0.43	-0.05	9.51*	0.15	0.22
α (WS)	0.00	-0.56	-0.40	16.63*	6.89*	2.56
α (MS)	0.00	-0.24	-0.20	2.73	1.75	0.79
α (PB)	0.00	-0.16	0.07	1.05	0.18	0.75
α (AG)	0.00	0.00	0.05	0.00	0.11	0.11

Anmerkungen. HV = Hörverstehen; PB = Phonologische Bewusstheit; WS = Wortschatz; MS = Morphosyntaktische Kenntnisse; AG = Arbeitsgedächtnis. * $p < .05$

Erwartungskonform zeigen sich für die phonologischen Bewusstheit sowie das Arbeitsgedächtnis keine Gruppenunterschiede zwischen den drei betrachteten Sprachgruppen. Für die Variablen Hörverstehen, Wortschatz und morpho-syntaktische Bewusstheit waren auf Basis der *time on task*-Hypothese Kompetenznachteile der L2-Schülerinnen und -Schüler gegenüber ihren L1-Peers erwartet worden, die auch unter Kontrolle des sozioökonomischen Status persistieren sollten. Die Ergebnisse des Mehrgruppenvergleichs zeigen erwartungskonform, dass sich im Wortschatz bei L2-Jugendlichen mit türkischer Herkunftssprache signifikante Kompetenzunterschiede zu den L1-Jugendlichen ergeben, die mit $d = -.56$ im mittleren Effektstärkenbereich liegen. Auch die L2_{Anderer}-Jugendlichen weisen unter Kontrolle des ESCS signifikant geringe Wortschatzkenntnisse auf als ihre L1-Peers, wobei der Effekt mit $d = -.40$ lediglich als gering bis mittel zu klassifizieren ist. Für das Hörverstehen zeigten sich Leistungs-nachteile gegenüber L1-Schülerinnen und -Schülern nur für jene L2-Jugendlichen, die Türkisch als Herkunftssprache angaben. Auch hier liegt die Effektstärke mit $d = -.43$ im geringen Bereich. Für L2_{Anderer}-Jugendliche fanden sich entgegen der Erwartungen unter Kontrolle des sozioökonomischen Status keine Kompetenznachteile im Hörverstehen gegenüber L1-Schülerinnen und -Schülern. Für die morpho-syntaktischen Kenntnisse zeigen sich entgegen der Hypothesen keinerlei Gruppenunterschiede zwischen den drei betrachteten Sprachgruppen. Unter Kontrolle des sozioökonomischen Status weisen L2-Jugendliche somit morpho-syntaktische Kenntnisse auf, die denen ihrer L1-Peers vergleichbar sind.

Die in Forschungsfrage 1.2 genannte Überprüfung auf Sprachgruppenunterschiede in den Zusammenhängen zwischen Prädiktoren des Hörverstehens und dem Hörverstehen wurde anhand von Wald- χ^2 -Tests auf Signifikanz vorgenommen. Die praktische Relevanz der Unterschiede in den korrelativen Beziehungen wurde mittels des von Cohen (1988) vorgeschlagenen Effektstärkenmaßes q beurteilt. q -Werte zwischen .10 und .30 können als kleine, q -Werte zwischen .30 und .50 als mittlere und q -Werte größer als .50 als starke Effekte interpretiert werden. Tabelle 4 zeigt eine Übersicht der standardisierten Pfad-koeffizienten. Die Ergebnisse zeigen, dass in allen Gruppen erwartungskonform ein signifikanter Zusammenhang zwischen dem Hörverstehen und dem Wortschatz bzw. der morpho-syntaktischen Bewusstheit besteht. Die Arbeitsgedächtnisleistung hängt hingegen entgegen der Annahmen in keiner der Sprachgruppen mit dem Hörverstehen zusammen. Die phonologische Bewusstheit hat nur in der Gruppe der L2_{Türk}-Jugendlichen einen Effekt auf das Hörverstehen.

Entgegen der Erwartungen finden sich keine Gruppenunterschiede in der Höhe der Pfadkoeffizienten. Bezüglich des relativen Einflusses der Prädiktoren des Hörverstehens wurden zusätzlich Analysen durchgeführt, um zu erfassen, ob einzelne Prädiktoren innerhalb der

Sprachgruppen differenziell bedeutsam sind. Berücksichtigt wurden dabei lediglich jene Prädiktoren des Hörverstehens, für die ein signifikanter Zusammenhang zur Hörverstehensleistung bestätigt worden war, d. h. der Wortschatz, die morpho-syntaktischen Kenntnisse sowie die phonologische Bewusstheit für die L2_{Türk}-Jugendlichen. Die Ergebnisse dieser Berechnungen erbrachten keine Hinweise auf eine unterschiedliche Relevanz der einzelnen Prädiktoren für das Hörverstehen (Wald- χ^2 -Werte zwischen $\chi^2 = 0.00$ und $\chi^2 = 0.16$). Innerhalb der Sprachgruppen scheinen die Prädiktoren demnach eine ähnliche Bedeutsamkeit zur Erklärung der Hörverstehensleistung zu besitzen.

Tabelle 4.

Standardisierte Pfadkoeffizienten der Sprachgruppen im messinvarianten Mehrgruppenmodell für das Hörverstehen in Teilstudie 1 (unter Kontrolle des ESCS)

	L1 _{Deutsch}	L2 _{Türk}	L2 _{Andere}	Wald- χ^2 -Werte für die Gruppenvergleiche					
				L1 _{Deutsch} – L2 _{Türk}		L1 _{Deutsch} – L2 _{Andere}		L2 _{Türk} – L2 _{Andere}	
				χ^2	<i>q</i>	χ^2	<i>q</i>	χ^2	<i>q</i>
β (WS \rightarrow HV)	.26*	.29*	.34*	0.05	.03	0.07	.09	0.33	.06
β (MS \rightarrow HV)	.35*	.27*	.53*	0.05	.09	0.81	.23	1.43	.31
β (PB \rightarrow HV)	.19	.31*	.03	1.06	.13	0.70	.16	2.84	.29
β (AG \rightarrow HV)	-.03	.01	-.09	0.09	.04	0.22	.06	0.53	.01

Anmerkungen. HV = Hörverstehen; PB = Phonologische Bewusstheit; WS = Wortschatz; MS = Morpho-syntaktische Kenntnisse; AG =Arbeitsgedächtnis. * $p < .05$

7.4.2 Teilstudie 2: Prädiktoren der Lesekompetenz in einer Erst- bzw. Zweitsprache (Marx, Stanat, Roick, Segerer, Marx & Schneider, 2015)

Für die Analyse der strukturellen Beziehungen des Hörverstehens zur Leseleistung wurde in Teilstudie 2 zunächst eine Messinvarianzanalyse unter Berücksichtigung des ESCS durchgeführt, um die Vergleichbarkeit der Messmodelle innerhalb der Gruppen zu verifizieren. Die Ergebnisse sowie die Angaben zum Fit der Modelle werden in Anhang B dargestellt. Zur Prüfung von Forschungsfrage 2.1 zu Gruppenunterschieden in den latenten Mittelwerten wurden Wald- χ^2 -Tests durchgeführt, bei denen die L1-Gruppe als Referenzgruppe diente. Für die

Prüfung auf Unterschiede zwischen den beiden Gruppen von L2-Schülerinnen und -Schülern wurde die Gruppe der L2_{Türk}-Jugendlichen als Referenz gewählt. Ohne Kontrolle des sozioökonomischen Hintergrunds zeigte sich, dass beide Gruppen von L2-Schülerinnen und -Schülern auf der Mehrheit der betrachteten Variablen schwächer als ihre L1-Peers abschneiden (vgl. Tabelle 5). Erwartungskonform gilt dies sowohl für das Textverstehen in den Modalitäten Lesen und Hören, aber auch für den Wortschatz und die morpho-syntaktischen Kenntnisse. Ebenfalls im Einklang mit den Annahmen fanden sich keine Unterschiede zwischen den Sprachgruppen in Bezug auf die Arbeitsgedächtniskapazität sowie die phonologische Bewusstheit. Für die Lesegeschwindigkeit waren keine Gruppenunterschiede erwartet worden, da diese eng mit der phonologischen Bewusstheit verbunden ist und somit als Aspekt metalinguistischer Bewusstheit gelten kann, der durch den Kontakt zu mehreren Sprachen eher gefördert werden sollte. Konträr zu dieser Annahme war die Lesegeschwindigkeit bei L2_{Türk}- sowie L2_{Andere}-Jugendlichen im Vergleich zu ihren monolingualen Peers jedoch geringer ausgeprägt.

Um die Effekte des Sprachhintergrunds der Schülerinnen und Schüler von jenen trennen zu können, die auf dem unterschiedlichen sozioökonomischen Status der Jugendlichen beruhen, wurden in einem zweiten Auswertungsschritt latente Mittelwertsdifferenzen unter Kontrolle des ESCS berechnet. Erwartet worden war, dass L2-Schülerinnen und -Schüler schwächere Leistungen im Hör- sowie im Leseverstehen aufweisen als ihre L1-Peers, wobei sich für L2_{Türk}-Jugendliche besonders stark ausgeprägte Leistungsdifferenzen gegenüber L1-Jugendlichen zeigen sollten. Erwartet wurde zudem, dass sich Disparitäten zwischen L1- und L2-Jugendlichen zu einem großen Teil durch den sozioökonomischen Status der Familien erklären lassen. Die Ergebnisse dieser Analysen liefern empirische Stützung für diese Annahme und zeigen, dass für die Gruppe der L2_{Andere}-Jugendlichen unter Kontrolle des ESCS nur die Mittelwertsdifferenzen im Wortschatz gegenüber L1-Jugendlichen signifikant werden. Der entsprechende Effekt ist mit $d = .28$ als klein zu klassifizieren. Die Differenzen in der Lesegeschwindigkeit, dem Hör- und Leseverstehen sowie den morpho-syntaktischen Kenntnissen zwischen L2_{Andere}-Jugendlichen und den L1-Jugendlichen ließen sich unter Berücksichtigung des sozioökonomischen Status nicht mehr finden.

Für die Gruppe der L2_{Türk}-Jugendlichen fanden sich hingegen auch unter Kontrolle des sozialen Hintergrunds schwächere Leistungen in Bezug auf Hör- und Leseverstehen, morpho-syntaktische Kenntnisse sowie den Wortschatz. Während der Effekt für die Unterschiede in den morpho-syntaktischen Kenntnissen dabei mit $d = -.20$ als gering bezeichnet werden kann,

zeigten sich für die übrigen Variablen mittlere Effekte. Leistungsdifferenzen zwischen den beiden Gruppen von L2-Jugendlichen konnten empirisch nicht gezeigt werden.

Tabelle 5.

Standardisierte latente Mittelwerte der Sprachgruppen im messinvarianten Mehrgruppenmodell für Teilstudie 2

	Gruppenvergleiche der latenten Mittelwerte					
	L1 _{Deutsch}	L2 _{Türk}	L2 _{Andere}	L1 _{Deutsch} – L2 _{Türk}	L1 _{Deutsch} L2 _{Andere}	L2 _{Türk} – L2 _{Andere}
<i>Ohne Kontrolle</i>						
α (LV)	0.00	-0.51	-0.29	18.98*	6.08*	0.22
α (HV)	0.00	-0.65	-0.33	34.83*	8.02*	0.16
α (LG)	0.00	-0.56	-0.31	11.60*	3.76*	0.03
α (WS)	0.00	-0.66	-0.39	39.03*	11.04*	1.99
α (MS)	0.00	-0.31	-0.29	5.31*	5.59*	1.29
α (PB)	0.00	-0.12	0.06	3.03	0.01	0.91
α (AG)	0.00	0.04	-0.06	0.56	0.18	0.63
<i>Unter Kontrolle des ESCS</i>						
α (LV)	0.00	-0.60	-0.27	18.23*	3.52	0.06
α (HV)	0.00	-0.69	-0.17	30.71*	0.88	0.73
α (LG)	0.00	-0.52	-0.23	5.22*	0.67	0.01
α (WS)	0.00	-0.52	-0.28	15.16*	3.87*	1.01
α (MS)	0.00	-0.20	-0.07	5.36*	0.22	0.27
α (PB)	0.00	-0.17	-0.17	0.73	0.65	0.08
α (AG)	0.00	-0.05	0.02	0.01	0.01	0.01

Anmerkungen. LV = Leseverstehen; HV = Hörverstehen; LG = Lesegeschwindigkeit; PB = Phonologische Bewusstheit; WS = Wortschatz; MS = Morpho-syntaktische Kenntnisse; AG = Arbeitsgedächtnis. * $p < .05$

Die im Rahmen von Forschungsfrage 2.2 vorgenommenen Prüfungen auf Gruppenunterschiede in den strukturellen Beziehungen zwischen Hör- und Leseverstehen zwischen L1- und L2-Jugendlichen erfolgten über Wald- χ^2 -Tests. Dabei wurden auch strukturelle Beziehungen zwischen einzelnen Prädiktoren des Hörverstehens und der Hörverstehensleistung modelliert (vgl. Anhang B). Da diese bereits in Teilstudie 1 berichtet wurden, wird im Folgenden nur auf die Effekte der Lesegeschwindigkeit und des Hörverstehens auf die Lesekom-

petenz eingegangen. Tabelle 6 zeigt die Ergebnisse dieser Analysen für die Beziehungen zwischen dem Hör- und dem Leseverstehen sowie der Lesegeschwindigkeit und dem Leseverstehen. Zur Beurteilung der praktischen Bedeutsamkeit von Gruppenunterschieden wurde das Effektmaß q angegeben (vgl. Cohen, 1988).

Tabelle 6.

Standardisierte Pfadkoeffizienten der Sprachgruppen im messinvarianten Mehrgruppenmodell für das Leseverstehen in Teilstudie 2

	L1 _{Deutsch}	L2 _{Türk}	L2 _{Andere}	Wald- χ^2 -Werte für die Gruppenvergleiche der Pfadkoeffizienten					
				L1 _{Deutsch} – L2 _{Türk}		L1 _{Deutsch} – L2 _{Andere}		L2 _{Türk} – L2 _{Andere}	
				χ^2	q	χ^2	q	χ^2	q
<i>Ohne Kontrolle des ESCS</i>									
β (LG \rightarrow LV)	.33*	.19*	.35*	1.12	.15	0.90	.02	3.11*	.17
β (HV \rightarrow LV)	.57*	.61*	.52*	0.80	.06	1.86	.07	0.45	.13
R^2 (LV)	.47	.44	.52						
<i>Unter Kontrolle des ESCS</i>									
β (LG \rightarrow LV)	.32*	.20*	.37*	0.80	.13	1.44	.06	3.58*	.19
β (HV \rightarrow LV)	.56*	.61*	.52*	0.70	.08	1.46	.06	0.31	.13
R^2 (LV)	.48	.44	.53						

Anmerkungen. LV = Leseverstehen; LG = Lesegeschwindigkeit; HV = Hörverstehen.

* $p < .05$

Erwartungskonform ergaben sich in der Höhe der Pfadbeziehungen kaum Unterschiede zwischen den Sprachgruppen. Hörverstehen und das Leseverstehen waren in allen drei Schülergruppen in ähnlicher Weise miteinander assoziiert. Zudem fand sich in den Wald- χ^2 -Tests für alle Gruppen ein im Vergleich zum Effekt der Lesegeschwindigkeit stärkerer Effekt des Hörverstehens auf die Lesekompetenz ($\chi^2_{(1)} = 7.37$; $p < .01$ für L1_{Deutsch}; $\chi^2_{(1)} = 13.32$; $p < .01$ für L2_{Türk}; $\chi^2_{(1)} = 10.93$; $p < .01$ für L2_{Andere}). Auch dieses Ergebnis ist erwartungsgemäß und entspricht den Annahmen des *Simple View of Reading* für ältere Schülerinnen und Schüler.

Ein unerwarteter Befund zeigte sich lediglich für die Effekte der Lesegeschwindigkeit auf die Lesekompetenz bei den L2-Schülerinnen und -Schülern: Sowohl ohne Kontrolle des sozioökonomischen Status als auch unter Berücksichtigung des sozialen Hintergrunds fand sich in der Gruppe der L2_{Andere}-Jugendlichen ein engerer Zusammenhang zwischen der Lesegeschwindigkeit und dem Leseverstehen als bei den L2_{Türk}-Jugendlichen. Der zugehörige Effekt ist als klein einzustufen.

7.4.3 Teilstudie 3: Bildungssprachliches und alltagssprachliches Hörverstehen in einer Erst- bzw. Zweitsprache (Marx, Heppt & Henschel; in Revision)

Für die Untersuchung der Dimensionalität des Hörverstehens in einer Erst- bzw. Zweitsprache wurden in Teilstudie 3 zunächst alle in die Analysen eingehenden Items auf das Vorliegen von adversen Item-DIF (vgl. Kapitel 7.2) in Abhängigkeit vom Sprachhintergrund überprüft. Die Ergebnisse zeigten, dass drei Items des Hörverstehentests, sieben Items des Wortschatztests und zwei Items des Tests zur Überprüfung der morpho-syntaktischen Kenntnisse problematische DIF-Werte aufwiesen. Aus den weiteren Analysen wurden diese Items entfernt.

Zur Überprüfung der Dimensionalität des Hörverstehens im Rahmen von Forschungsfrage 3.1 wurden im Rahmen einer konfirmatorischen Faktorenanalyse ein eindimensionales und ein zweidimensionales Modell des Hörverstehens verglichen, in dem jeweils für den sozioökonomischen Hintergrund der Schülerinnen und Schüler kontrolliert wurde. Die in Tabelle 7 dargestellten Ergebnisse zeigen, dass beide Modelle eine gute Anpassung an die Daten aufweisen, das zweidimensionale Modell jedoch, verglichen mit dem eindimensionalen Modell, einen besseren Fit aufweist. Mit $r = .84$ zeigte sich für dieses Modell allerdings eine starke Interkorrelation der beiden latenten Faktoren *Alltagssprachliches Hörverstehen* und *Bildungssprachliches Hörverstehen*. Aus diesem Grund wurde ein post hoc-Test durchgeführt, bei dem zunächst ein Modell geschätzt wurde, in dem die latente Interkorrelation auf Eins fixiert und ein anschließender Modellvergleich zu einem Modell mit unrestringierter Interkorrelation vorgenommen wurde. Ziel war es zu überprüfen, ob die Interkorrelation zwischen beiden Hörverstehensfacetten substanziell geringer ist als Eins, was zusätzlich für die Annahme zweier getrennter Faktoren des Hörverstehens sprechen würde. Die Ergebnisse dieser Überprüfung stützten die Annahme zweier getrennter Faktoren ($\Delta\chi^2(1) = 11.92, p < .01$).

Tabelle 7.

Vergleich des ein- und des zweidimensionalen Modells des Hörverstehens für Teilstudie 3

Modell	χ^2	df	p	RMSEA	CFI	TLI	Verglichen mit	$\Delta \chi^2$	df	p
A: eindimensional	492.04	405	<.01	.02	.94	.94	--	--	--	--
B: zweidimensional	473.49	403	.01	.02	.95	.95	Modell A	18.15	2	<.01

Anmerkungen. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; $N = 459$.

Zur weiteren Absicherung der Zweidimensionalität des Hörverstehens wurden Interkorrelationen zu den Außenkriterien *Wortschatz* und *morpho-syntaktische Kenntnisse* berechnet. Angenommen wurde, dass beide Außenkriterien stärker mit bildungssprachlichem als mit alltagssprachlichem Hörverstehen zusammenhängen, da ersteres höhere Anforderungen an die Wortschatz- und Grammatikkenntnisse stellt als alltagssprachliches Hörverstehen. Hypothesenkonträr hingen die morpho-syntaktischen Kenntnisse der Schülerinnen und Schüler enger mit dem alltagssprachlichen Hörverstehen zusammen als mit dem bildungssprachlichen Hörverstehen ($\chi^2_{(1)} = 11.33, p = .01$). Für den Wortschatz zeigte sich mit $r = .47$ zwar ein engerer Zusammenhang zum bildungssprachlichen als zum alltagssprachlichen Hörverstehen ($r = .28$), dieser ließ sich jedoch nicht gegen den Zufall absichern ($\chi^2_{(1)} = 0.38, p = .54$).

Für die Überprüfung auf Unterschiede in der Ausprägung bildungssprachlicher bzw. alltagssprachlicher Hörverstehenskompetenzen in einer Erst- bzw. Zweitsprache (Forschungsfrage 3.2) zeigte sich auf manifester Ebene zunächst, dass beide Schülergruppen schwächere Kompetenzen in Hinblick auf das bildungssprachliche ($M_{z-Wert} = 0.60, s = 0.18$ für L1-Jugendliche; $M_{z-Wert} = 0.50, s = 0.21$ für L2-Jugendliche) als im Hinblick auf das alltagssprachliche Hörverstehen aufwiesen ($M_{z-Wert} = 0.73, s = 0.15$ für L1-Jugendliche; $M_{z-Wert} = 0.65, s = 0.18$ für L2-Jugendliche). Diese Differenz ist allerdings für keine der Gruppen signifikant ($t_{(153)} = -0.86, p = .93$ für L1-Jugendliche; $t_{(262)} = 0.69, p = .95$ für L2-Jugendliche).

Zur Überprüfung auf Unterschiede in der Ausprägung alltagssprachlicher und bildungssprachlicher Hörverstehenskompetenzen in einer Erst- bzw. Zweitsprache wurde anschließend das zweidimensionale Modell auf Messinvarianz zwischen den Sprachgruppen getestet. Da keine Annahmen zu spezifischen Ausprägungen bildungssprachlicher Kompetenzen in Ab-

hängigkeit konkreter Herkunftssprachen vorlagen, wurde dabei neben der L1-Gruppe nur eine globale L2-Gruppe betrachtet. Die Ergebnisse der Überprüfung auf Messinvarianz sowie die Fitwerte der Modelle finden sich in Anhang C.

Anschließend wurden unter Kontrolle des sozioökonomischen Hintergrunds im messinvarianten zweidimensionalen Modell die latenten Mittelwerte für das bildungs- bzw. alltags-sprachliche Hörverstehen über Wald- χ^2 -Tests geprüft. Die Gruppe der L1-Jugendlichen diente dabei als Referenzgruppe, für die die Mittelwerte der latenten Variablen auf Null fixiert wurden. Die Differenzen in den standardisierten Mittelwerten zu den übrigen Sprachgruppen können bei dieser Modellierung analog zur Effektstärke d interpretiert werden. Tabelle 8 zeigt, dass L2-Jugendliche erwartungskonform auf beiden Dimensionen des Hörverstehens schwächere Kompetenzen aufweisen als ihre L1-Peers. Allerdings finden sich Unterschiede in der Ausprägung der Effekte: Für das bildungssprachliche Hörverstehen liegt der Unterschied im Bereich einer kleinen, für das alltags-sprachliche Hörverstehen im Bereich einer mittleren Effektstärke. Entgegen der Erwartungen sind Differenzen im alltags-sprachlichen Hörverstehen zwischen L1- und L2-Jugendlichen somit ausgeprägter als Differenzen im bildungssprachlichen Hörverstehen ($\chi^2_{(1)} = 8.72, p < .01$).

Tabelle 8.

Standardisierte latente Mittelwerte der Sprachgruppen im messinvarianten Mehrgruppenmodell für Teilstudie 3 (unter Kontrolle des ESCS)

	L1	L2	Wald- χ^2 -Werte für die Gruppenvergleiche ($df=1$)
α (HV-Alltags-sprachlich)	0.00	-0.65	12.98*
α (HV-Bildungssprachlich)	0.00	-0.45	4.68*

Anmerkung: HV = Hörverstehen. * $p < .05$

8. Zusatzanalyse: Die Bedeutung alltags-sprachlicher und bildungssprachlicher Wortschatzkenntnisse für das Hörverstehen in einer Erst- bzw. Zweitsprache

Die Ergebnisse der faktorenanalytischen Modellierung des Hörverstehens in Teilstudie 3 stützen im Wesentlichen die Annahme, dass sich alltags- und bildungssprachliches Hörverstehen voneinander abgrenzen lassen und somit zwei unterschiedliche Facetten des Sprachverstehens bilden. Die zusätzliche Validierung über die Korrelationen zu den Außen-kriterien lieferte

jedoch keine empirische Stützung für die postulierte Zweidimensionalität des Hörverstehens. So fand sich entgegen der Vermutung kein höherer Zusammenhang zwischen dem bildungssprachlichen Hörverstehen und dem Wortschatz als zwischen dem alltagsprachlichen Hörverstehen und dem Wortschatz. Eine mögliche Erklärung für das Ausbleiben differenzieller Zusammenhänge zum bildungs- bzw. alltagsprachlichen Hörverstehen liegt in der Erfassung des Wortschatzes: Möglicherweise werden vorhandene Unterschiede in den Zusammenhängen zum alltags- bzw. bildungssprachlichen Hörverstehen dadurch nivelliert, dass der Wortschatztest in Teilstudie 3 als globales Wortschatzmaß in die Analysen einging. Da sich alltags- und bildungssprachlicher Wortschatz differenzieren lassen (vgl. z. B. Snow, 2010), ist denkbar, dass sich die erwarteten Zusammenhänge nur zwischen den jeweils alltags- bzw. bildungssprachlichen Facetten zeigen, dass also der Zusammenhang zwischen bildungssprachlichem Wortschatz und bildungssprachlichem Hörverstehen höher ausfällt als der zwischen alltagsprachlichem Wortschatz und bildungssprachlichem Hörverstehen. Diese Hypothese soll in einer Zusatzanalyse überprüft werden. Im Folgenden werden zunächst wesentliche Annahmen zum bildungssprachlichen Wortschatz und zu dessen Bedeutung für das Hörverstehen in einer Erst- bzw. Zweitsprache vorgestellt, bevor anschließend die Frage nach differenziellen Zusammenhängen zwischen alltagsprachlichem und bildungssprachlichem Wortschatz und dem Hörverstehen empirisch überprüft wird.

Ein wesentliches Merkmal der in akademischen Kontexten verwendeten Sprache ist der Gebrauch eines bildungssprachlichen Wortschatzes (vgl. auch Kapitel 4). Dabei handelt es sich um ein Vokabular, das oftmals lateinischen oder griechischen Ursprungs ist und durch einen hohen Grad an Abstraktheit, eine geringe Bildhaftigkeit, eine semantische Mehrdeutigkeit sowie eine relativ geringe Auftretenshäufigkeit gekennzeichnet ist (Corson, 1997). Neben den spezifischen Fachtermini eines Unterrichtsfachs beinhaltet der bildungssprachliche Wortschatz auch lexikalische Merkmale, die der Verdichtung und Präzisierung von Informationen dienen, wie etwa sprachliche Mittel zur zeitlichen oder räumlichen Einordnung eines Sachverhalts oder Konnektoren (Gogolin, 2009). In der Vergangenheit wurde eine Reihe von Typologien postuliert, die den in Bildungskontexten verwendeten Wortschatz weiter spezifizieren, etwa in technisches und nicht-technisches Vokabular oder literarisches Vokabular (vgl. Bauman & Graves, 2010 für einen Überblick über entsprechende Typologien). Gemeinsam ist diesen Ansätzen, dass unterschieden wird zwischen einem domänenübergreifenden generellen bildungssprachlichen Wortschatz sowie einem spezifischen bildungssprachlichen Wortschatz, der überwiegend in einzelnen fachlichen Domänen, wie etwa dem Mathematikunterricht, Anwendung findet (Bauman & Graves, 2010).

Sowohl domänenspezifische als auch domänenübergreifende bildungssprachliche Wortschatzkenntnisse sind ein wesentlicher Bestandteil bildungssprachlicher Kompetenzen und können als eine wichtige Voraussetzung für den schulischen Erfolg von Schülerinnen und Schülern angesehen werden (vgl. Nagy & Twonson, 2012). Für den domänenspezifischen bildungssprachlichen Wortschatz merken Alexander und Jetton (2000) sowie Hyland und Tse (2007) beispielsweise an, dass akademische Domänen, d. h. einzelne Schulfächer, nicht nur über spezifische Inhalte verfügen, sondern darüber hinaus auch über spezifische Diskursformen und damit verbunden über spezielle Lexika, die für den Austausch über fachliche Themen genutzt werden. Die Beherrschung des Vokabulars einer Domäne ist somit zentral für den Erwerb domänenspezifischen Wissens und die Teilhabe am fachlichen Diskurs. Schülerinnen und Schüler „who do not become fluent in the ‘language’ of academic domains are unlikely to achieve competence“ (Jetton & Alexander, 2004; S. 17). Dies gilt in ähnlicher Weise auch für den domänenübergreifenden bildungssprachlichen Wortschatz. Allerdings sind Schülerinnen und Schüler hier zusätzlich mit der Schwierigkeit konfrontiert, dass sich die Bedeutung domänenübergreifender bildungssprachlicher Wörter je nach Kontext ändern kann. So hat das Wort „darstellen“ im Mathematikunterricht eine andere Bedeutung (z. B. einen mathematischen Zusammenhang grafisch darstellen) als im Literaturunterricht (z. B. die Wirkung des Gebrauchs bestimmter Metaphern auf den Rezipienten darlegen). Aufgrund dieser Ambiguität stellt der Erwerb fachübergreifenden bildungssprachlichen Vokabulars Schülerinnen und Schüler häufig vor besonders große Herausforderungen (Hiebert & Lubliner, 2008). Vor allem für den domänenübergreifenden bildungssprachlichen Wortschatz kann jedoch davon ausgegangen werden, dass dessen Beherrschung für das Textverstehen und die Aufgabenbearbeitung im schulischen Kontext bedeutsam ist und eine Grundlage für die aktive Teilnahme am Unterricht und den weiteren Wissenserwerb bildet. Entsprechend wird ein starker Zusammenhang zwischen dem bildungssprachlichen Wortschatz von Jugendlichen und der Ausbildung schulischer Kompetenzen postuliert (Leseaux, Kieffer, Kelley und Harris, 2014; Uccelli, Galloway, Barr, Meneses & Dobbs, 2015). Empirische Unterstützung für diese Position stammt beispielsweise aus Studien von Townsend und Kollegen (Townsend, Filippini, Collins & Biancarosa, 2012) sowie Cunningham und Moore (1993), in denen jeweils inkrementelle und über die Effekte des alltagsprachlichen Wortschatzes hinausgehende Effekte des bildungssprachlichen Wortschatzes auf das Leseverstehen sowie weitere schulische Kompetenzen bei L1- und L2-Kindern und -Jugendlichen mit unterschiedlichem sozioökonomischen Hintergrund gefunden wurden.

Der bildungssprachliche Wortschatz scheint demnach für das Leseverstehen sowohl in einer Erst- als auch in einer Zweitsprache eine bedeutsame Rolle zu spielen (z. B. Quian, 2002; Scarcella, 2003; Uccelli et al., 2015). Dabei wird von einem reziproken Zusammenhang zwischen der Entwicklung von Textverstehensfähigkeiten von Schülerinnen und Schülern und dem Erwerb des bildungssprachlichen Wortschatzes ausgegangen: Kinder mit gut ausgeprägten Lesekompetenzen lesen häufiger und verfügen über bessere Fähigkeiten, die Bedeutung von Wörtern innerhalb eines Textes zu entschlüsseln als Schülerinnen und Schüler mit schwächeren Lesekompetenzen (Ceci & Papierno, 2005). Sie verfügen somit über mehr Gelegenheiten zum Erwerb von bildungssprachlichen Wörtern als ihre Peers mit schwächeren Lesekompetenzen. Gleichzeitig erhöhen breite bildungssprachliche Wortschatzkenntnisse die Wahrscheinlichkeit, bildungssprachliche Texte verstehen zu können und in der Folge bessere schulische Leistungen zu erzielen (Uccelli et al., 2015). Inwiefern diese Annahmen auch für den Bereich des bildungssprachlichen Hörverstehens gelten, bleibt bislang offen. So existieren gegenwärtig kaum empirische Studien zu Effekten des bildungssprachlichen Wortschatzes auf das Hörverstehen und vice versa. Eine Studie von Vidal (2003) legt nahe, dass der Erwerb eines bildungssprachlichen Vokabulars innerhalb bildungssprachlich geprägter mündlicher Interaktionen auch durch L2-Lernende in den frühen Phasen des Spracherwerbs möglich ist, wobei die Effekte für den Erwerb domänenübergreifenden bildungssprachlichen Vokabulars schwächer ausfielen als für domänenspezifische technische Wörter. Insgesamt konnten L2-Lernende umso mehr bildungssprachliche Wörter aus dem Diskurs erlernen, je besser ihre sprachlichen Fähigkeiten in der Zweitsprache waren. Allerdings handelt es sich bei den Teilnehmerinnen und Teilnehmern dieser Studie um Studierende und es wurde keine monolinguale Vergleichsgruppe eingesetzt. Inwiefern die Befunde generalisierbar sind, bleibt unklar. Weitere Studien, die den Zusammenhang zwischen bildungssprachlichem Wortschatz und bildungssprachlichem Hörverstehen explizit thematisieren, finden sich bislang nicht. Diese Forschungslücke ist umso bedauerlicher, als nicht davon ausgegangen werden kann, dass Befunde aus dem Bereich der Leseforschung ohne weiteres auf das Hörverstehen übertragen werden können. So zeigt sich beispielsweise, dass gesprochene Sprache in Bildungskontexten weniger bildungssprachliches Vokabular beinhaltet als geschriebene Sprache (Corson, 1997; Nagy & Townsend, 2012) und somit weniger Lerngelegenheiten für den Erwerb bildungssprachlicher Wörter bietet. Umgekehrt könnte das Verständnis bildungssprachlicher Hörverstehenstexte aufgrund des geringeren Vorkommens von bildungssprachlichem Vokabular weniger hohe Anforderungen an den Umfang des bildungssprachlichen Wortschatzes stellen als das Verständnis entsprechender Lesetexte.

Da sich die in Bildungskontexten verwendete Sprache stark von jener unterscheidet, die Schülerinnen und Schüler im Alltag nutzen, kann davon ausgegangen werden, dass Kinder und Jugendliche einen bildungssprachlichen Wortschatz nicht notwendigerweise in ihrer Freizeit erlernen, sondern im schulischen Alltag Gelegenheiten benötigen, die Bedeutung bildungssprachlicher Wörter innerhalb eines Diskurses zu klären und zu erlernen (Edwards & Westgate, 1994). Allerdings nehmen im schulischen Alltag die Situationen, in denen Gespräche über gehörte oder gelesene Texte bzw. die Bedeutung darin enthaltener Wörter geführt werden können, im Lauf der Schuljahre ab, da breite Wortschatzkenntnisse, auch hinsichtlich des bildungssprachlichen Vokabulars, im Unterricht oftmals als gegeben vorausgesetzt werden (Corson, 1997). Insbesondere für den domänenübergreifenden bildungssprachlichen Wortschatz scheinen Lehrkräfte häufig davon auszugehen, dass die unterschiedlichen Bedeutungen fachübergreifender bildungssprachlicher Wörter bereits bekannt sind (Corson, 1997; Hiebert & Lubliner, 2008). So merkt Gogolin (2009) an, dass viele Lehrkräfte zwar sensibilisiert für die Relevanz sprachlicher Bildung, auch in Bezug auf bildungssprachliche Kompetenzen, sind. Allerdings konzentrierten sich die meisten Lehrkräfte nach wie vor lediglich auf die Vermittlung eines fachspezifischen Vokabulars, die in vielerlei Hinsicht einem Vokabeltraining ähneln.

Häufig wird angenommen, dass Zweitsprachlernende zwar in der Lage sind, die alltagssprachlichen Wörter zu verstehen, denen sie in den vergleichsweise einfachen und tendenziell alltagssprachlich gestalteten Textbüchern während der Grundschulzeit begegnen, aber Schwierigkeiten mit dem Verständnis des abstrakteren bildungssprachlichen Vokabulars haben, das in den späteren Schuljahren im schulischen Kontext verwendet wird (August & Shanahan, 2006; Goldenberg, 2011). Allerdings liefern empirische Befunde Hinweise darauf, dass Unterschiede im bildungssprachlichen Wortschatz zwischen L1- und L2-Schülerinnen und -Schülern eher geringer ausgeprägt sein könnten als Differenzen im alltagssprachlichen Vokabular. Bialystok, Luk, Peets und Yang (2010) beispielsweise fanden, dass L2-Kinder generell einen weniger breiten Wortschatz aufweisen als L1-Kinder. Die Unterschiede zwischen beiden Gruppen waren allerdings für Wortschatzitems, die sich auf den häuslichen Kontext der Kinder bezogen, stärker ausgeprägt als für Wortschatzitems, die sich auf Bildungskontexte bezogen. Zu bemerken ist, dass in dieser Studie nicht für den sozioökonomischen Hintergrund der Kinder kontrolliert wurde, so dass unklar bleibt, inwiefern die Befunde mit eventuellen Unterschieden im sozioökonomischen Status der Kinder konfundiert sind.

Dennoch kann der Befund von Bialystok et al. (2010) als Unterstützung für die Annahme gewertet werden, dass sich bestimmte sprachliche Kompetenzen sprachspezifisch ausbilden und entweder in der Erst- oder der Zweitsprache erworben werden, in Abhängigkeit davon, wie viel Zeit im Sinne einer *time on task* mit dem Erwerb linguistischer Regularitäten in einer bestimmten Sprache verbracht wird (Gathercole, 2002). Dieser Hypothese entsprechend scheint der Wortschatz bei L2-Schülerinnen und -Schülern in einer Weise über die beiden Sprachen verteilt zu sein, die der Verteilung des sprachlichen Inputs in den beiden Sprachen entspricht (Pearson, Fernandez, Lewedeg & Oller, 1997; Umbel, Pearson, Fernández & Oller, 1992). Entscheidend ist dabei nicht allein das Ausmaß des sprachlichen Inputs, sondern darüber hinaus auch der Kontext, in dem der Wortschatz erworben wird. So ist anzunehmen, dass einige der von L2-Lernenden zu erwerbenden Wörter nur in bestimmten Bereichen verwendet werden und in derjenigen Sprache erlernt werden, die in diesen Kontexten üblicherweise gesprochen wird (Bialystok et al., 2010; Umbel et al., 1992). Da viele L2-Schülerinnen und -Schüler in ihrer sozialen Umwelt vorwiegend ihre Erstsprache verwenden, die Zweitsprache hingegen oft vornehmlich im schulischen Umfeld erworben wird (vgl. Janssen, Bosman & Leseman, 2013), kann davon ausgegangen werden, dass lexikalische Begriffe aus dem schulischen Bereich vor allem in der Zweitsprache vorliegen, alltägliche lexikalische Begriffe hingegen eher in einer Erstsprache. Es ist demnach anzunehmen, dass L2-Lernende große Schwierigkeiten mit dem Erwerb zweitsprachlicher Wörter haben, die vornehmlich außerhalb des L2-Kontextes verwendet werden, wie es für alltagssprachliche Begriffe häufig der Fall ist. Die in der Studie von Bialystok et al. (2010) berichteten differenziellen Unterschiede im Erwerb alltagssprachlichen bzw. bildungssprachlichen Vokabulars sprechen für die Gültigkeit dieser Hypothese. Auch eine Studie für portugiesischsprachige L2- und L1-Kinder erbrachte ähnliche Befunde (Engel de Abreu, Baldassi, Puglisi & Befi-Lopes, 2013). Hier fanden sich für portugiesischsprachige Kinder in Brasilien (L1-Gruppe) und portugiesischsprachige Kinder in Luxemburg (L2-Gruppe) keine Unterschiede in dem auf den häuslichen Kontext bezogenen Wortschatz, den beide Gruppen in der Erstsprache, d. h. auf Portugiesisch erwarben. Für den schulbezogenen Wortschatz, dessen Erwerb für die L1-Gruppe ebenfalls auf Portugiesisch, für die L2-Gruppe in der Zweitsprache Luxemburgisch erfolgte, zeigten sich hingegen Nachteile der letztgenannten Gruppe. In dieser Studie waren die L1- und die L2-Gruppen in Hinblick auf ihren sozioökonomischen Status vergleichbar, so dass davon ausgegangen werden kann, dass Unterschiede im Erwerb bildungssprachlichen Vokabulars nicht auf den sozioökonomischen Hintergrund der Kinder zurückzuführen sind. Darüber hinaus liefert eine Studie von Lawrence (2012) Hinweise darauf, dass L2-

Schülerinnen und -Schüler während der Sommerferien in Bezug auf ihren bildungssprachlichen Wortschatz einen so genannten *summer setback*-Effekt zeigen, d. h. einen Regress ihres bildungssprachlichen Wortschatzes. Obwohl in dieser Studie ein ähnlicher Effekt auch für L1-Schülerinnen und -Schüler mit niedrigem sozioökonomischem Status gefunden wurde, war der *summer setback*-Effekt bei den L2-Schülerinnen und -Schülern auch unter Kontrolle sozialer Hintergrundvariablen deutlicher ausgeprägt als bei ihren L1-Peers (Lawrence, 2012). Den Autoren zufolge lässt sich dieser Effekt dadurch erklären, dass die L2-Schülerinnen und -Schüler in ihrem sozialen Umfeld weniger häufig bildungssprachlichem Vokabular begegnen als L1-Schülerinnen und -Schüler. Auch diese Studie liefert somit Hinweise darauf, dass der Erwerb des bildungssprachlichen Wortschatzes vornehmlich im schulischen Kontext erfolgt, d. h. in der Zweitsprache der L2-Kinder.

Bei der Interpretation der Befunde zum Erwerb bildungssprachlichen Vokabulars in einer Zweitsprache muss beachtet werden, dass Unterschiede im Erwerb lexikalischer Kompetenzen auch durch den sozioökonomischen Hintergrund der L2-Kinder beeinflusst werden könnten, der für diese Kinder häufig niedriger ist als für ihre L1-Peers (vgl. Kapitel 4). Der Erwerb generellen bildungssprachlichen Vokabulars scheint auch für L1-Schülerinnen und -Schüler aus Familien mit einem niedrigen sozioökonomischen Hintergrund eine bedeutsame Hürde darzustellen (Corson, 1997). Eine mögliche Ursache für diese Schwierigkeiten liegt darin, dass sich die diskursiven Praktiken im soziokulturellen Umfeld dieser Kinder häufig stark von jenen unterscheiden, die in Bildungskontexten verwendet werden (Corson, 1997; Curen-ton & Justice, 2004). Im Gegensatz zu Kindern mit hohem sozioökonomischem Status sind sie somit in besonderem Maße auf schulische Instruktionen angewiesen, um das spezifische Register der Bildungssprache zu erwerben. Vor diesem Hintergrund könnte vermutet werden, dass Unterschiede im Erwerb bildungssprachlichen Vokabulars zwischen L1- und L2-Kindern mit niedrigem sozioökonomischem Status aufgrund des ähnlichen soziokulturellen Hintergrunds nivelliert werden. Die Befunde der Studien von Lawrence (2012) sowie Engel de Abreu et al. (2013), in denen sich Unterschiede im bildungssprachlichen Wortschatz zwischen L1- und L2-Kindern auch unter Kontrolle des sozioökonomischen Status fanden, deuten allerdings darauf hin, dass mit dem Erwerb bildungssprachlichen Vokabulars in einer Zweitsprache zusätzliche Hürden verbunden sind.

Bisherige Studien zum bildungssprachlichen Wortschatz in der Erst- oder Zweitsprache fokussierten vorwiegend auf korpuslinguistische Untersuchungen zur Häufigkeit des Vorkommens und der Art der Nutzung bildungssprachlicher Wörter in verschiedenen akademischen Domänen (z. B. Corson, 1997; Coxhead, 2000; Hyland & Tse, 2007). Die grundsätzli-

che Frage nach der Abgrenzbarkeit des bildungssprachlichen Wortschatzes vom alltagsprachlichen Wortschatz war hingegen bislang nicht Gegenstand empirischer Untersuchungen. Neben den Studien von Townsend et al. (2012) und Cunningham und Moore (1993) liegt eine weitere Studie vor, die die Annahme einer Trennung von bildungssprachlichem und alltagsprachlichem Wortschatz empirisch stützt. In dieser Studie zeigten Townsend und Collins (2009), dass L2-Kinder, die an einem Training zur Verbesserung ihrer bildungssprachlichen Wortschatzkenntnisse teilnahmen, im Anschluss an die Intervention tatsächlich bessere Leistungen in einem Test zur Erfassung des bildungssprachlichen Vokabulars aufwiesen, in dem die Kenntnis anderer als der trainierten Wörter erhoben wurde. Die Leistung in einem alltagsprachlichen Wortschatztest blieb demgegenüber konstant. Unklar bleibt allerdings, ob diese Befunde darauf zurückzuführen sind, dass der bildungssprachliche Wortschatz eine eigenständige Facette von Wortschatzkenntnissen bildet oder ob es sich um Trainingseffekte handelt. Dennoch liefern diese Ergebnisse, ähnlich wie die Befunde der Studien von Townsend et al. (2012) und Cunningham und Moore (1993), erste Hinweise darauf, dass sich bildungssprachliche und alltagsprachliche Wortschatzkenntnisse voneinander abgrenzen lassen könnten.

8.1 Forschungsfragen

Bislang liegen keine empirischen Studien vor, die die Annahme einer Abgrenzung bildungssprachlicher und alltagsprachlicher Wortschatzkenntnisse strukturanalytisch untermauern. Darüber hinaus wurden in bisherigen Studien die Zusammenhänge zwischen bildungssprachlichen Wortschatzkenntnissen und der Textverstehensleistung von Schülerinnen und Schülern lediglich für die Modalität des Lesens untersucht. Diese Studien liefern Hinweise darauf, dass das Leseverstehen im schulischen Kontext eng mit dem bildungssprachlichen Wortschatz von Schülerinnen und Schülern zusammenhängt. Unklar bleibt bislang, ob sich ein ähnlicher Effekt auch für den Bereich der bildungssprachlichen Hörverstehensfähigkeiten findet. Da Wortschatzkenntnisse allgemein einen hohen Zusammenhang zum Hörverstehen aufweisen (vgl. z. B. Field, 2003; Stæhr, 2009), ist zu vermuten, dass sich auch hier ein Zusammenhang des bildungssprachlichen Wortschatzes mit dem bildungssprachlichen Hörverstehen zeigt. Vor diesen Hintergrund werden in Teilstudie 4 die folgenden Forschungsfragen untersucht:

- 4.1 Lassen sich bildungssprachlicher und alltagsprachlicher Wortschatz voneinander abgrenzen?

In Anlehnung an theoretische Annahmen sowie Befunde aus Studien, die zeigen, dass bildungssprachliche Wortschatzkenntnisse einen spezifischen Anteil der Leseverstehensleistung aufklären (Townsend et al. 2012; Cunningham & Moore, 1993) bzw. dass Fördereffekte eines bildungssprachlichen Wortschatztrainings spezifisch für das bildungssprachliche Vokabular zu sein scheinen (Townsend & Collins, 2009), wird vermutet, dass akademisches und alltagssprachliches Vokabular zwei unterschiedliche Facetten des Wortschatzes abbilden. Vermutet wird, dass diese Facetten spezifische Zusammenhänge zu den entsprechenden Facetten des Hörverstehens zeigen; bildungssprachlicher Wortschatz und bildungssprachliches Hörverstehen also höher mit einander korrelieren als bildungssprachlicher Wortschatz und alltagssprachliches Hörverstehen und vice versa.

4.2 Unterscheidet sich die Ausprägung des bildungssprachlichen bzw. alltagssprachlichen Wortschatzes zwischen L1- und L2-Jugendlichen?

Für beide Facetten des Wortschatzes wird vermutet, dass L2-Jugendliche schwächere Leistungen aufweisen als L1-Jugendliche. Ausgehend von der Annahme, dass sowohl L1- als auch L2-Jugendliche aus Familien mit einem geringen sozioökonomischen Hintergrund bildungssprachliches Vokabular vorwiegend im schulischen Kontext und somit in derselben Instruktionssprache erwerben, wird vermutet, dass die Leistungsunterschiede zwischen den Sprachgruppen für diesen Bereich geringer ausfallen als für den alltagssprachlichen Wortschatz, dessen Erwerb bei L2-Schülerinnen und -Schülern meist in der Zweitsprache stattfindet.

Die Annahmen zu Unterschieden in der Ausprägung des bildungssprachlichen bzw. alltagssprachlichen Wortschatzes in Teilstudie 4 unterscheiden sich somit von den Annahmen in Teilstudie 3, für die ausgeprägtere Differenzen im bildungssprachlichen als im alltagssprachlichen Hörverstehen zwischen L1- und L2-Jugendlichen angenommen worden waren (vgl. Kapitel 7.4.3). Hintergrund der unterschiedlichen Annahmen in beiden Teilstudien ist die Hypothese, dass die Vermittlung zumindest des fachspezifischen bildungssprachlichen Wortschatzes im Unterricht wesentlich häufiger erfolgt als die Vermittlung weiterer bildungssprachlicher Merkmale, die nur selten Lerngegenstand des Unterrichts sind (vgl. Ahrenholz, 2013; Haag et al, 2013; Köhne, Kronenwerth, Redder, Schuth & Weinert, 2015). Dieser Hypothese zufolge bietet der Unterricht L1- und L2-Schülerinnen und -Schülern zwar gleichermaßen Lerngelegenheiten zum Erwerb des bildungssprachlichen Wortschatzes auf Deutsch; der Erwerb weiterer bildungssprachlicher Merkmale des Hörverstehens (z. B. syntaktische

Regularitäten) erfolgt allerdings überwiegend außerhalb des schulischen Kontextes und für L2-Jugendliche somit häufig nicht in der Instruktionssprache. Auf Basis der differenziellen Lerngelegenheiten zum Erwerb bildungssprachlicher Merkmale in der Instruktionssprache kann vermutet werden, dass sich für den bildungssprachlichen Wortschatz geringer ausgeprägte Leistungsunterschiede zwischen L1- und L2-Schülerinnen und -Schülern ergeben als für das bildungssprachliche Hörverstehen.

8.2 Methode

In die Teilstudie 4 gingen die Daten aller 479 Schülerinnen und Schüler der DELKO-Stichprobe ein. Da für 20 Jugendliche auf keiner der betrachteten Variablen Daten vorlagen, wurden diese von den folgenden Analysen ausgeschlossen. Weitere 60 L2-Jugendliche gaben an, in ihrem sozialen Alltag vornehmlich oder ausschließlich die Zweitsprache Deutsch zu nutzen. Es kann davon ausgegangen werden, dass diese Jugendlichen sowohl alltagssprachliche als auch bildungssprachliche Kompetenzen in der Zweitsprache erwerben. Da die im Rahmen von Forschungsfrage 4.2 vermuteten Leistungsunterschiede in der Ausprägung alltagssprachlicher Kompetenzen zwischen L1- und L2-Jugendlichen auf der Annahme einer Kontextabhängigkeit des Spracherwerbs beruhen, wurden diese Schülerinnen und Schüler ebenfalls aus den Analysen ausgeschlossen. Insgesamt gingen in Teilstudie 4 somit die Daten von 399 Jugendlichen ein, davon $n = 173$ (43%) L1-Schülerinnen und -Schüler und $n = 226$ (57%) L2-Schülerinnen und -Schüler.

Alltagssprachliche und bildungssprachliche Hörverstehenskompetenzen wurden anhand der auch in Teilstudie 3 eingesetzten Units des in der DELKO-Studie entwickelten Hörverstehenstests operationalisiert. Zur Erfassung des Wortschatzes wurde wie in den Teilstudien 1 und 2 die Wortschatzerweiterung der Grundintelligenztest Skala 2 genutzt (Weiß, 2006). Für Teilstudie 4 wurde dieses Testverfahren zusätzlich in Hinblick auf die Bildungssprachlichkeit der vorgegebenen Wörter klassifiziert. Zu diesem Zweck beurteilten drei Raterinnen die Wörter anhand eines vom *National Center for Research on Evaluation, Standards, and Student Testing* (CRESST) in den USA entwickelten Kodierungsschemas (vgl. Butler, Bailey, Stevens & Huang, 2004) als bildungs- bzw. alltagssprachlich. Dabei wurden 11 Wörter (37%) als bildungssprachlich (z. B. *Gesetzesabschnitt*), 19 (63%) als alltagssprachlich (z. B. *Stimmung*) klassifiziert. Die Interraterreliabilität für diese Beurteilung lag bei $\kappa = .66$ ($p < .01$). Bei mangelnder Übereinstimmung der Ratings wurde diejenige Klassifizierung übernommen, die zwei von drei Raterinnen gewählt hatten. Sowohl für den Hörverstehenstest als auch für den Wortschatztest wurden Items, die adversen Item-DIF in Abhängigkeit vom Sprachhintergrund

aufwiesen, aus den Analysen entfernt (vgl. Kapitel 7.2). Dies betraf drei Items (ein bildungssprachliches, zwei alltagssprachliche Items) des Hörverstehenstests und sieben Items (vier bildungssprachliche, drei alltagssprachliche Items) des Wortschatztests.

Zur Beantwortung der Forschungsfrage nach der Abgrenzbarkeit alltags- bzw. bildungssprachlicher Wortschatzkenntnisse wurde zunächst auf Basis der dichotom kodierten Items eine konfirmatorische Faktorenanalyse mittels des Programms Mplus 7.11 (Muthén & Muthén, 1998-2012) durchgeführt. Dabei wurden ein eindimensionales und ein zweidimensionales Modell der Wortschatzkenntnisse jeweils unter Berücksichtigung des sozioökonomischen und kulturellen Status der Schülerinnen und Schüler geschätzt, der analog zu den Teilstudien 1 bis 3 anhand des *Index of Economic, Social and Cultural Status (ESCS)* erfasst wurde. Als Schätzer wurde aufgrund der kategorialen Daten der *weighted least square estimator (WLSMV)* verwendet. Der Vergleich des ein- mit dem zweidimensionalen Wortschatzmodell erfolgte analog zu Teilstudie 3 anhand der Parameter RMSEA, CFI und TLI sowie eines χ^2 -Differenzentests für genestete Modelle. Zur zusätzlichen Validierung der Dimensionalität des Wortschatzes wurden Korrelationen zum bildungssprachlichen bzw. alltagssprachlichen Hörverstehen berechnet, die als latente Variablen in die Analysen eingingen. Sprachgruppenunterschiede in der Ausprägung bildungssprachlicher bzw. alltagssprachlicher Wortschatzkenntnisse wurden analog zu dem in Teilstudie 1 und 2 beschriebenen Vorgehen über eine Prüfung der Messinvarianz des geschätzten Modells sowie eine anschließende Überprüfung auf Differenzen in den Intercepts zwischen L1- und L2-Jugendlichen mittels Wald- χ^2 -Tests erfasst.

8.3 Ergebnisse

Die faktorenanalytische Überprüfung der Dimensionalität des Wortschatzes zeigt eine gute Anpassung sowohl des ein- als auch des zweidimensionalen Wortschatzmodells auf die Daten (vgl. Tabelle 9). Die globalen und absoluten Fit-Indizes sowie das Verhältnis χ^2/df unterscheiden sich nicht zwischen dem ein- und dem zweidimensionalen Modell. Der χ^2 -Differenzentest deutet auf eine bessere Anpassung für das eindimensionale Modell. Dabei findet sich zwischen bildungssprachlichem und alltagssprachlichem Wortschatz eine latente Interkorrelation von $r = .89$. Allerdings ergibt eine Überprüfung mittels eines Wald- χ^2 -Tests, dass die entsprechende Korrelation signifikant von 1 abweicht ($\Delta\chi^2(1) = 4.67, p = .03$).

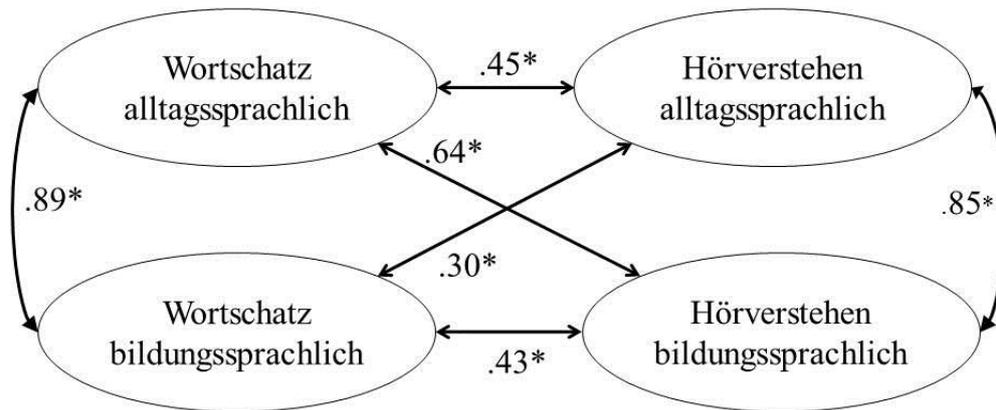
Tabelle 9.

Vergleich des ein- und des zweidimensionalen Modells des Wortschatzes in Teilstudie 4 (unter Kontrolle des ESCS)

Modell	χ^2	df	p	RMSEA	CFI	TLI	Verglichen mit	$\Delta \chi^2$	df	p
A: eindimensional	363.40	252	<.01	.03	.90	.89	--	--	--	--
B: zweidimensional	359.90	250	<.01	.03	.90	.89	Modell A	5.25	2	.07

Anmerkungen. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index.

Um die Zweidimensionalität des Wortschatzmodells weiter abzusichern, wurden in einem nächsten Schritt die latenten Korrelationen zum bildungssprachlichen bzw. alltagssprachlichen Hörverstehen berechnet. Unterschiede in der Höhe der Korrelationskoeffizienten wurden über Korrelationsdifferenzentests geprüft. Hinsichtlich der Stützung der Annahme einer Zweidimensionalität des Wortschatzes zeigen die Ergebnisse für den Zusammenhang der bildungssprachlichen Kompetenzen nur teilweise das erwartete Befundmuster (vgl. Abbildung 6): Der bildungssprachliche Wortschatz korreliert erwartungskonform signifikant stärker mit dem bildungssprachlichen Hörverstehen als mit dem alltagssprachlichen Hörverstehen ($z = 2.12$; $p = .03$). Für den alltagssprachlichen Wortschatz findet sich hingegen ein signifikant höherer Zusammenhang zum bildungssprachlichen als zum alltagssprachlichen Hörverstehen ($z = 3.85$; $p < .01$). Insgesamt liefern die Befunde somit eingeschränkt weitere empirische Stützung für die Annahme einer Zweidimensionalität der Wortschatzkenntnisse. Vor diesem Hintergrund sowie der Tatsache, dass der Vergleich des ein- mit dem zweidimensionalen Modells nur auf einem Signifikanzniveau von $p = .07$ für das eindimensionale Modell spricht, wurden die folgenden Analysen für das zweidimensionale Modell vorgenommen.



Anmerkungen. * $p < .01$

Abbildung 6. Zusammenhänge des bildungs- und des alltagssprachlichen Wortschatzes mit dem Außenkriterium Hörverstehen (unter Kontrolle des ESCS)

Zur Beantwortung von Forschungsfrage 4.2 nach Unterschieden in der Ausprägung der Wortschatzkenntnisse in einer Erst- bzw. Zweitsprache wurde zunächst das zweidimensionale Modell auf Messinvarianz zwischen den Gruppen überprüft. Obwohl die Wortschatzerweiterung der Grundintelligenztest Skala 2 für den in der Teilstudie 4 untersuchten Altersbereich normiert ist (vgl. Weiß, 2006), wiesen einige der Items in einer der beiden Sprachgruppen keine Varianz und extreme Schwierigkeiten auf. In der Folge resultierten für die Schätzung der Parameter zugrundeliegende Varianz-Kovarianzmatrix unbesetzte Zellen. Aufgrund der fehlenden Informationen könnte im Anschluss das skalare Modell der Messinvarianz nicht ohne Bias geschätzt werden. Um diese Problematik zu umgehen, wurden basierend auf Vorschlägen von Nasser-Abu Alhija und Wisenbaker (2006) je drei Parcels zur Operationalisierung des bildungssprachlichen bzw. alltagssprachlichen Wortschatzes gebildet. Um eine möglichst ausgeglichene Verteilung der Daten sicherzustellen, wurden dabei in jedem Parcel linksschiefe Items mit rechtsschiefen Items zusammengefasst (vgl. Nasser-Abu Alhija & Wisenbaker, 2006). Die auf Basis des beschriebenen Vorgehens vorgenommene Überprüfung auf Messinvarianz zeigt, dass das skalare Modell gegenüber den beiden anderen Modellen keine Verschlechterung im Fit aufweist, so dass von Messinvarianz zwischen den Sprachgruppen ausgegangen werden kann (vgl. Tabelle 10).

Tabelle 10.

Überprüfung des zweidimensionalen Wortschatzmodells in Teilstudie 4 auf Messinvarianz (unter Kontrolle des ESCS)

Modell	χ^2	df	p	RMSEA	CFI	TLI	SRMR	Vergleich mit	$\Delta \chi^2$	df	p
A: konfigural	40.74	26	.03	.05	.96	.94	.06	--	--	--	--
B: metrisch	40.80	28	.06	.05	.97	.95	.06	Modell A	0.06	2	.98
C: skalar	48.24	32	.03	.06	.95	.93	.06	Modell B	9.31	4	.06

Anmerkung. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; SRMR = Standardized Root Mean Square Residual.

Auf Basis des skalaren zweidimensionalen Modells wurden anschließend die latenten Mittelwerte der L1- und der L2-Schülerinnen und -Schüler für den alltags- bzw. bildungssprachlichen Wortschatz verglichen. Als Referenzgruppe dienten dabei die L1-Jugendlichen. Die Mittelwertsdifferenzen zwischen den Gruppen können in diesem Modell direkt als Einheiten der Effektstärke d interpretiert werden. Hypothesenkonform fanden sich für L2-Schülerinnen und -Schüler auf beiden Facetten des Wortschatzes schwächere Leistungen als für L1-Jugendliche (vgl. Tabelle 11). Allerdings sind die Effekte eher gering, so dass lediglich der Unterschied im alltagssprachlichen Wortschatz von statistischer Bedeutsamkeit ist. Der Befund besonders schwach ausgeprägter alltagssprachlicher Wortschatzkenntnisse entspricht zunächst dem erwarteten Muster. Eine Überprüfung der Ausprägung der Differenzen im alltagssprachlichen bzw. bildungssprachlichen Wortschatz mittels eines Wald- χ^2 -Tests erbrachte allerdings keine Hinweise darauf, dass die Leistungsunterschiede im alltagssprachlichen Wortschatz ausgeprägter sind als die im bildungssprachlichen Wortschatz (Wald- $\chi^2 = 0.00$; $p = .99$). Insgesamt finden sich somit nur eingeschränkt Hinweise auf besondere Leistungs-nachteile der L2-Jugendlichen für den alltagssprachlichen Wortschatz.

Tabelle 11.

Standardisierte latente Mittelwerte der Sprachgruppen im messinvarianten Mehrgruppenmodell für Teilstudie 4 (unter Kontrolle des ESCS)

	L1	L2	Wald- χ^2 -Werte für die Gruppenvergleiche
α (WS-Alltagssprachlich)	0.00	-0.38	6.04*
α (WS-Bildungssprachlich)	0.00	-0.31	3.01

Anmerkung: WS = Wortschatz. * $p < .05$

9. Abschließende Diskussion und Ausblick

Im Folgenden werden die Befunde der durchgeführten Studien zusammenfassend erörtert und in Hinblick auf die Anliegen dieser Arbeit diskutiert. Die Diskussion der Ergebnisse orientiert sich dabei an den in Kapitel 7 genannten Forschungsfragen. Zunächst wird dargestellt, inwiefern sich die strukturellen Beziehungen zwischen bottom-up-Prädiktoren des Hörverstehens in einer Erst- bzw. Zweitsprache ähneln. Betrachtet wird dabei ebenfalls, welche Prädiktoren der Hörverstehensleistung in einer Zweitsprache eine besondere Hürde darzustellen scheinen. In diesem Zusammenhang wird auch die Rolle des Hörverstehens für die Lesekompetenz der Jugendlichen diskutiert (Kapitel 9.1). Anschließend wird in Kapitel 9.2 basierend auf den Befunden der Teilstudien 3 bis 4 die Differenzierung und Ausprägung bildungssprachlicher und alltagsprachlicher Fähigkeiten von L1- und L2-Jugendlichen betrachtet. Die Limitationen der Studie werden in Kapitel 9.3 diskutiert. Abschließend zeigt Kapitel 9.4 Implikationen für die pädagogische Praxis bzw. zukünftige Forschungsarbeiten auf.

9.1 Hörverstehen in einer Erst- bzw. Zweitsprache: Linguistische Prädiktoren und Bedeutung für das Leseverstehen

Im Folgenden werden die Befunde für die Effekte und die Ausprägung von Prädiktoren des Hörverstehens in einer Erst- bzw. Zweitsprache diskutiert. Dabei wird auch auf die Bedeutung des Hörverstehens für das Leseverstehen eingegangen. Zunächst werden die Effekte der unterschiedlichen Prädiktoren auf die Hörverstehensleistung fokussiert, ehe auf die differenzielle Ausprägung linguistischer Prädiktoren des Hörverstehens in einer Erst- bzw. Zweitsprache eingegangen wird. Anschließend wird die Rolle des Hörverstehens für das Leseverstehen erläutert.

Effekte von Prädiktoren des Hörverstehens auf die Hörverstehenskompetenzen in einer Erst- bzw. Zweitsprache

Die dargestellten empirischen Befunde liefern Hinweise darauf, dass das Hörverstehen in einer Erst- bzw. Zweitsprache auf den unteren beiden Verarbeitungsebenen der Modelle von Anderson (1995) bzw. Kintsch und van Dijk (1987) im Wesentlichen von denselben Prädiktoren bestimmt wird. Die Ergebnisse sprechen somit gegen die Annahme sprachgruppenabhängiger Unterschiede in den Beziehungen einzelner linguistischer Prädiktoren mit den Hörverstehenskompetenzen von Schülerinnen und Schülern. In allen Gruppen zeigt sich gleichermaßen, dass die Hörverstehensleistung vorwiegend vom Wortschatz und den morpho-syntaktischen Kenntnissen der Schülerinnen und Schüler beeinflusst wird.

Dabei waren die Effekte der lexikalischen und morpho-syntaktischen Kenntnisse in den betrachteten Sprachgruppen gleichermaßen bedeutsam für die Hörverstehensleistung. Dieser Befund widerspricht dem *lexis first principle* (vgl. Rost, 1994), demzufolge die Worterkennung beim Hörverstehen den wichtigsten semantischen Verarbeitungsprozess bildet. Obwohl einige empirische Studien vorliegen, die diese Annahme bestätigen (z. B. Droop & Verhoeven, 2003; Felser & Clahsen, 2009), deuten die Befunde aus Teilstudie 1 eher darauf hin, dass morpho-syntaktische Kenntnisse eine gleichermaßen bedeutsame Rolle für das Hörverstehen spielen. Die differenziellen Befunde der Studien sind dabei möglicherweise auf die untersuchten Stichproben zurückzuführen: In bisherigen Studien zu Effekten von Wortschatz und morpho-syntaktischen Kenntnissen auf das zweitsprachliche Hörverstehen wurden vorwiegend jüngere L2-Lernende oder Lernende in den frühen Phasen des L2-Erwerbs betrachtet. Diese scheinen beim Hörverstehen häufig eine Verstehensstrategie zu wählen, bei der versucht wird, einzelne Wörter des Hörinhalts zu übersetzen (Lynch, 1998, Vandergrift, 2003). Für L2-Lernende, die bislang über wenig Kontakt zur Zweitsprache im Sinne einer *time-on task* verfügten, kann demnach vermutet werden, dass tatsächlich das vom *lexis first principle* postulierte Primat des Wortschatzes auf die Verarbeitung von Hörtexten Gültigkeit hat. Demgegenüber hatten die L2-Jugendlichen in der vorliegenden Studie zu einem großen Teil die gesamte Schullaufbahn in der Zweitsprache durchlaufen und somit eine längere Geschichte des Kontakts zur Zweitsprache. Obwohl sie teilweise dennoch über einen geringeren Wortschatz und geringere morpho-syntaktische Kenntnisse verfügten als die L1-Jugendlichen, scheinen ihre Zweitsprachkenntnisse doch ausreichend gewesen zu sein, um sich für die Bildung einer Textrepräsentation nicht allein auf eine Übersetzung der in den Hörtexten enthaltenen Wörter zu stützen. Vielmehr scheinen L2-Jugendliche in der vorliegenden Studie ebenso wie L1-Jugendliche auf der Ebene des Parsings die Analyse der syntaktischen und lexikalischen Informationen des Hörtextes zu integrieren, um eine mentale Repräsentation des Hörinhalts zu bilden.

Das Arbeitsgedächtnis und die phonologische Bewusstheit haben in der vorliegenden Studie keinen Effekt auf das Hörverstehen. Für die phonologische Bewusstheit ergibt sich allerdings ein sprachgruppenspezifischer Befund: Lediglich für die Gruppe der L2_{Türk}-Jugendlichen findet sich ein statistisch bedeutsamer Effekt der phonologischen Bewusstheit auf das Hörverstehen. Dieser kann als Hinweis darauf gewertet werden, dass in dieser Gruppe von Jugendlichen auch Merkmale auf der unteren, wahrnehmungsbasierten Ebene offenbar noch eine wesentlichere Rolle für das Hörverstehen spielen als in den beiden anderen Sprachgruppen. Eine mögliche Erklärung für diesen Befund wäre eine mangelhafte Automatisierung

der phonologischen Verarbeitung von Hörtexten in dieser Gruppe. Allerdings finden sich für die L2_{Türk}-Jugendlichen Ausprägungen der phonologischen Bewusstheit, die mit denen der anderen beiden Sprachgruppen vergleichbar sind, was gegen die Annahme einer geringer ausgeprägten Automatisierung der phonologischen Verarbeitung spricht. Eine weitere mögliche Ursache für den Effekt der phonologischen Bewusstheit auf das Hörverstehen bei türkischsprachigen Jugendlichen könnten sprachspezifische Verarbeitungsmechanismen sein: So erfordert das Türkische aufgrund seiner morphologischen Struktur die Beachtung einzelner Phoneme (Limbird & Stanat, 2006). Im Deutschen hingegen sind akzentuierte Silben und Wörter die Hauptinformationsträger, d. h. größere sprachliche Einheiten, die über die Ebene der Phoneme hinausgehen (Cauneau, 1992). Für die türkische Sprache ist es somit wesentlich als im Deutschen, eine Analyse des Sprachinputs auf Ebene einzelner Phoneme vorzunehmen. Für die L2_{Türk}-Jugendlichen könnte der Effekt der phonologischen Bewusstheit auf das Hörverstehen entsprechend dadurch zustande kommen, dass sie Segmentierungstechniken aus ihrer Erstsprache auch auf die Zweitsprache anwenden. Für L2_{Andere}-Jugendliche scheint es einen vergleichbaren Effekt nicht zu geben. Aufgrund der sprachlichen Heterogenität dieser Gruppe bleibt in der vorliegenden Studie allerdings unklar, ob dies aus einer phonologischen Ähnlichkeit der jeweiligen Erstsprachen mit der Zweitsprache Deutsch herrührt oder andere Ursachen hat.

Für das Arbeitsgedächtnis ließen sich in der vorliegenden Arbeit keine Effekte auf die Textverstehensleistung im Hören nachweisen. Dieser Befund steht im Widerspruch zu theoretischen Annahmen und empirischen Befunden, die die Arbeitsgedächtniskapazität als einen wesentlichen Prädiktor des Hörverstehens ausmachen (z. B. Call, 1985; Linck et al., 2014). Zur Erklärung der divergierenden Ergebnisse sind mehrere Ansätze denkbar. Erstens kann vermutet werden, dass das Ausbleiben von Effekten in der vorliegenden Studie ein Resultat des verwendeten Instrumentes zur Erfassung der Hörverstehensfähigkeiten ist, in dem nur sehr kurze und basale Hörtexte zum Einsatz kamen, deren Verarbeitung keine hohen Anforderungen an das Arbeitsgedächtnis stellt. Studien, in denen sich Effekte des Arbeitsgedächtnisses auf die Textverstehensleistung fanden, nutzen hingegen häufig Testverfahren, die die Verarbeitung längerer und komplexerer Hörtexte erforderten und somit höhere Anforderungen an die benötigten Arbeitsgedächtniskapazitäten stellen (vgl. Andringa, Olsthoorn, van Beuningen, Schoonen & Hulstijn, 2012; Daneman & Merikle, 1996). Zweitens könnten die zur Erfassung der Arbeitsgedächtniskapazität in der vorliegenden Studie verwendeten Ziffernspannen die geringen Zusammenhänge des Arbeitsgedächtnisses mit den Textverstehensmaßen erklären: Ziffernspannen können zwar als etablierte Maße zur Erfassung der Arbeitsgedäch-

niskapazität angesehen werden, zeigen aber in der Regel geringere Zusammenhänge zum Textverstehen als verbale Spannenmaße, die in der vorliegenden Studie aufgrund eines möglichen sprachlichen Bias nicht eingesetzt wurden (vgl. Daneman & Merikle, 1996). Drittens könnte der Effekt des Arbeitsgedächtnisses auf das Textverstehen über andere sprachliche Maße vermittelt werden. So fanden sich auch in anderen Studien zum Zusammenhang zwischen dem Arbeitsgedächtnis und Textverstehen nur geringe Korrelationen sobald weitere sprachliche Variablen berücksichtigt wurden (z. B. Andringa et al., 2012; Florit et al. 2014). Die Ergebnisse dieser Studien sowie der vorliegenden Arbeit deuten darauf hin, dass offenbar nicht die Kapazität des Arbeitsgedächtnisses per se entscheidend ist für das Sprachverstehen. Vielmehr scheint sich der Effekt des Arbeitsgedächtnisses über weitere sprachliche Maße zu vermitteln, wie etwa die Fähigkeit, morpho-syntaktische Regularitäten zur Sprachverarbeitung zu nutzen.

Zusammenfassend ist festzuhalten, dass in der vorliegenden Studie die Hörverstehensleistung bei L1- und L2-Jugendlichen gleichermaßen durch den Wortschatz sowie die morpho-syntaktischen Kenntnisse erklärt wird, während die phonologische Bewusstheit und das Arbeitsgedächtnis keine Effekte auf das Hörverstehen haben. Hinweise auf eine differenzielle Bedeutsamkeit einzelner linguistischer Prädiktoren auf das Hörverstehen in einer Erst- bzw. Zweitsprache finden sich nicht.

Unterschiede in der Ausprägung von Prädiktoren des Hörverstehens in einer Erst- bzw. Zweitsprache

Da sich in den Beziehungen der Prädiktoren des Hörverstehens mit den Hörverstehenskompetenzen keine Unterschiede zwischen den Sprachgruppen finden, kann vermutet werden, dass Unterschiede im erst- bzw. zweitsprachlichen Hörverstehen eher durch eine unterschiedliche Ausprägung der Prädiktoren bedingt werden. Dabei war vermutet worden, dass Prädiktoren, deren Entwicklung vom Kontakt zur Zweitsprache im Sinne einer *time-on-task* abhängt, d. h. Wortschatz und morpho-syntaktische Kenntnisse, bei L2-Schülerinnen und -Schülern schwächer ausgeprägt sein könnten als bei ihren L1-Peers. Aufgrund der hohen Bedeutung dieser Parsing-Komponenten für die Bildung einer mentalen Textrepräsentation sollten in der Folge auch schwächere Textverstehensleistungen im Hören und Lesen resultieren. Für das Arbeitsgedächtnis und die phonologische Bewusstheit waren hingegen keine Gruppenunterschiede angenommen worden.

Tatsächlich weisen L1- und L2-Jugendliche in der vorliegenden Studie ähnliche Leistungen in Bezug auf das Arbeitsgedächtnis auf. Während in einigen bisherigen Studien zur

Arbeitsgedächtniskapazität in einer Zweitsprache Hinweise auf Nachteile von L2-Lernenden gefunden wurden (z. B. Call, 1985; Grotjahn, 2005; Lesaux, Lipka & Siegel, 2006), liefern die Ergebnisse der vorliegenden Arbeit Unterstützung für die Annahme, dass das Auftreten von Unterschieden in der Arbeitsgedächtniskapazität zwischen L1- und L2-Lernenden vom verwendeten Maß zur Erfassung des Arbeitsgedächtnisses abhängen könnte. Die häufig zur Erfassung der Arbeitsgedächtniskapazität genutzten Satz- und Wortspannenmaße könnten L2-Schülerinnen und -Schüler benachteiligen, die mit den sprachlichen Strukturen der Zweitsprache weniger vertraut sind als ihre L1-Peers (Lanfranchi & Swanson, 2005; Olsthoorn, Andringa & Hulstijn, 2012). Verfahren, die geringere Anforderungen an die Zweitsprachkenntnisse stellen, können demgegenüber eher als kulturfaire Maße zur Erfassung der Arbeitsgedächtniskapazität gelten. Werden, wie in der vorliegenden Studie, Verfahren verwendet, die nicht differenziell valide sind, so scheinen die Arbeitsgedächtnisleistungen in einer Erst- und Zweitsprache ähnlich ausgeprägt zu sein.

Für die Ausprägung der phonologischen Bewusstheit waren ebenfalls keine Unterschiede zwischen L1- und L2-Jugendlichen erwartet worden, da die phonologische Bewusstheit als Aspekt metasprachlicher Bewusstheit durch den Kontakt zu unterschiedlichen lautlichen Systemen eher unterstützt werden sollte (Bialystok, 2009; Bruck & Genesee, 1995). Auch hier entsprechen die Ergebnisse den Erwartungen und weisen darauf hin, dass sich die phonologische Bewusstheit in einer Erst- bzw. Zweitsprache nicht unterscheidet.

Hinsichtlich von Wortschatz und morpho-syntaktischen Kenntnissen war vermutet worden, dass L2-Schülerinnen und -Schüler aufgrund des geringeren Kontakts zur Zweitsprache schwächere Kompetenzen zeigen als L1-Schülerinnen und -Schüler. Angenommen wurde, dass sich die entsprechenden Kompetenzunterschiede unter Kontrolle des sozioökonomischen Status zwar reduzieren, aber auch unter Berücksichtigung dieses Faktors persistieren. Tatsächlich lassen sich für die Gruppe der L2_{Andere}-Jugendlichen jedoch keine Nachteile im Hinblick auf die morpho-syntaktischen Fähigkeiten mehr feststellen, wenn für den sozioökonomischen Hintergrund kontrolliert wird. Schwächer ausgeprägte morpho-syntaktische Kenntnisse in dieser Gruppe scheinen mithin eher auf den tendenziell geringeren sozioökonomischen Status der Schülerinnen und Schüler zurückführbar zu sein als auf ihren Status als Zweitsprachlernende. L2_{Türk}-Jugendliche hingegen weisen zumindest den Ergebnissen aus Teilstudie 2 zufolge persistierende Schwächen im Bereich der morpho-syntaktischen Kenntnisse auf. Anzumerken ist, dass auch in Teilstudie 1 die latenten Mittelwerte der morpho-syntaktischen Kenntnisse zwischen den drei Sprachgruppen verglichen wurden. Anders als in Teilstudie 2 fanden sich dabei unter Kontrolle des ESCS auch für die Gruppe der L2_{Türk}-Jugendlichen kei-

ne Nachteile im Bereich der morpho-syntaktischen Kenntnisse gegenüber L1-Jugendlichen, wenn für den sozialen Hintergrund der Schülerinnen und Schüler kontrolliert wurde. Vor diesem Hintergrund kann vermutet werden, dass Schwächen im Bereich morpho-syntaktischer Fähigkeiten bei L2_{Türk}-Jugendlichen ähnlich wie bei L2_{Andere}-Jugendlichen primär ein Effekt des differenziellen sozio-ökonomischen Status sind. Gestützt wird diese Vermutung dadurch, dass die in Teilstudie 2 auch unter Berücksichtigung des sozialen Hintergrunds bestehenden Disparitäten in den morpho-syntaktischen Kenntnissen zwar statistisch bedeutsam, mit $d = .20$ allerdings als von eher geringer praktischer Bedeutsamkeit zu beurteilen sind.

Für den Wortschatz finden sich in beiden Gruppen der L2-Jugendlichen auch unter Kontrolle des sozioökonomischen Status Leistungsnachteile gegenüber L1-Jugendlichen. Die Ergebnisse aus Teilstudie 1 und 2 liefern somit empirische Hinweise auf die Gültigkeit der Annahme, dass Probleme im Wortschatzerwerb bei fortgeschrittenen Zweitsprachlernenden eine wesentliche Schwierigkeit für die Ausbildung von Hörverstehensfähigkeiten bilden (Kelly, 1991). Für L2_{Andere}-Jugendliche deuten die Ergebnisse darauf hin, dass die entsprechenden Kompetenzunterschiede mit $d = .28$ von eher geringer praktischer Bedeutsamkeit sind. Die Disparitäten zwischen L2_{Türk}-Schülerinnen und -Schülern und ihren mono-lingualen Peers liegen mit $d = .52$ hingegen im Bereich einer mittleren Effektstärke und können als Hinweis darauf gedeutet werden, dass in dieser Gruppe ein bedeutsamer Rückstand im Erwerb von Wortschatzkenntnissen besteht.

In allen Teilstudien zeigt sich, dass Unterschiede in den erzielten sprachlichen Kompetenzen zwischen L1- und L2-Jugendlichen zu einem großen Teil durch den differenziellen sozialen Hintergrund beider Gruppen erklärt werden können. So finden sich für die Gruppe der L2_{Andere}-Jugendlichen ohne Kontrolle familiärer Strukturmerkmale schwächere Leistungen gegenüber L1-Jugendlichen im Wortschatz und den morpho-syntaktischen Kenntnissen sowie in der Hör- und der Leseverstehenskompetenz. Unter Kontrolle des ESCS findet sich für diese Gruppe hingegen nur noch ein im Vergleich zu ihren L1-Peers signifikant niedrigerer Wortschatz. Für L2_{Andere}-Jugendliche scheinen Nachteile im Erwerb sprachlicher Kompetenzen demnach offenbar zum wesentlichen Teil auf ihren im Mittel geringeren sozioökonomischen Status zurückführbar zu sein. Dieser Befund entspricht den Ergebnissen der groß angelegten nationalen und internationalen Vergleichsstudien, wie etwa PISA (Baumert & Schümer, 2001) oder dem Ländervergleich des IQB (Haag et al., 2012), in deren Rahmen sich die Leistungsdifferenzen zwischen Kindern aus zugewanderten Familien und Kindern ohne Zuwanderungshintergrund ebenfalls substanziell verringerten, wenn soziale Hintergrundvariablen berücksichtigt wurden. Allerdings blieben in diesen Studien auch nach Kontrolle des

sozioökonomischen Hintergrunds noch Leistungsdifferenzen zwischen einem großen Teil der zugewanderten Schülerinnen und Schüler und Jugendlichen ohne Zuwanderungshintergrund bestehen. In der vorliegenden Studie hingegen zeigten sich für die L2_{Andere}-Jugendlichen unter Kontrolle familiärer Strukturmerkmale nahezu keine Unterschiede zu L1-Jugendlichen in Textverstehenskompetenzen bzw. deren Vorläuferfähigkeiten. Zu beachten ist bei der Einordnung der Befunde, dass in der vorliegenden Studie eine spezielle Substichprobe von Jugendlichen mit sehr geringem sozioökonomischem Hintergrund betrachtet wurde, während die genannten Studien Jugendliche mit unterschiedlichem sozioökonomischem Hintergrund einbeziehen.

Für L2_{Türk}-Jugendliche reduzierten sich, analog zu den Befunden aus den PISA-Studien bzw. dem IQB-Ländervergleich, zwar die Disparitäten in den Textverstehenskompetenzen bzw. deren Vorläuferfähigkeiten, wenn der sozioökonomische Hintergrund berücksichtigt wurde. Dennoch finden sich auch unter Kontrolle dieses strukturellen Merkmals noch signifikant schwächere Wortschatz-, morpho-syntaktische und Hör- bzw. Leseverstehensleistungen der L2_{Türk}-Jugendlichen im Vergleich zu ihren L1-Peers. Für diese Gruppe der L2-Jugendlichen können Disparitäten im Kompetenzerwerb offenbar in geringerem Ausmaß als bei L2_{Andere}-Jugendlichen durch den sozioökonomischen Hintergrund der Schülerinnen und Schüler erklärt werden. Dieser Befund entspricht Ergebnissen anderer Studien, in denen ebenfalls Disparitäten im Erwerb schulischer Leistungen zwischen L1-Schülerinnen und -Schülern und L2-Lernenden mit türkischer Herkunftssprache bestanden, wenn der unterschiedliche soziale und kulturelle Hintergrund der Schülerinnen und Schüler berücksichtigt wurde (Droop & Verhoeven, 2003; Haag et al., 2012; OECD, 2010; 2013; Stanat, Rauch & Segeritz, 2010).

Offenbar scheint demnach nicht der Erwerb sprachlicher Fähigkeiten in einer Zweitsprache per se ein Risikofaktor für die Ausbildung schwacher Hörverstehenskompetenzen zu sein. Vielmehr deuten die Ergebnisse auf sprachgruppenspezifische Effekte hin. Eine mögliche Erklärung für die auch unter Berücksichtigung familiärer Strukturmerkmale persistierenden schwächeren Leistungen der L2_{Türk}-Jugendlichen wären Unterschiede in familiären Prozessmerkmalen. Hier kommt insbesondere der in der Familie gesprochenen Sprache eine wichtige Rolle zu, die in einer Reihe von Studien als bedeutsamer Faktor für den Erwerb schulischer Kompetenzen in einer Zweitsprache identifiziert wurde (Haag et al., 2012; Stanat et al., 2010, Stanat et al., 2012). Die Familiensprache bildet dabei einen Indikator für die Gelegenheiten zum Erwerb der Zweitsprache im Sinne einer *time-on-task*: L2-Schülerinnen und -Schüler, die auch im familiären Kontext die Zweitsprache verwenden, verfügen demnach

über mehr Lerngelegenheit zum Erwerb zweitsprachlicher Kompetenzen als Kinder, die innerhalb der Familie die jeweilige Erstsprache sprechen.

In der vorliegenden Studie zeigt sich, dass sowohl die meisten L2-Jugendlichen mit türkischer Herkunftssprache als auch die meisten L2-Jugendlichen mit einer anderen Herkunftssprache als Türkisch mit ihren Eltern ausschließlich oder überwiegend die jeweilige Erstsprache sprechen (vgl. Tabelle 12). Signifikante Unterschiede in der Häufigkeit der Verwendung der Erst- bzw. Zweitsprache mit den Eltern finden sich zwischen den Sprachgruppen nicht. Trotz der überwiegenden Nutzung der jeweiligen Erstsprache innerhalb der Familie und den damit verbundenen geringeren Lerngelegenheiten zum Erwerb des Deutschen gelingt es L2_{Andere}-Jugendlichen offenbar, sprachliche Kompetenzen zu erzielen, die – sofern die unterschiedlichen sozioökonomischen Strukturmerkmale der Familie berücksichtigt werden – größtenteils denen von L1-Jugendlichen entsprechen. Für die L2_{Türk}-Jugendlichen zeigen sich hingegen schwächere sprachliche Kompetenzen. Da beide Gruppen von L2-Jugendlichen gleichermaßen die Erstsprache innerhalb der Familie verwenden, kann vermutet werden, dass der Effekt der Familiensprache allein nicht ausreichend ist, um den differenziellen Erwerb sprachlicher Fähigkeiten in den beiden Gruppen zu erklären.

Tabelle 12.

Unterschiede im familiären und außerfamiliären Gebrauch der Erst- bzw. Zweitsprache bei L2_{Andere} und L2_{Türk}-Jugendlichen in der DELKO-Stichprobe (unter Kontrolle des ESCS)

	nur oder überwiegend Erstsprache		nur oder überwiegend Deutsch		Überprüfung auf Gruppen unterschiede		
	L2 _{Türk}	L2 _{Andere}	L2 _{Türk}	L2 _{Andere}	<i>F</i>	<i>p</i>	η^2
Sprachgebrauch mit Eltern	90,6%	89,7%	9,4%	10,3%	1.64	.20	.01
Sprachgebrauch in der Nachbarschaft	29,6%	9,4%	70,4%	90,6%	8.71	.00	.05

Unterschiede im Ausmaß des Kontakts zur Zweitsprache im Sinne einer *time-on-task* ergeben sich in der vorliegenden Studie aus dem Sprachgebrauch im außerfamiliären sozialen Umfeld: In Hinblick auf die in der Wohnumgebung bzw. Nachbarschaft gesprochene Sprache geben L2-Jugendliche mit türkischer Herkunftssprache signifikant häufiger an, ihre Erstsprache zu nutzen als L2_{Andere}-Jugendliche (vgl. Tabelle 12). Die Unterschiede im Gebrauch der Herkunftssprache in der Nachbarschaft liegen im Bereich einer mittleren Effektstärke.

Der Befund einer gegenüber anderen Gruppen von L2-Jugendlichen weniger stark ausgeprägten Nutzung der Zweitsprache und damit verbundenen verstärkten Nutzung der Erstsprache Türkisch im außerfamiliären sozialen Kontext entspricht Annahmen von Backus (2005). Demnach ist davon auszugehen, dass die Beibehaltung der Erstsprache im sozialen Kontext für Familien mit türkischer Herkunftssprache aufgrund der großen türkischsprachigen Gemeinschaft sowie der hohen Verfügbarkeit türkischsprachiger Medien in stärkerem Maße möglich ist als für andere Gruppen mit nichtdeutscher Herkunftssprache. In der Folge ergeben sich für die Gruppe der Familien mit türkischer Herkunftssprache meist weniger Kontakte zur Zweitsprache Deutsch im Sinne einer *time-on-task*, die zu einer schwächeren Ausprägung sprachlicher Kompetenzen in der Zweitsprache beitragen könnten. Empirisch werden diese theoretischen Annahmen von Befunden aus Österreich gestützt, die zeigen, dass L2-Kinder mit türkischer Herkunftssprache weniger häufig deutschsprachige Medien nutzen und weniger deutschsprachige Freunde haben als L2-Kinder, deren Familien aus dem ehemaligen Jugoslawien zugewandert sind (Peltzer-Karpf et al., 2006). Allerdings wurde in dieser Studie nicht für den sozioökonomischen Hintergrund der Kinder kontrolliert, obwohl die Autoren angeben, dass zwischen beiden Zuwanderergruppen Unterschiede im Grad der elterlichen Bildung bestanden.

Auch Analysen von Daten des Nationalen Bildungspanels (NEPS) weisen darauf hin, dass unter Zuwanderern aus der Türkei vor allem in der Herkunftssprache kommuniziert wird. Der Anteil von türkischstämmigen Personen, die zu Hause ausschließlich oder meistens Türkisch sprechen liegt mit 72,1 % deutlich über den entsprechenden Anteilen von Zuwanderern aus Polen oder der ehemaligen Sowjetunion, die zuhause ihre jeweilige Herkunftssprache nutzen (Strobel & Kristen, 2015). Auch der Anteil eigenethnischer Bekannter, mit denen am ehesten in der jeweiligen Herkunftssprache kommuniziert werden kann, ist unter türkischstämmigen Zuwanderern besonders hoch und liegt im Mittel bei 80% (Strobel & Kristen, 2015). Beachtet werden muss dabei, dass in die NEPS-Analysen nur Daten von Personen eingehen, die selbst nach Deutschland zugewandert waren, d. h. der ersten Zuwanderergeneration angehörten. Aussagen über Muster des Sprachgebrauchs bei Zuwanderern der zweiten Generation lassen sich anhand der berichteten Daten demnach nicht treffen. Allerdings berichtet Esser (2001) Ergebnisse, die darauf hinweisen, dass in türkischen Familien über die Generationen hinweg eine Tendenz zur Segmentation besteht, so dass anzunehmen ist, dass sich auch für die zweite türkischstämmige Zuwanderergeneration eine starke Nutzung der Herkunftssprache im sozialen Alltag zeigen sollte. Die in der vorliegenden Studie berichteten Muster des Sprachgebrauchs innerhalb der Familie bzw. im sozialen Umfeld sind mit dieser

Annahme konform und deuten darauf hin, dass auch türkischstämmige Jugendliche, die mehrheitlich der zweiten Zuwanderergeneration angehören, sehr häufig die Herkunftssprache im Alltag nutzen. Dieser Fokus auf intraethnische soziale Kontakte bzw. den Gebrauch der Erstsprache könnte dazu führen, dass Kinder aus Familien mit türkischem Zuwanderungshintergrund über weniger Lerngelegenheiten zum Erwerb der deutschen Sprache verfügen als Kinder aus Familien mit einem anderen Zuwanderungshintergrund.

In Anbindung an diese Befunde kann auch für die vorliegende Studie vermutet werden, dass ein stärkerer Fokus auf die Erstsprache die schwächeren Kenntnisse in der Zweitsprache bei L2_{Türk}-Jugendlichen zumindest mitbedingt. Ergebnisse des IQB-Ländervergleichs 2011 (Haag et al., 2012) deuten zwar darauf hin, dass auch unter Kontrolle der Familiensprache noch substanzielle Kompetenznachteile bei türkischstämmigen Jugendlichen gegenüber Jugendlichen ohne Zuwanderungshintergrund bestehen bleiben. Die Befunde der vorliegenden Arbeit legen allerdings nahe, dass möglicherweise nicht nur der Sprachgebrauch in der Familie als Indikator für die Lerngelegenheiten zum Erwerb der Zweitsprache von Bedeutung sein könnte, sondern auch die für die Kommunikation im außerfamiliären sozialen Umfeld genutzte Sprache.

Betrachtet man beispielsweise die Effekte des Sprachgebrauchs auf den Wortschatz im Rahmen eines regressionsanalytischen Vorgehens, so zeigt sich, dass die Umgangssprache im Wohnviertel einen Effekt auf die Wortschatzentwicklung in der Zweitsprache aufweist, während kein bedeutsamer Effekt des Sprachgebrauchs mit den Eltern zu finden ist (vgl. Tabelle 13). Obwohl der Anteil der durch den Sprachgebrauch aufgeklärter Varianz im zweitsprachlichen Wortschatz insgesamt eher niedrig ist, können die Befunde der vorliegenden Studie als Hinweise darauf gedeutet werden, dass für den Erwerb zweitsprachlicher Kompetenzen nicht primär nur die Familiensprache von Bedeutsamkeit sein könnte, sondern die Summe der zweitsprachlichen Kontakte insgesamt eine Rolle spielt (vgl. auch Esser, 2006).

Tabelle 13.

Standardisierte Betakoeffizienten im Regressionsmodell zur Vorhersage des bildungs- und alltagssprachlichen Wortschatzes bei L2-Jugendlichen

	Alltagssprachlicher Wortschatz	Bildungssprachlicher Wortschatz
ESCS	.12	.27*
Sprachgebrauch mit Eltern	-.03	-.10
Sprachgebrauch im Viertel/Nachbarschaft	.17*	.15*
R^2	.05*	.10*

Anmerkung: Sprachgebrauch von 1 = nur Muttersprache bis 4 = nur Deutsch. *p < .05

Empirisch konnte entsprechend bereits gezeigt werden, dass türkischsprachige Kinder umso bessere zweitsprachliche Kompetenzen aufweisen, je mehr Kontakt – auch auf das außerfamiliäre soziale Umfeld bezogen – sie zur Zweitsprache Deutsch hatten (Röhr-Sendlmeier, 1985). Auch Esser (2006) geht davon aus, dass der ethnische Kontext, z. B. über die ethnische Konzentration im Wohngebiet, Einfluss auf den Zweitspracherwerb nehmen kann und berichtet entsprechende empirische Befunde. Dabei scheinen sich für die ethnische Konzentration im Wohngebiet sowie in den Schulklassen Interaktionseffekte auf den Zweitspracherwerb zu finden: Eine hohe ethnische Konzentration bzw. mangelnde Gelegenheiten zum Kontakt mit der Zweitsprache könnten dabei andere ungünstige Erwerbsvoraussetzungen, wie etwa ein geringes Bildungsniveau der Eltern, verstärken (Esser, 2006). Entsprechend dieser Annahme liegen Studien vor, die auf Effekte der ethnischen Komposition von Kindern in Kindergärten auf deren sprachlichen Kompetenzerwerb hinweisen (z. B. Biedinger & Becker, 2010; Niklas, Schmiedeler, Pröstler & Schneider, 2011). Befunde aus dem schulischen Bereich deuten demgegenüber eher darauf hin, dass sich weniger ethnische Kompositionsmerkmale als vielmehr die leistungsbezogene Zusammensetzung der Schülerschaft auf den Kompetenzerwerb auszuwirken scheint (Stanat, Schwippert & Gröhlich, 2010). Dennoch legen die Befunde der vorliegenden Studie nahe, dass mit dem Gebrauch der Herkunftssprache und dem daraus resultierenden geringeren Kontakt zur Zweitsprache offenbar Disparitäten im Erwerb sprachlicher Kompetenzen von L2-Schülerinnen und -Schülern gegenüber ihren L1-Peers einher gehen könnten.

Die Ergebnisse der vorliegenden Arbeit sind allerdings sicherlich nicht dahingehend zu interpretieren, dass zugewanderte Familien möglichst ausschließlich die Zweitsprache im sozialen Alltag gebrauchen sollten. Esser (2006) gibt hinsichtlich der Nutzung der jeweiligen Erstsprache an, „alles, was zählt ist der Zweitspracherwerb und dazu *zusätzliche* muttersprachliche Kompetenzen sind ohne Bedeutung.“ (S. 17). Tatsächlich gelingt es vielen L2-Schülerinnen und -Schülern vergleichbare schulische Kompetenzen zu erwerben wie ihre L1-Peers, wenn sie im außerschulischen Kontext häufig die Erstsprache gebrauchen (vgl. August & Shanahan, 2006). Umgekehrt zeigten sich beispielsweise in einer niederländischen Schule auch dann sprachliche Nachteile von L2-Kindern, wenn diese im außerschulischen Kontext häufig die Zweitsprache Niederländisch sprachen (Driessen, van der Slik & de Bot, 2002). Die Autoren der letztgenannten Studie empfinden die Empfehlung, innerhalb der Familien von L2-Schülerinnen und -Schülern solle möglichst häufig die Zweitsprache verwendet werden entsprechend als „a simplistic conception of the relations between languages used at home and at school“ (S. 191). Die Befunde der vorliegenden Studie deuten allerdings darauf hin,

dass mit einer generell – d. h. sowohl auf die Sprache innerhalb der Familie als auch die Sprache im sozialen Umfeld bezogenen – geringeren Nutzung der Zweitsprache offenbar Risiken für die Ausbildung sprachlicher Kompetenzen in der Zweitsprache einhergehen könnten, die im Einklang mit den Annahmen der *time-on-task* Hypothese wären.

Die Rolle des Hörverstehens für das Leseverstehen in einer Erst- bzw. Zweitsprache

Die vorliegenden Studien zeigen, dass das Hörverstehen in einer Erst- bzw. Zweitsprache vorwiegend von der auf der Ebene des Parsings angesiedelten Verarbeitung lexikalischer und syntaktischer Merkmale abzuhängen scheint. Die wahrnehmungsbasierte Verarbeitung auf der untersten Ebene der Modelle von Anderson (1995) bzw. Kintsch und van Dijk (1987) ist sowohl für die L2-Jugendlichen als auch für die L1-Schülerinnen und -Schüler demgegenüber weniger bedeutsam für die Hörverstehensleistung. Für den Bereich des Hörverstehens zeigen sich damit Befunde, die theoretischen Annahmen zum Leseverstehen, wie etwa dem *Simple View of Reading*, entsprechen. Postuliert wird dabei, dass die Textverstehensleistung mit zunehmender Automatisierung des Verstehensprozesses weniger stark von Merkmalen auf den unteren, wahrnehmungsbasierten Ebenen wie etwa der Verarbeitung des phonologischen Inputs beim Hörverstehen oder dem Dekodieren beim Leseverstehen bestimmt wird, während der Einfluss der semantischen und syntaktischen Analyse des Textinhalts auf das Textverstehen zunimmt (Hoover & Gough, 1990). Auch für den Bereich des Lesens liefern die Befunde der vorliegenden Arbeit erneut empirische Unterstützung für diese Annahme: In allen drei betrachteten Sprachgruppen hängt die Leseleistung jugendlicher Schülerinnen und Schüler stärker von ihren Hörverstehensfähigkeiten ab als von der Lesegeschwindigkeit als Aspekt des Dekodierens. Ein in der Forschungsliteratur postulierter engerer Zusammenhang zwischen dem Hörverstehen und dem Leseverstehen in einer Zweitsprache (vgl. z. B. August & Shanahan, 2006) findet sich dabei nicht. Vielmehr scheinen Unterschiede in der erzielten Lesekompetenz in einer Erst- bzw. Zweitsprache ebenso wie Unterschiede in den Hörverstehenskompetenzen primär ein Resultat schwächer ausgeprägter grundlegender sprachlicher Fähigkeiten zu sein und weniger aus einer unterschiedlichen Textverarbeitung zu resultieren. Sowohl für das Lesen als auch für das Zuhören weisen die Befunde der vorliegenden Studie somit darauf hin, dass Parsing-Komponenten wie der Wortschatz und die morphosyntaktischen Kenntnisse für das Textverstehen in den höheren Klassenstufen eine bedeutsamere Rolle spielen als die wahrnehmungsbasierte Verarbeitung. Dieser Befund ist dabei zumindest für L2-Jugendliche, die mehrheitlich in Deutschland geboren und aufgewachsen sind

und somit über eine vergleichsweise lange Kontaktdauer zur Zweitsprache verfügen, gleichermaßen gültig wie für L1-Jugendliche.

Unterschiede im Textverstehen in der Modalität Lesen ergeben sich zwischen L1- und L2-Jugendlichen demnach analog zum Textverstehen in der Modalität Hören aus einer differenziellen Ausprägung linguistischer Prädiktoren. Neben den schwach ausgeprägten Hörverstehensfähigkeiten in der Gruppe der L2_{Türk}-Schülerinnen und -schüler zeigt sich dabei in beiden Gruppen von L2-Schülerinnen und -Schülern auch eine im Vergleich zu L1-Jugendlichen geringere Lesegeschwindigkeit. Diese hängt als Aspekt des Dekodierens eng mit der Entwicklung der phonologischen Bewusstheit zusammen. Entsprechend waren wie für die phonologische Bewusstheit auch für das Dekodieren keine differenziellen Ausprägungen für L1- und L2-Jugendliche erwartet worden. Für L2_{Andere}-Jugendliche werden die Nachteile im Bereich des Dekodierens allerdings nivelliert, wenn für den sozioökonomischen Status kontrolliert wird, der als ein bedeutsamer Prädiktor des Dekodierens gelten kann (vgl. Noble, Farah, & McCandliss, 2006). Für diese Gruppe findet sich somit das erwartete Muster. Für L2_{Türk}-Jugendliche hingegen zeigt sich eine geringere Lesegeschwindigkeit gegenüber den L1-Jugendlichen auch unter Kontrolle des sozialen und kulturellen Hintergrunds. Möglicherweise ist dies eine Folge der generell eher schwach ausgeprägten linguistischen Kompetenzen in dieser Gruppe: Aufgrund der schwächer ausgeprägten Kompetenzen, z. B. im Bereich des Wortschatzes, dürften L2_{Türk}-Schülerinnen und -Schülern weniger gut als ihre Peers imstande sein, auch auf linguistisches Wissen zurückzugreifen, wenn sie Texte lesen. Dies würde beispielsweise dazu führen, dass sich aufgrund eines geringer ausgeprägten Sichtwortschatzes bei den L2-Jugendlichen eine gegenüber L1-Jugendlichen geringere Lesegeschwindigkeit ergibt (Geva & Farnia, 2012; Wolf & Katzir-Cohen, 2001).

Insgesamt lässt sich festhalten, dass sich ein theoretisch postulierter besonders enger Zusammenhang zwischen dem Hör- und dem Leseverstehen bei L2-Jugendlichen nicht bestätigen lässt. Allerdings ist das Leseverstehen bei allen Jugendlichen entsprechend der Annahmen des *Simple View of Reading* eng mit dem Hörverstehen von Schülerinnen und Schülern verbunden. Schwächen im Leseverstehen von L2-Schülerinnen und -Schülern könnten sich demnach direkt aus deren oftmals im Vergleich zu L1-Jugendlichen ebenfalls geringer ausgeprägten Hörverstehensfähigkeiten ergeben. Die vorliegende Studie verdeutlicht damit die Relevanz einer Betrachtung des zweitsprachlichen Hörverstehens für den Erwerb weiterer sprachlicher und schulischer Fähigkeiten von L2-Jugendlichen.

9.2 **Alltagssprachliches und bildungssprachliches Hörverstehen in einer Erst- bzw. Zweitsprache**

Die vorliegende Studie liefert erste Hinweise darauf, dass sich Bildungs- und Alltagssprache nicht nur in Bezug auf die zugrunde liegenden Charakteristika voneinander abgrenzen lassen, sondern auch mit unterschiedlichen Facetten sprachlicher Kompetenzen assoziiert sind. Die Zweidimensionalität des Hörverstehens und der Wortschatzkenntnisse entspricht dabei theoretischen Überlegungen zur Bildungssprache, von der angenommen wird, dass ihre Verarbeitung schülerseitig andere Fähigkeiten in Hinblick auf die Beherrschung von Genres, Strategien und linguistischen Merkmalen erfordert als die Verarbeitung alltagssprachlichen Inputs (z. B. Nagy & Townsend, 2012; Richards, 1983). Gleichzeitig finden sich sowohl für das Hörverstehen als auch für die Wortschatzkenntnisse hohe Interkorrelationen zwischen den beiden Facetten sprachlichen Verstehens. Dieser Zusammenhang entspricht der Vorstellung, dass beide Register eine Reihe von Prädiktoren gemeinsam haben (Scarcella, 2003) und sich auf einem Kontinuum sprachlicher Fähigkeiten verorten lassen (Koch & Oesterreicher, 1985; Tannen, 1980). Der Befund, dass bildungssprachliche und alltagssprachliche Fähigkeiten voneinander abgrenzbar sind, aber hoch miteinander zusammenhängen, entspricht somit den theoretisch postulierten Annahmen zum Konzept der Bildungssprache.

Während die Überprüfung der Dimensionalität sprachlicher Kompetenzen mittels einer konfirmatorischen Faktorenanalyse für das Hörverstehen die theoretisch postulierte Zweidimensionalität empirisch stützt, liefert die zusätzliche Validierung über Außenkriterien in Teilstudie 3 nur eingeschränkt Hinweise auf die Gültigkeit der entsprechenden Annahmen. So findet sich, wie erwartet, zwar ein stärkerer Zusammenhang zwischen dem bildungssprachlichen Hörverstehen und dem Wortschatz von Schülerinnen und Schülern als zwischen dem alltagssprachlichen Hörverstehen und dem Wortschatz, die entsprechende Differenz ist jedoch nicht statistisch signifikant. Für den Zusammenhang zwischen morpho-syntaktischen Kenntnissen und dem Hörverstehen findet sich erwartungskonträr sogar ein höherer Zusammenhang zum alltagssprachlichen als zum bildungssprachlichen Hörverstehen. Diese Befunde sprechen jedoch nicht notwendigerweise gegen die Annahme der postulierten Zusammenhänge zwischen Wortschatz bzw. morpho-syntaktischen Kenntnissen und bildungs- bzw. alltagssprachlichem Hörverstehen. Vielmehr könnten die unerwarteten Ergebnismuster ein Resultat der eingesetzten Instrumente sein: Das zur Erfassung der Wortschatzkenntnisse verwendete Verfahren erfordert sowohl die Kenntnis bildungssprachlicher als auch alltagssprachlicher Wörter. Der Test zur Erfassung morpho-syntaktischer Kenntnisse hingegen wurde als sehr basales Verfahren konzipiert und enthält entsprechend nur alltagssprachliche Merkmale. Da

die Verarbeitung bildungssprachlicher und alltagssprachlicher Hörverstehenstexte die Beherrschung unterschiedlicher linguistischer Merkmale erfordert (Nagy & Townsend, 2012; Richards, 1983), kann vermutet werden, dass sich Zusammenhänge vor allem zwischen bildungssprachlichen Prädiktoren und dem bildungssprachlichen Hörverstehen bzw. alltagssprachlichen Prädiktoren und dem alltagssprachlichen Hörverstehen zeigen sollten. Aus diesem Grund wurde in der vorliegenden Studie auch der Frage nach differenziellen Zusammenhängen zwischen unterschiedlichen Facetten des Hörverstehens mit den Außenkriterien *morpho-syntaktische Kenntnisse* und *bildungssprachlicher bzw. alltagssprachlicher Wortschatz* nachgegangen.

Für den Zusammenhang der morpho-syntaktischen Kenntnisse mit dem Hörverstehen in Teilstudie 3 zeigen sich dabei Befunde, die für die Annahme eines spezifischen Zusammenhangs zwischen alltags- bzw. bildungssprachlichen Fähigkeiten sprechen: Der rein alltagssprachlich konstruierte Test der morpho-syntaktischen Fähigkeiten korreliert stärker mit dem alltagssprachlichen als mit dem bildungssprachlichen Hörverstehen. In Hinblick auf den Wortschatz zeigen die Ergebnisse aus Teilstudie 4 tatsächlich einen stärkeren Zusammenhang des bildungssprachlichen Wortschatzes mit dem bildungssprachlichen als mit dem alltagssprachlichen Hörverstehen. Für den alltagssprachlichen Wortschatz findet sich allerdings ebenfalls eine höhere Korrelation zum bildungssprachlichen Hörverstehen als zum alltagssprachlichen Hörverstehen. Da die übrigen Zusammenhänge das erwartete Muster aufweisen, erscheint es unwahrscheinlich, dass dieser Befund gegen die Zweidimensionalität alltags- bzw. bildungssprachlicher Kompetenzen spricht. Vielmehr ist denkbar, dass zum Verstehen bildungssprachlicher Hörtexte aufgrund der höheren Dekontextualisierung breitere Wortschatzkenntnisse erforderlich sind als beim Verstehen alltagssprachlicher Texte (vgl. Staehr, 2009). Entsprechend könnte es notwendig sein, für das Verstehen bildungssprachlicher Hörtexte auch auf den alltagssprachlichen Wortschatz zurückzugreifen, während bildungssprachliche Wortschatzkenntnisse für die Verarbeitung alltagssprachlicher Hörverstehenstexte nicht benötigt werden.

Für die Ausprägung alltagssprachlicher und bildungssprachlicher Kompetenzen in einer Erst- bzw. Zweitsprache zeigt sich erwartungsgemäß, dass bildungssprachliche Hörverstehenskompetenzen sowohl bei L1- als auch bei L2-Schülerinnen und -Schülern geringer ausgeprägt sind als alltagssprachliche Hörverstehenskompetenzen. Dieser Befund entspricht der Erwartung, nach der Bildungssprache komplexer und schwieriger zu erwerben ist als die im Alltag verwendete Sprache. Darüber hinaus deutet die vorliegende Arbeit darauf hin, dass sich im Bereich des Wortschatzes und des Hörverstehens schwächere alltagssprachliche

Kompetenzen der L2-Jugendlichen gegenüber der L1-Gruppe finden. Für das Hörverstehen ließen sich zusätzlich auch schwächere bildungssprachliche Hörverstehensfähigkeiten der L2-Schülerinnen und -Schüler nachweisen, wobei die entsprechenden Leistungsdifferenzen etwas weniger stark ausgeprägt waren als im alltagssprachlichen Hörverstehen. Besondere Leistungsnachteile der L2-Jugendlichen im Bereich bildungssprachlicher Kompetenzen lassen sich somit nicht bestätigen. Dieses Befundmuster entspricht den Ergebnissen der Studien von Eckhardt (2008) sowie Heppt et al. (2014) mit Lernenden im Grundschulalter, in denen sich ebenfalls keine zusätzlichen Leistungsnachteile von L2-Kindern gegenüber L1-Kindern bei der Bearbeitung bildungssprachlicher Hörverstehensaufgaben fanden. Während die Autorinnen der genannten Studien vermuten, dass sich solche Effekte erst in den höheren Klassenstufen zeigen sollten, wenn zunehmend bildungssprachliche Texte im Unterricht eingesetzt werden, legen die Befunde der vorliegenden Studie nahe, dass auch bei älteren Schülerinnen und Schülern bildungssprachliche Kompetenzen gegenüber L1-Jugendlichen nicht disproportional schwächer ausgeprägt sind als alltagssprachliche Kompetenzen.

Dieser Befund widerspricht in zweifacher Hinsicht zentralen Annahmen zum Konzept der Bildungssprache. Zum einen erscheint die postulierte Dauer bis zur Beherrschung von Alltags- und Bildungssprache in einer Zweitsprache fraglich. So wird bisweilen angenommen, dass Schülerinnen und Schüler fünf bis sieben Jahre für den Erwerb bildungssprachlicher Kompetenzen in einer Zweitsprache benötigen, für den Erwerb alltagssprachlicher Kompetenzen drei bis fünf Jahre (Cummins, 2000). Die Ergebnisse der vorliegenden Arbeit deuten jedoch darauf hin, dass in Hinblick auf beide sprachlichen Register ältere L2-Schülerinnen und -Schüler, die mehrheitlich die gesamte Schullaufbahn in Deutschland durchlaufen und somit mindestens neun Jahre Kontakt zur Zweitsprache hatten, nicht über sprachliche Kompetenzen verfügen, die jenen von L1-Jugendlichen entsprechen.

Zum anderen weisen die vorliegenden Befunde darauf hin, dass der Erwerb bildungssprachlicher Kompetenzen für L2-Jugendliche nicht notwendigerweise mit größeren Schwierigkeiten verbunden sein muss als der Erwerb alltagssprachlicher Kompetenzen. So zeigen die Ergebnisse für den Wortschatz, dass der Erwerb eines dem von L1-Lernenden entsprechenden bildungssprachlichen Vokabulars offensichtlich auch dann gelingen kann, wenn in Bezug auf den alltagssprachlichen Wortschatz Leistungsnachteile der L2-Lernenden bestehen. Für das Hörverstehen ist dieser Befund zwar weniger eindeutig, jedoch zeigen sich auch hier stärker ausgeprägte Leistungsdifferenzen im alltagssprachlichen als im bildungssprachlichen Hörverstehen. Offenbar stellt demnach auch der Erwerb alltagssprachlicher Kompetenzen L2-Schülerinnen und -Schüler vor große Herausforderungen. Eine mögliche Ursache für diesen

Befund liegt im Ausmaß des Sprachkontakts im Sinne einer *time-on-task*, der für die Ausbildung sprachlicher Kompetenzen entscheidend sein könnte. So verwendet ein Großteil der L2-Jugendlichen in den vorliegenden Studien außerhalb des schulischen Kontextes die jeweilige Erstsprache (vgl. Tabelle 12) und verfügt für die außerhalb des schulischen Kontextes verwendete Alltagssprache somit nur über wenig Kontakt zur Zweitsprache. Der Erwerb bildungssprachlicher Kompetenzen hingegen verläuft sowohl für L1- als auch für L2-Jugendliche auf Deutsch, so dass die Lerngelegenheiten im Sinne einer *time-on-task* für den Erwerb eines bildungssprachlichen Wortschatzes bzw. bildungssprachlicher Hörverstehensfähigkeiten in beiden Gruppen vergleichbar sind. Es erscheint daher denkbar, dass die ausgeprägteren Leistungsnachteile der L2-Jugendlichen in Hinblick auf alltagssprachliche Zweitsprachkompetenzen auf einem Mangel an Lerngelegenheiten für den Erwerb dieses Registers in der Zweitsprache zurückzuführen sein könnten.

Diese Erklärung wäre konform mit Befunden aus Studien von Bialystok et al. (2010) und Engel de Abreu et al. (2013), in denen sich ebenfalls stärkere Leistungsdifferenzen zwischen L1- und L2-Lernenden für alltagssprachliche als für bildungssprachliche Kompetenzen fanden. Für den Wortschatz sind diese Befunde konform mit Annahmen, die davon ausgehen, dass zumindest der fachspezifische bildungssprachliche Wortschatz in der Regel in der Schule und somit in der Instruktionssprache vermittelt wird (Ahrenholz, 2013; Köhne, Kronenwerth, Redder, Schuth & Weinert, 2015), so dass L1- und L2-Schülerinnen und -Schüler über dieselben Lerngelegenheiten zum Erwerb dieses spezifischen Registers verfügen. Da weitere für das bildungssprachliche Hörverstehen bedeutsame bildungssprachliche Fähigkeiten demgegenüber im Unterricht nur selten explizit vermittelt werden, kann davon ausgegangen werden, dass diese größtenteils im Alltag der Schülerinnen und Schüler erworben werden müssen (vgl. Haag et al., 2013). Für L2-Schülerinnen und -Schüler, die in ihrem Alltag weniger häufig die Instruktionssprache nutzen, waren daher besondere Leistungsnachteile im Bereich bildungssprachlichen Hörverstehens gegenüber ihren L1-Peers vermutet worden, die sich allerdings nicht bestätigen ließen. Eine mögliche Ursache für das Ausbleiben der erwarteten besonderen Leistungsnachteile der L2-Jugendlichen im bildungssprachlichen Hörverstehen könnte im geringen sozioökonomischen Status der Gesamtstichprobe liegen. So scheint der Erwerb von Bildungssprache auch für L1-Lernende mit einem geringen sozioökonomischen Status eine bedeutsame Hürde darzustellen, da sich in ihrem soziokulturellen Umfeld häufig wenig Gelegenheiten zur Aneignung dieses Registers bieten (Corson, 1997; Curençon & Justice, 2004). Es ist daher denkbar, dass L2- und L1-Lernende in der vorliegenden Studie unabhängig von ihrem Sprachhintergrund nicht über ausreichende Lerngelegenheiten in Bezug auf Bildungs-

sprache verfügten und entsprechend auch die Sprachgruppenunterschiede im bildungssprachlichen Hörverstehen weniger stark ausgeprägt sind als im alltagssprachlichen Hörverstehen.

Eine weitere mögliche Erklärung für das Ausbleiben disproportionaler Schwierigkeiten der L2-Schülerinnen und -Schüler im Erwerb bildungssprachlicher Kompetenzen liegt in der Annahme, dass eine klare Trennung zwischen alltagssprachlichen Fähigkeiten im Sinne von BICS und bildungssprachlichen Fähigkeiten im Sinne von CALP, wie sie Cummins (2000) vornimmt, nicht ohne weiteres haltbar ist. Eine solche Sichtweise entspräche der Kritik Scarcellas (2003) an einer dichotomen Konzeptualisierung sprachlicher Fähigkeiten. Scarcella (2003) kritisiert dabei vor allem, eine solche Trennung sei “not useful for understanding the complexities of academic English or the multiple variables affecting its development” (S. 5). Ihrer Sichtweise nach bilden sowohl Bildungssprache als auch die im Alltag verwendete Sprache komplexe Register, die eine Reihe von Prädiktoren teilen. Die Annahme, Bildungssprache sei per se schwieriger zu verarbeiten und inhaltlich wie formal komplexer als Alltagssprache, sei vor diesem Hintergrund unzutreffend. Vielmehr nimmt Scarcella (2003) an, dass alltägliche Sprache und Bildungssprache eine Reihe von Überlappungen aufweisen, so dass einige Aspekte von Bildungssprache relativ früh erworben werden könnten, während einige Aspekte von Alltagssprache schwieriger zu beherrschen seien. Kontrolliert man in der vorliegenden Studie bei der Überprüfung von Unterschieden in Bezug auf das bildungssprachliche Hörverstehen für das alltagssprachliche Hörverstehen, so zeigen sich keine Unterschiede mehr zwischen L1- und L2-Schülerinnen und -Schülern ($Wald-\chi^2 = 0.55, df = 1, p = .46$). Ein ähnlicher Befund findet sich auch für Unterschiede im alltagssprachlichen Hörverstehen: Wird hier für das bildungssprachliche Hörverstehen kontrolliert, finden sich ebenfalls keine Unterschiede mehr im alltagssprachlichen Hörverstehen zwischen L1- und L2-Schülerinnen und -Schülern. Dieses Befundmuster spricht für die Annahme Scarcellas (2003), dass beide sprachliche Register jeweils für L2-Lernende mehr oder wenig schwierig zu beherrschende Merkmale aufweisen und Bildungssprache daher nicht notwendigerweise schwieriger zu beherrschen ist als Alltagssprache.

Insgesamt muss zudem konstatiert werden, dass die Evidenz für eine Differenzierung bildungssprachlicher und alltagssprachlicher Fähigkeiten in der vorliegenden Studie nicht eindeutig ist. Sowohl die Ergebnisse der faktorenanalytischen Untersuchungen von Hörverstehen und Wortschatz als auch die zusätzliche Validierung über die Außenkriterien liefern zwar erste Hinweise auf die Abgrenzbarkeit alltagssprachlicher und bildungssprachlicher Fähigkeiten, lassen aber aufgrund der teilweise widersprüchlichen Befunde keinen eindeutigen Schluss hinsichtlich der Differenzierung alltagssprachlicher und bildungssprachlicher Kompe-

tenzen zu. Während Befunde aus dem englischsprachigen Raum die Unterscheidung alltags-sprachlicher und bildungssprachlicher Merkmale in Texten bereits empirisch gestützt haben (vgl. z. B. Schleppegrell, 2004), kann für den Bereich der entsprechenden sprachlichen Kompetenzen auf Basis der vorliegenden Befunde lediglich konstatiert werden, dass beide Facetten sprachlichen Verstehens zumindest deutliche Überlappungen aufzuweisen scheinen. Auch zeigt sich, dass die Annahme, der Erwerb bildungssprachlicher Fähigkeiten sei für L2-Lernende aufgrund der erhöhten Komplexität dieses Registers per se besonders erschwert, keine Unterstützung erfährt. Wie die Ergebnisse aus Teilstudie 4 nahelegen, scheint vielmehr vor allem der Kontakt zu den jeweiligen sprachlichen Registern in der Zweitsprache entscheidend für die Ausbildung alltags-sprachlicher bzw. bildungssprachlicher Kompetenzen zu sein.

9.3 Grenzen der vorliegenden Arbeit

Bei der Interpretation der vorliegenden Befunde müssen einige Einschränkungen der Studien berücksichtigt werden, die in den einzelnen Manuskripten diskutiert werden (vgl. Anhang A bis C). Im Folgenden werden jene Limitationen dargestellt, die alle Teilstudien betreffen. Ein wesentlicher Aspekt ist dabei das querschnittliche Design der DELKO-Studie, die aufgrund ihrer korrelativen Anlage keine kausalen Interpretationen der berichteten Effekte zulässt. Somit bleibt unklar, inwiefern die berichteten Effekte sprachlicher Prädiktoren tatsächlich ursächlich für die Ausprägung von Hörverstehenskompetenzen sind. So wäre beispielsweise auf Basis bisheriger Forschung denkbar, dass zwischen dem Erwerb sprachlicher Prädiktoren sowie den Textverstehenskompetenzen komplexe Wechselwirkungen bestehen (z. B. Ceci & Papierno, 2005; Uccelli et al., 2015), die durch die strukturanalytischen Modellierungen in der vorliegenden Arbeit nicht abgebildet werden. Eine solche Klärung der entwicklungspsychologischen Zusammenhänge zwischen linguistischen Prädiktoren des Hörverstehens und den Textverstehenskompetenzen von Schülerinnen und Schülern war allerdings nicht Ziel der Studie. Vielmehr dienen die vorgestellten Teilstudien einer vertiefenden Analyse der Kompetenzprofile von Jugendlichen in einer Erst- bzw. Zweitsprache, mittels derer Hinweise auf spezifische Schwierigkeiten einzelner Schülergruppen und damit verbunden auf Ansatzpunkte für eine gezielte Förderung gewonnen werden können. Die Überprüfung der einzelnen Fragestellungen erfolgte dabei theoriegeleitet über regressionsanalytische Verfahren, innerhalb derer gerichtete Effekte postuliert wurden. Vorteil dieses Vorgehens im Rahmen querschnittlicher Untersuchungsdesigns ist, dass es die simultane Modellierung der Effekte mehrerer konkurrierender Faktoren erlaubt und somit eine valide Schätzung der jeweiligen Einzeleffekte ermöglicht (Shadish, Cook & Campbell, 2002). In der Folge kann durch die Berücksichtigung

weiterer theoretisch plausibler Einflussfaktoren die Anzahl kausaler Erklärungsalternativen reduziert werden, um auf diese Weise zu einer aussagekräftigen Schätzung der Einzeleffekte zu gelangen (vgl. Shadish et al., 2002). Dennoch bleiben die im Rahmen der Studie ermittelten Effekte mehrdeutig: Zwar liefern längsschnittliche Studien, beispielsweise zu Effekten des Wortschatzes auf die Hörverstehensleistung, empirische Unterstützung für die Annahme gerichteter Effekte einzelner sprachlicher Prädiktoren auf das Hörverstehen (z. B. Droop & Verhoeven, 2003). Allerdings existieren auch empirische Befunde, die bidirektionale Zusammenhänge, d. h. zusätzlich auch Effekte der Hörverstehenskompetenzen auf die Wortschatzentwicklung, nahelegen (z. B. Richards, 2005; Vidal, 2003). Vor diesem Hintergrund könnte angenommen werden, dass die gerichteten Zusammenhänge zwischen sprachlichen Prädiktoren und der Textverstehensleistung in den vorliegenden Studien eher überschätzt werden.

Eine weitere Einschränkung der Studien betrifft die Generalisierbarkeit der Befunde. Der Fokus der den Analysen zugrundeliegenden DELKO-Studie lag auf der Untersuchung von Prädiktoren der Leseleistung bei leseschwachen Jugendlichen, die vornehmlich aus Familien mit einem geringen sozioökonomischen Hintergrund stammten. Die berichteten Befunde sind somit nur für diese spezifische Population von Schülerinnen und Schülern interpretierbar und lassen sich nicht ohne weiteres auf andere Stichproben übertragen. So wäre auf Basis des engen Zusammenhangs zwischen dem sozioökonomischen Hintergrund und der Entwicklung sprachlicher Fähigkeiten (vgl. z. B. Hoff, 2013) beispielsweise denkbar, dass leseschwache Schülerinnen und Schüler aus Familien mit einem günstigeren sozioökonomischen Hintergrund andere Kompetenzprofile aufweisen als die hier betrachtete Stichprobe. Auch hinsichtlich der Zusammenhänge zwischen einzelnen Prädiktoren des Textverstehens könnte vermutet werden, dass sich für andere Populationen andere Ergebnismuster zeigen. Auf Grundlage des *Simple View of Reading* beispielsweise kann angenommen werden, dass der Zusammenhang zwischen dem Hörverstehen und den Lesekompetenzen bei jüngeren Schülerinnen und Schülern weniger eng ausfällt als in dem hier betrachteten Sample von Jugendlichen (Hoover & Gough, 1990; Rost & Hartmann, 1992).

Für die Analysen zur Differenzierung bildungssprachlicher und alltagsprachlicher Kompetenzen der Schülerinnen und Schüler ist zudem zu beachten, dass die Stichprobe der DELKO-Studie aus Jugendlichen zusammengesetzt war, die die niedrigeren Bildungsgänge des deutschen Schulsystems besuchten. Obgleich aus theoretischer Sicht davon ausgegangen werden kann, dass in den höheren Klassenstufen eine zunehmende Ausdifferenzierung bildungssprachlicher Kompetenzen stattfindet, bleibt bislang unklar, inwiefern dies für die ver-

schiedenen schulischen Bildungsgänge gleichermaßen der Fall ist. Befunde internationaler Studien legen nahe, dass sich bildungssprachliche Merkmale häufiger in den höheren Bildungsgängen des Schulsystems finden als in den niedrigen Bildungsgängen und Schülerinnen und Schüler somit über differenzielle Lerngelegenheiten zum Erwerb bildungssprachlicher Kompetenzen verfügen (z. B. Harklau, 1994). Erste Analysen eines Forschungsprojekts der Universität Tübingen implizieren, dass dies zumindest für die in deutschen Schulen verwendeten Textbücher ebenfalls gelten könnte (Meurers, Berendes, Vajjala, Bryant, 2015). Vor diesem Hintergrund kann vermutet werden, dass sich bildungs- und alltagsprachliche Fähigkeiten in einer Stichprobe von Schülerinnen und Schülern an Gymnasien deutlicher voneinander abgrenzen ließen, da sich für diese Jugendlichen mehr Gelegenheiten zum Erwerb und der Ausdifferenzierung bildungssprachlicher Fähigkeiten ergeben als in der hier betrachteten Stichprobe von Jugendlichen aus Haupt-, Real- und Gesamtschulen.

Die eingeschränkte Übertragbarkeit der berichteten Befunde auf andere Populationen muss allerdings nicht notwendigerweise ein Problem angesehen sein (vgl. Shadish et al., 2002). So war eine breite Generalisierbarkeit der Ergebnisse der DELKO-Studie a priori nicht angedacht. Ziel der Analysen war es vielmehr, Aussagen über spezifische Stärken und Schwächen in den Kompetenzprofilen einer Stichprobe von Schülerinnen und Schülern zu treffen, die als Risikogruppe für die Ausbildung schwacher Textverstehensleistungen gelten können. Trotz der eingeschränkten Übertragbarkeit der Ergebnisse auf andere Schülergruppen sind die Befunde somit von hoher praktischer Relevanz für die Erklärung von Textverstehensschwierigkeiten bei L1- und L2-Jugendlichen in den niedrigeren Bildungsgängen des Schulsystems, in denen Schülerinnen und Schüler mit schwachen Lese- und Hörverstehenskompetenzen häufig überrepräsentiert sind (vgl. Kapitel 1.1).

9.4 Ausblick

9.4.1 Implikationen für die pädagogische Praxis: Förderungen des Hörverstehens

Die vorliegende Arbeit zeigt auf, dass ein wesentlicher Anteil von L2-Schülerinnen und -Schülern auch gegen Ende ihrer Schullaufbahn noch schwächere sprachliche Kompetenzen aufweist als L1-Jugendliche. Die Befunde deuten somit darauf hin, dass zumindest für L2-Jugendliche mit türkischer Herkunftssprache die momentan implementierten Ansätze zur Sprachförderung nicht ausreichend sind, um die zweitsprachlichen Kompetenzen dieser Schülerinnen und Schüler auf ein Niveau zu heben, das dem ihrer monolingualen Peers entspricht (vgl. auch Gogolin, 2005; Söhn, 2005). Obwohl die Lage für die L2-Jugendlichen mit einer

anderen Herkunftssprache als Türkisch tendenziell etwas günstiger ist, kann auch die Situation dieser Schülerinnen und Schüler nicht als unproblematisch bewertet werden. Zwar zeigen sich für diese Gruppe nur für den Wortschatz noch Leistungsunterschiede zu L1-Jugendlichen, wenn für den sozioökonomischen Hintergrund kontrolliert wird. Damit erzielen L2_{Andere}-Jugendliche allerdings nur ähnliche Leistungen wie L1-Jugendliche aus sozial schwachen Familien, die ihrerseits gegenüber Jugendlichen aus sozioökonomisch besser gestellten Familien Benachteiligungen aufweisen (vgl. auch Stanat, Weirich & Radmann, 2012). Um diese Disparitäten im Erwerb sprachlicher Kompetenzen zwischen L1- und L2-Jugendlichen zu reduzieren, erscheint es notwendig, über die bestehenden Ansätze hinaus institutionelle Maßnahmen zu entwickeln, die für L2-Schülerinnen und -Schüler ausreichend Lerngelegenheiten bieten, um zweitsprachliche Kompetenzen zu erwerben, die denen monolingualer Schülerinnen und Schüler entsprechen.

Gegenwärtig findet sich in Deutschland eine erhebliche Heterogenität der Förderaktivitäten im Bereich sprachlicher Kompetenzen, die vorwiegend auf den Elementar- oder Grundschulbereich fokussieren (vgl. Stanat, Weirich & Radmann, 2012). Obwohl insgesamt ein Großteil der Schülerinnen und Schüler, die sprachliche Mindestkompetenzen nicht erreichen, gegenwärtig zusätzliche Sprachförderung erhält, scheint es angesichts der Befunde der vorliegenden Studie notwendig, die Aktivitäten im Bereich der Sprachförderung kontinuierlich zu optimieren. Dabei sollte neben der Förderung von Leseverstehens-fähigkeiten auch die Ausbildung von Hörverstehensfähigkeiten berücksichtigt werden, die bislang nur selten im Fokus institutioneller Fördermaßnahmen stehen (Imhof, 2010; Vandergrift 2007). Dieser bislang fehlende Fokus auf die Förderung von Hörverstehenskompetenzen ist problematisch, da das Hörverstehen eine wesentliche Determinante der Lesekompetenz sowie des weiteren schulischen Kompetenzerwerbs bildet. Auch in der vorliegenden Studie findet sich ein enger Zusammenhang zwischen Hör- und Leseverstehenskompetenzen bei älteren Schülerinnen und Schülern. Für die hier betrachtete Stichprobe von Jugendlichen mit schwachen Lesekompetenzen sollten daher die Hörverstehensfähigkeiten sowohl bei der Diagnose schwacher Leseleistungen, als auch bei deren Prävention zukünftig berücksichtigt werden (vgl. Language and Reading Research Consortium, 2015).

Zur Förderung von Vorläuferfähigkeiten des Textverstehens, wie etwa dem Wortschatz, der phonologischen Bewusstheit oder morpho-syntaktischen Kenntnissen, existiert bereits eine Reihe von Ansätzen (vgl. Marx, Wolf, Paetsch, Darsow & Felbrich, 2014). Bislang sind allerdings nur wenige dieser Förderansätze wissenschaftlich evaluiert worden. Insbesondere zur Effektivität sprachlicher Förderung von L2-Kindern im schulischen Bereich fehlen derzeit

kontrollierte Interventionsstudien, so dass es kaum möglich ist, verlässliche Aussagen über die „best practice“ bei der Förderung sprachlicher Kompetenzen von Kindern mit Zuwanderungshintergrund zu machen. Aufgrund der Komplexität von Textverstehensprozessen kann jedoch vermutet werden, dass eine reine Förderung von Vorläuferfähigkeiten des Hör- oder Leseverstehens nicht in allen Fällen ausreichend sein dürfte, um Textverstehensfähigkeiten zu verbessern, sondern dass diese selbst ebenfalls einer Förderung bedürfen (Hogan, Adlof, & Alonzo, 2014). Allerdings existieren neben Ansätzen zur Förderung des Hörverstehens im Fremdsprachenunterricht bislang kaum wissenschaftliche Erkenntnisse zu effektiven Maßnahmen zur Verbesserung des Hörverstehens.

Eine Reihe von theoretischen Überlegungen zu möglichen Förderansätzen im Bereich des Hörverstehens liegt bereits vor. Dabei lassen sich bottom-up basierte von top-down basierten Ansätzen unterscheiden, die jeweils unterschiedliche Teilkompetenzen des Hörverstehens fokussieren: Während top-down Ansätze Übungen beinhalten, die beispielsweise die Fähigkeiten fördern sollen, anhand von Schlüsselwörtern ein Diskursschema des Hörtextes zu entwickeln oder die Rolle einzelner Protagonisten des Hörtextes zu inferieren (vgl. Imhof, 2010), enthalten bottom-up Ansätze eher Übungen, anhand derer Hörende beispielsweise lernen sollen, Wort- und Satzgrenzen oder Schlüsselwörter zu erkennen (z. B. Gilakjani & Ahmadi, 2011). Einen integrierten Ansatz der Förderung von Zuhörkompetenzen in der Zweitsprache beschreibt Vandergrift (2007). Neben top-down- und bottom-up- Strategien soll in diesem Ansatz auch die metakognitive Bewusstheit für die beim Zuhören notwendigen kognitiven Prozesse gestärkt werden. Hörerinnen und Hörer werden dabei geschult, sowohl vor, als auch während und nach dem Zuhören den Hörprozess zu planen, zu überwachen und zu evaluieren. Erste Überprüfungen dieses integrierten Ansatzes zeigten, dass ältere Lernende, die Französisch als Zweitsprache erwarben, von einer entsprechenden Intervention tatsächlich profitierten (vgl. Vandergrift, 2007). Die Überprüfung für weitere Sprach- und Altersgruppen steht allerdings noch aus, so dass auch für diesen integrierten Ansatz derzeit keine gesicherten Aussagen zur Wirksamkeit für die Förderung von Hörverstehensfähigkeiten möglich sind.

Die Befunde der vorliegenden Studie deuten darauf hin, dass L2-Schülerinnen und -Schüler nicht nur im Bereich des alltäglichen Hörverstehens, sondern auch im bildungssprachlichen Hörverstehen schwächere Leistungen aufweisen als ihre monolingualen Peers. Vor diesem Hintergrund erscheint es notwendig, auch bildungssprachlichen Fähigkeiten bei der Förderung zweitsprachlicher Fähigkeiten Aufmerksamkeit zukommen zu lassen. Auch für diesen Bereich liegen bislang jedoch nur wenige empirische Studien vor, die die Wirksamkeit entsprechender Förderansätze belegen könnten. Programme zur Förderung des bildungs-

sprachlichen Wortschatzes haben sich im englischen Sprachraum allerdings bereits als erfolgreich gezeigt (Nagy & Townsend, 2012). Eine quasi-experimentelle Studie mit Schülerinnen und Schülern der 6. bis 8. Klassenstufe beispielsweise erbrachte Hinweise darauf, dass ein Training zur Verbesserung des bildungssprachlichen Wortschatzes einen positiven Effekt mittlerer Effektstärke auf die entsprechenden Kenntnisse hatte (Snow, Lawrence & White, 2009). L2-Schülerinnen und -Schüler scheinen von diesem Training stärker zu profitieren als L1-Jugendliche, so dass dieser Ansatz besonders geeignet scheint, sprachbedingte Disparitäten im bildungssprachlichen Wortschatz zu reduzieren. Allerdings existieren bislang kaum empirische Befunde, die Transfereffekte auf das Textverstehen nachweisen würden (Nagy & Townsend, 2012). Es bleibt somit unklar, inwiefern diese Förderprogramme geeignet sind, bildungssprachliche Kompetenzen in einer Zweitsprache nachhaltig zu stärken.

Vielversprechender scheint gegenüber Ansätzen, die isolierte Aspekte von Bildungssprache fokussieren, eine fortlaufende und konsequente Vermittlung bildungssprachlicher Fähigkeiten im Sinne einer *durchgängigen Sprachbildung* zu sein (Gogolin & Lange, 2011). Dabei wird die Notwendigkeit der Schaffung von Lernsituationen betont, in denen Schülerinnen und Schüler explizit und systematisch an Bildungssprache herangeführt werden. Hierfür sollen Lehrkräfte aller Fachbereiche im Unterricht entsprechende Lerngelegenheiten schaffen, etwa indem bewusst bildungssprachliche Mittel eingesetzt und thematisiert werden oder die Schülerinnen und Schüler Gelegenheit erhalten, bildungssprachliche Fähigkeiten aktiv einzusetzen. Sprache wird somit auch in den nichtsprachlichen Fächern zum Lerngegenstand (vgl. Gogolin & Lange, 2011). Insbesondere für das Hörverstehen bleibt allerdings unklar, inwiefern durchgängige Sprachbildung geeignet ist, L2-Schülerinnen und -Schüler zu befähigen, Kompetenzen zu entwickeln, die denen ihrer monolingualen Peers entsprechen.

Insgesamt lässt sich festhalten, dass bislang noch wenig über den Erfolg von Maßnahmen zur Förderung von Hörverstehensfähigkeiten, sowohl in einer Erst- als auch in einer Zweitsprache, bekannt ist. Konkrete Empfehlungen zu geeigneten Förderprogrammen lassen sich vor diesem Hintergrund derzeit nicht formulieren. Auf Basis der Befunde der vorliegenden Studie sowie theoretischer Überlegungen und empirischer Ergebnisse zur Förderung zweitsprachlicher Kompetenzen lässt sich allerdings mutmaßen, dass L2-Schülerinnen und -Schüler über die Teilnahme am Unterricht hinaus Unterstützung bei der Entwicklung ihrer zweitsprachlichen Hörverstehensfähigkeiten bzw. deren linguistischer Prädiktoren benötigen. Da in der vorliegenden Studie die Effekte linguistischer Prädiktoren auf das Textverstehen in einer Erst- bzw. Zweitsprache ähnlich stark ausgeprägt waren, ist davon auszugehen, dass L2-Schülerinnen und -Schüler dabei nicht grundsätzlich andere Fördermaßnahmen benö-

tigen als L1-Schülerinnen und -Schüler mit schwachen sprachlichen Fähigkeiten. Vielmehr sollte es Ziel einer Förderung sein, eventuell vorhandenen Schwächen im Bereich sprachlicher Prädiktoren des Textverstehens bei L1- und L2-Schülerinnen und -Schülern gleichermaßen zu fördern. Dabei scheint es bisherigen Befunden nach nicht ausreichend, nur einzelne Prädiktoren des Hörverstehens zu fokussieren, vielmehr sollten Förderansätze den gesamten Hörverstehensprozess betrachten und sowohl bottom-up als auch top-down-Prozesse thematisieren. Eine solche Förderung sollte nicht nur bildungssprachliche Aspekte ansprechen, von denen angenommen wird, dass sie für L2-Lernende besonders schwierig zu erwerben sind. Vielmehr sollten auch alltagssprachliche Kompetenzen der Schülerinnen und Schüler nicht als gegeben vorausgesetzt, sondern überprüft und bei Bedarf gefördert werden. Die bloße Teilnahme am zweitsprachlichen Unterricht sowie eine gegebenenfalls erfolgende zusätzliche Förderung im Förderunterricht oder ähnlichen, an Schulen bereits bestehenden Angeboten scheint dafür nicht ausreichend zu sein. Vor diesem Hintergrund ist es dringend notwendig, Sprachförderansätze für L2-Schülerinnen und -Schüler zu entwickeln und mit Hilfe von kontrollierten Evaluationsstudien auf ihre Wirksamkeit zu überprüfen.

9.4.2 Implikationen für die zukünftige Forschung

Die durchgeführten Studien liefern erste Hinweise auf Schwierigkeiten beim zweitsprachlichen Hörverstehen, die sich aus der unterschiedlichen Ausprägung von bottom-up Prädiktoren des Hörverstehens bei L1- und L2-Jugendlichen ergeben. Fokussiert wurden die unteren Verarbeitungsebenen der Textverstehensmodelle von Anderson (1995) bzw. Kintsch & van Dijk (1978), da häufig angenommen wird, dass L2-Lernende mit schwachen zweitsprachlichen Kompetenzen während des Hörverstehensprozesses verstärkt auf bottom-up Informationen zurückgreifen (z. B. Lynch, 1998; Rost, 2011, Vandergrift, 2003). Zukünftige Studien sollten diese Arbeit erweitern und auch die Bedeutung von top-down Informationen, wie etwa dem Einsatz metakognitiver Strategien oder dem Einbezug von Vorwissen im erst- bzw. zweitsprachlichen Hörverstehen vergleichend untersuchen. Es liegen zwar bereits erste Studien zur Nutzung top-down basierter Informationen beim zweitsprachlichen Hörverstehen vor, allerdings wird ein ausschließlicher Fokus auf die höheren Textverarbeitungsebenen der Komplexität des Hörverstehensprozesses nicht gerecht, für dessen Gelingen eine Integration von bottom-up und top-down verlaufenden Prozessen erforderlich ist. Eine integrative Betrachtung des zweitsprachlichen Hörverstehens und der auf den unterschiedlichen Verarbeitungsebenen auftretenden Probleme kann somit als eine Hauptaufgabe der Forschung im Bereich zweitsprachlichen Hörverstehens begriffen werden (Batova, 2013). Da angenommen wird, dass die

Integration von bottom-up und top-down verlaufenden Verarbeitungsprozessen auch vom erreichten Kompetenzniveau in einer Zweitsprache abhängt (vgl. Kapitel 2.5), sollten entsprechende Studien möglichst längsschnittlich angelegt sein und Stichproben aus L2-Lernenden mit unterschiedlich ausgeprägten Zweitsprachkenntnissen rekrutieren. Ein solches Vorgehen würde es erlauben, etwaige spezifische Schwierigkeiten in der Integration von bottom-up und top-down Informationen beim zweitsprachlichen Hörverstehen differenzierter zu bestimmen und somit weitere Ansatzpunkte für eine gezielte Förderung der Hörverstehenskompetenzen bei L2-Schülerinnen und -Schülern zu identifizieren.

Darüber hinaus wären weitere empirische Befunde zur Abgrenzbarkeit bildungssprachlicher und alltagssprachlicher Hörverstehenskompetenzen in einer Erst- und Zweitsprache wünschenswert. Die vorliegenden Ergebnisse liefern zwar insgesamt erste Hinweise auf die Gültigkeit einer solchen Trennung von Hörverstehensfacetten, allerdings sind die Befunde zur Ausprägung der entsprechenden Kompetenzen nicht vollumfänglich im Einklang mit theoretischen Annahmen zum Konzept der Bildungssprache. Zukünftige Studien sollten daher die Struktur des Hörverstehens in Hinblick auf den Erwerbskontext vertiefend untersuchen. Die vorliegenden Befunde legen nahe, dass entsprechende Studien dabei neben der Komplexität bildungssprachlichen Hörverstehens vor allem die Lerngelegenheiten zum Erwerb zweitsprachlicher Kompetenzen, auch in Hinblick auf das Ausmaß des Kontakts zu Zweitsprache, fokussieren sollten, die für alltagssprachliche und bildungssprachliche Kompetenzen häufig differenziell ausgeprägt sind. Wünschenswert wären vor allem längsschnittlich angelegte Studien, die die Ausdifferenzierung der beiden sprachlichen Facetten unter Berücksichtigung der differenziellen Lerngelegenheiten in einer Erst- und Zweitsprache vergleichend untersuchen.

Angesichts der besonders schwachen sprachlichen Kompetenzen der L2_{Türk}-Schülerinnen und -Schüler erscheint es notwendig, Bedingungsfaktoren des Zweitspracherwerbs in dieser Gruppe vertieft zu untersuchen. Da die vorliegenden Befunde darauf hindeuten, dass das Ausmaß des Kontakts zur Zweitsprache eine bedeutsame Rolle für den Erwerb von Textverstehensfähigkeiten bzw. deren Prädiktoren spielt, wäre wünschenswert, dass in zukünftigen Studien, in denen zweitsprachliche Kompetenzen von L2-Schülerinnen und -Schülern erfasst werden, neben der Familiensprache (vgl. Haag et al, 2012; Stanat, Rauch & Segeritz, 2010) auch die im außerfamiliären sozialen Alltag, wie etwa im Wohnumfeld genutzte Sprache berücksichtigt würde. Auf diese Weise ließe sich das Ausmaß des Kontakts zur Zweitsprache, der als wesentliche Grundlage für den Spracherwerb betrachtet

werden kann (vgl. Esser, 2006; Gathercole, 2002), valider abbilden als durch die ausschließliche Berücksichtigung der in der Familie verwendeten Sprache.

Die vorliegende Studie weist darauf hin, dass es derzeit vielen L2-Jugendlichen am Ende der Schullaufbahn nicht gelingt, Hörverstehenskompetenzen zu erzielen, die denen ihrer L1-Peers entsprechen. Gleichzeitig gilt Hörverstehen als wichtige Voraussetzung für den Bildungserfolg von Schülerinnen und Schülern (vgl. z. B. Imhof, 2010; KMK, 2005) und ist eng mit dem Erwerb weiterer schulischen Kompetenzen wie dem Leseverstehen verbunden. Ziel folgender Analysen sollte es daher sein, die Bedingungsfaktoren für einen erfolgreichen Erwerb zweitsprachlicher Kompetenzen möglichst umfassend zu klären und darauf aufbauend effektive Maßnahmen zur Unterstützung der Hörverstehensfähigkeiten von Schülerinnen und Schülern mit nichtdeutscher Herkunftssprache zu entwickeln.

10. Literatur

- Aarts, R., Demir, S. & Vallen, T. (2011). Characteristics of academic language register occurring in caretaker-child interaction: Development and validation of a coding scheme. *Language Learning*, 61, 1173–1221.
- Ahrenholz, B. (2008). Erstsprache – Zweitsprache – Fremdsprache. In B. Ahrenholz & I. Oomen-Welke (Hrsg.), *Deutsch als Zweitsprache* (S. 3-16). Baltmannsweiler: Schneider Hohengehren.
- Ahrenholz, B. (2013). Sprache im Fachunterricht untersuchen. In C. Röhner & B. Hövelbrinks (Hrsg.), *Fachbezogene Sprachförderung in Deutsch als Zweitsprache. Theoretische Konzepte und empirische Befunde zum Erwerb bildungssprachlicher Kompetenzen* (S. 87-98). Weinheim: Beltz Juventa
- Alexander, P. A. & Jetton, T. L. (2000). Learning from text: A multidimensional and developmental perspective. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson & R. Barr (Hrsg.), *Handbook of reading research* (S. 285–310). Mahwah, NJ: Erlbaum.
- Allison, P. D. (2001). *Missing Data*. Thousand Oaks: Sage Publications.
- Anderson, A. C. & Lynch, T. (1988). *Listening*. New York: Oxford University Press.
- Anderson, J. R. (1995). *Cognitive psychology and its implications*. New York: Freeman.
- Andringa, S., Olsthoorn, N., van Beuningen, C., Schoonen, R., & Hulstijn, J. (2012). Determinants of success in native and non-native listening comprehension: An individual differences approach. *Language Learning*, 62, 49-78.
- Anstrom, K., DiCerbo, P., Butler, F., Katz, A., Millet, J. & Rivera, C. (2010). *A review of the literature on academic English: Implications for K-12 English language learners*. Arlington, VA: The George Washington University Center for Equity and Excellence in Education.
- Asparouhov, T. & Muthén, B. (2010). *Weighted least squares estimation with missing data. Technical appendix*. Los Angeles: Muthén & Muthén.
- August, D. & Shanahan, T. (Hrsg.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum.
- Autorengruppe Bildungsberichterstattung. (2014). *Bildung in Deutschland 2014. Ein indikatorengestützter Bericht mit einer Analyse zur Bildung von Menschen mit Behinderungen*. Bielefeld: Bertelsmann.

- Babayigit, S. (2014). The role of oral language skills in reading and listening comprehension of text: a comparison of monolingual (L1) and bilingual (L2) speakers of English language. *Journal of Research in Reading*, 37, 22–47.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2006). *Multivariate Analysemethoden: eine anwendungsorientierte Einführung*. Berlin: Springer.
- Backus, A. (2005). Turkish as an immigrant language in Europe. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (S. 689–723). Oxford: Oxford University Press.
- 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, 417- 423.
- Baddeley, A. D., Gathercole, S. E. & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105, 158-173.
- Bailey, A. L. & Butler, F. (2003). *An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document*. Los Angeles, CA: University of California.
- Bailey, A. L., Butler, F. A., Stevens, R. & Lord, C. (2007). Further specifying the language demands of school. In A. L. Bailey (Hrsg.), *The language demands of school. Putting academic English to the test* (S. 103-156). New Haven, CT: Yale University Press.
- Batova, N. (2013). Academic listening: is there a place for bottom-up processing? *International Journal of Education and Research*, 1, 1-10.
- Baumann, J. F. & Graves, M. F. (2010). What is academic vocabulary? *Journal of Adolescent & Adult Literacy*, 54, 4-12.
- Baumert, J. & Schümer, G. (2001). Familiäre Lebensverhältnisse, Bildungsbeteiligung und Kompetenzerwerb. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider, P. Stanat, K.-J. Tillmann & M. Weiß (Hrsg.), *PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich* (S. 323-407). Opladen: Leske & Budrich.
- Baumgartner, H. & Steenkamp, J.-B. E. M. (1998). Multi-group latent variable models for varying numbers of items and factors with cross-national and longitudinal applications. *Marketing Letters*, 9, 21-35.
- Beck, I. L., McKeown, M. G. & Kucan, L. (2002). *Bringing words to life*. New York: Guilford Press.

- Becker, G. S. (1993). *Human Capital: A theoretical and empirical analysis, with special reference to education*. Chicago: University of Chicago Press.
- 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. Perspektiven aus Psychologie, Sprachwissenschaft und empirischer Bildungsforschung*. Münster: Waxmann.
- Bialystok, E. (2009). Consequences of bilingualism for cognitive development. In A. M. B. DeGroot & J. F. Kroll (Hrsg.), *Handbook of Bilingualism* (S. 417-432). Oxford: Oxford University Press.
- Bialystok, E., Luk, G., Peets, K. F. & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13, 525-531.
- Bialystok, E., Majumder, S. & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics*, 24, 27-44.
- Biedinger, N. & Becker, B. (2010). Frühe ethnische Bildungsungleichheit: Der Einfluss des Kindergartenbesuchs auf die deutsche Sprachfähigkeit und die allgemeine Entwicklung. In B. Becker & D. Reimer (Hrsg.), *Vom Kindergarten bis zur Hochschule. Die Generierung von ethnischen und sozialen Disparitäten in der Bildungsbiographie* (S. 49–79). Wiesbaden: VS Verlag.
- Bloomfield, A., Wayland, S. C., Rhoades, E., Blodgett, A., Linck, J., & Ross, S. (2010). What makes listening difficult? Factors affecting second language listening comprehension. Technical Report TTO 81434 E. 3.1. College Park, MD: University of Maryland Center for Advanced Study of Language.
- Böhme, K., Tiffin-Richards, S., Schipolowski, S. & Leucht, M. (2010). Migrationsbedingte Disparitäten bei sprachlichen Kompetenzen. In: O. Köller, M. Knigge & B. Tesch (Hrsg.), *Sprachliche Kompetenzen im Ländervergleich* (S. 203–225). Münster: Waxmann.
- Bonk, W. J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, 14, 14-31.
- Bonsen, M., Kummer, N. & Bos, W. (2008). Schülerinnen und Schüler mit Migrationshintergrund. In W. Bos, M. Bonsen, J. Baumert, M. Prenzel, C. Selter & G. Walther (Hrsg.), *TIMSS 2007. Mathematische und naturwissenschaftliche*

- Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 156–175). Münster: Waxmann.
- Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, 2, 331-350.
- Brownell, J. (1994). Teaching listening: Some thoughts on behavioral approaches. *The Bulletin*, 57, 19-26.
- Brownell, J. (2002). *Listening: Attitudes, principles, and skills*. Boston: Allyn and Bacon.
- Bruck, M. & Genesee, F. (1995). Phonological awareness in young second language learners. *Journal of Child Language* 22, 307-324.
- Buck, G. (1991). The testing of listening comprehension: an introspective study. *Language Testing*, 8, 67-91.
- Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- Butler, F. A., Bailey, A. L., Stevens, R. & Huang, B. H. (2004). *Academic English in fifth-grade mathematics, science, and social studies textbooks*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.
- Byrne, B. M., Shavelson, R. J. & Muthén, B. O. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin*, 105, 456-466.
- Call, M. E. (1985). Auditory short-term memory, listening comprehension, and the input hypothesis. *TESOL Quarterly*, 19, 765-781.
- Campbell, R. & Sais, E. (1995). Accelerated metalinguistic (phonological) awareness in bilingual children. *British Journal of Developmental Psychology*, 13, 61-68.
- Cauneau, I. (1992). Hören-Brummen-Sprechen. *Fremdsprache Deutsch*, 7, 28-30.
- Ceci, S. & Papierno, P. (2005). The rhetoric and reality of gap closing: When the “have-nots” gain but the “haves” gain even more. *American Psychologist*, 60, 149–160.
- Chall, J. S., Jacobs, V. A. & Baldwin, L. E. (1990). *The reading crisis. Why poor children fall behind*. Cambridge, MA: Harvard University Press.
- Chen, X., Ramirez, G., Luo, Y. C., Geva, E. & Ku, Y.-M. (2012). Comparing vocabulary development in Spanish and Chinese-speaking ELLs: the effects of metalinguistic and sociocultural factors. *Reading and Writing: An Interdisciplinary Journal*, 25, 1991-2020.
- Cheung, H. (2007). The role of phonological awareness in mediating between reading and listening to speech. *Language and Cognitive Processes*, 22, 130-154.

- Chiappe, P., Siegel, L. S. & Wade-Woolley, L. (2002). Linguistic diversity and the development of reading skills: A longitudinal study. *Scientific Studies of Reading*, 6, 369-400.
- Chiu, M. M., McBride-Chang, C. & Lin, D. (2012). Ecological, psychological, and cognitive components of reading difficulties: Testing the component model of reading in fourth graders across 38 countries. *Journal of Learning Disabilities*, 45, 391-405.
- Chudaske, J. (2012). *Sprache, Migration und schulfachliche Leistung*. Wiesbaden: VS Verlag.
- Chomsky, N. (1986). *Lectures on government and binding. The Pisa lectures*. Dordrecht: Foris Publications.
- Christman, U. & Groeben, N. (1999). Psychologie des Lesens. In B. Franzmann & G. Jäger (Hrsg.), *Handbuch Lesen* (S. 145 – 122). München: Saur.
- Clauser, B. E. & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice*, 17, 31-44.
- Cohen, J. (1988): *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Corson, D. (1997). The learning and use of academic English words. *Language Learning*, 47, 671-718.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213-238.
- Coxhead, A. (2012). Academic vocabulary, writing and English for academic purposes: Perspectives from second language learners. *RELC Journal*, 43, 137 –145.
- Cossu, G., Shankweiler, D., Liberman, I. Y., Katz, L. & Tola, G. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics*, 9, 1-16.
- Csomay, E. (2006). Academic talk in American university classrooms: crossing the boundaries of oral-literate discourse. *Journal of English for Academic Purposes*, 5, 117–135.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 222-251.
- Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 14, 175-187.
- Cummins, J. (1981). Age on arrival and immigrant second language learning in Canada. *Applied Linguistics*, 11, 132-149.
- Cummins, J. (2000). *Language, power, and pedagogy. Bilingual children in the crossfire*. Clevedon, UK: Multilingual Matters.

- 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* (S. 71-83). New York: Springer Science + Business Media LLC.
- Cummins, J. & Swain, M. (1986). *Bilingualism in education. Aspects of theory, research and practice*. London: Longman.
- Cunningham, J. W. & Moore, D. W. (1993). The contribution of understanding academic vocabulary to answering comprehension questions. *Journal of Reading Behavior*, 25, 171-180.
- Curenton, S. M. & Justice, L. M. (2014). African American and caucasian preschoolers' use of decontextualized language: Literate language features in oral narratives. *Language, Speech, and Hearing Services in Schools*, 35, 240-253.
- da Fontoura, H. A. & Siegel, L. S. (1995). Reading, syntactic, and working memory skills of bilingual Portuguese–English Canadian children. *Reading and Writing: An Interdisciplinary Journal*, 7, 139-153.
- Daneman, M. & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning & Verbal Behavior*, 19, 450-466.
- Daneman, M. & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. *Psychonomic Bulletin & Review*, 3, 422–433.
- Diefenbach, H. (2002). Bildungsbeteiligung und Berufseinmündung von Kindern und Jugendlichen aus Migrantenfamilien. Eine Fortschreibung der Daten des Sozio-Ökonomischen Panels (SOEP). In Sachverständigenkommission 11. Kinder- u. Jugendbericht (Hrsg.), *Migration und die europäische Integration - Herausforderungen für die Kinder- und Jugendhilfe* (S. 9-70). München: Verlag Deutsches Jugendinstitut.
- Douglas J., Roussos L. & Stout W. (1996). Item bundle DIF hypothesis testing: Identifying suspect bundles and assessing their DIF. *Journal of Educational Measurement*, 33, 465–484.
- Driessen, G., van der Slik, F. & de Bot, K. (2002). Home language and language proficiency: a large-scale longitudinal study in Dutch primary schools. *Journal of Multilingual and Multicultural Development*, 23, 175-194.
- Droop, M. & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly*, 38, 78-103.

- Dulay, H. C. & Burt, M. K. (1974). You can't learn without goofing. In J. C. Richards (Hrsg.), *Error analysis. Perspectives on second language acquisition* (S. 95-123). London: Longman.
- Duncan, L. G. & Seymour, P. H. K. (2000). Socio-economic differences in foundation-level literacy. *British Journal of Psychology*, 91, 145-166.
- Dunkel, P. (1986). Developing listening fluency in L2: Theoretical principles and pedagogical considerations. *The Modern Language Journal*, 70, 99-116.
- Duzy, D., Gold, A., Schneider, W. & Souvignier, E. (2013). Die Prädiktion von Leseleistungen bei türkisch-deutschsprachigen Kindern: Die Rolle der phonologischen Bewusstheit. *Zeitschrift für Pädagogische Psychologie*, 27, 41-50.
- Eckhardt, A. G. (2008). *Sprache als Barriere für den schulischen Erfolg: Potentielle Schwierigkeiten beim Erwerb schulbezogener Sprache für Kinder mit Migrationshintergrund*. Münster: Waxmann.
- Edwards, A. D. & Westgate, D. (1994). *Investigating classroom talk* (S. 65–73). Brighton: Falmer Press
- Ehmke, T. & Siegle, T. (2005). ISEI, ISCED, HOMEPOS, ESCS. Indikatoren der sozialen Herkunft bei der Quantifizierung von sozialen Disparitäten. *Zeitschrift für Erziehungswissenschaft*, 8, 521-540.
- Ehri, L. C. (1995). Stages of development in learning to read words by sight. *Journal of Research in Reading*, 18, 116-125.
- Engel de Abreu, P. M. J., Baldassi, M., Puglisi, M. L. & Befi-Lopes, D. M. (2013). Cross-linguistic and cross-cultural effects on verbal working memory and vocabulary: Testing language-minority children with an immigrant background. *Journal of Speech, Language, and Hearing Research*, 56, 630-642.
- Engel, P. M. J., Santos, F. H. & Gathercole, S. E. (2008). Are working memory measures free of socioeconomic influence? *Journal of Speech, Language, and Hearing Research*, 51, 1580–1587.
- Ennemoser, M., Pepouna, S. & Hartung, N. (2012). Kulturfaire und prognostisch valide Erfassung des phonologischen Arbeitsgedächtnisses bei Kindern mit Migrationshintergrund. *Unterrichtswissenschaft*, 40, 26-46.
- Esser, H. (2001). Integration und ethnische Schichtung. *Arbeitspapiere – Mannheimer Zentrum für Europäische Sozialforschung*, 40, 1-77.

- Esser, H. (2006). *Sprache und Integration: Konzeptionelle Grundlagen und empirische Zusammenhänge*. KMI Working Paper 7. Verfügbar unter:
http://www.oeaw.ac.at/kmi/Bilder/kmi_WP7.pdf
- Færch, C. & Kasper, G. (1986). The role of comprehension in second language learning. *Applied Linguistics*, 7, 257-274.
- Feilke, H. (2012). Bildungssprachliche Kompetenzen – fördern und entwickeln. *Praxis Deutsch*, 233, 1-13.
- Felser, C. & Clahsen, H. (2009). Grammatical processing of spoken language in child and adult language learners. *Journal of Psycholinguistic Research*, 38, 305-319.
- Ferris, D. & Tagg, T. (1996). Academic oral communication needs of EAP learners: What subject-matter instructors actually require. *TESOL Quarterly*, 30, 31-58.
- Field, J. (2003). Promoting perception: lexical segmentation in L2 listening. *ELT Journal* 57, 325-334.
- Field, J. (2004). An insight into listeners' problems: too much bottom-up or too much top-down? *System* 32, 363–377.
- Field, J. (2008). *Listening in the language classroom*. Cambridge: Cambridge University Press.
- Florit, E., Roch, M. & Levorato, M. C. (2014). Listening text comprehension in preschoolers: A longitudinal study on the role of semantic components. *Reading and Writing: An Interdisciplinary Journal*, 27, 793–817.
- Flowerdew, J. (1993). Concordancing as a tool in course design. *System*, 21, 231–244.
- Flowerdew, J. (1995). Research of relevance to second language lecture comprehension: An overview. In J. Flowerdew (Hrsg.), *Academic listening: Research perspectives* (S. 7-29). Cambridge: Cambridge University Press.
- Flowerdew, J. & Miller, L. (1997). The teaching of academic listening comprehension and the question of authenticity. *English for Specific Purposes*, 16, 27–46.
- Flowerdew, J. & Miller, L. (2005). *Second language acquisition. Theory and practice*. New York: Cambridge University Press.
- Fries, C. C. (1945). *Teaching and learning English as a second language*. Ann Arbor, MI: The University of Michigan Press.
- Gathercole, V. C. M. (2002). Monolingual and bilingual acquisition: Learning different treatments of that-trace phenomena in English and Spanish. In D. K. Oller & R. E. Eilers (Hrsg.), *Language and literacy in bilingual children* (S. 220-252). Clevedon, UK: Multilingual Matters.

- Gathercole, S. E. & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, 28, 200-213.
- Geva, E. & Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and EL1. *Reading and Writing: An Interdisciplinary Journal*, 25, 1819-1845.
- Gilakjani, A. P. & Ahmadi, M. A. (2011). A study of factors affecting EFL learners' English listening comprehension and the strategies for improvement. *Journal of Language Teaching and Research*, 2, 977-988.
- Gogolin, I. (2005). Bilingual education – The German experience and debate. In Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration (Hrsg.), *The effectiveness of bilingual school programs for immigrant children* (S. 133-145). Berlin: AKI.
- Gogolin, I. (2009). Zweisprachigkeit und die Entwicklung bildungssprachlicher Fähigkeiten. In I. Gogolin & U. Neumann (Hrsg.), *Streitfall Zweisprachigkeit – The Bilingualism Controversy* (S. 263-280). Wiesbaden: VS Verlag.
- Gogolin, I. & Lange, I. (2011). Bildungssprache und Durchgängige Sprachbildung. In S. Fürstenau & M. Gomolla (Hrsg.), *Migration und schulischer Wandel: Mehrsprachigkeit* (S. 107-128). Wiesbaden: VS Verlag.
- Goh, C. C. M. (1998). How ESL learners with different listening abilities use comprehension strategies and tactics. *Language Teaching Research*, 2, 124–147.
- Goh, C. C. M. (2000). A cognitive perspective on language learner's listening comprehension problems. *System*, 28, 55-75.
- Golberg, H., Paradis, J. & Crago, M. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied Psycholinguistics*, 29, 41-65.
- Goldenberg, C. (2008). Teaching English language learners. What the research does—and does not—say. *American Educator*, 32, 8-23.
- Goldenberg, C. (2011). Reading instruction for English language learners. In M. L. Kamil, P. D. Pearson, E. Birr Moje & P. P. Afflerbach (Hrsg.), *Handbook of reading research* (S. 684–710). New York, NY: Routledge.
- Goldenberg, C., Rueda, R. S. & August, D. (2006). Sociocultural influences on the literacy attainment of language-minority children and youth. In D. August & T. Shanahan (Hrsg.), *Developing literacy in second-language learners: Report of the national*

- literacy panel on language-minority children and youth* (S. 269-318). Mahwah, NJ: Lawrence Erlbaum.
- Gottardo, A. & Mueller, J. (2009). Are first- and second-language factors related in predicting second-language reading comprehension? A study of spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology, 101*, 330–344.
- Grießhaber, W. (1999). *Die relationierende Prozedur. Zur Grammatik und Pragmatik lokaler Präpositionen und ihrer Verwendung durch türkische Deutschlerner*. Münster: Waxmann.
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition, 1*, 131-140.
- Grotjahn, R. (2005). Testen und Bewerten des Hörverstehens. In M. O. Duill, R. Zahn & K. D. C. Höppner (Hrsg.), *Zusammenarbeiten: Eine Festschrift für Bernd Voss* (S. 115-144). Bochum: AKS-Verlag.
- Grotjahn, R. (2012). Hörverstehen: Konstrukt und Messung. *Fremdsprachen Lehren und Lernen 41*, 72-86.
- Gutierrez-Clellen, V. F. & Simon-Cereijido, G. (2007). The discriminant accuracy of a grammatical measure with Latino English-speaking children. *Journal of Speech, Language and Hearing Research, 50*, 968-981.
- Haag, N., Böhme, K. & Stanat, P. (2012). Zuwanderungsbezogene Disparitäten. 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* (S. 209–235). Münster: Waxmann.
- Haag, N., Heppt, B., Stanat, P., Kuhl, P. & Pant, H. A. (2013). Second language learners' performance in mathematics: Disentangling the effects of academic language features. *Learning and Instruction, 28*, 24-34.
- Harklau, L. (1994). Jumping tracks: How language-minority students negotiate evaluations of ability. *Anthropology and Education Quarterly, 25*, 347-363.
- Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore, MD: Brookes.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision: KFT 4-12+ R*. Weinheim: Beltz.

- Heppt, B., Haag, N., Böhme, K. & Stanat, P. (2015). The role of academic-language features for reading comprehension of language-minority students and students from low-SES families. *Reading Research Quarterly*, 50, 61-82.-
- Heppt, B., Stanat, P., Dragon, N., Berendes, K. & Weinert, S. (2014). Bildungssprachliche Anforderungen und Hörverstehen bei Kindern mit deutscher und nicht-deutscher Familiensprache. *Zeitschrift für Pädagogische Psychologie*, 28, 139-149.
- Hesse, H.G., Göbel, K. & Hartig, J. (2008). Sprachliche Kompetenzen von mehrsprachigen Jugendlichen und Jugendlichen nicht-deutscher Erstsprache. In DESI-Konsortium unter Leitung von Eckhard Klieme (Hrsg.). *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie (S.208-230)*. Weinheim: Beltz.
- Hiebert, E. H. & Lubliner, S. (2008). The nature, learning, and instruction of general academic vocabulary. In A. E. Farstrup & S. J. Samuels (Hrsg.), *What research has to say about vocabulary instruction* (S. 106–129). Newark, DE: International Reading Association
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, 74, 1368–1378.
- Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, 26, 55-88.
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, 49, 4-14.
- Hogan, T. P., Adlof, S. M. & Alonzo, C. N. (2014). On the importance of listening comprehension. *International Journal of Speech-Language Pathology*, 16, 199–207.
- Hoover, W. A. & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160.
- Hu, L.-T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Hufeisen, B. & Riemer, C. (2010). Spracherwerb und Sprachenlernen. In H.-J. Krumm, C. Fandrych, B. Hufeisen & C. Riemer (Hrsg.), *Handbuch Deutsch als Fremd- und Zweitsprache* (S. 737-752). Berlin: De Gruyter.
- Hyland, K. & Tse, P. (2007). Is there an “academic vocabulary”? *TESOL Quarterly*, 41, 235-253.

- Imhof, M. (1998). What makes a good listener? Listening behavior in instructional settings. *International Journal of Listening*, 12, 81–105.
- Imhof, M. (2004). *Zuhören und Instruktion – Empirische Zugänge zur Verarbeitung mündlich vermittelter Information*. Münster: Waxmann.
- Imhof, M. (2010). Zuhören lernen und lehren. Psychologische Grundlagen zur Beschreibung und Förderung von Zuhörkompetenzen in Schule und Unterricht. In V. Bernius (Hrsg.), *Zuhörkompetenz in Unterricht und Schule. Beiträge aus Wissenschaft und Praxis* (S. 15-30). Göttingen: Vandenhoeck & Ruprecht.
- Janssen, M., Bosman, A. M. T. & Leseman, P. P. M. (2013). Phoneme awareness, vocabulary and word decoding in monolingual and bilingual Dutch children. *Journal of Research in Reading*, 36, 1-13.
- Jean, M. & Geva, E. (2009). The development of vocabulary in English as a second language children and its role in predicting word recognition ability. *Applied Psycholinguistic*, 30, 153 - 185.
- Jöreskog, K. G. & Sörbom, D. (1982). Recent developments in structural equation modeling. *Journal of Marketing Research*, 19, 404-416.
- Johnson- Laird, P. N. (1983). A computational analysis of consciousness. *Cognition & Brain Theory*, 499-508.
- Kahn-Horwitz, J., Shimron, J. & Sparks, R. L. (2006). Weak and strong novice readers of English as a foreign language: Effects of first language and socioeconomic status. *Annals of Dyslexia*, 56, 161-185.
- Kang, J. Y. (2012). Do bilingual children possess better phonological awareness? Investigation of Korean monolingual and Korean-English bilingual children. *Reading and Writing: An Interdisciplinary Journal*, 25, 411–431.
- Kelly, P. (1991). Lexical ignorance: the main obstacle to listening comprehension with advanced foreign language learners. *International Review of Applied Linguistics* 29, 135–149.
- Kieffer, M. J. (2014). Morphological awareness and reading difficulties in adolescent Spanish-speaking language minority learners and their classmates. *Journal of Learning Disabilities*, 47, 44-53.
- Kieffer, M. J. & Lesaux, N. K. (2012). Knowledge of words, knowledge about words: Dimensions of vocabulary in first and second language learners in sixth grade. *Reading and Writing: An Interdisciplinary Journal*, 25, 347-373.

- Kieffer, M. J. & Vukovic, R. K. (2012). Components and context: Exploring sources of reading difficulties for language minority learners and native English speakers in urban schools. *Journal of Learning Disabilities, 20*, 1-20.
- Kieffer, M. J. & Vukovic, R. K. (2013). Growth in reading-related skills of language minority learners and their classmates: more evidence for early identification and intervention. *Reading and Writing: An Interdisciplinary Journal, 26*, 1159–1194.
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction integration model. *Psychological Review, 95*, 163-182.
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Kintsch, W. & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363-394.
- KMK - Sekretariat der ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. (2005). *Bildungsstandards im Fach Deutsch für den Hauptschulabschluss*. München: Luchterhand.
- Kniffka, G. & Siebert-Ott, G. (2009). *Deutsch als Zweitsprache. Lehren und Lernen*. Paderborn: Schöningh.
- Koch, P. & Oesterreicher, W. (1985). Sprache der Nähe - Sprache der Distanz. Mündlichkeit und Schriftlichkeit im Spannungsfeld von Sprachtheorie und Sprachgeschichte. *Romanistisches Jahrbuch, 36*, 15-43.
- Koda, K. (1998). The role of phonemic awareness in second language reading. *Second Language Research, 14*, 194-215.
- Köhne, J., Kronenwerth, S., Redder, A., Schuth, E. & Weinert, S. (2015). Bildungssprachlicher Wortschatz – linguistische und psychologische Fundierung und Itementwicklung. In A. Redder, J. Naumann & R. Tracy (Hrsg.), *Forschungsinitiative Sprachdiagnostik und Sprachförderung (FiSS) – Ergebnisse* (S. 67-92). Münster: Waxmann
- Kurita, T. (2012). Issues in second language listening comprehension and the pedagogical implications. *Accents Asia, 5*, 30-44.
- Lado, R. (1957). *Linguistics across cultures*. Ann Arbor, MI: University of Michigan Press.
- Lanfranchi, S. & Swanson, H. L. (2005). Short-term memory and working memory in children as a function of language-specific knowledge in English and Spanish. *Learning and Individual Differences, 15*, 299-319.
- Language and Reading Research Consortium. (2015). Learning to read: Should we keep things simple? *Reading Research Quarterly, 50*, 151-169.

- Lawrence, J. F. (2012). English vocabulary trajectories of students whose parents speak a language other than English: steep trajectories and sharp summer setback. *Reading and Writing: An Interdisciplinary Journal*, 25, 1113-1141.
- Lesaux, N. K., Kieffer, M. J., Faller, E. & Kelley, J. (2010). The effectiveness and ease of implementation of an academic vocabulary intervention for linguistically diverse students in urban middle schools. *Reading Research Quarterly*, 45, 198-230.
- Lesaux, N. K., Kieffer, M. J., Kelley, J. G. & Harris, J. R. (2014). Effects of academic vocabulary instruction for linguistically diverse adolescents: Evidence from a randomized field trial. *American Educational Research Journal*, 51, 1159–1194.
- Lesaux, N. K., Lipka, O. & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing: An Interdisciplinary Journal*, 19, 99-131.
- Lesaux, N. K. & Siegel, L. S. (2003). The development of reading in children who speak English as a second language. *Developmental Psychology*, 39, 1005–1019.
- Leucht, M., Retelsdorf, J., Möller, J. & Köller, O. (2010). Zur Dimensionalität rezeptiver englischsprachiger Kompetenzen. *Zeitschrift für Pädagogische Psychologie*, 24, 123-138.
- Limbird, C. (2007). *Phonological processing, verbal abilities, and second language literacy development among bilingual Turkish children in Germany*. Unveröffentlichte Dissertation, Freie Universität Berlin.
- Limbird, C. & Stanat, P. (2006). Prädiktoren von Leseverständnis bei Kindern deutscher und türkischer Herkunftssprache: Ergebnisse einer Längsschnittstudie. In A. Ittel & H. Merkens (Hrsg.), *Veränderungsmessung und Längsschnittstudien in der empirischen Erziehungswissenschaft* (S. 93-123). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Limbird, C. K., Maluch, J. A., Rjosk, C., Stanat, P. & Merkens, H. (2014). Differential growth patterns in emerging reading skills of bilingual and monolingual children: A longitudinal study of Turkish-German bilingual and German monolingual elementary school students. *Reading and Writing: An Interdisciplinary Journal*, 27, 945-968.
- Linck, J. A., Osthus, P., Koeth, J. T. & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 21, 861-883.
- Lindholm-Leary, K. & Borsato, G. (2006). Academic achievement. In F. Genesee, K. Lindholm-Leary, W. M. Saunders & D. Christian (Hrsg.), *Educating English language learners* (S. 176-222). Cambridge, NY: Cambridge University Press.

- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83, 1198–1202.
- Little, R. J. A. & Rubin, D. B. (2002). *Statistical Analysis with missing data*. New Jersey, NJ: John Wiley.
- Lonigan, C. J., Burgess, S. R., Anthony, J. L. & Barker, T. A. (1998). Development of phonological sensitivity in 2- to 5-year-old children. *Journal of Educational Psychology*, 90, 294-311.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen. *Psychologische Rundschau*, 58, 103-117.
- Lund, R. J. (1990). A taxonomy for teaching second language listening. *Foreign Language Annals*, 23, 105-115.
- Lundsteen, S. W. (1971). *Listening: Its impact on reading and other language arts*. Urbana, IL: National Council of Teachers of English.
- Lynch, T. (1998). Theoretical perspectives on listening. *Annual Review of Applied Linguistics*, 18, 3–19.
- Mancilla-Martinez, J. & Lesaux, N. K. (2011). Early home language use and later vocabulary development. *Journal of Educational Psychology*, 103, 535-546.
- Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, 78, 333-368.
- 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.
- Marx, A., Wolf, K. M., Paetsch, J., Darsow, A. & Felbrich, A. (2014). Förderung von Teilfertigkeiten der Lesekompetenz bei Kindern mit Zuwanderungshintergrund. In R. Valtin & I. Tarelli (Hrsg.), *Lesekompetenz nachhaltig stärken: Evidenzbasierte Maßnahmen und Programme* (S.206-224). Berlin: DGLS.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjær & J. Williams (Hrsg.), *Performance and competence in second language acquisition* (S. 35–53). New York: Cambridge University Press.
- Mecarty, F. H. (2000). Lexical and grammatical knowledge in reading and listening comprehension by foreign language learners of Spanish. *Applied Language Learning*, 11, 323-348.

- Meurers, D., Berendes, K., Vajjala, S. & Bryant, D. (2015, März). *Leseanforderungen in der Sekundarstufe: Ein Vergleich der linguistischen Komplexität von Schulbuchtexten*. Vortrag auf der 3. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF), 11. bis 13. März 2015 in Bochum. Online verfügbar unter <http://purl.org/dm/handouts/geb15.pdf>
- Müller, K. (2000). *Lernen im Dialog: Gestaltlinguistische Aspekte des Zweitspracherwerbs*. Tübingen: Gunter Narr Verlag.
- Muthén, B. O., Kaplan, D. & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, 52, 431-462.
- Muthén, L. K. & Muthén, B. O. (1998-2012). *MPlus Version 7.11*. Los Angeles, CA: Muthén & Muthén.
- Nagle, S. T. & Sanders, S. L. (1986). Comprehension theory and second language pedagogy. *TESOL Quarterly*, 20, 9-26.
- Nagy, W. & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47, 91-108.
- Nasser-Abu Alhija, F. & Wisenbaker, J. (2006). A Monte Carlo study investigating the impact of item parceling strategies on parameter estimates and their standard errors in CFA. *Structural Equation Modeling*, 13, 204-228.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63, 59-82.
- Naumann, J., Artelt, C., Schneider, W. & Stanat, P. (2010). Lesekompetenz von PISA 2000 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 23-71). Münster: Waxmann.
- Niklas, F., Schmiedeler, S., Pröstler, N. & Schneider, W. (2011). Die Bedeutung des Migrationshintergrunds, des Kindergartenbesuchs sowie der Zusammensetzung der Kindergartenengruppe für sprachliche Leistungen von Vorschulkindern. *Zeitschrift für Pädagogische Psychologie*, 25, 115-130.
- Noble, K. G., Farah, M. J. & McCandliss, B. D. (2006). Socioeconomic background modulates cognition-achievement relationships in reading. *Cognitive Development*, 21, 349-368.
- Ochs, E. (1979). Planned and unplanned discourse. In T. Givón (Hrsg.), *Discourse and syntax* (S. 51-80). New York, NY: Academic Press.

- OECD (2002). *Reading for change: Performance and engagement across countries. Results from PISA 2000*. Paris: OECD.
- OECD (2006). *PISA 2006: Technical report*. Paris: OECD.
- OECD (2010). *PISA 2009 results: Overcoming social background – equity in learning opportunities and outcomes (Volume II)*. Paris: OECD.
- OECD (2013). *PISA 2012 results: Excellence through equity. Giving every student the chance to succeed (Volume II)*. Paris: OECD.
- Oksaar, E. (2003). *Zweitspracherwerb. Wege zur Mehrsprachigkeit und interkulturellen Verständigung*. Stuttgart: Kohlhammer.
- Oller, D. K. (2005). The distributed characteristic in bilingual learning. In J. Cohen, K. T. McAlister, K. Rolstad & J. MacSwan (Hrsg.), *Proceedings of the 4th International Symposium on Bilingualism* (S. 1744–1749). Somerville, MA: Cascadilla Press.
- Oller, D. K., Pearson, B. Z. & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28, 191-230.
- Olsthoorn, N. M., Andringa, S. & Hulstijn, J. H. (2012). Visual and auditory digit-span performance in native and nonnative speaker. *International Journal of Bilingualism*, 18, 1-11.
- O'Malley, J. M., Chamot, A. U. & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics*, 10, 418-437.
- Paradis, J. (2010). Bilingual children's acquisition of English verb morphology: Effects of language exposure, structure complexity, and task type. *Language Learning*, 60, 651-680.
- Paschke, P. (2000). *Fremdsprachliches Hörverstehen: Grundlagen, Lernziele und Probleme der Leistungsmessung*. Unveröffentlichte Dissertation, University College Dublin.
- Pearson, B. Z., Fernandez, S. C., Lewedeg, V. & Oller, D. K. (1997). The relation of input factors to lexical learning by bilingual infants. *Applied Psycholinguistics*, 18, 41-58.
- Peltzer-Karpf, A., Schwab, B., Wurnig, V., Piwonka, D., Lederwasch, K., Griessler, M. A., Akkus, R., Blazevic, T., Karabinova, D., Kuemmel, R. & Sundl, P. (2006). *A kući sprecham Deutsch. Sprachstandserhebung in multikulturellen Volksschulklassen: bilingualer Spracherwerb in der Migration*. Wien: bm:bwk.
- Pöhlmann, C., Haag, N. & Stanat, P. (2013). Zuwanderungsbezogene Disparitäten. In H. A. Pant, P. Stanat, U. Schroeders, A. Roppelt, T. Siegle & C. Pöhlmann (Hrsg.), *IQB-Ländervergleich 2012. Mathematische und naturwissenschaftliche Kompetenzen am Ende der Sekundarstufe I* (S. 297-330). Münster: Waxmann.

- Proctor, C. P., Carlo, M., August, D. & Snow, C. (2005). Native spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology, 97*, 246-256.
- Purdy, M. (1997). What is listening? In M. Purdy & D. Borisoff (Hrsg.), *Listening in everyday life. A personal and professional approach* (S. 1-20). Lanham, MD: University Press of America.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning, 52*, 513–536.
- Reich, H. H. (2009). Entwicklungswege türkisch-deutscher Zweisprachigkeit. In U. Neumann & H. H. Reich (Hrsg.), *Erwerb des Türkischen in einsprachigen und mehrsprachigen Situationen* (S. 63-90). Münster: Waxmann.
- Richards, J. C. (1971). A non-contrastive approach to error-analysis. *English Language Teaching, 25*, 204-219.
- Richards, J. C. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly, 17*, 219-240.
- Richards, J. C. (2005). Second thoughts on teaching listening. *RELC Journal, 36*, 85-92.
- Röhr-Sendlmeier, U.M. (1985). *Zweitsprachenerwerb und Sozialisationsbedingungen*. Frankfurt am Main: Lang.
- Rost, D. H. & Hartmann, A. (1992). Lesen, Hören, Verstehen. *Zeitschrift für Psychologie, 200*, 345-361.
- Rost, M. (1990). *Listening in language learning*. Harlow: Longman.
- Rost, M. (1994). *Introducing listening*. London: Penguin.
- Rost, M. (2005). L2 listening. In E. Hinkel (Hrsg.), *Handbook of research on second language teaching and learning* (S. 503–528). Mahwah, NJ: Lawrence Erlbaum.
- Rost, M. (2011). *Teaching and researching listening (2nd edition)*. London: Routledge.
- Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal, 78*, 199-221.
- Rubin, H. & Turner, A. (1989). Linguistic awareness skills in grade one children in a French immersion setting. *Reading and Writing: An Interdisciplinary Journal, 1*, 73-86.
- Samuels, S. J. (1984). Factors influencing listening: Inside and outside the head. *Theory into Practice, 23*, 183-189.

- Savage, R. (2006). Reading comprehension is not always the product of nonsense word decoding and linguistic comprehension: Evidence from teenagers who are extremely poor readers. *Scientific Studies of Reading, 10*, 143-164.
- Scarcella, R. C. (2003). *Accelerating academic English: A focus on English language learners*. Oakland, CA: Regents of University of California.
- Scarcella, R. & Rumberger, R. W. (2000) Academic English key to long-term success in school. *UC Linguistic Minority Research Institute Newsletter, 9*, 1-2.
- Schleppegrell, M. J. (2004). *The language of schooling. A functional linguistics perspective*. Mahwah, NJ: Lawrence Erlbaum.
- Schneider, W., Schlagmüller, M. & Ennemoser, M. (2007). *Lesegeschwindigkeits- und Verständnistest für die Klassen 6-12 (LGVT 6-12)*. Göttingen: Hogrefe.
- Schwippert, K., Bos, W. & Lankes, E.-M. (2003). Heterogenität und Chancengleichheit am Ende der vierten Jahrgangsstufe im internationalen Vergleich. In W. Bos, E.-M. Lankes, M. Prenzel, K. Schwippert, G. Walther & R. Valtin (Hrsg.), *Erste Ergebnisse aus IGLU. Schülerleistungen am Ende der vierten Jahrgangsstufe im internationalen Vergleich* (S. 265–302). Münster: Waxmann.
- Schwippert, K., Hornberg, S., Freiberg, M. & Stubbe, T. C. (2007). Lesekompetenz von Kindern mit Migrationshintergrund im internationalen Vergleich. In W. Bos, S. Hornberg, K.-H. Arnold, G. Faust, L. Fried, E.-M. Lankes, K. Schwippert & R. Valtin (Hrsg.), *IGLU 2006. Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich* (S. 249–269). Münster: Waxmann.
- Selinker, L. (1992). *Rediscovering interlanguage*. London: Longman.
- Service, E., Simola, M., Metsänheimo, O. & Maury, S. (2002). Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology, 14*, 383-408.
- Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). *Experimental and quasiexperimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.
- Shaftel, J., Belton-Kocher, E., Glasnapp, D. & Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of English language learners and students with disabilities. *Educational Assessment, 11*, 105-126.
- Shealy, R. & Stout, W. (1993). A model-based standardization approach that separates true bias/DIF from group ability differences and detects test bias/DIF as well as item bias/DIF. *Psychometrika, 58*, 159-194.

- Shohamy, E. & Inbar, O. (1991). Validation of listening comprehension tests: the effect of text and question type. *Language Testing*, 8, 23-40.
- Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, 328, 450-452.
- Snow, C. E., Lawrence, J. F. & White, C. (2009). Generating knowledge of academic language among urban middle school students. *Journal of Research on Educational Effectiveness*, 2, 325-344.
- Snow, C. E. & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Hrsg.), *The Cambridge handbook of literacy* (S. 112-133). Cambridge: Cambridge University Press
- Söhn, J. (2005). *Zweisprachiger Schulunterricht für Migrantenkinder: Ergebnisse der Evaluationsforschung zu seinen Auswirkungen auf Zweitspracherwerb und Schulerfolg*. Berlin: Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration, Wissenschaftszentrum Berlin für Sozialforschung.
- Solmecke, G. (1992). Ohne Hören kein Sprechen. Bedeutung und Entwicklung des Hörverstehens im Deutschunterricht. *Fremdsprache Deutsch. Zeitschrift für die Praxis des Deutschunterrichts*, 7, 4-11.
- Solmecke, G. (1993). *Texte hören, lesen, verstehen. Eine Einführung in die Schulung der rezeptiven Kompetenz mit Beispielen für den Unterricht Deutsch als Fremdsprache*. München: Langenscheidt.
- Song, M.-Y. (2011). Note-taking quality and performance on an L2 academic listening test. *Language Testing*, 29, 67-89.
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31, 577-607.
- Stanat, P., Pant, H. A., Böhme, K. & Richter, D. (Hrsg.). (2012). *Kompetenzen von Schülerinnen und Schülern am Ende der vierten Jahrgangsstufe in den Fächern Deutsch und Mathematik. Ergebnisse des IQB-Ländervergleichs 2011*. Münster: Waxmann.
- Stanat, P., Rauch, D. & Segeritz, M. (2010). Schülerinnen und Schüler mit Migrationshintergrund. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009: Bilanz nach einem Jahrzehnt* (S. 200-230). Münster: Waxmann.
- Stanat, P., Schwippert, K. & Gröhlich, C. (2010). Der Einfluss des Migrantenanteils in Schulklassen auf den Kompetenzerwerb: Längsschnittliche Überprüfung eines umstrittenen

- Effekts. In C. P. Allemann-Ghionda, P. Stanat, K. Göbel, & C. Röhner (Hrsg.), *Migration, Sprache, Identität* [55. Beiheft]. *Zeitschrift für Pädagogik*, 56, 147–164.
- Stanat, P., Weirich, S. & Radmann, S. (2012). Sprach- und Leseförderung. 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* (S. 251-276). Münster: Waxmann.
- Steenkamp, J.-B. E. M. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research,” *Journal of Consumer Research*, 25, 78-90.
- Steil, L. K., Barker, L. L. & Watson, K. W. (1983). *Effective listening: Key to your success*. Reading: Addison-Wesley.
- Stein, S. K. (1999). *Uncovering listening strategies: Protocol analysis as a means to investigate student listening in the basic communication course*. College Park, MY: University of Maryland.
- Strobel, B. & Kristen, C. (2015). Erhalt der Herkunftssprache? – Muster des Sprachgebrauchs in Migrantenfamilien. *Zeitschrift für Erziehungswissenschaft*, 18, 125–142.
- Tannen, D. (1980). Implications of the oral/literate continuum for cross-cultural communication. In J. E. Alatis (Hrsg.), *Current issues in bilingual education* (S. 326-347). Washington, DC: Georgetown University Press.
- Tannen, D. (1982). The oral literate continuum of discourse. In D. Tannen (Hrsg.), *Spoken and written language* (S. 1-6). Norwood, NJ: Ablex Publishing.
- Taylor, L. & Geranpayeh, A. (2011). Assessing listening for academic purposes: Defining and operationalising the test construct. *Journal of English for Academic Purposes*, 10, 89–101.
- Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P. & Rapp, D. (2009). Simple but complex: components of the simple view of reading across grade levels. *Journal of Research in Reading*, 32, 383–401.
- Townsend, D. & Collins, P. (2009). Academic vocabulary and middle school English learners: an intervention study. *Reading and Writing: An Interdisciplinary Journal*, 22, 993–1019.
- Townsend, D. R., Filippini, A., Collins, P. & Biancarosa, G. (2012). Evidence for the importance of academic word knowledge for the academic achievement of diverse middle school students. *Elementary School Journal*, 112, 497-518.
- Tsui, A. M. B. & Fullilove, J. (1998). Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics* 19, 432-451.

- Tyler, M. D. (2001). Resource consumption as a function of topic knowledge in nonnative and native comprehension. *Language Learning*, 51, 257–280. doi: 10.1111/1467-9922.00155
- Tyler, R. & Marslen-Wilson, W. (1977). The on-line effects of semantic context on syntactic processing. *Journal of Verbal Learning and Verbal Behavior*, 16, 683-692.
- Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A. & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly*, 50, 337–356.
- Umbel, V. M. & Oller, D. K. (1994). Developmental changes in receptive vocabulary in Hispanic bilingual school children. *Language Learning*, 44, 221–242.
- Umbel, V. M., Pearson, B. Z., Fernández, M. C. & Oller, D. K. (1992). Measuring bilingual children's receptive vocabularies. *Child Development*, 63, 1012-1120.
- Van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Vandergrift, L. (2003). From prediction through reflection: Guiding the students through the process of L2 listening. *Canadian Modern Language Review*, 59, 425 –440.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching* 40, 191-210.
- Verhoeven, L. (1990). Acquisition of reading in a second language. *Reading Research Quarterly*, 25, 90-114.
- Verhoeven, L. & van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing: An Interdisciplinary Journal*, 25, 1805-1818.
- Vidal, K. (2003). Academic listening: A source of vocabulary acquisition? *Applied Linguistics*, 24, 56-89.
- von Goldammer, A., Mähler, C. & Hasselhorn, M. (2011). Determinanten von Satzgedächtnis-Leistungen bei deutsch- und mehrsprachigen Vorschulkindern. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43, 1-15.
- Voss, B. (1984). *Slips of the ear*. Tübingen: Gunter Narr Verlag.
- Wade-Woolley, L. & Siegel, L. S. (1997). The spelling performance of ESL and native speakers of English as a function of reading skill. *Reading and Writing: An Interdisciplinary Journal*, 9, 387–406.
- Wallace, M. P. (2012). Searching for a new approach to listening. *Accents Asia*, 5, 8-22.

- Warren, R. M. & Warren, R. P. (1970). Auditory illusions and confusions. *Scientific American*, 223, 30-36.
- Weiß, R. H. (2006). *Grundintelligenztest Skala 2, CFT 20 - Revision (CFT 20-R)*. Göttingen: Hogrefe.
- Wiley, H. I. & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education*, 26, 207–214.
- Wilson, M. (2003). Discovering listening - Improving perceptual processing. *ELT Journal*, 57, 335-343.
- Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. A. (2007). *ACER ConQuest Version 2: Generalised item response modelling software*. Camberwell, Victoria: ACER Press.
- Wolter, B. (2001). Comparing the L1 and L2 mental lexicon. *Studies in Second Language Acquisition*, 23, 41-69.
- Wolvin, A. D. & Coakley, C. G. (1979). *Listening instruction*. Urbana, IL: ERIC Clearinghouse on Reading and Other Communication Skills.
- Wolvin, A. D. & Coakley, C. G. (1996). *Listening (5th ed)*. Dubuque, IA: Brown and Benchmark.
- Zwaan, R. A. (1993). *Aspects of literary comprehension: A cognitive approach*. Amsterdam: John Benjamins.
- Zwaan, R. A. & Singer, M. (2003). Text comprehension. In A. C. Graesser, M. A. Gernsbacher & S. R. Goldman (Hrsg.), *Handbook of discourse processes* (S. 83–121). Mahwah, NJ: Lawrence Erlbaum Associates.

Anhang A: Prädiktoren des Hörverstehens bei Jugendlichen deutscher und Jugendlichen nichtdeutscher Herkunftssprache (Teilstudie 1)

Marx, A. & Roick, T. (2012). Prädiktoren des Hörverstehens bei Jugendlichen deutscher und Jugendlichen nichtdeutscher Herkunftssprache. *Zeitschrift für Pädagogische Psychologie*, 26, 121-134.² <http://dx.doi.org/10.1024/1010-0652/a000067>

(*Zeitschrift für Pädagogische Psychologie*, 26/2 © 2012 by Hogrefe)

Zusammenfassung. Trotz der zentralen Bedeutung des Hörverstehens für weitere sprachliche Leistungen existieren bislang nur wenige Untersuchungen, die Hörverstehensleistungen und deren Prädiktoren bei Schülerinnen und Schülern unterschiedlicher Herkunftssprache vergleichend analysieren. Die vorliegende Studie untersucht deutschsprachiges Hörverstehen und dessen Determinanten bei Jugendlichen deutscher (L1) und Jugendlichen nichtdeutscher (L2) Herkunftssprache. Im Rahmen der Erhebung bearbeiteten 424 Neuntklässler Aufgaben zum deutschsprachigen Hörverstehen sowie zur phonologischen Bewusstheit, zum Arbeitsgedächtnis, zu morphosyntaktischen Fähigkeiten und zum Wortschatz. Ausgehend von theoretischen Annahmen und empirischen Befunden wird angenommen, dass L2-Jugendliche schwächere Hörverstehensleistungen aufweisen sowie über einen geringeren Wortschatz verfügen als L1-Jugendliche. Weiterhin wird ein im Vergleich zu L1-Jugendlichen stärkerer Einfluss des Wortschatzes, der phonologischen Bewusstheit und des Arbeitsgedächtnisses auf die Hörverstehensleistung bei L2-Jugendlichen angenommen. Die strukturanalytische Modellierung dieser Determinanten des Hörverstehens zeigt, dass L2-Jugendliche geringere Wortschatzkenntnisse aufweisen als L1-Jugendliche. Schwächen im Bereich des Hörverstehens finden sich nur für Jugendliche, die Türkisch als Erstsprache gelernt haben. Bei allen Jugendlichen wird die Hörverstehensleistung durch die morphosyntaktischen Kompetenzen sowie den Wortschatz determiniert; bei Jugendlichen mit Türkisch als Erstsprache findet sich zusätzlich ein Effekt der phonologischen Bewusstheit auf die Hörverstehensleistung. Die Implikationen der Befunde für theoretische Annahmen zu Hörverstehensleistungen in einer Zweitsprache werden diskutiert.

Schlagwörter: Hörverstehen – Sprachkompetenz (Zweitsprache) – Zweisprachigkeit – Schüler der Sekundarstufe

² Diese Artikelfassung entspricht nicht vollständig dem in der Zeitschrift veröffentlichten Artikel. Dies ist nicht die Originalversion des Artikels und kann daher nicht zur Zitierung herangezogen werden.

Abstract. Numerous studies indicate a strong relationship between students' reading comprehension and listening skills. Despite the important role of listening comprehension for other language competencies, only a few studies compared listening comprehension and its determinants in first-language (*L1*) and second-language (*L2*) students. The present study examines listening comprehension and its predictors for students with German as a first or as a second language. Based on theoretical assumptions and empirical results from previous studies, we expected that listening comprehension and vocabulary would be less developed in *L2* students than in *L1* students. Furthermore, we predicted that vocabulary, phonological awareness, and working memory would play a more important role for *L2* students' listening comprehension than for *L1* students' listening comprehension. In our study, we assessed listening comprehension, phonological awareness, working memory, morphosyntactic skills, and vocabulary of 424 ninth-grade students. SEM analyses show that *L2* students' vocabulary was less developed than *L1* students' vocabulary. However, only Turkish-speaking *L2* students showed weaker listening comprehension than *L1* students. For all students, listening comprehension was affected by vocabulary and morphosyntactic knowledge. In addition, phonological awareness had an effect on Turkish-speaking *L2* students' listening comprehension. The results are discussed in terms of theoretical assumptions on listening comprehension in a second language.

Keywords: Listening Comprehension – Language Proficiency – Bilingualism – High School Students

Traditionell wird davon ausgegangen, dass Hören und Sprechen in der Erstsprache dem Erwerb des Lesens und Schreibens vorausgehen (Danks & End, 1987; Stanovich, Cunningham & Freeman, 1984). Befunde aus dem englischsprachigen Raum belegen, dass die Hörverstehensleistung auch in einer Zweitsprache ein signifikanter Prädiktor der Leseverstehensleistung ist (Royer & Carlo, 1991). In Deutschland konnte beispielsweise für Jugendliche mit Deutsch als Erstsprache gezeigt werden, dass Hörverstehen und Leseverstehen beim Erwerb einer Fremdsprache hoch miteinander korreliert sind und über einen Faktor zweiter Ordnung modelliert werden können (Leucht, Retelsdorf, Möller & Köller, 2010).

Trotz der Bedeutung des Hörverstehens für die Leseleistung in einer Zweitsprache existieren bislang nur wenige Studien, die das Hörverstehen und dessen Determinanten bei Schülerinnen und Schülern in der Erst- (*L1*) und Zweitsprache (*L2*) untersuchen (Eckhardt, 2008). Es fehlen zum einen Analysen zu zentralen theoretischen Vorstellungen über den Prozess des Hörverstehens in einer Zweitsprache, zum anderen aber auch die differenzielle Betrachtung des Einflusses einzelner sprachlicher Determinanten auf die Hörverstehensleistung in einer Erstsprache und in einer Zweitsprache. Im Folgenden werden bisherige Befunde zu Determinanten des Hörverstehens in einer Erst- und Zweitsprache dargestellt. Aufgrund der dürftigen Befundlage zum Hörverstehen im Zweitspracherwerb wird dabei auch Literatur zum Fremdspracherwerb zitiert.

1. Das Prozessmodell des Hörverstehens

Ähnlich wie Leseverstehen ist Hörverstehen ein konstruktiver Prozess, bei dem unter Zuhilfenahme linguistischen Wissens und Weltwissens eine mentale Repräsentation des Textes gebildet wird. So genannte *PDP (Parallel Distributed Processing)-Modelle* postulieren einen multidimensionalen Hörverstehensprozess, in dessen Verlauf phonologische, syntaktische, semantische und pragmatische Informationen simultan genutzt werden, um die Bedeutung eines Hörtextes zu erfassen (Buck, 1994; Flowerdew & Miller, 2005). Anderson (1995) hat ein PDP-Modell des Textverstehens vorgelegt, das sowohl auf das Lesen als auch auf das Hören angewendet werden kann und den Verstehensprozess in drei Ebenen unterteilt: *Wahrnehmung, syntaktische und semantische Analyse (sog. Parsing)* sowie *Verwendung des Gehörten* (vgl. Abbildung 1).

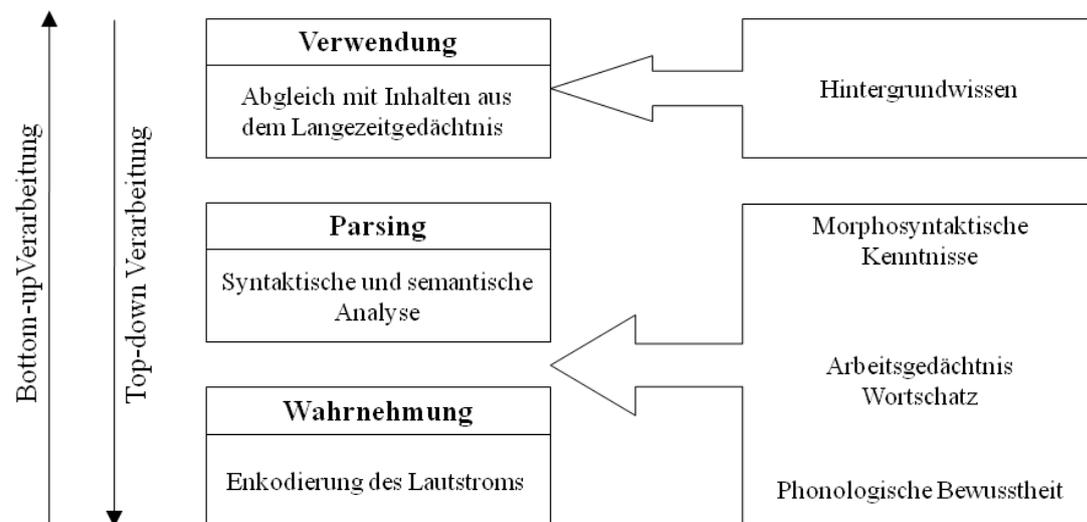


Abbildung 1. Modell des Hörverstehens und seiner Determinanten, angelehnt an Anderson (1995)

In diesem Modell wird auf der *Wahrnehmungsebene* die gesprochene Nachricht zunächst durch das Segmentieren von Phonemen aus dem Lautstrom enkodiert. Hierfür ist eine gut ausgeprägte phonologische Bewusstheit, d. h. die Einsicht in die phonologischen Strukturen der Sprache sowie die Analyse und Synthese phonologischer Einheiten, von zentraler Bedeutung (Flowerdew & Miller, 2005; Jansen, 1992; Osada, 2004). Die enkodierten Laute verbleiben im echoischen Gedächtnis, um anschließend unter Nutzung syntaktischer und lexikalischer Kenntnisse auf der *Ebene des Parsing* in eine mentale Repräsentation der sprachlichen Äußerung überführt zu werden. Dieser Prozess erfordert eine unmittelbare und fortlaufende Interpretation des Gehörten sowie die Speicherung der gewonnenen Informationen im Speichersystem des sprachlichen Arbeitsgedächtnisses (Call, 1985; Shohamy & Inbar, 1991). Auf der *Verwendungsebene* wird die mentale Repräsentation des Textinhalts schließlich mit bereits bestehenden Gedächtnisinhalten abgeglichen und in Form von Schemata oder Propositionen im Langzeitgedächtnis abgespeichert (vgl. Anderson, 1995; Goh, 2000). Die verschiedenen Verarbeitungsebenen agieren dabei nicht isoliert voneinander, sondern verlaufen parallel sowohl ausgehend von den hierarchieniedrigen Ebenen zu den hierarchiehöheren Ebenen

(*bottom up-Verarbeitung*), als auch in die umgekehrte Verarbeitungsrichtung (*top down-Verarbeitung*) (Anderson, 1995; Flowerdew & Miller, 2005; Vandergrift, 2007).

Obwohl das grundlegende Hörverstehensmodell für Erst- und Zweitsprachen als gleichermaßen gut empirisch belegt gelten kann (vgl. O'Malley, Chamot & Küpper, 1989), scheinen Hörerinnen und Hörer in einer Zweitsprache spezifische Hörverarbeitungsprofile aufzuweisen. So greifen sie offenbar häufiger auf bottom-up Prozesse zurück als Muttersprachler und versuchen anhand von isolierten Einheiten des Sprachinputs (z. B. einzelnen Wörtern) die Bedeutung des Hörtextes zu entschlüsseln (O'Malley et al., 1989; Stæhr, 2009; Vandergrift, 2007). Diese Fokussierung der unteren Verarbeitungsebenen kann in der Folge zu einer schwachen Hörverstehensleistung führen (Hasan, 2000).

Trotz der zentralen Bedeutung von bottom-up Prozessen für die Hörverstehensleistung in einer Zweitsprache wurden bislang hauptsächlich top-down Prozesse der höheren Verarbeitungsebenen, wie etwa der Einfluss von Hintergrundwissen, untersucht (z. B. Field, 2003; Long, 1990). Determinanten des L2-Hörverstehens auf den unteren Verarbeitungsebenen waren bislang kaum Gegenstand empirischer Studien (vgl. Bonk, 2000; Eckhardt, 2008). Es bleibt daher unklar, inwiefern Schwächen im Hörverstehen bei Schülerinnen und Schülern in einer Zweitsprache ein Resultat von geringer ausgeprägten Determinanten des Hörverstehens sind bzw. ob die einzelnen Determinanten bei Schülerinnen und Schülern unterschiedlicher Herkunftssprache möglicherweise differenziell bedeutsam für die Hörverstehensleistung sind.

2. Hörverstehen in einer Zweitsprache

2.1 Determinanten des Hörverstehens in einer Zweitsprache

Während einige Determinanten des Hörverstehens bei Schülerinnen und Schülern in einer Zweitsprache geringer ausgeprägt sind als bei Schülerinnen und Schülern in einer Erstsprache, lassen sich für andere Bereiche keine Nachteile mehrsprachig aufwachsender Kinder feststellen. So weisen empirische Studien beispielsweise darauf hin, dass phonologische Fähigkeiten auch bei unterschiedlichen Schriftsprachsystemen von der Erst- auf die Zweitsprache übertragen werden können (z. B. Schwartz, Geva, Share & Leikin, 2007). Teilweise verfügen mehrsprachig aufwachsende Kinder im Vergleich zu einsprachig aufwachsenden Kindern sogar über eine ausgeprägtere phonologische Bewusstheit, die ein Resultat des Kontakts mit mehreren phonologischen Systemen sein dürfte. Dieser Vorteil scheint jedoch mit Beginn des schulischen Lesenlernens, wenn der Lautanalyse im Rahmen des Unterrichts vermehrt Aufmerksamkeit gewidmet wird, zu verschwinden (Bialystok, Luk & Kwan, 2005; Stewart, 2004). Insgesamt existieren keine eindeutigen Belege dafür, dass sich die Ausprägung der

phonologischen Bewusstheit bei mehrsprachig aufwachsenden Schülerinnen und Schülern von der einsprachig aufwachsender Kinder dauerhaft unterscheidet (Bialystok et al., 2005).

Auch die Arbeitsgedächtniskapazität mehrsprachig aufwachsender Kinder scheint sich nicht von der monolingualer Schülerinnen und Schüler zu unterscheiden (Da Fontoura & Siegel, 1995). Insbesondere bei fortgeschrittenen zweitsprachlichen Kenntnissen und hohem Automatisierungsgrad des Stimulusmaterials (z. B. Ziffern oder einfache einsilbige Wörter) sind in Hinblick auf die Arbeitsgedächtnisleistung keine Differenzen zwischen L1- und L2-Schülerinnen und Schülern zu erwarten (vgl. Call, 1985).

Im Bereich der lexikalischen Kenntnisse hingegen weisen Schülerinnen und Schüler in einer Zweitsprache häufig schwächere Leistungen auf als Schülerinnen und Schüler in einer Erstsprache (Limbird, 2007; Verhoeven, 1990). So ist das Hörvokabular in einer Zweitsprache oftmals begrenzt und weniger automatisiert verfügbar als in einer Erstsprache (Grotjahn, 2005).

Neben den Wortschatzkenntnissen ist auch das notwendige syntaktische Wissen in einer Zweitsprache häufig geringer als in einer Erstsprache (Eckhardt, 2008; Lesaux, Lipka & Siegel, 2006). Syntaktische Bereiche, die potenziell zu Schwierigkeiten im Umgang mit der Zweitsprache Deutsch führen könnten sind u.a. Passivkonstruktionen, komplexe Satzkonstruktionen und Strukturwörter, wie etwa Konjunktionen oder Pronomen (Eckhardt, 2008).

2.2 Bedeutsamkeit von Determinanten des Hörverstehens in einer Zweitsprache

Generell sind für das Hörverstehen in einer Erst- und in einer Zweitsprache dieselben Determinanten von Bedeutung. Allerdings existieren Hinweise darauf, dass einzelne Determinanten in einer Zweitsprache einen stärkeren Effekt auf das Hörverstehen haben als in einer Erstsprache. So vermutet Field (2003) etwa, dass Schwierigkeiten bei der Identifikation einzelner Wörter im Sprachfluss die häufigste perzeptuelle Ursache für schwach ausgebildete L2-Hörverstehensleistungen seien. Tatsächlich gab in einer von Goh (2000) durchgeführten Interview-Studie ein Großteil der L2-Probanden an, Schwierigkeiten bei der Identifizierung von Phonemen und dem Erkennen von Wortgrenzen zu haben (vgl. auch Fishman, 1980). Die Segmentation des Gehörten anhand phonologischer Marker scheint demnach einer der Bereiche zu sein, der die Hörverstehensleistung in einer Zweitsprache stark beeinflusst.

Auch die Arbeitsgedächtniskapazität dürfte in einer Zweitsprache besonders eng mit der Hörverstehensleistung verbunden zu sein. Ebenso wie muttersprachliche Hörerinnen und Hörer behalten Lernende einer Zweitsprache auditiv dargebotene Wörter nur so lange im Speichersystem des sprachlichen Arbeitsgedächtnisses, bis auf der Ebene der syntaktischen

und semantischen Analyse deren Bedeutungsinhalt extrahiert wurde (Call, 1985). Allerdings kann es bei L2-Schülerinnen und Schülern aufgrund der im Vergleich zu muttersprachlichen Hörerinnen und Hörern geringeren Wortschatz- und Syntaktikkenntnisse zu einer Überlastung des Arbeitsgedächtnissystems kommen (Vandergrift, 2007). Diese ergibt sich daraus, dass eine automatisierte Verarbeitung des Hörtextes nicht möglich ist und eine arbeitsgedächtnisintensive aktive Analyse des Hörinhalts nötig wird. Als Folge einer differenziellen Ressourcenallokation kann eine im Vergleich zu L1-Schülerinnen und Schülern schwächere Hörverstehensleistung bei Schülerinnen und Schülern in einer Zweitsprache resultieren.

Empirische Studien belegen auch einen engen Zusammenhang zwischen zweitsprachlichem Wortschatz und Hörverstehen. So korrelierten der Umfang des Wortschatzes und das englische Hörverstehen bei Schülerinnen und Schülern mit Englisch als Zweitsprache in einer Studie von Stæhr (2009) zu $r = .70$. Ebenso weisen Fehleranalysen zu Hörtexttranskriptionen auf die besondere Bedeutung des Wortschatzes bei L2-Lernenden hin: Bei Studierenden, die Englisch als Zweitsprache lernten, kamen auch bei fortgeschrittenen Zweitsprachkenntnissen noch 42% der während eines Diktats gemachten Transkriptionsfehler zustande, weil das gehörte Wort unbekannt war (Kelly, 1991). Angesichts dieses hohen Anteils an wortschatzbedingten Fehlern kommt Kelly (1991) zu dem Schluss, dass geringe lexikalische Kenntnisse hauptverantwortlich für schwache Hörverstehensleistungen sind. Für Studierende, die Englisch als Zweitsprache lernten und weniger als 80% der Wörter in den Inputtexten erkannten, zeigte sich entsprechend eine geringere Hörverstehensleistung als für L2-Studierende mit sehr guten Wortschatzkenntnissen (Bonk, 2000).

Obwohl aus theoretischen Gründen auch eine starke Beeinflussung der Hörverstehensleistung in einer Zweitsprache durch die syntaktischen Fähigkeiten der Schülerinnen und Schülern zu erwarten wäre, ist die empirische Befundlage hierzu eher widersprüchlich. So zeigte sich für Studierende mit Englisch als Zweitsprache und geringen L2-Kenntnissen kein Einfluss syntaktischer Variationen des Sprachinputs auf das zweitsprachliche Hörverstehen (Blau, 1990). Mecarty (2000) berichtet einen mittleren, statistisch aber nicht signifikanten Zusammenhang ($r = .26$) zwischen den syntaktischen Kenntnissen von Studierenden, die Englisch als Zweitsprache lernten, und deren Hörverstehensleistungen. Ähnliche Ergebnisse zeigten sich in einer Studie für Grundschülerinnen und -schüler mit Deutsch als Zweitsprache, die keine spezifischen, auf die syntaktische Komplexität der Hörtexte zurückführbaren Nachteile im Hörverstehen von L2-Schülerinnen und -Schülern belegen konnte (Eckhardt, 2008). Allerdings wurden weder bei Blau (1990) noch bei Eckhardt (2008) die syntaktischen Kenntnisse direkt erfasst, sondern lediglich die syntaktische Komplexität des Inputs variiert. Aufgrund

des Fehlens eines direkten Maßes bleibt der Effekt der grammatikalischen Kenntnisse auf die Hörverstehensleistung in einer Zweitsprache somit unklar.

Insgesamt weist der aktuelle Forschungsstand darauf hin, dass dieselben Determinanten für Hörverstehensprozesse in der Erst- sowie Zweitsprache bedeutsam sind. Gleichzeitig scheinen einzelne Prädiktoren für die Hörverstehensleistung in einer Zweitsprache von besonderer Bedeutung zu sein. Da nur wenige Studien vorliegen, die Prädiktoren des Hörverstehens in einer L1 und einer L2 vergleichend untersuchen, sind belastbare Aussagen über die relativen Effekte von Determinanten der Hörverstehensleistung in Abhängigkeit vom Sprachhintergrund bislang nicht möglich.

3. Fragestellung

Die vorliegende Studie beschäftigt sich mit der Frage, ob die im Vergleich zu Schülerinnen und Schülern in einer Erstsprache häufig geringere Ausprägung des Hörverstehens in einer Zweitsprache auf entsprechende Unterschiede in basalen Determinanten des Hörverstehens zurückgeführt werden kann. Weiterhin wird gefragt, ob die einzelnen Determinanten die Hörverstehensleistung in einer Erst- bzw. Zweitsprache unterschiedlich stark beeinflussen.

Auf der Grundlage der beschriebenen theoretischen Annahmen und der empirischen Befundlage ist bezüglich der *Unterschiede* in den Leistungsausprägungen zu erwarten, dass L2-Jugendliche im Vergleich zu deutschsprachigen L1-Jugendlichen schwächere Hörverstehensleistungen zeigen sowie über einen geringeren Wortschatz verfügen. Hinsichtlich der syntaktischen Kenntnisse werden ebenfalls Nachteile der L2-Jugendlichen gegenüber L1-Jugendlichen erwartet. Für die Arbeitsgedächtniskapazität und die phonologische Bewusstheit hingegen werden keine Gruppenunterschiede angenommen.

Hinsichtlich der *differenziellen Bedeutsamkeit* der Determinanten ist auf Basis bisheriger Befunde zu erwarten, dass die Effekte der phonologischen Bewusstheit und des Wortschatzes auf die Hörverstehensleistung bei L2-Jugendlichen stärker ausgeprägt sind als bei L1-Jugendlichen (Eckhardt, 2008; Field, 2003; Fishman, 1980; Mecarty, 2000). Für das Arbeitsgedächtnis wird für die L2-Jugendlichen ebenfalls ein im Vergleich zu den L1-Jugendlichen stärkerer Effekt auf die Hörverstehensleistung angenommen, der sich aus der Notwendigkeit einer aktiven Verarbeitung des Hörinhalts ergibt (Goh, 2000; Vandergrift, 2007). Syntaktische Kenntnisse dürften sowohl für das Hörverstehen in einer Erst- als auch für das Hörverstehen in einer Zweitsprache von Bedeutung sein, um die eingehenden Informationen zu verarbeiten (vgl. Anderson, 1995). Bisherige Studien konnten aber für Schülerinnen und Schüler in einer Zweitsprache keinen ausgeprägten Effekt des syntaktischen Wis-

sens auf das Hörverstehen nachweisen (vgl. Eckhardt, 2008; Mecarty, 2000), so dass zu erwarten ist, dass der Effekt der morphosyntaktischen Kompetenzen auf das Hörverstehen in der Gruppe der L2-Jugendlichen eher schwächer ausfällt als in der Gruppe der L1-Schülerinnen und Schüler. Eine Übersicht der vermuteten Bedeutsamkeit von am Hörverstehen beteiligten Faktoren für die beiden Sprachgruppen findet sich in Tabelle 1.

Tabelle 1.

Theoriegeleitete Annahmen zu Niveauunterschieden und der Bedeutsamkeit von Determinanten des Hörverstehens in einer Erst- bzw. Zweitsprache

	Niveauunterschiede in Determinanten des Hör- verstehens		Bedeutsamkeit von Determinan- ten des Hörverstehens	
	Erstsprache	Zweitsprache	Erstsprache	Zweitsprache
Arbeitsgedächtnis		=	+	++
Phonologische Be- wusstheit		=	+	++
Morphosyntaktische Kenntnisse		>	+	○
Wortschatz		>	+	++

Anmerkungen. = = ähnliche Ausprägung der Determinanten; > = stärkere Ausprägung der Determinanten; ○ = schwacher Einfluss auf das Hörverstehen; + = Einfluss auf das Hörverstehen; ++ = starker Einfluss auf das Hörverstehen

Aufgrund der besonders benachteiligten Position türkischstämmiger Schülerinnen und Schüler im deutschen Bildungssystem (Stanat, Rauch & Segeritz, 2010) liegt der Fokus der Analysen auf einem Vergleich der Hörverstehensprozesse bei Jugendlichen deutscher und türkischer Herkunftssprache. Um Anhaltspunkte darauf zu gewinnen, inwieweit etwaige Schwierigkeiten im Hörverstehensprozess der L2-Schülerinnen und Schüler spezifisch für die jeweilige Erstsprache sind, wird zusätzlich zu Jugendlichen mit Türkisch als Erstsprache eine weitere L2-Gruppe mit anderen Erstsprachen als Türkisch untersucht. Weiter ist anzunehmen, dass sich familiäre Strukturmerkmale, wie etwa der sozioökonomische Status, sowohl mittelbar über familiäre Prozessmerkmale als auch unmittelbar auf die Kompetenzen von Schülerinnen und Schülern auswirken. Für geschriebene Texte konnte ein entsprechendes Wirkmodell direkter und indirekter Effekte auf die Lesekompetenz und den Wortschatz von Schülerinnen und Schülern bereits bestätigt werden (z. B. McElvany, Becker & Lüdtke, 2009). Für

den Zweispracherwerb zeigen sich Zusammenhänge zwischen sozioökonomischem Status der Jugendlichen und der Lesekompetenz bzw. dem Hörverstehen (Leucht, Retelsdorf, Möller, Köller, 2010). Aber auch in grundlegenden Leistungen des Sprachverstehens, wie der phonologischen und morphologischen Bewusstheit, unterscheiden sich L2-Schülerinnen und Schüler in Abhängigkeit vom sozioökonomischen Status (vgl. Kahn-Horwitz & Shimron, 2006). Um sicherzustellen, dass gefundene Unterschiede in den sprachlichen Kompetenzen tatsächlich auf den sprachlichen Hintergrund zurückzuführen sind und nicht auf mögliche konfundierende Variablen, wurde in der vorliegenden Studie der sozioökonomische und kulturelle Hintergrund der Schülerinnen und Schüler kontrolliert.

4. Methode

4.1 Stichprobe

Die im Folgenden beschriebenen Analysen wurden im Rahmen des DFG-Projekts „Determinanten der Lesekompetenz“ (DELKO)³ durchgeführt. DELKO ist als querschnittliche Studie konzipiert und erfasst Leseverstehensleistungen und deren Determinanten bei Jugendlichen der 9. Klassenstufe. Die DELKO-Studie war an die PISA 2009-Erhebung angeknüpelt und nutzt zur Erfassung der Lesekompetenz sowie der Hintergrunddaten der Schülerinnen und Schüler die entsprechenden PISA-Instrumente. Da das Hauptaugenmerk von DELKO auf der Analyse schwacher Leseleistungen lag, wurde die Studie vor allem in Hauptschulen durchgeführt. Die Auswertung der PISA-Lesekompetenztests für die DELKO-Stichprobe zeigt, dass 49% der an DELKO teilnehmenden Schülerinnen und Schüler nicht die PISA-Lesekompetenzstufe II erreicht haben.

Jugendliche mit diagnostizierten Hörschwierigkeiten ($n = 6$) wurden aus den Analysen ausgeschlossen. Die resultierende Stichprobe bestand aus 424 Schülerinnen und Schülern (45% Mädchen) aus 20 Schulen. Von den Jugendlichen besuchten 43% Hauptschulen, 25% Realschulen und 32% Gesamtschulen. Alle Schülerinnen und Schüler befanden sich zum Zeitpunkt der Untersuchung in der 9. Klassenstufe, das mittlere Alter lag bei 16.03 Jahren ($SD = 0.75$).

Die Jugendlichen wurden anhand der Antworten auf die in einem Fragebogen vorgegebene Frage nach ihrer Erstsprache („Welche Sprache hast du in deiner Familie *zuerst* gelernt?“) einer von drei Sprachgruppen (deutschsprachig, türkischsprachig, anderssprachig) zugeordnet. Von den Schülerinnen und Schülern gaben 157 (37%) Deutsch als Herkunftssprache an, 161 Schülerinnen und Schüler (38%) nannten Türkisch als Herkunftssprache, wei-

³ Projektleiter: Prof. Dr. Petra Stanat, Institut zur Qualitätsentwicklung im Bildungswesen (IQB) an der Humboldt-Universität zu Berlin und Prof. Dr. Wolfgang Schneider, Julius-Maximilians-Universität Würzburg

tere 106 Schülerinnen und Schüler (25%) gaben eine andere Sprache als Herkunftssprache an. Die in dieser Gruppe am häufigsten vertretenen Sprachgruppen waren Italienisch (15%), Arabisch (14%) und Polnisch (11%). Die überwiegende Mehrheit sowohl der türkischsprachigen als auch der anderssprachigen Schülerinnen und Schüler hatte die deutsche Sprache bereits vor dem Schuleintritt gelernt (93% bzw. 80%).

Hinsichtlich des Alters ($F(2;421)=1.02, p=.36$), der Geschlechterzusammensetzung ($\chi^2(2)=1.44, p=.49$) und des Besuchs der einzelnen Schulformen ($\chi^2(4)=2.51, p=.64$) finden sich keine Unterschiede zwischen den drei Sprachgruppen. In der Betrachtung von Unterschieden in der kognitiven Grundfähigkeit, die über den Subtest *Figurale Analogien (N2)* des Kognitiven Fähigkeitstest (KFT 4-12+R; Heller & Perleth, 2000) erfasst wurden, zeigten sich nur tendenzielle Unterschiede zum Nachteil der L2-Gruppen ($F(2;390)=2.90, p=.06$). Derartige Differenzen können für Grundintelligenztests, die kulturfrees Testmaterial nutzen, als akzeptabel gelten (vgl. z. B. Weiß, 2006).⁴

4.2 Instrumente

4.2.1 Hörverstehen in Deutsch

Die basalen Hörverstehensleistungen in Deutsch wurden mit einem computergestützten Testverfahren erfasst (Marx & Stanat, 2009). Den Schülerinnen und Schülern wurden acht kurze deutschsprachige Hörtexte mit einer Textlänge von jeweils 100 bis 150 Wörtern vorgespielt. Die Schwierigkeit der Hörtexte wurde über den Wortschatz und die syntaktische Komplexität variiert. Syntaktisch und lexikalisch anspruchsvollere Texte zeichnen sich dabei durch die vermehrte Anzahl von Strukturwörtern, Komposita, Relativsätzen und Passivkonstruktionen aus (vgl. hierzu auch Eckhardt, 2008). Thematisch behandeln die Hörtexte sowohl alltägliche (z. B. Einkauf in einem Kiosk) als auch schulbezogene Inhalte (z. B. Besprechung einer Hausaufgabe). Im Rahmen einer Pilotstudie konnte sichergestellt werden, dass die behandelten Themen Schülerinnen und Schülern der verschiedenen Sprachgruppen in ähnlichem Maße vertraut sind.

Unmittelbar nach dem Anhören eines Textes wurden den Jugendlichen am Computerbildschirm Fragen dazu präsentiert. Insgesamt enthält der Hörverstehenstest 32 Fragen zu acht Hörtexten, davon 19 Items im Multiple-Choice-Format mit je vier Antwortalternativen sowie 13 offene Fragen, die eine kurze schriftliche Antwort erfordern. Die Auswertungen der offenen Aufgaben wurden von einer geschulten Mitarbeiterin anhand einer Kodieranweisung vor-

⁴ Aus diesen Gründen wurde in der vorliegenden Analyse die kognitive Grundfähigkeit nicht als Kovariate einbezogen. Unter Einbezug der kognitiven Grundfähigkeit als Kovariate ändert sich das Befundmuster nicht.

genommen. Die Gesamtskala der Hörverstehensleistung weist eine interne Konsistenz von $\alpha=.78$ auf.

4.2.2 Phonologische Bewusstheit

Da kein geeignetes Verfahren zur Erfassung der phonologischen Bewusstheit für die untersuchte Altersgruppe zur Verfügung stand, wurden in Anlehnung an einen Test zur Erfassung der phonologischen Informationsverarbeitung älterer Schulkinder (Leibert, 2001) sowie des Testverfahrens „Basiskompetenzen für Lese-Rechtschreibleistungen (BAKO)“ (Stock & Schneider, in Vorb.) Skalen zur Lautkategorisierung und Vokallängenbestimmung entwickelt und in einer Pilotstudie erprobt. Die Skala *Lautkategorisierung* erfordert von den Schülerinnen und Schülern die Identifikation eines akustisch vorgegebenen Wortes mit abweichendem Konsonanten am Wortende oder abweichendem Vokal in der Wortmitte in einer Reihe von Pseudowörtern (Itembeispiel: *Welches dieser Worte hört sich am Ende anders an als die anderen? pat - kut - pit - pal*).

Bei der Skala zur *Vokallängenbestimmung* wurden die Schülerinnen und Schüler gebeten, aus einer Reihe von vier akustisch vorgegebenen Pseudowörtern das Wort mit dem längsten bzw. dem kürzesten Vokal in der Wortmitte zu identifizieren (Itembeispiel: *Bei welchem dieser Worte ist der Laut in der Mitte etwas länger oder kürzer als bei den anderen? maar - raas - dack - laat*). Die Aufgaben zur Lautkategorisierung und Vokallängenbestimmung wurden akustisch über einen Kopfhörer dargeboten und computergestützt ohne Zeitvorgabe administriert. Insgesamt bearbeiteten die Jugendlichen 26 Items zur phonologischen Bewusstheit, davon 11 Items zur Lautkategorisierung ($\alpha=.71$) und 15 Items zur Vokallängenbestimmung ($\alpha=.77$). Die beiden Subskalen korrelieren zu $r=.39$.

4.2.3 Wortschatz

Zur Erfassung des deutschsprachigen Wortschatzes der Schülerinnen und Schüler wurde die Wortschatz-Erweiterung des Grundintelligenztests Skala 2 (CFT 20 R-WS; Weiß, 2008) eingesetzt. Den Jugendlichen wurden 30 Testaufgaben mit Wörtern vorgelegt, zu denen aus einer Reihe von jeweils fünf Auswahlwörtern dasjenige Wort mit der gleichen oder einer ähnlichen Bedeutung ausgewählt werden sollte. Bei diesem Test handelt es sich um ein Papier-und-Bleistift-Verfahren, das ohne Zeitvorgabe administriert wird. Die interne Konsistenz der Skala liegt in der vorliegenden Untersuchung bei $\alpha=.81$.

4.2.4 Morphosyntaktische Kenntnisse

Zur Erfassung sprachstruktureller Kenntnisse bearbeiteten die Schülerinnen und Schüler einen neu entwickelten computergestützten Test zur Erfassung syntaktischer und morphosyntaktischer Kenntnisse (Segerer, Marx & Schneider, 2009). Die Subskala zu *syntaktischen Kenntnissen* erfordert von den Jugendlichen die Beurteilung schriftlich vorgegebener Sätze auf ihre syntaktische Richtigkeit. Um die Ratewahrscheinlichkeit zu minimieren, wird neben den ja/nein-Antwortkategorien zusätzlich die Option „unklar“ vorgegeben, die die Jugendlichen ankreuzen sollen, wenn sie sich hinsichtlich der Beurteilung des Satzes unsicher sind (Itembeispiel: *Ist der folgende Satz grammatikalisch richtig oder falsch? „Dort, wo viele Menschen sind, auch sind Diebe viele.“ richtig - falsch - unklar*). Bei der Subskala *morphologische Kenntnisse* werden den Schülerinnen und Schülern unvollständige Sätze schriftlich dargeboten, die durch ein morphologisch passendes Wort ergänzt werden sollen (Itembeispiel: *Vervollständige den Satz! „Ich weiß nicht, ____ von all diesen Spielern ich am liebsten mag.“ welcher - welchen - welchem - wer*). Insgesamt bearbeiteten die Jugendlichen 47 Items, davon 22 zu syntaktischen und 25 Items zu morphologischen Kenntnissen. Die interne Konsistenz der Skala zur Erfassung syntaktischer Kenntnisse liegt bei $\alpha=.86$, die der Skala zur Erfassung morphologischer Kenntnisse bei $\alpha=.83$. Die Interkorrelation der beiden Subskalen zur Erfassung morphosyntaktischer Fähigkeiten beträgt $r=.59$.

4.2.5 Arbeitsgedächtnis

Zur Feststellung der Arbeitsgedächtnisleistung im Sinne von Baddeley (1986) wurden zwei Varianten von Ziffernspannungsaufgaben in einer computerisierten Fassung visuell dargeboten. Bei der *Ziffernspanne vorwärts* mussten die Schülerinnen und Schüler die auf dem Computerbildschirm präsentierten Ziffern in der gesehenen Reihenfolge über den Tastaturnummernblock eingeben. Für die *Ziffernspanne rückwärts* waren die Ziffern im Anschluss an die Darbietung in der umgekehrten Reihenfolge in den Nummernblock einzugeben. Die Ziffern wurden jeweils für eine Dauer von 1500 ms im Abstand von 1000 ms präsentiert. Die Eingabe der Lösung erfolgte ohne Zeitvorgabe, nach erfolgter Eingabe wurde 1500 ms später die nächste Ziffernspannungsaufgabe präsentiert. Insgesamt wurden jeweils 12 Ziffernspannen vorwärts und rückwärts mit einer Spannenlänge von 3 bis 7 Ziffern vorgegeben, wobei die Spannenlänge im Untersuchungsverlauf sukzessive anstieg. Die interne Konsistenz der Ziffernspanne vorwärts liegt bei $\alpha=.66$, die der Ziffernspanne rückwärts bei $\alpha=.84$. Die Interkorrelation der beiden Subskalen beträgt $r=.44$.

4.2.7 *Ökonomischer, sozialer und kultureller Status (ESCS)*

Der Sozialstatus der Schülerinnen und Schüler wird mit dem *Index des ökonomischen, sozialen und kulturellen Status* (ESCS; vgl. OECD, 2006) abgebildet, in den Angaben über den höchsten Bildungsabschluss und den höchsten Berufsstatus der Eltern sowie die ökonomischen und kulturellen Ressourcen der Familie eingehen. Der Index wurde anhand des Mittelwerts der OECD-Population zentriert und an deren Standardabweichung standardisiert.

4.3 *Analysestrategie*

Die strukturanalytische Überprüfung der Hypothesen erfolgte anhand von Mehrgruppenmodellen (vgl. Abbildung 2). Zur Schätzung der latenten Variablen wurden entsprechend der Vorschläge von Little, Cunningham, Shahar und Widaman (2002) je Konstrukt zwei Indikatoren über eine Item-Parcelung gebildet und auf Gleichheit der Residualvarianz restringiert. Das Modell wurde zunächst ohne Einbezug von Kovariaten geschätzt. Anschließend wurde der Index des ökonomischen, sozialen und kulturellen Status (ESCS) als manifeste Kovariate in der Analyse berücksichtigt (zum Vorgehen vgl. Muthén & Muthén, 1998-2010).

Sämtliche Modellschätzungen erfolgten mit dem Programm Mplus Version 6.1 (Muthén, & Muthén, 1998-2010). Zur Prüfung auf Messinvarianz zwischen den drei Sprachgruppen wurden auf Basis eines unrestringierten Modells sukzessive Parameter für die Gruppen gemeinsam geschätzt (vgl. Steenkamp & Baumgartner, 1998). In einem ersten Schritt erfolgte die Schätzung eines Baseline-Modells, in dem die Faktorstruktur für die Gruppen auf Gleichheit restringiert wurde, Faktorladungen und Mittelwerte der Indikatoren jedoch frei variieren konnten (*Modell 1: konfigurale Invarianz*). Im nächsten Schritt wurden neben der Faktorstruktur auch die Faktorladungen auf Gleichheit restringiert (*Modell 2: metrische Invarianz*) sowie anschließend zusätzlich die Mittelwerte der Indikatoren (*Modell 3: skalare Invarianz*). Die Verbesserung der Anpassungsgüte eines weniger restriktiven Modells kann über χ^2 -Differenzentests beurteilt werden. Die Messinvarianz gilt als bestätigt, wenn die weniger restriktiven Modelle keine bessere Modellanpassung als das Modell der skalaren Invarianz bieten. In diesem Fall können die latenten Mittelwerte sowie die Pfadkoeffizienten zwischen den Gruppen verglichen werden (vgl. Baumgartner & Steenkamp, 1998).

sowie deren SRMR und RMSEA unter .05 liegen, werden als sehr gute Approximation an die Daten bewertet (Hu & Bentler, 1999).

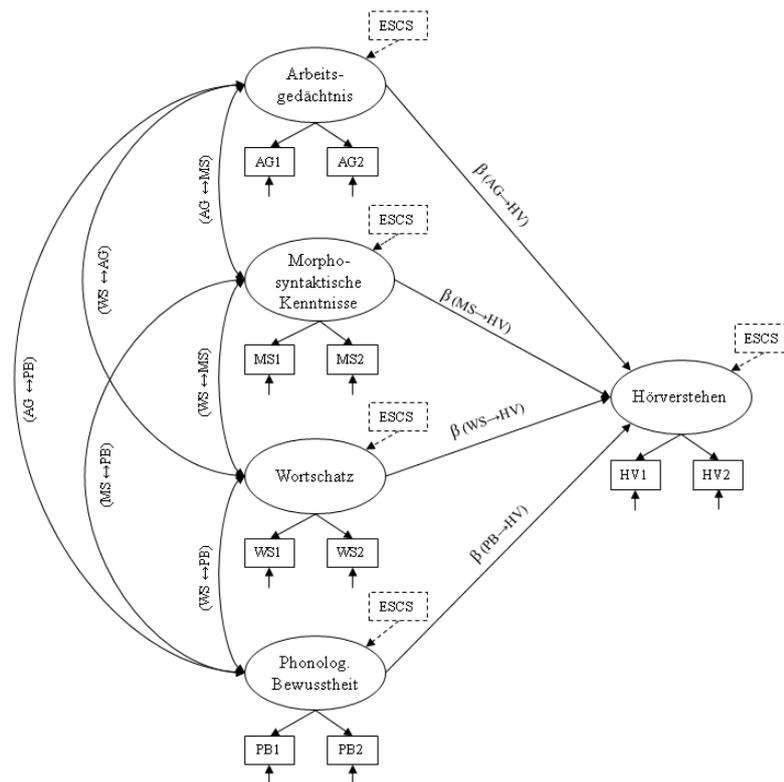


Abbildung 2. Strukturgleichungsmodell des Hörverstehens

HV = Hörverstehen, WS = Wortschatz, MS = Morphosyntaktische Kenntnisse, PB = Phonologische Bewusstheit, AG = Arbeitsgedächtnis, ESCS = Index des ökonomischen, sozialen und kulturellen Status.

Das Modell wird zunächst ohne Kovariaten geschätzt. In einem zweiten Schritt erfolgt eine Modellschätzung unter Einbezug der Kovariaten ESCS.

Fehlende Werte in den einzelnen Skalen wurden unter Einbezug aller Variablen mithilfe des Programms SPSS 18.0 multipel imputiert (vgl. Lüdtke, Robitzsch, Trautwein & Köller, 2007). Neben der χ^2 -Statistik wurde die Anpassungsgüte der Modelle anhand der Kennwerte CFI, TLI, RMSEA und SRMR bewertet. Modelle, deren TLI und CFI über .96

Zur inferenzstatistischen Überprüfung von Gruppenunterschieden auf latenter Ebene wurden ebenfalls Modellvergleiche mittels Wald- χ^2 -Tests vorgenommen, bei denen die interessierenden Parameter entweder über die Gruppen frei geschätzt oder gleich gesetzt wurden.

5. Ergebnisse

5.1 Deskriptive Befunde

Tabelle 2 enthält die z-standardisierten Mittelwerte und Standardabweichungen für alle Skalen differenziert nach dem Sprachhintergrund der Jugendlichen. Auf manifester Ebene zeigen sich gegenüber deutschsprachigen Jugendlichen signifikant schwächere Leistungen der Jugendlichen mit türkischer bzw. anderer Herkunftssprache im Hörverstehen, im Wortschatz

und in den morphologischen Kenntnissen. Im Hinblick auf den sozioökonomischen Status unterscheiden sich die drei Sprachgruppen ebenfalls signifikant voneinander. Jugendliche nichtdeutscher Herkunftssprache weisen einen geringeren sozioökonomischen Status auf als Jugendliche deutscher Herkunftssprache. Auch innerhalb der Gruppe von Jugendlichen nichtdeutscher Herkunftssprache zeigen sich signifikante Unterschiede: Schülerinnen und Schüler türkischer Herkunftssprache stammen aus Familien mit einem niedrigeren sozioökonomischen Status als Jugendliche mit anderer Herkunftssprache.

Tabelle 2.

z-standardisierte Mittelwerte und Standardabweichungen der Skalen in Abhängigkeit vom Sprachhintergrund der Jugendlichen

Instrument	N	deutschsprachig		türkischsprachig		anderssprachig	
		M	SD	M	SD	M	SD
HV Hörverstehen	417	.32 _{b,c}	.87	-.29 _a	.98	-.02 _a	1.08
PB Lautkategorisierung	363	-.08	1.00	.02	.97	.08	1.05
PB Vokallängenbestimmung	373	.10	1.00	-.14	.98	.06	1.01
WS Wortschatztest	392	.38 _{b,c}	.78	-.38 _{a,c}	1.06	.01 _{a,c}	.99
MS Syntaktische Kenntnisse	382	.09	1.00	-.06	.88	-.05	1.16
MS Morphologische Kenntnisse	388	.25 _{b,c}	.92	-.17 _a	.97	-.10 _a	1.09
AG Ziffernsprache vorwärts	373	-.02	.92	.09	1.07	-.12	1.01
AG Ziffernsprache rückwärts	360	.00	1.02	-.01	.96	.02	1.05
KFT Kognitive Grundfähigkeiten	390	.13	.98	-.14	1.00	.07	1.01
ESCS Index des ökonomischen, sozialen und kulturellen Status	328	.35 _{b,c}	-.85	-.35 _{a,c}	1.08	.03 _{a,b}	.89

Anmerkungen. HV = Hörverstehen; PB = Phonologische Bewusstheit; WS = Wortschatz; MS = Morphosyntaktische Kenntnisse; AG = Arbeitsgedächtnis; KFT = Kognitive Grundfähigkeiten; ESCS = Index des ökonomischen, sozialen und kulturellen Status.

z-standardisierte Mittelwerte mit unterschiedlichen Indizes (a, b, c) unterscheiden sich mit $p < .05$ statistisch signifikant. a = Unterschied zu deutschsprachigen Jugendlichen signifikant; b = Unterschied zu türkischsprachigen Jugendlichen signifikant; c = Unterschied zu anderssprachigen Jugendlichen signifikant.

5.2 Prüfung auf Messinvarianz

Die Ergebnisse der Modellvergleiche zur Überprüfung auf Messinvarianz zwischen den Gruppen sind in Tabelle 3 aufgeführt. Wie die χ^2 -Statistik und die Modellfitmaße zeigen, weisen alle Modelle eine zufriedenstellende Anpassungsgüte auf. Der χ^2 -Differenzentest belegt, dass sich die Anpassungsgüte der Gruppenmodelle trotz sukzessiv zunehmender Restriktionen nicht bedeutsam verschlechtert, so dass sowohl für das Modell ohne Kovariate als auch für das Modell mit ESCS als Kovariate von Messinvarianz über die Gruppen ausgegangen werden kann.

Tabelle 3.

Überprüfung der Mehrgruppenmodelle auf Messinvarianz

Invarianz-Modell	χ^2	df	p	RMSEA	CFI	TLI	SRMR	Vergleich mit ...	$\Delta \chi^2$	df	p
Ohne Kovariate											
A: konfigural	81.02	75	.30	.02	>.99	>.99	.03	--	--	--	--
B: metrisch	89.37	85	.35	.02	>.99	>.99	.03	Modell A	8.35	10	.59
C: skalar	93.62	95	.32	<.01	>.99	>.99	.03	Modell B	4.25	10	.94
ESCS als Kovariate											
A: konfigural	90.09	90	.48	<.01	>.99	>.99	.03	--	--	--	--
B: metrisch	98.43	100	.53	<.01	>.99	>.99	.03	Modell A	8.34	10	.60
C: skalar	102.72	110	.68	<.01	>.99	>.99	.03	Modell B	4.29	10	.93

Anmerkungen. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; SRMR = Standardized Root Mean Square Residual.

5.3 Multipler Gruppenvergleich

Abbildung 2 stellt das den Gruppenvergleichen zugrundeliegende Strukturgleichungsmodell dar. Dabei wurde zunächst ein Modell ohne Einbezug von Kovariaten geschätzt, anschließend ein Modell, in das der ESCS als manifeste Kovariate einging. Tabelle 4 zeigt die standardisierten Mittelwerte, Interkorrelationen und Pfadgewichte der latenten Variablen für beide Modelle getrennt für Schülerinnen und Schüler deutscher, türkischer und anderer Herkunftssprache.

Tabelle 4.

Standardisierte Modellparameter der messinvarianten Mehrgruppenmodelle (Modell C)

	Ohne Kovariate			ESCS als Kovariate		
	deutsch- sprachig (n=157)	türkisch- sprachig (n=161)	anders- sprachig (n=106)	deutsch- sprachig (n=157)	türkisch- sprachig (n=161)	anders- sprachig (n=106)
Mittelwerte der latenten Variablen						
α (HV)	0.00 _b	-0.37 _a	-0.08	0.00 _b	-0.43 _a	-0.05
α (WS)	0.00 _{b,c}	-0.74 _a	-0.43 _a	0.00 _{b,c}	-0.56 _a	-0.40 _a
α (MS)	0.00 _{b,c}	-0.35 _a	-0.28 _a	0.00	-0.24	-0.20
α (PB)	0.00	-0.17	0.07	0.00	-0.16	0.07
α (AG)	0.00	0.04	-0.03	0.00	0.00	0.05
Pfadgewichte der latenten Variablen						
β (WS \rightarrow HV)	.27*	.27*	.34*	.26*	.29*	.34*
β (MS \rightarrow HV)	.34*	.27*	.54*	.35*	.27*	.53*
β (PB \rightarrow HV)	.19	.31*	.03	.19	.31*	.03
β (AG \rightarrow HV)	-.03	.02	-.09	-.03	.01	-.09
Interkorrelationen der latenten Variablen						
(WS \leftrightarrow MS)	.54*	.57*	.59*	.54*	.55*	.59*
(WS \leftrightarrow PB)	.25*	.27*	.38*	.23*	.27*	.37*
(WS \leftrightarrow AG)	.01	.15	.19	.02	.18	.19
(MS \leftrightarrow PB)	.56*	.53*	.48*	.56*	.53*	.48*
(MS \leftrightarrow AG)	.33*	.26*	.41*	.34*	.28*	.42*
(PB \leftrightarrow AG)	.25 _c *	.30*	.51 _a *	.26 _c *	.31*	.51 _a *

Anmerkungen. HV = Hörverstehen; PB = Phonologische Bewusstheit; WS = Wortschatz; MS = Morphosyntaktische Kenntnisse; AG =Arbeitsgedächtnis.

Parameter mit unterschiedlichen Indizes (a, b, c) unterscheiden sich mit $p < .05$ statistisch signifikant. a = Unterschied zu deutschsprachigen Jugendlichen signifikant; b = Unterschied zu türkischsprachigen Jugendlichen signifikant; c = Unterschied zu anderssprachigen Jugendlichen signifikant.

* $p < .05$

5.3.1 Überprüfung auf Niveauunterschiede in der Ausprägung der Determinanten

Ohne Kontrolle der Kovariate ergeben sich für beide Gruppen von Jugendlichen nicht-deutscher Herkunftssprache im Vergleich zu Jugendlichen deutscher Herkunftssprache ein geringerer Wortschatz sowie geringere morphosyntaktische Kenntnisse. Zusätzlich findet sich bei türkischsprachigen Jugendlichen eine im Vergleich zu deutschsprachigen Schülerinnen

und Schülern schwächere Hörverstehensleistung. Die Leistungen der drei Sprachgruppen in der phonologischen Bewusstheit und im Arbeitsgedächtnis unterscheiden sich hingegen wie erwartet nicht überzufällig voneinander (vgl. Tabelle 4).

Unter Kontrolle des ESCS bleiben die Nachteile der Jugendlichen nichtdeutscher Herkunftssprache im Bereich des Wortschatzes bestehen. Auch zeigt sich weiterhin eine im Vergleich zu deutschsprachigen Jugendlichen geringere Hörverstehensleistung der türkischsprachigen Jugendlichen. Die signifikanten Gruppenunterschiede in den morphosyntaktischen Kenntnissen verschwinden hingegen unter Kontrolle des ESCS (vgl. Tabelle 4).

5.3.2 *Interkorrelationen zwischen den Determinanten*

Ohne Kontrolle von sozialen Hintergrundvariablen zeigen sich in allen Sprachgruppen signifikante Korrelationen zwischen den betrachteten Determinanten des Hörverstehens. Lediglich Wortschatz und Arbeitsgedächtnis korrelieren in keiner der Sprachgruppen miteinander. Signifikante Gruppenunterschiede in der Stärke der Zusammenhänge finden sich nur für die Korrelation zwischen phonologischer Bewusstheit und Arbeitsgedächtnis, die in der Gruppe der anderssprachigen Schülerinnen und Schüler höher ausfällt als in der Gruppe der deutschsprachigen Jugendlichen. Die Berücksichtigung des ESCS als Kovariate verändert diese Ergebnisse nicht (vgl. Tabelle 4).

5.3.3 *Überprüfung auf Unterschiede in den Pfadkoeffizienten*

Ohne Einbeziehung der Kovariate zeigen sich in allen drei Sprachgruppen signifikante Effekte des Wortschatzes und der morphosyntaktischen Kenntnisse auf die Hörverstehensleistung. Zusätzlich findet sich in der Gruppe der Jugendlichen mit türkischer Herkunftssprache ein signifikanter Effekt der phonologischen Bewusstheit auf das Hörverstehen. Die Arbeitsgedächtnisleistung beeinflusst das Hörverstehen in keiner der Sprachgruppen. Statistisch bedeutsame Gruppenunterschiede in der Ausprägung der Pfadkoeffizienten lassen sich nicht belegen. Ebenso wenig finden sich Unterschiede in der Ausprägung der Pfadkoeffizienten innerhalb der Gruppen, d. h. die Effekte der einzelnen Determinanten auf das Hörverstehen scheinen annähernd gleich stark zu sein.

Insgesamt ist der Anteil der aufgeklärten Varianz der Hörverstehensleistung mit $R^2=.60$ in der Gruppe der anderssprachigen Jugendlichen deutlich höher als in der Gruppe der rein deutschsprachigen Jugendlichen ($R^2=.42$) und der türkischsprachigen Jugendlichen ($R^2=.47$). Die Unterschiede in der Varianzaufklärung lassen sich jedoch nicht gegen den Zufall absichern.

Unter Kontrolle des ESCS ändert sich das Ergebnismuster kaum (vgl. Tabelle 4). Es zeigen sich weiterhin in allen drei Gruppen signifikante Effekte des Wortschatzes und der morphosyntaktischen Kenntnisse auf die Hörverstehensleistung, bei türkischsprachigen Jugendlichen findet sich zudem ein Effekt der phonologischen Bewusstheit auf das Hörverstehen. Signifikante Gruppenunterschiede in der Bedeutsamkeit der Pfadkoeffizienten lassen sich auch unter Kontrolle des ESCS nicht finden. Auch die jeweiligen Anteile an aufgeklärter Varianz ändern sich durch Aufnahme des ESCS in das Modell kaum.

6. Diskussion

In der vorliegenden Untersuchung von Determinanten der Hörverstehensleistung in einer Erst- bzw. Zweitsprache finden sich entgegen der Erwartungen nur für jene Jugendlichen nichtdeutscher Herkunftssprache signifikante Schwächen im Bereich des Hörverstehens, die als Erstsprache Türkisch gelernt haben. Offenbar sind die Verstehensprobleme spezifisch für diese Sprachgruppe und treten nicht generell bei allen Schülerinnen und Schülern nichtdeutscher Herkunftssprache auf. Dieser Befund deckt sich mit Ergebnissen der PISA 2009-Studie, die für den Bereich der Lesekompetenz ebenfalls eine im Vergleich zu anderen Migrantengruppen besonders schwache Leistung türkischsprachiger Jugendlicher zeigte (Stanat et al., 2010). Vor dem Hintergrund der strukturellen Ähnlichkeit von Hör- und Leseverstehensprozessen kann somit vermutet werden, dass Schülerinnen und Schüler türkischer Herkunftssprache beim deutschsprachigen Textverstehen generell Probleme haben.

Allerdings lassen sich für die betrachteten Determinanten des Hörverstehens keine spezifischen Schwächen von Schülerinnen und Schülern mit türkischer Herkunftssprache erkennen. Im Hinblick auf die Arbeitsgedächtniskapazität sowie die phonologische Bewusstheit unterscheiden sich beide Gruppen von Jugendlichen mit nichtdeutscher Herkunftssprache erwartungskonform nicht von Jugendlichen mit deutscher Herkunftssprache. Wird der sozioökonomische Status kontrolliert, weisen Jugendliche unterschiedlicher Herkunftssprache auch vergleichbar ausgeprägte morphosyntaktische Kenntnisse auf. Dies entspricht zunächst nicht Befunden, die für Lernende einer Zweisprache ein schwächeres syntaktisches Wissen belegen als für Schülerinnen und Schüler in einer Erstsprache (da Fontoura & Siegel, 1995; Verhoeven, 1990). Allerdings wurde in diesen Studien der sozioökonomische Hintergrund der Schülerinnen und Schüler nicht kontrolliert, so dass der Effekt des sozialen Hintergrunds auf die syntaktischen Kenntnisse von L1- und L2- Lernenden bislang unklar blieb. Die vorliegende Studie weist darauf hin, dass Unterschiede in den morphosyntaktischen Kenntnissen zwischen Jugendlichen deutscher und Jugendlichen nichtdeutscher Herkunftssprache offenbar auf

den unterschiedlichen sozioökonomischen Hintergrund der Schülerinnen und Schüler zurückzuführen sind. Für den Wortschatz zeigen sich hingegen auch unter Kontrolle des sozioökonomischen Status die vorhergesagten Nachteile der beiden Gruppen nichtdeutscher Herkunftssprache im Vergleich zu Schülerinnen und Schülern mit deutscher Herkunftssprache. Obwohl ein Großteil der untersuchten Jugendlichen nichtdeutscher Herkunftssprache in Deutschland geboren ist und die Schullaufbahn bis zum Ende der Regelschulzeit in Deutschland durchlaufen hat, gelingt es diesen Jugendlichen anscheinend nicht, Wortschatzkenntnisse zu entwickeln, die denen von deutschsprachigen Jugendlichen entsprechen. Die Ergebnisse entsprechen somit Befunden früherer Studien, die eine schwache Wortschatzleistung bei Schülerinnen und Schülern in einer Zweitsprache belegten (Limbird, 2007; Verhoeven, 1990).

Hinsichtlich der Effekte sprachlicher Determinanten auf die Hörverstehensleistung lassen sich entgegen der Erwartungen keine signifikanten Gruppenunterschiede in der Ausprägung der jeweiligen Pfadkoeffizienten feststellen. Für alle drei Sprachgruppen haben sowohl die morphosyntaktischen Kenntnisse als auch der Wortschatz einen Effekt auf die Hörverstehensleistung. Entscheidend für das Hörverstehen sind demnach bei allen Jugendlichen die Faktoren, die auf der Ebene des Parsing benötigt werden, um durch die semantische und syntaktische Analyse eines Satzes eine bedeutungsbasierte Repräsentation seines Inhalts zu generieren.

In diesem Prozess scheint das Arbeitsgedächtnis bei allen Sprachgruppen eine weniger starke Rolle zu spielen als vermutet. Das Ausbleiben signifikanter Effekte ist möglicherweise auf den verwendeten basalen Hörverstehenstest zurückzuführen, der aus kurzen Hörtexten mit überwiegend einfacher lexikalischer und syntaktischer Struktur bestand. Für die Verarbeitung von einfachen Hörtexten werden weniger Arbeitsgedächtnisressourcen benötigt als bei längeren komplexeren Texten. Es ist daher zu vermuten, dass sich für einen längeren Hörverstehenstest, der vermehrt syntaktisch bzw. semantisch anspruchsvolle Sätze enthält, ein Effekt der Arbeitsgedächtniskapazität auf die Hörverstehensleistung nachweisen ließe.

Im Gegensatz zu Jugendlichen deutscher und anderer Herkunftssprache zeigt sich für Schülerinnen und Schüler mit türkischer Herkunftssprache auch ein Effekt der phonologischen Bewusstheit auf das Hörverstehen. Dies könnte auf die unterschiedliche phonologische Struktur der jeweiligen Erstsprache der Jugendlichen zurückzuführen sein. Auch fortgeschrittene L2-Lernende haben aufgrund der Tendenz, L1-Segmentierungsstrategien auch in einer Zweitsprache anzuwenden, häufig Schwierigkeiten mit der Wahrnehmung von Phonemen in der L2 (vgl. Cutler, Kim & Otake, 2006). Die türkische Sprache verfügt über eine klar abgrenzbare Silbenstruktur und enthält im Gegensatz zur deutschen Sprache kaum Konsonan-

tencluster, so dass einzelne Phoneme leicht zu identifizieren sind (Öney & Durgunoğlu, 1997). Darüber hinaus erfordert die morphologische Struktur des Türkischen die Beachtung der letzten Phoneme in den einzelnen Wörtern (Limbird & Stanat, 2006). Demgegenüber sind im Deutschen akzentuierte Silben und Wörter die Hauptinformationsträger, zwischen denen Intonationstäler liegen, die für eine schnelle Informationserfassung weniger zentral sind (Cauneau, 1992). Die Unterschiede in der phonologischen Struktur beider Sprachen und die resultierende Notwendigkeit der Beachtung größerer phonetischer Einheiten im Deutschen könnten dazu führen, dass Jugendliche türkischer Herkunftssprache der phonologischen Verarbeitung des Hörtextes verstärkt Aufmerksamkeit widmen müssen. Für anderssprachige Jugendliche tritt dieser Effekt möglicherweise nicht auf, da die phonologische Struktur der jeweiligen Erstsprache der des Deutschen ähnlicher ist. Aufgrund der sprachlichen Heterogenität der Gruppe der anderssprachigen Jugendlichen und der geringen Fallzahlen in den einzelnen Sprachgruppen kann dies in der vorliegenden Studie jedoch nicht überprüft werden.

Insgesamt scheint der Hörverstehensprozess bei Jugendlichen unterschiedlicher Herkunftssprache durch dieselben Faktoren bestimmt zu werden. Allerdings weisen Schülerinnen und Schüler nichtdeutscher Herkunftssprache gegenüber Jugendlichen mit deutscher Herkunftssprache Schwächen im Bereich des Wortschatzes auf. Obwohl der Wortschatz in allen Gruppen ein signifikanter Prädiktor der Hörverstehensleistung ist, scheinen die geringeren Wortschatzkenntnisse nur bei Jugendlichen türkischer Herkunftssprache im Vergleich zu deutschsprachigen Jugendlichen auch zu einer geringeren Hörverstehensleistung zu führen. Dass in dieser Gruppe neben den lexikalischen und morphosyntaktischen Kenntnissen auch die phonologische Bewusstheit einen Effekt auf das Hörverstehen hat, kann als Hinweis darauf gewertet werden, dass für Jugendliche mit türkischer Herkunftssprache auch basale Lautwahrnehmungsprozesse noch bedeutsam für das Hörverstehen sind. Möglicherweise interferiert diese Fokussierung auf kleinere sprachliche Einheiten mit der syntaktischen und semantischen Analyse der Hörinhalte und führt somit zu einer schwachen Hörverstehensleistung der türkischsprachigen Jugendlichen. Angesichts der im Vergleich zu anderssprachigen Jugendlichen geringeren Varianzaufklärung des Hörverstehens in der Gruppe der türkischsprachigen Jugendlichen ist zudem denkbar, dass weitere, in dieser Studie nicht erfasste Faktoren, wie etwa Aufmerksamkeit oder metakognitives Strategiewissen, die Hörverstehensleistung bei türkischsprachigen Jugendlichen bedingen (Goh, 1997; Vandergrift, 2007).

Bei der Interpretation der Befunde ist anzumerken, dass aufgrund der Anlage der DELKO-Studie nur Schülerinnen und Schüler mit schwachen Leseverstehensleistungen untersucht wurden. Die Ergebnisse der Studie sind daher nur für diese Gruppe von Schülerinnen

und Schülern gültig. Die Untersuchung erfolgte zudem im Rahmen einer Querschnittstudie, so dass trotz aller theoretisch begründeten Plausibilität keine Aussagen über kausale Mechanismen getroffen werden können. Dennoch verdeutlichen die vorliegenden Ergebnisse die Notwendigkeit einer gezielten Förderung des Hörverstehens von Schülerinnen und Schülern nichtdeutscher Herkunftssprache, insbesondere bei Jugendlichen mit türkischem Sprachhintergrund. Auf Basis der vorliegenden Befunde empfiehlt sich v. a. eine Förderung grundlegender Determinanten des Hörverstehens, wie etwa des Wortschatzes. Inwiefern darüber hinaus auch die Vermittlung elaborierterer Fähigkeiten wie z. B. kognitiver und metakognitiver Hörstrategien (vgl. Vandergrift, 1999) hilfreich sein kann, um die Hörverstehensleistung in einer Zweitsprache zu fördern, müssen weitere Studien zeigen.

Literatur

- Anderson, J. R. (1995). *Cognitive psychology and its implications*. New York: Freeman.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.
- Baumgartner, H. & Steenkamp, J.-B. E. M. (1998). Multi-group latent variable models for varying numbers of items and factors with cross-national and longitudinal applications. *Marketing Letters*, 9, 21-35.
- Bialystok, E., Luk, G. & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: interactions among languages and writing systems. *Scientific Studies of Reading*, 9, 43-61.
- Blau, E. K. (1990). The effect of syntax, speed, and pauses on listening comprehension. *TESOL Quarterly*, 24, 746-753.
- Bonk, W. J. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, 14, 14-31.
- Buck, G. (1994). The appropriacy of psychometric measurement models for testing second language listening comprehension. *Language Testing*, 11, 145-170.
- Call, M. E. (1985). Auditory short-term memory, listening comprehension, and the input hypothesis. *TESOL Quarterly*, 19, 765-781.
- Cauneau, I. (1992). Hören-Brummen-Sprechen. *Fremdsprache Deutsch*, 7, 28-30.
- Cutler, A., Kim, J. & Otake, T. (2006). On the limits of L1 influence on non-L1 listening: Evidence from Japanese perception of Korean. In P. Warren & C. I. Watson (Eds.), *Proceedings of the 11th Australian International Conference on Speech Science & Technology* (pp. 106-111).
- Da Fontoura, H. A. & Siegel, L. S. (1995). Reading, syntactic, and working memory skills of bilingual Portuguese-English Canadian children. *Reading & Writing: An Interdisciplinary Journal*, 7, 139-153.
- Danks, J. H. & End, L. J. (1987). Processing strategies for reading and listening. In R. Horowitz & S. J. Samuels (Eds.), *Comprehending oral and written language* (pp. 271-293). San Diego: Academic Press.
- Eckhardt, A. G. (2008). *Sprache als Barriere für den schulischen Erfolg: Potentielle Schwierigkeiten beim Erwerb schulbezogener Sprache für Kinder mit Migrationshintergrund*. Münster: Waxmann.
- Field, J. (2003). Promoting perception: Lexical segmentation in L2 listening. *ELT Journal*, 57, 325-334.

- Fishman, M. (1980). We all make the same mistakes: A comparative study of native and non-native errors in taking dictation. In J. W. Oller & K. Perkins (Eds.), *Research in Language Testing* (pp. 187-197). Rowley, MA: Newbury House.
- Flowerdew, J. & Miller, L. (2005). *Second language listening: Theory and practice*. Cambridge: Cambridge University Press.
- Goh, C. (1997). Metacognitive awareness and second language listeners. *ELT Journal*, 51, 361-369.
- Goh, C. (2000). A cognitive perspective on language learner's listening comprehension problems. *System*, 28, 55-75.
- Grotjahn, R. (2005). Testen und Bewerten des Hörverstehens. In M. O Duill, R. Zahn & K. D. C. Höppner (Hrsg.), *Zusammenarbeiten: Eine Festschrift für Bernd Voss* (S. 115-144). Bochum: AKS-Verlag.
- Hasan, A. S. (2000). Learners' perceptions of listening comprehension problems. *Language, Culture and Curriculum*, 13, 137-153.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision: KFT 4-12+ R*. Göttingen: Beltz Test.
- Hu, L.-T. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Jansen, H. (1992). *Untersuchungen zur Entwicklung lautsynthetischer Verarbeitungsprozesse im Vorschul- und frühen Grundschulalter*. Egelsbach: Hänsel-Hohenhausen.
- Kahn-Horwitz, J. & Shimron, J. (2006). Weak and strong novice readers of English as a second language: Effects of first language and socio-economic status. *Annals of Dyslexia*, 56, 161-185.
- Kelly, P. (1991). Lexical ignorance: the main obstacle to listening comprehension with advanced foreign language learners. *IRAL - International Review of Applied Linguistics in Language Teaching*, 29, 135-149.
- Leibert, D. (2001). *Zur Erfassung der phonologischen Informationsverarbeitung älterer Schulkinder: Entwicklung eines Tests zur phonologischen Bewusstheit für die fünfte bis neunte Jahrgangsstufe*. Unveröffentlichte Diplomarbeit, Psychologisches Institut der Universität Heidelberg.
- Lesaux, N. K., Lipka, O. & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing: An Interdisciplinary Journal*, 19, 99-131.

- Leucht, M., Retelsdorf, J., Möller, J. & Köller, O. (2010). Zur Dimensionalität rezeptiver Kompetenzen im Fach Englisch. *Zeitschrift für Pädagogische Psychologie*, 24, 123-138.
- Limbird, C. (2007). *Phonological processing, verbal abilities, and second language literacy development among bilingual Turkish children in Germany*. Freie Universität Berlin, Berlin.
- Limbird, C. & Stanat, P. (2006). Prädiktoren von Leseverständnis bei Kindern deutscher und türkischer Herkunftssprache: Ergebnisse einer Längsschnittstudie. In A. Ittel & H. Merkens (Hrsg.), *Veränderungsmessung und Längsschnittstudien in der empirischen Erziehungswissenschaft* (S. 93-123). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Little, T. D., Cunningham, W. A., Shahar, G. & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal*, 9, 151-173.
- Long, D. R. (1990). What you don't know can't help you. An exploratory study of background knowledge and second language acquisition. *Studies in Second Language Acquisition*, 12, 65-80.
- Lüdtke, O., Robitzsch, A., Trautwein, U. & Köller, O. (2007). Umgang mit fehlenden Werten in der psychologischen Forschung. Probleme und Lösungen. *Psychologische Rundschau*, 58, 103-117.
- Marx, A. & Stanat, P. (2009). *Entwicklung eines Hörverstehenstests für Jugendliche*. Poster präsentiert auf der 72. Tagung der Arbeitsgruppe für empirische pädagogische Forschung (AEPF), 23.-25. März 2009 in Landau.
- McElvany, N., Becker, M. & Lüdtke, O. (2009). Die Bedeutung familiärer Merkmale für Lesekompetenz, Wortschatz, Lesemotivation und Leseverhalten. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41, 121-131.
- Mecartty, F. H. (2000). Lexical and grammatical knowledge in reading and listening comprehension by foreign language learners of Spanish. *Applied Language Learning*, 11, 323-348.
- Muthén, L. K. & Muthén, B. O. (1998-2010). *Mplus 6.1*. Los Angeles, CA: Muthén & Muthén.
- O'Malley, J. M., Chamot, A. U. & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics*, 10, 418-437.
- OECD. (2006). *PISA 2006: Technical Report*. Paris: OECD.

- Öney, B. & Durgunoğlu, A. Y. (1997). Beginning to read in Turkish: A phonologically transparent orthography. *Applied Psycholinguistics*, 18, 1-15.
- Osada, N. (2004). Listening comprehension research: A brief review of the past thirty years. *Dialogue*, 3, 53-66.
- Royer, J. M. & Carlo, M. S. (1991). Transfer of comprehension skills from native to second language. *Journal of Reading*, 34, 450-455.
- Schwartz, M., Geva, E., Share, D. L. & Leikin, M. (2007). Learning to read in English as third language: The cross-linguistic transfer of phonological processing skills. *Written Language & Literacy*, 10, 25-52.
- Segerer, R., Marx, P. & Schneider, W. (2009). *Test zur Erfassung sprachstruktureller Kompetenzen bei Jugendlichen*. Unveröffentlichtes Testverfahren, Julius-Maximilians-Universität Würzburg.
- Shohamy, E. & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question type. *Language Testing*, 8, 23-40.
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31, 577-607.
- Stanat, P., Rauch, D. & Segeritz, M. (2010). Schülerinnen und Schüler mit Migrationshintergrund. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Hrsg.), *PISA 2009: Bilanz nach einem Jahrzehnt* (S. 200-230). Münster: Waxmann.
- Stanovich, K. E., Cunningham, A. E. & Freeman, D. J. (1984). Intelligence, cognitive skills and early reading progress. *Reading Research Quarterly*, 19, 278-303.
- Steenkamp, J.-B. E. M. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78-90.
- Stewart, M. R. (2004). Phonological awareness and bilingual preschoolers: Should we teach it and, if so, how? *Early Childhood Education Journal*, 32, 31-37.
- Stock, C. & Schneider, W. (in Vorb.). *BAKO 5-9: Basiskompetenzen für Lese-Rechtschreibleistungen*. Göttingen: Hogrefe.
- Vandergrift, L. (1999). Facilitating second language listening comprehension: Acquiring successful strategies. *ELT Journal*, 53, 168-176.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40, 191-210.
- Verhoeven, L. (1990). Acquisition of reading in a second language. *Reading Research Quarterly*, 25, 90-114.

Weiß, R. H. (2008). *Grundintelligenztest Skala 2, CFT 20 - Revision (CFT 20-R). Mit Wortschatz- und Zahlenfolgentest (WS/ZF)*. Göttingen: Hogrefe.

Anhang B: Prädiktoren der Lesekompetenz in einer Erst- bzw. Zweitsprache (Teilstudie 2)

Marx, A., Stanat, P., Roick, T., Segerer, R., Marx, P. & Schneider, W. (2015). Components of reading comprehension in adolescent first-language and second-language students from low-track schools. *Reading and Writing: An Interdisciplinary Journal*, 28, 891-914. ⁵

<http://dx.doi.org/10.1007/s11145-015-9554-3>

Abstract. Based on the *Simple View of Reading*, the present study investigates whether component skills of reading differ between adolescents from low-track schools speaking German as a first (L1) versus as a second language (L2). In our study, 479 9th-grade students completed the *Programme for International Student Assessment (PISA)* reading test as well as measures of working memory, reading fluency, reading vocabulary, morpho-syntactic skills, phonological awareness, and listening comprehension. Results indicate that L2 students lagged behind their L1 peers on most measures. After controlling for students' socioeconomic status, only L2 students with Turkish as a first language continued to show weaknesses in linguistic skills. Linguistic comprehension as well as reading fluency were associated with reading comprehension in all groups of students. We observed only a few group differences in the relationships between component skills of reading and reading comprehension. The results suggest that reading fluency is important for reading comprehension of L1 and L2 readers from low-track schools and that socioeconomic background partially accounts for group differences in component skills of reading.

Keywords: Reading Comprehension, Simple View of Reading, PISA, L2, Socioeconomic Background, Reading Fluency

⁵ The final publication is available at <http://link.springer.com>

Reading skills are an essential prerequisite for lifelong learning and participation in almost all aspects of society. Findings from large-scale-assessment studies, such as the *Programme for International Student Assessment* (PISA), however, suggest that a substantial proportion of students fails to develop sufficient reading skills. In Germany, 19% of students did not reach proficiency level II in PISA 2009 which, for example, indicates that they were unable to identify the main idea of texts. Within this group of poor readers, students with an immigrant background are highly overrepresented (cf. OECD, 2010).

Recent research shows that L2 reading performance of students with immigrant background is associated with speaking the language of instruction at home. In the German PISA 2009 sample, not speaking the test language at home was linked with an achievement disadvantage of a quarter standard deviation on the composite reading literacy scale, even when the students were comparable in terms of social and educational background characteristics (OECD, 2010). The role language use at home seems to play beyond the effects of social and educational background suggests that a lack of opportunities for second-language (L2) acquisition impairs reading development in L2.

Findings like these certainly do not imply that students with immigrant background who speak their L1 at home are unable to achieve adequate reading skills. Successful literacy development in more than one language is possible and may be supported by a set of skills underlying both, L1 and L2 (Riches & Genesee, 2006). Yet, students need to reach some level of linguistic proficiency in L2 in order to become fluent readers in this language. The question, then, becomes whether certain linguistic components of reading may pose a special challenge for reading comprehension in L2. As a theoretical starting point, the following section describes a well-established model of reading comprehension and the component skills of reading comprehension it includes.

1. The Simple View of Reading

According to the *Simple View of Reading* (SVR) model, reading comprehension is determined by decoding and linguistic comprehension (Gough & Tunmer, 1986). The model defines decoding as efficient word recognition which strongly relies on students' phonological awareness. Linguistic comprehension represents the ability to process lexical information and to derive sentence or discourse interpretations (Hoover & Gough, 1990). Typically, linguistic comprehension is assessed with listening comprehension measures. These measures capture more basic aspects of linguistic comprehension as well, including vocabulary and syntactic skills (Gough & Tunmer, 1986; Hoover & Gough, 1990). The SVR assumes a mul-

tuplicative link between decoding and linguistic comprehension – if either skill is insufficiently developed, reading comprehension cannot be attained. In addition, Hoover and Gough (1990) suggest that decoding is the more important component at early stages of reading acquisition whereas linguistic comprehension becomes more important in later grades, when decoding is typically fluent and texts become increasingly demanding in terms of their linguistic properties. Various studies have lent support to the basic assumptions of the SVR for L1 and L2 students (e. g., Droop & Verhoeven, 2003; Gottardo & Mueller, 2009; Hoover & Gough, 1990).

Although the SVR does not clearly indicate how to measure decoding, most studies have operationalized this component skill with word-reading accuracy (Kirby & Savage, 2008). However, efficient decoding not only needs to be accurate but has to be fluent as well: If decoding is too slow, key information will have decayed by the time subsequent information is decoded (Kirby & Savage, 2008). Reading fluency, i.e. “the ability to read connected text rapidly, smoothly, effortlessly, and automatically with little conscious attention to the mechanics of reading such as decoding” (p. 284; Meyer & Felton, 1999) should therefore constitute an important predictor of reading comprehension. This should be especially the case at the later stages of reading acquisition when students already master accurate and automatized word decoding and can allocate cognitive resources to fluent text reading (Wolf & Katzir-Cohen, 2001). Evidence from a number of studies with L1 students suggests that text reading fluency does indeed have a positive relationship with reading comprehension at the later stages of reading acquisition (Cutting & Scarborough, 2006; Jenkins, Fuchs, van den Broek, Espin, & Deno, 2003) and that this association is stronger than the relationship between word reading fluency and reading comprehension (Jenkins et al., 2003). It thus seems reasonable to use measures of reading fluency rather than word-level decoding measures in studies investigating the SVR in older readers (e.g., Savage, 2006).

A number of researchers suggest that the achievement levels for some component skills of reading comprehension as well as their relationships with reading comprehension may differ between L1 and L2 learners. The next section discusses theoretical considerations underlying this assumption and the available evidence for the predicted differences.

2. Differences and Similarities in L1 and L2 Reading

2.1 *Component Skills of Reading in L1 and L2*

Reading in an L2 poses specific challenges that are, for the most part, related to linguistic comprehension. As the time-on-task hypothesis posits, the more exposure children have to

a specific language, the greater the likelihood that they will become proficient in that language (Gathercole, 2002). Children who do not speak the L2 at home have less exposure to this language and therefore limited opportunities to develop linguistic skills in L2. In line with this assumption, research has consistently shown that, for each of the languages involved, the vocabulary of L2 speakers tends to be smaller than that of L1 speakers of the respective languages, although L2 students' combined lexicon in both languages may exceed monolingual students' L1 lexicon (Bialystok, 2001; Lesaux & Kieffer, 2010). The morpho-syntactic awareness of bilingual students in their L2 also tends to be more limited than that of monolingual students. Lesaux, Lipka, and Siegel (2006), for example, showed that even five years of L2 immersion were insufficient for L2 elementary school children to develop syntactic skills comparable to those of monolingual children. L2 learners also tend to show distinct listening comprehension profiles: Research indicates that students with limited L2 proficiency predominantly try to decipher the meaning of an orally presented text based on isolated units, such as single words, and that this focus on lower-level processing may impair their listening comprehension (Verhoeven & van Leeuwe, 2012).

While skills associated with linguistic comprehension tend to be less developed in L2 than in L1, L2 learners may exhibit heightened levels of phonological awareness due to their contact with different phonological systems. As phonological awareness is an important prerequisite for reading fluency, this advantage may have positive effects on L2 learners' fluent reading (Ehri, 1995). Research indicates that L2 students do, indeed, show more acute phonological awareness than L1 students (Bialystok, Luk, & Kwan, 2005). Findings from a few studies show that text reading fluency of English language learners in elementary school years is close to national norms (Crosson & Lesaux, 2010) and does not differ from those of L1 students (Geva & Farnia, 2012). Thus, it seems that L2 students are likely to develop reading fluency comparable to L1 students after they have moved beyond the early stages of learning the L2.

2.2 Relationships Between Component Skills of Reading and Reading Comprehension in L1 and L2

Although it seems that phonological awareness and text reading fluency are, for the most part, comparable in L1 and L2 students, L2 students typically reach lower levels of reading comprehension (August & Shanahan, 2006). Based on this pattern of findings, it has been suggested that L2 students' disadvantages in linguistic comprehension should be the main cause of their reading difficulties. In line with this assumption, a number of studies in-

dicates that linguistic comprehension may be a more important predictor of reading comprehension in L2 than in L1. Findings by Droop and Verhoeven (2003) as well as Gottardo and Mueller (2009), for example, show that linguistic comprehension is a more important predictor of reading comprehension for L2 children in elementary school than for their L1 peers. Similarly, Geva and Farnia (2012) demonstrated that syntactic skills and listening comprehension predict L2 learners' reading comprehension in upper elementary grades, whereas no such effect was found for L1 students.

Text reading fluency seems to emerge as an important predictor of reading comprehension for both L1 and L2 learners in higher elementary grades (Geva & Farnia, 2012; Wiley & Deno, 2005). However, as results from a study by Wiley and Deno (2005) suggest, the relationship between text reading fluency and reading comprehension may be somewhat stronger for L1 students than for L2 students. Similar results were obtained by Buly and Valencia (2002) who reported that the association between text reading fluency and reading comprehension was weaker in L2 elementary school children than in L1 students. Crosson and Lesaux (2010) found text reading fluency to predict reading comprehension in fifth-grade English language learners, even after word reading fluency and oral language proficiency had been taken into account. However, text reading fluency seemed to influence reading comprehension only for those L2 students who had already developed good oral language skills. Based on such findings, several researchers suggest that students with limited language proficiency focus on word decoding rather than on fluent text reading (Buly & Valencia, 2002; Geva & Farnia, 2012; Wolf & Katzir-Cohen, 2001). Thus, the relationship between reading fluency and reading comprehension in L2 students may be moderated by the level of linguistic proficiency these students have reached.

To date, most research on reading comprehension and its components has been conducted with English language learners. For languages with deep orthographies, such as English or French, that do not have a one-to-one correspondence between sounds and the letters that represent them, accurate word recognition is relatively difficult. In languages with more transparent orthographies, like German, word recognition is automatized more quickly and reading fluency rather than reading accuracy may therefore be a better predictor of reading comprehension (Share, 2008).

Past studies that compared reading comprehension in struggling L1 and L2 readers mainly used word reading fluency as a measure of decoding and focused primarily on younger students in elementary grades (e. g. Lesaux & Kieffer, 2010; Lesaux et al., 2006). As text reading fluency may be a better predictor of reading comprehension in older L2 students and

as the relationship between text reading fluency and reading comprehension seems to be shaped by students' language proficiency, it is important to investigate the contribution of text reading fluency to reading comprehension in adolescent struggling readers as well.

Besides decoding and skills associated with linguistic comprehension working memory has also been found to be an important determinant of reading comprehension. According to Baddeley (1986), working memory entails a central executive supervisory system, a phonological subsystem that stores verbal and auditory information temporarily, and a visual-spatial subsystem that retains visual information. The central executive and phonological loop capacities have been shown to influence the development of reading comprehension as well as various of its component skills (e.g. Adams, Bourke, & Willis, 1999; Cain, Oakhill, & Bryant, 2004; Gathercole & Baddeley, 1989). Thus far, however, few studies have examined the impact of working memory on reading comprehension together with other component skills of reading. As the effect of working memory on reading comprehension may depend on other component skills of reading and vice versa (cf. Florit, Roch, & Levorato, 2014), further research is needed that simultaneously examines the impact of various components of reading comprehension.

3. Reading Comprehension, Language Background, and Socioeconomic Status

Studies investigating the impact of students' language background on their reading comprehension need to take potentially confounded variables into account that may contribute to this relationship. According to the *Componential Model of Reading* (Aaron, Joshi, Gooden, & Bentum, 2008), reading performance is influenced by the factors included in the SVR framework and psychological factors, yet also by factors from an ecological domain. These ecological factors include characteristics of students' home environment, such as the families' sociocultural status or the number of books in the home, but also factors from other ecological settings, such as school or social communities.

In line with the Componential Model of Reading, students' socioeconomic background has been found to be associated with language development and a variety of component skills of reading, such as vocabulary and listening comprehension (Hart & Risley, 1995). In a longitudinal study on the emergence of reading comprehension problems in L1 and L2, Kieffer (2010) found L2 students and students from low-SES homes to be overrepresented in the group of struggling readers. After students' SES was taken into account, the difference between L1 and L2 students' risk of becoming struggling readers was substantially reduced.

Kieffer (2010) concluded that L2 students' SES may be more predictive of reading difficulties than their status as L2 learners.

In most OECD countries, students with immigrant background have a more disadvantaged socioeconomic background than their native peers (OECD, 2010). In Germany, this difference is particularly pronounced (e.g., OECD, 2010). The lower SES of students with an immigrant background, then, may partially explain their lower performance levels in reading and other domains.

4. The Present Study

The present study builds on theoretical assumptions and empirical studies on cognitive components of reading comprehension included in the SVR. To date, most studies investigating reading comprehension of L2 students focused on elementary school students or older L2 students at the beginning stages of learning an L2. Little is known about the applicability of the SVR to adolescent L2 students who have completed compulsory schooling in their L2. This is the focus of the present study.

Furthermore, most studies on the SVR have used word reading accuracy or fluency as measures of decoding. Yet, the research evidence suggests that text reading fluency may be more relevant for adolescent students' reading comprehension as they have, for the most part, already developed fluent word reading skills. Following the suggestion by Savage (2006), we therefore used reading fluency as a measure for the decoding subcomponent in the SVR framework in our study. This is also important because research suggests that the impact of reading fluency on reading comprehension may differ between orthographies. We therefore examine whether results obtained for English language learners also apply to learners of a more transparent orthography like German.

The current study also addresses the need to isolate effects of language background from effects of socioeconomic background. A number of studies have already documented pronounced relationships between L1 and L2 students' socioeconomic background and their reading comprehension. However, most studies on reading in L1 and L2 that have taken students' socioeconomic background into account focused on a subset of component skills, such as decoding and vocabulary or listening comprehension. The present study adds to this research by simultaneously investigating multiple component skills of reading comprehension incorporated in the SVR as well as working memory, while controlling for socioeconomic background.

Thus, the study explores the extent to which the pattern of associations between component skills and reading comprehension depends on students' language background rather than SES, and it provides further insights into the nature of reading comprehension problems of older readers from low-track schools where the proportion of struggling readers is likely to be especially high. More specifically, it explores the following research questions and hypotheses for this student group:

1. Do L2 readers exhibit particularly pronounced weaknesses in component skills of reading comprehension? Based on the theoretical assumptions and previous research discussed above, we expect that L2 students reach lower performance levels than L1 students on measures of skills associated with linguistic comprehension, namely listening comprehension, vocabulary, and morpho-syntactic awareness. No group differences are expected for phonological awareness and reading fluency.
2. Can differences in reading comprehension and its components between German L1 and L2 readers be attributed to differences in their socioeconomic background? As socioeconomic background shapes the context for reading acquisition and is confounded with students' immigration background in Germany, we expect that differences in reading comprehension and its components between L1 and L2 students are largely accounted for by their families' socioeconomic status.
3. Are relationships between component skills of reading and reading comprehension equally pronounced in adolescent L1 and L2 students? Previous research on differences and similarities in relationships between linguistic comprehension and reading comprehension for L1 and L2 readers has yielded mixed results. As no sound theoretical assumptions predicting differences in the relationships between linguistic comprehension and reading comprehension exist, we expect these relationships to be equally pronounced in L1 and L2 students. Yet, differences between L1 and L2 students in the association of reading fluency with reading comprehension should emerge due to L2 students' limited language proficiency. More specifically, we expect that the relationship between reading fluency and comprehension is weaker in L2 students than in L1 students.

5. Method

5.1 Context

Although it is currently undergoing structural changes, at the time of data collection the German secondary school system was largely divided into three school types: *Hauptschule*

imparts a basic general education as the basis for vocational training, *Realschule* leads to an intermediary qualification, and *Gymnasium* leads to the qualification required for university attendance. The recommendation for track attendance at the end of elementary school is based on grades in all German federal states, but the extent to which this recommendation is binding varies.

Promoting language proficiency of students with immigrant background is considered to be a task of all teachers in all school subjects. First-generation students without knowledge of German sometimes take part in preparatory classes or intensive courses in German as a foreign language before they attend regular classroom instruction. Second-generation immigrant students are typically placed in regular German classrooms. Students with immigrant background and native students who need to improve their German skills may also receive remedial teaching (*Förderunterricht*) outside regular school hours. The organization as well as the content of these remedial teaching classes varies considerably between schools. Bilingual approaches or approaches that support students with immigrant background in transferring their L1 skills to their L2 are not common in Germany (Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (KMK), 2013).

5.2 *Sample*

The data of this investigation were collected in the research project “Determinanten der Lesekompetenz (DELKO)” (*Determinants of Reading Comprehension*). The cross-sectional DELKO project explored reading comprehension and its components in students with German as a first and students with German as a second language. As children of Turkish immigrants form the largest group of L2 students in Germany, the main focus of the study was on monolingual German and Turkish-German bilingual students. In addition, we included a group of L2 students with another first language than Turkish in the analyses to determine whether the results generalize across different second languages. Because the study focuses on adolescents with low levels of reading achievement, it was mainly conducted at schools from the lowest track of the German educational system (*Hauptschulen*) where the number of students with poor reading skills is especially high. All schools were public schools with 60% or more students with an immigration background. Most schools (59%) offered remedial support and 41% offered preparatory courses in German for students with another first language than German. Only one school offered lessons that were held in students’ L1. Of all

schools, 35% reported to assign L2 students to small classes such that the teacher-student ratio was lower for them.

Overall, 479 students (44.9 % female) from 20 schools took part in the study. All students attended the 9th grade, and their mean age was 16;1 years ($SD = 0.77$). Students' first language was determined with a single item in a questionnaire (“Which language did you learn first within your family?”). Based on their answers, we divided participants into three language groups. A total of 179 (37.4%) students reported German as their L1 (L1 German), 186 (38.8%) had Turkish as L1 ($L2_{\text{Turkish}}$), and 114 (23.8%) indicated that they speak another L1 than German or Turkish ($L2_{\text{Other}}$). The most frequent first languages reported in the latter group were Italian (11.4%), Arabic (11.4%), and Polish (9.6%). The majority of L2 students was born in Germany (86% for $L2_{\text{Turkish}}$ and 61% of $L2_{\text{Other}}$ students; $\chi^2_{(1)} = 20.73, p = <.01$) and most L2 students reported that they acquired German language skills before entering school (93% in the $L2_{\text{Turkish}}$ and 80% in the $L2_{\text{Other}}$ group, respectively). The majority of L2 students indicated that they presently spoke another language than German at home (70.5% of $L2_{\text{Turkish}}$ students and 60.3% of $L2_{\text{Other}}$ students). This proportion did not differ between the two groups of L2 students ($\chi^2_{(1)} = 2.04, p = .20$). Only a few students reported to receive remedial support in German (13.3% for L1 students, 22.5% for $L2_{\text{Turkish}}$ students, 26.1% for $L2_{\text{Other}}$ students; $\chi^2_{(2)} = 6.42, p = .04$).

The three language groups did not differ in terms of age ($F_{(2;421)} = -1.02, p = .36$) or gender distribution ($\chi^2_{(1)} = 1.44, p = .49$). We also found no group differences ($F_{(2;391)} = 2.63, p = .08$) for students' basic cognitive abilities as measured with a subtest that requires students to detect figural analogies and was taken from an intelligence test battery (KFT 4-12+ R; Heller & Perleth, 2000).

5.3 Instruments

5.3.1 Reading comprehension

We assessed students' reading comprehension with the PISA 2009 test. Students read several stimulus texts and subsequently answered a number of multiple-choice and short constructed-response questions about each text. In addition to the regular PISA booklets, the international PISA-consortium provided a so-called “easy set” of items that was anchored on the general PISA scale while allowing for a better differentiation in the lower range of reading performance.

The reading items were presented to students in thirteen test booklets with overlapping sets of items. The booklets contained 9 units and 31 reading items on average. Following the

PISA-procedure, students' responses were scaled with item response theory techniques (IRT). Five plausible values were calculated for each student and then used to estimate their reading comprehension scores (cf. OECD, 2012). The scores were anchored on the PISA metric, with an international mean of 500 and a standard deviation of 100. The reliability of the reading comprehension test based on the plausible values was .86.

5.3.2 *Listening comprehension*

To assess listening comprehension, we developed a computer-based test. Students listened to German texts each with a length of 100 to 150 words and varying between 0:42 min and 1:05 min in duration. All texts were listening passages specifically written for the test and spoken by native speakers of German. Four texts (50%) were dialogues between two persons (a male and a female), two texts were narratives delivered by one speaker (25%), and two texts were expository passages (25%) delivered by one speaker. After listening to each of the eight texts, students answered three to five items, resulting in a total of 32 questions. The questions were either in multiple-choice format with four response options (19 items) or in constructed-response format requiring a short written response (13 items). Items either targeted the texts' main idea, such as deciding on where a conversation may have taken place (17 items, 53%), or local details from the listening passage, such as remembering a provided definition (15 items, 47%). All items were presented acoustically and simultaneously in written form on the computer screen. In our sample, the listening comprehension test has an internal consistency of $\alpha = .78$.

5.3.3 *Phonological awareness*

To assess phonological awareness, we developed computer-based items assessing sound categorization and vowel length. The scale capturing *sound categorization* requires the identification of a word with different end sounds or middle sounds from a series of four pseudo-words (sample item: *Which of these words differs from the others in its ending sound? pat - kut - pit - pal*).

For the assessment of students' ability to determine *vowel length*, they were asked to identify a word with a longer/shorter vowel (in the instruction referred to as "sound") from a series of four pseudo-words (sample item: *In which of these words is the middle sound shorter or longer than in the others? maar - raas - dack - laa*). Overall, the phonological awareness test consisted of 27 items, with 12 items on sound categorization ($\alpha = .71$ for the current

sample) and 15 items on vowel length ($\alpha = .77$ for the current sample). The correlation between the subscales is $r = .39$ ($p < .01$).

5.3.4 Reading vocabulary

We assessed reading vocabulary using the CFT 20-R/WS, version A (Weiß, 2006). The subscale consists of 30 multiple-choice items which require students to find a correct synonym from a list of five words (sample item: *One of the following words matches the bolded word. Please tick off this word. **Moped** a) traffic b) driver's license c) motor vehicle d) engine e) bicycle.*) The tests' internal consistency for the current sample is $\alpha = .81$.

5.3.5 Working memory

Measures of working memory based on verbal material may be biased against L2 students because they are often less familiar with the words or language structures used in these measures (Olsthoorn, Andringa, & Hulstijn, 2012). We therefore applied visual digit spans to assess students' working memory. The test consisted of digits that were visually presented on a computer screen. To capture storage capacity of the phonological loop, students were instructed to reproduce a series of digit spans on the computer keyboard in the given order (*digit span forward*). Subsequently, students reproduced series of digits in reverse order (*digit span backward*) to assess central-executive functioning. Digits were presented for 1500 ms with an interval of 1000 ms in between. Overall, students worked on 12 digit spans forward and 12 digit spans backward, with tasks varying in length between 3 to 7 digits. The internal consistency was $\alpha = .66$ for *digit span forward* and $\alpha = .84$ for *digit span backward*. The correlation between the two subscales is $r = .44$ ($p < .01$).

5.3.6 Morpho-syntactic awareness

To assess students' morpho-syntactic awareness, we developed two computer-based scales. The first scale – a grammaticality judgment task (cf. Bialystok, 1979) – consisted of written sentences, the syntactical correctness of which students were asked to check. To control for guessing, students could choose between the answers “incorrect”, “correct”, and “unclear”, with the latter option indicating that they were uncertain in their judgment (sample item: *Is the following sentence grammatically correct or incorrect? “I to bed went late yesterday.” correct – incorrect – unclear*). “Unclear” ratings were coded as incorrect.

As the German language contains a large amount of inflectional morphemes, we also constructed a cloze test that captures *morphological awareness*. The scale is based on previ-

ous work by Siegel and Ryan (1988) and consists of written sentences that students were asked to complete with a morphologically matching word (sample item: “Now I’m _____ than yesterday.” *calm – calmer – calmest – calmly*). Overall, the test of morpho-syntactic awareness included 47 items, with 22 items in the grammaticality judgment task ($\alpha = .86$ for the current sample) and 25 items in the cloze test ($\alpha = .83$ for the current sample). The correlation between the subscales is $r = .59$ ($p < .01$).

5.3.7 Text reading fluency

To assess students’ text reading fluency, we used the *Lesegeschwindigkeits- und Verständnistest* (LGVT 6-12; Test of reading fluency and accuracy; Schneider, Schlagmüller, & Ennemoser, 2007) which requires students to silently read a text of 1.727 words. At 23 positions throughout the text, a word is missing and has to be selected from a list of three alternatives (sample item: “The giraffe is one of the world’s largest mammals, it may become as high as six [centimeters – meters – kilometers].” The number of words students had read after eight minutes was used as an outcome measure. To ensure that students reported the correct number of words they had read, we weighted this information with the correctness of selected words. For each of the 23 missing words, students were given 2 points if the correct word was chosen, 0 points if no word was chosen, and -2 points if a wrong alternative was chosen. The scale’s internal consistency is $\alpha = .72$.

5.3.8 Socioeconomic status

As the indicator of students’ socioeconomic status, we used the Index of Social, Economic, and Cultural Status (ESCS) that is also employed in the PISA studies (cf. OECD, 2012). The ESCS has a mean of 0 and a standard deviation of 1. This measure is a composite indicator comprised of families’ home possessions (such as books in the home, cars, televisions, computers), parental occupation as measured by the International Socioeconomic Index of Occupational Status, and parental education in years of schooling. If levels of parental occupation or education differ for mother and father, the highest level is used to indicate a families’ occupational or educational status. The data for these variables were obtained with student and parent questionnaires.

5.4 Analyses

To test the hypotheses of the present investigation, we employed structural equation modeling (SEM) using Mplus 7.11 (Muthén & Muthén, 1998-2012). For each of the latent

constructs included in the analyses, we constructed indicators using item parceling, with the residuals constrained to be equal (Little, Cunningham, Shahar, & Widaman, 2002). For the reading comprehension measure, we used the five plausible values as item parcels. For the test of text reading fluency, we entered the weighted score of words had read in eight minutes and a weighted score of words students had read in four minutes as parcels. For all other measures, items were odd-even parceled (cf. Kim & Hagtvet, 2003). We used SPSS 18.0 missing data analysis to ensure that missingness was at random which was confirmed. Missing data were then imputed using the SPSS 18.0 multiple imputation algorithm.

In a first step, we tested for measurement invariance to ensure that the measurement model could be applied to all groups of students (Baumgartner & Steenkamp, 1998). We initially estimated a baseline model, with the factorial structure set to be equal for the three language groups but factor loadings and means allowed to vary (*model I: configural invariance*). Subsequently, factor loadings (*model II: metric invariance*) and, finally, means were restricted to be equal across groups as well (*model III: scalar invariance*). We evaluated goodness of fit using a χ^2 -difference test. If a more restrictive model does not yield a poorer fit than a less restrictive model, measurement invariance can be assumed. In addition to the χ^2 -statistic, we also used the Comparative Fit Index (CFI), the Tucker-Lewis-Index (TLI), the Root Mean Square Error of Approximation (RMSEA), and the Standardized Root Mean Square Residual (SRMR) to evaluate goodness of fit. Models with a TLI and CFI higher than .95, a RMSEA smaller than .06, and a SRMR smaller than .08 are regarded as good approximations of the data (Hu & Bentler, 1999).

As a second step, we tested our first hypothesis on differences in L1 and L2 students' component skills of reading comprehension by comparing latent variable means among the language groups. More specifically, we compared models estimated with confirmatory factor analyses (CFA) using Wald- χ^2 -tests with the respective parameter either constrained to be equal or allowed to vary across groups. To disentangle effects of students' socioeconomic background from their language status, as necessary for testing the second hypothesis, we first estimated a model without covariates and subsequently included the ESCS as a manifest covariate in the analysis. The relationships between components of reading and reading comprehension predicted by our third hypothesis were investigated by comparing path coefficients for L1 and L2 students. Again, we used Wald- χ^2 -statistics to test whether a given path coefficient differed among groups.

6. Results

6.1 Descriptive Results

Table 1 shows descriptive statistics for all measures. Participants in the DELKO study showed relatively low levels of reading achievement. The reading comprehension mean in DELKO was $M = 396$ ($SD = 86$) which is one standard deviation below the German PISA average ($M = 497$, $SD = 95$; cf. Klieme et al., 2010). Results also indicate that students in the DELKO sample were mainly from families with low socioeconomic status. The ESCS mean in the DELKO sample was $M = -0.63$ ($SD = 0.92$), which is more than half a standard deviation below the German average (OECD, 2010). The three language groups differed in terms of their socioeconomic background ($F_{(2; 363)} = 12.09$; $p < .01$; $\eta^2 = .06$). Families' socioeconomic status was lower for L2_{Turkish} students ($M = -0.86$; $SD = 0.92$) than for L1 students ($M = -0.43$; $SD = 0.79$) and L2_{other} students ($M = -0.62$; $SD = 0.85$).

Table 1

Descriptive Data for all Measures

	<i>N</i>	<i>Percentage of missing data</i>	<i>min</i>	<i>max</i>	<i>M</i>	<i>SD</i>
ESCS (<i>WLE</i>)	364	24.0	-3.18	2.09	-0.63	0.92
RC (<i>M=500; SD=100</i>)	479	0.0	172	591	396	86
LC (<i>min=0; max=32</i>)	369	23.0	6	30	19.49	4.89
RF (<i>min=-79; max=79</i>)	369	23.0	-29	68	13.54	14.24
VO (<i>min=0; max=30</i>)	392	18.2	6	29	22.21	4.47
MS (<i>min=0; max=47</i>)	395	17.5	10	47	36.16	7.90
PA (<i>min=0; max=27</i>)	406	15.2	2	26	14.12	5.23
WM (<i>min=0; max=24</i>)	406	15.2	3	24	12.39	4.56

Note. ESCS = Index of Economic, Social, and Cultural Status; RC = Reading Comprehension; LC = Listening Comprehension; RF = Reading Fluency; VO = Reading Vocabulary; MS = Morpho-syntactic Awareness; PA = Phonological Awareness; WM = Working Memory.

6.2 Test of Measurement Invariance

The model we used to test our hypotheses is depicted in Figure 1. Table 2 shows the results for the test of measurement invariance across the three language groups. Goodness-of-fit indices as well as χ^2 -statistics indicate that all models fit the data well. The χ^2 -difference test shows that goodness-of-fit does not decrease for the more restrictive models, hence measurement invariance can be assumed. One caveat, however, is that the SRMR for the models did not meet the criteria proposed by Hu and Bentler (1999). Yet, given that the SRMR may be biased in small sample and the other fit indices were satisfactory, it seems justified to conclude that the general assumptions of the model are supported by the data.

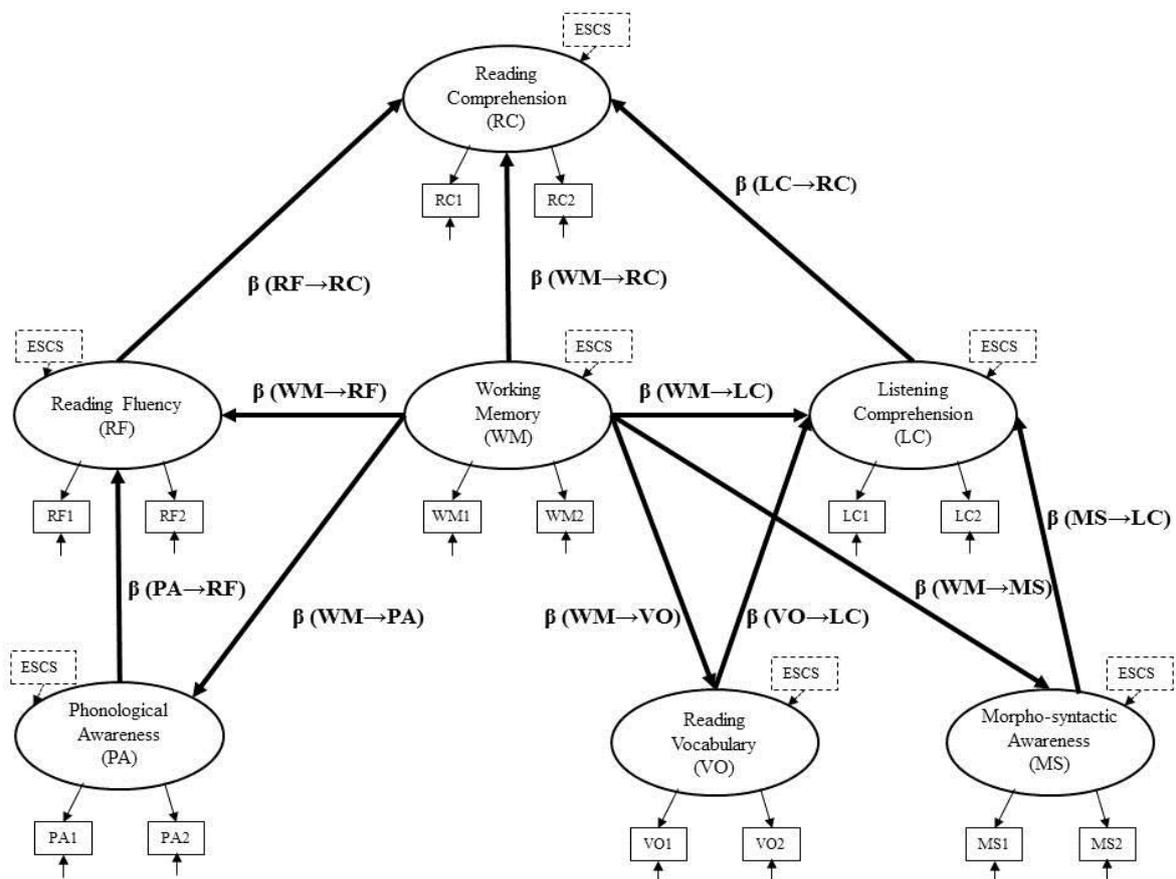


Fig. 1. Model of Reading Comprehension

Table 2.

Test of Measurement Invariance

Model	χ^2	df	p	RMSEA	CFI	TLI	SRMR	Compared to ...	$\Delta \chi^2$	df	p
<i>No Controls</i>											
A: configural	545.58	342	<.01	.06	.96	.95	.17	--	--	--	--
B: metric	552.50	362	<.01	.06	.96	.96	.17	model A	6.92	20	.99
C: scalar	573.69	382	<.01	.06	.96	.96	.17	model B	21.19	20	.39
<i>Controlling for ESCS</i>											
A: configural	556.81	372	<.01	.06	.96	.96	.17	--	--	--	--
B: metric	565.81	392	<.01	.05	.97	.96	.17	model A	9.00	20	.98
C: scalar	587.32	412	<.01	.05	.97	.96	.17	model B	21.52	20	.37

Note. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; SRMR = Standardized Root Mean Square Residual.

6.3 Differences in Latent Variable Means

In order to test the first hypothesis on differences and similarities in L1 and L2 learners' component skills of reading comprehension, we compared their latent variable means using a CFA. Table 3 shows the standardized model parameters for the three language groups. Latent variable means of the two L2 groups were first compared with the reference group of students with German as L1 for which we fixed the latent means to zero. Subsequently, L2_{Turkish} students were compared to L2_{Other} students, with latent means fixed to zero for the former group. The results indicate that both groups with German as L2 reached significantly lower levels of reading comprehension than German L1 students. Furthermore, all L2 students exhibited poorer results than German monolingual students for the skills associated with linguistic comprehension, namely reading vocabulary, morpho-syntactic skills, and listening comprehension. L2 students also performed less well than L1 students on the text reading fluency measure. No differences were found among the three groups on measures of working memory and phonological awareness.

Table 3.

Latent Variable Means

	L1 German (<i>n</i> = 179)	L2 Turkish (<i>n</i> = 186)	L2 Other (<i>n</i> = 114)	Group Comparisons					
				L1 German vs. L2 Turkish		L1 German vs. L2 Other		L2 Turkish vs. L2 Other	
				<i>Wald</i> χ^2 (<i>df</i> = 1)	<i>p</i>	<i>Wald</i> χ^2 (<i>df</i> = 1)	<i>p</i>	<i>Wald</i> χ^2 (<i>df</i> = 1)	<i>p</i>
<i>No Controls</i>									
α (RC)	0.00	-0.51	-0.29	18.98	<.01	6.08	.01	0.22	.64
α (LC)	0.00	-0.65	-0.33	34.83	<.01	8.02	<.01	0.16	.69
α (RF)	0.00	-0.56	-0.31	11.60	<.01	3.76	.05	0.03	.86
α (VO)	0.00	-0.66	-0.39	39.03	<.01	11.04	<.01	1.99	.16
α (MS)	0.00	-0.31	-0.29	5.31	.02	5.59	.02	1.29	.26
α (PA)	0.00	-0.12	0.06	3.03	.10	0.01	.93	0.91	.34
α (WM)	0.00	0.04	-0.06	0.56	.46	0.18	.68	0.63	.43
<i>Controlling for ESCS</i>									
α (RC)	0.00	-0.60	-0.27	18.23	<.01	3.52	.06	0.06	.80
α (LC)	0.00	-0.69	-0.17	30.71	<.01	0.88	.35	0.73	.39
α (RF)	0.00	-0.52	-0.23	5.22	.02	0.67	.35	0.01	.92
α (VO)	0.00	-0.52	-0.28	15.16	<.01	3.87	.05	1.01	.32
α (MS)	0.00	-0.20	-0.07	5.36	.01	0.22	.64	0.27	.60
α (PA)	0.00	-0.17	-0.17	0.73	.39	0.65	.42	0.08	.78
α (WM)	0.00	-0.05	0.02	0.01	.94	0.01	.93	0.01	.91

Note. ESCS = Index of Economic, Social, and Cultural Status; RC = Reading Comprehension; LC = Listening Comprehension; RF = Reading Fluency; VO = Reading Vocabulary; MS = Morpho-syntactic Awareness; PA = Phonological Awareness; WM = Working Memory.

6.4 Disentangling Effects of Students' Language Background and their Socioeconomic Background on Component Skills of Reading Comprehension

To determine whether the differences between L1 and L2 learners in reading comprehension and its components are accounted for by differences in socioeconomic background as predicted by the second hypothesis, we entered the ESCS as a manifest covariate into the analyses. Controlling for the ESCS caused most mean differences between L1 students and L2_{Other} students to disappear (Table 3). For the L2_{Other} group, only the difference to L1 readers in vocabulary remained significant. The L2_{Turkish} students, in contrast, continued to lag behind

their L1 peers in terms of reading comprehension and component skills associated with linguistic comprehension, namely listening comprehension, morpho-syntactic skills, and reading vocabulary, even after their socioeconomic background was controlled. No group differences were found between L2_{Turkish} and L2_{Other} students.

6.5 Structural Relationships Between Reading Comprehension and Component Skills of Reading

To investigate differences and similarities between L1 and L2 students in the structural relationships between reading comprehension and its components as predicted by the third hypothesis, we tested differences in the path coefficients among language groups. More specifically, we compared the models using Wald- χ^2 -tests, with the respective parameter either constrained to be equal across groups or allowed to vary. The model comparisons revealed only a few differences in structural relationships between the three language groups (Table 4). Differences were found for the relationship between text reading fluency and reading comprehension which was stronger in the L2_{Other} than in the L2_{Turkish} group. The associations between working memory on the one hand and morpho-syntactic awareness as well as phonological awareness on the other hand were also significantly stronger for L2_{Other} than for L2_{Turkish} students. Furthermore, the association between reading fluency and phonological awareness was stronger for L1 students than for L2_{Turkish} students (Table 4). All other relationships did not differ among the language groups. Entering students' ESCS as a covariate did not change the results substantially (Table 4). Only the association between morpho-syntactic awareness and working memory no longer differed between L2_{Other} and L2_{Turkish} students when ESCS was entered into the analysis.

For all groups, Wald- χ^2 -tests revealed that reading comprehension was more strongly related to listening comprehension than to reading fluency ($\chi^2_{(1)} = 7.37$; $p < .01$ for L1 German; $\chi^2_{(1)} = 13.32$; $p < .01$ for L2_{Turkish}; $\chi^2_{(1)} = 10.93$; $p < .01$ for L2_{Other}).

Table 4.

Standardized Parameters of the Multi-Group Model (Model C)

	L1 German (n = 179)	L2 Turkish (n = 186)	L2 Other (n = 114)	Group Comparisons						
				L1 German vs. L2 Turkish		L1 German vs. L2 Other		L2 Turkish vs. L2 Other		
				<i>Wald</i> χ^2 (<i>df</i> = 1)	<i>p</i>	<i>Wald</i> χ^2 (<i>df</i> = 1)	<i>p</i>	<i>Wald</i> χ^2 (<i>df</i> = 1)	<i>p</i>	
<i>No Controls</i>										
β (RF \rightarrow RC)	.33*	.19*	.35*	1.12	.29	0.90	.34	3.11	.08	
β (LC \rightarrow RC)	.57*	.61*	.52*	0.80	.37	1.86	.17	0.45	.50	
β (PA \rightarrow RF)	.41*	.20*	.32*	3.98	.05	0.50	.48	1.17	.28	
β (VO \rightarrow LC)	.40*	.33*	.30*	0.54	.46	1.02	.31	0.16	.69	
β (MS \rightarrow LC)	.56*	.50*	.66*	0.02	.89	0.59	.44	0.76	.38	
β (WM \rightarrow RC)	.08	.07	.14	0.38	.54	0.08	.78	0.86	.35	
β (WM \rightarrow LC)	-.08	.04	-.13	0.41	.52	0.13	.72	0.85	.36	
β (WM \rightarrow RF)	-.04	-.04	.13	0.04	.83	0.00	.99	0.05	.83	
β (WM \rightarrow VO)	.06	.15	.32*	0.91	.34	2.47	.12	0.50	.48	
β (WM \rightarrow MS)	.37*	.27*	.52*	0.37	.54	1.47	.23	3.14	.08	
β (WM \rightarrow PB)	.30*	.33*	.55*	0.61	.44	1.22	.27	3.96	.05	
R^2 (RC)	.47	.44	.52							

Table 4 (continued).

Standardized Parameters of the Multi-Group Model (Model C)

	L1 German (n=179)	L2 Turkish (n=186)	L2 Other (n=114)	Group Comparisons					
				L1 German vs. L2 Turkish Wald χ^2 (df=1)	p	L1 German vs. L2 Other Wald χ^2 (df=1)	p	L2 Turkish vs. L2 Other Wald χ^2 (df=1)	p
<i>Controlling for ESCS</i>									
β (RF \rightarrow RC)	.32*	.20*	.37*	0.80	.37	1.44	.23	3.58	.05
β (LC \rightarrow RC)	.56*	.61*	.52*	0.70	.44	1.46	.23	0.31	.48
β (PA \rightarrow RF)	.40*	.20*	.33*	3.97	.05	0.45	.50	1.30	.25
β (VO \rightarrow LC)	.40*	.36*	.31*	0.21	.64	0.96	.33	0.38	.54
β (MS \rightarrow LC)	.55*	.50*	.64*	0.06	.81	0.44	.51	0.75	.39
β (WM \rightarrow RC)	.08	.07	.13	0.49	.53	0.06	.80	0.82	.36
β (WM \rightarrow LC)	-.08	.02	-.12	0.30	.59	0.09	.76	0.62	.43
β (WM \rightarrow RF)	-.04	-.02	.11	0.00	.95	0.01	.92	0.00	.95
β (WM \rightarrow VO)	.06	.18	.30*	1.41	.23	2.18	.14	0.17	.68
β (WM \rightarrow MS)	.38*	.29*	.50*	0.27	.60	1.08	.30	2.31	.13
β (WM \rightarrow PB)	.31*	.33*	.55*	0.52	.47	1.18	.28	3.63	.05
R^2 (RC)	.48	.44	.53						

Note. ESCS = Index of Economic, Social, and Cultural Status; RC = Reading Comprehension; LC = Listening Comprehension; RF = Reading Fluency; VO = Reading Vocabulary; MS = Morpho-syntactic Awareness; PA = Phonological Awareness; WM = Working Memory.

*p < .05

7. Discussion

The present study investigated component skills of reading comprehension in L1 and L2 adolescents from low-track schools. Based on theoretical assumptions and previous research, we expected that L1 readers would outperform L2 readers on component skills of reading associated with linguistic comprehension but not on phonological awareness and reading fluency (Hypothesis 1). We further expected that differences in socioeconomic background as an ecological factor would at least partially account for the group differences in reading comprehension and its components (Hypothesis 2). The associations between linguistic comprehension and reading comprehension were predicted to be comparable for L1 and

L2 students, yet we expected reading fluency to be more strongly associated with reading comprehension in L1 students than in L2 students (Hypothesis 3).

In line with the first hypothesis, the results of the present study indicated that, without controlling for students' socioeconomic background, adolescent L2 readers lagged behind their monolingual peers on measures associated with linguistic comprehension, namely listening comprehension, reading vocabulary, and morpho-syntactic awareness, but performed similarly to L1 students on measures of working memory and phonological awareness. Contrary to our expectation, we found that both groups of L2 students performed less well than their L1 peers on the reading fluency test. Yet, L1 and L2 students differed in terms of their socioeconomic background which has been shown to be a strong predictor of decoding (e.g. Noble, Farah, & McCandliss, 2006). In line with this general finding, text reading fluency was associated with socioeconomic background in our study, and the performance differences in reading fluency between L1 and L2_{Other} readers disappeared when their socioeconomic background was entered into the analysis as a control variable. L2_{Turkish} students, however, continued to show weaker performance on the reading fluency test than L1 students after we controlled for ESCS. This persistent performance disadvantage of the L2_{Turkish} group could be attributable to the relationships between text reading fluency and linguistic proficiency: the higher students' linguistic proficiency, the more they are able to rely on linguistic knowledge rather than on word recognition when processing text (Geva & Farnia, 2012; Wolf & Katzir-Cohen, 2001). The weaker performance of the L2_{Turkish} group on the measure of text reading fluency may thus result from their less developed linguistic proficiency they showed in our study.

With regard to our second hypothesis, we found that controlling for socioeconomic background caused most differences between L1 and L2_{Other} students' component skills of reading comprehension to disappear. As socioeconomic background and performance were correlated and L2_{Other} students in our sample were from families with a more disadvantaged socioeconomic background than L1 students, it seems likely that the performance gap between these groups reflects this difference. L2_{Turkish} readers, in contrast, showed persistent weaknesses in component skills of reading comprehension after socioeconomic background was taken into account. Although the Turkish-speaking students were, for the most part, born and raised in Germany, they apparently did not develop language skills in their L2 comparable to those of L1 students with a similar socioeconomic background. This finding is consistent with results from previous studies documenting especially weak language skills in Turkish-speaking L2 learners as compared to other L2 learners even when their socioeco-

conomic background was held constant (Droop & Verhoeven, 2003; Stanat, Rauch, & Segeritz, 2010).

To date, the cause for L2_{Turkish} students' particularly low educational outcomes is unclear and a variety of linguistic but also social factors as well as their interplay may cause this pattern (Glock & Krolak-Schwerdt, 2013). One possible explanation is that the German educational system disadvantages L2_{Turkish} students. Among the factors that might be of importance in this context are teachers' expectations which may be lower for L2_{Turkish} students and could affect the development of students' competences as well as educational transitions. A number of studies indicate that teachers' expectations tend to be lower for children from families with low socioeconomic status and families with immigrant background even when their performance is taken into account (e.g., Jussim & Harber, 2005; Ready & Wright, 2011). Negative stereotypes against certain groups of students have been discussed as an important source of such biased expectations (Jussim & Harber, 2005; Nelson, 2002). Stereotypes are most likely to affect teacher judgments when they have limited information about the performance of a given student (Fiske, Lin, & Neuberg, 1999). Inaccurate teacher expectations are then assumed to result in self-fulfilling prophecies (Jussim & Harber, 2005). In Germany, only a few studies have investigated the negative stereotypes about students with immigration background and differential teacher expectations. The available findings indicate that negative stereotypes against Turkish immigrants in Germany exist (e.g., Kahraman & Knoblich, 2000). Results from two experimental studies further suggest that, due to such stereotypical expectations, teachers' judgments are negatively biased against students from Turkish immigrant families (Glock & Krolak-Schwerdt, 2013; Sprietsma, 2013). Similar results were obtained in a recent field study that also documented systematic bias in teacher expectations against children from Turkish immigrant families (Lorenz, Gentrup, Kristen, Stanat, & Kogan, submitted). It is therefore conceivable that lower teacher expectancies for children with Turkish immigrant background may, at least partially, account for the lower educational outcomes of this student group. Yet, direct tests of this assumption are not available yet.

Another possible explanation for L2_{Turkish} students' poorer performance compared to L2_{Other} students is that Turkish immigrants tend to have less frequent second-language contacts than other immigrant groups (Backus, 2005). The time-on-task hypothesis maintains that the more children are exposed to a given language, the more likely they are to develop proficiency in this language (Gathercole, 2002), although this effect is likely to also depend on the quality of the input (Hart & Risley, 1995). L2_{Turkish} students' disadvantages in im-

portant L2 linguistic skills may thus also result from their limited exposure to the L2. As L2_{Turkish} adolescents in our sample were, for the most part, enrolled in German schools since the beginning of formal schooling, they should have had ample exposure to their L2. However, the quality of L2 instruction currently implemented in Germany may not match these students' needs, as several authors have suggested (e.g., Becker & Beck, 2011; Gogolin, 2005; Söhn, 2005). Yet, again, the extent to which this explains L2_{Turkish} students' disadvantages in L2 proficiency needs to be tested in future research.

For the reading vocabulary measure, we found differences between both L2 groups on the one hand and the L1 group on the other hand even after controlling for socioeconomic status. This finding parallels previous studies showing that L2 vocabulary is often more limited than L1 vocabulary and that this disadvantage is common even in L2 learners who use their L2 routinely (Lesaux & Kieffer, 2010). Our results add to this research and indicate that, even at the end of formal schooling, L2 students may not reach reading vocabulary skills that are comparable to those of their monolingual peers.

In line with our third hypothesis, we did not find substantial differences in the relationships between reading comprehension and linguistic comprehension for L1 and L2 adolescents. Furthermore, the association between listening comprehension and reading comprehension was stronger than the association between text reading fluency and reading comprehension for all students. This finding for German is in line with results from research conducted for opaque orthographies, such as English, which has found listening comprehension to be especially important for reading comprehension of older students (Florit & Cain, 2011). Thus, the pattern seems to hold for more transparent orthographies like German as well.

While reading fluency was not included in the original version of the SVR, several studies have indicated that text reading fluency might be more important for reading comprehension in older students than word reading accuracy (Geva & Farnia, 2012; Share, 2008). Although we did not assess word reading, our findings add to this research and support the notion that text reading fluency is an important component skill of reading comprehension in adolescent readers with weak reading skills. Contrary to our assumptions, however, we found text reading fluency to be associated with reading comprehension to a similar degree for all students. This contradicts previous research revealing a stronger relationship between reading fluency and comprehension in L1 students than in L2 students (Buly & Valencia, 2002; Wiley & Deno, 2005). Our research differs from these studies by focusing on reading comprehension in a sample of predominantly weak readers from the lowest track of the German educational system. As several researchers have suggested, the development of text reading

fluency may depend on students' proficiency in language skills, such as vocabulary or syntactic knowledge (Crosson & Lesaux, 2010; Geva & Farnia, 2012). In general, students in our sample performed poorly on measures of skills associated with linguistic comprehension. For the reading vocabulary measure, for example, their performance was equivalent to a percentile rank of 24 for age norms and a percentile rank of 34 for grade-level norms. Our findings thus are in line with an assumption set forth by Crosson and Lesaux (2010), namely that monolingual children with weak linguistic skills should show similar profiles in the relationship between text reading fluency and comprehension as L2 learners.

Working memory was associated with reading vocabulary and phonological awareness for all groups but with morpho-syntactic awareness only for L2_{Other} students. These findings contradict previous research highlighting the role of working memory for comprehension processes in both L1 and L2 learners (e.g. Cain et al., 2004). Most of these studies used other measures of working memory than the digit spans, which we employed in the present study, and this may explain the differences in results. Measures of working memory using verbal material, for example, are likely to show stronger associations with language-related variables than digit spans. Yet, measures with verbal material (e.g., sentence span tasks) may be biased against L2 learners. Another possible explanation for the lack of associations between component skills of reading and working memory in our study is that these relationships may be masked by the inclusion of further component skills of reading. Recent research indicates that working memory does not emerge as a predictor of reading or listening comprehension when other components of comprehension, such as word knowledge, are taken into account (Florit, Roch, & Levorato, 2014).

The present study has a few limitations. First of all, its cross-sectional design does not permit causal inferences. Further research from a developmental perspective is needed to identify causal links between reading comprehension and its components longitudinally. Furthermore, the proportion of variance explained by the SVR was relatively low in the present study. This may result from the fact that our sample consisted of students from the lower track of the German school system and the variability in achievement was thus restricted. Also, internal consistencies of some of the measures used in our study were only satisfactory. These measures need further development to improve their psychometric properties. Furthermore, additional variables that were not considered in our analyses may play a role in explaining reading comprehension. These could include further components of reading from the ecological or psychological domain of the Componential Model of Reading, such as students' reading motivation or home reading resources (Netten, Droop, & Verhoeven, 2011).

Despite its limitations, our study indicates that an SVR model including text reading fluency is a valid model of reading comprehension for adolescent L1 and L2 readers in a transparent orthography. Our results further suggest that differences in socioeconomic background between L1 and L2_{Other} students are associated with differences in reading comprehension and its component skills. Studies that aim at investigating reading comprehension and its components in L1 and L2 students should therefore control for students' socioeconomic background. For reading vocabulary, however, we found L2 learners to have lower scores than L1 learners even after accounting for socioeconomic background. As vocabulary strongly influences students' listening comprehension skills which, in turn, affect their reading comprehension, the weaker vocabulary may put L2 students at risk for developing poor reading skills. Approaches to promoting L2 students' second-language proficiency should therefore target L2 vocabulary skills.

L2_{Turkish} students in our study exhibited persistent weaknesses in important linguistic skills even after we accounted for socioeconomic background. Apparently, nine years of immersion in the German school system were not sufficient for these students to develop a level of proficiency comparable to that of their L1 peers. To date, it remains subject of considerable debate whether the lower performance of L2_{Turkish} students reflects problems with the acquisition of the German language or is a result of social processes that disadvantage students with Turkish immigrant background. Notwithstanding the need of further studies that explore possible explanations for L2_{Turkish} students' achievement disadvantages, results of our study indicate that instructional approaches currently implemented in Germany to enhance German language proficiency of students with an immigrant background may not be successful for all students. Therefore, effective instructional methods are needed to foster these students' linguistic skills and, in turn, their reading comprehension.

8. References

- Aaron, P. G., Joshi, R. M., Gooden, R., & Bentum, K. E. (2008). Diagnosis and treatment of reading disabilities based on the component model of reading. *Journal of Learning Disabilities, 41*, 67-84. doi: 10.1177/0022219407310838
- Adams, A.-M., Bourke, L., & Willis, M. (1999). Working memory and spoken language comprehension in young children. *International Journal of Psychology, 34*, 364-373. doi: 10.1080/002075999399701
- August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth*. Mahwah, NJ: Lawrence Erlbaum.
- Backus, A. (2005). Turkish as an immigrant language in Europe. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 689–723). Oxford: Oxford University Press.
- Baddeley, A. D. (1986). *Working memory*. Oxford: Clarendon Press.
- Baumgartner, H., & Steenkamp, J.-B. E. M. (1998). Multi-group latent variable models for varying numbers of items and factors with cross-national and longitudinal applications. *Marketing Letters, 9*, 21-35. doi: 10.1023/A:1007911903032
- Becker, R. & Beck, M. (2011). Migration, Sprachförderung und soziale Integration. Eine Evaluation der Sprachförderung von Berliner Schulkindern mit Migrationshintergrund anhand von ELEMENT-Panel-Daten [Migration, language promotion, and social integration. An evaluation of language promotion of immigrant students in Berlin using the ELEMENT panel data]. In R. Becker (Ed.), *Integration durch Bildung. Bildungserwerb von jungen Migranten in Deutschland* (pp. 121-138). Wiesbaden: VS Verlag.
- Bialystok, E. (1979). Explicit and implicit judgements of L2 grammaticality. *Language Learning, 29*, 81-103. doi: 10.1111/j.1467-1770.1979.tb01053.x
- Bialystok, E. (2001). *Bilingualism in development. Language, literacy & cognition*. Cambridge: Cambridge University Press.
- Bialystok, E., Luk, G., & Kwan, E. (2005). Bilingualism, biliteracy, and learning to read: Interactions among languages and writing systems. *Scientific Studies of Reading, 9*, 43-61. doi: 10.1207/s1532799xssr0901_4
- Buly, M. R., & Valencia, S. W. (2002). Below the bar: Profiles of students who fail state reading. *Educational Evaluation and Policy Analysis, 24*, 219-239. doi: 10.3102/01623737024003219

- Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal activity, and component skills. *Journal of Educational Psychology, 96*, 31-42. doi: 10.1037/0022-0663.96.1.31
- Crosson, A. C., & Lesaux, N. K. (2010). Revisiting assumptions about the relationship of fluent reading to comprehension: Spanish-speakers' text-reading fluency in English. *Reading and Writing: An Interdisciplinary Journal, 23*, 475-494. doi: 10.1007/s11145-009-9168-8
- Cutting, L. E., & Scarborough, H. S. (2006). Prediction of reading comprehension: Relative contributions of word recognition, language proficiency, and other cognitive skills can depend on how comprehension is measured. *Scientific Studies of Reading, 10*, 277-299. doi: 10.1207/s1532799xssr1003_5
- Droop, M., & Verhoeven, L. (2003). Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly, 38*, 78-103. doi: 10.1598/RRQ.38.1.4
- Ehri, L. C. (1995). Stages of development in learning to read words by sight. *Journal of Research in Reading, 18*, 116-125. doi: 10.1111/j.1467-9817.1995.tb00077.x
- Fiske, S. T., Lin, M., & Neuberg, S. (1999). The continuum model. In S. Chaiken, & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 321-254). New York: Guilford Press.
- Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review, 23*, 553-576. doi: 10.1007/s10648-011-9175-6
- Florit, E., Roch, M., & Levorato, M. C. (2014). Listening text comprehension in preschoolers: A longitudinal study on the role of semantic components. *Reading and Writing: An Interdisciplinary Journal, 27*, 793-817. doi: 10.1007/s11145-013-9464-1
- Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*, 200-213. doi: 10.1016/0749-596X(89)90044-2
- Gathercole, V. C. M. (2002). Monolingual and bilingual acquisition: Learning different treatments of that-trace phenomena in English and Spanish. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children* (pp. 220-252). Clevedon, UK: Multilingual Matters.
- Geva, E., & Farnia, F. (2012). Developmental changes in the nature of language proficiency and reading fluency paint a more complex view of reading comprehension in ELL and

- EL1. *Reading and Writing: An Interdisciplinary Journal*, 25, 1819-1845. doi: 10.1007/s11145-011-9333-8
- Glock, S., & Krolak-Schwerdt, S. (2013). Does nationality matter? The impact of stereotypical expectations on student teachers' judgments. *Social Psychology of Education*, 16, 111–127. doi: 10.1007/s11218-012-9197-z
- Gogolin, I. (2005). Bilingual education – the German experience and debate. In Arbeitsstelle Interkulturelle Konflikte und gesellschaftliche Integration (Ed.), *The effectiveness of bilingual school programs for immigrant children* (pp. 133-145). Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).
- Gottardo, A., & Mueller, J. (2009). Are first- and second-language factors related in predicting second-language reading comprehension? A study of spanish-speaking children acquiring English as a second language from first to second grade. *Journal of Educational Psychology*, 101, 330–344. doi: 10.1037/a0014320
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial & Special Education*, 7, 6-10. doi: 10.1177/074193258600700104
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experiences of young American children*. Baltimore, MD: Brookes.
- Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision: KFT 4-12+ R* [Test of cognitive abilities for grades 4 to 12]. Weinheim: Beltz.
- Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127-160. doi: 10.1007/BF00401799
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55. doi: 10.1080/10705519909540118
- Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C., & Deno, S. L. (2003). Sources of individual differences in reading comprehension and reading fluency. *Journal of Educational Psychology*, 95, 719–729. doi: 10.1037/0022-0663.95.4.719
- Jussim, L., & Harber, K. D. (2005). Teacher expectations and self-fulfilling prophecies: Knowns and unknowns, resolved and unresolved controversies. *Personality and Social Psychology Review*, 9, 131-155. doi: 10.1207/s15327957pspr0902_3
- Kahraman, B., & Knoblich, G. (2000). Stechen statt sprechen: Valenz und Aktivierbarkeit von Stereotypen über Türken [Pricking instead of talking. Valence and activability of stereotypes about Turks]. *Zeitschrift für Sozialpsychologie*, 31, 31-43. doi: 10.1024//0044-3514.31.1.31

- Kieffer, M. J. (2010). Socioeconomic status, English proficiency, and late-emerging reading difficulties. *Educational Researcher*, 39, 484-486. doi: 10.3102/0013189X10378400
- Kim, S., & Hagtvet, K. A. (2003). The impact of misspecified item parceling on representing latent variables in covariance structure modeling: A simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, 10, 101-127. doi: 10.1207/S15328007SEM1001_5
- Kirby, J. R., & Savage, R. S. (2008). Can the simple view deal with the complexities of reading? *Literacy*, 42, 75-82. doi: 10.1111/j.1741-4369.2008.00487.x
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M., Schneider, W., & Stanat, P. (Eds.). (2010). *PISA 2009 - Bilanz nach einem Jahrzehnt* [PISA 2009 - balance after a decade]. Münster: Waxmann.
- Lesaux, N. K., & Kieffer, M. J. (2010). Exploring sources of reading comprehension difficulties among language minority learners and their classmates in early adolescence. *American Educational Research Journal*, 47, 596-632. doi: 10.3102/0002831209355469
- Lesaux, N. K., Lipka, O., & Siegel, L. S. (2006). Investigating cognitive and linguistic abilities that influence the reading comprehension skills of children from diverse linguistic backgrounds. *Reading and Writing: An Interdisciplinary Journal*, 19, 99-131. doi: 10.1007/s11145-005-4713-6
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling: A Multidisciplinary Journal*, 9, 151-173. doi: 10.1207/S15328007SEM0902_1
- Lorenz, G., Gentrup, S., Kristen, C., Stanat, P., & Kogan, I. (submitted). Stereotype bei Lehrkräften? Eine Untersuchung systematisch verzerrter Lehrererwartungen [Stereotypes among teachers? A study of systematic bias in teacher expectations]. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*.
- Meyer, M. S., & Felton, R. H. (1999). Repeated reading to enhance fluency: Old approaches and new directions. *Annals of Dyslexia*, 49, 283-306. doi: 10.1007/s11881-999-0027-8
- Muthén, L. K., & Muthén, B. O. (1998-2012). *MPlus Version 7.11*. Los Angeles, CA: Muthén & Muthén.
- Nelson, T. D. (2002). *The psychology of prejudice*. Needham Heights, MA: Allyn & Bacon.

- Netten, A., Droop, M., & Verhoeven, L. (2011). Predictors of reading literacy for first and second language learners. *Reading and Writing: An Interdisciplinary Journal*, *24*, 413–425. doi: 10.1007/s11145-010-9234-2
- Noble, K. G., Farah, M. J., & McCandliss, B. D. (2006). Socioeconomic background modulates cognition–achievement relationships in reading. *Cognitive Development*, *21*, 349-368. doi: 10.1016/j.cogdev.2006.01.007
- OECD. (2010). *PISA 2009 results: Overcoming social background – equity in learning opportunities and outcomes (volume II)*. Paris: OECD Publishing.
- OECD. (2012). *PISA 2009 Technical Report*. Paris: OECD Publishing.
- Olsthoorn, N. M., Andringa, S., & Hulstijn, J. H. (2012). Visual and auditory digit-span performance in native and nonnative speakers. *International Journal of Bilingualism*, *0*, 1-11. doi: 10.1177/1367006912466314
- Ready, D. D., & Wright, D. L. (2011). Accuracy and inaccuracy in teachers’ perceptions of young children’s cognitive abilities. The role of child background and classroom context. *American Educational Research Journal*, *48*, 335-360. doi: 10.3102/0002831210374874
- Riches, C., & Genesee, F. (2006). Literacy: Crosslinguistic and crossmodal issues. In F. Genesee, K. Lindholm-Leary, W. M. Saunders & D. Christian (Eds.), *Educating English language learners. A synthesis of research evidence* (pp. 64-109). Cambridge: Cambridge University Press.
- Savage, R. (2006). Reading comprehension is not always the product of nonsense word decoding and linguistic comprehension: Evidence from teenagers who are extremely poor readers. *Scientific Studies of Reading*, *10*, 143-164. doi: 10.1207/s1532799xssr1002_2
- Schneider, W., Schlagmüller, M., & Ennemoser, M. (2007). *Lesegeschwindigkeits- und Verständnistest für die Klassen 6-12 (LGVT 6-12)* [Test of reading fluency and accuracy]. Göttingen: Hogrefe.
- Share, D. L. (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin*, *134*, 584–615. doi: 10.1037/0033-2909.134.4.584
- Siegel, L. S., & Ryan, E. B. (1988). Development of grammatical-sensitivity, phonological, and short-term memory skills in normally achieving and learning disabled children. *Developmental Psychology*, *24*, 28-37. doi: 10.1037/0012-1649.24.1.28

- Söhn, J. (2005). *Zweisprachiger Schulunterricht für Migrantenkinder. Ergebnisse der Evaluationsforschung zu seinen Auswirkungen auf Zweitspracherwerb und Schulerfolg* [Bilingual education for immigrant students. Results from evaluation research concerning effects on second-language acquisition and educational outcomes]. Berlin: Wissenschaftszentrum für Sozialforschung.
- Sprietsma, M. (2013). Discrimination in grading: Experimental evidence from primary school teachers. *Empirical Economics*, 45, 523-538. doi: 10.1007/s00181-012-0609-x
- Stanat, P., Rauch, D., & Segeritz, M. (2010). Schülerinnen und Schüler mit Migrationshintergrund [Immigrant students]. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel, W. Schneider & P. Stanat (Eds.), *PISA 2009: Bilanz nach einem Jahrzehnt* (pp. 200-230). Münster: Waxmann.
- Stanat, P., Segeritz, M., & Christensen, G. S. (2010). Schulbezogene Motivation und Aspiration von Schülerinnen und Schülern mit Migrationshintergrund [School-related motivation and aspirations of immigrant students]. In W. Bos, E. Klieme & O. Köller (Hrsg.), *Schulische Lerngelegenheiten und Kompetenzentwicklung. Festschrift für Jürgen Baumert* (pp. 31-57). Münster: Waxmann.
- Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany (KMK). (2013). *The education system in the Federal Republic of Germany 2011/2012*. Bonn: Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany.
- Verhoeven, L., & van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing: An Interdisciplinary Journal*, 25, 1805-1818. doi: 10.1007/s11145-011-9346-3
- Weiß, R. H. (2006). *Grundintelligenztest Skala 2, CFT 20 - Revision (CFT 20-R)* [Test of general cognitive abilities - Scale 2]. Göttingen: Hogrefe.
- Wiley, H. I., & Deno, S. L. (2005). Oral reading and maze measures as predictors of success for English learners on a state standards assessment. *Remedial and Special Education*, 26, 207-214. doi: 10.1177/07419325050260040301
- Wolf, M., & Katzir-Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading*, 5, 211-239. doi: 10.1207/S1532799XSSR0503_2

Anhang C: Bildungssprachliches und alltagssprachliches Hörverstehen in einer Erst- bzw. Zweitsprache (Teilstudie 3)

Marx, A., Heppt, B. & Henschel, S. (in Revision). *Listening comprehension of academic and everyday language in first-language and second-language students*. Applied Psycholinguistics.

Abstract. Listening comprehension in academic contexts is an important prerequisite for the acquisition of school-related skills. However, there is no consensus yet on what constitutes academic listening comprehension and the concept of academic language is generally subject of much debate. To date, it remains unclear whether academic listening may be regarded as a distinct construct. The present study investigates the distinction of academic and everyday listening comprehension in 459 ninth grade students speaking German as a first (L1) or as a second language (L2). Empirical results confirm that academic listening forms a separate facet of listening comprehension. Contrary to theoretical assumptions, L2 students' listening comprehension was not differentially affected by language demands of academic texts.

Keywords: Listening Comprehension, Academic Language, CALP, BICS, Second-language students

Learning in any school subject depends on students' proficiency in listening comprehension in academic contexts (Rost, 1994). Classroom discourse has, for example, repeatedly been found to facilitate the acquisition of reading comprehension (cf. Nystrand, 2006; Wolf, Crosson, & Resnick, 2005). Listening comprehension in academic contexts thus may be regarded as basis for the acquisition of knowledge and academic achievement.

Several definitions of academic listening comprehension and its components have been suggested (cf. Ferris & Tagg, 1996; Taylor & Geranpayeh, 2011). Most of these characterizations emphasize that the purpose of listening in academic contexts is primarily to impart knowledge and communicate information. A variety of taxonomies were proposed to describe comprehension subskills that are necessary for this particular type of listening, such as "differentiation between relevant and irrelevant information," "command of strategies that support understanding (i.e., taking notes)," "the ability to identify the topic, purpose and scope of a lecture," or "making inferences" (e.g. Buck, 2001; Flowerdew, 1995; Richards, 1983). In academic listening, the listener needs to extract the direct and exact meaning of the message that is presented and has to concentrate on as well as understand long stretches of speech without being able to use the facilitating functions of interactive discourse (Flowerdew, 1994).

For the most part, communication in academic contexts is marked by a high level of "writtenness", whether information is presented written or orally. Academic listening texts possess characteristics of formal and planned discourse as they are lexically dense, context-reduced and characterized by textual structures that resemble those of written texts (Flowerdew & Miller, 1997). Classroom talk may thus be localized at the interface of an oral-literate continuum where „oral" refers to prototypical speaking in everyday conversations and "literate" refers to prototypical writing in academic prose (Biber, 1995; Tannen, 1982). Although oral academic language is often less complex and contains fewer academic language features than written academic language (cf. Corson, 1997; Nagy & Townsend, 2012), they share important features, such as an informational focus and an abstract style (Csomay, 2006). Furthermore, oral academic language is, for the most part, context-reduced: The goal of academic discourse is to share information about abstract objects, events, and situations that are removed from the immediate context. To convey this information, special linguistic devices are applied, such as precise vocabulary and formal syntactic markers of the temporal and causal nature of events (Curenton & Justice, 2004; Collier, 1987). Oral language in academic settings is therefore more complex and more difficult to process than oral language in everyday conversations (cf. Flowerdew, 1995). Accordingly, Richards (1983) has suggested that

everyday listening and academic listening comprehension may rely on different sets of sub-skills. Yet it remains unclear how skills associated with listening in academic contexts, such as the “ability to recognize key lexical items related to subject/topic“ (cf. Richards, 1983, p. 229), may be operationalized and to what extent they are specific for listening in academic contexts. To describe challenges of academic listening comprehension it thus seems useful to resort to more general conceptualizations of academic language.

1. Challenges of Academic Language

The general differentiation between language skills needed to interact with others in informal contexts and those needed to understand school-related language has been addressed by various researchers (e.g., Bailey & Butler, 2003; Schleppegrell, 2004). An important starting point of this line of research was the distinction between *Basic Interpersonal Communication Skills* (BICS) and *Cognitive Academic Language Proficiency* (CALP) proposed by Cummins (1980). According to Cummins (1980), BICS are defined as the language skills needed in social situations. When people interact with each other, their communication usually occurs in a meaningful context that supports understanding. Cummins (1980) assumes that the language used in such situations is, on average, relatively easy to understand and cognitively not very demanding; it therefore only requires BICS. CALP, in contrast, is needed to understand language in context-reduced settings and is cognitively demanding, such as when abstract ideas or concepts are presented. CALP thus primarily pertains to language typically used in academic contexts and can be defined as “the extent to which an individual has access to and command of the oral and written academic registers of schooling” (Cummins, 2000, p. 67).

The differentiation between BICS and CALP served as an important starting point for further conceptualizations of academic language. To date, there is no single definition of what constitutes academic language and a variety of terms have been suggested to label the construct, such as “language of schooling”, “scientific language”, or “advanced literacy” (cf. Snow & Uccelli, 2009). The US-based National Center for Research on Evaluation, Standards, and Student Testing (CRESST) has developed a conceptualization of academic language which adds additional features to the CALP definition of academic language (Bailey & Butler, 2003). According to this view, academic language can be defined as “the language that is used by teachers and students for the purpose of acquiring new knowledge and skills, imparting new information, describing abstract ideas, and developing students’ conceptual understanding” (Chamot & O’Malley, 1994, p. 40). On the content level, academic language is

characterized by cognitively demanding school-related topics, such as lectures on abstract concepts or ideas (Bailey & Butler, 2003; Eckhardt, 2008). It is also, to a certain extent, context-reduced: the information conveyed is often removed from its immediate context, i.e. there is a very limited amount of shared knowledge or context between speaker and listener (cf. Carenton & Justice, 2004).

Academic language is associated with the mastery of specific genres, command of specific reasoning and argumentative strategies (Snow & Uccelli, 2009) as well as with language features that pertain to specific cognitive procedures, such as generalization or abstraction (cf. Gottlieb & Ernst-Slavit, 2014; Nagy & Townsend, 2012). It is therefore marked by a more sophisticated use of grammar and vocabulary than everyday language, thus requiring mastery of an extensive range of linguistic features (Scarcella & Rumberger, 2000). On the semantic level, academic language is characterized by functional academic vocabulary (e. g., *synthesize, denominator*) that students encounter at school but rarely in less formal social settings (Bailey & Butler, 2003). Academic language also contains a high proportion of nouns which add to the lexical density of academic texts (Schleppegrell, 2004). On the syntactic level, academic language entails a variety of structures that are typical for written discourse. These include verb forms in passive voice, long and complex noun and prepositional phrases, nominalizations, and participial modifiers (Abedi & Lord, 2001; Schleppegrell, 2004). There are currently no empirical studies on what constitutes academic language in German, however it is assumed that linguistic features of English academic language are also relevant for German academic language (cf. Eckhardt, 2008).

Researchers at CRESST have developed a framework that allows to differentiate academic language from everyday language and to develop tasks to operationalize them (Bailey, Huang, Shin, Farnsworth, & Butler, 2007). This framework, characterizes academic language by a frequent use of the features described above. By contrast, these features rarely occur in everyday language (cf. Bailey, 2007), which is characterized by the use of high-frequency vocabulary and relatively simple syntactical structures (Schleppegrell, 2004). Several researchers have argued that texts may be located on a continuum, with everyday language and academic language at the extremes (Berendes, Dragon, Weinert, Heppt, & Stanat, 2013; Snow, 2010).

2. Academic Language in L2 Learners

Both second-language (L2) and first-language (L1) students need to acquire academic language when they enter school. However, L2 children face a double challenge: They need

to learn the school-specific register and they need to do so in a language that is not their native one (Aarts, Demir, & Vallen, 2011). Various researchers assume that L2 speakers acquire conversational language faster than academic language (Collier, 1989; Roessingh, Kover, & Watt, 2005). Cummins (2008), for example, reports a gap of several years between the attainment of conversational fluency, which is usually acquired within two years of exposure to a particular language, and the attainment of academic language proficiency, which may take five to seven years. Analyzing data from four different school districts in the US, Hakuta, Butler, and Witt (2000) confirmed this assumption and showed that it took English language learners three to five years to develop everyday language skills in English while the acquisition of academic English proficiency took four to seven years. However, the rate of academic language attainment in L2 may depend on the age of arrival and prior schooling in L1: older children from ages eight to twelve and adolescents who have already completed several years of schooling in their L1 tend to acquire L2 academic language faster than younger children who did not attend school in their L1 (Collier, 1989). However, even the academic language skills of advanced L2 learners may not reach those of L1 students (Mason, 1995).

3. Criticism of and Empirical Results on Academic Language

The differentiation between BICS and CALP proposed by Cummins (2000) has been criticized by several researchers (see Cummins, 2008 for an overview). First, it does not specify which language skills are encompassed in CALP (Snow & Uccelli, 2009). A simple differentiation between BICS and CALP that does not include a description of the necessary skills does not suffice to construct tasks and promote learning, such that the concept is of limited practical utility (Scarcella, 2003). In absence of a detailed description of CALP skills, Edelsky et al. (1983) propose that academic language is an artifact of the way it is measured and hence it represents nothing but a form of test-wiseness. Furthermore, it is questionable whether everyday language is always contextualized and less cognitively demanding than academic language. Persuasion, deceiving, or winning over others in everyday life, for example, require cognitive abilities as complex as those needed to comprehend a persuasive argument in a social studies textbook (Bailey & Huang, 2011). The exact boundary between BICS and CALP thus remains somewhat unclear.

Only few studies have sought to differentiate academic language and everyday language empirically. Evidence concerning the assumed differences between everyday conversational listening and academic listening is particularly scarce (Buck, 2001). Research that focused on written language has, however, provided some empirical support for the assumption

that academic and everyday language might be separable. Schleppegrell (2004) analyzed linguistic features of academic texts and transcripts of everyday conversations and found evidence for the theoretically postulated differences between academic and everyday language. In a similar vein, a number of studies indicated that the linguistic complexity of test items in mathematics or science tests impacted item difficulties differentially (e.g., Martiniello, 2008; Shaftel, Belton-Kocher, Glasnapp & Poggio, 2006). The group specific difficulty of test items containing academic language features compared to items containing everyday language features provides some support for the assumption of different language registers.

For academic listening comprehension, Shohamy and Inbar (1991) investigated the effect of different types of listening texts (orally-oriented vs. literal-oriented) on L2 students in twelfth grade. The authors did not explicitly mention that they were measuring academic and everyday listening comprehension but the operationalization of the oral/literate listening tasks is in line with central assumptions on academic and everyday listening. Results indicate that listening texts containing academic language features were significantly more difficult to process for the L2 listeners than everyday listening texts, thus supporting the notion that academic listening poses special demands on L2 listeners. As Shohamy and Inbar (1991) did not include a L1 comparison group in their study, however, it remains unclear whether differences in academic vs. everyday listening comprehension may also be found in native students. Furthermore, the main focus of this study was on oral vs. literate texts, such that the authors did not include all features that may be relevant to differentiate academic from everyday listening.

A number of studies has challenged the assumption that academic language is particularly difficult for L2 students (Eckhardt, 2008; Heppt, Stanat, Dragon, Berendes, & Weinert, 2014). Eckhardt (2008) reported for German elementary students that L1 and L2 students performed less well on listening comprehension tasks that contained complex grammatical features and academic vocabulary. Furthermore, she found weaker performance for both groups of students on texts with school-related contents as opposed to texts with everyday contents. In a third study, Eckhardt (2008) varied the contextual embedding of the listening comprehension tasks. Texts that were context-reduced (verbally described film scenes) were clearly more difficult to understand for elementary school students than texts that were supported by visual context (acted film scenes). Contrary to the author's expectation, these characteristics did not generally affect the performance of L2 students more than the performance of L1 students. Only for texts that contained school-related content, Eckhardt (2008) found L2 students' performance to be more strongly affected than L1 students' performance, yet, this effect disappeared when students' socioeconomic background was controlled.

This finding adds to a number of studies indicating that children from different socioeconomic backgrounds differ considerably in terms of language proficiency (Hoff, 2013). Aarts, Demir, and Vallen (2011), for example, point out that mothers from families of high socioeconomic status (SES) tend to use more academic language features when talking to their children than mothers from families of lower socioeconomic status. Given that L2 students are often disadvantaged in terms of their socioeconomic background (OECD, 2010), it is necessary to disentangle effects of language background and SES when comparing academic language skills in L1 and L2 students.

Taken together, empirical research on academic language has found support for the assumption that texts containing academic language features can be more difficult to process than texts containing everyday language. Contrary to theoretical assumptions on academic language, however, studies have not consistently shown that L2 students are more strongly affected by texts containing academic language features than L1 students. This may result from the fact that most studies have focused on beginning-level L2 learners with limited language skills, or on students in elementary grades. As academic language closely resembles everyday language at elementary grades, it may be assumed that disadvantages of L2 students with academic language texts do not become visible until the later grades, when language becomes increasingly demanding in terms of academic features (Chudaske, 2012).

Furthermore, previous research has focused on academic language in terms of classification of texts and tasks. Yet, academic language is associated with specific language characteristics, such as sophisticated grammatical and lexical features that pertain to specific cognitive features and demand a more pronounced language proficiency than that needed to process everyday language (cf. Gottlieb & Ernst-Slavit, 2014; Scarcella & Rumberger, 2000). It may therefore be assumed that the processing of academic language requires a specific facet of language proficiency. However, it remains unclear whether listening comprehension of academic language and everyday language are associated with distinct facets of language proficiency.

4. Research Questions and Hypotheses

To date, findings on academic language proficiency in older students are inconclusive and empirical evidence on the differentiation between academic and everyday language proficiency is scarce. Our study varied features of academic language and investigated whether listening comprehension of academic language and listening comprehension of everyday language may be separable on the basis of these features in a sample of adolescent L1 and L2

students. As grammatical features and, even more so, vocabulary knowledge may be regarded as the most important aspects of academic oral proficiency in learning another language (Saville-Troike, 1984; Nagy & Townsend, 2012), we based our study on the CRESST definition of academic language that puts a special emphasis on linguistic features of academic language. The following research questions and hypotheses are addressed:

- I. Can listening proficiency of academic language and everyday language be differentiated on the basis of features of academic language suggested in the literature (school-based content, reduced context, linguistic complexity)? We expect that academic language proficiency and everyday language proficiency may be distinguished empirically and that they show differential correlations to students' vocabulary and syntactic skills. We also assume students' vocabulary and syntactic skills to be more closely associated with their academic language proficiency than with their everyday language proficiency because academic language is assumed to entail vocabulary and syntactic features that are more difficult to process.
- II. Are there performance differences between adolescent L1 and L2 students in listening comprehension of academic and everyday language? We expect L2 students to generally perform lower than L1 students on measures of listening comprehension. Based on theoretical assumptions that assume academic language to be more challenging for L2 adolescents than for L1 adolescents, we predict that these performance differences are more pronounced for listening comprehension of academic language than for listening comprehension of everyday language.

5. Method

5.1 Participants

Overall, 479 students from 20 schools took part in the study. Of these, 20 students did not complete any of the test items and were therefore excluded from the analyses, resulting in a total of 459 students (45% female, 55% male). All students attended the ninth grade, and their mean age was 16.3 years ($SD = 0.9$). The majority of students (43%) attended the lowest secondary school track of the German educational system (*Hauptschule*), 25% attended the intermediate track (*Realschule*), and 32% attended a comprehensive school (*Gesamtschule*) which provides several courses of education under one roof. All schools were public schools with 60% or more students with an immigration background and were located in urban districts with a high proportion of low-income or immigrant background population.

Students' first language was determined with a single item in a questionnaire (“*Which language did you first learn within your family?*”). Based on their answers, we classified them into two language groups with either German as their first language (L1 group) or a language other than German as their first language (L2 group). A total of 173 (37.7%) students reported German as their L1 and 286 students (62.3%) indicated a language other than German as their L1. The most frequent first languages in the latter group were Turkish (70.0%), Arabic (7.2%), and Italian (6.7%). Although most L2 students were born in Germany (76.6%), the majority indicated speaking a language other than German in their homes (67.9%). Most of the L2 students (89.9%) reported to have begun acquiring German before the age of six years, such that they had been speaking German for at least 10 years. Another 3.2% of L2 students had acquired German for more than 7 years, and the remaining 6.9% of L2 students had started acquiring German after the age of ten years. The latter students' teachers evaluated their German language skills as being sufficiently developed, hence, none of these students was excluded from the analyses. All L1 students were born in Germany and had parents that were born in Germany. L1 and L2 students of our sample were distributed across the educational tracks in a similar manner ($\chi^2(2) = 0.44, p = .80$). The language groups did not differ in terms of age ($F(1,421) = 1.28, p = .26$) or gender distribution ($\chi^2(1) = 0.10, p = .75$). Furthermore, we found no group differences ($F(1,391) = 2.43, p = .12$) in students' basic cognitive abilities as measured with a test on figural analogies (KFT 4-12+ R; Heller & Perleth, 2000). The z-scores for the test of figural analogies were $M = 0.11$ ($SD = 0.98$) for the L1 group and $M = -0.05$ ($SD = 1.01$) for the L2 group.

5.2 *Instruments*

5.2.1 *Listening comprehension of academic and everyday language*

To assess listening comprehension of academic and everyday language, we developed a computer-based test. Students listened to German texts with a length of 100 to 150 words varying between 42 seconds and 65 seconds in duration. All texts were listening passages that were written for the test and were spoken by German native speakers. Four texts (50%) were dialogues between two persons (1 male, 1 female), two texts were narratives delivered by one speaker (25%), and two texts were expository texts (25%) delivered by one speaker. After listening to each of the eight texts, students answered between three and five items, resulting in a total of 32 questions. Questions were either in multiple-choice format with four response options (19 items, 59%), or in constructed-response format requiring a short written response (13 items, 41%). Items either pertained to the texts' main idea, for example, deciding on

where a conversation may have taken place (17 items, 53%) or aimed at local details included in the listening passage, for example, remembering a given definition (15 items, 47%). Students were given 1 point if they chose the correct answer from the multiple choice questions and 0 points if they chose a wrong option. Constructed response items were coded following a coding scheme and were also scored with 1 point when the correct answer was given and 0 points when the answer was wrong. The texts were only presented acoustically while all items were presented acoustically as well as in written form on the computer screen. As the focus of our study was on verbal aspects of academic language, we did not use visual input (e.g., videos) to accompany the listening texts.

The listening comprehension test consisted of two subtests: One was designed to capture listening comprehension of academic language and the other listening comprehension of everyday language. Both subtests included four listening texts with 16 items, respectively. Texts that were designed to be academic had contents pertaining to more formal settings (e.g., a job interview) and were context-reduced, that is, provided less information on the characters and the setting of the plot than the everyday texts. On the linguistic level, academic texts had more syntactic features that are typical for written discourse, namely passive voice, complex subordinate clauses, and complex noun phrases. Two associate professors for German linguistics who were blind to the research questions and the designated classification of the texts rated the difficulty of the syntactic features, the vocabulary, and the context on a scale ranging from 1 = *very easy* to 6 = *very difficult*. Units with academic listening texts were rated to be more difficult in terms of syntactic features ($M = 3.88$; $SD = 1.79$), vocabulary ($M = 3.25$; $SD = 0.71$), and content ($M = 3.25$; $SD = 0.50$) than units with everyday language ($M = 2.75$; $SD = 0.89$ for syntactic features, $M = 2.75$; $SD = 0.71$ for vocabulary, and $M = 1.5$; $SD = 1.03$ for content). Interrater reliability for the two raters was $\kappa = .69$ ($p < .01$) for the syntactic features, $\kappa = .70$ ($p < .01$) for vocabulary and $\kappa = .66$ ($p < .01$) for the content.

Besides the variation of syntax and content, academic listening texts were constructed such that they contained more academic vocabulary (e.g. *syllable*, *to curtail*) than everyday listening texts. To classify vocabulary as academic or everyday vocabulary, three raters classified each word in the listening texts following the CRESST criteria to assess academic language (Butler, Bailey, Stevens, & Huang, 2004). Interrater reliability for the three raters was $\kappa = .66$ ($p < .01$). Appendix A provides an overview of the listening text features for academic and everyday units. Sample texts and items can be found in Appendix B.

As described above, academic listening texts and everyday listening *texts* differed in terms of their linguistic features, their content, and their contextualization. However, *items* for

academic units as well as those for everyday units were constructed such that they were comparable in terms of their linguistic difficulty. Academic language and everyday language units were comparable in terms of the number of items in multiple-choice versus constructed-response format, as well as in terms of items on main ideas versus local details from the text. The tests' internal consistency was $\alpha = .71$ for *academic listening comprehension* and $\alpha = .59$ for *everyday listening comprehension*. The correlation between the two subtests was $r = .54$ ($p < .01$).

5.2.2 Vocabulary

We assessed vocabulary using the CFT 20-R/WS, version A (Weiß, 2006). The test consists of 30 multiple-choice items that require students to find a synonym for a given German word from a list of five words. Its internal consistency was $\alpha = .81$.

5.2.3 Syntactic knowledge

To assess students' syntactic knowledge, we developed a computer-based test that consists of written sentences with various syntactical structures, such as relative clauses or passive voice. Students had to decide for each sentence whether it was syntactically correct or not. To control for guessing, students could choose between the answers *incorrect*, *correct*, and *unclear*, with the latter option indicating that they were uncertain in their judgment (sample item: *Is the following sentence grammatically correct or incorrect? "I to bed went late yesterday." correct – incorrect – unclear*). The *unclear*-option was later coded as wrong. The test includes 22 items and had an internal consistency of $\alpha = .86$ in the study sample.

5.2.4 Socioeconomic status

We used the ESCS (Index of Economic, Social, and Cultural Status; OECD, 2010) as an indicator of students' socioeconomic status. This measure is also employed in the PISA studies and has a mean of 0 and a standard deviation of 1 in the PISA student population. The components comprising the ESCS are the families' home possessions (such as books at home, cars, televisions), the highest parental occupation as measured by the International Socioeconomic Index of Occupational Status (Ganzeboom, De Graaf, & Treiman, 1992), and the highest parental education in years of schooling. When levels of parental occupation or education differ for mother and father, the highest level is used to indicate a families' occupational and educational status, respectively. All data were obtained with student and parent questionnaires.

5.3 Analyses

We first tested the measures of listening comprehension, vocabulary, and syntactic knowledge for items displaying *differential item functioning* (DIF) as a function of language background. DIF occurs when students with comparable abilities but from different groups have different probabilities of success on an item (Clauser & Mazor, 1998). Item DIF can be adverse or benign (e.g. Gierl, 2005). Benign item DIF is typically due to a multidimensionality of the construct and should therefore not be eliminated from the test. Adverse DIF can be attributed to the presence of nuisance dimensions and therefore threatens construct validity, such that items displaying adverse DIF between L1 and L2 students should be omitted.

DIF for each item was operationalized as the difference of item logit parameters between L1 and L2 students. The values were judged not only in terms of statistical significance, but also in terms of effect size. Logit differences above 0.64 are considered as large magnitude DIF (cf. Wilson, 2005, p. 167). DIF which persists in a two-dimensional Rasch model of listening comprehension is considered to be adverse. Items that displayed adverse and large magnitude DIF for the language groups were excluded from the following analyses. The DIF analyses were conducted using the computer program ConQuest 2.0 (Wu, Adams, Wilson, & Haldane, 2007).

To test the first hypothesis on the separability of everyday listening proficiency and academic listening proficiency, we conducted a confirmatory factor analysis (CFA) using Mplus 7.11 (Muthén & Muthén, 1998-2012). Students' economic, social, and cultural status (ESCS) was entered as a manifest covariate into the analyses. We first estimated a two-dimensional model with items loading either on an academic or on an everyday language factor and then compared this model to a unidimensional model with a global listening comprehension factor. Parameter estimates for the models are based on the robust *weighted least square estimator* (WLSMV) for categorical outcomes (Muthén & Muthén, 1998-2012). We used the *Comparative Fit Index* (CFI), the *Tucker-Lewis-Index* (TLI), and the *Root Mean Square Error of Approximation* (RMSEA) to evaluate goodness of fit. Models with a TLI and CFI coefficient higher than .95 and a RMSEA coefficient smaller than .06 are regarded as good approximations to the data (Hu & Bentler, 1999). The models were then compared using χ^2 -difference tests. A significant χ^2 -difference of nested models suggests that the less restrictive model fits the data better than the more restrictive model.

To further investigate dimensionality, we estimated correlations between listening comprehension of academic and everyday language and receptive vocabulary and syntactic

knowledge within the models. As listening comprehension of academic language is more complex than listening comprehension of everyday language in terms of vocabulary and syntax, listening comprehension of academic language should be more strongly associated with vocabulary and syntax than listening comprehension of everyday language. Both external criteria were modeled as unidimensional latent variables. We used Wald χ^2 -tests to determine differences in the correlations between the external criteria and listening comprehension of academic and everyday language (Wald, 1943).

In order to compare latent variable means between L1 and L2 students, as mentioned in our second hypothesis, we first tested the resulting listening comprehension model for measurement invariance to ensure applicability of the model for both groups of students (Baumgartner & Steenkamp, 1998). We initially estimated a baseline model, with the factorial structure set to be equal for the language groups but factor loadings and means allowed to vary (*configural invariance*). Subsequently, factor loadings (*metric invariance*) and, finally, means were restricted to be equal across groups as well (*scalar invariance*). Again, we evaluated goodness of fit according to model fit parameters and χ^2 -difference tests. We then compared latent variable means between L1 and L2 students using Wald χ^2 -tests (Wald, 1943).

6. Results

6.1 Descriptive Results

Table 1 shows the means, standard deviations, and intercorrelations for all measures. Results indicate that the majority of students in our sample grew up in low-SES families. The samples' ESCS mean was $M = -0.63$ ($SD = 0.92$), which is more than half a standard deviation below the German average (OECD, 2010b). The language groups differed in terms of their socioeconomic background with L2 students coming from families with lower SES than L1 students ($M = -0.33$, $SD = 0.84$ for L1 students; $M = -0.79$, $SD = 0.94$ for L2 students; $F(1,363) = 17.50$, $p < .01$). Table 1 also shows that ESCS is only correlated with syntactic knowledge, vocabulary, and academic listening comprehension, but not with everyday language comprehension. Therefore, it is appropriate to include ESCS as a covariate.

Table 1

Means and Standard Deviations of all Measures

	N	M	SD	Intercorrelations ²				
				2	3	4	5	6
1) ESCS	363	-0.63	0.92	.26*	.16*	.16*	.09	.20*
2) VO ¹	392	.75	.15		.33*	.44*	.33*	.43*
3) SK ¹	382	.84	.18			.45*	.42*	.37*
4) LC (total) ¹	417	.61	.17				.85*	.90*
5) LC (EL) ¹	417	.68	.18					.54*
6) LC (AL) ¹	417	.54	.20					

Note. ESCS = Index of Economic, Social, and Cultural Status, reported as WLE; VO = Vocabulary; SK = Syntactic knowledge; LC = Listening Comprehension; AL = Academic Language; EL = Everyday Language; ¹Reported as Percentage of Correct Answers; ²Zero-order Correlation Coefficients with Sample Sizes of $357 \leq N \leq 417$. * $p < .01$

We did not find any language-group differences in the syntactic knowledge measure which appeared to be relatively easy for all students: The L1 and L2 students solved 88% resp. 84% of the tasks ($F(1,382) = 3.21, p = .07$). Although students also performed well on the vocabulary test overall, L2 students showed significantly weaker vocabulary skills than L1 students ($M = 0.80, SD = 0.12$ for L1 students and $M = 0.72, SD = 0.16$ for L2 students; $F(1,392) = 14.30, p < .01$). L2 students also performed less well than L1 students on the listening comprehension test ($M = 0.66, SD = 0.14$ for L1 students and $M = 0.58, SD = 0.17$ for L2 students; $F(1,417) = 29.48, p < .01$). Both groups of students showed weaker performance on the academic listening subtest ($M = 0.60, SD = 0.18$ for L1 students; $M = 0.50, SD = 0.21$ for L2 students) than on the everyday listening subtest ($M = 0.73, SD = 0.15$ for L1 students, $M = 0.65, SD = 0.18$ for L2 students), yet, the difference was not significant within either of the groups ($t(153) = -0.86, p = .93$ for L1 students, $t(262) = 0.69, p = .95$ for L2 students).

6.2 Detecting Differential Item Functioning

Results of the DIF analyses indicated that the listening comprehension test as well as the vocabulary test and the test of syntactic knowledge contained several items displaying significant DIF as a function of language background: We identified three items (two everyday listening items, one academic listening item) as displaying adverse and large magnitude DIF for the listening comprehension test (9.4%), seven items (23.3%) for the vocabulary test, and two items (9.1%) for the test of syntactic knowledge. These items were excluded from further analyses.

6.3 Dimensionality of Listening Comprehension

Table 2 shows the fit data for the unidimensional and the two-dimensional model of listening comprehension. The fit indices suggest that both models fit the data well. However, the comparison of the models using χ^2 -difference testing indicates that the fit differs: The less restrictive two-dimensional model fits the data significantly better than the more restrictive unidimensional model. Furthermore, the fit indices CFI and TLI are higher for the two-dimensional than for the unidimensional model, thus providing further support for the assumption that the two-dimensional model better fits the data. At the same time, the latent correlation between listening comprehension for academic language and listening comprehension for everyday language is substantial and significant ($r = .84$; $p < .01$), suggesting that there is substantial overlap between the two facets. We therefore conducted a post-hoc test to analyze whether this correlation is significantly lower than one. For this purpose, we fixed the correlation of the two facets to one and compared this model with a model where the correlation was not restricted. The nested model test indicates that the correlation between the latent variables is significantly different from one ($\Delta\chi^2(1) = 11.92$, $p < .01$).

Table 2

Test of Dimensionality for the Listening Comprehension Test (Controlling for ESCS)

Model	χ^2	df	p	RMSEA	CFI	TLI	Compared to ...	$\Delta\chi^2$	df	p
A: unidimensional	492.04	405	<.01	.02	.94	.94	--	--	--	--
B: two-dimensional	473.49	403	.01	.02	.95	.95	Model A	18.15	2	<.01

Note. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; $N = 459$.

We further assured dimensionality by estimating the correlations between the two listening comprehension scores and vocabulary resp. syntactic knowledge as external criteria. Figure 1 shows the concurrent correlations between listening comprehension of academic and everyday language on the one hand and students' receptive vocabulary and syntactic knowledge on the other hand controlling for ESCS. The results indicate that syntactic knowledge is more strongly associated with listening comprehension of everyday language than with listening comprehension of academic language ($\chi^2(1) = 11.33$, $p = .01$). Receptive vocabulary, in contrast, correlates somewhat more strongly with listening comprehension of

academic language than with listening comprehension of everyday language, yet the difference is not significant ($\chi^2(1) = 0.38, p = .54$). Thus, we did not find that vocabulary and syntactic knowledge were more strongly associated with academic language proficiency as suggested by our first hypothesis. The proportion of explained variance was $R^2 = .46$ for listening comprehension of everyday language and $R^2 = .49$ for listening comprehension of academic language.

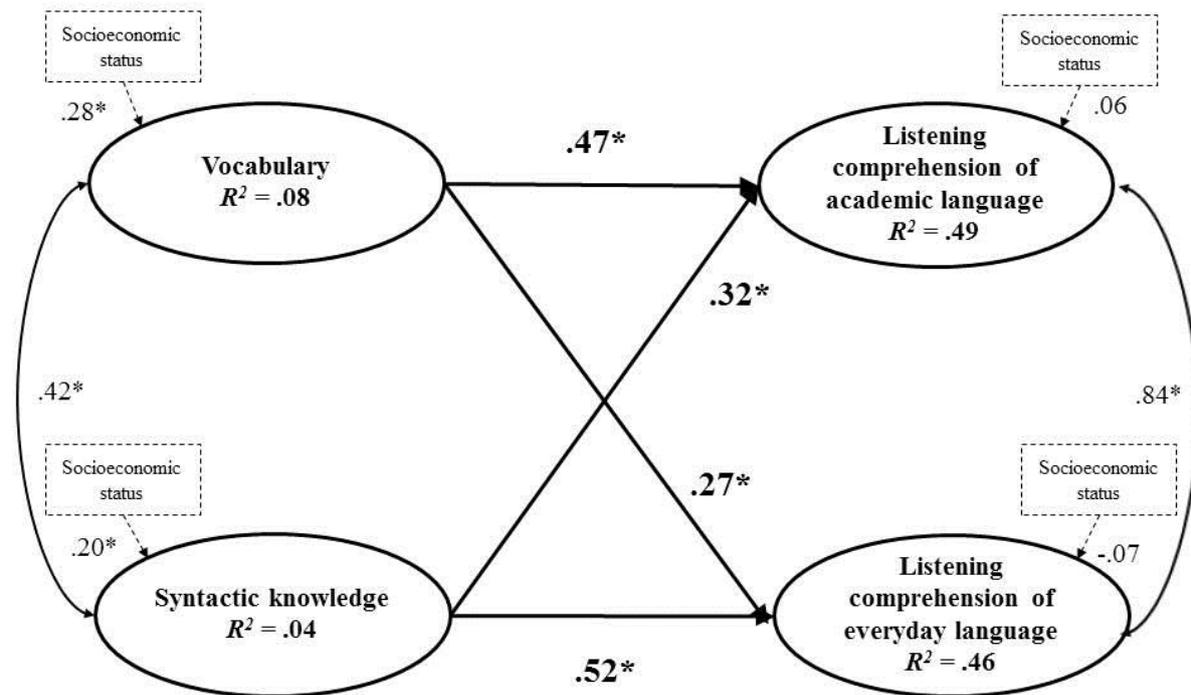


Figure 1. Correlations between listening comprehension of academic and everyday language and vocabulary and syntactic knowledge. Indicators are not depicted. $\chi^2 = 2720.58, df = 2546, p < .01, RMSEA = .01, CFI = .95; TLI = .95. (N = 459)$

6.4 Group Differences in Listening Comprehension of Academic Language and Listening Comprehension of Everyday Language

To test our second hypothesis on group differences between L1 and L2 students in terms of listening comprehension of everyday and academic language, we compared latent variable means between the groups. Table 3 shows the results for the test of measurement invariance for the two groups. Goodness-of-fit indices as well as χ^2 -statistics indicate that all models fit the data well. The nested model tests show that goodness-of-fit does not decrease in the more restrictive models. Measurement invariance for the two language groups may therefore be assumed.

Table 3

Test of Measurement Invariance for the Two-Dimensional Model (Controlling for ESCS)

model	χ^2	df	p	RMSEA	CFI	TLI	Compared to ...	$\Delta \chi^2$	df	p
configural	859.73	806	.09	.02	.95	.94	--	--	--	--
metric	870.46	833	.18	.01	.96	.96	configural	21.52	27	.76
scalar	897.40	860	.18	.01	.96	.96	metric	28.00	27	.41

Note. RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; $N = 459$.

We then tested our second hypothesis on differences in listening comprehension of academic and everyday language scores between L1 and L2 students by comparing latent variable means between the two language groups using Wald χ^2 -tests (Wald, 1943). Latent variable means of L2 students were compared to the reference group of L1 students for which the latent means were fixed to zero. The results in Table 4 indicate that L2 students performed less well than L1 students on listening comprehension of academic language comprehension as well as on listening comprehension of everyday language. Contrary to what we expected, differences between L1 and L2 students' listening comprehension of everyday language were more pronounced than differences between listening comprehension of academic language ($\chi^2(1) = 8.72, p < .01$).

Table 4

Variable Means for L1 and L2 Students

	L1 ($n = 173$)		L2 ($n = 286$)		Difference Testing for Latent Variable Means		
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	Wald χ^2	<i>p</i>	<i>d</i> ¹
Listening Comprehension (EL)	.73	.15	.65	.18	12.98	<.01	-0.65
Listening Comprehension (AL)	.60	.18	.50	.21	4.68	.03	-0.45

Note. AL = Academic Language; EL = Everyday Language. ¹Difference Between Latent Means of the L1 Group (reference group) and the L2 Group.

7. Discussion

The study investigated listening comprehension of academic and everyday language in students speaking German as a first or as a second language. It was hypothesized that listening comprehension of academic language can be separated empirically from listening comprehension of everyday language. We expected that L2 students would generally perform less well than L1 students, yet that the performance gap would be more pronounced for listening comprehension of academic language than for listening comprehension of everyday language. For both groups of students, we assumed that listening comprehension of academic language would be more difficult than listening comprehension of everyday language.

7.1 *Dimensionality of Listening Comprehension in L1 and L2 Learners*

Concerning the first research question on the structure of listening comprehension proficiency, findings of a confirmatory factor analysis indicated that a two-dimensional model fits the data better than a unidimensional model. Our results thus provide preliminary evidence that listening comprehension of academic and everyday language may be distinguished from each other empirically. This differentiation is in line with theoretical assumptions by Richards (1983) who postulated that conversational listening and academic listening rely on different skills as well as with assumptions that suppose that academic language proficiency is associated with mastery of specific genres, strategies, and linguistic features (Gottlieb & Ernst-Slavit, 2014; Nagy & Townsend, 2012; Scarcella & Rumberger, 2000; Snow & Uccelli, 2009). While previous studies have not empirically validated whether academic listening proficiency and everyday listening proficiency are distinguishable factors of listening comprehension, results of our study indicate that everyday listening and academic listening may indeed be associated with different facets of listening comprehension. It must be noted, however, that our study does not allow to differentiate whether our two-dimensional model reflects different aspects of language proficiency or differences in listening texts' difficulty. However, academic language is by definition more complex and difficult to process than everyday language, such that a greater difficulty of academic language demands is inherent to the construct and is hence also reflected in our operationalization of academic listening comprehension. Our findings thus may also be seen to corroborate results by Shohamy and Inbar (1991) who found that the amount of oral language features in listening texts had an impact on L2 test takers' listening comprehension. The present study adds to this research and indicates that differences in the processing of academic vs. everyday listening texts may not only be a mat-

ter of text type/difficulty, but that listening texts that are located near the extremes of the oral-literate continuum also may be associated with different facets of listening comprehension.

However, an additional validation procedure using students' vocabulary and their syntactic knowledge as external criteria did not support our assumption. Contrary to our predictions, students' listening comprehension of everyday language correlated more strongly with their syntactic knowledge than students' listening comprehension of academic language. We did not find significantly different correlations for the vocabulary measure although the association with vocabulary was somewhat stronger for listening comprehension of academic language than for listening comprehension of everyday language, as hypothesized. This absence of significantly differential relationships may be an artifact of the vocabulary measure applied in the present study, given that the CFT 20 R/WS does not distinguish between academic and everyday vocabulary. However, it is possible to classify the remaining target words used in the test as either academic vocabulary or everyday vocabulary. Based on the coding scheme by the CRESST (Butler, et al., 2004), we classified seven (30.4%) of the words as academic vocabulary and 16 words (69.6%) as everyday vocabulary ($\kappa = .66$; $p < .01$ for three raters). The vocabulary measure thus entails words from both facets of language comprehension which may explain the lack of differences in correlations between the vocabulary test and listening comprehension of academic and everyday language respectively. This assumption is supported by the fact that removing items containing everyday language from the vocabulary measure resulted in a significant correlation between vocabulary and academic listening comprehension ($r = .30$; $p = .02$) but a non-significant association between vocabulary and everyday language proficiency ($r = .16$; $p = .23$). The difference between the correlation coefficients is significant ($z = 2.24$; $p = .03$). The lack of expected associations between the vocabulary measure and listening comprehension of academic language may thus be due to the instrument we used to assess students' vocabulary. The syntactic knowledge measure only included everyday language. It may therefore be assumed that the higher correlation between syntactic knowledge and everyday listening as compared to the correlation between syntactic knowledge and academic listening also results from the operationalization.

7.2 Differences in Academic and Everyday Listening Proficiency Between L1 and L2 Learners

Concerning our second hypothesis on group differences in latent variable means for listening comprehension of academic and everyday language, our data failed to support the assumption that academic listening comprehension is especially demanding for L2 students.

Both groups of students showed weaker performance on the test of academic listening comprehension than on the test of everyday listening comprehension and L1 students generally outperformed their L2 peers. However, contrary to what was hypothesized, we did not find more pronounced disadvantages for adolescent L2 students for academic listening texts. We thus could not find support for a central assumption by Cummins (1980) and other researchers (cf. Collier, 1989; Roessingh, et al., 2005), namely that academic language is disproportionately more difficult for L2 students than for L1 students. Furthermore, our results question the postulated longer period of time it takes to master academic language in a second language. Various researchers have hypothesized that it may take L2 students five to seven years to acquire proficiency in the academic language register (e.g., Cummins, 1981; Collier, 1989). Our results suggest that even a longer period may not be sufficient for L2 students to obtain native-like levels of performance in academic language. The differential results on the time it takes to master L2 may again be explained to some extent by students' socioeconomic background which was very low in the present study. As research indicates, there is a significant correlation between SES and the rate of acquisition of academic language (e.g., Hoff, 2013) and it may be assumed that it may take somewhat longer than the postulated period for L2 students from low-SES families to master academic language.

There are several possible explanations for the lack of differences in academic listening proficiency between L1 and L2 students in the present study. One such explanation is that facets of academic vs. everyday language may be more clearly distinguishable in written than in oral texts, as oral academic language typically contains fewer academic language features than written academic language (cf. Corson, 1997; Nagy & Townsend; 2012). Another explanation pertains to the restricted range of students' socioeconomic background. The majority of students in our sample were from low-income and low-SES families. As research has indicated, the attainment of academic language is closely associated with students' socioeconomic background (Aarts, Demir, & Vallen, 2011; Hoff, 2013). Students that do rarely encounter an "academic culture" of literacy (Corson, 1997, p. 672) outside school are thus less likely to develop academic language competencies than students from more privileged families. It is thus possible that low-SES L1 students in our study did not have sufficient opportunities to acquire academic language and were therefore also compromised in their academic language skills.

Moreover, all students from our sample were from the lowest tracks of the German education system and tend to perform well below students from other school tracks. Although there are currently no empirical studies on the amount of academic language features in

German schools, international studies indicate that academic language is more prevalent in the higher tracks of the educational system, thus providing high-track students with more opportunities to gain proficiency in this language register (e.g. Harklau, 1994). Everyday and academic listening comprehension may therefore be more clearly distinguishable in higher tracks, where the materials used are more demanding in terms of academic language features.

Contrary to what was expected, differences between L1 students' and L2 students' everyday listening proficiency were more pronounced than differences in their academic listening proficiency. Given this finding, it seems questionable that everyday language is generally less complex and therefore easier to master for L2 students than academic language. While Cummins (1980) assumes that it takes L2 students approximately two years to master everyday language skills in their L2, our results indicate that everyday language proficiency may not be fully acquired even by adolescent L2 students who were born and raised in an L2-speaking environment. As Scarcella (2003) points out, both academic language and everyday language are complex registers and share a number of determinants (e.g., processing strategies) that affect their development. Everyday language may thus also entail features that are particularly difficult to master for L2 students and, in turn, lead to weaker performance on measures of everyday language as compared to L1 students. Our results are consistent with this assumption and show that even older, advanced L2 learners may exhibit difficulties not only with the academic but also with the more informal registers of the L2.

7.3 *Limitations*

The present study has some limitations. First, our sample was restricted to low-SES students from the lowest tracks of the educational system. Further studies are needed that investigate whether listening comprehension of academic and everyday language are more clearly distinguishable in students who have had more opportunities to gain proficiency in the register of academic language. Furthermore, the aim of the present study was to compare listening comprehension of academic and everyday language with a special emphasis on lexical and grammatical features of academic language (cf. Butler et al., 2004). However, academic language is not only characterized by linguistic features, but also includes cognitive and socio-cultural aspects, such as the use of listening strategies, higher-order thinking skills, or knowledge about social and cultural norms, values, and habits that shape the context of academic language (Scarcella, 2003). These aspects were not considered in our operationalization of academic language. Therefore, internal validity is ensured but ecological and external validity may be limited.

7.4 *Conclusion*

Despite these limitations, the present study indicates that listening comprehension of academic and everyday language may be separable facets of listening comprehension. However, as the external validation procedure did not fully substantiate this finding, and as we did not find the expected performance differences between L1 and L2 students, the present results only partially support the notion that academic language and everyday language are associated with distinct facets of listening proficiency. Nonetheless they may serve as a starting point to further examine whether listening comprehension of academic language and everyday language are clearly distinct. Our results provide first evidence that the processing of academic vs. everyday listening texts, at least when located somewhere near the extremes of the oral-literate continuum, may require different facets of listening comprehension proficiency.

As L2 students performed significantly below their native peers not only when listening to academic language texts but also when processing everyday language, our results further indicate that L2 instruction should focus on academic language *and* everyday language skills to help L2 students gain proficiency in all registers of their L2. In Germany, second-generation L2 students are typically placed in regular German classrooms. However, as our results suggest, simply exposing L2 students to an immersion context seems not to be sufficient for developing listening competencies that are comparable to those of their monolingual peers. Further research is needed to investigate which approaches can effectively help L2 students gain proficiency in all aspects of L2 listening comprehension.

8. References

- Aarts, R., Demir, S., & Vallen, T. (2011). Characteristics of academic language register occurring in caretaker-child interaction: Development and validation of a coding scheme. *Language Learning, 61*, 1173-1221. doi:10.1111/j.1467-9922.2011.00664.x
- Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education, 14*, 219-234. doi:10.1207/S15324818AME1403_2
- Bailey, A. L. (Ed.). (2007). *The language demands of school. Putting academic English to the test*. New Haven, CT: Yale University Press.
- Bailey, A. L., & Butler, F. (2003). *An evidentiary framework for operationalizing academic language for broad application to K-12 education: A design document*. Los Angeles, CA: University of California.
- Bailey, A. L., & Huang, B. H. (2011). Do current English language development/proficiency standards reflect the English needed for success in school? *Language Testing, 28*, 343-365. doi:10.1177/0265532211404187
- Bailey, A. L., Huang, B. H., Shin, H. W., Farnsworth, T., & Butler, F. A. (2007). *Developing academic English language proficiency prototypes for 5th grade reading: Psychometric and linguistic profiles of tasks. CRESST report 727*. Los Angeles, CA: The National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Baumgartner, H., & Steenkamp, J.-B. E. M. (1998). Multi-group latent variable models for varying numbers of items and factors with cross-national and longitudinal applications. *Marketing Letters, 9*, 21-35. doi:10.1023/A:1007911903032
- Berendes, K., Dragon, N., Weinert, S., Heppt, B., & Stanat, P. (2013). Hürde Bildungssprache? Eine Annäherung an das Konzept „Bildungssprache“ unter Einbezug aktueller empirischer Forschungsergebnisse [Academic language as an obstacle? Approaching the concept of "academic language" considering recent empirical findings.]. In A. Redder & S. Weinert (Eds.), *Sprachförderung und Sprachdiagnostik. Perspektiven aus Psychologie, Sprachwissenschaft und empirischer Bildungsforschung* (pp. 17-41). Münster: Waxmann.
- Biber, D. (1995). *Dimensions of register variation*. New York: Cambridge University Press.
- Buck, G. (2001). *Assessing listening*. Cambridge: Cambridge University Press.
- Butler, F. A., Bailey, A. L., Stevens, R., & Huang, B. H. (2004). *Academic English in fifth-grade mathematics, science, and social studies textbooks*. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

- Chamot, A. U., & O'Malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. Reading, MA: Addison-Wesley Publishing Company.
- Chudaske, J. (2012). *Sprache, Migration und schulfachliche Leistung*. [Language, migration, and school performance.]. Wiesbaden: VS Verlag.
- Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. *Educational Measurement: Issues and Practice*, 17, 31-44.
- Collier, V. P. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly*, 21, 617-641. doi: 10.2307/3586986
- Collier, V. P. (1989). How long? A synthesis of research on academic achievement in a second language. *TESOL Quarterly*, 23, 509-531. doi: 10.2307/3586923
- Corson, D. (1997). The learning and use of academic English words. *Language Learning*, 47, 671-718. doi: 10.1111/0023-8333.00025
- Csomay, E. (2006). Academic talk in American university classrooms: Crossing the boundaries of oral-literate discourse. *Journal of English for Academic Purposes*, 5, 117-135. doi: 10.1016/j.jeap.2006.02.001
- Cummins, J. (1980). The construct of language proficiency in bilingual education. In J. E. Alatis (Ed.), *Current issues in bilingual education* (pp. 81-103). Washington, D.C.: Georgetown University Press.
- Cummins, J. (2000). *Language, power, and pedagogy. Bilingual children in the crossfire*. Clevedon, UK: Multilingual Matters.
- Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. In B. Street & N. H. Hornberger (Eds.), *Encyclopedia of Language and Education, 2nd Edition* (Volume 2: Literacy, pp. 71-83). New York: Springer Science + Business Media LLC.
- Curenton, S. M., & Justice, L. M. (2014) African American and caucasian preschoolers' use of decontextualized language: Literate language features in oral narratives. *Language, Speech, and Hearing Services in Schools*, 35, 240-253. doi: 10.1016/j.lshs.2014.04.004
- Eckhardt, A. G. (2008). *Sprache als Barriere für den schulischen Erfolg: Potentielle Schwierigkeiten beim Erwerb schulbezogener Sprache für Kinder mit Migrationshintergrund*. [Language-related barriers in educational success: Potential difficulties in the acquisition of academic language in immigrant students.]. Münster: Waxmann.

- Edelsky, C., Hudelson, S., Flores, B., Barkin, F., Altwerger, B., & Jilbert, K. (1983). Semilingualism and language deficit. *Applied Linguistics*, 4, 1-22. doi: 10.1093/applin/4.1.1
- Ferris, D., & Tagg, T. (1996). Academic oral communication needs of EAP learners: What subject-matter instructors actually require. *TESOL Quarterly*, 30, 31-58. doi: 10.2307/3587606
- Flowerdew, J. (1995). Research of relevance to second language lecture comprehension: An overview. In J. Flowerdew (Ed.), *Academic listening: Research perspectives* (pp. 7-29). Cambridge: Cambridge University Press.
- Flowerdew, J. (1994). *Academic listening: Research perspectives*. Cambridge: Cambridge University Press.
- Flowerdew, J., & Miller, L. (1997). The teaching of academic listening comprehension and the question of authenticity. *English for Specific Purposes*, 16, 27-46. doi: 10.1016/S0889-4906(96)00030-0
- Ganzeboom, H. B. G., De Graaf, P. M., & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1-56. doi:10.1016/0049-089X(92)90017-B
- Gierl, M. J. (2005). Using dimensionality-based DIF analyses to identify and interpret constructs that elicit group differences. *Educational Measurement: Issues and Practice*, 24, 3-14. doi:10.1111/j.1745-3992.2005.00002.x
- Gottlieb, M., & Ernst-Slavit, G. (2014). *Academic language in diverse classrooms: Definitions and contexts*. Corwin Press.
- Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* Santa Barbara, CA: The University of California Linguistic Minority Research Institute.
- Harklau, L. (1994). Jumping tracks: How language-minority students negotiate evaluations of ability. *Anthropology and Education Quarterly*, 25, 347-363. doi: 10.1525/aeq.1994.25.3.04x0149s
- Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision: KFT 4-12+ R*. [Test of cognitive abilities for grades 4 to 12.]. Weinheim: Beltz.
- Heppt, B., Stanat, P., Dragon, N., Berendes, K., & Weinert, S. (2014). Bildungssprachliche Anforderungen und Hörverstehen bei Kindern mit deutscher und nicht-deutscher Familiensprache [Academic language demands and listening comprehension in

- children with German and Non-German home language]. *Zeitschrift für Pädagogische Psychologie*, 28, 139-149. doi: 10.1024/1010-0652/a000130
- Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: Implications for closing achievement gaps. *Developmental Psychology*, 49, 4-14. doi: 10.1037/a0027238
- Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55. Doi:10.1080/10705519909540118
- Mason, A. (1995). By dint of: Student and lecturer perceptions of lecture comprehension strategies in first-term graduate study. In J. Flowerdew (Ed.), *Academic listening: Research perspectives* (pp. 199-218). Cambridge: Cambridge University Press.
- Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, 78, 333-368.
- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide. Seventh edition*. Los Angeles, CA: Muthén & Muthén.
- Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47, 91-108. doi: 10.1002/RRQ.011
- Nystrand, M. (2006). Research on the role of classroom discourse as it affects reading comprehension. *Research in the Teaching of English*, 40, 392-412.
- Organization for Economic Cooperation and Development (2010). *PISA 2009 results: Overcoming social background – equity in learning opportunities and outcomes (Volume II)*. Paris: OECD.
- Richards, J. C. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly*, 17, 219-240. doi: 10.2307/3586651
- Roessingh, H., Kover, P., & Watt, D. (2005). Developing cognitive academic language proficiency: The journey. *TESL Canada Journal*, 23, 1-27.
- Rost, M. (1994). *Introducing listening*. London: Penguin.
- Saville-Troike, M. (1984). What really matters in second language learning for academic achievement? *TESOL Quarterly*, 18, 199-219. doi: 10.2307/3586690
- Scarcella, R. (2003). *Academic English: A conceptual framework*. Santa Barbara, CA: The University of California Linguistic Minority Research Institute.
- Scarella, R., & Rumberger, R. W. (2000) Academic English key to long-term success in school. *UCL MRI Newsletter*, 9, 1-2.

- Schleppegrell, M. J. (2004). *The language of schooling. A functional linguistics perspective*. Mahwah, NJ: Lawrence Erlbaum.
- Shaffel, J., Belton-Kocher, E., Glasnapp, D., & Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of English language learners and students with disabilities. *Educational Assessment, 11*, 105-126. doi:10.1207/s15326977ea1102_2
- Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question type. *Language Testing, 8*, 23-40. doi: 10.1177/026553229100800103
- Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science, 328*, 450-452. doi:10.1126/science.1182597
- Snow, C. E., & Uccelli, P. (2009). The challenge of academic language. In D. R. Olson & N. Torrance (Eds.), *The Cambridge handbook of literacy* (pp. 112-133). Cambridge: Cambridge University Press
- Tannen, D. (1982). The oral literate continuum of discourse. In D. Tannen (Ed.), *Spoken and written language* (pp. 1-6). Norwood, NJ: Ablex Publishing.
- Taylor, L., & Geranpayeh, A. (2011). Assessing listening for academic purposes: Defining and operationalising the test construct. *Journal of English for Academic Purposes, 10*, 89-101. doi: 10.1016/j.jeap.2011.03.002
- Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society, 54*, 426-482. doi:10.2307/1990256
- Weiß, R. H. (2006). *Grundintelligenztest Skala 2, CFT 20 - Revision (CFT 20-R)*. [Culture Fair Intelligence Test - Scale 2.]. Göttingen: Hogrefe.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Erlbaum.
- Wolf, M. K., Crosson, A. C., & Resnick, L. B. (2005). Classroom talk for rigorous reading comprehension instruction. *Reading Psychology, 26*, 27-53. doi: 10.1080/02702710490897518
- Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. A. (2007). *ACER ConQuest Version 2: Generalised item response modelling software*. Camberwell, Victoria: ACER Press.

Appendix A.

Variation of linguistic features for academic and everyday listening texts

Feature	EL ¹	AL ²
	<i>M (SD)</i>	<i>M (SD)</i>
1 Number of words per text	138.25 (21.78)	108.25 (14.08)
Lexical features		
2 Number of nominalizations per text	0.75 (0.96)	1.00 (0.00)
3 Number of academic vocabulary per text	2.50 (2.08)	8.50 (3.42)
Syntactic features		
4 Sentence length (words per sentence)	10.50 (4.68)	15.18 (2.13)
5 Number of complex subordinate clauses per text	0.75 (0.96)	1.00 (0.00)
6 Number of passive constructions per text	0.00 (0.00)	3.25 (2.87)
7 Number of structural words per text	58.08 (2.82)	56.32 (6.29)
Mean number of items correct		
Total sample	.70 (.17)	.53 (.20)
L1 students	.74 (.15)	.59 (.19)
L2 students	.68 (.18)	.49 (.20)

Note: AL = academic listening; EL = everyday listening;

¹contents: going shopping, doing homework, getting the driver's license, friendship;

²contents: haikus, ocelots, doing a job interview, planning a hike

Appendix B.

HAIKU

Short poems have been known since ancient Japan, where they were called “haikus”. The society at the emperor’s court developed them from a sort of tournament because they were frequently bored. Back then, the high society liked to compete in rhyming verses, however, strict rules had to be obeyed. With regard to content, for instance, a haiku had to portray an image from nature. Furthermore, it was allowed to include only a certain amount of syllables and always had to exhibit the same length. The haiku, therefore, was the shortest poetic form in the world. Even today lots of people engage in composing haikus in their leisure time. For the most part, the old rules were adopted, so that a lot of modern haikus appear as if they were written in ancient Japan. Thus, the haiku is often declared to be a timeless form of poetry.

Question 1

What do you need to make a good haiku?

- A You need to be good at writing.
- B You need to be good at listening.
- C You need to be good at competing.
- D You need to be good at drawing.

Code 1: A: You need to be good at writing.

Code 0: All other responses

Question 2

What is a haiku?

- A A Japanese game.
- B A short poem.
- C A name for ancient Japan.
- D A Japanese form of song.

Code 1: B: A short poem.

Code 0: All other responses

Question 3

In the text, something is said about why haikus are special. Which kind of „record“ do they hold?

Code 1: Mentions the fact that haikus are the shortest poetic form in the world

- They are the shortest.
- They are very short.
- The shortest poetic form (of the world).

Code 0: Answer is insufficient or vague. Shows insufficient understanding of the text or gives an implausible or irrelevant answer.

- A high record.
- They have strict rules.
- They are a timeless form of poetry.

Shopping

Salesperson: May I help you?

Client: Oh...well...yes, you can. I am looking for a.. this magazine for teens. That has a lot on music inside.

Salesperson: „Melody“, right?

Client: Yeah, exactly, that was it. Unfortunately, I forgot my shopping list at home, it was on it. Do you have it, „Melody“?

Salesperson: Look at your left, next to the box with the lighters.

Client: Where? I only see comics.

Salesperson: No, one box further to the left.

Client: Ah, now I got it! Usually, my daughter gets the money from me and then buys her stuff by herself, you know. That's why I'm not familiar here. But now she's broken her leg.

Question 1

What does the client want to buy?

- A A lighter
- B A newspaper
- C A comic
- D A magazine for teens

Code 1: D: A magazine for teens

Code 0: All other responses

Question 2

The client has problems finding what she is looking for. Why is that? Use the information you have heard to answer the question.

Code 1: Refers to the fact that the client is in the shop for the first time OR on the fact that the daughter is usually going shopping herself OR to the shopping list that has been forgotten

- Because, normally, the daughter goes shopping.
- She is in the shop for the first time and doesn't know where things are.
- She forgot her shopping list.
- Because she is not familiar with the shop.
- She's there for the first time.

Code 0: Answer is insufficient or vague. Shows insufficient understanding of the text or gives an implausible or irrelevant answer.

- The daughter has broken her leg.
- She is looking for a magazine.
- Because there are so many comics.
- Because it is somewhere else.

Question 3

In what kind of shop this conversation may have taken place? Use the information you have heard to answer the question.

Code 1: Mentions at least one sort of shop that usually sells magazines, comics, and tobacco supplies.

- At a kiosk
- At a lottery shop
- At a newspapers stand
- At a supermarket

Code 0: Answer is insufficient or vague. Gives an implausible or irrelevant answer or mentions shops that do not sell magazines and tobacco supplies at the same time.

- A shop where comics and magazines are sold.
- In Germany.
- In a comic store.

Question 4

In the text, the clients' daughter is mentioned. What do you think she does?

- A She goes to the kindergarten.
- B She visits highschool.
- C She's studying in another city.
- D She lives abroad.

Code 1: B: She visits high school.

Code 0: All other responses.

Erklärung

Hiermit versichere ich, die vorliegende Arbeit mit dem Titel „Hören und Verstehen: Struktur und Determination des Hörverstehens bei Schülerinnen und Schülern mit Deutsch als Erst- bzw. Zweitsprache“ selbstständig verfasst zu haben. Alle Hilfsmittel und sinngemäße sowie wörtliche Zitate, die bei der Erstellung der Arbeit verwendet wurden, sind unter Angabe der Quellen kenntlich gemacht. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Berlin, 30. November 2015