

Britta Juska-Bacher, Ladina Brugger, Annina Fischer  
& Martina Röthlisberger (éds.)

- Wortschatzkompetenzen definieren, erheben und fördern
- Defining, assessing and fostering vocabulary skills



# bulletin vals-asla

## numéro 113

Bulletin suisse  
de linguistique appliquée  
Été 2021

Vereinigung für angewandte  
Linguistik in der Schweiz

Associazion svizra  
da linguistica applitgada

Association suisse  
de linguistique appliquée

Associazione svizzera  
di linguistica applicata

Vereinigung für Angewandte  
Linguistik in der Schweiz

Association Suisse de  
Linguistique Appliquée

Associazione Svizzera di  
Linguistica Applicata

Associazion Svizra da  
Linguistica Applitgada



Le *Bulletin suisse de linguistique appliquée* est l'organe de publication de l'Association Suisse de Linguistique Appliquée (VALS/ASLA).

#### **Comité de relecture pour ce numéro**

Karin Aguado (Universität Kassel), Christine Beckert (Pädagogische Hochschule Bern), Raphael Berthele (Universität Freiburg), Noah Bubenhofer (Universität Zürich), Helen Christen (Universität Freiburg), Catherine Diederich (Pädagogische Hochschule St. Gallen), Andrea Ender (Universität Salzburg), Agnes Groba (Universität Leipzig), Astrid Jurecka (Goethe-Universität Frankfurt), Irina Hertel (Europa-Universität Flensburg), Irmtraud Kaiser (Universität Salzburg), Annegret Klassert (Fachhochschule Clara Hoffbauer Potsdam), Simone Lehrl (Universität Bamberg), Urs Maurer (University of Hongkong), Maik Philipp (Pädagogische Hochschule Zürich), John Read (University of Auckland), Pascale Schaller (Pädagogische Hochschule Bern), Hansjakob Schneider (Pädagogische Hochschule Zürich), Ophélie Tremblay (Université du Québec à Montréal), Jan Vanhove (Universität Freiburg), Letizia Volpin (Université de Neuchâtel), Anna Zimmermann-Stübe (Schweizer Hochschule für Logopädie).

Nous remercions vivement les membres du comité de lecture pour leur contribution précieuse à la qualité de ce volume.

Publié avec le soutien financier de l'Académie suisse des sciences humaines et sociales (ASSH / SAGW), le *Bulletin suisse de linguistique appliquée* paraît deux fois par an.

Chaque article paru dans le bulletin a été soumis à un processus de "peer review" anonyme.

Abonnement personnel	Suisse CHF 30.-	Etranger CHF 35.-
Abonnement institutionnel	Suisse CHF 50.-	Etranger CHF 55.-
Prix au numéro	Suisse CHF 25.-	Etranger CHF 30.-
Pas de réduction librairie		

#### **Rédaction**

Dr. Sara Cotelli Kureth (Université de Neuchâtel)  
E-mail: sara.cotelli@unine.ch  
Centre de langues, Av. du 1<sup>er</sup>-Mars 26, CH-2000 Neuchâtel

#### **Responsables comptes rendus**

Dr. Gilles Merminod (Université de Lausanne)  
E-mail: gilles.merminod@unil.ch

#### **Administration**

Florence Waelchli (Université de Neuchâtel)

#### **Abonnements, commandes**

Institut des sciences du langage  
Rue Pierre-à-Mazel 7, CH-2000 Neuchâtel  
E-mail: bulletin.valsasla@unine.ch / CCP: 20-7427-1

© Cette revue est distribuée selon la licence Creative Commons CC-BY 4.0.



# Table des matières

## **Wortschatzkompetenzen definieren, erheben und fördern / Defining, assessing and fostering vocabulary skills**

Britta JUSKA-BACHER, Ladina BRUGGER, Annina FISCHER & Martina RÖTHLISBERGER Wortschatz definieren, erheben, fördern .....	1-7
Wolfgang LENHARD & Alexandra LENHARD Bedeutung und Diagnostik des Wortschatzes am Beispiel des Peabody Picture Vocabulary Test (PPVT-IV) .....	9-27
Christian W. GLÜCK Der WWT 6-10 und seine Testitems .....	29-47
Britta JUSKA-BACHER & Martina RÖTHLISBERGER Das Konstrukt Wortschatz: Dimension(en) Umfang und Tiefe? Empirische Ergebnisse aus der Unterstufe .....	49-68
Christine COX ERIKSSON Researching vocabulary depth in a multilingual Swedish elementary school.....	69-87
Livia HAAG, Sarah V. DI PIETRO, Martina RÖTHLISBERGER, Carina WICK, Rita FÜZÉR & Silvia BREM Impact of parental reading skills and children's reading environment and interest at pre-school age on the vocabulary and reading skills in elementary school .....	89-116
Julia WINKES Die Bedeutung des mentalen Lexikons beim Wortlesen und -schreiben: Einsichten anhand der semantisch- lexikalischen Verarbeitungskompetenzen von Kindern mit und ohne Schriftspracherwerbsstörungen.....	117-135
Marie-Anne MORAND, Sandra SCHWAB & Stephan SCHMID Standarddeutsche Interferenzen im Dialektwortschatz Schweizer Jugendlicher: Lexikalische und lautliche Entlehnungen.....	137-154

Christina KAUSCHKE	
Der Wortschatz im Kindesalter: Entwicklung und Intervention .....	155-174
Christoph TILL	
Semantisch-lexikalische Diagnostik bei sukzessiv-bilingualen Kindern .....	175-192
Kerstin ALBER	
Wortschatzkompetenz erheben – eine synoptische Zusammenstellung wortschatzdiagnostischer Verfahren .....	193-212
Ladina BRUGGER & Britta JUSKA-BACHER	
Assessing primary grade children's lexical inferencing strategies while reading – A review .....	213-232

# Wortschatz definieren, erheben, fördern

## **Britta JUSKA-BACHER**

PHBern

Institut Primarstufe & Institut für Forschung, Entwicklung und Evaluation  
Fabrikstrasse 8, CH-3012 Bern  
britta.juska@phbern.ch

## **Ladina BRUGGER**

PHBern

Institut für Forschung, Entwicklung und Evaluation  
Fabrikstrasse 2a, CH-3012 Bern  
ladina.brugger@phbern.ch

## **Annina FISCHER**

Selbstständig erwerbend

Bachgässchen 6, CH-4125 Riehen  
anninafischer@hotmail.com

## **Martina RÖTHLISBERGER**

PHBern

Institut für Forschung, Entwicklung und Evaluation  
Fabrikstrasse 2a, CH-3012 Bern  
martina.roethlisberger@phbern.ch

Individuelles Wortwissen stellt einen zentralen Baustein sprachlicher Kommunikation dar, der grundlegend für rezeptive wie produktive Sprachhandlungen ist. In der Sprachdidaktik wird Wortschatz daher auch als "Schaltstelle des schulischen Spracherwerbs" bezeichnet (Steinhoff 2009). Die Frage, wie Wortwissen definiert wird, welche Dimensionen und Aspekte zu unterscheiden sind und wie diese erhoben werden, ist für dessen Erforschung grundlegend. In der Forschung gibt es diesbezüglich bisher kein Einvernehmen, so dass Operationalisierung und Messung von Wortschatz ein breites Spektrum aufweisen und eine Vergleichbarkeit von Studien schwierig ist. Der umfassendste Ansatz einer Spezifizierung von Wortwissen stammt aus der Fremdsprachdidaktik von Paul Nation (2001), der Form, Bedeutung und Gebrauch unterscheidet und diese jeweils in drei weitere Aspekte unterteilt. In der neueren Forschung werden häufig – nicht ohne Diskussion und Kritik – die zwei Dimensionen Wortschatzumfang und Wortschatztiefe voneinander abgegrenzt, wobei letztere alle Aspekte umfassen kann, die beschreiben, wie gut einem Sprechenden ein Wort bekannt ist (u.a. formales, semantisches und pragmatisches Wortwissen, siehe Nation 2001).

In der Linguistik steht das individuelle Wortwissen bisher vor allem im Fokus der Spracherwerbsforschung bei Kindern bis zum Alter von etwa vier Jahren. Dabei interessiert die Entwicklung des rezeptiven und produktiven Wortschatzumfangs und seine Zusammensetzung (nach semantischen Domänen und Wortarten), Aufbau und Entwicklung von Bedeutungswissen, die Organisation des kindlichen mentalen Lexikons, das Zusammenspiel von Wortschatz und Phonetik/Phonologie und Grammatik u.v.a.m. (für einen Überblick z.-B. Clark 2005; Kauschke 2012). Der fortgesetzte Erwerb im Schulalter ist in der Spracherwerbsforschung bisher immer noch eher zurückhaltend erforscht. Sprachdidaktische Stimmen hingegen betonten in den letzten Jahren vermehrt die zentrale Bedeutung des Wortschatzes als Grundlage aller Sprachlernbereiche sowie des Schulerfolgs im Allgemeinen (z.-B. Steinhoff 2013; Jurecka et al. 2019). Sie fanden abgesehen von wenigen Ausnahmen aber bisher in der linguistischen Forschung wenig Gehör. Wenn Wortschatz im Schulalter untersucht wird, stehen sehr oft seine Bedeutung für das Lesen und Schreiben oder eine Diagnose von Defiziten im Wortschatz im Vordergrund. Die Grundlagenforschung im Bereich Lesen (häufig aus der Psychologie) untersucht Wortschatz (meist den Umfang) als einen unter einer ganzen Reihe von Prädiktoren des Lesens. Die Diagnose und Förderung von Wortschatzdefiziten, sei es bei Kindern mit Sprachentwicklungsverzögerungen oder Kindern mit anderer Erstsprache, ist Thema der Sprachheilkunde und Fremdsprachdidaktik. Eine Beschreibung von Wortschatz und seiner Entwicklung als eigenständiger Kompetenzbereich im fortgesetzten ungestörten Erstspracherwerb – parallel zum frühen Spracherwerb – ist bisher im deutschsprachigen Raum selten erfolgt und stellt ein wichtiges Desiderat dar (z.-B. Kleinbub 2011; Steinhoff 2013; Leimbrink 2014).

Ziel des vorliegenden Bulletins ist es, Einblicke in die aktuelle Definition, Erhebung, Erforschung und Förderung von Wortschatz in verschiedenen wissenschaftlichen Disziplinen und damit in verschiedene Sichtweisen zu gewähren und einige daraus abgeleitete offene Forschungsfragen aufzuzeigen. Die Beitragenden der Artikel in diesem Band stammen aus der Linguistik, Logopädie, Psychologie, Sprachdidaktik und Sprachheilpädagogik und nehmen sowohl Wortschatzkompetenzen in Deutsch als Erst- wie auch als Zweitsprache in den Blick. Fokussiert werden Kinder im Vorschul- und Primarschulalter sowie Jugendliche. Die Beiträge lassen sich in vier Blöcke gliedern.

Der erste Block enthält Fokusartikel zu zwei zentralen Wortschatz-Testinstrumenten des deutschsprachigen Raums für das Primar- (und Sekundar-)Schulalter. Den Auftakt bildet ein Beitrag von **Wolfgang Lenhard** und **Alexandra Lenhard**, die den ursprünglich englischsprachigen, international breit eingesetzten "Peabody Picture Vocabulary Test" (PPVT-IV für 3- bis 16-Jährige) gemeinsam mit Robin Segerer und Sebastian Suggate ins Deutsche übertragen haben. Der Beitrag geht einleitend auf die Bedeutung

des Wortschatzes, u.a. für den schulischen Wissenserwerb, ein und stellt seinen Zusammenhang mit der kognitiven Entwicklung und der Intelligenz dar. Die beiden Psycholog(inn)en geben einen Einblick in den Ablauf und das Aufgabenmaterial des PPVT, der auszugsweise den rezeptiven Wortschatzumfang von Kindern und Jugendlichen misst. Anhand der Normierungsdaten werden die Entwicklung des Wortschatzumfangs sowie die Effekte verschiedener Einflussfaktoren beschrieben.

Der zweite Beitrag des Sprachheilpädagogen **Christian W. Glück** stellt den von ihm entwickelten "Wortschatz- und Wortfindungstest für 6- bis 10-Jährige" (WWT 6–10) für deutsch- und türkischsprachige Kinder vor. Dieser testet den rezeptiven und den produktiven Wortschatzumfang und erlaubt darüber hinaus Aussagen zur Organisation des mentalen Lexikons. Über eine Itemanalyse wird ein Einblick in die Testanlage und -konstruktion gewährt. Dabei spricht der Autor verschiedene Faktoren an, die einen zentralen Einfluss auf den Abruf von Wörtern aus dem mentalen Lexikon und damit den Testerfolg haben. Dies sind einerseits die spezifischen Eigenschaften der Wörter wie ihre Länge, Konkretheit und Wortart sowie ihre Verwendung in der Sprachgemeinschaft (Benennungsübereinstimmung, Häufigkeit). Andererseits sind es individuelle Faktoren der Kinder wie das Erwerbsalter, die Familiarität mit den Items, Sprachlernbiographie sowie der Testaufbau (Auswahl und Abfolge der Items).

Den zweiten Block bilden fünf empirische Studien zu Wortschatz und seiner Entwicklung auf der Primar- und Sekundarstufe I. Die Beiträge sind nach dem Alter der Probandinnen und Probanden angeordnet. Der erste Artikel von **Britta Juska-Bacher** und **Martina Röthlisberger** beschreibt aus der linguistischen Perspektive die Dimensionen Wortschatzumfang und -tiefe von Erst- bis Drittklässlern mit Deutsch als Erstsprache. Für die Operationalisierung von Wortschatztiefe werden semantisches und relationales Wortwissen analysiert. Ziel ist es, die für diese Altersstufe in der Literatur diskutierte Frage nach der Anzahl Dimensionen des Wortwissens und die angemessenen Testverfahren zu beleuchten. Mittels Faktorenanalysen wird für die ersten drei Schuljahre belegt, dass eine Trennung der Dimensionen Umfang und Tiefe wichtig ist, um Variation im individuellen Wortwissen zu erklären und wichtige Hinweise für die Förderung zu liefern.

Auch der Beitrag der Sprachdidaktikerin **Christine Cox Eriksson** beschäftigt sich mit der Messung von Wortschatztiefe in Form von semantischem Wortwissen. Die methodischen Möglichkeiten und Herausforderungen dieser Messungen durch Wortdefinitionen und Assoziationsaufgaben in der internationalen und schwedischen Forschung werden aufgezeigt und diskutiert. Empirische Daten von Zweit- und Fünftklässlern mit Schwedisch als Erst- bzw. Zweitsprache zeigen die Kompetenzentwicklung in diesem Zeitraum sowie die Unterschiede zwischen erst- und zweitsprachigen Kindern auf.

Die Studie von **Livia Haag, Sarah Di Pietro, Martina Röthlisberger, Carina Wick, Rita Füzér** und **Silvia Brem** aus der psychologischen Leseforschung untersucht den Einfluss des vorschulischen Leseumfelds, des IQ sowie elterlicher Lesekompetenzen und des sozioökonomischen Status auf den Wortschatz und die Lesefertigkeiten von Fünftklässlern. Korrelationsanalysen und Regressionen zeigen, dass wichtige Prädiktoren für den Wortschatzumfang der Fünftklässler insbesondere der non-verbale IQ und die Lesefertigkeiten der Eltern, aber auch das vorschulische Leseumfeld, d.h. das gemeinsame Lesen mit den Eltern, sind.

**Julia Winkes** vergleicht in ihrer logopädischen Studie semantisch-lexikalische Verarbeitungsfähigkeiten von drei Gruppen von Sechstklässlern: einer Gruppe mit durchschnittlichen Lese- und Rechtschreibfähigkeiten, einer mit Lese- und Rechtschreibstörungen und einer Gruppe mit isolierten Rechtschreibstörungen. Die statistischen Analysen belegen für die beiden letzteren Gruppen signifikant schlechtere Leistungen im semantisch-lexikalischen Bereich, d.h. des Wortschatzumfangs, des wortspezifischen orthographischen Wissens sowie der morphematischen Bewusstheit, nicht jedoch der strategischen Abrufflüssigkeit. Leseschwierigkeiten sind mit einem Defizit bei der Benennungsgeschwindigkeit verbunden.

Mit Sekundarschülerinnen und -schülern befassen sich **Marie-Anne Morand, Sandra Schwab** und **Stephan Schmid** in ihrer linguistischen Untersuchung von Multiethnolekt und dessen Wahrnehmung, indem sie die von Jugendlichen mit Schweizerdeutsch als Erst- oder Zweitsprache in einer Bildbeschreibung verwendeten lexikalischen und phonetischen Innovationen im Vergleich zum traditionellen Zürichdeutsch auswerten. Als Instrument wurde eine Bewertung der Multiethnolektalität durch eine weitere Gruppe Jugendlicher eingesetzt. Die Ergebnisse dieser Momentaufnahme zeigen, dass der Dialektwortschatz von Zürcher Jugendlichen mehr standarddeutsche Interferenzen aufweist als vor dem Hintergrund der Diglossiesituation im traditionellen Zürichdeutschen zu erwarten wäre und dass diese Interferenzen vor allem lautlicher Art sind. Bei Jugendlichen mit anderer Erstsprache als Schweizerdeutsch zeigt sich der Ansatz einer Vermischung von Dialekt und Standard.

An den empirischen Block schliessen zwei Beiträge aus sprachheilkundlicher Perspektive zur Diagnose und Förderung von Wortschatzdefiziten bei Kindern mit lexikalisch-semantischen Störungen an. **Christina Kauschke** gibt in ihrem Artikel einen Überblick über Aufbau und Entwicklungen des (klein-)kindlichen mentalen Lexikons und geht dabei sowohl auf einen ungestörten wie einen gestörten Verlauf ein. Im Zentrum des Beitrags stehen Vorschulkinder mit lexikalischen Beeinträchtigungen und Interventionsansätze zu ihrer Förderung. Diese hat das Ziel, Wortschatzumfang und -tiefe zu verbessern, um mögliche negative Konsequenzen für das weitere Lernen und die Entwicklung zu reduzieren. Die Autorin spricht wesentliche Bestandteile – wie



die Auswahl und Häufigkeit der Präsentation der Zielwörter – für eine erfolgreiche Intervention an. Eine solche sollte immer auf die individuellen Bedürfnisse des Kindes zugeschnitten sein.

**Christoph Till** thematisiert die Diagnose des beeinträchtigten kindlichen Wortschatzerwerbs in einer Zweitsprache. Bei sukzessiv-bilingualen Kindern, d.h. Kindern, die mit drei Jahren oder später eine Zweitsprache erwerben, können die gleichen Wortschatzdefizite auftreten wie bei Kindern mit Wortfindungsstörungen. Eine Unterscheidung zwischen diesen beiden Gruppen ist notwendig, da sie eine unterschiedliche Therapie benötigen: Während für letztere eine Wortfindungstherapie in beiden Sprachen nötig ist, ist für erstere u.U. eine in der Zweitsprache ausreichend. Im Beitrag werden verschiedene Diagnoseverfahren für Wortfindungsstörungen bei bilingualen Kindern, ihre Stärken und Schwächen vorgestellt.

Den vierten und letzten Block bilden zwei Reviewpapers zu Wortschatztests bzw. zum lexikalischen Inferieren. **Kerstin Alber** steuert mit ihrem Text einen synoptischen Überblick über Wortschatz-Diagnoseinstrumente für Kinder im Alter von 0 bis 12 Jahren bei. Sie beginnt den Beitrag mit theoretischen Grundlagen zur Operationalisierung von Wortschatzkompetenzen und nimmt dabei besonders die drei Dimensionen Wortschatzbreite, -tiefe und -geläufigkeit in den Blick, die sie mit der Unterscheidung in repräsentatives Wissen und prozedurale Fähigkeiten kombiniert. In Anlehnung an die Klassifikationen für Sprachentwicklungsstörungen und den Zweitspracherwerb wird eine neue Klassifikation der lexikalischen Kompetenz abgeleitet, welche die Grundlage der Zusammenstellung bildet. Die Autorin geht auf eine Reihe verschiedener deutschsprachiger Instrumente wie Elternfragebögen, Sprachbeobachtungsbögen, Screenings und Sprachtests ein und zeigt auf, welche Bereiche der lexikalischen Kompetenz in der Sprachdiagnostik bisher nicht ausreichend berücksichtigt werden.

Den Abschluss des Bandes bildet das Paper von **Ladina Brugger** und **Britta Juska-Bacher**, das sich mit lexikalischem Inferieren, d.h. der Bedeutungserschliessung unbekannter Wörter beim Lesen von Primarschülerinnen und Primarschülern beschäftigt. Das Inferieren bildet die Voraussetzung für den Ausbau des rezeptiven Wortschatzes. Während sich die Forschung zu lexikalischen Inferenzstrategien und -prozessen mehrheitlich auf zweitsprachliche Kontexte und ältere Lernende konzentriert, stellt der vorliegende Artikel Unterstufenkinder und Inferenzprozesse in der Erstsprache in den Fokus und diskutiert für die Zielstufe angemessene Erhebungsmethoden, ihre Grenzen und Möglichkeiten und schlägt weitere Forschungsperspektiven vor.

## Fazit und Ausblick

Die in diesem Bulletin enthaltenen Beiträge vermitteln einen Eindruck der Bemühungen aus verschiedenen fachlichen Perspektiven, das komplexe Phänomen Wortschatz zu definieren, messbar und damit auch förderbar zu machen.

Von angewandter Seite her zeigt sich klar der Wunsch, Wortschatz in verschiedenen Dimensionen möglichst differenziert zu beschreiben und zu erheben. Während es hinsichtlich des Wortschatzumfangs kaum Diskussionen gibt und auch im Deutschen verschiedene standardisierte Testinstrumente vorliegen, gibt es hinsichtlich der Wortschatztiefe bisher keine Einigkeit. Wenn auch seit Nation (2001) eine umfassende theoretische Übersicht vorliegt, was Wortwissen beinhaltet, fehlen jedoch in der Umsetzung Erhebungs-Diagnoseinstrumente, die semantisches und formales Wortwissen sowie Wissen zum Gebrauch von Wörtern angemessen berücksichtigen. Die Entwicklung von Testinstrumenten, die möglichst breit auch verschiedene Aspekte der Tiefe trennscharf und altersgemäss abdecken und auf die jeweilige Sprache abgestimmt sind, ist dringend nötig, einerseits um Grundlagen des Wortwissens und seiner Entwicklung im Schulalter zu erforschen und andererseits um die Voraussetzung für eine differenzierte und auf das einzelne Kind abgestimmte Förderung zu schaffen.

Die empirischen Beiträge aus den verschiedenen Disziplinen zeigen, dass die vorhandenen Testinstrumente zwar eine gute Grundlage für die Erhebung des Wortschatzumfangs bieten, eine umfassende Erhebung der Wortschatzqualität derzeit aber noch nicht möglich ist. Wortschatz wird im deutschsprachigen Raum daher oft eher "miterhoben" statt im Fokus der Aufmerksamkeit zu stehen. Umfassende Testinstrumente könnten der Wahrnehmung von Wortschatz als eigenständiger und komplexer Kompetenzbereich über das Kleinkindalter hinaus auch in der linguistischen Forschung Vorschub leisten, seine differenzierte Erhebung ermöglichen und davon ausgehend die "Schaltstelle des schulischen Spracherwerbs" gezielt fördern.

Neben diesem angewandten Bereich harren einige Grundlagen individuellen Wortwissens sowie des Wortschatz-Erwerbsprozesses im Schulalter und in der Adoleszenz der Erforschung: Wie entwickeln sich die verschiedenen Wortschatzdimension Umfang und Tiefe und damit das mentale Lexikon im Laufe der Schulzeit? Wie hängen die verschiedenen Bereiche des Wortwissens und ihre Entwicklung zusammen? Wie sieht der Erwerbsprozess neuer Wörter in Mündlichkeit und Schriftlichkeit aus? Und welche Bedeutung haben die verschiedenen Wortschatzbereiche für rezeptive und produktive mündliche wie schriftliche Sprachhandlungen?

Wir danken den in- und ausländischen Kolleginnen und Kollegen für ihre Beiträge, die das Thema Wortschatz ins Licht rücken, den anonymen Gutachtenden für ihre konstruktive Kritik zu den Beiträgen sowie Sara Cotelli, der Chefredakteurin des Bulletin Suisse de Linguistique Appliquée, für ihre hilfreiche Unterstützung bei der Entstehung des Heftes.

Bern, im Oktober 2021

## LITERATUR

- Clark, E. V. (2005). Semantic categories in acquisition. In H. Cohen & C. Lefebvre (Hgg.), *Handbook of Categorization in Cognitive Science* (S. 459-479). London: Elsevier.
- Jurecka, A., Cinar, M. & Hardy, I. (2019). Messung von Wortschatztiefe und -breite bei mono- und bilingualen Vorschulkindern. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 24(1), 145-163.
- Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen: Verläufe, Forschungsmethoden, Erklärungsansätze*. Berlin: De Gruyter.
- Kleinbub, I. (2011). Empirische Forschungsansätze im Bereich Wortschatzkompetenz. In I. Pohl & W. Ulrich (Hgg.), *Wortschatzarbeit* (S. 503-506). Baltmannsweiler: Schneider.
- Leimbrink, K. (2015). Wortschatzerwerb. In U. Haß & P. Storjohann (Hgg.), *Handbuch Wort und Wortschatz* (S. 27-52). Berlin: Mouton de Gruyter.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Steinhoff, T. (2009). Der Wortschatz als Schaltstelle des schulischen Spracherwerbs. *Didaktik Deutsch*, 27, 33-52.
- Steinhoff, T. (2013). Wortschatz – im Zentrum von Sprachgebrauch und Kompetenzförderung. In S. Gailberger & F. Wietzke (Hgg.), *Handbuch Kompetenzorientierter Deutschunterricht* (S. 12-29). Weinheim/Basel: Beltz.



# Bedeutung und Diagnostik des Wortschatzes am Beispiel des Peabody Picture Vocabulary Test (PPVT-IV)

## Wolfgang LENHARD

Universität Würzburg  
Institut für Psychologie  
Wittelsbacherplatz 1, D-97074 Würzburg  
wolfgang.lenhard@uni-wuerzburg.de

## Alexandra LENHARD

Psychometrica  
Am Kreuz 14, D-97337 Dettelbach  
lenhard@psychometrica.de

Vocabulary is a fundamental determinant of language development. In the Cattell-Horn-Carroll model of intelligence (CHC) it represents a narrow ability loading on the broad comprehension-knowledge (Gc) factor of intelligence. Furthermore, it is also an important predictor of reading comprehension. The Peabody Picture Vocabulary Test (Version 4) aims at assessing receptive vocabulary. Here, we describe the German adaption of the test. In the construction, we matched word frequencies and item complexity of the original form. The normative data is based on a representative sample of  $N = 3550$  children and adolescents with an age range from 2.59 to 17.99 years. The test features excellent reliability. The raw scores display a strong curvilinear development during childhood and a tremendous heterogeneity within each age group. The best performing children at the age of 5 achieved raw scores which the poor performing children did not achieve until the age of 15. While effects of sex were negligible, migration background of the family had a strong effect. In sum, the PPVT-IV is an economic and reliable instrument to assess the receptive vocabulary from an early age until adulthood.

### Keywords:

vocabulary assessment, bilingualism, cognitive abilities, word knowledge, language development.

### Stichwörter:

Wortschatztest, Bilingualismus, kognitive Fähigkeiten, Wortschatzwissen, Sprachentwicklung.

## 1. Die Bedeutung des Wortschatzes für menschliche Kognition

Der Begriff "Wortschatz" ist sehr schillernd, da er selbst auf seine enorme Bedeutung verweist. Ein Schatz ist etwas Kostbares, das man hegt und pflegt und zu vergrößern trachtet. Während materielle Schätze auch schlicht gefunden oder geerbt werden können, verhält es sich mit dem Wortschatz auf individueller Ebene anders. Der Erwerb des Wortschatzes stellt eine Form von Wissenserwerb dar, welcher sich über einen langen Zeitraum hinzieht und erst im mittleren Erwachsenenalter den Höhepunkt erreicht (Wechsler 2008). Er findet entweder vorwiegend implizit im Rahmen des Erstspracherwerbs oder explizit im Rahmen von formellem (Zweit- oder Fremd-)Sprachunterricht statt. Welche spezifischen Wörter gelernt werden, hängt nicht nur von der erlernten Sprache, sondern auch vom Kontext des Erwerbs ab, also davon, ob die Wörter

in häuslichen, akademischen, beruflichen oder anderen Situationen erworben werden. Wortbedeutungen können sich dabei mit der Zeit nicht nur gesamtgesellschaftlich, sondern auch auf individueller Ebene verändern, d. h. sie werden durch neue Bedeutungsfacetten bereichert, verändern ihren semantischen Gehalt, werden von anderen Wörtern verdrängt und möglicherweise zu einem späteren Zeitpunkt erneut reaktiviert. Auf individueller Ebene übersteigt der Wortschatz einer Sprache das Wortschatzwissen der einzelnen Personen deutlich. Anders als andere Fähigkeits- und Wissensbereiche wie beispielsweise das Phoneminventar einer Sprache kann der Wortschatz also fortlaufend erweitert und vertieft werden und ist somit ein lohnenswertes Ziel für jegliche Form von Bildungsprozessen.

### *1.1 Wortschatz und Intelligenz*

Die Bedeutung des Wortschatzes liegt nicht alleine im Umfang des verfügbaren Vokabulars, sondern auch in seiner Tiefe, also der Reichhaltigkeit der verfügbaren Wortbedeutungen, der Flexibilität der Anwendung und den Querverbindungen oder Relationen zwischen sprachlichen Konzepten. Wissen organisiert sich meist um eine spezifische Terminologie, bzw. beinhaltet diese, sodass sich i. d. R. hohe Korrelationen zwischen bereichsspezifischem Wissen und Wortschatz ergeben (z. B. Cromley & Azvedo 2007). Auch beeinflusst die Größe des Wortschatzes maßgeblich das Ziehen von Schlussfolgerungen beim Lesen von Sachtexten und damit das gesamte Textverständnis.

Die Bedeutung des Wortschatzes geht also weit über eine rein sprachliche Ebene hinaus und kann als ein konstitutives Merkmal menschlicher Kognition angesehen werden. Die Stellung des Wortschatzes innerhalb des Spektrums menschlicher Intelligenzleistungen ist durch weithin anerkannte faktorenanalytische Modelle der Intelligenz empirisch hervorragend belegt (Schneider & McGrew 2018; vgl. Abbildung 1). So enthalten viele wichtige Intelligenztests (z. B. die Wechsler-Tests und die Kaufman-Tests) auch Untertests zur Erfassung des Wortschatzes. Der Anteil sprachlicher Inhalte an Intelligenztests wurde im Laufe der Jahrzehnte tendenziell sogar aufgewertet, da sich der Umgang mit sprachlichem Material als ein fast unverzichtbarer Prädiktor zur Vorhersage der allgemeinen Intelligenz einer Person herausstellte. Testverfahren, die auch sprachliche Aufgaben beinhalten, sagen schulische Leistungen besser vorher als rein non-verbale Tests (Roth et al. 2015; A. Lenhard 2020). Das gilt nicht alleine für sprachliche Fächer, sondern auch für das Schulfach Mathematik.

Eines der aktuell etabliertesten Intelligenzmodelle ist das Cattell-Horn-Carroll-Modell (CHC; z. B. Schneider & McGrew 2018), das auf Grundlage hunderter empirischer Primärstudien konzipiert wurde. Das Modell listet über 80 sog. engere Intelligenzfähigkeiten auf, die anhand des Kovarianzmusters zu breiteren Intelligenzfaktoren gruppiert werden. Die gemeinsame Varianz aller Intelligenzleistungen wird als allgemeine Intelligenz G bezeichnet. Unter den

breiten Faktoren finden sich beispielsweise die *Arbeitsgedächtniskapazität*  $G_{wm}$ , die *Verarbeitungsgeschwindigkeit*  $G_s$ , die sogenannte *fluide Intelligenz*  $G_f$  (d. h. die Fähigkeit zum Erkennen und Anwenden von Mustern und Gesetzmäßigkeiten), aber auch das *Sprachverständnis und -wissen* ( $G_c$ ). Der letztgenannte Faktor wird dabei durch die engen Fähigkeiten *Sprachentwicklung, allgemeines Wissen, Hörverständnis, Kommunikationsfähigkeit, Grammatikempfinden*, aber eben auch *lexikalisches Wissen* – sprich: Wortschatz – aufgespannt. Die Bedeutung der engen Fähigkeiten für die breiten Faktoren lässt sich in den statistischen Modellen über die Faktorladungen spezifizieren, mit denen die Einzelleistungen auf dem übergeordneten latenten Konstrukt laden. Betrachtet man die bestehenden Modelle, so zeigt sich, dass Wortschatzleistungen unter den verschiedenen Einzelleistungen die höchsten Ladungen auf  $G_c$  aufweisen und dieser Faktor lädt wiederum hoch auf der allgemeinen Intelligenz (z. B. im Testverfahren WISC-V; Wechsler 2014). Der Wortschatz ist somit nicht nur einer der besten Indikatoren für das Sprachverständnis, sondern auch für die allgemeine Intelligenz selbst. In Large-Scale-Studien wird der rezeptive Wortschatz deshalb zuweilen sogar als alleiniger Indikator für  $G_c$ , manchmal sogar auch als Schätzer für die allgemeine Intelligenz  $G$  erhoben (z. B. British Cohort Study, Bynner & Parsons 2005). Auch Carroll (1993) zog in seinen Arbeiten den Schluss, dass Wortschatz eine sehr hohe Überlappung mit der breiteren Dimension "verbale Intelligenz" aufweist und unterstrich damit die Bedeutung des Wortschatzes für die kognitiven Fähigkeiten einer Person.

Wortschatz und Intelligenz hängen vermutlich auf vielfältige Weise miteinander zusammen und bedingen sich gegenseitig (vgl. A. Lenhard et al. 2015). Beispielsweise können Personen mit höherer fluider Intelligenz die Bedeutung von Wörtern besser aus dem Kontext erschließen und implizit einen größeren Wortschatz erwerben. Auf der anderen Seite erleichtert ein besser vernetzter Wortschatz vermutlich wiederum das Erkennen von sprachlichen Konzepten und Zusammenhängen, wie die Identifikation von Gemeinsamkeiten und Bedeutungsunterschieden, d. h. das logische Denken mit Sprache. Die Herausbildung von Unter- und Oberbegriffen wird unterstützt und neues Wissen besser organisiert und schneller erworben. Ein großer Wortschatz erleichtert zudem viele andere akademische Tätigkeiten, sodass Personen mit einem größeren Wortschatz tendenziell mehr lesen und beim Lesen wiederum mehr neuen Wortschatz erwerben (Pfost et al. 2013). Sie verfügen über einen größeren Umfang an sprachlichen Konzepten, auf die sie als Werkzeug beim Denken zugreifen können. Es lässt sich somit festhalten, dass der Wortschatz zwar einen abgrenzbaren Bereich innerhalb der Palette kognitiver Fähigkeiten und Fertigkeiten aufspannt, aber mit vielen weiteren Intelligenzfacetten interagiert und deshalb zugleich Ursache und Wirkung der Intelligenzentwicklung darstellt.

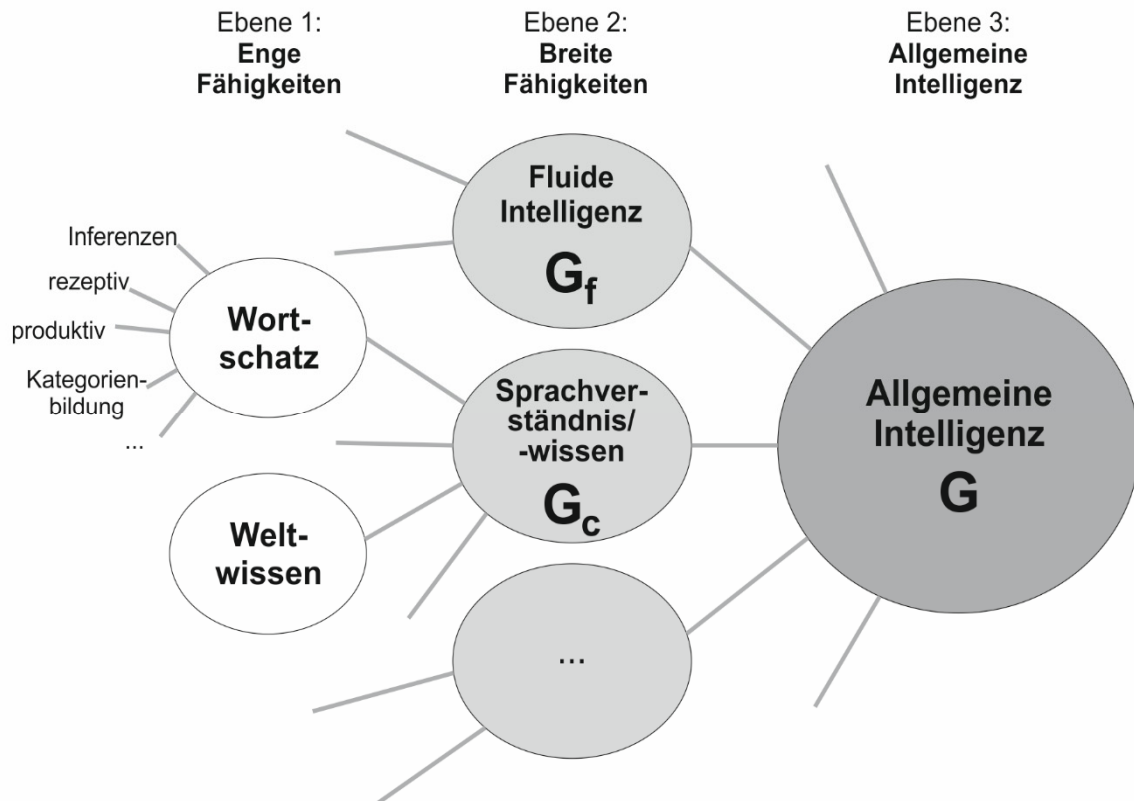


Abbildung 1: Das CHC-Rahmenmodell der Intelligenz. Das Cattell-Horn-Carroll-Modell (CHC; z. B. Schneider & McGrew 2018) definiert Intelligenz als hierarchische Struktur mit eng umrissenen kognitiven Fähigkeiten, die sich auf Ebene 2 des Modells zu breiten Faktoren gruppieren. Die allgemeine Intelligenz ist in diesem Modell die gemeinsame Varianz, die sich aus den Intelligenzfaktoren extrahieren lässt. Manche Aufgaben, wie z. B. das Auffinden von Gemeinsamkeiten oder Unterschieden zwischen Begriffen, hängen sowohl mit  $G_c$  als auch mit  $G_f$  eng zusammen.

## 1.2 Wortschatz, Schriftsprache und schulischer Wissenserwerb

Je jünger Kinder sind, desto weniger formell erfolgt in der Regel der Wortschatzerwerb. Doch bereits im ersten Lebensjahr werden die Grundlagen für die spätere Größe des Wortschatzes gelegt. So wirkt sich gemeinsames Bücherlesen vermutlich bereits Ende des ersten Lebensjahres förderlich aus (Lenhart et al., im Druck). Zunächst besteht dieses gemeinsame "Lesen" natürlich v. a. im Benennen und Zeigen von Objekten in Bilderbüchern, später dann im (dialogischen) Vorlesen oder dem freien Geschichtenerzählen. Je höher der Grad an Interaktion ist, desto mehr neuer Wortschatz wird erworben. In eigenen Untersuchungen erwies sich insbesondere das freie Geschichtenerzählen im Vorschulalter dem Vorlesen von Geschichten als überlegen, da es dabei ein größeres Ausmaß an Interaktionen gibt (Suggate et al. 2021). Der Lerneffekt kann noch gesteigert werden, indem Kinder die Geschichten nicht nur hören, sondern anschließend selbst nacherzählen ("Interactive Elaborative Story Retelling"; Vaahtoranta et al. 2019).



Ein sehr starker Zusammenhang besteht natürlich nicht nur zwischen Wortschatz und allgemeiner Intelligenz, sondern auch spezifisch zwischen Wortschatz und Leseverständnis (Beck et al. 1982; Verhoeven et al. 2011). Zusammen mit anderen Intelligenzleistungen wie z. B. Arbeitsgedächtniskapazität und grammatikalischen Fähigkeiten ist die Kenntnis von Wortbedeutungen von zentraler Bedeutung beim verstehenden Lesen von Texten. Cromley und Azevedo (2007) quantifizierten die direkten und indirekten Einflüsse verschiedener Determinanten des Leseverständnisses, wie z. B. Wortschatz, Leseflüssigkeit, Hintergrundwissen, Strategiewissen und schlussfolgerndes Denken. Der Wortschatz erwies sich in dieser Untersuchung als der mit weitem Abstand bedeutendste Einflussfaktor. Wie bereits dargestellt, erfassen Menschen die Bedeutung von neuen Wörtern unter anderem, indem sie den Sinn unbekannter Wörter, denen sie im Text begegnen, aufgrund des Kontexts erschließen (siehe Beitrag von Brugger & Juska-Bacher 2021). Allerdings ist es für Personen mit einem geringen Wortschatz oftmals überproportional anstrengend, Texte zu verstehen, da der Verständnisprozess umso brüchiger wird, je größer die Verständnislücken sind. In der Folge kann sich ein Teufelskreis aus Lesevermeidung, mangelnder Automatisierung, und geringerem Wissens- und Wortschatzerwerb entwickeln, durch den ein weiterer Wortschatzaufbau verzögert und die Leseentwicklung weiter unterbunden wird.

Lesefähigkeiten werden jedoch nicht automatisch erworben, indem man viel liest (Hattie 2010: 140). So führt außerschulisches Lesen v. a. bei bildungsnahen Schichten zu einer Zunahme der Lesekompetenz (Pfof et al. 2010). Ähnlich verhält es sich mit dem Erwerb neuer Wörter: Die Lesekompetenz muss überhaupt erst hinreichend gut ausgeprägt sein, damit Kinder neuen Wortschatz implizit beim Lesen erlernen können. In einer eigenen Untersuchung erwarben beispielsweise Kinder im Grundschulalter zunächst die Bedeutung unbekannter Wörter v. a. beim Geschichtenerzählen, aber nicht, wenn sie die Texte selbst lasen (Suggate et al. 2013). Erst wenn Kinder sich am Ende der Grundschulzeit befanden, gelang der implizite Erwerb von Wortschatz ähnlich gut wie beim Geschichtenerzählen.

## **2. Spektrum der Wortschatzdiagnostik**

Wortschatzdiagnostik war bereits in den Frühphasen der Psychometrie Gegenstand der Forschung (z. B. Sims 1929). Die Erfassung des Wortschatzes kann nicht nur nach Aufgabenformen, sondern auch danach gruppiert werden, wie tief das erfasste Wissen reicht, nämlich vom bloßen Wiedererkennen der phonetischen oder visuellen Gestalt bis zum situationsangemessenen Anwenden und der Fähigkeit, das sprachliche Konzept zu erläutern und Querverbindungen zu anderen Wörtern zu ziehen. Im Folgenden werden die verschiedenen Ansätze charakterisiert und im Hinblick auf ihre Vor- und Nachteile diskutiert.

Auf der Ebene des bloßen Wiedererkennens von Wörtern entstanden in den letzten Jahren verschiedene Variationen der sog. *Lexical Decision Tasks*. Wörter werden dabei entweder akustisch oder visuell präsentiert. Die Testperson muss entscheiden, ob es sich dabei um ein reales Wort handelt oder nicht, bzw. ob die Person das Wort kennt (z. B. Lemhöfer & Broersma 2012 und die darauf aufbauende Online-Plattform <http://www.lextale.com/>; Trautwein & Schroeder 2018). Solche Tests sind einfach zu konstruieren und sie weisen eine hohe Ökonomie auf. Es existieren sogar Ansätze zur automatisierten Generierung von falschen Antwortalternativen (Hamed & Zesch 2018). Denn um reales Wissen von einer reinen Ja-Sage-Tendenz abzugrenzen, müssen zu solchen Tests auch Pseudowörter beigemischt werden. Gerade hierbei ergibt sich jedoch eine Reihe an Fallstricken. Aus methodischer Sicht stellt sich nämlich nicht nur die Frage, wie man die korrekten Wörter auswählt und nach welchen Kriterien das Aufgabenmaterial ausbalanciert wird, sondern auch, wie man Pseudowörter generiert, die realen Wörtern hinreichend ähnlich sind. Beispielsweise ist es möglich, einfach Anfangsbuchstaben zu ersetzen. Bei der schriftlichen Darbietung liegt allerdings eine Gefahr darin, Pseudohomophone zu generieren, also Wörter, deren lautliche Realisation einem realen Wort entspricht, das aber lediglich falsch geschrieben ist. Berücksichtigt man solche Probleme nicht, dann wird eher orthographisches Wissen erfasst. Dies kann durchaus auch das Ziel einer Untersuchung sein, aber man sollte sich der Konfundierung bewusst sein. Zum anderen stellen solche Tests auch eine sehr oberflächliche Überprüfung des Wortschatzes dar, da sie das bloße Wiedererkennen der Wortgestalt ohne irgendeine Form tiefergehenden lexikalischen Wissens bzw. Kenntnis der Semantik erfassen. Auch eine reine Stichwortsuchmaschine würde einen solchen Test also mit Bestergebnissen bestehen.

Validere Ergebnisse erzielt man deshalb in der Regel mit Tests, bei denen eine Form des semantischen Verständnisses erfasst wird. Dies geschieht beispielsweise mit Wort-Bild-Zuordnungen (z. B. der Peabody Picture Vocabulary Test 4: PPVT-IV; deutschsprachige Adaptation: A. Lenhard et al. 2015; s. u.) oder indem Synonyme, Antonyme oder Unter- und Oberbegriffe identifiziert werden müssen. Letzteres wird häufig in Intelligenztestbatterien eingesetzt (z. B. BIS-HB, Jäger et al. 2006; KFT 4-12+, Heller & Perleth 2000). Die erfassten Fähigkeiten reichen hier also von einer bloßen Wiedererkennung der Wortbedeutung bis zu einem tieferen Verständnis der Bezüge zwischen Wörtern. Zwar liefern auch diese Tests noch keinen Nachweis dafür, dass eine Testperson die Wörter auch kontextangemessen einsetzen kann. Außerdem ist der Aufwand bei der Testkonstruktion generell höher als bei lexical decision tasks. Der große Vorteil liegt jedoch in der hohen Reliabilität und Objektivität der geschlossenen Aufgabenformate. Zudem kann auch hier angesichts der hohen Ökonomie in der Anwendung ein breites Wortschatzwissen abgeprüft werden.

Und schließlich kann auch der produktive Wortschatzgebrauch überprüft werden. Dies wird beispielsweise mit Lückentexten wie den C-Tests gemacht, die v. a. im Rahmen der Fremdsprachdiagnostik eingesetzt werden (methodisch Diskussion siehe Eckes & Grotjahn 2006; Khodadady 2014). Allerdings werden hierbei über das Wortschatzwissen hinaus auch syntaktische (und bei automatisierter Darbietung orthografische) Fähigkeiten mitefassen.

Besonders tiefgehend wird das Verständnis für die Konzepte hinter den Wörtern in Tests erfasst, bei denen Personen die Wortbedeutung in eigenen Worten erklären müssen. Dies ist beispielsweise bei Untertests der Skala *Sprachverständnis* der Wechsler-Reihe (z. B. WAIS-IV, Wechsler 2008; WISC-V, Wechsler 2014) der Fall. Beim Untertest *Gemeinsamkeiten finden* müssen Personen beispielsweise Vergleiche zwischen verschiedenen, teilweise sehr abstrakten sprachlichen Konstrukten ziehen, beim *Wortschatz-Test* müssen sie die zentralen Aspekte eines sprachlichen Konzepts erklären. Die abgefragten Bedeutungsfacetten sind dabei z. T. sehr komplex und abstrakt. Die erfassten Fähigkeiten gehen deshalb stark in Richtung allgemeine Intelligenz. Eine Schwierigkeit hierbei besteht außerdem in der Bewertung der Aufgabenlösungen, da diese nicht ausschließlich mit richtig vs. falsch bewertet werden können. Bei der Testauswertung besteht deshalb ein größerer subjektiver Bewertungsspielraum als bei einfacheren Wortschatztests.

### 3. Peabody Picture Vocabulary Test IV (PPVT-IV)

Ein im internationalen Bereich sehr weit verbreitetes, aus dem englischen Sprachraum stammendes, standardisiertes Testverfahren liegt mit dem PPVT vor, der aktuell in der vierten Ausgabe verfügbar ist. Das Verfahren geht ursprünglich auf das Ehepaar Lloyd und Leona Dunn zurück und hat seitdem eine Reihe an Überarbeitungen erfahren, an denen auch deren Sohn Douglas Dunn mitgewirkt hat. Das Verfahren wurde in zahlreiche Sprachen übertragen bzw. für diese anhand des ursprünglichen Konstruktionsprinzips neu entwickelt. Die Adaptionen sind so weit wie möglich zur Originalfassung parallelisiert, sodass sprachübergreifend ein Untersuchungsinstrument zur Erfassung des rezeptiven Wortschatzes zur Verfügung steht. Auch die deutschsprachige Adaption nahm ihren Ausgangspunkt in einer interkulturellen Vergleichsstudie zum Spracherwerb (Suggate et al. 2014), bei dem der vorschulische Spracherwerb in Deutschland und Neuseeland und die Bedeutung verschiedener Prädiktoren auf den schulischen Schriftspracherwerb längsschnittlich verglichen wurden. Die folgenden Ausführungen beziehen sich schwerpunktmäßig auf die deutschsprachige Adaption (Lenhard et al. 2015) und die Ergebnisse der zugehörigen Normierungsstudie.

### 3.1 Testlogik und Ablauf

Ziel des PPVT ist die Erfassung des rezeptiven Wortschatzes. Hierzu wird jeweils eine Bildkarte mit vier Bildern gezeigt und dazu ein Wort vorgesprochen. Die Testperson muss auf das Bild deuten, das am besten zum Wort passt. Insgesamt gibt es im Testverfahren 228 Wörter und zugehörige Bildkarten, die der Schwierigkeit nach angeordnet und in 12er-Blöcke gruppiert sind. Es existieren folglich insgesamt 19 dieser Aufgabenblöcke, die jeweils mit einem altersspezifischen Einstiegskriterium versehen sind. So beginnen beispielsweise Kinder im Alter von 5 Jahren bei Aufgabenset 4, Jugendliche ab 14 Jahren dagegen erst mit Aufgabenblock 11. Werden in diesem Block mindestens 11 der 12 Wörter korrekt erkannt, so gelten alle Aufgaben in niedrigeren Blöcken automatisch als bestanden. Ist dies nicht der Fall, so wird zunächst das Bodenset ermittelt, indem Block für Block zurückgegangen wird, bis schließlich ein Aufgabenblock mit mindestens 11 korrekten Lösungen gefunden wurde. Anschließend wird die Testung mit schwierigeren Aufgaben fortgesetzt, bis schließlich in einem Block acht oder mehr Fehler aufgetreten sind.

Wenn eine Testperson bei einer Aufgabe nicht innerhalb von etwa 10 Sekunden antwortet, so wird eine Ermutigung gegeben, z. B. "Versuche es einfach. Zeige auf dasjenige, von dem Du denkst, es könnte richtig sein". Wenn die Testperson immer noch nicht antwortet, wertet man das Item als Fehler, protokolliert die fehlende Antwort und geht zügig zur nächsten Aufgabe über. Die Gesamtpunktzahl ergibt sich schließlich aus der Anzahl richtig gelöster Aufgaben plus der Anzahl an automatisch als korrekt gewerteten Aufgaben unterhalb des Bodensets. Die Testdarbietung erfolgt ohne Zeitbegrenzung. Da Testpersonen unterschiedlich schnell arbeiten, hängt die Dauer der Untersuchung von der untersuchten Person ab. Im Schnitt dauert die Testung allerdings weniger als 15 Minuten, in nur 10% aller Durchführungen werden 20 Minuten oder mehr benötigt.

### 3.2 Auswahl und Adaption des Aufgabenmaterials

Die ursprüngliche Fassung des PPVT von 1953 hatte zunächst kein linguistisch orientiertes Konstruktionsmodell, sondern die Auswahl der Wörter geschah durch Sichtung aller Wörter in Standardlexika, die sich gut für eine Visualisierung eigneten. In die ursprünglichen Datenerhebungen gingen 600 Bildkarten mit 2400 Begriffen ein, aus denen schließlich 300 Aufgaben, aufgeteilt auf zwei Testfassungen, ausgewählt wurden. Fassung 2 und 3 zeichneten sich durch eine Verbesserung der psychometrischen Eigenschaften aus. Es entstand dabei ein Kategoriensystem mit 19 Inhaltskategorien (z. B. Tiere, Berufe, Spielzeug, Aktionen ...), anhand derer eine möglichst gleichmäßige inhaltliche Verteilung der Aufgaben und eine breite Abdeckung des Wortschatzes erzielt wurden. Die dritte Fassung enthielt je Testfassung 204 Aufgaben.

In Fassung 4 wurden schließlich Boden- und Deckeneffekte durch Hinzunahme sehr leichter und sehr schwerer Aufgaben weitgehend reduziert. Außerdem erfolgte eine komplette Überarbeitung des Bildmaterials und eine bessere Ausbalancierung hinsichtlich der Inhaltskategorien. So wurde der Anteil an Aufgaben in den Kategorien *Körperteile*, *Kleidung und Zubehör*, *Obst und Gemüse*, *Musikinstrumente und Spielzeug* und *Erholung* erhöht, der Anteil der Kategorien *Aktionen* und *Adjektive* hingegen reduziert. Beide Maßnahmen dienen dem Zweck, das Testverfahren stärker für die Anwendung bei jüngeren Kindern zu optimieren.

Bei der Adaptation ins Deutsche wurde die Übersetzung jedes einzelnen Wortes von Personen mit jeweils deutscher oder englischer Muttersprache gemeinsam vorgenommen. Bei jenen Zielwörtern, bei denen keine stimmige Übertragung möglich war, wurde eines der Distraktorbilder als Target verwendet, die Wortart geändert (z. B. *coniferous* → Konifere) oder auf einen unterschiedlichen Aspekt des Targetbildes fokussiert. Das Bildmaterial wurde also komplett beibehalten.

Ein besonderer Fokus lag sowohl in der amerikanischen als auch in der deutschen Version in der gerechten Ausgestaltung der Stimuli. So sind Menschen beiderlei Geschlechts und verschiedener Ethnien gleichmäßig auf alle Berufsgruppen und Tätigkeiten aufgeteilt. Die Testfairness wurde mittels Item Response Theory (IRT)- und Differential Item Functioning (DIF)-Analysen untersucht. Mit Letzteren lässt sich überprüfen, ob einzelne Aufgaben für bestimmte Bevölkerungsgruppen unverhältnismäßig leicht oder schwer sind, wodurch die Fairness des Verfahrens negativ beeinflusst werden könnte. Dabei geht es nicht darum, ob eine Bevölkerungsgruppe bei einem Wort im Schnitt besser oder schlechter abschneidet als eine andere Gruppe, sondern ob die Lösungswahrscheinlichkeit für dieses Wort bei einer Bevölkerungsgruppe niedriger oder höher liegt, als auf der Basis des durchschnittlichen Fähigkeitsniveaus dieser Gruppe erwartet werden würde. Ist dies der Fall, so kann man schließen, dass die Lösung außer der Fähigkeit durch weitere Merkmale (z. B. dem Sprachhintergrund) beeinflusst wird (vgl. Swaminathan & Rogers 1990). In der US-Fassung wurde diese Analyse für die Variablen Geschlecht, ethnische Herkunft, Sozialstatus und Wohnort innerhalb der USA durchgeführt, in der deutschsprachigen Adaptation für die Variablen Geschlecht und Sprachhintergrund. Damit sollte unter anderem ein möglichst hoher Grad an Kulturunabhängigkeit erreicht werden.

Da die maßgebliche Motivation für die Erstellung der deutschsprachigen Fassung die Vergleichbarkeit der Ergebnisse zwischen der englischen und der deutschen Sprache war, wurden die einzelnen Wörter mit möglichst gleichfrequenten Wörtern der deutschen Sprache übersetzt. Dabei kam ein relatives Maß der Worthäufigkeit zum Einsatz, bei dem die Frequenz des häufigsten Wortes der jeweiligen Sprache durch die Auftretenshäufigkeit des Zielwortes geteilt und der Quotient logarithmiert wird. Zwar wird auf diese Weise

eine gemeinsame, sprachübergreifende Metrik etabliert, jedoch gibt es dennoch prinzipielle Unterschiede zwischen der deutschen und der englischen Sprache, die beispielsweise in den Wurzeln beider Sprachen begründet sind. Obwohl beide Sprachen zu den germanischen Sprachen zählen und erhebliche Verwandtschaft aufweisen, kommen in der amerikanischen Version des PPVT-4 einige Wörter vor, für die das Deutsch keine passende Übersetzung mit ähnlicher Wortschwierigkeit bereithält. Ein Beispiel ist das englische Wort "cairn", das aus dem Gälischen stammt und pyramidenförmig aufgeschichtete Steinhügel des Neolithikums bezeichnet, die nur auf den britischen Inseln vorkommen. Man könnte den Begriff zwar schlicht mit "Steinhügel" übersetzen, würde damit jedoch eine deutlich andere Schwierigkeit erzielen, da "cairn" fast ausschließlich fachsprachliche Anwendung findet.

Weiterhin verfügt die englische Sprache über ein größeres Vokabular und damit einhergehend über mehr Synonyme unterschiedlicher Schwierigkeit. So führt beispielsweise das Oxford English Dictionary aktuell über 600.000 Schlagwörter auf, während die Duden-Redaktion den deutschen Wortschatz nur auf 300.000 bis 400.000 beziffert (Kunkel-Razum 2000). Die Ursache hierfür liegt mutmaßlich in der Eroberung der britischen Inseln durch die Normannen 1066, wodurch das heutige Englisch sowohl einen großen germanischen als auch romanischen Wortschatz hat. Wörter mit romanischer Wurzel wurden zunächst jedoch eher im höfischen Kontext gebraucht, wohingegen Wörter mit germanischer Wurzel in die Alltagssprache niedrigerer gesellschaftlicher Klassen abgedrängt wurden. Auch heute noch werden Wörter mit germanischer Wurzel tendenziell eher in alltäglichen oder profanen Situationen gebraucht, während ihre romanischen Entsprechungen einen höheren Grad an Elaboriertheit und eine niedrigere Frequenz aufweisen. Im PPVT-4 betrifft dies beispielsweise das Wort "beverage" (= Getränk), welches mit dem französischen Verb boire (= trinken) verwandt ist (vgl. auch franz. "beuverie" = Zecherei). Das germanischstämmige Synonym dazu lautet "drink". Für den deutschsprachigen Begriff "Getränk" stehen dagegen kaum vergleichbare Alternativen zur Verfügung. Dadurch ist die Schwierigkeit jedes einzelnen Wortes im Englischen höher als diejenige des deutschen Begriffs "Getränk", da jedes Synonym jeweils weniger häufig verwendet wird als sein deutsches Pendant. Deutschsprachige Kinder lösen aufgrund der geringeren lexikalischen Diversität der Sprache deshalb bei der deutschen Adaptation des PPVT-4 im Schnitt mehr Aufgaben richtig als gleichaltrige englischsprachige Kinder in der US-Version. Die deutschsprachige Testversion differenziert in den höheren Altersstufen und im hohen Leistungsbereich also etwas weniger gut als die amerikanische. Die Adaption des PPVT-IV fand in Deutschland statt, sodass die Testgüte für die Schweiz nicht systematisch untersucht wurde, jedoch stehen mittlerweile weiterführende Erfahrungen zur Verfügung (siehe Juska-Bacher & Röthlisberger 2021). Mit leichten Anpassungen (z.B. Item 65: Schornstein –

Kamin, Item 70: Umschlag – Kuvert) lässt sich das Testverfahren auch erfolgreich in der Deutschschweiz einsetzen.

### 3.3 Normierung und Reliabilität

Die deutsche Übersetzung des PPVT-4 wurde Anfang 2013 an einer Stichprobe von  $N = 389$  Kindern der Klassenstufe 3 bis 10 aus Grund-, Hauptschulen und Gymnasien überprüft. Die Darbietung der vier Auswahlbilder pro Aufgabe erfolgte ohne Abbruchkriterien mittels eines dafür erstellten Computerprogramms, welches auch die Antwortzeiten protokollierte. Das zugehörige Wort, das richtig zugeordnet werden musste, wurde auditiv über Kopfhörer vorgespielt. Für die Neunormierung des Tests wurden die Items gemäß der im Deutschen erzielten Schwierigkeiten und – bei gleicher Schwierigkeit – der Reaktionszeiten aufsteigend gereiht.

Diese endgültige Testversion wurde schließlich zur Erhebung der deutschen Normierungsdaten verwendet. Hierfür führten 42 Testleiter\_innen das Verfahren gemäß der standardisierten Testdurchführung in 20 Kindergärten und 47 Schulen in ganz Deutschland durch. Insgesamt nahmen 4880 Kinder und Jugendliche im Alter zwischen 2;7 Jahren und 18;0 Jahren an der Untersuchung teil. Die Stichprobe wurde schließlich für jede Klassenstufe sowie für Kinder aus dem Kindergarten hinsichtlich Geschlecht, Schularart und Anteil an Kindern mit Migrationshintergrund stratifiziert. Bei Überrepräsentation von Merkmalen infolge der geclusterten Datenerhebung erfolgte eine zufällige Ziehung von Fällen bis die Zielquote erreicht wurde. In der repräsentativen Normstichprobe verblieben schließlich  $N = 3555$  Kinder und Jugendliche aus den Bundesländern Baden-Württemberg, Bayern, Brandenburg, Hessen, Niedersachsen, Nordrhein-Westfalen und Rheinland-Pfalz. Der Anteil an Kindern und Jugendlichen mit Migrationshintergrund im Sinne mindestens eines im Ausland geborenen Elternteils betrug 29.3 %, was gemäß des Zensus als repräsentativ angesehen werden kann. Die endgültige Testversion liefert Normen für Kinder zwischen 3;0 und 16;11 Jahren. Die einzelnen Altersgruppen umfassen dabei zwischen 102 und 431 Kinder pro Jahrgang, wobei die Altersnormen mit einem kontinuierlichen Normierungsverfahren modelliert wurden (A. Lenhard et al. 2018).

Für die Beurteilung der Testgüte wurden die Rohdaten mittels eines 1PL-IRT-Modells skaliert. Eine Analyse der tatsächlich bearbeiteten Aufgaben (Aufgaben unterhalb des Boden- und oberhalb des Deckensets wurden dabei als Missing behandelt) ergab sehr hohe Reliabilitätskennwerte mit einer EAP- und RLE-Reliabilität von  $r = .965$ . Auch die Odd-Even-Split-Half-Reliabilitäten lagen aufgeteilt nach Altersgruppe oder Klassenstufe durchgehend bei mindestens  $r_{tt} = .92$ , d. h. im sehr guten Bereich (siehe A. Lenhard et al. 2015: 76). Dies galt ebenso für die Retestreliabilität, die nach 6 bis 12 Monaten  $r_{tt} = .91$  betrug.

### *3.4 Altersentwicklung und Heterogenität des rezeptiven Wortschatzes*

Den absoluten Umfang des Wortschatzes einer Person zu bestimmen, ist messtechnisch nur sehr schwierig möglich, insbesondere da es nicht nur auf die Kenntnis eines Wortes, sondern auch auf die Tiefe des Wissens ankommt. Die Wortschatztiefe spiegelt sich beispielsweise nicht nur in der Kenntnis, sondern auch in der Organisation des Wortschatzes, in der Fähigkeit zum rezeptiven und produktiven Einsatz, in der Schnelligkeit beim Zugriff auf die Wortbedeutung usw. wider (Schmitt 2014). Diese Aspekte sind unscharf und überlappend und deswegen schwer voneinander zu trennen. Die Verfügbarkeit eines Wortes im mentalen Lexikon ist nicht dichotom (verfügbar versus nicht verfügbar), sondern graduell, beginnend bei der Wiedererkennung der phonetischen Form oder visuellen Gestalt des geschriebenen Wortes ohne Kenntnis der Bedeutung bis hin zu einem gereiften Verständnis und tiefergehenden Wissen, möglicherweise sogar der etymologischen Wurzeln. In den meisten anwendungsbezogenen Fällen wird das rezeptive Wiedererkennen von Wörtern im Sinne der Zuordnung zu einem Bedeutungsgehalt als ein hinreichend guter Indikator für die Größe und Tiefe des Wortschatzes angesehen. So korrelierte beispielsweise in der Untersuchung von Vermeer (2001) zum Wortschatz bei Erst- und Zweitsprache in einer Gruppe niederländischer Kinder der rezeptive und produktive Wortschatz zu  $r = .80$ , auch wenn in absoluten Zahlen der rezeptive Wortschatz – erhoben mit einem Instrument, dessen Aufbau dem PPVT-4 entsprach – erheblich größer war als der produktiv eingesetzte. Auch beim PPVT-4 spezifizieren die Rohwerte die absolute Größe des Wortschatzes zwar nicht direkt, dienen aber als Indikator für die relative Größe des Wortschatzes im Vergleich zu Personen des gleichen Alters. Um diesen Vergleich zu ermöglichen, verfügt der PPVT-4 analog zu anderen psychometrischen Untersuchungsinstrumenten über Normwerte, welche die Ausprägung der latenten Wortschatzgröße und die relative Position der Person im Vergleich zur Normierungsstichprobe wiedergeben (siehe Kap. 3.3). Der Verlauf dieser Normwerte ist in Form einiger ausgewählter Perzentilbänder in Abbildung 2 dargestellt.



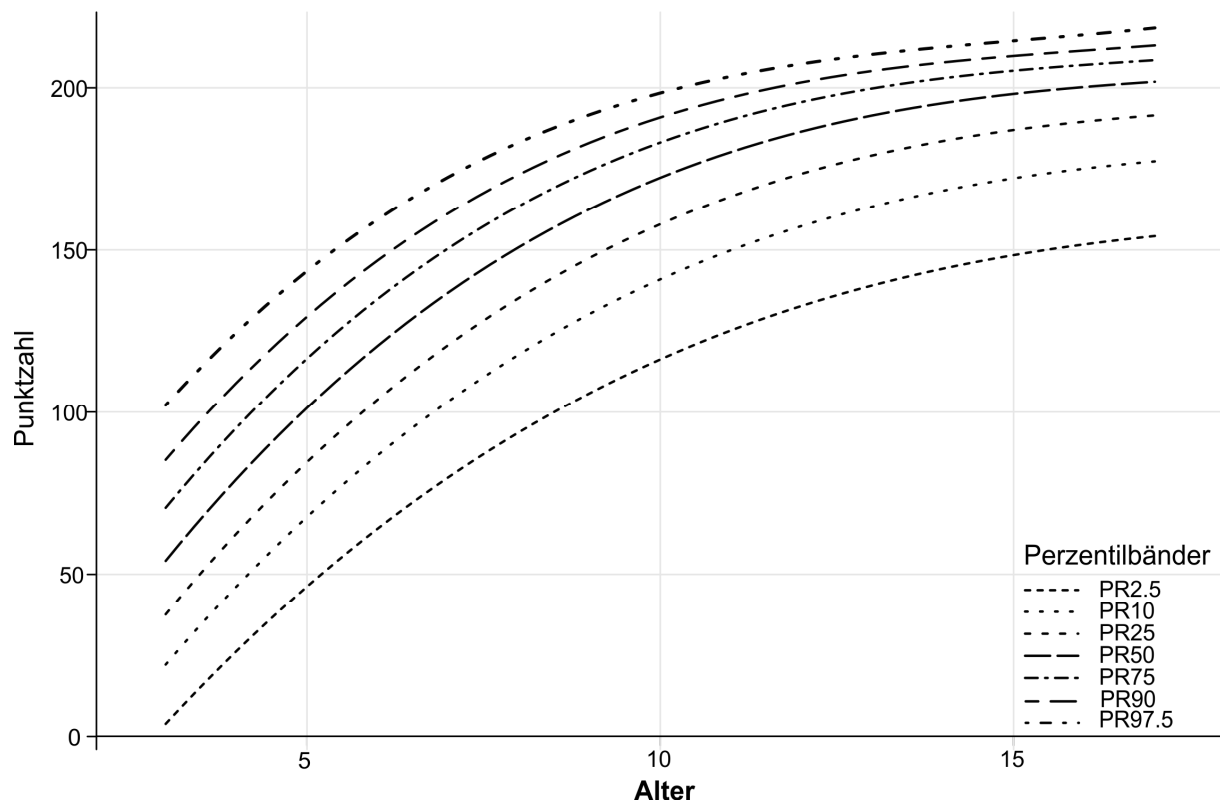


Abbildung 2: Perzentilbänder der Normierungsstichprobe im Altersverlauf von 3 bis 17 Jahren

Der Verlauf der Perzentilbänder zeigt erwartungsgemäß einen starken Anstieg des mittleren Wortschatzes, der besonders im Altersbereich von 3 bis 10 Jahren steil verläuft (siehe auch Tab. 1, Haupteffekte Alter). Die anschließende Abflachung spiegelt nicht ausschließlich eine geringere Entwicklungsgeschwindigkeit ab der frühen Pubertät wider, sondern ist auch der Verteilung der Aufgabenschwierigkeiten im Testverfahren geschuldet. Man beachte allerdings, dass es sich bei der Darstellung nicht um längsschnittliche Verläufe, sondern um die querschnittliche Modellierung der Normierungsdaten handelt. Nichtsdestotrotz kann im erfassten Altersbereich unabhängig von der Fähigkeitsausprägung von einem kontinuierlichen Zuwachs des Wortschatzes ausgegangen werden. Der Alterseffekt ist sehr stark ausgeprägt,  $F(3, 3351) = 2916.2$ ,  $p < .001$  und erklärt mit  $\eta^2_{\text{partial}} = .72$  einen erheblichen Teil der Varianz der Rohwerte.

Prädiktor	Schätzwert	Punktzahl			
		CI	t	SE	p
<b>Haupteffekte</b>					
(Intercept)	173.21	170.82 – 175.60	1.22	141.96	<0.001
Alter (linear)	1707.03	1586.61 – 1827.45	61.42	27.79	<0.001
Alter (quadratisch)	-559.10	-596.89 – -521.31	19.27	-29.01	<0.001
Alter (kubisch)	114.42	77.05 – 151.78	19.06	6.00	<0.001
MI einfach	-12.91	-22.20 – -3.62	4.74	-2.73	0.006
MI beide	-35.68	-44.35 – -27.00	4.42	-8.07	<0.001
Geschlecht	-3.07	-7.56 – 1.41	2.29	-1.34	0.179
<b>Interaktionseffekte</b>					
MI einfach x Alter	0.67	0.02 – 1.33	0.33	2.02	0.043
MI beide x Alter	0.55	-0.04 – 1.14	0.30	1.83	0.067
Geschlecht x Alter	-0.01	-0.42 – 0.40	0.21	-0.06	0.954
MI einfach x Geschlecht	-1.53	-5.49 – 2.43	2,021	-0,76	0.449
MI beide x Geschlecht	5.46	1.90 – 9.02	1,815	3,01	0.003
Anzahl Fälle	3363				
R <sup>2</sup>	0.736				

Tabelle 1: Ergebnisse einer polynomialen Regression der Haupteffekte Alter, Geschlecht und Migrationshintergrund (MI) auf die erzielte Punktzahl<sup>1</sup>

Ins Auge fällt zudem die sehr breite Verteilung der Ergebnisse innerhalb jeder Altersstufe. So entspricht ein Rohwert von 49 im Alter von 5 Jahren einem Prozentrang von 2.5 und ein Rohwert von 145 einem Prozentrang von 97.5. Die letztgenannte Leistung erreichen die leistungsschwächsten 2.5 % erst im Alter von 15 Jahren. Anders ausgedrückt besitzen die sehr leistungsstarken fünfjährigen Kinder einen Entwicklungsvorsprung von etwa 10 Jahren im Vergleich zu den sehr leistungsschwachen!

<sup>1</sup> MI bezeichnet den Migrationshintergrund der Eltern (einfach = ein Elternteil ist migriert; beide = beide Elternteile sind migriert) mit Personen ohne Migrationshintergrund der Eltern als Referenzgruppe. Signifikante Einflussfaktoren sind hervorgehoben. Die Schätzwerte geben den mittleren Einfluss der unabhängigen Variablen auf den Gesamtrawwert an. Positive Schätzwerte erhöhen folglich im Schnitt das erzielte Ergebnis, negative Werte senken es. Bei Kodierung des Geschlechts bildeten die Jungen die Referenzkategorie.

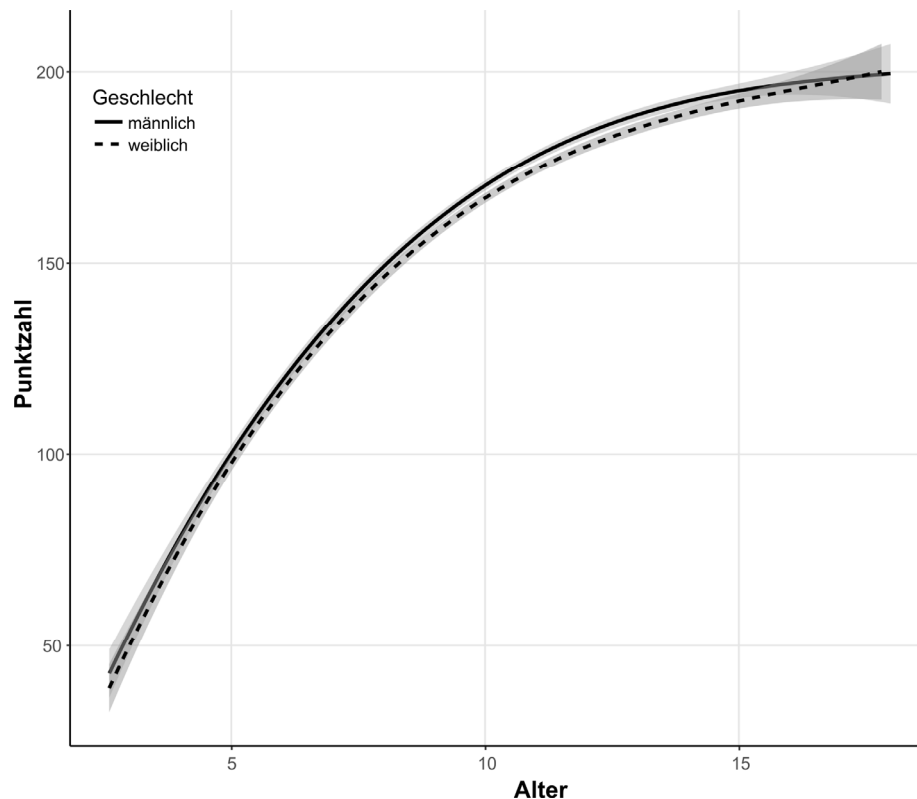


Abbildung 3: Mittlere Leistung im Altersverlauf nach Geschlecht<sup>2</sup>

### 3.5 Effekte von Geschlecht, Migrationshintergrund und Alter

Jungen und Mädchen der Normierungsstichprobe zeigten im Schnitt vergleichbare Leistungen (Abbildung 3, Haupteffekt Geschlecht, Tab. 1), mit einem minimalen Nachteil der Mädchen. Dieser Unterschied war allerdings so gering, dass er keine statistische Bedeutsamkeit erreichte und lediglich 1.9 ‰ der Varianz aufklärt. Auch fanden sich keine relevanten Interaktionen von Alter und Geschlecht, sodass der Unterschied über die Altersspanne im wesentlich konstant blieb. Diese Befunde galten sowohl für Menschen ohne Migrationshintergrund als auch für Personen mit einem migrierten Elternteil. Bei Personen, bei denen beide Eltern eingewandert waren, wechselwirkte das Geschlecht hingegen mit den Sprachfertigkeiten dahingehend, dass Jungen einen niedrigeren Wortschatz aufwiesen (Tab. 1, Interaktionseffekt MI beide x Geschlecht) und dieser Nachteil vor allem im leistungsschwachen Bereich stark ausgeprägt war. Der Unterschied zwischen Mädchen und Jungen betrug in dieser Gruppe im Schnitt 5.45 Rohpunkte. Mädchen wiesen also, anders als auf der Basis von Geschlechtsstereotypen zu erwarten gewesen wäre, nicht generell einen größeren Wortschatz auf – ein Befund, der im Rahmen der Intelligenzforschung bereits als gut gesichert gilt (Toivainen et al. 2017). Allem

<sup>2</sup> Die Linien geben die mittlere Leistung von Jungen (durchgezogene Linie) und Mädchen in Abhängigkeit vom Alter wieder, die grauen Bänder stellen die 95%-Konfidenzintervalle dar.

Anschein nach existieren jedoch sehr wohl einzelne Gruppen unter den Jungen, die besonders benachteiligt sind.

Eine Migrationserfahrung der Eltern hat einen sehr starken Einfluss auf die Wortschatzleistung,  $F(2, 3351) = 288.8$ ,  $p < .001$ ,  $\eta^2_{\text{partial}} = .15$ . Personen, bei denen ein Elternteil eingewandert war und die i. d. R. bilingual aufwuchsen, erzielten eine im Schnitt 12.9 niedrigere Punktzahl. Besonders stark fiel der Unterschied bei Personen aus, bei denen beide Eltern eingewandert waren. Sie lagen im Schnitt 35.7 Punkte hinter den Personen ohne Migrationshintergrund – ein Unterschied, der je nach Altersgruppe zwischen 1.5 und 2.0 Standardabweichungen beträgt und als sehr groß eingestuft werden kann. In beiden Personengruppen reduzierte sich der Unterschied zu Personen ohne Migrationshintergrund in den höheren Altersgruppen, wobei diese Interaktion nur bei den Personen mit einem migrierten Elternteil Signifikanz erreicht. Insbesondere die Altersverläufe in dieser Gruppe zeigen eine sehr spannende Form: Während in jüngeren Kohorten die Leistung stärker den Kindern mit zwei migrierten Elternteilen gleicht, reduziert sich der Leistungsnachteil zum Schuleintritt und bleibt dann lange konstant. In der Pubertät schließlich nähern sich Personen mit einem migrierten Elternteil immer stärker der Kohorte ohne Migrationserfahrung an und überflügeln diese am Ende der Jugend deskriptiv sogar leicht (Abbildung 4). Die Gruppe an Kindern mit zwei migrierten Elternteilen holt dagegen nur bis zum Alter von 10 Jahren leicht auf, weist danach aber einen konstanten Abstand zu Personen ohne Migrationshintergrund auf, sodass sich in der Interaktion von Migrationshintergrund und Alter (siehe Tabelle 1) nur ein Trend zeigt.

Die Verteilung der Rohdaten der Person mit zwei migrierten Elternteilen zeigt innerhalb der einzelnen Altersgruppen eine größere Schiefe als in den anderen Gruppen. Ähnlich wie bei den Geschlechtseffekten fallen leistungsschwache Personen überproportional zurück. Der Abstand der Perzentile der Personen mit zwei migrierten Elternteilen zu den entsprechenden Referenzperzentilen der monolingual deutschsprachigen Personen nimmt also mit sinkender Fähigkeit immer weiter zu.

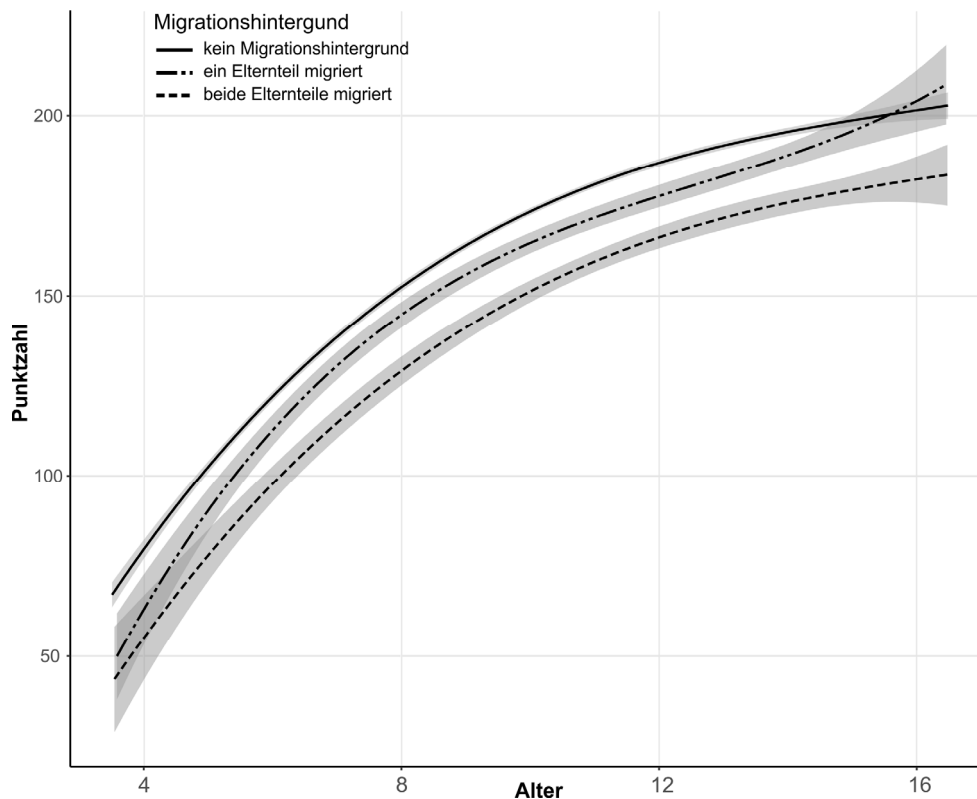


Abbildung 4: Der Einfluss der Familiensprache auf die im Mittel erzielten Rohpunkte<sup>3</sup>

#### 4. Abschluss

Abschließend möchten wir auf die Ziele und Grenzen einer Wortschatzdiagnostik mit dem PPVT-IV eingehen. Eine offensichtliche Grenze liegt darin, dass natürlich keineswegs überprüft wird, welche Wörter im Wortschatz fehlen, sodass das Verfahren nicht für die Feindiagnostik zur Planung von Wortschatzeinheiten im Rahmen von formellem Deutschunterricht geeignet ist. Hierfür müssten in der Tat jene Wörter überprüft werden, deren Beherrschung erreicht werden soll. Stattdessen zielt das Verfahren darauf ab, über die Normierung die relative Position einer Person innerhalb der Altersgruppe zu bestimmen. Es dient damit als Diagnoseinstrument zur Ermittlung des Sprachstandes, als international anerkannter psychometrischer Test für Forschungszwecke, also beispielsweise zum Vergleich des Wortschatzerwerbs in verschiedenen Sprachen, oder auch um die Ursache schulischen Scheiterns genauer zu untersuchen. So lässt sich beispielsweise mit Hilfe des PPVT-4 differenzialdiagnostisch überprüfen, ob schlechte Leistungen im Lesen und Schreiben bei einem Kind, das nicht oder nicht ausschließlich Deutsch als Familiensprache spricht, auf die Schriftsprache

<sup>3</sup> Die Linien geben die mittlere Leistung von Personen ohne Migrationshintergrund (Referenzgruppe, Code 1), Personen mit einem migrierten Elternteil und Personen mit zwei migrierten Elternteilen in Abhängigkeit vom Alter wieder, die grauen Bänder stellen die 95%-Konfidenzintervalle dar.

beschränkt sind, oder ob diese Probleme eventuell durch allgemeine Sprachrückstände bedingt sind.

## LITERATUR

- Beck, I. L., Perfetti, C. A. & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology*, 74(4), 506-521.
- Brugger, L. & Juska-Bacher, B. (2021). Assessing primary grade children's lexical inferencing strategies while reading – A review. *Bulletin suisse de linguistique appliquée*, 113, 213-232.
- Bynner, J. & Parsons, S. (2005). *New light on literacy and numeracy*. London: National Research and Development Centre for adult literacy and numeracy.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.
- Cromley, J. G. & Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 99(2), 311-325.
- Eckes, T. & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23(3), 290-325.
- Hamed, O. & Zesch, T. (2018). The automatic generation of nonwords for lexical recognition tests. In Z. Vetulani, J. Mariani, & M. Kubis (eds.), *Lecture notes in computer science. Human language technology. Challenges for computer science and linguistics* (Vol. 10930, pp. 321-331). Cham: Springer.
- Hattie, J. (2010). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Heller, K. A. & Perleth, C. (2000). *KFT 4-12+ R kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision*. Göttingen: Beltz.
- Jäger, A.O., Holling, H., Preckel, F., Schulze, R., Vock, M., Süß, H.-M. & Beauducel, A. (2006). *BIS-HB. Berliner Intelligenzstruktur-Test für Jugendliche*. Göttingen: Hogrefe.
- Juska-Bacher, B. & Röthlisberger, M. (2021). Das Konstrukt Wortschatz: Dimension(en) Umfang und Tiefe? Empirische Ergebnisse aus der Unterstufe. *Bulletin suisse de linguistique appliquée*, 113, 49-68.
- Khodadady, E. (2014). Construct validity of c-tests: A factorial approach. *Journal of Language Teaching and Research*, 5(6), 1353-1362.
- Kunkel-Razum, K. (2000). *Wie viele Wörter hat die deutsche Sprache?* Goethe-Institut. Verfügbar unter: <https://www.goethe.de/ins/be/de/kul/prj/ssk/21784921.html>
- Lemhöfer, K. & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44(2), 325-343.
- Lenhard, A. (2020). *Das aufgezeichnete Video-Seminar zur WISC-V [Videoseminar]*. Frankfurt a. M.: Pearson.
- Lenhard, A., Lenhard, W., Segerer, R. & Suggate, S. (2015). *Peabody Picture Vocabulary Test - Revision 4 (PPVT-4), deutsche Version*. Frankfurt a. M.: Pearson Assessment.
- Lenhard, A., Lenhard, W., Suggate, S. & Segerer, R. (2018). A continuous solution to the norming problem. *Assessment*, 25(1), 112-125.
- Lenhart, J., Suggate, S. & Lenhard, W. (im Druck). Shared-reading onset and emergent literacy development. *Early Education and Development*.

- Pfost, M., Dörfler, T. & Artelt, C. (2010). Der Zusammenhang zwischen außerschulischem Lesen und Lesekompetenz. *Zeitschrift Für Entwicklungspsychologie Und Pädagogische Psychologie*, 42(3), 167-176.
- Pfost, M., Dörfler, T. & Artelt, C. (2013). Students' extracurricular reading behavior and the development of vocabulary and reading comprehension. *Learning and Individual Differences*, 26, 89-102.
- Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F. & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, 53, 118-137.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913-951.
- Schneider, W. J. & McGrew, K. S. (2018). The Catell-Horn-Carroll theory of cognitive abilities. In D. P. Flanagan & McDonough, E. M. (eds.), *Contemporary intellectual assessment* (pp. 73-163). New York: the Guilford Press.
- Sims, V. M. (1929). The reliability and validity of four types of vocabulary tests. *The Journal of Educational Research*, 20(2), 91-96.
- Suggate, S. P., Reese, E., Lenhard, W. & Schneider, W. (2014). The relative contributions of vocabulary, decoding, and phonemic awareness to word reading in English versus German. *Reading and Writing*, 27, 1395-1412.
- Suggate, S., Lenhard, W., Neudecker, E. & Schneider, W. (2013). Incidental vocabulary acquisition from stories: Second and fourth graders learn more from listening than reading. *First Language*, 33(6), 551-571.
- Suggate, S., Lenhart, J., Vaahtorana, E. & Lenhard, W. (2021). Interactive elaborative story-telling fosters vocabulary in preschoolers compared to repeated-reading and phonemic awareness interventions. *Cognitive Development*, 57, Artikel 100996.
- Swaminathan, H. & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement*, 27(4), 361-370.
- Toivainen, T., Papageorgiou, K. A., Tosto, M. G. & Kovas, Y. (2017). Sex differences in non-verbal and verbal abilities in childhood and adolescence. *Intelligence*, 64, 81-88.
- Trautwein, J. & Schroeder, S. (2019). WOR-TE: Ein Ja / Nein-Wortschatztest für Kinder verschiedener Altersgruppen. *Diagnostica*, 65(1), 37-48.
- Vaahtoranta, E., Lenhart, J., Suggate, S. & Lenhard, W. (2019). Interactive elaborative storytelling: Engaging children as storytellers to foster vocabulary. *Frontiers in Psychology*, 10, 1534.
- Verhoeven, L., van Leeuwe, J. & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, 15(1), 8-25.
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, 22, 217-234.
- Wechsler, D. (2008). *Wechsler intelligence scale for adults (4th ed.)*. Bloomington, MN: NCS Pearson.
- Wechsler, D. (2014). *Wechsler intelligence scale for children (5th ed.)*. Bloomington, MN: NCS Pearson.





# Der WWT 6-10 und seine Testitems

**Christian W. GLÜCK**

Universität Leipzig, Pädagogik  
Marschnerstraße 29d/e, D-04109 Leipzig  
christian.glueck@uni-leipzig.de

The WWT 6-10 (test of vocabulary and wordfinding for children aged 6 to 10) has been in use since 2007. The WWT is well established in speech-language pathology, it contains various subtests, a short form exists, and it comes in two languages (German and Turkish). Apart from clinical assessments, the WWT is widely used in research on language or second language development and language development in special conditions (among others ASD, ADHD, childhood aphasia) as well as in research on education, learning instruction, and migration. This article focuses on the criteria for the selection of words included in the naming task. The analysis of the WWT items considers German norm databases that had been published prior to the development of WWT: norms for nouns (Schröder et al. 2003), adjectives (Groba & De Houwer 2018) and 3.200 different lexemes (Birchenough et al. 2017). Significant correlations of age of acquisition and frequency with item difficulty confirm the suitability of the WWT item set. This in-depth analysis of the WWT items provides valuable suggestions for future improvements of the WWT 6-10.

**Keywords:**

vocabulary, language assessment, language developmental impairment, age of acquisition, frequency.

**Stichwörter:**

Wortschatz, Sprachdiagnostik, Sprachentwicklungsstörung, Erwerbssalter, Häufigkeit.

## 1. Einleitung

Einschließlich der Pilotierungserhebungen liegen nun bald zwei Jahrzehnte Erfahrungen mit dem Wortschatz- und Wortfindungstest für sechs- bis zehnjährige Kinder (WWT 6-10) in zwei Auflagen vor (Glück 2007 und 2011). In diesem Zeitraum hat sich der WWT als diagnostisches Instrument sowohl in der Forschung als auch im Kontext der sprachentwicklungsdiagnostischen Praxis in Logopädie, Sprachtherapie und Sonderpädagogik von Sprache und Kommunikation etabliert.

Im Folgenden wird zuerst der WWT im Überblick vorgestellt. Es werden die Einsatzbereiche beschrieben, wie sie sich in der Fachliteratur darstellen. Der Hauptteil dieses Beitrags widmet sich jedoch der Analyse des Itemsets. Diese ergänzt die im Handbuch zum Test gegebenen Informationen erheblich, da mit den zwischenzeitlich entstandenen lexikographischen Tools heute deutlich umfangreichere und neue Informationen für die linguistische Analyse des Wortmaterials genutzt werden können. Ein kurzer Ausblick auf mögliche Veränderungen für eine neue Auflage schließt den Beitrag ab.

## 2. Der WWT 6-10

Der WWT liefert diagnostische Aussagen zu semantischen und lexikalischen Aspekten des mentalen Lexikons von Kindern im Alter von 5;6 bis 10;11 Jahren. Er ermöglicht sowohl statusdiagnostische als auch therapeuleitende Aussagen zum expressiv und rezeptiv aktivierbaren Lexem-Bestand und über Detailanalysen korrekter und nicht-korrekt beantworteter Antworten auch zur semantischen und lexikalischen Organisation des Lexikons. Die Einschätzung der Wortfindung kann über mehrere Indikatoren vorgenommen werden: über die mit der expressiven Leistung kontrastierende Leistung in der rezeptiven Bildauswahlaufgabe, die Erfassung der Antwortzeit bei der Bildbenennungsaufgabe und die Auswertung der Antworten bei zweifacher Vorlage der Testitems. Um diese differenzierten Aussagen erhalten zu können, wurde der WWT als komplexes Untersuchungsinstrument entwickelt. Die Durchführung als Einzelverfahren mit dem Windows-PC-Programm hat gegenüber der Papier-Durchführung erhebliche Vorzüge. Allerdings entstanden in den letzten Jahren – insbesondere bei Versionswechseln im Microsoft Windows Betriebssystem – immer wieder teils größere, technische Probleme, sodass viele Praktiker\_innen auf die Vorzüge der computerbasierten Version verzichten und mit der Papier-Version arbeiten.

Die Komplexität besteht im Aufbau, in der Durchführung und Auswertung bzw. Protokollierung. Ein großer Vorzug der PC-Version ist es, dass die hohe Komplexität und der Testaufwand mit Blick auf das diagnostische Ziel, z.B. eine Indikationsdiagnostik, auch leicht reduziert werden können.

In der Standarddurchführung werden nur zwei Subtests angeboten: *WWTexpressiv* – ein Subtest zur expressiven Leistung – und *WWTrezeptiv* – ein ergänzender rezeptiver Subtest. Für die Erfassung der expressiven Leistung werden 71 Bildbenennungsaufgaben und 24 Benennungsaufgaben mit verbalem Stimulus vorgelegt. Erfasst werden mindestens die Antwortkorrektheit, optional auch die Antwortzeit (gemessen durch Stopp-Uhr-Funktionalität in der PC-Version) und die Analyse der Falschantworten. Auf der Basis der Unterstellung, dass expressiv abruhbare Items auch rezeptiv wiedererkennbar seien, werden im rezeptiven Subtest nur noch die Eins-aus-Vier-Bildauswahlaufgaben zu den Items vorgelegt, die im expressiven Subtest nicht oder nicht-korrekt beantwortet worden sind, um hier die rezeptive Antwortgenauigkeit sowie die gewählten Distraktoren zu erfassen. Dahingehend handelt es sich um eine adaptive Testung. Eine minimale Auswertung erfasst die Anzahl der korrekten Antworten im expressiven Subtest und im rezeptiven Subtest (= Summe aus expressiv korrekten Antworten und den korrekten Antworten der rezeptiven Bildauswahlaufgabe).

Weitergehende, meist vom PC automatisch erstellte Analysen berücksichtigen die Antwortzeit, die Relation der Falschantwort zum Zielwort (vgl. Kategorien zur Spontansprachanalyse SemLexKrit, Glück 2008) sowie die Art des

(phonologisch-, semantisch-, nicht-relatierten) Distraktors in der Bildauswahlaufgabe. Die Auswertung der Antwortkorrektheit wird auch für vier Itemsubsets vorgenommen: für Nomen, Verben, Adjektive/Adverbien und kategoriale Nomen (Erläuterung s.u.).

Neben der Standard-Durchführung bietet der WWT optionale, expressive Subtests an, die je nach diagnostischem Ziel (therapieleitende Diagnostik) hinzugewählt werden können und sich in den adaptiven Testablauf eingliedern: *WWTexpressiv-Wiederholung* und *WWTexpressiv-Abrufhilfen*. Während in dem einen Subtest die Items in aufsteigender Schwierigkeit geordnet erneut angeboten werden, um die Stabilität und Konsequenz des Abrufs zu untersuchen, wird in dem anderen Subtest die Begrenzung der Antwortzeit aufgehoben und es können kontrollierte Abrufhinweise verschiedener Qualität (unspezifische Ermunterung, semantische Hinweise, Anlautvorgabe) gegeben werden.

Die genannten Subtests können als Langform mit 95 Testitems und als Kurzform mit 40 Items durchgeführt werden. Die PC-Version bietet auch noch Ergänzungsformen an, mit denen nachträglich eine vorherige Kurzform-Durchführung zu einer Langform aufgefüllt werden kann. Die genannten Subtests und Testformen liegen auch als türkischsprachige Version einschließlich türkischsprachiger Instruktionen vor – allerdings nur in der PC-Version.

Die Normdaten zum WWT wurden in verschiedenen Regionen des Bundeslandes Bayern erhoben und sind nach Geschlecht, Bildungsstand der Bezugspersonen und Wohnortgröße kontrolliert. Sie beziehen sich auf eine Normstichprobe von monolingualen Kindern mit der Erstsprache Deutsch ohne, aber auch (in einem gewichteten Anteil) mit Sprachentwicklungsstörungen.

Mit seinem Erscheinen schloss der WWT 6-10 eine wichtige, diagnostische Lücke im altersmäßigen Anschluss an den Aktiven Wortschatztest (AWST, später AWST-R, Kiese-Himmel 2005) und ermöglichte die Diagnostik expressiver Wortschatzleistung auch im Grundschulalter. Auch wenn der später hinzugekommene Sprachstandserhebungstest (SET 5-10, Petermann 2010) mit seinen Subtests Bildbenennung, Kategorienbildung und Sternensuche teilweise auf ähnliche Konstrukte abzielt wie der WWT, gibt es dennoch einige Alleinstellungsmerkmale des WWT, die in der diagnostischen Praxis geschätzt werden, u.a.:

- Der WWT liefert zum expressiven Subtest eine rezeptive Ergänzung, die im Sinne eines Screenings auch eine Einschätzung der rezeptiven lexikalischen Fähigkeiten ermöglicht.
- Der WWT bietet in seiner PC-Version die Möglichkeit einer Erfassung der Antwortzeit der Testperson, sodass auch quantitative Aussagen zur Effizienz des Zugriffs auf das mentale Lexikon möglich werden.

- Der WWT ermöglicht durch Variationen des expressiven Subtests die Gewinnung von therapeuleitenden Informationen über Modifikatoren des Wortabrufs (Zeit, wiederholte Darbietung, Wirksamkeit von Abrufhinweisen).
- In seiner Langform mit 95 Testitems erreicht der WWT sehr hohe Reliabilitätswerte, die von kürzeren (SET-Bildbenennung: 40 Items) oder gekürzten Verfahren (Kurzform des WWT: 40 Items) nicht erreicht werden können. Damit verkleinert sich das Vertrauensintervall und die Interpretation der Ergebnisse wird erleichtert.
- Durch die Vorgabe der in einer Datenbank im PC hinterlegten korrekten und typischen nicht-korrekten Antworten erhöht sich die Auswertungsobjektivität.

Neben der Auswertung der korrekten Antworten im WWT wird eine qualitative Analyse der Falschantworten vorgeschlagen, zu der Montanari & Nagi (2021) ein alternatives Auswertungsmodell vorschlagen.

Für die Erfassung der semantisch-lexikalischen Fähigkeiten werden also unterschiedliche Aufgabentypen unter unterschiedlichen Bedingungen eingesetzt, was eine aufgabenangepasste, flexible Verwendung des WWT 6-10 ermöglicht.

### 3. Einsatzbereiche des WWT

Der Wortschatz- und Wortfindungstest WWT 6-10 beansprucht, wichtige Teilaspekte semantischer und lexikalischer Fähigkeiten abbilden zu können (Juska-Bacher & Jakob 2014; Mayer 2012) und war mit dem Ziel konstruiert worden, zur Identifikation von Kindern mit auffälliger oder gestörter semantisch-lexikalischer Entwicklung etwa im Rahmen einer Sprachentwicklungsstörung bei Kindern im Grundschulalter eingesetzt zu werden. Tatsächlich wird er hierzu auch empfohlen (Spreer & Achhammer 2018; Ptok et al. 2014; Siegmüller 2008; im schulischen Kontext: Schönauer-Schneider & Reber 2014). Für die diagnostische Aufklärung von Sprachentwicklungsauffälligkeiten bei Kindern mit Mehrsprachenerwerb ist es essenziell, neben den Kontextvariablen auch die sprachlichen Fähigkeiten im Erwerb des deutschen Wortschatzes und am besten auch noch im Erwerb der anderen Sprache(n) einschätzen zu können. Für das Türkische bietet der WWT diese Informationen an, weshalb er für die Sprachdiagnostik bei entsprechender Mehrsprachigkeit empfohlen wird (Groba 2018; Bastians et al. 2017). Auch zur Erfassung lexikalischer Fähigkeiten bei Kindern mit speziellen Bedingungen im Spracherwerb wird der WWT eingesetzt (Bartha-Doering et al. 2018 zu kindlichen Aphasien, Schöfl et al. 2015 zu ADHS, Snippe 2015 zu Autismus).

Jenseits dieser Aufgabe in Logopädie/Sprachtherapie und Sonderpädagogik Sprache und Kommunikation ist der WWT auch als Forschungsinstrument, etwa

in der Migrationsforschung, empfohlen (Rauch et al. 2016) und eingesetzt (Schastak et al. 2017) worden. Er dient der Erfassung des Wortschatzes als Kontrollvariable in Studien, z.B. zur Entwicklung von Emotionsbegriffen (Kauschke et al. 2017), zum Einfluss der exekutiven Funktionen beim Rechtschreiben (Czapka et al. 2019), zu didaktischen Aspekten und Sprachverstehen im Mathematikunterricht (Berg & Werner 2014) oder zur Feststellung der Prävalenz von Sprachentwicklungsstörungen (Tiede & Braun 2017). Auch bei der Erfassung von Veränderungen in der Effektvariable in Interventionsstudien bei (mehrsprachigen) Kindern mit semantisch-lexikalischem Fokus (Paul et al. 2018; Motsch & Ulrich 2012) oder mit Schwerpunkt auf dem Leseerwerb (Seifert et al. 2016) wird der WWT eingesetzt. In anderen Studien ging das Material des WWT in die Instrumentenentwicklung ein (z.B. bei einer Studie zum Einfluss des Wortschatzes auf das Lesen Juska-Bacher et al. 2016) oder zur Partizipation im Unterricht (Decristan et al. 2020) oder die Durchführung wurde adaptiert, z.B. als Gruppenversion (Paetsch et al. 2015).

Gerade auch mit Blick auf methodenkritische Anmerkungen (Paul et al. 2018) und die Idee, Itemmaterial aus dem WWT für die Instrumentenentwicklung zu überentnehmen, lohnt sich eine genauere Analyse der Items des WWT.

#### **4. Das Itemset des WWT**

Im WWT werden unterschiedliche Aufgabentypen und Antwortbedingungen eingesetzt. Während für die Beurteilung des expressiven Wortschatzes Bildbenennungsaufgaben, Aufgaben zum Benennen nach verbaler Gegenteilvorgabe, vorgelegt werden, erfolgt die Prüfung des rezeptiven Wortschatzes über Bildauswahlaufgaben. Die Ermittlung der Antwortgenauigkeit erfolgt in expressiven Subtests unter den Bedingungen mit bzw. ohne Beschränkung der Antwortzeit und mit bzw. ohne Abrufhilfen. Gleichzeitig wird die Antwortzeit durch ein von der/dem Testleiter\_in vorgenommenen Tastendruck registriert und ausgewertet. Hinweise zur Strukturierung des Wortschatzes geben Teilauswertungen für verschiedene Items subsets (Wortarten) und die Analyse der Falschantworten.

Damit bilden die Aufgaben isoliert Prozesse aus psycholinguistischen inkrementellen (Levitt 2001) oder konnektionistischen (Dell & O'Seaghdha 1992) Modellen des Wortabrufs ab: die Aktivierung semantisch-konzeptuellen Wissens durch die visuelle Vorgabe und die damit verbundene Aktivierung lexikalischen Wissens. Damit werden aber auch mit dem WWT Merkmale typischer Sprechsituationen nicht oder nicht systematisch erfasst, wie etwa die Auswahl und Anpassung lexikalischer Einheiten entsprechend syntaktisch-morphologischer oder pragmatischer Erfordernisse eines Kontextes (z.B. beim Geschichtenerzählen, Nicoladis & Jiang 2018). Auch gelingt es mit dem WWT nicht, semantische Strukturierungen von Wortarten (Pafel & Reich 2016)

systematisch zu erfassen, semantische Feindifferenzierungen in Wortfeldern oder auch morphologisch-lexikalischen Phänomenen wie Komposition und Derivation nachzuspüren.

Benenn- und Worterkennungsaufgaben, mit denen Genauigkeit und Geschwindigkeit der Bearbeitung in schriftlichen oder lautsprachlichen Aufgaben erfasst werden, gehören zum Standardrepertoire psycholinguistischer Methoden (Dietrich & Gerwien 2017; Höhle 2012). Als wesentliche Einflussfaktoren auf den Abruf von lexikalischen Einheiten aus dem mentalen Lexikon gelten Häufigkeit/Frequenz, das Erwerbssalter, die Familiarität/Vertrautheit, Vorstellbarkeit/Imageability, Konkretheit, Wortlänge, Benennungsübereinstimmung, Wortartzugehörigkeit, das *Priming* durch vorangegangene Items und die Sprachlernbiographie. Aktuelle Forschungen arbeiten daran, die Verwobenheit dieser Faktoren (bspw. Frequenz als eigenständiger oder durch Kumulation als abhängig vom Erwerbssalter anzusehender Faktor) aufzulösen bzw. in neuen Konstrukten (wie der *semantic richness*, Pexman 2012) zu fassen (Perret & Bonin 2019). Wesentliche Voraussetzung für die Einbeziehung dieser Faktoren in die Itemauswahl eines Testverfahrens ist das Vorliegen entsprechender Normen-Lexika. Da diese zum Zeitpunkt der Entstehung des WWT kaum vorlagen, werden in diesem Beitrag entsprechende Post-hoc-Analysen vorgestellt.

#### 4.1 Umfang des Itemsets

Die Auswahl von Wörtern in einem Wortschatztest ist in Bezug auf Umfang und Auswahlkriterien kritisch hinsichtlich ihrer Repräsentativität, die wegen der Gebundenheit semantisch-lexikalischer Fähigkeiten an Entwicklung und Erfahrung und damit an Alter und Kultur keine absolute Größe darstellen kann. Auch die Festlegung von Günther (1988), der einen Testumfang von 100 Items empfiehlt, kann mit Hinblick auf die mehrere Tausend Einträge im mentalen Lexikon eines Kindes und unter Berücksichtigung o.g. semantisch-lexikalischer Kriterien nur eine Stichprobenziehung darstellen. Größere Testumfänge sind in einer für Kinder akzeptablen Bearbeitungszeit nur im rezeptiven Modus von Auswahlaufgaben (z.B. PPVT 4 mit bis zu 228 Items, Dunn & Dunn 2007) zu erreichen.

Der WWT 6-10 nähert sich mit 95 Testitems dem empfohlenen Umfang an und kann in allen Altersgruppen für die Langformen hervorragende Reliabilitätswerte in Cronbach's Alpha  $\geq .9$  erreichen. Die Kurzform mit 40 Items ist in der Praxis – der geringeren Bearbeitungszeit wegen – beliebt. Die geringere Itemzahl spiegelt sich jedoch auch in geringeren, immer noch als gut zu interpretierenden Reliabilitätswerten zwischen .87 und .88 wider.

Die nach der Pilotierung ausgewählten Items sind zur Testkonstruktion nach den Kriterien Verwendungshäufigkeit, Wortlänge, Wortartzugehörigkeit und semantische Bereiche (Sachgruppen) kontrolliert worden. Für weitere der oben

genannten Kriterien wie Erwerbssalter, Vorstellbarkeit etc. standen zum Zeitpunkt der Entwicklung des WWT (1. Auflage 2004) nicht für alle Wortarten entsprechende Daten zur Verfügung (Snodgrass & Vanderwart 1980, Stadthagen-Gonzalez & Davis 2006).

#### 4.2 Benennübereinstimmung

Während für psycholinguistische Experimente die Berücksichtigung der Benennübereinstimmung (*naming agreement*) eine wichtige Voraussetzung zur Kontrolle der Itemsets bildet, stellt sie in einem expressiven Wortschatztest das zu erfassende Merkmal dar und wird teststatistisch als Itemschwierigkeit operationalisiert, also als Anteil der Probanden mit korrekter Benennung eines Items an allen Probanden. Da Items mit einem Schwierigkeitsindex  $> 95\%$  keine Varianz bei den Testpersonen erzeugen können, sind diese Items in der Revision entfernt worden bzw. werden als Beispielitems genutzt. Die Benennübereinstimmung, gemessen an der Itemschwierigkeit, steigt im Mittel erwartungsgemäß über die Altersgruppen an: von 0.10 (SD: 0.3) bei den 5;6 bis 6;0-Jährigen bis 0.78 (SD: 0.2) bei den 10;0 bis 10;11-Jährigen. In einer Studie mit 44 Studierenden der Sprachheilpädagogik konnte eine weit geringere mittlere Itemschwierigkeit in der Benennaufgabe von 98% festgestellt werden (hohe Werte zeigen eine geringe Schwierigkeit an). Dabei ist zu bedenken, dass die Itemschwierigkeit nur eine grobe Schätzung der Benennübereinstimmung darstellt, da die Elizitationsbilder korrekterweise mit verschiedenen Lexemen ausgedrückt werden können. Bei 29% der Items wird von den Studierenden nur eine korrekte Antwort gegeben, 51% der Items provozierten 2 oder 3 verschiedene korrekte Antworten und nur 5% der Items wurden durch mehr als 5 verschiedene korrekte Antworten benannt.

Grundlage der Benennübereinstimmung ist die übereinstimmende Interpretation des Bildes (*image agreement*). Diese wird wiederum in einem rezeptiven Wortschatztest mit Bildauswahl indirekt geprüft. Hier ist die Übereinstimmung direkt an der Itemschwierigkeit ablesbar, da es in der Auswahlaufgabe nur eine korrekte Lösung gibt. Auch hier steigt die Itemschwierigkeit mit dem Alter an, allerdings von einem viel höheren Wertniveau aus (MW der Itemschwierigkeit bei 5;6 bis 6;0-Jährigen: 0.83 (SD: 0.2)). In der Normierungsstudie erreicht die Altersgruppe der 9;6 bis 9;11-Jährigen bereits eine Itemschwierigkeit von 0.97 (SD  $< 0.1$ ), was anzeigt, dass die Erkennung des Bildes einen sehr hohen Übereinstimmungsgrad aufweist.

#### 4.3 Häufigkeit

Für die Berücksichtigung der Häufigkeit, mit der einem Kind ein bestimmtes Wort begegnet, standen nur sehr eingeschränkt korpusanalytisch gewonnene Frequenzlexika zum Kinderwortschatz zur Verfügung (Pregel & Rickheit 1987; Augst 1984; zusammenfassend Naumann 1999). Da sich in der Pilotierung zeigte, dass Wörter, die nicht aus der niedrigsten dort angegebenen

Häufigkeitsklasse stammen, Deckeneffekte hervorrufen, konnte auf diese Quellen zur Ermittlung der Frequenz nicht zurückgegriffen werden – da innerhalb dieser niedrigsten Häufigkeitsklasse keine Differenzierung vorgenommen werden konnte. Ersatzweise wurde zur Entstehung des WWT auf das damals verfügbare Korpus "Deutscher Wortschatz" zurückgegriffen – heute erreichbar unter: [wortschatz.uni-leipzig.de](http://wortschatz.uni-leipzig.de). Dieses Korpus umfasst schriftsprachliche, an Erwachsene gerichtete Texte. Die 95 Testitems erreichen hier eine annähernde Normalverteilung über die Häufigkeitsklassen 7 bis 18. Eine nachträgliche Analyse mit dem auf Kinder ausgerichteten Schriftsprach-Korpus childLex (Schroeder et al. 2015) zeigt, dass 6 Items (7%) keinen Eintrag im Korpus haben (z.B. "einfarbig"), 15% der Items eine normalisierte Frequenz annotierter Types <10 je 1 Mio. Tokens im Korpus aufweisen und damit als infrequent gelten (z.B. "demonstrieren"), während 42% der Items eine Typefrequenz von 10 bis 100/Mio. aufweisen und damit frequent sind, während 35% der Items hochfrequent mit einer Frequenz von bis 1000/Mio. sind und 3% sogar Häufigkeitswerte darüber aufweisen (z.B. "einfach") (Abruf 20.12.2018). Bezogen auf diese Datengrundlage zeigt sich, dass 80% der Items in an Kinder gerichteten, schriftsprachlichen Texten häufig oder sehr häufig auftreten, so dass dem Itemset des WWT eine hohe, ökologische Validität zukommt.

Der Frequenzeffekt drückt sich auch in der Itemschwierigkeit aus, also im Anteil der korrekt beantworteten Items an allen Items. Der Schwierigkeitsindex weist dabei hohe Werte bei geringer Schwierigkeit aus und umgekehrt. Dieser Zusammenhang ist auch bei den in der Normierungsstudie (N=880) gegebenen Antworten mit kleinem Effekt nachweisbar, da die Typefrequenz (nach childLex) mit der Itemschwierigkeit in einem positiven, korrelativen Zusammenhang nach Pearson steht ( $r = .21$  bei  $p = .045$ ).

#### 4.4 Erwerbssalter, Imageability, Familiarität und Konkretheit

Als einer der stärksten Einflussfaktoren auf die Geschwindigkeit des Abrufs (für eine Übersicht Perret & Bonin 2019) gilt das Erwerbssalter (*age of acquisition*), also der Zeitpunkt, zu dem ein Wort typischerweise gelernt wird, wobei sich die Operationalisierungen des Faktors und die Erhebungsmethoden deutlich unterscheiden (vgl. Groba & De Houwer 2018). Auch weitere Faktoren scheinen jeweils eigene Beiträge zur Varianz des Abrufprozesses aus dem mentalen Lexikon zu leisten: die Imageability als Maß dafür, wie leicht und vielfältig innere Vorstellungsbilder durch Worte hervorgerufen werden können, Familiarität, also die Vertrautheit mit Objekten, Ereignissen und Zuständen im Lebensumfeld und die Konkretheit, die als Einschätzung auf dem Kontinuum von abstrakt zu konkret vorgenommen wird. Diese Faktoren sind jedoch auch stark konfundiert (Perret & Bonin 2019; Bird et al. 2001; Brysbaert et al. 2014; Alario et al. 2004).

Seit der Entwicklung des WWT hat sich die Datenbasis durch Lexika mit Angaben zu den Einflussfaktoren auf semantische und lexikalische Abrufprozesse erheblich erweitert (vgl. Kuperman et al. 2012; Scott et al. 2019



für sehr umfangreiche Lexikonprojekte). Für psycholinguistische schrift- und lautsprachliche Experimente ist dadurch eine stärkere Kontrolle der Items möglich. Auch für die Testentwicklung können diese Kriterien zu einer theoretisch begründeten Komposition des Itemsets beitragen. Wenn ein Itemset nach den Kriterien der klassischen Testtheorie die erforderliche Varianz und Güte erreicht, aber nicht nach diesen Kriterien konstruiert wurde, dann sollte sich dennoch die Varianz auch bei den Einflussfaktoren zeigen, da diese in dem Itemset dann implizit wirken.

Daher gilt es, soweit Daten für das Deutsche vorliegen, für das Itemset des WWT 6-10 zu prüfen, inwieweit die empirische Itemschwierigkeit in der Normstichprobe mit Erwerbssalter und Imageability korreliert. Als deutschsprachige Datenbasis dieser Itemmerkmale kommen in Betracht die Arbeiten von Schröder et al. (2003), Schröder et al. (2012) beziehungsweise auf Snodgrass und Vanderwart (1980) mit 222 Einträgen. Da diese Lexika nur konkrete Nomen erfassen, könnte ein Vergleich zum Itemset des WWT, der 26 Nomen enthält, vorgenommen werden. Tatsächlich überlappen sich die Itemsets jedoch nur in drei Exemplaren ("Krone", "Mantel", "Hocker"). Damit ist hieraus kein Vergleich für Erwerbssalter und Imageability ableitbar.

Für 258 deutsche Adjektive liegen Normen zum Erwerbssalter von Groba und De Houwer (2018) vor. Von den 24 im WWT vertretenen Items dieser Wortklasse sind immerhin 12 Items in dieser Zusammenstellung enthalten. Für diese nun insgesamt 15 Items ergibt sich ein Korrelationskoeffizient nach Pearson von  $r = -.58$ ,  $p = .025$ . Normen für Verben liegen zusätzlich im 3200 deutsche Wörter umfassenden Lexikon von Birchenough et al. (2017) vor, das mit der erwähnten Erhebung Schröder et al. (2012) immerhin im Rating einen hochsignifikanten Übereinstimmungskoeffizienten von .91 aufweist.

Mit dem Normen-Lexikon von Birchenough et al. (2017) stimmen über ein Drittel der Zielwörter im WWT überein (16 Nomen, 5 Verben, 10 Adjektive, 8 Kategorienomen). Für diese Items kann das Erwerbssalter angegeben werden (MW 5.2 Jahre, SD 1.7). Die Übereinstimmung des Erwerbssalters mit der Itemschwierigkeit wird mittels Korrelationskoeffizient nach Pearson ermittelt und beträgt  $r = .49$  bei  $p = .002$ . Diese Übereinstimmung weist damit nahezu einen großen Effekt auf und zeigt, dass die Varianz der Itemschwierigkeit auch die Varianz des geschätzten Erwerbssalters sehr gut abbildet, was insbesondere für die Itemkategorien Adjektive und Verben gilt ( $r > .68$  bei  $p < .05$ ).

Aufgrund der oben beschriebenen, teilweise geringen Schnittmenge der WWT-Items mit Normen-Lexika lassen sich die weiteren Kriterien (Imageability, Konkretheit, Familiarität) für das Itemset nicht darstellen. Die visuelle Darstellbarkeit der Items spielte insofern in der Konstruktion eine Rolle, als dass sich in der Pilotierung zeigte, dass bildlich gut darstellbare Adjektive (z.B. "sauber" vs. "schmutzig") mit einem zu niedrigen Schwierigkeitsniveau in der Benennung durch Grundschul Kinder einhergehen. Das ist der Grund, warum für

diese Wortart der Aufgabentyp in das deutlich abstraktere Format der verbalen Gegenteilvorgabe geändert wurde.

#### 4.5 Wortlänge

Wortlänge wird in Untersuchungen unterschiedlich operationalisiert (Schrift, Lautsprache, Phonem-, Silbenebene). Welchen Einfluss sie auf die Antwortzeit bei Benennungsaufgaben hat, scheint u.a. durch die Art der Aufgabenstellung moderiert zu sein und drückt sich uneinheitlich aus (Alario et al. 2004). Für das Itemset des WWT ist der Zusammenhang zwischen Silbenzahl und Itemschwierigkeit dokumentiert ( $r = -.27$  bei  $p = .008$  – Glück 2011: 31). Auch mit der Lexemfrequenz der childLex-Datenbank (Schroeder et al. 2015) ergibt sich eine sogar stärker ausgeprägte, negative Korrelation mit moderatem Effekt ( $r = -.381$ ,  $p < .001$ ). Daraus kann geschlussfolgert werden, dass die im WWT geprüften Wörter höherer Silbenzahl zugleich seltenere Wörter in der Umgebung der Kinder sind und diese Wörter damit tendenziell eine höhere Itemschwierigkeit aufweisen.

#### 4.6 Itemkategorien

Für die Auswahl der Items musste die Entscheidung getroffen werden, welche Wortartklassen aufgenommen werden sollen. Da im Verfahren hauptsächlich die Benennungsaufgabe eingesetzt werden sollte, blieben Wortarten mit schwacher oder stark kontextabhängiger Semantik unberücksichtigt. Entsprechend ihrer semantischen und syntaktischen Eigenschaften wurden vier Itemkategorien gebildet. Diese umfassen: (N) 27 Nomen, (V) 24 Verben, (A) 24 Adjektive und Adverbien sowie (K) 24 Nomen, die auf übergeordnete Kategorien ("kategoriale Nomen") referieren. Damit wird sowohl von der Komposition des Deutschen (74,9% Substantive, 13,5% Adjektive, 9,8% Verben, 1,2% Adverbien, 0,6% sonstige – Duden-Redaktion 2021) als auch in Bezug auf die Wortarten von Types oder Tokens verwendeter Sprache bei Kindern (Kauschke 2000; Schulze 2017) abgewichen. Damit im umfangsbegrenzten Itemset genügend Repräsentanten vorhanden sind, wurden Verben und Adjektive/Adverbien auf Kosten der Nomen verstärkt aufgenommen. Dadurch konnten für die vier Itemkategorien altersspezifische Orientierungsnormen aus der Normstichprobe abgeleitet werden, die einem kategoriespezifischen Alterstrend unterliegen.

<b>Altersbereich / Itemkategorie</b>	<b>5;6-6;11</b>	<b>7;0-7;11</b>	<b>8;0-8;11</b>	<b>9;0-9;11</b>	<b>10;0-10;11</b>
Nomen (MW / SD)	11,4 / 4,2	14,1 / 4,4	16,2 / 4,5	19,4 / 3,9	20,3 / 3,8
Verben (MW / SD)	12,8 / 3,1	14,4 / 3,1	15,9 / 3,0	17,9 / 2,9	19,0 / 2,5
Adjektive/Adverbien (MW / SD)	8,5 / 4,5	12,6 / 5,1	15,9 / 4,0	18,1 / 3,1	19,0 / 2,7
Kategoriale Nomen (MW / SD)	8,7 / 3,6	11,2 / 3,3	13,4 / 3,5	16,2 / 3,5	17,4 / 3,2

Tabelle 1: Mittelwert (MW) und Standardabweichung (SD) der Itemschwierigkeit für die vier Itemkategorien in Altersgruppen

Die steilste Altersentwicklung ist in der Kategorie der Adjektive/Adverbien zu finden. Dies kann allerdings auch als Reflex der unterschiedlichen Häufigkeiten bewertet werden, mit denen Wörter der verschiedenen Itemkategorien auftreten. Insbesondere an der unterschiedlichen Streuung der Typefrequenz im childLex-Korpus (annotierte Typefrequenz absolut bezogen auf 1 Mio. laufende Wörter im Korpus von Schroeder et al. 2015) sind die wortartspezifischen Unterschiede gut zu erkennen, so ist die Spanne bei den Adjektiven/Adverbien von "wolkenlos" Typefrequenz = 9 bis "schnell" Typefrequenz = 4720 besonders groß:

- Nomen: MW: 66,52; SD: 75,76 MIN: 1; MAX: 327
- Kategoriale Nomen: MW: 107,82; SD: 142,159; MIN: 1; MAX: 626
- Verben: MW: 140,96; SD: 249,205; MIN: 3; MAX: 915
- Adj./Adv.: MW: 642,09; SD: 1004,563; MIN: 9; MAX: 4720

In gewisser Weise werden so im WWT-Itemset die Kompositionsveränderungen im kindlichen Lexikon nachvollzogen (Schulze 2017).

#### 4.7 Thematische Bereiche

Bei der Auswahl der Items wurde darauf geachtet, dass diese verschiedenen thematischen Bereichen zugeordnet werden können. So soll sichergestellt werden, dass persönliche Erfahrungen bzw. Erfahrungsdefizite in einzelnen semantischen Bereichen nicht entscheidend auf das Gesamtergebnis durchschlagen. Dabei folgt die semantische Kategorisierung nicht einem Ansatz semantischer (geteilter bzw. distinkter) Merkmale (sensu Collins & Quillian 1969), sondern erfasst vielmehr in einem ontologischen Zugriff Ausschnitte der Welt, die Referenten versammeln, die untereinander in dichter Beziehung stehen (thematische Felder). Die dabei entstehende Gliederung des Wortschatzes ist wesentlich vom angenommenen Abstraktionsniveau und kultureller Prägung abhängig und unterliegt bei Kindern den Veränderungen ihrer semantischen Entwicklung (Schulze 2017).

Für die Zuordnung der Items im WWT zu solchen thematischen Feldern wurde auf die Sachgruppen-Gliederung nach Dornseiff (Dornseiff et al. 2004)

zurückgegriffen. Die Items sind 17 Bereichen mit einer Besetzung von 1 – 16 Exemplaren (im Mittel 5,6 je Gruppe) zuzuordnen, wodurch eine breite Abdeckung, wenn auch mit Schwerpunkten in den Sachgruppen (SG) "Geräte/Technik" und "Raum/Lage/Form" entsteht. Ein genauerer Blick in die Zuordnung offenbart Schwierigkeiten. Diese entstehen zum einen durch Polysemie. Das Item "Gipfel" lässt sich, das Elizitationsbild legt es (fast) nahe, zur SG 4 "Größe, Menge, Zahl, Grad" und Subkategorie (SK) 12 "Hoch" zuordnen, gehört aber je nach Konnotation in dieser SG auch zu SK 41 "Gesamtheit" und zu SG 3 "Raum, Lage, Form", SK 3 "Anwesenheit" und SK 48 "Wölbung", zu SG 11 "Fühlen, Affekte, Charakter", SK 36 "Wunsch" und SK "Abneigung" sowie zu zwei weiteren SG. Nicht immer lässt sich zweifelsfrei zuordnen, welche Bedeutungsmerkmale im Vordergrund stehen und eine Zuordnung begründen.

Zum anderen bildet die Sachgruppengliederung nicht die bei Kindern psychologisch reale, eher thematisch geprägte Organisationslogik ihres semantischen Wissens ab. So finden sich in der stark besetzten Sachgruppe "Geräte/Technik" auch Möbel- und Kleidungsstücke, während alle Fahrzeuge in die Sachgruppe "Ort/Ortsveränderung" fallen. Daher wird der Versuch unternommen, in einer eigenen, thematischen Gliederung in 15 Sachbereiche die Erfahrungswelt der Zielgruppe angemessener abzubilden. Die hieraus entstehende Zuordnung zeigt eine stärkere Gleichverteilung der Items auf die Sachbereiche. Fünf Bereiche umfassen jeweils zwischen 12% und 9% der Items ("Küche und Kochen", "Schule und Freizeit", "Natur", "Körperteile, Körperpflege, Gesundheit", "Soziales und Gefühle"). Sieben Bereiche werden jeweils durch 5% bis 7% der Items repräsentiert und nur drei Bereiche ("Zeit", "Kunst", "Form") versammeln zwischen 2% bis 4% der Items. In dieser Gliederung wird stärker deutlich, wie lebensnahe Bereiche der Zielgruppe (Schule, Freizeit, Küche, Körperteile, Kleidung...) im Itemset des WWT ausgewogen repräsentiert sind.

Und doch schließt sich hier die Frage an, inwieweit die Items an kulturelle Erfahrungen gebunden sind und damit einerseits dem gesellschaftlichen Wandel unterliegen und andererseits der gebotenen kulturellen Sensibilität entgegenstehen. Seit der Entstehung des WWT sind durch Veränderungen (mit Schlaglichtbeispielen) in Umwelt ("Jahrhunderthochwasser"), Gesellschaft ("Rechtsterrorismus"), Kommunikation ("Soziale Netzwerke"), Technik ("Smartphone") sowie für die Zielgruppe gegenwärtig höchst bedeutsam im Bereich Bildung ("Distanzlernen") Veränderungen im Lexikon sowohl im Lexem-Bestand als auch in der semantischen Organisation anzunehmen. Auch wenn naturgemäß im WWT-Itemset keine neuen Lexeme aufgenommen worden sind, so lässt sich doch auch für fast alle Items behaupten, dass sie Konzepte bezeichnen, deren Üblichkeit sich im Erfahrungsraum von Kindern in dieser Zeit kaum geändert haben dürften (Ausnahmen evtl.: "stricken", "Riegel").

Auch wenn Wert darauf gelegt wurde, Konzepte aus einer allgemein zugänglichen Lebenswelt anzusprechen, so sind dennoch Items bzw. die zugeordneten Bilder z.T. nicht frei von kultureller Gebundenheit (z.B. Bild zum Zielwort "früh" als Gegenteil von "spät" stellt ein "deutsches Frühstück" dar, das nicht der Alltagserfahrung aller Kinder entspricht). Das Item "Schleier" wird durch ein Bild mit Hochzeitsschleier elizitiert, was ggf. mit Schleier-Polysemien anderer Sprachräume im Konflikt steht. Diese Kritik gilt auch für das Item "Feste" für das Bilder verschiedener, in Deutschland traditionell üblicher Festanlässe präsentiert werden. Für Ansätze, dies zukünftig besser zu kontrollieren, kann auf Ergebnisse crosslinguistischer Studien zurückgegriffen werden (Vulić et al. 2021).

#### *4.8 Reihung der Items*

Durch Abfolge der Items könnten sowohl in der expressiven als auch in der rezeptiven Durchführung Priming-Effekte entstehen, die eine Abruffazilitation bewirken könnten. Phonologisches Priming könnte dann auftreten, wenn aufeinanderfolgende Wörter sich einander phonologisch im Anlaut oder Reim ähneln. Semantisches Priming, wenn Wörter mit geteilten Bedeutungsmerkmalen aus dem gleichen thematischen Feld aufeinander folgen und syntaktisches Priming, wenn Wörter gleicher Wortart aufeinander folgen.

Für die Durchführung des expressiven Subtests in der Standard-Durchführung kann ein Priming durch die festgelegte Itemreihenfolge weitgehend vermieden werden. Da alle anderen Subtests, einschließlich des rezeptiven Subtests in der Standard-Durchführung aber adaptiv angelegt sind und nur noch die Items enthalten, die im ersten expressiven Subtest nicht korrekt benannt worden sind, können die oben beschriebenen Priming-Effekte in diesen Subtests aus Gründen des Testaufbaus nicht ausgeschlossen und auch nicht systematisch berücksichtigt werden. Da der Fokus des Testverfahrens jedoch auf diesem ersten, sorgfältig zusammengestellten, expressiven Subtest liegt, kann dieses Problem als nachrangig hingenommen werden.

#### *4.9 Items der türkischsprachigen Version*

Der WWT 6-10 ist immer noch eines der wenigen sprachdiagnostischen Verfahren, das explizit für den mehrsprachigen Kontext entwickelt wurde (vgl. Groba 2018). Trotzdem bestehen beachtenswerte Einschränkungen: bislang liegt nur eine Version für das Türkische vor, die für das Türkische in einem Mehrsprachen-Erwerbskontext nicht gesondert normiert ist. Dennoch kann die Testdurchführung wichtige diagnostische Hinweise geben. Ausgangspunkt bei der Erstellung der türkischsprachigen Version sind die Aufgaben der deutschsprachigen Version, für die passende Lexeme der türkischen Sprache gesucht wurden. Dies hat bestimmte Implikationen und Folgeentscheidungen notwendig gemacht.

Mit der zusätzlich zur deutschsprachigen Version durchführbaren türkischsprachigen Version lassen sich die semantisch-lexikalischen Fähigkeiten Deutsch-Türkisch mehrsprachiger Kinder angemessener erfassen. Dabei erfolgen alle Instruktionen durch eine türkischsprachige Testleiter\_in oder/und über die Audiofiles in türkischer Sprache vom PC.

So kann bereits bei adaptiver Testung (beginnend mit Subtests auf Deutsch und türkischsprachiger Nachtestung auf Deutsch nicht korrekt benannter bzw. ausgewählter Items) der Konzeptwortschatz ermittelt werden (Groba 2018). Mit einer nicht-adaptiven Testung (vollständige Durchführung von Subtests auf Deutsch und auf Türkisch) lassen sich Vergleiche im Entwicklungsstand (bzw. in der Aktivierbarkeit) des Wortschatzes im Deutschen und Türkischen vornehmen. Denn für eine Diagnostik einer Sprachentwicklungsstörung ist es erheblich, ob eine Auffälligkeit nur eine Sprache oder beide Sprachen betrifft (siehe Beitrag von Till 2021). Außerdem wird die Klassifikationsleistung durch die Testung beider Sprachen verbessert (Peña et al. 2016). Allerdings müssen hierbei neben den individuellen Erwerbsbedingungen auch die sprachlichen und kulturellen Besonderheiten berücksichtigt werden, die sich im Itemset des WWT niederschlagen. So werden im Türkischen Verben häufig mit "machen" gebildet ("malen"/"resim yapmak", "demonstrieren"/"gösteri yapmak", "trösten"/"teselli etmek"). Für stark kulturgebundene Items gibt es keine direkten ("Wappen") oder mehrere differenzierte Entsprechungen ("Schleier"/"örtü, yaşmak, duvak").

Da sich die ergänzende türkischsprachige Version des WWT insbesondere an Kinder in Deutschland richtet, erfolgte die Orientierung in der Übertragung am in Deutschland gesprochenen Türkisch. Muttersprachliche Gewährsfrauen und Sprachtherapeutinnen haben die umfangreiche Übertragungsarbeit vorgenommen. Denn neben den als korrekte Antworten erwarteten Testwörtern auch die häufig auftretenden Falsch-Antworten bereits in der Datenbank des Tests hinterlegt. Noch dazu mussten für die Gestaltung der Distraktoren in der rezeptiven Auswahlaufgabe neue Wörter und Bilder gefunden werden. Die drei Distraktoren sind zum vorgegebenen Zielwort entweder semantisch, phonologisch oder nicht related. Insbesondere die phonologisch (Anlaut) relateden Items mussten entlang der türkischen Aussprache gefunden werden. Wenn semantische oder nicht-related Ablenker in der türkischen Sprache zufällig auch phonologische Ähnlichkeit mit dem Zielwort aufwiesen, so wurden auch diese ersetzt. Somit steht zumindest für das Türkische eine linguistisch kontrollierte Ergänzung zum deutschsprachigen WWT für die Sprachentwicklungsdiagnostik zur Verfügung.

## 5. Ausblick

Der WWT 6-10 ist ein in Praxis und Forschung gut etabliertes Instrument. Allerdings widmet es sich mit seinem semantisch-lexikalischen Fokus der sprachlichen Ebene, die am stärksten vom synchronen Sprachwandel betroffen ist (Bechmann 2016). Daher muss zumindest eine Prüfung der Normierung, besser eine neue Normierung erfolgen. Bei dieser Gelegenheit könnten Veränderungen im Itemset auch die in der Analyse des Itemsets bereits benannten Kritikpunkte bspw. der kulturellen Sensibilität aufgreifen. Es wäre zu prüfen, inwieweit moderne Konzepte semantischer Modellierung wie das der semantischen Reichhaltigkeit bzw. Dichte (*semantic richness*: Pexman 2020, und *semantic density*: Freedman 2013) bei der Zusammenstellung der Items zur Varianz der Itemschwierigkeit beitragen oder gar als Gruppierungsmerkmale für Itemsubsets geeignet sind, da sie sowohl das Wortlernen als auch den Wortabruf beeinflussen können.

Auch im Bereich der Psychometrie und der notwendigen testtheoretischen Absicherung der Normdaten sind Entwicklungen zu beobachten. Raschskalierte Verfahren und solche, die nach dem *continuous norming*-Ansatz (Lenhard et al. 2018) normiert wurden, gewinnen an Bedeutung.

Auch wenn in diesem Beitrag die linguistischen und psycholinguistischen Perspektiven überwogen, lässt sich vor dem Hintergrund der bald zwei Jahrzehnte langen Erfahrungen mit dem WWT 6-10 feststellen, dass die Qualität eines sprachentwicklungsdiagnostischen Verfahrens in der Beachtung und Verschränkung von drei Blickwinkeln entsteht: der (psycho-)linguistischen Perspektive, der Perspektive der Spracherwerbsforschung sowie der psychometrisch-methodologischen Perspektive.

## LITERATUR

- Alario, F. X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H. & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods*, 36(1), 140-155.
- Augst, G. (1984). *Kinderwort*. Frankfurt am Main, Bern, New York, Nancy: Lang.
- Bartha-Doering, L., Novak, A., Kollndorfer, K., Schuler, A.-L., Kasprian, G., Langs, G., Schwartz, E., Fischmeister, F. P. S., Prayer, D. & Seidl, R. (2018). Atypical language representation is unfavorable for language abilities following childhood stroke. *European Paediatric Neurology Society*, 23, 102-116. <https://www.sciencedirect.com/science/article/pii/S1090379818301855>
- Bastians, E., Eiba, C., Mühlenhoff, H., Papenbrock, S. & Wulff, U. (2017). Orientierungsrahmen (fach-)spezifischer Sprachförderung mehrsprachiger Kinder und Jugendlicher im Förderschwerpunkt Sprache. *VDS Sonderpädagogische Förderung in NRW*. [https://www.verband-sonderpaedagogik-nrw.de/fileadmin/user\\_upload/Orientierungsrahmen\\_Sprache.pdf](https://www.verband-sonderpaedagogik-nrw.de/fileadmin/user_upload/Orientierungsrahmen_Sprache.pdf)
- Bechmann, S. (2016). *Sprachwandel – Bedeutungswandel. Eine Einführung*. Tübingen: A. Francke Verlag.

- Berg, M. & Werner, B. (2014). PRIMA@Sprache. vergleichende Analysen zum Sprachverständnis bei Schülern der Klasse  $\frac{3}{4}$  an Grund-, Sprachheil- und Förderschulen. In S. Sallat, M. Spreer & C. W. Glück (Hgg.), *Sprache professionell fördern: kompetent, vernetzt, innovativ*. Idstein: Schulz-Kirchner. [https://www.pedocs.de/volltexte/2016/11868/pdf/Berg\\_Werner\\_2014\\_PRIMA\\_Sprache\\_vergleichende\\_Analysen.pdf](https://www.pedocs.de/volltexte/2016/11868/pdf/Berg_Werner_2014_PRIMA_Sprache_vergleichende_Analysen.pdf)
- Birchenough, J. M. H., Davies, R. & Connelly, V. (2017). Rated age-of-acquisition norms for over 3,200 German words. *Behavior Research Methods*, 49(2), 484-501.
- Bird, H., Franklin, S. & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods*, 33(1), 73-79.
- Brysbaert, M., Warriner, A. B. & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3), 904-911.
- Collins, A. M. & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8(2), 240-247.
- Czapka, S., Klassert, A. & Festman, J. (2019). Executive functions and language: Their differential influence on mono- vs. multilingual spelling in primary school. *Frontiers in Psychology*, 10, 97.
- Decristan, J., Fauth, B., Heide, E. L., Locher, F. M., Troll, B., Kurucz, C. et al. (2020). Students' differential participation in classroom discourse in primary schools. Who participates (not), and what are the consequences for student learning? *Zeitschrift für Pädagogische Psychologie*, 171-186. <https://econtent.hogrefe.com/doi/ref/10.1024/1010-0652/a000251>
- Dell, G. S. & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42(1-3), 287-314.
- Dietrich, R. & Gerwien, J. (2017). *Psycholinguistik. Eine Einführung* (3. Auflage). Stuttgart: J. B. Metzler Verlag.
- Dornseiff, F., Quasthoff, U. & Wiegand, H. E. (2004). *Der deutsche Wortschatz nach Sachgruppen* (8. Aufl.). Berlin: de Gruyter.
- Duden-Redaktion. (2021). *Sprachwissen: Die Verteilung der Wortarten im Rechtschreibduden*. Zugriff am 10.03.2021. <https://www.duden.de/sprachwissen/sprachratgeber/Die-Verteilung-der-Wortarten-im-Rechtschreibduden>
- Dunn, L. M. & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test IV. Deutschsprachige Bearbeitung: A. Lenhard, W. Lenhard, R. Segerer, S. Suggate*. Bloomington, MN: Pearson Assessments.
- Freedman, S. E. (2013). Children's naming as a function of neighbourhood density. *CPSLP* 15(1), 19-24.
- Glück, C. W. (2007). *Wortschatz- und Wortfindungstest für 6- bis 10-Jährige. WWT 6-10*. München: Elsevier, Urban & Fischer.
- Glück, C. W. (2008). Diagnostik semantisch-lexikalischer Störungen. In M. Grohnfeldt (Hg.), *Lehrbuch der Sprachheilpädagogik und Logopädie, Bd. 3 Diagnostik, Prävention und Evaluation*. Stuttgart: Kohlhammer.
- Glück, C. W. (2011). *Wortschatz- und Wortfindungstest für 6- bis 10-Jährige. WWT 6-10* (2., überarb.). München: Elsevier, Urban & Fischer.
- Groba, A. (2018). Diagnostik der Sprachentwicklung bei mehrsprachigen Kindern. In M. Spreer & B. Achhammer (Hgg.), *Diagnostik von Sprach- und Kommunikationsstörungen im Kindesalter. Methoden und Verfahren* (S. 156-184). München: Ernst Reinhardt Verlag.
- Groba, A. & De Houwer, A. (2018). Einschätzungsdaten zum rezeptiven Erwerbsalter von 258 deutschen Adjektiven mit Implikationen für die kindliche Adjektiventwicklung. In C. Baumann, V. Dabóczy & S. Hartmaier (Hgg.), *Adjektive* (S. 350-392). Berlin, München, Boston: De Gruyter.
- Günther, K.-B. (1988). Probleme der Diagnostik semantisch-lexikalischer Entwicklungsstörungen am Beispiel des aktiven Wortschatztests für drei- bis sechsjährige Kinder (AWST 3-6). In K.-B. Günther (Hg.), *Sprachstörungen* (S. 117-166). Heidelberg: Winter.



- Höhle, B. (2012). *Psycholinguistik*. Berlin: Akademie Verlag.
- Juska-Bacher, B., Beckert, C., Stalder, U. & Schneider, H. (2016). Die Bedeutung des Wortschatzes für basale Lesekompetenzen. *Halbjahresschrift für die Didaktik der deutschen Sprache und Literatur* 21(40), 20-40. [https://www.pedocs.de/volltexte/2019/16938/pdf/Didaktik\\_Deutsch\\_2016\\_40\\_Juska-Bacher\\_Die\\_Bedeutung\\_des\\_Wortschatzes.pdf](https://www.pedocs.de/volltexte/2019/16938/pdf/Didaktik_Deutsch_2016_40_Juska-Bacher_Die_Bedeutung_des_Wortschatzes.pdf)
- Juska-Bacher, B. & Jakob, S. (2014). Wortschatzumfang und Wortschatzqualität und ihre Bedeutung im fortgesetzten Spracherwerb. *Zeitschrift für angewandte Linguistik*, 61(1), 49-75.
- Kauschke, C. (2000). *Der Erwerb des frühkindlichen Lexikons. Eine empirische Studie zur Entwicklung des Wortschatzes im Deutschen* (Bd. 27). Tübingen: G. Narr Verlag.
- Kauschke, C., Bahn, D., Vesker, M. & Schwarzer, G. (2017). Die semantische Repräsentation von Emotionsbegriffen bei Kindern im Grundschulalter. *Kindheit und Entwicklung*, 26(4), 251-260.
- Kiese-Himmel, C. (2005). *AWST-R. Aktiver Wortschatztest für 3- bis 5-jährige Kinder - Revision*. Göttingen: Hogrefe.
- Kuperman, V., Stadthagen-Gonzalez, H. & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, 44(4), 978-990.
- Lenhard, A., Lenhard, W., Suggate, S. & Segerer, R. (2018). A continuous solution to the norming problem. *Assessment*, 25(1), 112-125.
- Levelt, W. J. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences of the United States of America*, 98(23), 13464-13471.
- Mayer, A. (2012). Semantisch-lexikalische Störungen - Basisartikel. *Sprachförderung und Sprachtherapie*, (2), 5462. [https://www.edu.lmu.de/shp/team/lehrstuhlinhaber/andreas\\_mayer/publikationen/mayer\\_2012\\_semantisch.pdf](https://www.edu.lmu.de/shp/team/lehrstuhlinhaber/andreas_mayer/publikationen/mayer_2012_semantisch.pdf)
- Montanari, E. & A. Nagi (2021). Über Antworten, die keine Antworten auf Testfragen sind und trotzdem etwas über das Wissen sagen können. *Zielsprache Deutsch*, 48(1), 3-19.
- Motsch, H.-J. & Ulrich, T. (2012). Effects of the strategy therapy 'exicon pirate' on lexical deficits in preschool age: A randomized controlled trial. *Child Language Teaching and Therapy*, 28(2), 150-175.
- Naumann, C. L. (1999). *Orientierungswortschatz. Die wichtigsten Wörter und Regeln für die Rechtschreibung Klasse 1 bis 6* (Beltz Praxis, Überarb., 4. Aufl). Weinheim, Basel: Beltz.
- Nicoladis, E. & Jiang, Z. (2018). Language and cognitive predictors of lexical selection in storytelling for monolingual and sequential bilingual children. *Journal of Cognition and Development*, 19(4), 413-430.
- Paetsch, J., Felbrich, A. & Stanat, P. (2015). Der Zusammenhang von sprachlichen und mathematischen Kompetenzen bei Kindern mit Deutsch als Zweitsprache. *Zeitschrift für Pädagogische Psychologie*, 19-29.
- Pafel, J. & Reich, I. (2016). *Einführung in die Semantik. Grundlagen, Analysen, Theorien* (Lehrbuch). Stuttgart: J.B. Metzler Verlag.
- Paul, C., Brensing, M., Frickenhaus, C., Frisch, A., Heinrich, J., Normann, A., Peters, J. & Schüller, T. (2018). *Sprachdiagnostik und Förderung im Regel- und DaZ-Unterricht. Ein Kooperationsprojekt der Universität zu Köln und der Maria Montessori Grundschule Köln-Ossendorf*. Universität zu Köln. [https://www.hf.uni-koeln.de/data/fgpsych/File/Paul/Paper\\_MonteProjekt\\_Paul.pdf](https://www.hf.uni-koeln.de/data/fgpsych/File/Paul/Paper_MonteProjekt_Paul.pdf)
- Peña, E. D., Bedore, L. M. & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one. *International Journal of Language & Communication Disorders / Royal College of Speech & Language Therapists*, 51(2), 192-202.
- Perret, C. & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, 51(6), 2533-2545.

- Petermann, F. (2010). *Sprachstandserhebungstest für Fünf- bis Zehnjährige (SET 5-10)*. Göttingen: Hogrefe.
- Pexman, P. M. (2012). Meaning based influences on visual word recognition. In *Visual word recognition*, 1, 24-43. London u.a.: Psychology Press.
- Pexman, P. M. (2020). How does meaning come to mind? Four broad principles of semantic processing. *Canadian Journal of Experimental Psychology = Revue Canadienne De Psychologie Experimentale*, 74(4), 275-283.
- Pregel, D. & Rickheit, G. (1987). *Der Wortschatz im Grundschulalter. Häufigkeitswörterbuch zum verbalen, substantivischen und adjektivischen Wortgebrauch*. Hildesheim, New York: G. Olms.
- Ptok, M., Kühn, D. & Miller, S. (2014). Wortschatzerwerb. Konstruktion verschiedener in der Praxis eingesetzter Wortschatztests. *HNO*, 62(4), 258-265.
- Rauch, D. P., Schastak, M. & Richter, K. (2016). Diagnostik sprachlicher Kompetenzen bei Personen mit Migrationshintergrund. In D. Maehler & H. U. Brinkmann (Hgg.), *Methoden der Migrationsforschung* (S. 97-133). Wiesbaden: Springer Fachmedien Wiesbaden.
- Schastak, M., Reitenbach, V., Rauch, D. P. & Decristan, J. (2017). Türkisch-deutsch bilinguale Interaktion beim Peer-Learning in der Grundschule. Selbstberichtete Gründe für die Annahme oder Ablehnung bilingualer Interaktionsangebote. *Zeitschrift für Erziehungswissenschaft*, 20, 213-235.
- Schöfl, M., Kloo, D. & Kaufmann, L. (2015). Planungsleistungen bei Grundschulern mit ADHS und LRS. Ein Vergleich von Fremdbeurteilungsverfahren und psychometrischen Testverfahren. *Lernen und Lernstörungen*, 4, 9-29.
- Schönauer-Schneider, W. & Reber, K. (2014). Schüler im Blick. Bausteine zur sprachheilpädagogischen Diagnostik IM Unterricht. In S. Sallat, M. Spreer & C. W. Glück (Hgg.), *Sprache professionell fördern: Kompetent, vernetzt, innovativ* (S. 327-334). Idstein: Schulz-Kirchner.
- Schröder, A., Gemballa, T., Ruppin, S. & Wartenburger, I. (2012). German norms for semantic typicality, age of acquisition, and concept familiarity. *Behavior Research Methods*, 44(2), 380-394.
- Schröder, A., Kauschke, C. & Bleser, R. de. (2003). Messungen des Erwerbsalters für konkrete Nomina. *Neurolinguistik*, 17(2), 83-114.
- Schroeder, S., Würzner, K.-M., Heister, J., Geyken, A. & Kliegl, R. (2015). childLex – Eine lexikalische Datenbank zur Schriftsprache für Kinder im Deutschen. *Psychologische Rundschau*, 66(3), 155-165.
- Schulze, J. (2017). *Lexikalischer Spracherwerb im Vorschulalter eine Annäherung an die Komposition des Lexikons von deutschsprachigen Kindern im Alter von drei bis fünf Jahren Erstellung eines Korpuslexikons auf der Basis spontaner Sprachdaten (Dissertationsschrift)*. Berlin. [https://depositonce.tu-berlin.de/bitstream/11303/6192/5/schulze\\_julia.pdf](https://depositonce.tu-berlin.de/bitstream/11303/6192/5/schulze_julia.pdf)
- Scott, G. G., Keitel, A., Becirspahic, M., Yao, B. & Sereno, S. C. (2019). The Glasgow norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, 51(3), 1258-1270.
- Seifert, S., Schwab, S. & Gasteiger-Klicpera, B. (2016). Effects of a whole-class reading program designed for different reading levels and the learning needs of L1 and L2 children. *Overcoming Learning Difficulties*, 32(6), 499-526.
- Siegmüller, J. (2008). Therapie von Kindlichen Wortfindungsstörungen nach dem Patholinguistischen Therapieansatz. *Forum Logopädie*, 22(5), 6-13.
- Snippe, K. (2015). *Autismus. Wege in die Sprache*. Idstein: Schulz-Kirchner Verlag. [https://books.google.de/books?hl=de&lr=&id=t4UNhNMGpKQC&oi=fnd&pg=PA5&ots=6JAJy0VskW&sig=TS3BnQyJDVrBq\\_IFAJyB9FLZ1yU&redir\\_esc=y#v=onepage&q=WWT&f=false](https://books.google.de/books?hl=de&lr=&id=t4UNhNMGpKQC&oi=fnd&pg=PA5&ots=6JAJy0VskW&sig=TS3BnQyJDVrBq_IFAJyB9FLZ1yU&redir_esc=y#v=onepage&q=WWT&f=false)

- Snodgrass, J. G. & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174-215.
- Spreer, M. & Achhammer, B. (Hgg.). (2018). *Diagnostik von Sprach- und Kommunikationsstörungen im Kindesalter. Methoden und Verfahren* (UTB, Bd. 4946). München: Ernst Reinhardt Verlag.
- Stadthagen-Gonzalez, H. & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods*, 38(4), 598-605.
- Tiede, S. & Braun, J.-U. (2017). Ist Chancengerechtigkeit für Kinder mit Sprachentwicklungsstörungen schon Realität? Eine empirische Querschnittstudie zur Quantifizierung des Bedarfs sprachtherapeutischer Interventionen im Primarbereich. *Forschung Sprache*, (1). [https://www.forschung-sprache.eu/fileadmin/user\\_upload/Dateien/Heftausgaben/2017-1/forschungSprache\\_Tiede.pdf](https://www.forschung-sprache.eu/fileadmin/user_upload/Dateien/Heftausgaben/2017-1/forschungSprache_Tiede.pdf)
- Till, C. (2021). Semantisch-lexikalische Diagnostik bei sukzessiv-bilingualen Kindern. *Bulletin suisse de linguistique appliquée*, 113, 175-191.
- Vulić, I., Baker, S., Ponti, E. M., Petti, U., Leviant, I., Wing, K., Majewska, O., Bar, E., Malone, M., Poibeau, T., Reichart, R. & Korhonen, A. (2021). Multi-SimLex: A large-scale evaluation of multilingual and crosslingual lexical semantic similarity. *Computational Linguistics*, 46(4), 847-897.



# Das Konstrukt Wortschatz: Dimension(en) Umfang und Tiefe? Empirische Ergebnisse aus der Unterstufe

**Britta JUSKA-BACHER**

PHBern

Institut Primarstufe & Institut für Forschung, Entwicklung und Evaluation

Fabrikstrasse 8, CH-3012 Bern

britta.juska@phbern.ch

**Martina RÖTHLISBERGER**

PHBern

Institut Forschung, Entwicklung und Evaluation

Fabrikstrasse 2a, CH-3012 Bern

martina.roethlisberger@phbern.ch

Vocabulary breadth and depth are considered as two dimensions of individual word knowledge. However, it is still unclear which aspects the vocabulary depth encompasses and whether in practice the dimensions can be clearly distinguished. In this article, we approach these questions using data from a longitudinal study with about 300 children from first to third grade. In addition to vocabulary breadth, we measured semantic word knowledge and network knowledge as aspects of vocabulary depth. Correlations between the three vocabulary measures increase slightly over time, but are still only moderately strong, indicating that the three variables should be kept separate. Factor analyses indicated two distinct vocabulary factors for all grades, with semantic word knowledge loading on a different factor than the other two measures. For children of early primary school age, a separation of the dimensions vocabulary breadth and depth and a breakdown of depth into different aspects is helpful to explain variation in individual word knowledge.

**Keywords:**

vocabulary size, vocabulary breadth, vocabulary depth, network knowledge, precision of meaning, semantic word knowledge, vocabulary test.

**Stichwörter:**

Wortschatzumfang, Wortschatzbreite, Wortschatztiefe, relationales Wortwissen, semantisches Wortwissen, Wortschatztest.

## 1. Wortwissen, seine Dimensionen und seine Bedeutung

Individuelles Wortwissen setzt sich aus einer Vielzahl von Informationen über Wörter und Wortverbindungen zusammen. Dazu gehört bspw. semantisches Wissen, Wissen zur Aussprache und Schreibung sowie zum Gebrauch von Wörtern u.a. in verschiedenen Registern. Da eine Abbildung des im mentalen Lexikon gespeicherten Wortwissens als Gesamtkonstrukt<sup>1</sup> nicht möglich ist, wurde vorgeschlagen, es empirisch über mehrere messbare Teilkompetenzen

---

<sup>1</sup> Für eine Diskussion der Begriffe "Konstrukt" und "Dimension" sowie weitere Dimensionen des Konstrukts Wortschatz siehe Gyllstad (2013).



zu erheben. Welche davon dieses Konstrukt gut abbilden, darüber gibt es bisher keine Einigkeit. Ausgehend von der Zweit- und Fremdsprachforschung werden im angloamerikanischen Raum seit den 1980er Jahren zwei bis vier verschiedene Dimensionen diskutiert. Den kleinsten gemeinsamen Nenner dieser Dimensionierungen bilden einerseits die *Wortschatzbreite*, d.h. die Zahl der einer Person bekannten Wörter, andererseits die *Wortschatztiefe*. Was genau letztere beinhaltet, auch darüber gehen die Meinungen auseinander, vorgeschlagen wurden etwa drei relevante Aspekte, nämlich semantisches, relationales und formales Wortwissen (z.B. Read 2004; Juska-Bacher & Jakob 2014).

Die Notwendigkeit der Unterscheidung dieser zwei Dimensionen wird aufgrund des hohen Zusammenhangs der Messdaten für Breite und Tiefe gelegentlich angezweifelt und stattdessen eine Beschränkung auf die Erhebung der Wortschatzbreite vorgeschlagen (z.B. Vermeer 2001). Reicht aber die Erhebung des Umfangs tatsächlich aus, um das Wortwissen eines Individuums angemessen abzubilden? Oder können zusätzliche Wortschatzdimensionen und -aspekte das Verständnis des mentalen Lexikons und des Zusammenspiels des Wortschatzes mit anderen Kompetenzbereichen wie dem Lesen erhöhen? Eine detaillierte Untersuchung des individuellen Wortwissens von Sprechenden scheint besonders angezeigt, wenn es darum geht, Grundlagen für eine Diagnose und gezielte Förderung zu schaffen.

Die Notwendigkeit einer gezielten Wortschatzförderung lässt sich aus seiner Bedeutung als "Schaltstelle" jeglicher schulischer Sprachhandlungen (Steinhoff 2009) herleiten. Bereits bei Schuleintritt bestehen grosse individuelle Unterschiede im Wortwissen (Moser et al. 2005) und Mängel in diesem Bereich werden als Risikofaktor für die weitere Schullaufbahn eingestuft (ebd.). Dass sich der Zusammenhang von Wortschatz und Schulerfolg durch die Schulzeit hindurchzieht, zeigen bspw. die Ergebnisse der DESI-Studie. Die fast 40% der Schülerinnen und Schüler im 9. Schuljahr, die über unzureichende Wortschatzkenntnisse verfügen (frequente Wörter des Grundwortschatzes sind nicht bekannt), besuchen in der Regel das unterste Schulniveau, d.h. die Hauptschule (Willenberg 2008).

Die Erforschung von Wortschatzkompetenzen sollte daher vorzugsweise frühzeitig und differenziert erfolgen, um mit abgestimmten Fördermassnahmen möglichst schnell einleiten zu können. Aus diesem Grunde wurde die im Folgenden beschriebene Studie mit Kindern am Beginn ihrer Schulzeit durchgeführt.

Während Wortschatz üblicherweise in seiner Funktion als Einflussvariable auf andere Kompetenzbereiche, insbesondere das Lesen, untersucht wird (siehe Schmitt 2014), steht er in diesem Beitrag selbst im Fokus. In einem ersten Schritt wird ausgeführt, was sich hinter den Dimensionen Wortschatzumfang und -tiefe verbirgt, wie sich diese im Schuleingangsalter messen lassen und

wie sie im Lichte der bisherigen Forschung zusammenhängen. Aus diesem Theorieteil werden Forschungsfragen für den empirischen Teil abgeleitet. Darin werden die Entwicklungsstände verschiedener Dimensionen und Aspekte von Wortschatz nachgezeichnet und ihre Zusammenhänge auf der Unterstufe beleuchtet, um Hinweise zu liefern, ob eine Unterscheidung von Wortschatzumfang und -tiefe in diesem Alter messbar und sinnvoll ist. Grundlage der vorgestellten Untersuchung bildet ein Datenauszug aus der SNF-Langzeitstudie "Wortschatz und Lesen. Eine Untersuchung auf der Unterstufe" (kurz: EnWoLe) mit insgesamt rund 350 Erst- bis Drittklässlern.

## 2. Wortschatzdimensionen und ihre Erhebung

Die Unterscheidung einer quantitativen Dimension Wortschatzbreite (vocabulary breadth) und einer qualitativen Dimension Wortschatztiefe (vocabulary depth) bestimmt seit einigen Jahrzehnten die Wortschatzforschung (z.B. Anderson & Freebody 1981; Vermeer 2001; Read 2004; Ouellette 2006; Tannenbaum et al. 2006; Schmitt 2014; Cox Eriksson 2021; im deutschsprachigen Raum Juska-Bacher 2014; Alber 2016; Jurecka et al. 2019).<sup>2</sup> Wortschatzbreite oder -umfang zielt auf die Gesamtzahl lexikalischer Einträge oder Knoten im Netz des mentalen Lexikons, deren Bedeutung dem Individuum zu einem Mindestgrad bekannt ist (Anderson & Freebody 1981). Die Messung des Wortschatzumfangs erfolgt dichotom (bekannt vs. nicht-bekannt), ohne dass eine Abstufung der Tiefe des semantischen Wissens erfolgt. International wie auch im deutschsprachigen Raum liegen für diese häufig erhobene Dimension verschiedene Testinstrumente vor (siehe Beiträge von Glück, Lenhard & Lenhard und Alber 2021).

Weniger deutlich ist die Dimension Wortschatztiefe oder -qualität definiert (Anderson & Freebody 1981), welche die einzelnen Worteinträge im mentalen Lexikon sowie ihre Verbindungen untereinander betrachtet und danach fragt, wie gut ein Individuum Wörter kennt. Wortschatztiefe umfasst als Sammelbegriff verschiedene Aspekte (z.B. Gyllstad 2013; Cain & Oakhill 2014; Schmitt 2014), zu denen sich eine Übersicht bei Read (2004) findet. Er

---

<sup>2</sup> Eine umfassende Operationalisierung des Wortwissens stammt von Nation (2001), der das Wissen zu einem Wort nach Form, Bedeutung und Gebrauch unterteilt, zu denen er je drei weitere Aufteilungen vorschlägt (ähnlich Schmitt 2010). Diese Klassifizierung wurde vor dem Hintergrund der Fremdsprachendidaktik entwickelt und ist teilweise für Kinder im Schuleingangsalter weniger geeignet (z.B. im Hinblick auf deren orthografisches Wissen, Wissen zu grammatischen Funktionen und Gebrauchsrestriktionen). Aus diesem Grunde wurde in der hier beschriebenen Studie eine schlichtere, altersangemessene Auslegung von Wortschatztiefe gewählt, die versucht, die Charakteristika des mentalen Lexikons durch die Zahl der Einträge (Wortschatzumfang), das semantische Wissen zu den einzelnen Einträgen sowie Verbindungen zwischen den Einträgen (relationales Wortwissen) einzubeziehen (Juska-Bacher & Jakob 2014; siehe auch Schoonen & Verhallen 2008). Zu verschiedenen Dimensionen siehe bspw. Gyllstad (2013).

nennt als erste Auslegungsmöglichkeit den Umfang semantischer Repräsentationen zu Wörtern (precision of meaning; im Folgenden: semantisches Wortwissen). Die zweite Auslegungsmöglichkeit ist breiter und umfasst neben semantischem auch formales Wortwissen, d.h. grammatisches, phonologisches und orthografisches Wissen (comprehensive word knowledge). Die dritte Auslegung schliesslich fokussiert die Struktur des mentalen Lexikons und thematisiert die Verknüpfungen der Wörter untereinander (nach Read 2004: network knowledge; im Folgenden: relationales Wortwissen). Für den unbedingten Einbezug dieses Aspekts plädieren beispielsweise Meara & Wolter (2004), wenn auch nicht unter der Sammelbezeichnung Wortschatztiefe.

Die komplexe Dimension der Wortschatztiefe lässt sich mit Hilfe von Tests, die nur einzelne Aspekte erheben, nicht angemessen fassen (Cain & Oakhill 2014; Schmitt 2014; Yanagisawa & Webb 2020). Bisher besteht aber keine Einigkeit darüber, welche der genannten Aspekte zentral sind und wie sich diese am besten operationalisieren lassen. Zudem sind die einzelnen Aspekte nicht distinkt, sondern weisen Überschneidungen auf (Schmitt 2014).

Semantisches Wortwissen wird häufig durch Definitionsaufgaben getestet, die anhand einer Skala bewertet werden. Ein standardisiertes Testinstrument für die Schuleingangsstufe gibt es im Deutschen nicht. Leichter messbar ist relationales Wortwissen (z.B. Hyperonym- und Antonymkenntnis). Im Deutschen steht dafür u.a. der WWT von Glück (2011, siehe Beitrag von Glück 2021) zur Verfügung. Eine andere Möglichkeit für das Testen relationalen Wortwissens stellen Assoziationsaufgaben dar (Überblick siehe Fitzpatrick 2013 und Cox Eriksson 2021).

Die Komplexität dieser Dimension dürfte dafür verantwortlich sein, dass Wortschatztiefe weit weniger häufig untersucht wird als der Umfang. Auch neuere Studien verwenden immer noch unterschiedliche Auslegungen der hier vorgestellten Aspekte von Wortschatztiefe und stellen damit jeweils verschiedene mögliche Annäherungen an die Messung dieser Dimension dar (Yanagisawa & Webb 2020). Beispielsweise operationalisierten auf der Primarschulstufe Karlsen et al. (2017) Wortschatztiefe anhand von semantischem Wortwissen (Definitionen), während Castillo & Tolchinsky (2018) relationales Wortwissen (Synonym- und Antonymkenntnis) erhoben. Bei älteren Versuchspersonen kamen auch umfangreichere Kombinationen zum Einsatz (z.B. Lawrence et al. 2019 und Tran et al. 2020).

### **3. Zusammenhang von Wortschatzumfang und -tiefe**

Eine Reihe von Studien hat den engen Zusammenhang zwischen Wortschatzumfang und Aspekten der -tiefe bei Kindern im Kindergarten- und Primarschulalter aufgezeigt (z.B. Vermeer 2001; Ouellette 2006; Tannenbaum et al. 2006; Ouellette & Beers 2010; im Deutschen: Alber 2016; Juska-Bacher



et al. 2016; Jurecka et al. 2019). Dieser Zusammenhang wird jedoch kontrovers diskutiert (siehe Schmitt 2014). Die zentrale Frage ist, ob eine Unterscheidung zwischen Wortschatzumfang und -tiefe für die Abbildung der Wortschatzkompetenzen eines Individuums überhaupt neue Erkenntnisse bringt oder ob sie nicht prinzipiell dasselbe messen. Vermeer (2001) interpretiert hohe Korrelationen zwischen Umfang und Tiefe bei Kindergartenkindern dahingehend, dass sich eine solche Unterscheidung nicht rechtfertigen lasse: Je mehr Wörter das Kind kenne, desto mehr Verbindungen zwischen Wörtern gäbe es in seinem mentalen Lexikon und desto tiefer sei sein Wissen zu Wörtern (Vermeer 2001: 231). Zu seiner Studie sei allerdings kritisch angemerkt, dass beide Wortschatzdimensionen mit Definitionsaufgaben gemessen wurden, wobei für den Umfang lediglich weniger detaillierte Angaben notwendig waren. Hier setzt Schmitts (2014) Kritik an, der zufolge nur eine trennscharfe Operationalisierung das Verhältnis von Wortschatzumfang und -tiefe angemessen bestimmen lässt. Auf Untersuchungen verweisend, die mit Hilfe von Regressionsanalysen gezeigt haben, dass Wortschatztiefe zusätzliche Erklärungskraft für das Lesen hat, das deutlich über den Wortschatzumfang hinausgeht (für Primarschulkinder lassen sich hierzu bspw. Studien von Ouellette 2006 mit Viertklässlern, Tannenbaum et al. 2006 mit Drittklässlern, Proctor et al. 2012 mit Zweit- bis Viertklässlern und Swart et al. 2017 mit Viertklässlern anführen), geht Schmitt (2014) zwar nicht von unabhängigen, aber dennoch unterschiedlichen Entwicklungen der beiden Dimensionen aus.

Read (2004) spricht eine mögliche Veränderung des Verhältnisses von Wortschatzumfang und -tiefe mit zunehmendem/n Alter bzw. Wortschatzkompetenzen an. Diese Annahme wird gestützt durch eine Untersuchung von Nurweni & Read (1999) mit älteren Fremdsprachenstudierenden, in der nachgewiesen werden konnte, dass Wortschatzumfang und -tiefe (relationales Wortwissen getestet durch Assoziationsaufgaben) bei fortgeschrittenen Lernenden deutlich stärker korrelierten. Einen gegensätzlichen Befund lieferte Henriksen (2008), deren Ergebnisse bei zunehmendem Sprachstand einen geringeren Zusammenhang zwischen Umfang und -tiefe (ebenfalls gemessen mit Hilfe von Assoziationen) zeigten. Bisher ist damit also offen, ob mit zunehmenden Sprachkenntnissen (steigendem Wortschatzumfang) die Verbindung zwischen Umfang und relationalem Wortwissen stärker oder schwächer wird (Schmitt 2014).

Als weitere für die Wortschatzentwicklung relevante Variablen wurden Alter, Geschlecht, sozioökonomische und kognitive Merkmale des Kindes untersucht. Hinsichtlich des Alters zeigen Kinder im Vorschul- wie im Schulalter im Zeitverlauf einen erheblichen und stetigen Zuwachs (Anglin 2008). Für das Geschlecht hingegen konnten in empirischen Studien sowie bei den Normierungsstichproben gängiger Wortschatztests bisher allenfalls

leichte Vorteile der Mädchen, aber keine statistisch gesicherten Unterschiede aufgezeigt werden (z.B. Lange et al. 2016, siehe auch Beiträge von Lenhard & Lenhard und Glück 2021). Der Einfluss des sozioökonomischen Status der Eltern (Rindermann & Baumeister 2015) und der kognitiven Leistungen auf den Wortschatz hingegen wurde bereits verschiedentlich nachgewiesen (Lenhard & Lenhard 2021). Die Frage, wie und ob diese Faktoren ggf. verschiedene Wortschatzdimensionen beeinflussen, ist ungeklärt.

#### 4. Forschungsfragen

Ausgehend von dem in Kapitel 2 und 3 zusammengefassten Forschungsstand sollen mit Hilfe der empirischen Daten folgende Fragen zu den Zusammenhängen zwischen den Dimensionen und Aspekten von Wortschatz geklärt werden:

1. Sind semantisches und relationales Wortwissen Aspekte einer gemeinsamen Dimension Wortschatztiefe?
2. Sind Wortschatzumfang und -tiefe – obwohl sie korrelieren – klar unterscheidbare Dimensionen?
3. Beeinflussen persönliche, sozioökonomische und kognitive Variablen die unterschiedlichen Dimensionen des Wortschatzes?

#### 5. Studiendesign

In der Langzeitstudie EnWoLe wurden jeweils von März bis Mai 2018 und 2019 mit 348 Erstklässlern bzw. 361 Zweitklässlern<sup>3</sup> sowie 2020 – wegen Corona bedingter Schulschliessungen in einem etwas verlängerten Zeitraum – von März bis September 2020 mit 336 Drittklässlern eine Reihe von Wortschatz- und Lesetests durchgeführt (Ausfallrate t1 zu t2: 8%, t2 zu t3: 7%).

Für diesen Beitrag wurden ausschliesslich die Wortschatztests, die IQ-Testergebnisse sowie persönliche Angaben der Kinder berücksichtigt, von denen vollständige Datensätze aller drei Erhebungszeitpunkte vorlagen,  $N=292$  aus 38 Klassen, davon 170 Mädchen und 122 Jungen mit einem Durchschnittsalter von 7;6 Jahren ( $SD = 4.5$  Monate) zu t1, 8;5 Jahren ( $SD = 4.4$  Monate) zu t2 und 9;7 Jahren zu t3 ( $SD = 4.9$  Monate). Die Kinder besuchten eine Primarschule in den Kantonen Bern, Basel-Stadt oder -land oder Luzern, sprachen Deutsch als Erstsprache und wiesen keine diagnostizierten Sprachentwicklungsbeeinträchtigungen auf.

Im Projekt wurden in Einzelsettings folgende Tests für die Messung von Wortwissen eingesetzt, die alle dekontextualisiert Wortwissen abfragen (siehe

---

<sup>3</sup> Die Zunahme der Zahl der teilnehmenden Kinder ist darauf zurückzuführen, dass zum zweiten Erhebungszeitpunkt Kinder aus drei weiteren Klassen gewonnen werden konnten.

Tabelle 1, für eine ausführliche Beschreibung der Testinstrumente siehe Juska-Bacher et al. 2021a).

Für die Messung des (rezeptiven) Wortschatzumfangs kam die deutsche Version des PPVT-4 (Lenhard et al. 2015; siehe Beitrag Lenhard & Lenhard 2021) zum Einsatz. Aus Zeitgründen wurde mit einer Testhalbierung (nur ungerade Items) gearbeitet. Da diese Übertragung auf den bundesdeutschen Raum abgestimmt ist, das Projekt aber in der Schweiz durchgeführt wurde, wurden einige deutsche Varianten durch Schweizerische ersetzt (z.B. Item 65: "Schornstein" – "Kamin", Item 70: "Umschlag" – "Kuvert").

Als Aspekte von Wortschatztiefe wurden semantisches Wortwissen und relationales Wortwissen erfasst. Da für das semantische Wortwissen für die Zielstufe im Deutschen bisher kein standardisiertes Testinstrument vorliegt, wurde wie bereits in einigen anderen Sprachen<sup>4</sup> auf den aus dem Englischen übersetzten Intelligenztest HAWIK-IV (Petermann & Petermann 2007, Untertest 6 Wortschatz) zurückgegriffen. Auch dieser Test wurde halbiert (Aufgaben 7-35, nur die 15 ungeraden Items). In diesem Teilttest werden den Testpersonen 15 Inhaltswörter genannt (z.B. "Regenschirm", "zwingen" und "mutig"), um eine Definition zu elizitieren. Das im Benutzerhandbuch vorgegebene dreistufige Bewertungssystem mit 0 bis 2 Punkten pro Definition<sup>5</sup> (zu den begrenzten Differenzierungsmöglichkeiten einer dreistufigen Skala siehe Tran et al. 2020) schien für die Beurteilung des semantischen Wortwissens, das in EnWoLe erhoben werden wollte, nicht ausreichend. Aus diesem Grund wurden vom Projektteam in einer semantischen Analyse eigene Bewertungskriterien für die einzelnen Items erarbeitet<sup>6</sup>, die je nach Wortart etwas unterschiedlich aussahen, und die Punkteskala auf vier Stufen (0 bis 3 Punkte) erhöht (siehe Juska-Bacher et al. 2021a), was eine grössere Differenzierung ermöglicht, zugleich aber bei der Bewertung noch handhabbar ist (mit der gleichen Skala arbeitet bspw. der englischsprachige Test von Wiig & Secord 1992).<sup>7</sup> Im Projekt wurden die Antworten der Kinder von jeweils zwei unabhängigen Raterinnen beurteilt. Mit Cohen's Kappa  $\kappa = 0.75$ ,  $SD = 0.1$ , Spannweite 0.6 – 1 (Mittelwerte aller drei Zeitpunkte) war die Interrater-

<sup>4</sup> Im Englischen z.B. von Cain & Oakhill (2014), im Niederländischen von Swart et al. (2017), im Norwegischen von Karlsen et al. (2017) und im Spanischen von Strasser & del Rio (2013).

<sup>5</sup> Eine falsche Antwort erhält 0 Punkte, ein treffendes Synonym, die Angabe des hauptsächlichen Verwendungszwecks oder ein Hyperonym das Maximum von 2 Punkten. Die übrigen Antworten, die korrekt, aber nicht ausreichend sind, werden mit 1 Punkt bewertet (Petermann & Petermann 2007).

<sup>6</sup> Diese Bewertungskriterien berücksichtigen bspw. Polysemien, die im Testmanual nicht aufgeführt sind (z.B. "fortziehen" sowohl intransitiv für einen Ortswechsel als auch transitiv im Sinne von 'jmdm. etwas wegnehmen') oder stufen Bedeutungsnuancen feiner ab als dies im Manual der Fall ist.

<sup>7</sup> Zur Problematisierung solcher Bewertungen von semantischem Wissen sowie zu im Englischen vorhandenen Testinstrumenten siehe Schmitt (2010: 217-224).

Reliabilität zufriedenstellend. Bei Abweichungen in der Bewertung wurden Definitionen von einer dritten Raterin unabhängig beurteilt.

Relationales Wortwissen wurde mit der Kurzform 2 des WWTexpressiv (Glück 2011; siehe auch Glück 2021) für das Alter von 7;0 bis 8;11 erfasst. Diese Fassung besteht aus 40 Items, von denen die Hälfte mit je zehn Antonymen und Hyperonymen als relationales Wortwissen in die Auswertung eingeht. Das Kind soll bildunterstützt Hyperonyme benennen bzw. ohne bildliche Unterstützung Antonyme zu Adjektiven/Adverbien bilden. Da die Stichprobe nur Kinder ohne diagnostizierte Sprachentwicklungsverzögerungen bestand, wurde auf Abrufhilfen verzichtet.

Variable	Testinstrument
Wortschatzumfang	Lenhard et al. (2015): PPVT (Testhalbierung)
Semantisches Wortwissen	Petermann & Petermann (2007): WISC-IV (Untertest 6 Wortschatztest mit Testhalbierung)
Relationales Wortwissen	Glück (2011): WWT (Kurzform 2 expressiv), Teil Antonyme und Hyperonyme

Tabelle 1: Untersuchte Variablen und die verwendeten Testinstrumente

Als Kontextvariablen wurden die demographischen Variablen Alter, Geschlecht und Beruf der Eltern nach dem standardisierten internationalen sozioökonomischen Index des beruflichen Status ISEI (siehe Ganzeboom et al. 1992; verwendet wurde der höhere Wert der Elternteile) sowie die nonverbale Intelligenz anhand des CFT 20-R (Weiß 2006, Untertest 3, Matrizen<sup>8</sup>) erfasst.

## 6. Ergebnisse

### 6.1 Masszahlen zu den Einzelvariablen

In Tabelle 2 sind die Masszahlen der einzelnen Wortschatzvariablen dargestellt. Die Verbesserung der Kinderleistungen wurde mit Varianzanalysen mit Messwiederholungen geprüft. Als abhängige Variable wurde das jeweilige Testresultat (Wortschatzumfang, semantisches und relationales Wortwissen) in Form der Rohwerte definiert. Die Klassenstufe wurde als Innersubjektvariable eingefügt. Der Fortschritt ist bei allen Wortschatzmassen zwischen t1 und t2 wie zwischen t2 und t3 signifikant, alle  $F > 236$ , alle  $p < .001$ , alle  $\eta^2 > 0.4$ . Es ist bei allen Aspekten eine grosse

<sup>8</sup> Dieser Testteil wurde ausgewählt, weil er am höchsten mit dem Gesamtergebnis korreliert (Weiß 2006).

Streuung der Ergebnisse festzustellen, die nur beim Wortschatzumfang von t1 zu t3 zu-, bei den anderen Aspekten aber abnimmt.

	1. Schuljahr				2. Schuljahr				3. Schuljahr			
	<i>Min</i>	<i>Max</i>	<i>Mittelwert</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>Mittelwert</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>Mittelwert</i>	<i>SD</i>
Wortschatzumfang	90	197	142.1	19.0	100	211	154.6	19.2	102	215	168.7	21.7
semantisches Wortwissen	2	29	14.5	4.8	0	31	20.0	5.7	4	36	22.9	5.2
relationales Wortwissen	1	19	7.8	3.8	2	19	11.3	3.5	5	20	14.1	3.0

*N* = 292 (170 w, 122 m)

Tabelle 2: Mittelwerte (Rohwerte) der Einzelvariablen

## 6.2 (Ko-)Varianzanalysen Einflussfaktoren

Um den Effekt des relativen Alters pro Zeitpunkt zu prüfen, wurde je eine multivariate Kovarianzanalyse pro Zeitpunkt durchgeführt. Als abhängige Variablen wurden die standardisierten Resultate der drei Wortschatzmasse bestimmt, und als Kovariaten das jeweilige Alter (mittelwertzentriert). Das Alter zeigte zu allen drei Zeitpunkten einen signifikanten Effekt auf den Wortschatzumfang. Zudem fand sich zu t1 ein Effekt des Alters auf das relationale Wortwissen. Alle anderen Effekte blieben insignifikant (siehe Tabelle 3).

Der Einfluss des Geschlechts wurde mit drei Varianzanalysen zu jedem Zeitpunkt geprüft. Als abhängige Variablen wurden die standardisierten Wortschatzmasse, als Zwischensubjektfaktor das Geschlecht definiert. Ein signifikanter Effekt des Geschlechts zugunsten der Jungen (t1:  $M = 0.15$ ,  $SE = 0.1$ ; t2:  $M = 0.83$ ,  $SE = 0.1$ ) gegenüber den Mädchen (t1:  $M = -0.12$ ,  $SE = 0.08$ ; t2:  $M = 0.53$ ,  $SE = 0.08$ ) auf den Wortschatzumfang zeigte sich zu t1,  $F(1, 290) = 5.1$ ,  $p = .024$ ,  $\eta^2 = 0.02$ , und zu t2,  $F(1, 290) = 6.7$ ,  $p = .010$ ,  $\eta^2 = 0.02$ , während dieser Effekt zu t3 nicht mehr signifikant war,  $F(1, 290) = 1.11$ ,  $p = .299$ ,  $\eta^2 < 0.01$ . Für die übrigen Wortschatzvariablen war kein Geschlechtseffekt feststellbar (alle  $F < 3.6$ , alle  $p > .061$ ). Da zudem in einer Varianzanalyse mit Messwiederholung keine dreifache Interaktion zwischen Wortschatzmasse, Geschlecht und Zeitpunkt sichtbar war,  $F(4, 1160) = 0.8$ ,  $p = .529$ ,  $\eta^2 = 0.003$ , wurden für die nachfolgenden Analysen beide Geschlechter als eine Gruppe behandelt.

Die Effekte des sozio-ökonomischen Status (SES) und der non-verbale Intelligenz (CFT) wurden mit Kovarianzanalysen pro Zeitpunkt getestet. Die Wortschatzmasse wurden als abhängige Variablen definiert, der SES respektive der CFT als Kovariablen. Die Wirkung des SES auf die abhängigen

Variablen war zu jedem Zeitpunkt signifikant, alle  $F > 5.4$ , alle  $p < .021$ . Der CFT zeigte hingegen recht unterschiedliche Effekte je nach Zeitpunkt und Wortschatzvariable, vgl. Tabelle 3.

Kovariate	Wortschatzmass	1. Schuljahr			2. Schuljahr			3. Schuljahr		
		F	p-Wert	$\eta^2$	F	p-Wert	$\eta^2$	F	p-Wert	$\eta^2$
Alter	Wortschatzumfang	9.21	<b>.003</b>	0.03	4.09	<b>.044</b>	0.01	13.19	< <b>.001</b>	0.04
	relat. Wortwissen	5.69	<b>.016</b>	0.02	1.65	.200	< 0.01	0.80	.373	< 0.01
	semant. Wortwissen	3.54	.061	0.01	0.89	.345	< 0.01	1.38	.242	< 0.01
SES	Wortschatzumfang	11.79	<b>.001</b>	0.04	8.88	<b>.003</b>	0.03	11.90	<b>.001</b>	0.04
	relat. Wortwissen	12.00	<b>.001</b>	0.04	17.28	< <b>.001</b>	0.06	22.54	< <b>.001</b>	0.07
	semant. Wortwissen	5.45	<b>.020</b>	0.02	12.17	<b>.001</b>	0.04	13.19	< <b>.001</b>	0.04
CFT	Wortschatzumfang	10.15	<b>.002</b>	0.03	7.78	<b>.006</b>	0.03	18.59	< <b>.001</b>	0.06
	relat. Wortwissen	15.36	< <b>.001</b>	0.05	13.48	< <b>.001</b>	0.04	11.52	<b>.001</b>	0.04
	semant. Wortwissen	0.88	.350	< 0.01	4.58	<b>.033</b>	0.02	6.68	<b>.010</b>	0.02

*N = 293 (122 m, 171 w), df = 291; signifikante p-Werte sind fettgedruckt*

Tabelle 3: Kovarianzanalysen

Der Einfluss von SES und CFT auf den Leistungszuwachs wurde mit einer Kovarianzanalyse mit Messwiederholungen geprüft. Als abhängige Variablen wurden alle Wortschatzresultate eingefügt, als Innersubjektsfaktoren die drei Klassenstufen und die drei Tests (Wortschatzumfang, semantisches und relationales Wortwissen) und als Kovariaten der ISEI und die Resultate des CFT. Die Analyse zeigte weder zwischen ISEI, Klassenstufe und Test eine signifikante Interaktion,  $F(3.81, 1100.29) = 0.65$ ,  $p = .617$ ,  $\eta^2 < 0.01$ , noch zwischen CFT, Klassenstufe und Test,  $F(3.81, 1100.29) = 1.47$ ,  $p = .212$ ,  $\eta^2 < 0.01$ .

### 6.3 Korrelationen der Wortschatzmasse

Allfällige lineare Zusammenhänge zwischen den standardisierten Variablen wurden mit Korrelationen berechnet. Die Verteilung der Daten wurde angesichts der Stichprobengrösse mit dem Normalitäts-Test Kolmogorov-Smirnov überprüft, der für alle drei Variablen zu allen Zeitpunkten signifikant ausfiel, alle  $p < .009$ . Angesichts der fehlenden Normalverteilung wurde ein nicht-parametrisches Verfahren gewählt.

Um die Vergleichbarkeit der Variablen zu gewährleisten, wurden sie auf der Basis der Werte aus der 1. Klasse gemäss der folgenden Formel standardisiert:

$$(X_{tn} - \mu_{t1}) / \sigma_{t1}$$

wobei gilt, dass  $X_{tn}$  für den Wert der Datenzelle zu jedem Zeitpunkt,  $\mu_{t1}$  für den Mittelwert von  $t1$  und  $\sigma_{t1}$  für die Standardabweichung zu  $t1$  steht.

Die Korrelationen zeigten über die drei Zeitpunkte eine zunehmende Tendenz, die einzig zu t3 zwischen relationalem und semantischem Wortwissen nicht auftritt (siehe Tabelle 4).

	1. Schuljahr			2. Schuljahr			3. Schuljahr		
	1.	2.	3.	1.	2.	3.	1.	2.	3.
1. Wortschatzumfang	–			–			–		
2. semantisches Wortwissen	.301**	–		.374**	–		.390**	–	
3. relationales Wortwissen	.367**	.324**	–	.467**	.493**	–	.514**	.463**	–

*N* = 292; \*\* *p* < .001; Cronbach's alpha zwischen allen Variablen:  $\alpha_{t1} = .61$ ;  $\alpha_{t2} = .70$ ;  $\alpha_{t3} = .69$

Tabelle 4: Korrelationen der standardisierten Variablen (Spearman's rho)

#### 6.4 Faktorenanalysen

Um zu prüfen, ob sich die manifesten Variablen auf eine oder zwei latente Variable(n) zurückführen lassen, wurden Faktorenanalysen pro Zeitpunkt evaluiert. Als Kriterien für die Validität solcher Modelle wurden die positiven Korrelationen zwischen den Wortschatzmassen sowie die üblichen Tests berücksichtigt. Mit einem Kaiser-Meyer-Olkin Kriterium über dem Schwellenwert von .6 und dem Bartlett's Test der Sphärität, alle  $\chi^2 > 86.5$ , alle *p* < .001, waren die Kovarianzen ausreichend für eine Faktorisierung. Um Korrelationen zwischen den Faktoren zuzulassen, wurde eine oblique Rotationsmethode gewählt (oblimin; Fabrigar et al. 1999). Die Resultate zeigen zu allen drei Zeitpunkten, dass das semantische Wortwissen in einer 2-Faktorenlösung einen eigenen Faktor bildet, während Wortschatzumfang und relationales Wortwissen in einen gemeinsamen Faktor laden (siehe Tabelle 5 und Abbildung 1).

Zeitpunkt	Modell	Variablen	RMSR	$\chi^2$	F1 Load	F2 Load	Kom	F1 Prp Var	F2 Prp Var
t1	1-Faktor	Wortschatz- umfang			0.74	–	0.45		
		Relationales Wortwissen	0.22	83.35	0.77	–	0.40	0.56	–
		Semantisches Wortwissen			0.74	–	0.46		
	2-Faktor	Wortschatz- umfang			0.94	-0.06	0.86		
		Relationales Wortwissen	0.18	58.56	0.56	0.39	0.60	0.41	0.38
		Semantisches Wortwissen				0.97	0.92		
t2	1-Faktor	Wortschatz- umfang			0.77	–	0.40		
		Relationales Wortwissen	0.19	60.16	0.84	–	0.29	0.63	–
		Semantisches Wortwissen			0.77	–	0.41		
	2-Faktor	Wortschatz- umfang			0.97	-0.05	0.91		
		Relationales Wortwissen	0.14	33.49	0.52	0.49	0.71	0.43	0.42
		Semantisches Wortwissen				0.97	0.92		
t3	1-Faktor	Wortschatz- umfang			0.78	–	0.39		
		Relationales Wortwissen	0.18	58.95	0.83	–	0.31	0.64	–
		Semantisches Wortwissen			0.78	–	0.40		
	2-Faktor	Wortschatz- umfang			0.97	-0.05	0.91		
		Relationales Wortwissen	0.14	35.54	0.52	0.47	0.70	0.43	0.41
		Semantisches Wortwissen				0.97	0.92		

*RMSR: Root Mean Square of Residuals; Load: Faktorladungen; Kom: Kommunalitäten; Prp Var: Proportion an Varianzaufklärung*

Tabelle 5: Kennzahlen 1- und 2-Faktorlösungen über alle drei Zeitpunkte



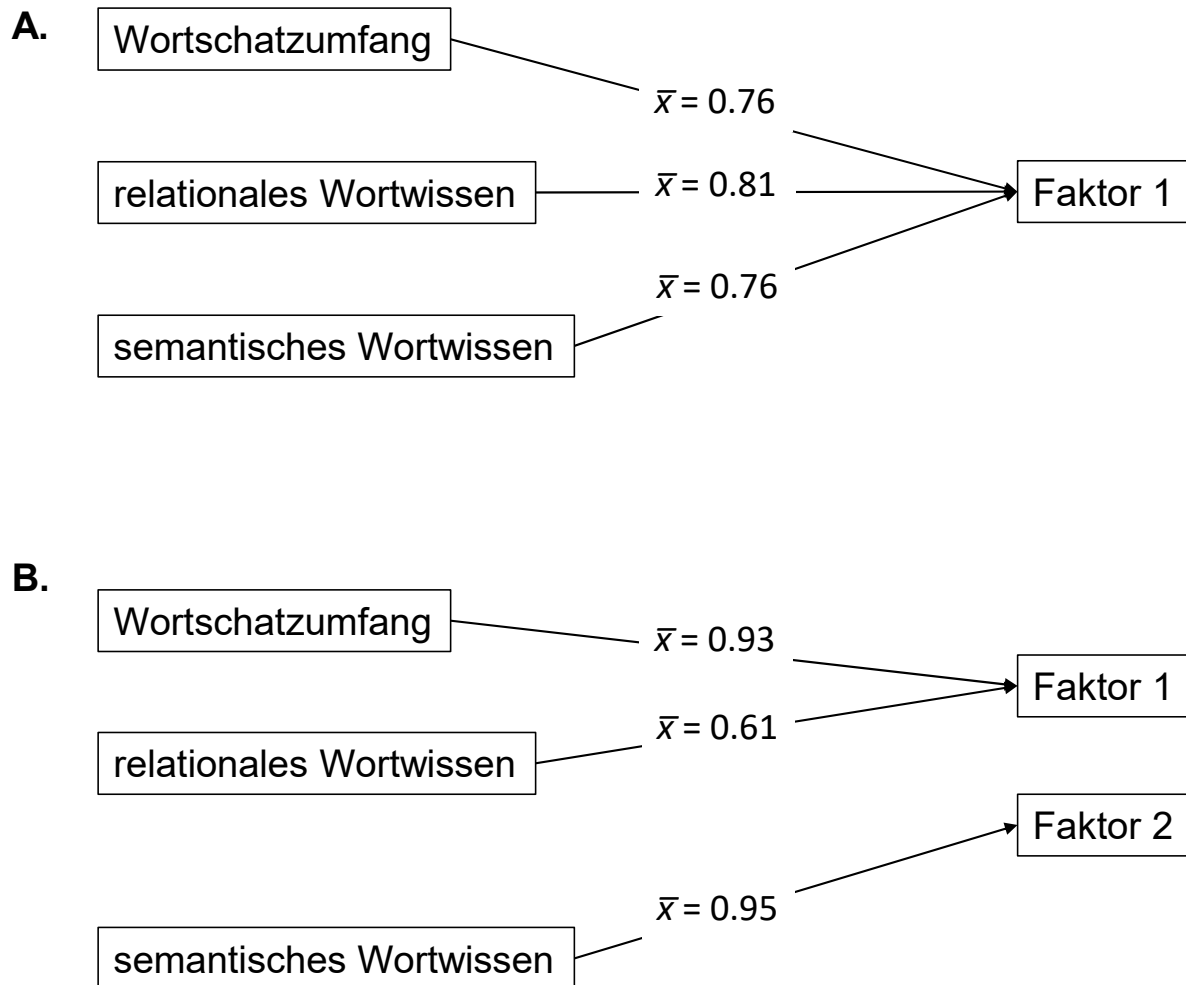


Abbildung 1: A. 1-Faktormodell; B. 2-Faktorenmodell; die Ladungen der Variablen auf den Faktor sind als Durchschnitt der drei Zeitpunkte angegeben.

## 7. Diskussion der Ergebnisse

Ziel dieses Beitrags war es, Zusammenhänge zwischen verschiedenen Dimensionen und Aspekten von Wortwissen sowie ihren Einflussfaktoren zu untersuchen und zu prüfen, ob auf der Unterstufe eine Unterscheidung verschiedener Wortschatzdimensionen angemessen ist.

### 7.1 *Semantisches und relationales Wortwissen als Aspekte einer gemeinsamen Dimension Wortschatztiefe (Forschungsfrage 1)*

Sowohl beim semantischen als auch beim relationalen Wortwissen ist der stetige Zuwachs über die drei ersten Schuljahre hinweg signifikant und vergleichbar. Die Standardabweichungen zeigen eine abnehmende Tendenz<sup>9</sup>,

<sup>9</sup> Dies gilt nicht für den Wortschatzumfang, bei dem sich die Standardabweichungen leicht vergrößern. Die Entwicklung der Standardabweichungen ist auch in anderen Studien unterschiedlich, longitudinal zunehmend bei Kindern mit durchschnittlichen bis

die sowohl eine sprachliche wie eine kognitive Entwicklung widerspiegeln (vgl. Lenhard und Lenhard 2021). Diese Hypothese wird unterstützt durch die prominente Wirkung der nonverbalen Intelligenz auf das relationale Wortwissen und die Zunahme dieser Wirkung auf das semantische Wortwissen. Der Effekt der von Lenhard & Lenhard 2021 beschriebenen Wechselwirkung zwischen Wortschatz und logischem Denken mit Sprache nimmt stetig zu, es kann von einem "Matthäus"-Effekt bei der Wortschatztiefe analog zu dem beim Lesen ausgegangen werden.

Der Zusammenhang zwischen semantischem und relationalem Wortwissen ist nicht sehr ausgeprägt und liegt mit einem Korrelationskoeffizienten zwischen .33 bis .49 niedriger als bei Schoonen & Verhallen (2008; mit einem Wert von .74), und ist als moderat zu qualifizieren. Der deutlich höhere Wert in der niederländischen Studie wurde möglicherweise dadurch unterstützt, dass die Hälfte der Wörter in beiden Tests identisch war (Schoonen & Verhallen 2008). Die Ladungen des semantischen Wortwissens auf den Faktor Wortschatz (siehe Faktoranalysen in Abbildung 1) sprechen zu allen drei Zeitpunkten dafür, das semantische Wortwissen als gesonderte Teilkompetenz eines Wortschatzkonstrukts im Rahmen eines Wortschatz-Tests zu prüfen. Das relationale Wortwissen hingegen lässt sich in dieser Studie aufgrund der Faktorenladung in den drei ersten Grundschuljahren nicht gemeinsam mit dem semantischen Wortwissen einer Dimension Wortschatztiefe zuordnen.

## *7.2 Korrelationen zwischen Wortschatzumfang und semantischem bzw. relationalem Wortwissen als Aspekte der Wortschatztiefe (Forschungsfrage 2)*

Generell sind die Korrelationen zwischen Wortschatzumfang und semantischem Wortwissen mit Werten zwischen .30 und .39 zwar statistisch signifikant, aber nicht in einem Bereich, der eine Zusammenfassung der beiden Kompetenzbereiche rechtfertigen würde. Die hier vorgestellten Ergebnisse entsprechen denen von Ouellette (2006) für Viertklässler (.38), bleiben hingegen tiefer als die Korrelationen von Tannenbaum et al. (2006) für Drittklässler (.49), Ouellette & Beers (2010) für Erst- und Sechstklässler (.75 bzw. .60) sowie auch von denen Tran's et al. (2020) für Erwachsene mit geringen Lesekenntnissen (.83). Die Resultate dieser Studien sprechen für eine Zunahme der Korrelationsstärke zwischen Wortschatzumfang und dem semantischen Wortwissen mit steigendem Alter. Die vergleichsweise hohen Korrelationen von bis zu .85, die Vermeer (2001) bereits für das Kindergartenalter fand und aus denen er folgerte, dass eine Unterscheidung verschiedener Wortschatzdimensionen nicht notwendig und die Messung von Wortschatzumfang ausreichend sei, dürften darauf zurückzuführen sein, dass

---

überdurchschnittlichen Sprachkompetenzen (Viesel-Nordmeyer et al. 2019), aber abnehmend in einem cross-over design (Protopapas et al. 2012).

die von ihm durchgeführten Wortschatztests nicht trennscharf waren (siehe Kap. 3).

Etwas höher als die Korrelation zwischen Wortschatzumfang und semantischem Wortwissen, aber immer noch in einem ähnlichen Rahmen, bewegen sich die Korrelationen zwischen Umfang und relationalem Wortwissen (.37 bis .51). Studien mit Erwachsenen haben deutlich höhere Werte zwischen .7 und .86 ergeben (siehe Schmitt 2014; Tran et al. 2020). Schmitt (2014) betont einen engen Zusammenhang zwischen Umfang und relationalem Wortwissen und hebt hervor, dass bisher nicht klar zu entscheiden ist, ob die Korrelationen beim Wachsen des Wortschatzumfangs stärker oder schwächer werden (2014: 941, siehe Kap. 3). Die beschriebenen Ergebnisse aus EnWoLe weisen für die Unterstufe auf einen stetigen Anstieg der Korrelationsstärken hin. Zwischen t1 und t2 ist dieser Anstieg signifikant, was den Vorschlag von Read (2004, siehe auch Nurweni & Read 1999) unterstützt, dass Wortschatzumfang und -breite bei jüngeren Kindern bzw. beginnenden Sprachlernenden klarer unterscheidbar sind und sich einander mit fortschreitendem Sprachwissen annähern.

### *7.3 Die Dimensionen Wortschatzumfang und Aspekte der Tiefe als klar unterscheidbare Dimensionen (Forschungsfrage 2)*

Die Tatsache, dass die Faktorenanalysen mit einer 1-Faktorlösung eine befriedigende Modellgüte und eine substantielle Varianzaufklärung erzielten, zeigt die Notwendigkeit, alle drei untersuchten Wortschatzbereiche für ein Konstrukt Wortschatz vorauszusetzen. Allerdings ist dabei zu beachten, dass in der vorliegenden Konstellation das relationale Wortwissen nicht klar vom Wortschatzumfang abgegrenzt werden kann. Die Faktorbildung des semantischen Wortwissens in den 2-Faktorenlösungen verdeutlicht zwar, dass bei Kindern im Primarschulalter die Annahme einer getrennten Dimension "Wortschatztiefe" sinnvoll ist, aber nur in Bezug auf das semantische Wortwissen. Das relationale Wortwissen korreliert in der untersuchten Zeitspanne stärker mit dem Wortschatzumfang als mit dem semantischen Wortwissen.

In der Entwicklung des Wortschatzes insgesamt legen semantisches und relationales Wortwissen an Bedeutung zu, ihre Ladungen in der 1-Faktorlösung werden stärker. Die zusätzliche Varianzaufklärung, die in einer 2-Faktorenlösung erzielt wird, ist ganz im Sinne Schmitts (2014), dass die Aussagekraft von Regressionsanalysen des Wortschatzes durch den Einbezug der Wortschatztiefe gewinnt. Die Resultate sind aber immer abhängig von der Wahl der Wortschatztiefe-Tests, die wiederum von der Definition der unterschiedlichen Aspekte der Wortschatztiefe abhängig sind. Wortdefinitionen – so komplex ihre Bewertung sein mag – sind auf der Basis der vorliegenden Studienresultate bewertet – ganz offensichtlich ein probates Mittel zur Messung semantischen Wissens.

#### *7.4 Einfluss persönlicher, sozioökonomischer und kognitiver Merkmale (Forschungsfrage 3)*

Das Alter der Kinder hatte – in Übereinstimmung mit den anderen hier zitierten Studien – zu allen drei Zeitpunkten einen statistisch auffälligen positiven Einfluss auf den Wortschatzumfang. Das relationale Wortwissen (nur zu t1 war ein Einfluss festzustellen) und das semantische Wortwissen hingegen zeigten sich relativ unbeeinflusst. Im Falle des relationalen Wortwissens ist der Grund ein Bodeneffekt zu t1, bei dem die Anzahl unbeantworteter Items bei den jüngsten Kindern am grössten war, während zu t2 und t3 ungeachtet des relativen Alters nur noch wenige Items überhaupt nicht verstanden wurden. Im Falle des semantischen Wortwissens konnte bereits andernorts gezeigt werden, dass Definitionen auf individueller Ebene innerhalb von wenigen Schuljahren keine zielgerichtete Entwicklung darzustellen scheinen, sondern Schwankungen aufweisen (Juska-Bacher et al. 2021).

Ein Einfluss des Geschlechts zeigte sich ausschliesslich beim Wortschatzumfang im ersten und zweiten Schuljahr in Form eines Vorteils für die Jungen, im dritten Schuljahr war der Effekt nicht mehr vorhanden. Ersteres deckt sich mit den Beobachtungen von Lenhard & Lenhard (siehe Beitrag 2021), obwohl dort die Signifikanzgrenze nicht erreicht wurde, zweiteres widerspricht der Konstanz des Unterschieds, den Lenhard & Lenhard feststellen.

Der SES wies eine hohe Streuung auf, die sich in Form eines signifikanten Einflusses auf alle Variablen zu allen Zeitpunkten äusserte. Alle Wortschatzdimensionen sind ganz offensichtlich ein Spiegel der familiären Bildungsnähe (hier ausgedrückt durch den aktuellen Beruf der Eltern) mit all ihren Konsequenzen (siehe Beitrag Haag et al. 2021).

Die nonverbale Intelligenz wiederum zeigte den stärksten Einfluss beim relationalen Wortwissen. Die Kenntnis von Hyper- und Antonymen scheint enger verknüpft mit den exekutiven Funktionen als der Wortschatzumfang. Im Bereich des semantischen Wortwissens dürfte der relativ geringe Zusammenhang zu t1 und die Steigerung danach der Tatsache geschuldet sein, dass jüngere Kinder mehr Probleme damit haben, in dieser dekontextualisierten Form ihr semantisches Wissen abzurufen und präzise zu verbalisieren. Um den Fähigkeiten von Kindern am Schulanfang gerecht zu werden, muss dieses Erhebungsinstrument überprüft und gegebenenfalls durch ein geeigneteres ersetzt werden.

## **8. Das Konstrukt Wortschatz**

Die Resultate dieser Studie weisen darauf hin, dass Wortwissen in den frühen Primarschuljahren als mehrdimensionales Konstrukt mit separaten Massen für Umfang und Tiefe behandelt werden sollte (siehe auch Yanagisawa & Webb

2020). Zumindest der Wortschatzumfang und das semantische Wortwissen liessen sich statistisch eindeutig unterscheiden, wenn auch eine solche klare Unterscheidung zwischen Umfang und relationalem Wortwissen anhand der hier präsentierten Daten nicht möglich war. Eine Beschränkung auf die Erhebung des Wortschatzumfangs, wie Vermeer sie postuliert, reicht damit nicht aus, um Wortschatzkompetenzen in diesem Alter angemessen abzubilden. Sowohl im Forschungs- als auch im Förderungskontext bedarf es insbesondere einer Berücksichtigung verschiedener Aspekte der komplexen und bisher sehr unterschiedlich ausgelegten Dimension Wortschatztiefe. Bisher liegen dafür im Deutschen auch für die hier berücksichtigten zwei Aspekte noch keine spezifischen Testinstrumente vor. Als weiterer Aspekt scheint zudem formales Wortwissen vielversprechend, denn Studien, die dieses einbeziehen (z.B. Proctor et al. 2012 oder für Erwachsene Tran et al. 2020), legen nahe, dass der Einbezug morphologischer Kompetenzen auch bei Kindern das Konstrukt Wortschatz noch besser abbilden dürfte.

Die hier präsentierten Daten der Unterstufe zeigten einen Trend, dass die Wortschatzdimensionen mit zunehmendem Alter näher zusammenrücken. Um diesen Trend zu bestätigen, sind Longitudinaldaten aus höheren Klassenstufen notwendig, die derzeit nicht vorliegen.

Um eine weitere systematische Erforschung wie Förderung von Wortschatzwissen anzuregen, postulieren wir – im Anschluss an Yanagisawa & Webb (2020) – erstens eine intensiviertere Erforschung der Zusammenhänge der Dimensionen und -aspekte über die ganze Schulzeit, zweitens im Wechselspiel mit der Forschung eine Entwicklung standardisierter Tests, insbesondere für das semantische Wortwissen, und drittens schliesslich das Fruchtbarmachen der empirischen Ergebnisse für die Wortschatzdidaktik. Nur durch eine differenzierte Diagnose lässt sich die Voraussetzung für eine gezielte Förderung von Wortschatz als "Schaltstelle des schulischen Spracherwerbs" (Steinhoff 2009) bei erst- wie zweitsprachigen Kinder schaffen.

## Danksagung

Das Projekt "Die Entwicklung von Wortschatz und Wortlesen. Eine Untersuchung auf der Unterstufe" wird vom Schweizerischen Nationalfonds gefördert (<http://p3.snf.ch/Project-173245>). Bedanken möchten wir uns besonders bei den am Projekt beteiligten Kindern und ihren Lehrpersonen für ihre Teilnahme am Projekt. Wir danken ausserdem Dr. Gorka Fraga González und Dr. Christoph Zangger für hilfreiche Anregungen zur statistischen Auswertung sowie den anonymen Gutachtenden für ihre konstruktive Kritik an einer vorangehenden Version dieses Artikels.

## LITERATUR

- Alber, K. (2016). Wortschatzumfang, Wortschatztiefe und Verarbeitungsgeschwindigkeit – Analyse der lexikalischen Kompetenz. *Zeitschrift für Angewandte Linguistik*, 65, 107-128.
- Alber, K. (2021). Wortschatzkompetenz erheben – eine synoptische Zusammenstellung wortschatzdiagnostischer Verfahren. *Bulletin suisse de linguistique appliquée*, 113, 193-210.
- Anderson, R. & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Hg.), *Comprehension and teaching: research reviews* (S. 72-117). Newark, DE: International Reading Association.
- Anglin, J. M. (2008). The acquisition of word meaning II: Later lexical and semantic development. In A. D. Cruse, F. Hundsnurscher, M. Job & P. R. Lutzeier (Hgg.), *Lexikologie. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen* (S. 1789-1800). Berlin & New York: De Gruyter.
- Cain, K. & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'année psychologique/Topics in Cognitive Psychology*, 114, 647-662.
- Castillo, C. & Tolchinsky, L. (2018). The contribution of vocabulary knowledge and semantic orthographic fluency to text quality through elementary school in Catalan. *Reading and Writing*, 31, 293-323.
- Cox Eriksson, C. (2021). Researching vocabulary depth in a multilingual Swedish elementary school. *Bulletin suisse de linguistique appliquée*, 113, 69-87.
- Fabrigar, L., Wegener, D., MacCallum, R. & Strahan, E. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.
- Fitzpatrick, T. (2013). Word associations. In C. A. Chapelle (Hg.), *The encyclopedia of applied linguistics* (S. 6193-6199). Oxford, UK: Wiley-Blackwell.
- Ganzeboom, H. G. B., De Graaf, P. M. & Treiman, D. J. (1992). A standard international socio-economic index of occupational status. *Social Science Research*, 21, 1-56.
- Glück, C. W. (2011). *Wortschatz- und Wortfindungstest für 6- bis 10-Jährige (WWT 6–10)*. München: Urban & Fischer.
- Glück, C. W. (2021). Der WWT 6-10 und seine Testitems. *Bulletin suisse de linguistique appliquée*, 113, 29-47.
- Gyllstad, H. (2013). Looking at L2 vocabulary knowledge dimensions from an assessment perspective – challenges and potential solutions. In C. Bardel, B. Laufer & C. Lindqvist (Hgg.), *L2 Vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (S. 11-28). Eurosla Monographs Series: Eurosla.
- Haag, L. N., Di Pietro, S. V., Röthlisberger, M. Wick, C., Füzér, R. & Brem, S. (2021). Impact of parental reading skills and children's reading environment and interest at pre-school age on the vocabulary and reading skills in elementary school. *Bulletin suisse de linguistique appliquée*, 113, 89-116.
- Henriksen, B. (2008). Declarative lexical knowledge. In D. Albrechtsen, K. Haastrup & B. Henriksen (Hgg.), *Vocabulary and writing in a first and second language* (S. 22-66). Basingstoke, UK: Palgrave Macmillan.
- Jurecka, A., Cinar, M. & Hardy, I. (2019). Messung von Wortschatztiefe und -breite bei mono- und bilingualen Vorschulkindern. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 24(1), 145-163.
- Juska-Bacher, B., Beckert, C., Stalder, U. & Schneider, H. (2016). Die Bedeutung des Wortschatzes für basale Lesekompetenzen. *Didaktik Deutsch*, 40, 20-39.
- Juska-Bacher, B., Brugger, L., Korthus, R. & Zangger, C. (2021). Definitionskompetenzen von Erst- und Zweitklässlern. Mit einem Ausblick auf die weitere Entwicklung. *Muttersprache*, 2, 117-137.

- Juska-Bacher, B. & Jakob, S. (2014). Wortschatzumfang und Wortschatzqualität und ihre Bedeutung im fortgesetzten Spracherwerb. *Zeitschrift für Angewandte Linguistik*, 61(1), 49-75.
- Juska-Bacher, B., Röthlisberger, M., Brugger, L. & Zangger, C. (2021a). Lesen im 1. Schuljahr: Die Bedeutung von Vorläuferkompetenzen und Wortschatz. In S. Gailberger & C. Sappok (Hgg.), *Weiterführende Grundlagenforschung in der empirischen Leseforschung und Lesedidaktik. Theorie – Empirie – Didaktik*, (S. 11-26). SLLD-B. <https://omp.ub.rub.de/index.php/SLLD/catalog/book/189>
- Karlsen, J., Halaas Lyster, S.-A. & Lervåg, A. (2017). Vocabulary development in Norwegian L1 and L2 learners in the kindergarten–school transition. *Journal of Child Language*, 44, 402-426.
- Lange, B. P., Euler, H. A. & Zaretsky, E. (2016). Sex differences in language competence of 3- to 6-year-old children. *Applied Psycholinguistics*, 37, 1417-1438.
- Lawrence, J. F., Hagen, A. M., Kyoung Hwang, J., Lin, G & Lervåg, A. (2019). Academic vocabulary and reading comprehension: Exploring the relationships across measures of vocabulary knowledge. *Reading and Writing*, 32, 285-306.
- Lenhard, W. & Lenhard, A. (2021). Bedeutung und Diagnostik des Wortschatzes am Beispiel des Peabody Picture Vocabulary Test (PPVT-IV). *Bulletin suisse de linguistique appliquée*, 113, 9-27.
- Lenhard, A., Lenhard, W., Segerer, R. & Suggate, S. (2015). *Peabody Picture Vocabulary Test – Revision IV* (Deutsche Adaption). Frankfurt am Main: Pearson Assessment.
- Meara, P. & Wolter, B. (2004). V\_LINKS: Beyond vocabulary depth. In D. Albrechtsen, K. Hastrup, & B. Henriksen (Hgg.), *Angles on the English-speaking world 4* (S. 85-96). Copenhagen, Denmark: Museum Tusulanum Press.
- Moser, U., Stamm, M. & Hollenweger, J. (2005). *Für die Schule bereit? Lesen, Wortschatz, und soziale Kompetenzen bei Schuleintritt*. Oberentfelden: Sauerländer.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Nurweni, A. & Read, J. (1999). The English vocabulary of Indonesian university students. *English for Specific Purposes*, 18, 161-175.
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98(3), 554-566.
- Ouellette, G. P. & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, 23, 189-208.
- Petermann, F. & Petermann, U. (2007). *Hamburg-Wechsler-Intelligenztest für Kinder – IV*. Übersetzung und Adaptation des WISC-IV® von David Wechsler. Bern: Huber.
- Proctor, C. P., Silverman, R. D., Harring, J. R. & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish–English bilingual children in elementary school. *Reading and writing*, 25, 1635-1664.
- Protopapas, A., Simos, P.G., Sideridis, G.D. & Mouzaki, A. (2012). The components of the simple view of reading: A confirmatory factor analysis. *Reading Psychology*, 33(3), 217-240.
- Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (Hgg.), *Vocabulary in a second language* (S. 209-227). Amsterdam: John Benjamins.
- Rindermann, H. & Baumeister, A. E. E. (2015). Parent's SES vs. parental behaviour and children's development: A reanalysis of the Hart and Risley study. *Learning and Individual Differences*, 37, 133-138.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Houndmills, Basingstoke, Hampshire, New York, NY: Palgrave Macmillan.

- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913-951.
- Schoonen, R. & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing*, 25(2), 211-236.
- Steinhoff, T. (2009). Der Wortschatz als Schaltstelle des schulischen Spracherwerbs. *Didaktik Deutsch*, 27, 33-52.
- Strasser, K. & del Río, F. (2013). The role of comprehension monitoring, theory of mind, and vocabulary depth in predicting story comprehension and recall of kindergarten children. *Reading Research Quarterly*, 49(2), 169-187.
- Swart, N. M., Muijselaar, M. M. L., Steenbeek-Planting, E. G., Droop, M., De Jong, P. F. & Verhoeven, L. (2017). Differential lexical predictors of reading comprehension in fourth graders. *Reading and Writing*, 30, 489-507.
- Tannenbaum, K. R., Torgesen, J. K. & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading*, 10(4), 381-398.
- Tran, A. H., Tremblay, K. A & Binder, K. S. (2020). The factor structure of vocabulary: An investigation of breadth and depth of adults with low literacy skills. *Journal of Psycholinguistic Research*, 49, 335-350.
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, 22, 217-234.
- Viesel-Nordmeyer, N., Schurig, M. & Ritterfeld, U. (2019). Auswirkungen vorschulischer sprachlicher Disparitäten auf die sprachliche und mathematische Kompetenzentwicklung im Grundschulalter. *Empirische Sonderpädagogik*, 11(4), 294-309.
- Weiß, R. H. (2006). *CFT 20-R: Grundintelligenztest Skala 2-Revision*. Göttingen: Hogrefe.
- Wiig, E.H. & Secord, W. (1992). *TOWK: Test of word knowledge*. Toronto, Ontario, Canada: Psychological Corporation.
- Willenberg, H. (2008). Wortschatz Deutsch. In Klieme, E. (Hg.), *Unterricht und Kompetenzerwerb in Deutsch und Englisch. Ergebnisse der DESI-Studie* (S. 72-80). Weinheim u.a.: Beltz.
- Yanagisawa, A. & Webb, S. (2020). Measuring depth of vocabulary knowledge. In S. Webb (Hg.), *The Routledge handbook of vocabulary research* (S. 371-386). Oxon/New York: Routledge.



# Researching vocabulary depth in a multilingual Swedish elementary school

**Christine COX ERIKSSON**

Dalarna University  
SE-791 88 Falun  
cce@du.se

Ziel dieses Beitrags ist es, Möglichkeiten und Herausforderungen bei der Untersuchung von Wortschatztiefe im Kontext mehrsprachiger Primarschulen aufzuzeigen. Im theoretischen Teil wird ein Überblick über die Messung von Wortschatztiefe anhand von Wortdefinitionen und Assoziationsaufgaben in der internationalen und schwedischen Forschung gegeben. Der empirische Teil beschreibt Datenerhebungen mit einem neu entwickelten Definitionstest und dem Kent-Rosanoff-Assoziationstest mit 92 Zweit- bzw. Fünftklässlern an einer schwedischen Schule mit hohem Anteil an Zweitsprachlern. Thematisiert wird die Bewertung der Definitionen von Schülerinnen und Schülern sowie damit zusammenhängende Herausforderungen. Die Ergebnisse, die nach Alter, Aufenthaltsdauer in Schweden und Sprachhintergrund analysiert wurden, werden im Hinblick auf die bestehende Forschung und methodische Überlegungen diskutiert. Angesprochen werden der Mangel an Forschung und verfügbaren Testinstrumenten in Bezug auf Wortdefinitionen sowie an geeigneten Worthäufigkeitslisten, die auf mündlicher Sprache in einem Grundschulkontext basieren. Das in diesem Artikel beschriebene Vorgehen und die Herausforderungen sowie die Ergebnisse stellen einen Beitrag zu unserem Wissen über die Wortschatztiefe bei mehrsprachigen Schülerinnen und Schülern dar und sind für Forscher und Pädagogen gleichermaßen von Interesse.

## **Stichwörter:**

Wortschatztiefe, Erstspracherwerb, Zweitspracherwerb, Wortdefinitionen, Wortassoziationen, Primarschule, Mehrsprachigkeit.

## **Keywords:**

vocabulary depth, L1 acquisition, L2 acquisition, word definitions, word associations, elementary school, multilingualism.

## **1. Introduction**

This article focuses on issues related to investigating vocabulary knowledge in Swedish, specifically lexical depth, in an elementary school where approximately 75% of pupils have a migration background and only 30% of parents have post-secondary education. Understanding how words are stored in the mental lexicon is a complex undertaking, even within the context of first language acquisition. Vocabulary depth, a measure of how well words are known, has alternatively been referred to as lexical quality (Anderson & Freebody 1981) or lexical organization (Meara 1996). According to Meara, the vocabularies of second language (L2) speakers may not be as well structured as those of first language (L1) speakers. Thus, organization may be a way to distinguish learners at different proficiency levels. Meara views organization as a characteristic of a learner's entire vocabulary, not of individual words. How vocabulary depth is measured, as well as empirical results, depend on how the construct is conceptualized (Schmitt 2014).

Receptive word knowledge is easier to measure and show, for example through picture recognition. Productive knowledge involves at least some mastery of a word's spoken or written form, as well as meaning and use (Nation 2013). Vocabulary production and grammatical ability in children approximately 1 to 4 years of age can be measured using the Communicative Development Inventories (CDI; SECDI in Sweden; e.g. Cox Eriksson 2014), which represent parents' records of their children's early language development. For older children, however, other tools must be used to describe vocabulary development in terms of lexical quality.

### 1.1 Aim

The present article aims to discuss methodological choices made and challenges experienced in assessing pupils' Swedish vocabulary depth in a multilingual elementary school. The research project from which these experiences are taken aimed to describe and attempt to explain the variation in vocabulary development in pupils attending a multilingual school and, together with the teachers, devise new classroom approaches that facilitate vocabulary learning, regardless of the pupils' language backgrounds. Participants include 92 pupils in grades 2 and 5 (mean ages 8.4 and 11.4 years), with at least 15 different first languages. Choices discussed in this article include deciding how to measure vocabulary depth and what instruments to use. In particular, the article will explore assessing vocabulary depth through a word definition task and a word association task.

## 2. Background

This section presents an overview of international research on investigating vocabulary depth in both L1 and L2 language contexts, with adults and younger learners.

### 2.1 Conceptualizing and measuring vocabulary depth

Nation's (2013) well-known conceptualization of what it means to know a word breaks down lexical knowledge into three main categories: *form* (spoken and written forms, as well as word parts), *meaning* (concepts, referents and associations) and *use* (grammatical functions, collocations and constraints on use). Each category includes receptive as well as productive aspects. Using most of Nation's terms, Figure 1 illustrates a learner's knowledge of a specific word after a certain number of exposures, with the various aspects being mastered at different rates. In studies with multilingual individuals, it is especially relevant to consider that there may be varied mastery of word knowledge.

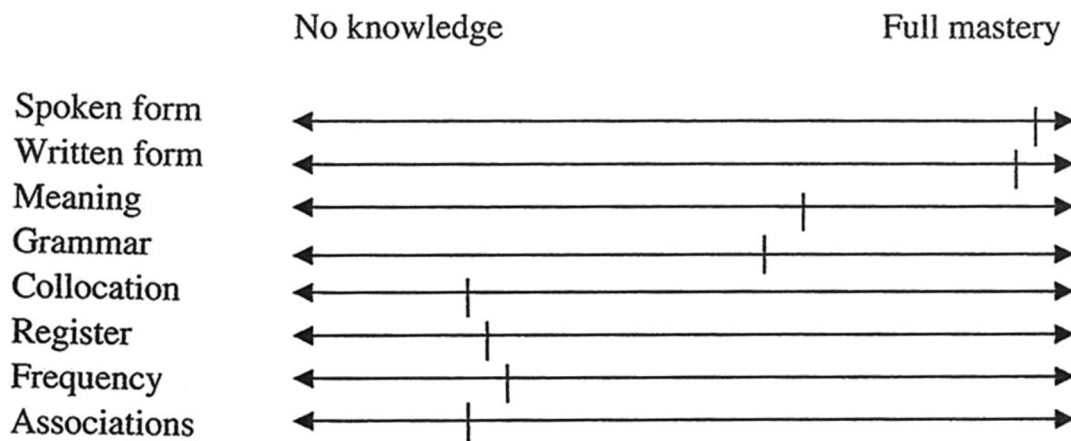


Figure 1: Developing knowledge of a word (Schmitt 2010: 38)

Meara and Wolter (2004: 87) have criticized the above conceptualization of vocabulary as components because it regards depth as knowledge of individual words, rather than a "test taker's entire vocabulary". Meara and Wolter compare two different ways of looking at vocabulary in Figure 2.

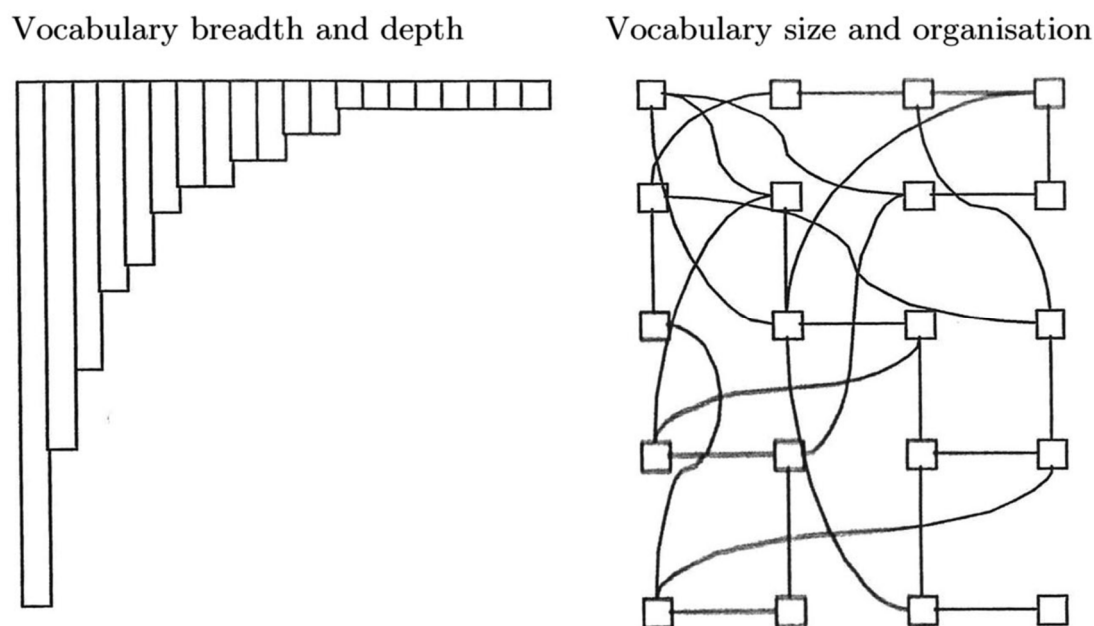


Figure 2: Two ways of looking at vocabulary (Meara & Wolter 2004: 89)

The image on the left shows that items learned earlier have more depth, as opposed to newly acquired words. However, this would imply that words are not related to one another. Therefore, Meara and Wolter provide the image on the right as a better model of vocabulary depth, illustrating how words are stored in

the mental lexicon as a network, with more links between individual words resulting in greater lexical organization.

Read (2004) outlines three different approaches that have been applied in research on L2 vocabulary acquisition. The first is *precision of meaning*, which implies the difference between a vague understanding of what a word means and more specific knowledge of that meaning. Secondly, there is *comprehensive word knowledge*, which involves many different aspects of a word, including orthographic, phonological, syntactical, collocational, and pragmatic attributes. This is comparable to Nation's component model described above. Thirdly, Read lists *network knowledge*, which means the incorporation of a word into the mental lexicon, where words are linked together in different ways. Read points out that any kind of vocabulary test will only reflect what a learner knows about the target words at the point the test is taken. Moreover, vocabulary tests often assess learners' declarative knowledge, which is descriptive and easier to verbalize, as opposed to procedural knowledge. The latter is more implicit and associated with word recognition or competent listening comprehension.

Precision of meaning has been exemplified by describing the progression in word learning in categories or stages. One classic scale is Dale's (1965: 898) four-stage description of how well a word is known, originally created to be used with elementary age L1 learners:

Stage 1: "I never saw it before"; Stage 2: "I've heard of it, but I don't know what it means";  
Stage 3: "I recognize it in context-it has something to do with ..."; Stage 4: "I know it".

This is still a useful tool in developing awareness of word meanings among young learners.

## 2.2 Word definition tasks

Precision of knowledge can be investigated through definition tasks, often used in assessing children's L1 or L2 vocabulary knowledge. Benelli et al. (1988) conducted three studies to investigate the development of word definitions among children (5- and 7-year-olds) and adults. The task included defining nine common nouns that fall in three categories. Responses were classified on a scale ranging from perceptual appearance and functional definitions to different types of categorical definitions. The latter included either definitions with superordinates alone, those with some kind of specification added to the superordinate, or generic categorical definitions (i.e. 'thing' or 'object'). The youngest children gave the fewest definitions containing superordinates, the 7-year-olds used more, and the adults used most. A new group of adults judged the best definitions to be those including both categorical items and specific information about the target word. It was also found that by the time children are 10 years old, their definitions resemble adult models.

Kurland & Snow (1997) assessed the definitional skill of children in low-income families over time (from 5;3 to 10;10<sup>1</sup>) and compared children's definitions to those of their mothers. The children were asked to define eight nouns and answers were assessed on a 7-point scale for definitional quality, including various kinds of clauses and descriptors. Like Benelli et al., Kurland & Snow found that 9 to 10-year-olds had also reached "adult levels of definitional skill", with approximately two thirds of the children giving better definitions than their low-income mothers. The authors conclude that "definitional skill is related to being part of an academic culture" (Kurland & Snow 1997: 603).

In a study aiming to investigate the role of oral vocabulary in reading skills (Ouellette 2006), 4<sup>th</sup> grade pupils were asked to define 32 words, including nouns, verbs and adjectives. Answers to the definition test were scored on a 0–3-point scale, based on the number of semantic features given. Other definition tasks for younger children (preK-1<sup>st</sup> grade) have been scored along a continuum, with fewer points for contextual responses and more for decontextualized responses (e.g. Coyne et al. 2009). More recently, Juska-Bacher et al. (2021) analyzed functional, descriptive and categorical word definitions in both children and adults on a scale from contextualized to decontextualized.

Two studies by Vermeer (2001) with L1 and L2 Dutch-speaking children illustrate different ways of conceptualizing how to assess vocabulary depth. The first tested breadth with a picture vocabulary test and a description task, where kindergarten children were asked to explain or describe the meaning of 27 words. Depth was measured through an association task where children provided formal definitions of 10 words, external features and component parts, functional and material characteristics, as well as use. Children's responses were described qualitatively using an association network for each word. L1 children received higher scores than L2 children, who were able to provide characteristics, but lacked the verbal ability to describe them. The second study included a picture vocabulary test (receptive breadth) and a task asking for descriptions, characteristics or definitions of words. Results indicated that the input frequency of the words strongly correlated with the probability of knowing a word, especially in the description task. Vermeer's operationalization of vocabulary breadth and depth with young children exemplifies difficulties in comparing results across studies (Schmitt 2014).

In other research from the Netherlands, Verhallen & Schoonen (1993, 1998) compared lexical knowledge in L1 and L2 elementary school pupils using an extended word definition task. In a structured interview situation, bilingual Turkish-Dutch 3<sup>rd</sup> and 5<sup>th</sup> graders all born in the Netherlands were asked to give

---

<sup>1</sup> Within the field of child language research, age is often denoted for example as 5;3 (five years and three months).

as many meaning aspects as possible to six stimulus words. Results indicated that the pupils were better able to explain the words in Dutch than in Turkish. Comparisons with L1 pupils showed that the L2 pupils gave fewer meaning aspects as well as less well-developed answers, which included fewer paradigmatic meaning aspects. The relationship between paradigmatic (same-class, as in *sit-stand*) and syntagmatic (different class, *sit-chair*) word knowledge was similarly investigated in an American study (Ordóñez et al. 2002), where bilingual (Spanish-English) 4<sup>th</sup> and 5<sup>th</sup> graders provided superordinates, definitions and rich object descriptions of six familiar concrete nouns. Results indicated that pupils' ability to produce paradigmatic responses in one language often transferred to the other.

### 2.3 Word association tasks

The classic Kent-Rosanoff word association test (Kent & Rosanoff 1910) assesses the ability to make hierarchical associations, essentially measuring network knowledge (Read 2004). The original list is comprised of 100 nouns and adjectives and has been used in studies with both children and adults. The test involves presenting participants with a single stimulus word, such as *cat*, and asking them to give the first word they associate with that word. Scholars have systematized responses into categories, including clang associations, based on phonological similarity to the stimulus word (*hat-cat*), syntagmatic associations, based on semantic principles (*cat-meows*), and paradigmatic associations (*cat-animal*), often denoting the hierarchical category in which the stimulus word belongs. The so-called syntagmatic-paradigmatic shift has been found to occur at some point between 6 and 10 years of age and is an indication of increased lexical knowledge with the ability to form paradigmatic associations (Nelson 1977). According to Nelson (1977), the shift may not only indicate conceptual reorganization, but also task capacity. One study with young first and second language learners has indicated that the shift to more paradigmatic associations may occur later in multilingual children (Schoonen & Verhallen 2008).

Read (1993) created the *Word Associates Format* (WAF), where responses to a stimulus word are selected, rather than supplied. This was thought to be a more practical option for L2 university learners than the classic Kent-Rosanoff test. In this test, a target word (for example *contract* is followed by six or eight other words, which may or may not be related to the stimulus word (i.e., *agreement, confident, formal, notice, sign, special*). Answers can be scored with respect to the type of relationship between the target word and associated words, e.g., *agreement* (paradigmatic; synonym), *sign* (syntagmatic; collocate) and *formal* (analytic; word which represents a key element in the meaning of the target word). Research has generally shown that scores on tests using WAF are predicative of skills such as reading comprehension (Fitzpatrick & Thwaite 2020). Schoonen & Verhallen (2008) adapted Read's WAF for use with upper elementary pupils (grades 3-6) and developed the Word Association Task

(WAT). Pupils choose three words out of six that best go together with the stimulus word and draw connecting lines (see Figure 3 for correct answers). *Monkey* is a semantically less related word and therefore, a less decontextualized answer.

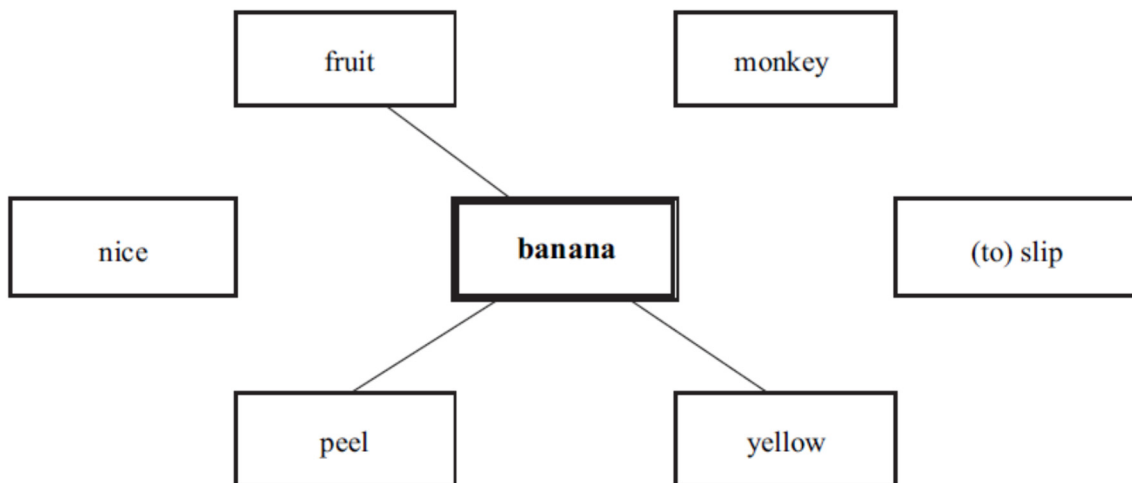


Figure 3: Sample item from the WAT (translated from Dutch; Schoonen & Verhallen 2008: 219)

In a recent review article, Fitzpatrick & Thwaite (2020) highlight several methodological challenges involved in comparing results of word association research, including specific definitions of the paradigmatic/syntagmatic/clang categories, as well as the number of cue words used and elicited responses per cue. They also caution against using word association results to accurately reflect the organization of the mental lexicon, as no single method can be used to do this.

### 3. Research in the Swedish context

Many Swedish studies on children's vocabulary development are from the field of speech therapy. Andersson et al. (2019) investigated the role of multilingualism and socioeconomic status in Swedish L1 and L2 first and second graders' results on the Clinical Evaluation of Language Fundamentals (CELF4; Semel et al. 2013). A composite of social deprivation measures was found to account for a much greater proportion of variance than the children's multilingualism. As one of few Swedish studies on word definitions, Erlandsson & Yhlen (2019) investigated Swedish L1 and L2 8-year-olds' definitions of 10 words common in a school context. They followed the scoring principle (0-3 points) used by McGregor et al. (2012): 0 points for no correct information; 1 point for some meaningful relationship to the target word; 2 points for minimal definitions; 3 points for more than minimal accurate information. Erlandsson & Yhlen found that while L1 pupils' definitions gave more information than their L2 peers, participants rarely received full scores for their definitions. There is a fair

amount of research on grammatical development in Sweden, for example work by Gisela Håkansson. She has recently been involved in a project using the new assessment tool Cross-linguistic Lexical Tasks (LITMUS-CLT; Haman et al. 2017), which explores young learners' productive and receptive knowledge of nouns and verbs in multiple European languages. A recent doctoral thesis investigating oral narrative competence in L1 and L2 Swedish 4-6-year-olds (Lindgren 2018) included vocabulary production scores using CLTs.

Namei (2002), who translated the original 100-word Kent-Rosanoff list into Persian and Swedish, compared the L2 and L1 mental lexicon in terms of syntagmatic-paradigmatic development and similar/dissimilar organization in 100 bilinguals aged 6 to 22 and 50 L1 speakers of both languages. Clang responses were most common with unknown or less frequent words, syntagmatic responses were given when words were partially known, and paradigmatic responses were given when words were highly frequent or well-integrated in the lexicon. Mikoczy & Nyman (2008) used Namei's Swedish translation with bilingual Swedish-Arabic 4<sup>th</sup> graders, calling for a shorter version of the list. Johansson & Wahlstrand (2010) shortened the list to 50 words and provided an extended assessment guide. The shortened version, also translated into Arabic, has been used in many studies, including one comparing lexical development in bilingual 4<sup>th</sup> graders who had received bilingual instruction as opposed to those who only received instruction in Swedish (Salameh 2011). A significantly higher proportion of pupils receiving bilingual instruction showed hierarchical lexical organization in both languages, but this was not correlated to lexical size. Nilsson & Svenbe (2017) used the 50-word Swedish list, comparing lexical organization and size in three groups of teenagers and young adults. Results indicated that paradigmatic responses dominated in all groups, followed by syntagmatic responses.

Other Swedish studies have used the Kent-Rosanoff test to compare lexical organization in monolingual and bilingual children with and without developmental language impairment. Holmström et al. (2016b) used conceptual scoring in the word association task for bilinguals, where knowledge of words in both languages was combined. Bilingual children with language impairment received significantly higher paradigmatic conceptual scores compared to scores based on one of their languages alone. Their scores were also higher than those of monolingual children with language impairment. However, when both productive and receptive measures were included, smaller vocabulary sizes were evident in the bilinguals with language impairment as opposed to typically developing bilinguals (Holmström 2015). Holmström et al. (2016a) point out that the Kent-Rosanoff list was not originally constructed to study lexical organization, calling for further development of a word association list considering word frequency, part of speech and mean age of acquisition of words. A more recent study (Sandgren et al. 2020) also aimed to investigate vocabulary breadth, depth, and fluency, and used a semantic depth score in



comparing word associations in children aged 6 to 9 with developmental language disorder (DLD) to those with typical development. To calculate the semantic depth score, paradigmatic (2 points each) and syntagmatic (1 point) associations were summed. The typically developing group had significantly higher semantic depth scores than the DLD group.

In summary, this literature review has shown that vocabulary depth has been conceptualized and operationalized in different ways. This, along with great variety in the scoring of definition and association tasks, makes comparisons across studies a challenge. However, the variety of tools available do allow researchers to tailor assessment instruments to a specific student population. As the current study included pupils with varying reading ability in Swedish, we wanted to make sure that we were assessing vocabulary knowledge and not reading ability. Thus, we chose two test instruments that were administered orally, even if the WAT has already been used internationally with young pupils.

#### 4. Measuring depth in the current study

Data for the current project was collected from November 2018 to early April 2019 at a linguistically diverse elementary school in central Sweden. Participants were 92 pupils (47 girls; 45 boys) in grades 2 ( $n = 54$ ; mean age 8.4 years) and 5 ( $n = 38$ ; mean age 11.4 years). Pupils were diverse in terms of how long they had lived in Sweden as well as whether they had Swedish as L1 or L2 (see Table 1).

Grade	Years in Sweden				Total
	1–4 (SL2)	5–7 (SL2)	Born in Sweden (SL2)	Born in Sweden (SL1)	
2	14	11	18	11	54
5	7	8	13	10	38
<b>Total</b>	<b>21</b>	<b>19</b>	<b>31</b>	<b>21</b>	<b>92</b>

Table 1: Years in Sweden and language background for all participants ( $N = 92$ ) in grades 2 and 5.

##### 4.1 Word definitions

Due to the lack of available normed instruments for measuring vocabulary depth in Swedish, a newly constructed word definition task (Edquist 2021) was developed for use in the current project. Following Edquist's previous trials, items include words of varying frequency and difficulty level, as well as a mix of everyday and academic words. The current project used a list of 12 items comprising four verbs (*applådera* [applaud], *inbilla sig* [imagine], *fantisera* [fantasize] and *prioritera* [prioritize]), five nouns (*lön* [salary], *ved* [firewood], *fe*

[fairy], *kapitel* [chapter] and *dygn* [one day and night/24 hours]) and three adjectives (*generös* [generous], *fridlyst* [protected, as in a protected species] and *stolt* [proud]).

The definition task attempted to assess pupils' ability to spontaneously define words with only a limited amount of prompting. After the original question *What does xxx mean?*, additional prompting was limited to questions such as: *Can you say more/add anything/explain?* or *What do you mean?* The pupils were told to imagine that the researcher came from another planet and did not know the meaning of many words. The goal of this test was to obtain some measure of depth as "precision of meaning" (Read 2004) or the "extent of semantic representation" (Ouellette 2006). Inspiration was also taken from categorizations of definitions outlined in several previous studies (Benelli et al. 1988; Kurland & Snow 1997). Like Ouellette (2006), our study included words from different parts of speech (although fewer words) and used a 3-point scoring scale, in contrast to studies using only concrete nouns and employing different scoring scales (Benelli et al. 1988; Kurland & Snow 1997; Ordóñez et al. 2002). Our 3-point scale (see below) is somewhat more specific than that used by McGregor et al. (2012) and Erlandsson & Yhlen (2019).

#### 4.1.1 Scoring

Pupils' answers were scored on a scale of 0-3 points, depending on how much and what kind of information was provided. A new scoring rubric, which was continually revised, varied slightly for the different parts of speech. For example, for the noun *fe* [fairy], one point was awarded for a description (*a fairy is a fairy tale character*), one point for an example (*the tooth fairy*), and one point for a characteristic (*a fairy can fly*).

To further illustrate scoring, answers (author's translation) are shown for a verb (*applådera*), a noun (*ved*) and an adjective (*generös*):

*Applådera* (3 points): *When you dance or sing on the stage (example), if people think it is good (purpose - why) they do it, they clap their hands (description).*

(2<sup>nd</sup> grader, Tigrinya speaker, 2 years in Sweden)

*Ved* (2 points): *It's wood (description), thick, that you can heat with (purpose).*

(5<sup>th</sup> grader, Kurmanji speaker, born in Sweden)

*Generös* (0 points): *If someone is good-looking, or? Generous is like if someone is nice, courageous and intelligent, and ehm what else, smart, funny.*

(2<sup>nd</sup> grader, Kurmanji speaker, born in Sweden)

One of the many challenges involved in the scoring process regarded how to systematize pupils' answers in the rubric and decide what constituted a full 3-point answer. For *dygn* [one day and night], one point each was awarded for using the word in an example (*it's like around the clock*) or providing a description (*24 hours*). In the end, we decided that providing both *24 hours* and *one day and night* constituted a full understanding of the word and was worth 3 points, even if this could only be considered two elements. We went through a similar process with the word *lön* [salary]. After deciding to award 1 point for a description (*it's money*), 1 point for an example (*maybe you work as a teacher*)

and 1 point for the purpose, or reason why (*when you work/from your job*), we decided that a relatively concise response (*you get money because you work*) could also be considered a 3-point answer. Wordy answers were not always worth more points. However, some talkative 5<sup>th</sup> graders gave very thorough answers with more examples or characteristics than were needed for the 3-point criteria. Other less talkative pupils (in both grades) gave shorter answers, even when the researcher asked follow-up questions. Individual personality traits may thus affect the answers given by participants, as well as factors such as language background, proficiency level, years in Sweden, and reasonably, socioeconomic status.

#### 4.1.2 Results: Word definitions

Figure 4 shows the mean number of points awarded for the 12 items in the definition test, with information on frequency and difficulty levels in the legend (see below). Scores for 2<sup>nd</sup> graders and 5<sup>th</sup> graders are shown separately. Although there was great individual variation among the groups, 5<sup>th</sup> graders' definition scores were significantly higher ( $p = .000$ ; two-tailed;  $t = -7.941$ ) than those of the 2<sup>nd</sup> graders.

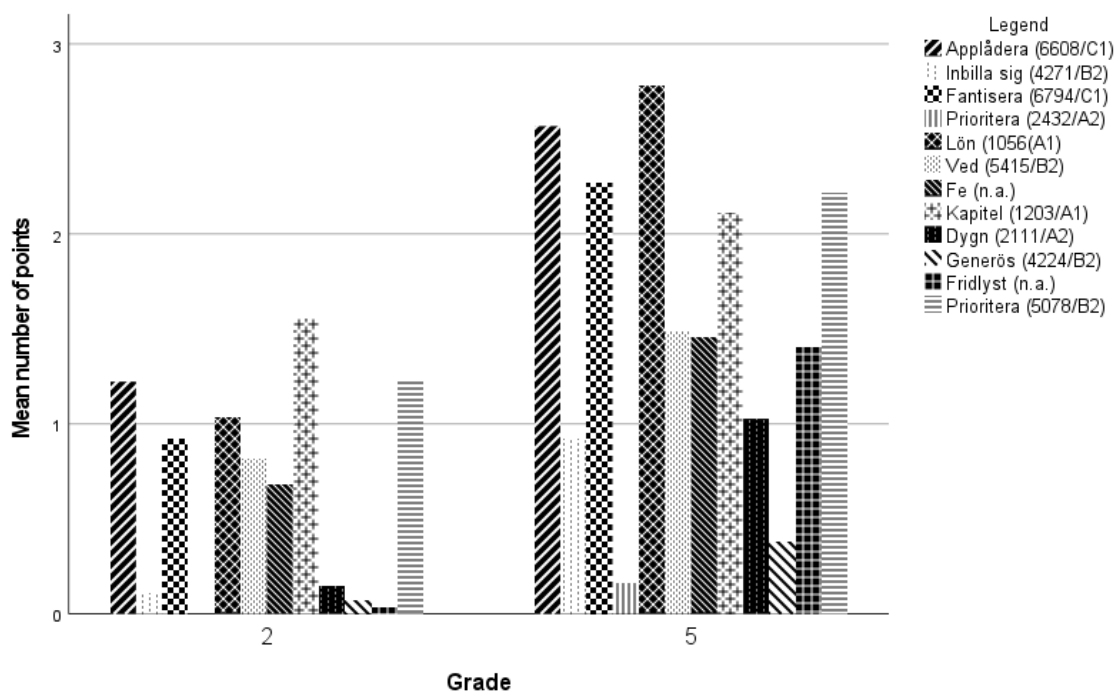


Figure 4: Mean number of points awarded for the 12 items in the definition test for all pupils grouped according to grade level. Frequency statistics and difficulty level are given for each word (KELLY/CEFR; the Swedish KELLY list and the Council of Europe's difficulty list are explained below).

As previous research shows that children's definitional ability approaches adult levels by approximately age 10 (Benelli et al. 1988; Kurland & Snow 1997), this is a reasonable result. In our study, the same words were easiest to define for both groups (*applådera*, *fantisera*, *lön*, *kapitel* and *stolt*), although to varying

degrees for younger and older pupils. The words that were most difficult to define for both ages were *inbilla sig*, *prioritera*, *generös*, *dygn* and *fridlyst*. None of the 2<sup>nd</sup> graders, and only 11% of the 5<sup>th</sup> graders could define *prioritera*. *Generös* was difficult for both 2<sup>nd</sup> and 5<sup>th</sup> graders (with 6% and 24%, respectively, receiving points at all). Only one 2<sup>nd</sup> grader could give a partial definition for *fridlyst*, while almost half of the 5<sup>th</sup> graders received a full score (16 of 37). Of the 19 5<sup>th</sup> graders who could not define the word, 17 were L2 Swedish speakers.

In the evaluation of results, comparisons were made with the Swedish frequency-based vocabulary list, KELLY (KEYwords for Language Learning for Young and adults alike; Kilgarrieff et al. 2014), which is based on a web-acquired corpus (SweWAC<sup>2</sup>) of 114 million words and incorporates the Common European Framework of Reference for Languages (CEFR; Council of Europe 2001) levels of difficulty. Frequency statistics for both the Swedish KELLY and CEFR levels are based on web texts aimed at L1 speakers. Despite efforts to create frequency-based word lists for Swedish L2 speakers, none of these are appropriate for research with children. A few comparisons between word frequency and our results illustrate the need for suitable lists. For example, the word *applådera* ranks 6608 in difficulty on the KELLY list and is a C1 word according to CEFR. However, it is commonly used in oral language, as well as in a school context. Therefore, it was one of the easier words for the pupils to define. *Prioritera*, on the other hand, ranks 2432 on the KELLY list and is an A2 word in accordance with CEFR. However, none of the second graders and only three of the fifth graders could provide any aspect of a definition. The verbs *inbilla sig* and *fantisera* are semantically related, but more different in Swedish than what they may seem to be in English. *Fantisera* is ranked 6794 in the KELLY list and is a C1 word according to CEFR. However, it is more common in a school context, especially regarding story writing. Pupils are encouraged to think up stories and use their imagination, which is *fantasi* in Swedish. *Inbilla sig*, on the other hand, is ranked somewhat more frequent and easier than *fantasize* (4271 KELLY /B2 CEFR). Nevertheless, the meaning is less related to children's everyday classroom experience, as it refers to something imagined in the sense of "to delude oneself". Our results underscore the need for more appropriate frequency lists, as previous research indicates that input frequency correlates with children's word knowledge (e.g., Vermeer 2001), but also that word frequency itself may vary in different contexts (Nation 2013; Schmitt 2014).

Figure 5 contrasts definition results for Swedish L1 speakers with those for pupils with Swedish as L2. Group means are higher at each grade level for Swedish L1 pupils (14.9 compared to 6.1 for grade 2; 24.3 compared to 16.7 for grade 5). In grade 2, the easiest words for L2 pupils to define were *kapitel* and

<sup>2</sup> The Swedish Web Acquired Corpus (SweWAC) was the main corpus, consisting of texts from the internet, which was used in the construction of the Swedish KELLY list (Kilgarrieff et al. 2014).

*stolt*, the only two words with means greater than 1 point. Apart from *prioritera*, Swedish L1 pupils in grade 2 received more points than L2 pupils for all words.

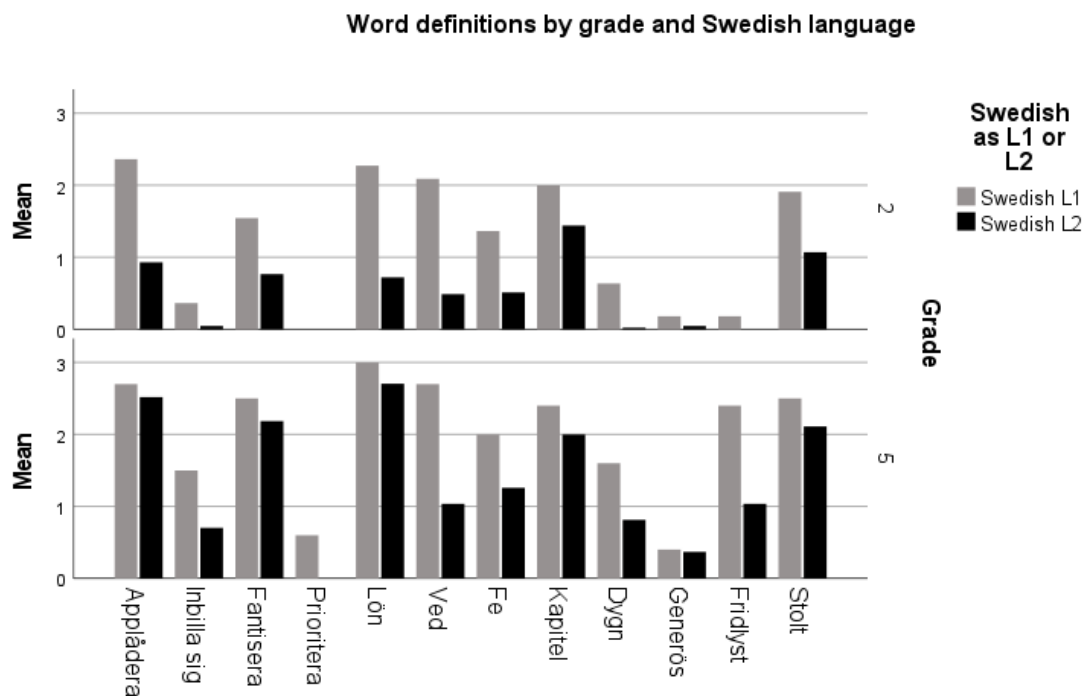


Figure 5: Mean number of word definition points according to grade and Swedish L1/L2

In grade 5 the Swedish L1 pupils received on average more than 2 points for eight of the 12 items, whereas those with other language backgrounds only had means above 2 points for five words: *applådera*, *fantisera*, *lön*, *kapitel* and *stolt*. Words such as *ved* and *fe* may have been culturally very different for the pupils. Although frequency statistics are missing for the word *fridlyst*, it represents a subject-specific word and was much more difficult for Swedish L2 5<sup>th</sup> graders.

#### 4.2 Kent-Rosanoff word association test

In our study, we used the 50-word Kent-Rosanoff (K-R) list in Swedish only as a measure of vocabulary depth. Responses were coded as clang, syntagmatic, paradigmatic, other or no answer. Two groups of pupils were tested by a student, and three groups were tested by the present author. Administration and coding of the test was thoroughly discussed with the student together with a project colleague. As part of the student's essay (Norman 2019) inter-rater reliability was calculated between three independent scorers of the 50 responses from six pupils. As we found our score of 87% to be less than adequate, lengthy discussions involving scoring guides from two different sources (Nilsson & Svenbe 2017; Språkens hus, n.d.) were conducted to obtain more uniform coding.

### 4.2.1 Scoring difficulties

Assessment of pupils' answers was very time-consuming. Despite instructions to give one word in response to the stimulus word, many pupils answered with longer phrases instead, which could be classified as paradigmatic, syntagmatic or 'other', depending on the type of phrase and the part of speech of the headword: different part of speech for syntagmatic responses (*doctor – sometimes I'm afraid of shots*); same part of speech for paradigmatic responses (*doctor – someone who gives medicine*). Multiword responses without semantic connection to the stimulus word were classified as 'other' (*woman – one who stands somewhere*). In general, as noted by Fitzpatrick & Thwaite (2020), scoring difficulties often regarded choices between syntagmatic and paradigmatic responses, or whether a word had a strong enough semantic relation to the stimulus word.

### 4.2.2 Results for Kent-Rosanoff test

Figure 6 presents results for all 92 pupils on the Kent-Rosanoff test, grouped according to grade level and time in Sweden. Pupils in grade 2 with language backgrounds other than Swedish used a larger percentage of clang responses compared to the L1 pupils, while pupils in grade 5 used proportionately very few or none overall, showing both age and proficiency effects (see Namei 2002).

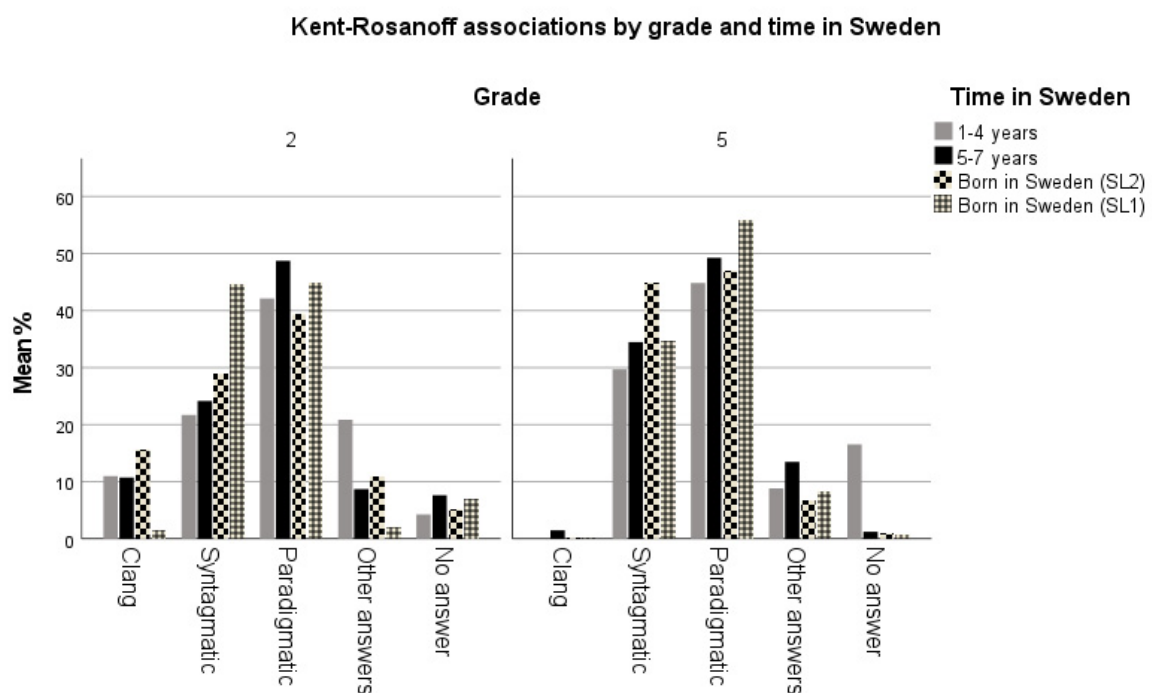


Figure 6: Proportions of clang, syntagmatic, paradigmatic, other and no answers for all pupils in grades 2 and 5 (N = 92), grouped according to time in Sweden. For numbers of pupils in each group, see Table 1.

Swedish L1 pupils in grade 2 responded with equal percentages of paradigmatic and syntagmatic responses, while L1 pupils in grade 5 used considerably more paradigmatic responses, which may be evidence for the syntagmatic/paradigmatic shift. Grade 5 pupils with shorter length of residence in Sweden also gave more paradigmatic responses, but the trend was not as clear for the L2 pupils born in Sweden, indicating that the shift may occur later in these multilingual children (Schoonen & Verhallen 2008). The percentages of responses classified as 'other' or no answers at all can be due to both age and proficiency level.

Following Sandgren et al. (2020) we calculated a semantic depth score for each pupil based on the number of syntagmatic (1 point) and paradigmatic (2 points) associations. As a group, 5<sup>th</sup> graders had significantly higher semantic depth scores than the 2<sup>nd</sup> graders ( $p = .011$ ;  $t = -2.586$ ), which may also indicate a higher level of lexical organization in their vocabulary.

## 5. Discussion

The aim of this article has been to discuss choices made and challenges experienced in assessing pupils' Swedish vocabulary depth in a multilingual elementary school. As highlighted by Schmitt (2014), empirical results depend first on how depth is conceptualized, and then on how it is measured. We see word definitions and associations as complementary measures of vocabulary depth, illustrating different aspects of pupils' vocabulary knowledge. Our study presents results using a new definition test, providing valuable new knowledge, considering the paucity of appropriately normed tests in Sweden. We chose to use a definition test with indirect prompting, where pupils were told to imagine that the researcher came from another planet. This contrasts with other methods such as Verhallen & Schoonen's (1993, 1998) extended interview model. We found that older pupils and those with Swedish as L1 had higher definition scores at group level, while L2 pupils often lacked the verbal ability to describe words (see Verhallen & Schoonen 1998; Vermeer 2001). These results are perhaps not surprising, as age and language proficiency are important factors influencing vocabulary development. However, many individual pupils in grade 2, including speakers of Swedish as L1 and L2, received higher scores than older pupils.

We also chose to use the Kent-Rosanoff association test, which had previously been used in Sweden. Our results seemed to indicate differences in the organization of L2 learners' vocabularies compared to L1 learners (Meara 1996; Namei 2002; Schoonen & Verhallen 2008) as well as between younger and older pupils, although this may also be related to task capacity (Nelson 1977). In addition, critics question the use of any one task to reflect lexical organization and call for more large-scale research focused on L2 word association networks (Fitzpatrick & Thwaites 2020). Moreover, our two assessments give an

indication of pupils' lexical knowledge based on the words included, and capture what pupils know at that specific time point (Read 2004). Choosing to use a different association task would likely have produced different results. If we had tested pupils' L1 as well and used conceptual scoring on the Kent-Rosanoff test as did Holmström et al. (2016b), we would have obtained a better picture of the multilingual pupils' lexical knowledge. Due to the large number of first languages among the pupils and our limited time frame, this was not possible.

Participants were not only linguistically diverse but had varying lengths of residency in Sweden. This contributed to the great overall variation in our results. We lacked specific information regarding pupils' previous educational experiences and exposure to Swedish which confounds the categorization of all multilingual pupils born in Sweden as L2 learners, even though most of them had two foreign-born parents. It is likely that some of these pupils may have two first languages. In Sweden today, where input in English is pervasive, very few young pupils, including those with Swedish language background, can be considered strictly monolingual. We also lacked background information regarding parental education and other socioeconomic indicators, which have been shown to account for a large degree of variation in pupils' vocabulary knowledge (Andersson et al. 2019). These factors, as well as the use of a newly constructed definition test without a set scoring rubric, influence the reliability of our results.

We had originally planned to carry out our battery of tests, which also included a newly constructed naming test and a test of academic word knowledge, at two time points. This would have given us a picture of development over time. However, our schedule was disrupted by the Corona pandemic. Still, we were able to provide a snapshot of what pupils in a linguistically diverse elementary school know or can articulate. Our results also demonstrate great individual variability in pupils' vocabulary knowledge, which is to be expected among multilingual learners, but also in young learners' language development overall.

Despite uncertainties in our methods, our qualitative and quantitative analyses represent valuable knowledge regarding vocabulary development among multilingual pupils and are of interest for researchers and educators alike. We hope that our study may inspire the development of methods to investigate vocabulary depth, as well as the creation of frequency lists appropriate for pupils in an elementary school context. This would deepen our understanding of age, frequency, and proficiency effects on word learning.



## REFERENCES

- Anderson, R. C. & Freebody, P. (1981). Vocabulary knowledge. I: 1. In T. Guthrie (ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). Newark, DE: International Reading Association.
- Andersson, K., Hansson, K., Rosqvist, I., Lyberg Åhlander, V., Sahlén, B. & Sandgren, O. (2019). The contribution of bilingualism, parental education, and school characteristics to performance on the clinical evaluation of language fundamentals: Swedish. *Frontiers in Psychology, 10*, 1586.
- Benelli, B., Arcuri, L. & Marchesini, G. (1988). Cognitive and linguistic factors in the development of word definitions. *Journal of Child Language, 15*(3), 619-635.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Press Syndicate of the University of Cambridge.
- Cox Eriksson, C. (2014). *Children's vocabulary development: The role of parental input, vocabulary composition and early communicative skills* (Doctoral dissertation, Department of Special Education, Stockholm University).
- Coyne, M.D., McCoach, D.B., Loftus S., Zipoli, R. & Kapp, S. (2009). Direct vocabulary instruction in kindergarten: Teaching for breadth versus depth. *The Elementary School Journal 110*(1), 1-18.
- Dale, E. (1965). Vocabulary measurement: Techniques and major findings. *Elementary English, 42*, 895-902, 948.
- Edquist, G. (2021). *Expressivt ordförråd hos barn med lätt till måttligt svår hörselnedsättning*. [Masteruppsats]. Institutionen för neurovetenskap, logopedi, Uppsala universitet.
- Erlandsson, J. & Yhlen, T. (2019). "Skillnad betyder att det kan vara en skillnad mellan en person och en annan person" *En jämförelse mellan en-och flerspråkiga skolbarns definitioner av skolrelaterade ord*. Lund, Sweden: Lund University.
- Fitzpatrick, T., & Thwaites, P. (2020). Word association research and the L2 lexicon. *Language Teaching, 53*(3), 237-274.
- Haman, E., Łuniewska, M., Hansen, P. Gram Simonsen, H., Chiat, S., Bjekić, J., Blažienė, A., Chyl, K., Dabašinskienė, I., Engel de Abreu, P., Gagarina, N., Gavarró, A., Håkansson, G., Harel, E., Holm, E., Kapalková, S., Kunnari, S., Levorato, C. Lindgren, J., Mieszkowska, K., Montes Salarich, L., Potgieter, A., Ribu, I., Ringblom, N., Rinker, T., Roch, M., Slančová, D., Southwood, F., Tedeschi, R., Müge Tuncer, A., Ünal-Logacev, Ö., Vuksanović, J. & Armon-Lotem, S. (2017). Noun and verb knowledge in monolingual preschool children across 17 languages: Data from Cross-linguistic Lexical Tasks (LITMUS-CLT). *Clinical Linguistics & Phonetics, 31*(11-12), 818-843.
- Holmström, K. (2015). *Lexikal organisation hos en-och flerspråkiga skolbarn med språkstörning*. (Doctoral dissertation, Lund University).
- Holmström, K., Salameh, E.K. Nettelbladt, U. & Dahlgren-Sandberg, A. (2016a). A descriptive study of lexical organisation in bilingual children with language impairment: Developmental changes. *International Journal of Speech-language Pathology, 18*(2), 178-189.
- Holmström, K., Salameh, E.K., Nettelbladt, U. & Dahlgren-Sandberg, A. (2016b). Conceptual scoring of lexical organization in bilingual children with language impairment. *Communication Disorders Quarterly, 38*(1), 24-34.
- Johansson, S. & Wahlstrand, E. (2010). *Lexikal organisation hos barn i förskoleklass och tredje klass. En omarbetning av Kent och Rosanoffs associationstest*. Lund, Sweden: Lund University.
- Juska-Bacher, B., Brugger, L., Korthus, R. & Zangger, C. (2021). Definitionskompetenzen von Erst- und Zweitklässlern. Mit einem Ausblick auf die weitere Entwicklung. *Muttersprache 2*, 117-137.
- Kilgarrieff, A., Charalabopoulou, F., Gavrilidou, M., Bondi Johannessen, J., Kahklil, S., Johansson Kokkinakis, S., Lew, R., Sharoff, S., Vadlapudi, R. & Voldina, E. (2014). Corpus-based vocabulary lists for language learners for nine languages. *Language Resources and Evaluation, 48*, 121-163.

- Kent, G. H. & Rosanoff, A.J. (1910). A study of association in insanity. *American Journal of Psychiatry*, 67(1), 37-96.
- Kurland, B.F. & Snow, C.E. (1997). Longitudinal measurement of growth in definitional skill. *Journal of child language*, 24(3), 603-625.
- Lindgren, J. (2018). *Developing narrative competence: Swedish, Swedish-German and Swedish-Turkish children aged 4-6* (Doctoral dissertation, Acta Universitatis Upsaliensis).
- McGregor, K.K., Berns, A.J., Owen, A.J., Michels, S. A., Duff, D., Bahnsen, A. J. & Lloyd, M. (2012). Associations between syntax and the lexicon among children with or without ASD and language impairment. *Journal of autism and developmental disorders*, 42(1), 35-47.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer & J. Williams (eds.), *Performance and competence in second language acquisition* (pp. 35-53). Cambridge: Cambridge University Press.
- Meara, P. & Wolter, B. (2004). V\_LINKS: Beyond vocabulary depth. In D. Albrechtsen, K. Haastруп, & B. Henriksen (eds.), *Angles on the English speaking world 4* (pp. 85-96). Copenhagen, Denmark: Museum Tusulanum Press.
- Mikoczy, T. & Nyman, A. (2008). *Lexikal förmåga hos flerspråkiga svensk-arabiska barn i skolår 4. En undersökning med PPVT-III och Kent & Rosanoffs associationstest*. Lund, Sweden: Lund University.
- Namei, S. (2002). *The bilingual lexicon from a developmental perspective: A word association study of Persian-Swedish bilinguals* (Doctoral dissertation, Stockholms universitet, Centrum för tvåspråkighetsforskning).
- Nation, I.S.P. (2013). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological bulletin*, 84(1), 93.
- Nilsson, A. & Svenbe, J. (2017) *Lexikal organisation hos ungdomar och unga vuxna*. Magisteruppsats. Linköpings universitet.
- Norman, J. (2019). *Ordförråd hos elever i en mångkulturell skola: Lexikal organisation hos två klasser i årskurs 2 och 5*. Falun, Sweden. Högskolan Dalarna.
- Ordóñez, C.L., Carlo, M.S., Snow, C.E. & McLaughlin, B. (2002) Depth and breadth of vocabulary in two languages: Which vocabulary skills transfer? *Journal of Educational Psychology* 94(4), 719-728.
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98(3), 554.
- Read, J. (1993). The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10, 355-371.
- Read, J. (2004). Plumbing the depths: How should the construct of vocabulary knowledge be defined? In P. Bogaards & B. Laufer (eds.), *Vocabulary in a second language: Selection, acquisition and testing* (pp. 209-227). Amsterdam: John Benjamins.
- Salameh, E.K. (2011). Lexikal utveckling på svenska och arabiska vid tvåspråkig undervisning. *Educare*, 3, 205-227.
- Sandgren, O., Salameh, E.K., Nettelblatt, U., Dahlgren-Sandberg, A. & Andersson, K. (2020). Using a word association task to investigate semantic depth in Swedish-speaking children with developmental language disorder. *Logopedics Phoniatics Vocology*, 46(3), 1-7.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Houndmills, Basingstoke, Hampshire, New York, NY: Palgrave Macmillan.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913-951.

- Schoonen, R. & Verhallen, M. (2008). The assessment of deep word knowledge in young first and second language learners. *Language Testing*, 25(2), 211-236.
- Semel, E., Wiig, E.H. & Secord, W.A. (2013). *CELF-4. Clinical evaluation of language fundamentals - fourth edition*. Bloomington, MN: NCS Pearson, Inc.
- Språkhus (n.d.) Material. Kent-Rosanoff list. [www.sprakenshus.se/material](http://www.sprakenshus.se/material).
- Verhallen, M. & Schoonen, R. (1993). Lexical knowledge of monolingual and bilingual children. *Applied linguistics*, 14(4), 344-363.
- Verhallen, M. & Schoonen, R. (1998). Lexical knowledge in L1 and L2 of third and fifth graders. *Applied Linguistics*, 19(4), 452-470.
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, 22(2), 217-234.



# Impact of parental reading skills and children's reading environment and interest at pre-school age on the vocabulary and reading skills in elementary school

**Livia N. HAAG, Sarah V. DI PIETRO, Martina RÖTHLISBERGER, Carina WICK, Rita FÜZÉR & Silvia BREM**

University of Zurich / University Hospital of Psychiatry  
Department of Child and Adolescent Psychiatry and Psychotherapy  
Neumünsterallee 9, CH-8032 Zürich  
livia.haag@kjpd.uzh.ch, sarahvalerie.dipietro@kjpd.uzh.ch,  
martina.roethlisberger@kjpd.uzh.ch, carina.wick@bluewin.ch,  
rita.fuezer@kjpd.uzh.ch, silvia.brem@kjpd.uzh.ch

Das Ziel dieser Studie ist es, den Einfluss der mütterlichen und väterlichen Lesekompetenzen sowie des Leseumfelds von Kindern im Vorschulalter auf deren Wortschatz, Lesefertigkeiten und Textverständnis am Ende der Grundschule zu klären. Dazu wurden der rezeptive Wortschatz, die Leseflüssigkeit und weitere Lesekompetenzen von 91 Schulkindern erhoben und mit der elterlichen Lesekompetenz, dem sozioökonomischen Status und der non-verbale Intelligenz sowie dem Leseumfeld ebendieser Kinder im Vorschulalter in Beziehung gesetzt. Die Resultate zeigen, dass das vorschulische Leseumfeld und der non-verbale IQ Prädiktoren des Wortschatzes von Kindern der 5. Klasse sind, und dass die Lesefertigkeiten der Eltern und das vorschulische Interesse an Büchern die Lesefertigkeiten der Kinder in der 5. Klasse vorhersagen. Diese Studie zeigt auf, welche Faktoren des vorschulischen Leseumfelds als Prädiktoren des Wortschatzes und der Lesekompetenzen am Ende der Primarschule besonders wichtig sind.

## **Stichwörter:**

Wortschatz, Textverständnis, Lesekompetenz, Dekodierfähigkeit, vorschulisches Leseumfeld, elterliche Lesekompetenz, non-verbaler IQ, sozioökonomischer Status.

## **Keywords:**

vocabulary, reading comprehension, reading fluency, decoding fluency, preschool literacy environment, parental reading level, non-verbal IQ, socioeconomic standard.

## **1. Introduction**

The word is the central pillar of human language and communication because it conveys complete semantic and grammatical meaning. Vocabulary – the number of words an individual can recall (definition of the Cambridge dictionary (Cambridge University Press, 2014)) – plays an important role for a child's school and academic success, and is closely linked to general cognitive abilities (Biemiller 2006; Kievit et al. 2017; Lenhard & Lenhard 2021; Pace et al. 2019). Vocabulary has been recognised as a critical factor to promote learning in various domains, such as verbal concept formation, reading and also fluid intelligence skills (e.g. matrix reasoning) (Kievit et al. 2017). It is thereby essential not only for a successful scholastic career but also for children's overall cognitive and social development (Ebert et al. 2013). Young children's verbal



production begins around the age of one to one and a half. Between the age of one and a half and two years, children master about 50 words actively and have a receptive vocabulary of around 200 words (Kauschke 2021; Szagun 2006). Between the age of 2 and 6, vocabulary is massively boosted, such that at school onset, the beginning of literacy acquisition, the average receptive vocabulary consists of about 14,000 words, with extraordinarily wide variation among individuals (Clark 2009).

There is considerable evidence that home literacy environment (HLE) is an important predictor for children's literacy and language development (Frijters et al. 2000; Levy et al. 2006; Sénéchal & LeFevre 2002). HLE includes several formal and informal literacy activities such as shared parental reading, teaching the child how to write his or her name but also passive influences such as parental attitudes towards literacy or their literacy resources (e.g. number of books available at home) (Puglisi et al. 2017). A large part of early childhood vocabulary acquisition simply occurs through everyday language contact. Child-directed speech, storytelling and reading aloud by adult caregivers critically support children in this process (Clark 2009). Particularly the amount of child-directed speech has a considerable impact on the size of the expressive vocabulary in toddlers (Huttenlocher et al. 1991; Weisleder & Fernald 2013), and its further development (Rowe et al. 2012). The important contribution of shared reading on young children's vocabulary growth (Marulis & Neuman 2010) can be explained by the fact that storybooks contain up to 50% more words than are commonly used during TV shows, in everyday conversations or in child-directed speech (Hayes & Ahrens 1988). The benefits are further enhanced when shared reading or storytelling is structured as an interactive process (Vaahtoranta et al. 2019). Shared reading not only has a positive effect on vocabulary growth, but also on other literature related abilities, such as recognising letters, linking print to speech but also on learning how to handle a book (Bus et al. 1995). According to a meta-analysis, shared storybook reading of preschool children with their parents has, among other factors, an important positive effect on reading achievement independent of the socio-economic status of the family and is thus considered an important prerequisite for successful reading acquisition (Bus et al. 1995).

While shared reading is important for vocabulary development, vocabulary itself has a long-lasting impact on reading acquisition (Ouellette 2006; Sénéchal et al. 2006; Whitehurst & Lonigan 1998). It is an important predictor of several literacy-related abilities such as spelling, reading comprehension and accuracy (Fricke et al. 2009; Nation & Snowling 2004).

Moreover, vocabulary expands and enhances the phonological representation of words and sublexical items in the mental lexicon, thus fostering phonological awareness (Metsala 1999). Phonological awareness, for its part, is typically impaired in poor readers (Goswami 2000; Vellutino et al. 2004), which is why

promoting vocabulary through shared reading is particularly beneficial for these children. Accordingly, children with good reading skills and a rich vocabulary usually read more frequently, thereby learn the meaning of more words and consequently improve their reading skills even more (Stanovich 1986). Conversely this also means that the poor get poorer. This may be especially critical for those 3-11% of children suffering from developmental dyslexia (DD), a neurodevelopmental learning disorder characterised by severe impairments in reading and/or spelling (Galuschka & Schulte-Körne 2016). In a recent study of Snowling and colleagues (2020), more than a third of the children with DD showed impairments in vocabulary, which may at least partly stem from less reading experience. If, in addition, children with specific language impairments are taken into account, the proportion of those showing deficits in vocabulary increases to more than 80% (Snowling et al. 2020).

Yet, findings from studies on the role of HLE in children with specific language impairments or DD vary widely. While Skibbe and colleagues (2008) did not find any significant association between HLE and children's emerging reading skills, HLE was the only significant predictor of print knowledge after controlling for socio-economic status (SES) in the study of McGinty & Justice (2009). Importantly, HLE also showed a positive effect on vocabulary in children at risk for dyslexia (Caglar-Ryeng et al. 2020). A possible account for these inconsistencies is that the relation between HLE and children's literacy skills is mediated by genetic factors (Puglisi et al. 2017; van Bergen et al. 2017). A meta-analysis in twins estimated the heritability of reading at 73% and of reading comprehension at 49%. Environmental influences accounted only for around 10-13% of the variance in reading and reading comprehension (de Zeeuw et al. 2015). Slightly lower effects of heritability (57%) but similar effects of environment (13%) were reported for vocabulary (verbal IQ) (Rowe et al. 1999). Interestingly, the influence of heritability was especially high and environmental influences low in highly educated families. In families with lower educational background, the environmental influences increased (Rowe et al. 1999). However, the contribution of genetic and environmental contributions changes with increasing literacy skills and age (Samuelsson et al. 2007; Verhoef et al. 2021).

In the present study, we aimed to clarify how the HLE at preschool age impacts literacy development on the long-term in a sample of 91 native German-speaking school children coming from middle to high SES families, the majority having a moderate to high familial risk for developmental dyslexia. Specifically, we aimed to determine how the current reading performance of both parents, as well as HLE and literacy interests of the children at preschool age account for children's receptive vocabulary, reading fluency, decoding fluency and reading comprehension skills. The mothers of the children showed on average better performance in reading and decoding fluency skills than the fathers (see Table 1). To account for possible differences in the involvement and effect of

each parents' preschool reading activities with their children on children's later language performance, we recorded joint reading activities separately for each parent in our questionnaires. At the time of the study, children were attending the 5<sup>th</sup> grade of Swiss elementary school and were about 11 years old. In accordance with previous studies (Ebert et al. 2013), we expected to find a moderate impact of children's HLE on the long-term outcome of vocabulary skills and a strong familial transmission of reading skills shown by a strong association of parents' and children's current reading skills (Lyytinen et al. 2004; Snowling et al. 2003).

## 2. Methods

### 2.1 Participants

The sample included 91 (37 boys, 54 girls, 80 right-handed) mostly native (Swiss-)German-speaking children aged between 10 and 12 years in the 5<sup>th</sup> (n=90) or in the 6<sup>th</sup> (n=1) grade of elementary school. 17 children spoke another language next to (Swiss-)German. Most children were from middle- to upper-class income families (mean SES=62; Min.= 10, Max.= 89 according to ISEI-08, see below). Exclusion criteria were physical, neurological, or psychiatric disorders. Based on parents' report (questionnaire), our sample included children with delayed language development (n=13), diagnosis of developmental dyslexia (n=23) and attention deficit hyperactivity disorder (ADHD, n=9). The 9 children diagnosed with ADHD were either receiving no medication or we asked for medication to be discontinued before the behavioural tests. One child with an autism spectrum disorder was excluded. The Adult Reading History Questionnaire (ARHQ) of both parents served to identify children with a familial risk for dyslexia (Lefly & Pennington 2000). The ARHQ is a self-assessment questionnaire, in which a score of > 0.40 indicates a history of reading disability. Children in these families met criteria of heightened risk for dyslexia. Accordingly, a familial risk was presumed when at least one parent scored > 0.40 in the ARHQ (one parent n=39, both parents n=15, mothers n=26, fathers n=40) or if a sibling had reading problems (n=18). Hence children at risk and with poor reading skills are overrepresented in this study (see Table 1). Furthermore, 15 mothers and 35 fathers of our children performed below 1 SD in the standardised decoding fluency test, indicating current impairments in reading.

The participants were recruited by flyers, letters and advertisements distributed in schools in the context of a larger project on neural correlates of reading development and dyslexia. The research protocol was approved by the local ethics committee in Zurich, and all parents gave written informed consent in accordance with the Declaration of Helsinki.



	N	Minimum	Maximum	Mean	SD
Age	91	9.97	12.38	11.37	0.41
ARHQ mother	91	0.03	0.80	0.32	0.16
ARHQ father	89	0.06	0.79	0.40	0.15
ARHQ mean (both parents)	91	0.09	0.65	0.36	0.11
Vocabulary (PPVT4) <sup>a*</sup>	90	4.50	98.90	55.34	26.98
Reading fluency (SLRT-II W) <sup>b*</sup>	91	<1.00	95.00	37.19	31.32
Decoding fluency (SLRT-II PW) <sup>b*</sup>	91	<1.00	98.00	34.80	32.10
Text comprehension (ELFE II) <sup>c*</sup>	91	1.10	99.40	51.07	31.56
Non-verbal IQ (RIAS) <sup>d</sup>	91	80	123	104.65	6.96
SES	91	31	89	62.32	12.90
Reading fluency mother (SLRT-II W) <sup>b*</sup>	90	<1.00	93.50	41.99	26.20
Decoding fluency mother (SLRT-II PW) <sup>b*</sup>	90	<1.00	97.50	44.36	26.94
Reading fluency father (SLRT-II W) <sup>b*</sup>	83	<1.00	90.5	32.84	25.08
Decoding fluency father (SLRT-II PW) <sup>b*</sup>	83	<1.00	89.50	29.24	23.23

Note: N = Number, SD = Standard deviation; ARHQ = Adult Reading History Questionnaire; SES = Socio-economic Status; PW = Pseudoword; W = Word; <sup>a</sup>Peabody picture vocabulary test (PPVT4); <sup>b</sup>Salzburger Lese- und Rechtschreibtest-II (SLRT-II); <sup>c</sup>Ein Leseverständnistest für Erst- bis Siebtklässler Version II (ELFE-II); <sup>d</sup>Reynolds Intellectual Assessment Scales and Screening (RIAS), \*percentile scores; Mothers and fathers differed significantly with regard to their percentiles in reading fluency ( $t=3.66$ ,  $p<0.001$ ), decoding fluency ( $t=3.98$ ,  $p<0.001$ ) and their ARHQ risk scores ( $t=-3.35$ ,  $p<0.001$ ).

Table 1: Mean, standard deviation (SD), minima, and maxima of the different reading, vocabulary, IQ, SES measures of the children, risk assessment and reading scores of their mothers and fathers.

The current study focussed on the analysis of reading skills, vocabulary, non-verbal IQ, and the questionnaires related to reading history and reading environment of the children.

## 2.2 Testing procedure and materials

Behavioural tests to assess children's reading and vocabulary skills, as well as other measurements (Table 1), took place at the children's home ( $n=61$ ) or in an office at the Department of Child and Adolescent Psychiatry and Psychotherapy ( $n=30$ ) and are detailed in the next section. In addition, the children's parents filled in several questionnaires and also conducted a timed overt word and pseudoword reading test (SLRT-II: Moll & Landerl 2014)<sup>1</sup>. Data on vocabulary skills was missing for one child, reading tests were missing for one mother and eight fathers and questionnaires were available for 90 families.

<sup>1</sup> For a detailed test description see below.

### 2.2.1 Behavioural Tests

*Reading fluency and decoding:* Word and pseudoword reading fluency of the participating children and their parents was assessed using the one-minute-reading fluency tests of the "Salzburger Lese- und Rechtschreibtest" (SLRT-II) (Moll & Landerl 2014). The percentile rank (PR) of the number of correctly read words/pseudowords per minute served as a measure for reading fluency, respectively decoding fluency.

*Reading comprehension:* To measure reading comprehension, we administered the short version of the text comprehension subtest of ELFE-II ("Ein Leseverständnistest für Erst- bis Siebtklässler" Version II) by Lenhard et al. (2018). The children had six minutes to read as many short texts and subsequently select the correct answer out of four given options as possible. The number of correct answers was used to derive the PR of reading comprehension for each child.

*Vocabulary:* To estimate the receptive vocabulary of each child, the German version of the "Peabody Picture Vocabulary Test" Volume 4 (PPVT-4) by Lenhard et al. (2015) was conducted. The participant must choose out of four pictures the one that best matches to the word read aloud by the examiner. The PR was derived using the total number of correct answers.

*IQ estimate:* The non-verbal IQ of each child was estimated by use of the two non-verbal subtests "Odd Man Out" (German: Unpassendes ausschliessen) and "What's Missing?" (German: Was fehlt?) of the German edition of the "Reynolds Intellectual Assessment Scales and Screening" (RIAS) by Hagmann-von Arx & Grob (2014).

Further behavioural tests were carried out (not reported in the present study) and included examinations of verbal IQ (RIAS verbal part, Hagmann-von Arx & Grob 2014), rapid automatized naming of objects, colours, letters, numbers (subtests of the Test zur Erfassung der phonologischen Bewusstheit und der Benennungsgeschwindigkeit, Mayer 2013), rapid automatized naming of animals (Moll et al. 2005), math skills (Heidelberger Rechentest, Hafner et al. 2005), text reading accuracy (Zürcher Lesetest-II, Petermann & Daseking 2015), short term and working memory (Wechsler Intelligence Scale for Children-Fourth Edition, Wechsler 2011), reading speed (Salzburger Lese-Screening für die Schulstufen 2-9, Mayringer & Wimmer 2014), spelling (Deutscher Rechtschreibtest für 5. Klassen, Grund et al. 2015) and phonological awareness by phoneme blending and phoneme reversal tasks (Basiskompetenzen für Lese-Rechtschreibleistungen Stock et al. 2017).

### 2.2.2 Questionnaires

To get an estimate of children's preschool HLE and to have an index of parents' reading history, the parents filled in several questionnaires about their child's reading environment and interest in books as well as their own reading history.

The Adult Reading History Questionnaire (ARHQ) by Lefly & Pennington (2000) consists of 23 questions about the reading history and the current reading behaviour. It is a widely used measure to self-report the risk of reading disability (dyslexia) in adults. Both parents filled in this questionnaire to determine if there was a familial risk of dyslexia for the child (cf. Table 1). ARHQ values and parent's individual reading skills were highly correlated as expected (word reading fluency mother:  $r = -.53$ ,  $p < .001$ ; decoding mother:  $r = -.61$ ,  $p < .001$ ; word reading fluency father:  $r = -.40$ ,  $p < .001$ ; decoding father:  $r = -.66$ ,  $p < .001$ ). We chose pseudoword decoding fluency as the core measure to define parents' reading skills, because it was highly correlated with ARHQ values, and because this measure is more sensitive than word reading fluency to potential compensated reading problems.

Further questionnaires, which were developed within our group, enquired about children's reading interest, reading environment and childcare situation at home at preschool and kindergarten age (questions listed in Table 2), and about parents' occupation and education to obtain an estimate of the family's SES. The ISEI-08 score of the occupation as defined by the International Standard Classification of Occupations (Ganzeboom et al. 1992) was used to derive a SES estimate ranging between 10 and 89. In total, 14 questions (see Appendix Table A1) were entered to the subsequent principal component analysis (PCA, chapter 2.3).

### 2.3 Statistical analyses

For all statistical analyses, the IBM SPSS Statistics 25 software was used. First, a PCA with promax rotation was applied to the data to reduce the number of measures. The initially 14 measures resulting from the questionnaires were reduced to a small set of informative, independent summary factor scores. To justify the use of a PCA, the Kaiser–Meyer–Olkin measure of sampling adequacy ( $KMO = .61$ ) and Bartlett's test of Sphericity ( $p < .001$ ) were checked first. The four factors indicated in the scree-plot, all exceeding an Eigenvalue of 1 and accounting for 69% of the total variance, were considered for further analyses (cf. results).

For the reading measures of both parents, the SES and children's non-verbal IQ, Z-scores were computed and subsequently used for all analyses. Shapiro-Wilk tests showed that all variables were non-normally distributed, therefore Spearman's correlations were used to analyse the intercorrelations between all independent and dependent variables.

Finally, four multiple regression analyses were conducted to explain the variance in the four outcome measures vocabulary, reading fluency, decoding fluency and text comprehension in the 5<sup>th</sup> grade by eight independent variables. These variables included the four factor scores derived by the PCA and the Z-scores of the SES, the non-verbal IQ and the decoding fluency of father and

mother. For the multiple regression analyses, we first checked for outliers in the dependent and independent variables. Data with standardised residuals exceeding  $\pm 3$  were excluded from further analyses (this applied to one case for the multiple regression with the dependent variable vocabulary). Afterwards, based on the guidelines by Huber (1981), three more subjects with a centred leverage value  $>0.2$  in one of the independent variables were excluded for all the multiple regression analyses. Values of  $p < .050$  were considered to be statistically significant and values of  $p < .100$  as trends.

### 3. Results

#### 3.1 *Principal component analysis (PCA)*

The factor loadings after rotation are summarised in Table 2. According to Stevens (2002), loadings greater than .51 were considered important, given the sample size of almost 100 children. The results from the pattern matrix indicated four components: component 1 represents the preschool childcare situation, component 2 the children's early interest in books, component 3 shared reading with the father & children's preschool reading skills and component 4 shared reading with the mother. 12 out of the 14 questions showed a significant loading on one of the four components (printed in bold type in Table 2).

#### 3.2 *Intercorrelation Analyses*

The results of the Spearman's correlational analyses between all dependent and independent variables are illustrated and summarised in Appendix Figure A1.

Item	Rotated Factor Loadings			
	<i>Preschool childcare</i>	<i>Early interest in books</i>	<i>Shared reading father &amp; preschool reading</i>	<i>Shared reading mother</i>
<b>Preschool childcare 0-2 years old</b>	<b>.93</b>	-.04	-.27	-.06
<b>Preschool childcare 2-4 years old</b>	<b>.96</b>	-.08	-.32	-.07
<b>Preschool childcare 4-6 years old</b>	<b>.93</b>	-.08	-.39	-.10
<b>Was your child particularly interested in book(s) between the ages of 0 and 4?</b>	-.09	<b>.80</b>	.34	.18
<b>Was your child particularly interested in book(s) between the ages of 4 and 6?</b>	.06	<b>.82</b>	.30	.05
<b>Did your child have a favourite book that he/she wanted to read several times a week between the ages of 0 and 4?</b>	-.23	<b>.60</b>	-.10	.45
<b>Did your child have a favourite book that he/she wanted to read several times a week between the ages of 4 and 6?</b>	-.14	<b>.71</b>	-.08	.48
<b>How often did you (father) read together with your child between the ages of 0 and 4?</b>	-.44	.14	<b>.91</b>	.09
<b>How often did you (father) read together with your child between the ages of 4 and 6?</b>	-.41	.13	<b>.91</b>	.08
<b>Could your child already read simple texts at kindergarten age (4-6 years old)?</b>	.03	.31	<b>.55</b>	.06
<b>How often did you (mother) read together with your child between the ages of 0 and 4?</b>	-.07	.28	.06	<b>.91</b>
<b>How often did you (mother) read together with your child between the ages of 4 and 6?</b>	-.06	.31	.11	<b>.90</b>
How often did your child talk during shared reading between the ages of 4 and 6?	-.28	.09	.26	.17
How often did your child ask questions during shared reading between the ages of 4 and 6?	-.27	.30	.22	.39
Eigenvalues	4.18	2.58	1.61	1.36
% of variance	29.83	18.41	11.52	9.73

Note: Factor loadings >.51 appear in bold.

Table 2: Summary of PCA results Note: Factor loadings >.51 appear in bold.

### 3.3 Predictors of the vocabulary

The multiple regression model explained 43% of the variance (adjusted  $R^2 = .43$ ) in children's vocabulary skills in the 5<sup>th</sup> grade ( $F(8,68) = 8.27$ ,  $p < .001$ , Table 3). From the eight independent variables, *non-verbal IQ* ( $t = 3.14$ ,  $p = .003$ ), *shared reading father & preschool reading* ( $t = 2.76$ ,  $p = .008$ ) and *shared reading mother* ( $t = 2.02$ ,  $p = .048$ ) showed a significant effect and *early interest*

*in books* showed a trend ( $t = 1.74$ ,  $p = .087$ ). The corresponding partial regression plots are shown in Appendix Figure A2.

	$\beta$	t	Sig.
<b>Non-verbal IQ</b>	<b>.32</b>	<b>3.14</b>	<b>.003</b>
SES	.05	0.47	.643
Decoding fluency mother	.11	1.12	.265
Decoding fluency father	-.09	-0.92	.363
Preschool childcare	.14	1.45	.152
<i>Early interest in books</i>	.20	1.74	.087
<b>Shared reading father &amp; preschool reading</b>	<b>.30</b>	<b>2.76</b>	<b>.008</b>
<b>Shared reading mother</b>	<b>.21</b>	<b>2.02</b>	<b>.048</b>

Note: Significant effects appear in bold letters; trends appear in italic letters;  $\beta$  = standardised beta coefficient

Table 3: Summary of the multiple regression model for vocabulary

### 3.4 Predictors of reading fluency

34% of the variance (adjusted  $R^2 = .34$ ) in children's reading fluency in 5<sup>th</sup> grade ( $F(8,70) = 5.98$ ,  $p < .001$ , Table 4) was explained in the multiple regression model. *Decoding fluency father* ( $t = 2.65$ ,  $p = .010$ ) and children's *early interest in books* ( $t = 2.67$ ,  $p = .009$ ) showed a significant effect, *decoding fluency mother* ( $t = 1.99$ ,  $p = .051$ ) showed a trend. The corresponding partial regression plots are shown in Appendix Figure A3.

	$\beta$	t	Sig.
Non-verbal IQ	-.01	-0.11	.912
SES	-.01	-0.11	.912
<i>Decoding fluency mother</i>	.22	1.99	.051
<b>Decoding fluency father</b>	<b>.27</b>	<b>2.65</b>	<b>.010</b>
Preschool childcare	.07	0.71	.479
<b>Early interest in books</b>	<b>.33</b>	<b>2.67</b>	<b>.009</b>
Shared reading father & preschool reading	.17	1.52	.133
Shared reading mother	-.00	-0.01	.994

Note: Significant effects appear in bold letters; trends appear in italic letters;  $\beta$  = standardised beta coefficient

Table 4: Summary of the multiple regression model for reading fluency

### 3.5 Predictors of decoding fluency

39% of the variance (adjusted  $R^2 = .39$ ) of children's decoding fluency in the 5<sup>th</sup> grade ( $F(8,70) = 7.21$ ,  $p < .001$ , Table 5) was explained in the multiple regression model. The *mother's* ( $t = 2.28$ ,  $p = .026$ ) and the *father's* ( $t = 3.46$ ,  $p = .001$ ) *decoding fluency* showed a significant effect as well as children's *early interest in books* ( $t = 2.50$ ,  $p = .015$ ) and *shared reading father & preschool reading* ( $t = 2.24$ ,  $p = .029$ ). The corresponding partial regression plots are shown in Appendix Figure A4.

	$\beta$	t	Sig.
Non-verbal IQ	-.10	-0.92	.363
SES	-.11	-1.09	.279
<b>Decoding fluency mother</b>	<b>.24</b>	<b>2.28</b>	<b>.026</b>
<b>Decoding fluency father</b>	<b>.34</b>	<b>3.46</b>	<b>.001</b>
Preschool childcare	.08	0.85	.400
<b>Early interest in books</b>	<b>.30</b>	<b>2.50</b>	<b>.015</b>
<b>Shared reading father &amp; preschool reading</b>	<b>.25</b>	<b>2.24</b>	<b>.029</b>
Shared reading mother	.01	0.12	.904

Note: Significant effects appear in bold letters;  $\beta$  = standardised beta coefficient

Table 5: Summary of the multiple regression model for decoding fluency

### 3.6 Predictors of reading comprehension

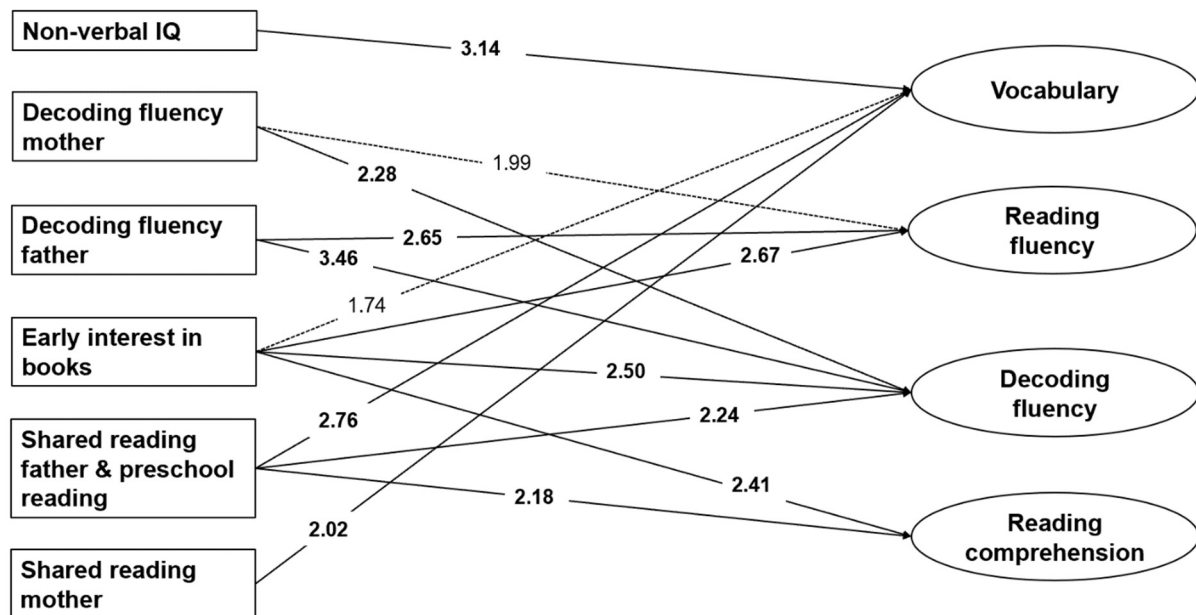
31% of the variance (adjusted  $R^2 = .31$ ) of children's reading comprehension in the 5<sup>th</sup> grade ( $F(8,70) = 5.39$ ,  $p < .001$ , Table 6) was explained in the multiple regression model. Only children's *early interest in books* ( $t = 2.41$ ,  $p = .019$ ) and *shared reading father & preschool reading* ( $t = 2.18$ ,  $p = .033$ ) significantly contributed to explaining the variance in reading comprehension. The corresponding partial regression plots are shown in Appendix Figure A5.

	$\beta$	t	Sig.
Non-verbal IQ	.17	1.52	.133
SES	-.07	-0.68	.500
Decoding fluency mother	.12	1.04	.301
Decoding fluency father	.13	1.22	.227
Preschool childcare	.08	0.73	.471
<b>Early interest in books</b>	<b>.30</b>	<b>2.41</b>	<b>.019</b>
<b>Shared reading father &amp; preschool reading</b>	<b>.25</b>	<b>2.18</b>	<b>.033</b>
Shared reading mother	.02	0.14	.887

Note: Significant effects appear in bold letters;  $\beta$  = standardised beta coefficient

Table 6: Summary of the multiple regression model for reading comprehension

All the significant effects and trends of the four multiple linear regression models are summarised and presented in Figure 1. The figure nicely illustrates that preschool reading interest and home reading environment had significant positive effects on later reading, vocabulary and comprehension skills, while the reading skills of the parents exclusively relate to children's current reading skills.



Note: T-values are given on the arrows between variables. Solid arrows with bold numbers indicate a significant effect; dotted arrows indicate a trend. Only variables showing a significant (or trend) association are considered for this illustration.

Figure 1: Overall model, summarising the contributions of the different measures on the four outcome variables

## 4. Discussion

The goal of this study was to identify the influence of several preschool factors such as shared reading activities and children's interest in books as well as parents' current reading performance, the SES and the current non-verbal IQ on children's vocabulary, reading and decoding fluency and reading comprehension towards the end of elementary school. All regression analyses achieved a moderate to good prediction with explained variances ranging from 31% for text comprehension up to 43% for vocabulary size in 5<sup>th</sup> grade. Shared reading with parents at preschool age, children's preschool reading skills and children's early interest in books contributed to the prediction of most outcome variables. The non-verbal IQ of the children, in contrast, was exclusively relevant for explaining variance in vocabulary, and the parents current reading skills only contributed to explaining variance in children's fluency measures (word reading and decoding fluency). In the current sample, the estimated SES did not show any significant effect on reading skills or vocabulary. This is in accordance with previous evidence that the SES shows a greater influence on children's achievements and ability-related outcome in low-income than in high-income families (Duncan et al. 1998). Furthermore, the preschool childcare situation at children's homes did not account for any variance on reading and vocabulary skills.



#### *4.1 Preschool HLE and early interest in books impact on children's later vocabulary and reading outcome*

Our overall definition of preschool HLE included three factors, namely 'shared reading with the mother' and 'shared reading with the father and preschool reading skills', and 'children's early interest in books'. Shared reading with the parents was a significant predictor of vocabulary. It is, however, important to note that the factor of shared reading with the father also included children's preschool reading skills. Early interest in books was not significantly related to the development of children's vocabulary.

There is a large consensus that shared reading with parents at preschool age is beneficial for the development of vocabulary size in young children (Bracken & Fischel 2008; Bus et al. 1995; Sénéchal & Lefevre 2014) including children at risk for dyslexia (Caglar-Ryeng et al. 2020). Our findings of the strong impact of shared reading at preschool on later vocabulary outcome is therefore convergent with previous literature and confirms the beneficial effects of shared reading at home for building up a rich vocabulary. Moreover, our findings confirm that shared storybook reading is highly beneficial for children coming from middle to high SES backgrounds with a family background of dyslexia and children developing specific reading impairments. In contrast to the impact of shared storybook reading on vocabulary development, its impact on reading skills is less direct. According to the home literacy model (Sénéchal & LeFevre 2002, 2014), informal and meaning-related literacy activities such as storybook reading have only an indirect effect on reading through enhancing children's language skills (e.g. vocabulary). Direct effects on word reading may in contrast be expected through formal, print-related literacy activities such as letter teaching at preschool age (Sénéchal & Lefevre 2014). Our findings are in accordance with this model in that shared reading with the mother, an informal meaning-related activity, was not associated with reading. The effect of shared reading with the father and preschool reading on both decoding and reading comprehension in the 5<sup>th</sup> grade might be driven by the impact of preschool reading skills of the children on their further reading development, rather than shared reading with the father itself (Bowey 2005; Leppänen et al. 2008). In this study we did not specifically assess print-related teaching activities at children's homes. But it is likely that such activities fostered early preschool reading skills, which loaded on the factor shared reading with father. Such early preschool reading skills thus may explain the beneficial influence on children's long-term reading development. Of note, also shared reading may have indirect effects on reading comprehension in the long run: It has been shown, for example, that shared preschool storybook reading had a beneficial effect on reading comprehension in grade 4 (Sénéchal 2006).

While the effect of shared reading at preschool age most likely only indirectly influences future reading skills through early preschool literacy knowledge,

children's early interest in books was a strong predictor for all reading measures in the 5<sup>th</sup> grade. In this study, the factor '*early interest in books*' is formed by a questionnaire filled in by the parents. The questions that loaded on the factor asked on whether the child was particularly interested in books, and if the child had a favourite book at preschool and kindergarten age. Our results thus converge with the strong relationship between the attitude towards reading and reading achievement in primary school children as reported in a meta-analysis (Petscher 2010). Investigating reading self-concept, Katzir and colleagues (2009) obtained similar results for reading comprehension. In their study, children's attitude towards reading, as one of three dimensions considered for reading self-concept, showed a significant effect on reading comprehension in 4<sup>th</sup>-grade pupils (Katzir et al. 2009). In comparison to the strong effect on reading scores, the effect of reading interest on later vocabulary size in our sample was rather marginal. Moreover, the literature on the effect of early interest on vocabulary size or growth is inconsistent. Several studies show a beneficial effect of reading interest at preschool age on vocabulary growth (Bracken & Fischel 2008; Crain-Thoreson & Dale 1992) and others show no association (Caglar-Ryeng et al. 2020; Frijters et al. 2000; Sparks & Reese 2013).

While our results on the impact of HLE or early interest in books on vocabulary and reading outcomes largely coincide with previous literature, partly diverging results may be explained by a lack of standardised questionnaires or definitions on how to assess such factors. The assessment of "HLE" for example varies considerably among studies and includes different variables, such as shared storybook reading (Sénéchal 2006; Silinskas et al. 2020), parents' report of their own reading frequency and number of subscribed newspapers (Johnson et al. 2008; van Bergen et al. 2017), number of books at home (Katzir et al. 2009; Sénéchal et al. 1996), frequency of library visits (Caglar-Ryeng et al. 2020; Sénéchal & LeFevre 2002) among others. As with HLE, there are also different approaches to assess interest in books such as through a self-report by the children (Frijters et al. 2000) or questionnaires filled in by the parents (Bracken & Fischel 2008). Furthermore, different types of questions were used to assess reading interest, which again may explain diverging results. Moreover, reading and decoding fluency as well as vocabulary are often assessed at different times during reading development in primary school, which renders such comparisons difficult (Sénéchal et al. 1998; Silinskas et al. 2020).

#### ***4.2 Parental reading skills are strongly related to children's reading skills***

In agreement with most previous studies (Lyytinen et al. 2004; Pennington & Lefly 2001; Snowling et al. 2003; van Bergen et al. 2012, 2015), our data provide evidence that children's reading and decoding fluency are strongly associated with the decoding fluency of their parents. This strong link between children's and parents' reading skills is likely to be mediated by genetic factors, as

indicated by a meta-analysis reporting a high heritability of 73% between reading scores of the parents and their children (de Zeeuw et al. 2015), and the finding that only the reading level of non-adopted, biological children, but not of adopted children is correlated with parental reading fluency (Wadsworth et al. 2002). Both parents contributed similarly to children's reading fluency skills in our study. In accordance with previous studies, our sample showed a higher prevalence of impaired reading in fathers than mothers (Hawke et al. 2009; Moll et al. 2014; Rutter et al. 2004). 47% of the fathers and 31% of the mothers exceeded the threshold for a high dyslexia risk as estimated by the adult reading history questionnaire (ARHQ) and 42% of the fathers and 17% of the mothers performed worse than 1SD below norms in the oral decoding fluency test. Despite the strong association between parents' and children's reading fluency skills, parents' decoding fluency did not show any significant effect on children's reading comprehension or vocabulary skills. Reading comprehension and vocabulary were better explained by literacy activities at children's homes as described in the former section.

#### *4.3 Non-verbal IQ is associated with vocabulary in 5<sup>th</sup> grade*

The non-verbal IQ showed a strong effect on children's vocabulary in the 5<sup>th</sup> grade, but it is not related to children's reading and decoding fluency nor to comprehension skills. Accordingly, children with better non-verbal abilities have on average a larger vocabulary. Even though both vocabulary and non-verbal IQ are often seen as stable and unrelated traits (Blaga et al. 2009), recent models emphasise the mutualistic coupling of basic cognitive abilities across development (Kievit et al. 2017). Apart from the study of Kievit and colleagues (2017), in which adolescents and adults were examined, studies in children with and without language impairments also indicate that a higher non-verbal IQ has a significant positive effect on receptive vocabulary growth (Lervåg et al. 2019; Rice & Hoffman 2015). The strong relation of non-verbal IQ and vocabulary findings in our study are thus in accord with such a mutualist model account (Kievit et al. 2017).

In contrast to vocabulary, we found no indication that non-verbal abilities of children are associated with their reading skills in 5<sup>th</sup> grade. This partly agrees with the results of Ribeiro and colleagues (2016). In their study, no significant relation between non-verbal reasoning and reading comprehension in 2<sup>nd</sup> grade school children were found after controlling for demographic variables and several reading related factors. For reading comprehension in 4<sup>th</sup> grade, however, they found a significant correlation. Similarly, in a study of Tighe and Schatschneider (2014), non-verbal IQ only explained 1% of the variance of reading comprehension in 3<sup>rd</sup>-grade students. Later, in the 10<sup>th</sup> grade, reasoning was the best predictor of reading comprehension, which may indicate that reasoning becomes more important for reading comprehension in later grades. The reasoning factor in their study, however, also included verbal IQ and thus

cannot be directly compared with our results. Finally, we need to consider that the dynamic interrelation between IQ (full scale) and reading skills may also differ between children with and without dyslexia, as shown by Ferrer and colleagues (2009). In their study, IQ was associated with reading skills in children with typical development, but the influence of IQ was minor for children with dyslexia (Ferrer et al. 2009). A direct comparison to our results is, however, difficult because we used non-verbal and not the full IQ measure and studied a sample including both children with and without dyslexia.

#### *4.4 Limitations*

This study contributes to the existing knowledge about early predictors of vocabulary, reading fluency, decoding fluency and reading comprehension by means of preschool HLE, preschoolers' reading interests as well as parental reading skills. Although our results largely correspond to previous studies, there are some limitations to be mentioned. First, our rather restricted sample included mostly families with a middle- to upper-class socio-economic background. Secondly, children with poor reading skills together with children at familial risk for reading impairments are strongly overrepresented in our sample, but the sample is still too small to study differences in subgroups of impaired or non-impaired, at-risk or no-risk children. The current data have been collected in the context of a larger ongoing neuroimaging study on reading acquisition and reading impairments (Karipidis et al. 2018; Mehringer et al. 2020; Pleisch et al. 2019). Thus, the relatively restricted sample, the over-representation of families with a high educational background and a high SES, and the increased number of children with familial risk for dyslexia and poor reading skills can be explained by the study design. It is important to note that the results of the multiple regression coefficients are conditional on the other predictors in the model and do not allow any causative inferences (Vanhove 2021). Therefore, the results of the multiple regression analyses are not directly transferrable to children of low SES families or with low familial risk for reading impairments.

An additional problem in the field is the absence of standardised and comparable questionnaires to assess HLE and children's reading interests. While we had several questions from which to derive HLE, we did not ask for library visits or number of books at home. Such measures are, however, often part of HLE questionnaires (Caglar-Ryeng et al. 2020; Sénéchal & LeFevre 2002). Further, for HLE we only collected informal, meaning-related literacy activities but not formal parental print-related teaching activities, which would have allowed more detailed examination of differential effects of such HLE factors on literacy and vocabulary outcomes (Sénéchal & Lefevre 2014). Finally, as usual with retrospective assessments, it is difficult to judge on how well the parents remembered the HLE situation that took place 5 or 10 years ago.

#### 4.5 Conclusion

In line with previous literature, our study showed that shared storybook reading with both parents at preschool is an important predictor of the long-term development of children's vocabulary skills but has only a weak relation to the development of reading skills. This clearly indicates that despite the differences in how HLE is assessed across studies, shared reading at preschool age is an important factor in a child's environment to develop a rich vocabulary which in turn may support reading comprehension in advanced stages of reading development. Finally, children's reading outcomes are strongly related to parental reading, confirming the strong genetic transmission of these skills and are, furthermore, associated with children's early interest in books and preschool reading skills.

The results of our sample of German-speaking children, with reading skills ranging from severely impaired to excellent, coming from moderate to high SES families and most of them carrying a risk for developmental dyslexia, thus emphasise again previous accounts of the important influence of preschool HLE and reading interests on children's subsequent development of vocabulary and reading skills.

#### 4.6 Acknowledgments

We thank all the families who participated in this study, Simon Foster for the statistical counselling and Daniela Benz, Nadine Kuhn and Lisa Zikeli for their help in data acquisition and analyses. This study was supported by Fondation Botnar (project AllRead).

## REFERENCES

- Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. *Handbook of Early Literacy Research*, 2, 41-51.
- Blaga, O. M., Shaddy, D. J., Anderson, C. J., Kannass, K. N., Little, T. D., & Colombo, J. (2009). Structure and continuity of intellectual development in early childhood. *Intelligence*, 37(1), 106-113. \*
- Bowey, J. A. (2005). Predicting individual differences in learning to read. In M. J. Snowling & C. Hulme (eds.), *The science of reading: A handbook* (pp. 155-172). Oxford: Blackwell. \*
- Bracken, S. S. & Fischel, J. E. (2008). Family reading behavior and early literacy skills in preschool children from low-income backgrounds. *Early Education and Development*, 19(1), 45-67. \*
- Bus, A. G., van IJzendoorn, M. H. & Pellegrini, A. D. (1995). Joint book reading makes for success in learning to read: A meta-analysis on intergenerational transmission of literacy. *Review of Educational Research*, 65(1), 1-21. \*
- Caglar-Ryeng, Ø., Eklund, K. & Nergård-Nilssen, T. (2020). The effects of book exposure and reading interest on oral language skills of children with and without a familial risk of dyslexia. *Dyslexia*, 26(4), 394-410. \*
- Clark, E. V. (2009). *First language acquisition*. Cambridge: Cambridge University Press. \*

- Crain-Thoreson, C. & Dale, P. S. (1992). Do early talkers become early readers? Linguistic precocity, preschool language, and emergent literacy. *Developmental Psychology*, 28(3), 421-429. \*
- de Zeeuw, E. L., de Geus, E. J. C. & Boomsma, D. I. (2015). Meta-analysis of twin studies highlights the importance of genetic variation in primary school educational achievement. *Trends in Neuroscience and Education*, 4(3), 69-76. \*
- Duncan, G. J., Yeung, W. J., Brooks-Gunn, J. & Smith, J. R. (1998). How much does childhood poverty affect the life chances of children? *American Sociological Review*, 63(3), 406-423. \*
- Ebert, S., Lockl, K., Weinert, S., Anders, Y., Kluczniok, K. & Rossbach, H.-G. (2013). Internal and external influences on vocabulary development in preschool children. *School Effectiveness and School Improvement*, 24(2), 138-154. \*
- Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K. & Shaywitz, S. E. (2010). Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. *Psychological Science*, 21(1), 93-101. \*
- Fricke, S., Szczerbinski, M. & Stackhouse, J. (2009). Predicting individual differences in early literacy acquisition in German: The role of speech and language processing skills and letter knowledge. *Written Language and Literacy*, 11(2), 103-146. \*
- Frijters, J. C., Barron, R. W. & Brunello, M. (2000). Direct and mediated influences of home literacy and literacy interest on prereaders' oral vocabulary and early written language skill. *Journal of Educational Psychology*, 92(3), 466-477. \*
- Galuschka, K. & Schulte-Körne, G. (2016). The diagnosis and treatment of reading and/or spelling disorders in children and adolescents. *Deutsches Ärzteblatt International*, 113(16), 279-286. \*
- 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), 1-56. \*
- Goswami, U. (2000). Early phonological development and the acquisition of literacy. *Phonological Development & the Acquisition of Literacy*, 1-38. \*
- Grund, M., Leonhart, R. & Naumann, C. L. (2015). *Diagnostischer Rechtschreibtest für 5. Klassen (DRT 5)* (3., aktual). Göttingen: Hogrefe.
- Hafner, J., Baro, K., Parzer, P. & Resch, F. (2005). *Heidelberger Rechentest (HRT 1-4), Erfassung mathematischer Basiskompetenzen im Grundschulalter*. Göttingen: Hogrefe.
- Hagmann-von Arx, P. & Grob, A. (2014). *RIAS. Reynolds Intellectual Assessment Scales and Screening. Deutschsprachige Adaptation der Reynolds Intellectual Assessment Scales (RIASTM) & des Reynolds Intellectual Screening Test (RISTTM)* von Cecil R. Reynolds und Randy W. Kamphaus. Huber. Göttingen: Hogrefe.
- Hawke, J. L., Olson, R. K., Willcutt, E. G., Wadsworth, S. J. & DeFries, J. C. (2009). Gender ratios for reading difficulties. *Dyslexia*, 15(3), 239-242. \*
- Hayes, D. P. & Ahrens, M. G. (1988). Vocabulary simplification for children: A special case of motherese? *Journal of Child Language*, 15(2), 395-410. \*
- Huber, P. J. (1981). *Robust Statistics*. John Wiley & Sons.
- Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M. & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, 27(2), 236-248. \*
- Johnson, A. D., Martin, A., Brooks-Gunn, J. & Petrill, S. A. (2008). Order in the house! Associations among household chaos, the home literacy environment, maternal reading ability, and children's early reading. *Merrill-Palmer Quarterly*, 54(4), 445-472.
- Karipidis, I. I., Pleisch, G., Brandeis, D., Roth, A., Röthlisberger, M., Schneebeli, M., Walitza, S. & Brem, S. (2018). Simulating reading acquisition: The link between reading outcome and multimodal brain signatures of letter-speech sound learning in prereaders. *Scientific Reports*, 8(1), 7121. \*

- Katzir, T., Lesaux, N. K. & Kim, Y. S. (2009). The role of reading self-concept and home literacy practices in fourth grade reading comprehension. *Reading and Writing*, 22(3), 261-276. \*
- Kauschke, C. (2021). Der Wortschatz im Kindesalter: Entwicklung und Intervention. *Bulletin suisse de linguistique appliquée*, 113, 155-173.
- Kievit, R. A., Lindenberger, U., Goodyer, I. M., Jones, P. B., Fonagy, P., Bullmore, E. T. & Dolan, R. J. (2017). Mutualistic coupling between vocabulary and reasoning supports cognitive development during late adolescence and early adulthood. *Psychological Science*, 28(10), 1419-1431. \*
- Lefly, D. L. & Pennington, B. F. (2000). Reliability and validity of the adult reading history questionnaire. *Journal of Learning Disabilities*, 33(3), 286-296. \*
- Lenhard, A., Lenhard, W., Segerer, R. & Suggate, S. (2015). *Peabody Picture Vocabulary Test - 4. Ausgabe (PPVT 4)*. Frankfurt am Main: Pearson Assessment & Information GmbH.
- Lenhard, W. & Lenhard, A. (2021). Bedeutung und Diagnostik des Wortschatzes am Beispiel des Peabody Picture Vocabulary Test (PPVT-IV). *Bulletin suisse de linguistique appliquée*, 113, 9-27.
- Lenhard, W., Lenhard, A. & Schneider, W. (2018). *Ein Leseverständnistest für Erst- bis Siebtklässler - Version II (ELFE II)* (2. korrig). Göttingen: Hogrefe.
- Leppänen, U., Aunola, K., Niemi, P. & Nurmi, J.-E. (2008). Letter knowledge predicts Grade 4 reading fluency and reading comprehension. *Learning and Instruction*, 18(6), 548-564. \*
- Lervåg, A., Dolean, D., Tincas, I. & Melby-Lervåg, M. (2019). Socioeconomic background, nonverbal IQ and school absence affects the development of vocabulary and reading comprehension in children living in severe poverty. *Developmental Science*, 22(5), e12858. \*
- Levy, B. A., Gong, Z., Hessels, S., Evans, M. A. & Jared, D. (2006). Understanding print: Early reading development and the contributions of home literacy experiences. *Journal of Experimental Child Psychology*, 93(1), 63-93. \*
- Lyytinen, H., Aro, M., Eklund, K., Erskine, J., Guttorm, T., Laakso, M. L., Leppänen, P. H. T., Lyytinen, P., Poikkeus, A. M., Richardson, U. & Torppa, M. (2004). The development of children at familial risk for dyslexia: Birth to early school age. *Annals of Dyslexia*, 54(2), 184-220. \*
- Marulis, L. M. & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research*, 80(3), 300-335. \*
- Mayer, A. (2013). *Test zur Erfassung der phonologischen Bewusstheit und der Benennungsgeschwindigkeit (TEPHOBE)* - Manual. Ernst Reinhardt Verlag.
- Mayringer, H. & Wimmer, H. (2014). *Salzburger Lese-Screening für die Schulstufe 2-9 (SLS 2-9)*. Göttingen: Hogrefe.
- McGinty, A. S. & Justice, L. M. (2009). Predictors of print knowledge in children with specific language impairment: Experiential and developmental factors. *Journal of Speech, Language, and Hearing Research*, 52(1), 81-97. \*
- Mehringer, H., Fraga-González, G., Pleisch, G., Röthlisberger, M., Aepli, F., Keller, V., Karipidis, I. I. & Brem, S. (2020). (Swiss) GraphoLearn: An app-based tool to support beginning readers. *Research and Practice in Technology Enhanced Learning*, 15(1). \*
- Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, 91(1), 3-19. \*
- Moll, K., Hutzler, F. & Wimmer, H. (2005). Developmental dyslexia in a regular orthography: A single case study. *The Neural Basis of Cognition*, 6(11), 433-440. \*
- Moll, K., Kunze, S., Neuhoff, N., Bruder, J. & Schulte-Körne, G. (2014). Specific learning disorder: Prevalence and gender differences. *PLoS ONE*, 9(7). \*
- Moll, K., & Landerl, K. (2014). *Lese- und Rechtschreibtest (SLRT-II). Weiterentwicklung des Salzburger Lese- und Rechtschreibtests (SLRT) [Reading and writing test (SLRT-II). Advanced version of Salzburg's reading and writing test (SLRT)]*. Bern: Verlag Hans Huber.

- Nation, K. & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, 27(4), 342-356. \*
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98(3), 554-566. \*
- Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M. & Hirsh-Pasek, K. (2019). Measuring success: Within and cross-domain predictors of academic and social trajectories in elementary school. *Cross-Domain Development of Academic and Cognitive Skills*, 46, 112-125. \*
- Pennington, B. F. & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development*, 72(3), 816-833. \*
- Petermann, F. & Daseking, M. (2015). *Zürcher Lesetest - II (ZLT-II) Weiterentwicklung des Zürcher Lesetests (ZLT) von Maria Linder und Hans Grissemann 3., überarbeitete Auflage*. Bern: Verlag Hans Huber.
- Petscher, Y. (2010). A meta-analysis of the relationship between student attitudes towards reading and achievement in reading. *Journal of Research in Reading*, 33(4), 335-355. \*
- Pleisch, G., Karipidis, I. I., Brauchli, C., Röthlisberger, M., Hofstetter, C., Stämpfli, P., Walitza, S. & Brem, S. (2019). Emerging neural specialization of the ventral occipitotemporal cortex to characters through phonological association learning in preschool children. *NeuroImage*, 189, 813-831. \*
- Puglisi, M. L., Hulme, C., Hamilton, L. G. & Snowling, M. J. (2017). The home literacy environment is a correlate, but perhaps not a cause, of variations in children's language and literacy development. *Scientific Studies of Reading*, 21(6), 498-514. \*
- Revelle, W. & Revelle, M. W. (2015). Package 'psych'. *The comprehensive R archive network*, Vol. 337, 338.
- Ribeiro, I., Cadime, I., Freitas, T. & Viana, F. L. (2016). Beyond word recognition, fluency, and vocabulary: The influence of reasoning on reading comprehension. *Australian Journal of Psychology*, 68(2), 107-115. \*
- Rice, M. L. & Hoffman, L. (2015). Predicting vocabulary growth in children with and without specific language impairment: A longitudinal study from 2;6 to 21 years of age. *Journal of Speech, Language, and Hearing Research*, 58(2), 345-359. \*
- Rowe, D. C., Jacobson, K. C. & den Oord, E. J. C. G. (1999). Genetic and environmental influences on vocabulary IQ: Parental education level as moderator. *Child Development*, 70(5), 1151-1162. \*
- Rowe, M. L., Raudenbush, S. W. & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development*, 83(2), 508-525. \*
- Rutter, M., Caspi, A., Fergusson, D., Horwood, L. J., Goodman, R., Maughan, B., Moffitt, T. E., Meltzer, H. & Carroll, J. (2004). Sex differences in developmental reading disability: New findings from 4 epidemiological studies. *Journal of the American Medical Association*, 291(16), 2007-2012. \*
- Samuelsson, S., Olson, R., Wadsworth, S., Corley, R., Defries, J. C., Willcutt, E., Hulslander, J. & Byrne, B. (2007). Genetic and environmental influences on prereading skills and early reading and spelling development in the United States, Australia, and Scandinavia. *Reading and Writing*, 20(1-2), 51-75. \*
- Sénéchal, M. (2006). Testing the home literacy model: Parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. *Scientific Studies of Reading*, 10(1), 59-87. \*
- Sénéchal, M. & LeFevre, J. A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development*, 73(2), 445-460. \*
- Sénéchal, M. & Lefevre, J. A. (2014). Continuity and change in the home literacy environment as predictors of growth in vocabulary and reading. *Child Development*, 85(4), 1552-1568. \*



- Sénéchal, M., LeFevre, J. A., Hudson, E. & Lawson, E. P. (1996). Knowledge of storybooks as a predictor of young children's vocabulary. *Journal of Educational Psychology*, 88(3), 520-536. \*
- Sénéchal, M., Lefevre, J.-A., Thomas, E. M. & Daley, K. E. (1998). Differential effects of home literacy experiences on the development of oral and written language. *Reading Research Quarterly*, 33(1), 96-116. \*
- Sénéchal, M., Ouellette, G. & Rodney, D. (2006). *The misunderstood giant: On the predictive role of early vocabulary to future reading*. New York: Guilford Press.
- Silinskas, G., Sénéchal, M., Torppa, M. & Lerkkanen, M. K. (2020). Home literacy activities and children's reading skills, independent reading, and interest in literacy activities from kindergarten to grade 2. *Frontiers in Psychology*, 11. \*
- Skibbe, L. E., Justice, L. M., Zucker, T. A. & McGinty, A. S. (2008). Relations among maternal literacy beliefs, home literacy practices, and the emergent literacy skills of preschoolers with specific language impairment. *Early Education and Development*, 19(1), 68-88. \*
- Snowling, M. J., Gallagher, A. & Frith, U. (2003). Family risk of dyslexia is continuous: Individual differences in the precursors of reading skill. *Child Development*, 74(2), 358-373. \*
- Snowling, M. J., Hayiou-Thomas, M. E., Nash, H. M. & Hulme, C. (2020). Dyslexia and developmental language disorder: Comorbid disorders with distinct effects on reading comprehension. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 61(6), 672-680. \*
- Sparks, A. & Reese, E. (2013). From reminiscing to reading: Home contributions to children's developing language and literacy in low-income families. *First Language*, 33(1), 89-109. \*
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360-407. \*
- Stevens, J. (2002). *Applied multivariate statistics for the social sciences*. Mahwah, NJ: Lawrence Erlbaum Associates. \*
- Stock, C., Marx, P. & Schneid, W. (2017). *BAKO 1-4 - Basiskompetenzen für Leserechtschreibleistungen Ein Test zur Erfassung der phonologischen Bewusstheit vom ersten bis vierten Grundschuljahr 2.*, ergänzte und aktualisierte Auflage. Göttingen: Hogrefe.
- Szagan, G. (2006). Variabilität im frühen Spracherwerb: Normal-nicht pathologisch. *Kinder-Und Jugendarzt*, 55(11), 1-4.
- Team, R. C. (2020). *R* (p. <https://www.R-project.org>). <https://www.r-project.org>
- Tighe, E. L. & Schatschneider, C. (2014). A dominance analysis approach to determining predictor importance in third, seventh, and tenth grade reading comprehension skills. *Reading and Writing*, 27(1), 101-127. \*
- Vaahoranta, E., Lenhart, J., Suggate, S. & Lenhard, W. (2019). Interactive elaborative storytelling: Engaging children as storytellers to foster vocabulary. *Front Psychology*, 10, 1534. \*
- van Bergen, E., Bishop, D., van Zuijen, T. & de Jong, P. F. (2015). How does parental reading influence children's reading? A study of cognitive mediation. *Scientific Studies of Reading*, 19(5), 325-339. \*
- van Bergen, E., de Jong, P. F., Plakas, A., Maassen, B. & van der Leij, A. (2012). Child and parental literacy levels within families with a history of dyslexia. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 53(1), 28-36. \*
- van Bergen, E., van Zuijen, T., Bishop, D. & de Jong, P. F. (2017). Why are home literacy environment and children's reading skills associated? What parental skills reveal. *Reading Research Quarterly*, 52(2), 147-160. \*
- Vanhove, J. (2021). Towards simpler and more transparent quantitative research reports. *ITL - International Journal of Applied Linguistics*, 172(1), 3-25. \*
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J. & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 1(45), 2-40. \*

- Verhoef, E., Shapland, C. Y., Fisher, S. E., Dale, P. S. & St Pourcain, B. (2021). The developmental origins of genetic factors influencing language and literacy: Associations with early-childhood vocabulary. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 62(6), 728-738. \*
- Wadsworth, S. J., Corley, R. P., Hewitt, J. K., Plomin, R. & DeFries, J. C. (2002). Parent-offspring resemblance for reading performance at 7, 12 and 16 years of age in the Colorado Adoption Project. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43(6), 769-774. \*
- Wechsler, D. (2011). *Wechsler Intelligence Scale for Children - Fourth Edition (WISC-IV)*, Deutsche Bearbeitung von F. Petermann / U. Petermann. Frankfurt am Main: Pearson Assessment & Information GmbH.
- Weisleder, A. & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143-2152. \*
- Whitehurst, G. J. & Lonigan, C. J. (1998). Child development and emergent literacy. *Child Development*, 69(3), 848-872. \*

\* These references are available online.

## Appendix

### Figure A1

Significant correlations (surviving correction for multiple comparisons) were found between children's vocabulary and reading comprehension, children's vocabulary and non-verbal IQ, children's vocabulary and *shared reading father & preschool reading*, reading fluency and decoding fluency, reading fluency and reading comprehension, decoding fluency and decoding fluency of the mother and of the father, reading fluency and *early interest in book*, decoding fluency and reading comprehension, decoding fluency and decoding fluency of the father, decoding fluency and *early interest in books*, decoding fluency and *shared reading father & preschool reading*, reading comprehension and non-verbal IQ, reading comprehension and *early interest in books*, reading comprehension and *shared reading father & preschool reading*, SES and non-verbal IQ, non-verbal IQ and *shared reading father & preschool reading* and between mothers decoding fluency and *early interest in books*. The scatter plots were produced in R (R Core Team 2020) with the package psych, function pairs.panels (Revelle & Revelle 2015).

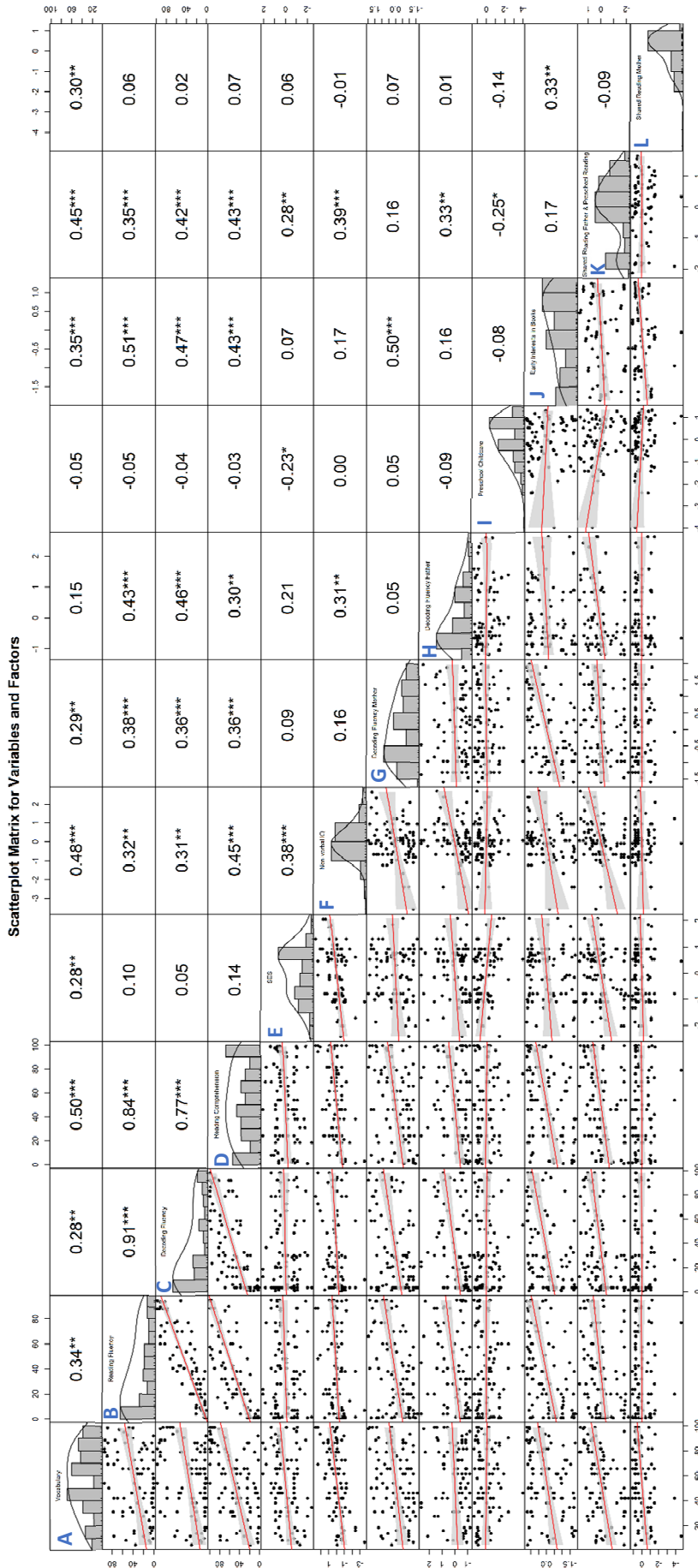


Figure A1: On the diagonal, a histogram of each variable shows the data distribution. Scatter plots of the relationship between each pair of variables can be found below the diagonal with x-axis scaled to the column variable and y-axis to the row variable. Spearman's Rho values and their statistical significances are given in the corresponding boxes on the right hand side; \*p<0.05, \*\*p<0.01, \*\*\*p<0.001. Labelling (blue letters) of variables is as follows: A: vocabulary; B: reading fluency; C: decoding fluency; D: reading comprehension; E: SES; F: non-verbal IQ; G: decoding fluency mother; H: decoding fluency father; I: preschool childcare; J: early interest in books; K: shared reading father & preschool reading; L: shared reading mother

**Figure A2**

*Vocabulary: Partial regression plots of the factors predicting vocabulary in 5th grade school children.*

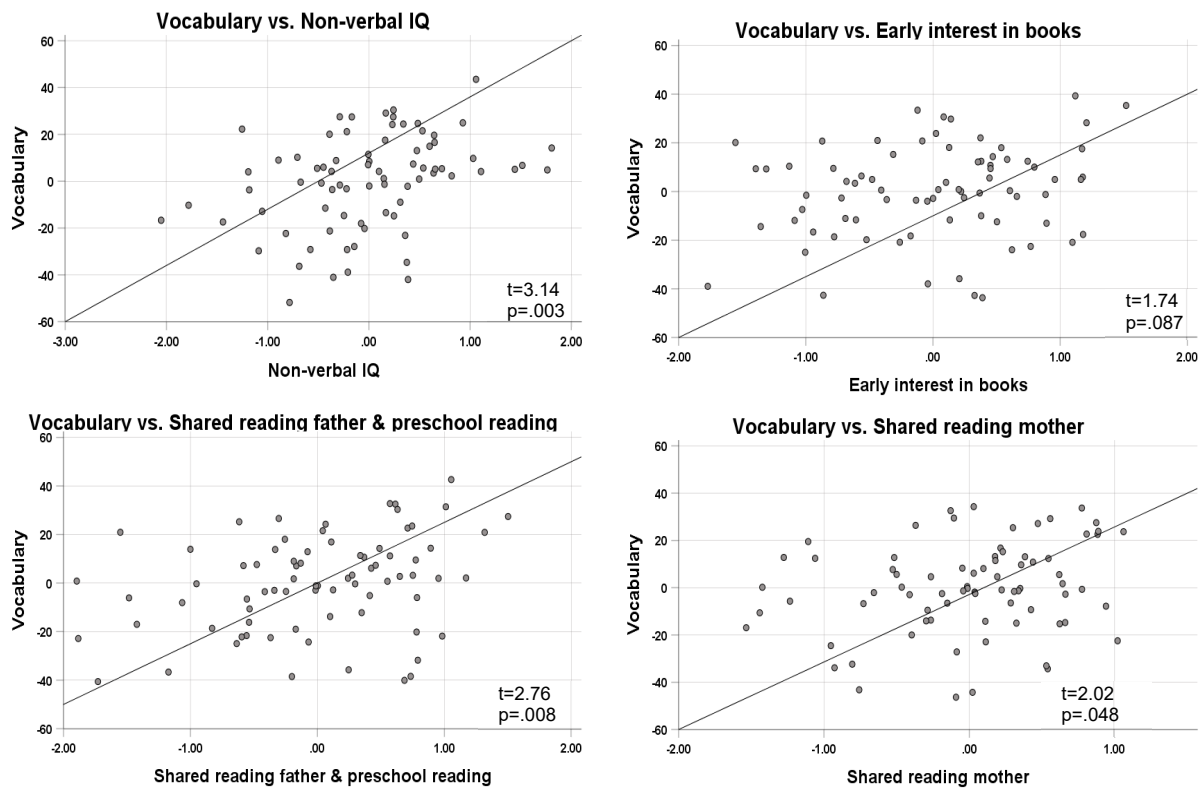


Figure A2: Partial regression plots of the multiple regression model to explain vocabulary

**Figure A3**

*Reading fluency: Partial regression plots of the factors predicting reading fluency in 5th grade school children.*

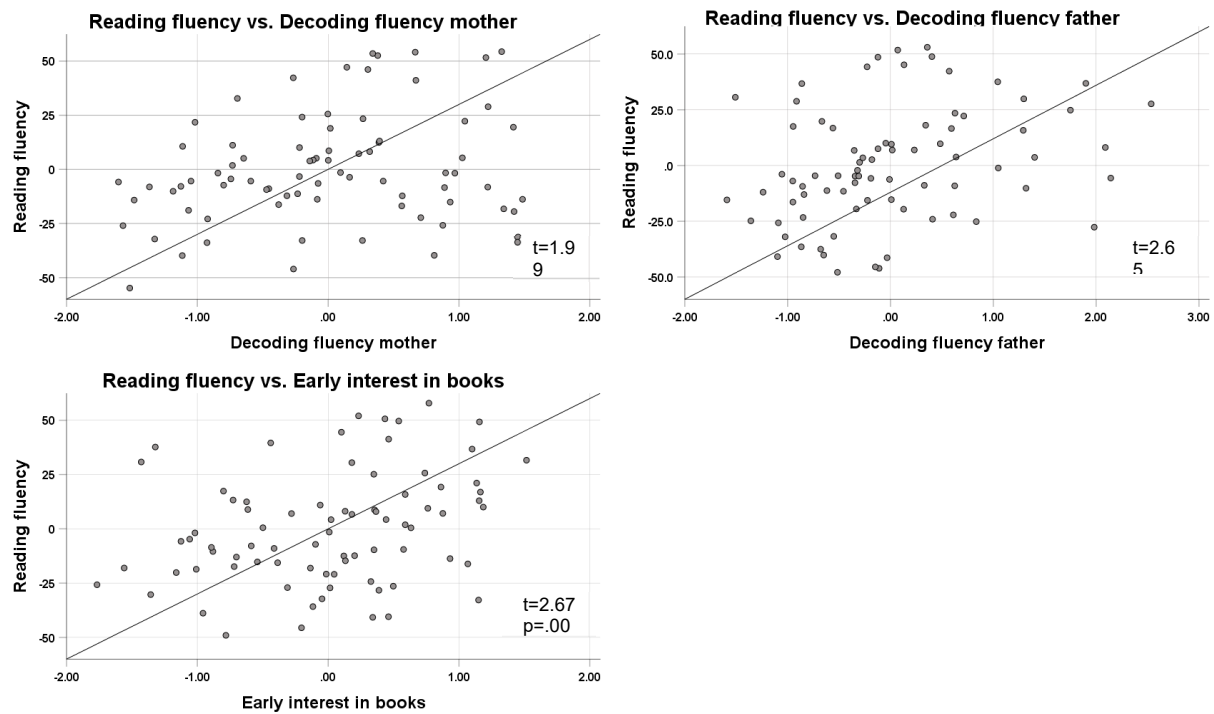


Figure A3: Partial regression plots of the multiple regression model to explain reading fluency

## Figure A4

*Decoding fluency: Partial regression plots of the factors predicting decoding fluency in 5th grade school children.*

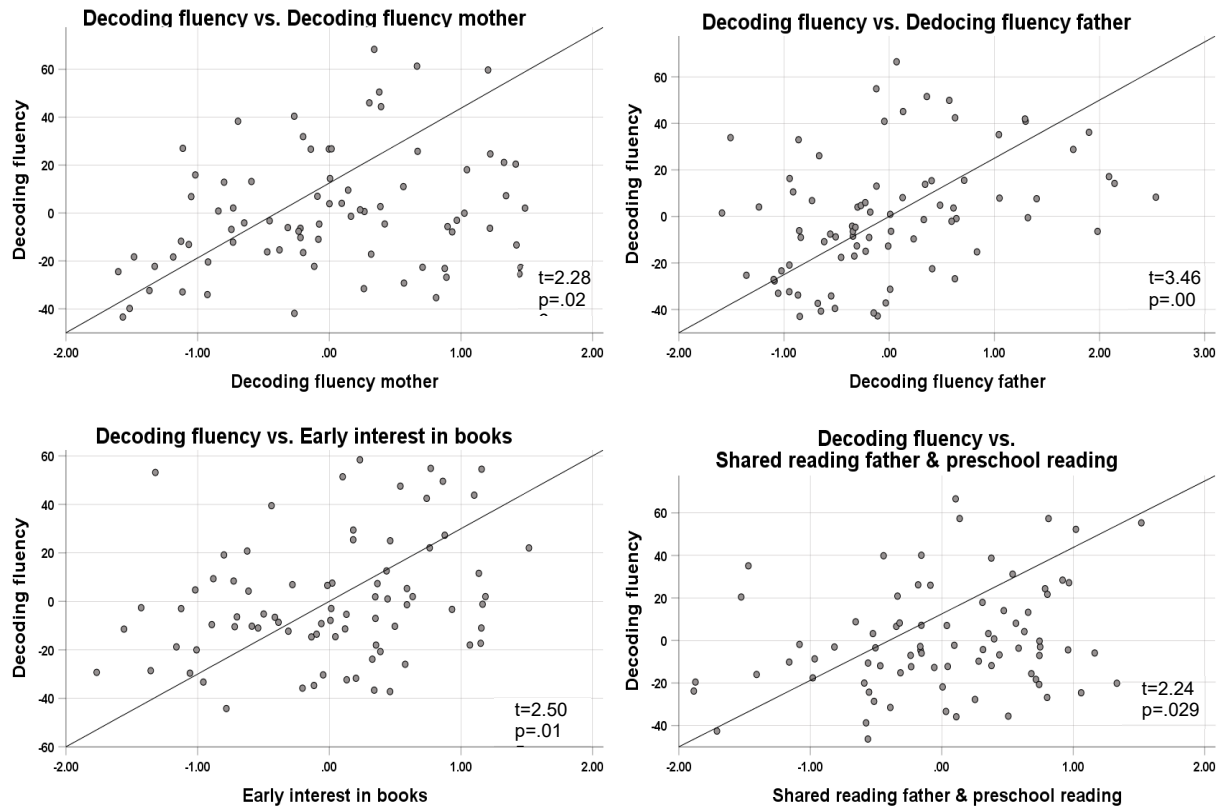


Figure A4: Partial regression plots of the multiple regression model to explain decoding fluency

## Figure A5

*Comprehension: Partial regression plots of the factors predicting reading comprehension in 5th grade school children.*

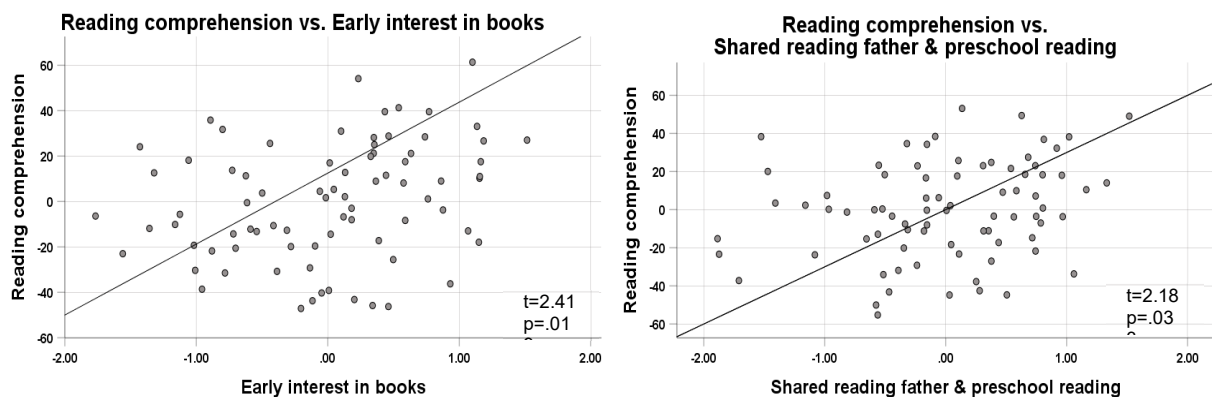


Figure A5 Partial regression plots of the multiple regression model to predict reading comprehension

**Table A1**

*Compilation of questions from the two questionnaires used in this study.*

<b>Vorschulalter (0-4) Jahre</b>					
Hat sich Ihr Kind speziell für (Bilder-) Bücher interessiert?			<input type="radio"/> Ja <input type="radio"/> Nein		
Hatte Ihr Kind ein Lieblingsbuch? (ein Buch/mehrere Bücher, welche das Kind jeden Tag/mehrmals in der Woche lesen möchte?)			<input type="radio"/> Ja <input type="radio"/> Nein		
<b>Kindergarten (4-6 Jahre)</b>					
Hat sich Ihr Kind speziell für Bücher interessiert?			<input type="radio"/> Ja <input type="radio"/> Nein		
Hatte Ihr Kind ein Lieblingsbuch? (ein Buch/mehrere Bücher, welche das Kind jeden Tag/mehrmals in der Woche lesen möchte?)			<input type="radio"/> Ja <input type="radio"/> Nein		
<b>Vorschulalter (0-4) Jahre</b>					
Wie oft haben Sie (Mutter) mit Ihrem Kind gemeinsam Bücher angeschaut/vorgelesen/erzählt?			<input type="radio"/> Selten <input type="radio"/> Monatlich <input type="radio"/> Wöchentlich <input type="radio"/> Täglich		
<b>Kindergarten (4-6 Jahre)</b>					
Wie oft haben Sie (Mutter) mit Ihrem Kind gemeinsam Bücher angeschaut/vorgelesen/erzählt?			<input type="radio"/> Selten <input type="radio"/> Monatlich <input type="radio"/> Wöchentlich <input type="radio"/> Täglich		
<b>Vorschulalter (0-4) Jahre</b>					
Wie oft haben Sie (Vater) mit Ihrem Kind gemeinsam Bücher angeschaut/vorgelesen/erzählt?			<input type="radio"/> Selten <input type="radio"/> Monatlich <input type="radio"/> Wöchentlich <input type="radio"/> Täglich		
<b>Kindergarten (4-6 Jahre)</b>					
Wie oft haben Sie (Vater) mit Ihrem Kind gemeinsam Bücher angeschaut/vorgelesen/erzählt?			<input type="radio"/> Selten <input type="radio"/> Monatlich <input type="radio"/> Wöchentlich <input type="radio"/> Täglich		
Konnte Ihr Kind im Kindergarten schon selber einfache Texte lesen?			<input type="radio"/> Ja <input type="radio"/> Nein		
Wie oft stellte Ihr Kind Fragen beim gemeinsamen Lesen?			<input type="radio"/> Nie/kaum <input type="radio"/> Eher selten <input type="radio"/> Manchmal <input type="radio"/> Häufig <input type="radio"/> (fast) immer		
Wie oft redet Ihr Kind während dem gemeinsamen Vorlesen/Bücher anschauen?			<input type="radio"/> Nie/kaum <input type="radio"/> Eher selten <input type="radio"/> Manchmal <input type="radio"/> Häufig <input type="radio"/> (fast) nie		
<b>Ausbildung der Eltern</b>					
Beruf Mutter (Bei Hausfrau/-mann: Zuletzt ausgeübte Erwerbstätigkeit)					
Beruf Vater (Bei Hausfrau/-mann: Zuletzt ausgeübte Erwerbstätigkeit)					
<b>Zeitliche Aufteilung der Kinderbetreuung (ausserhalb der Fremdbetreuungszeiten)</b>					
Alter Kind zwischen 0-2 Jahre					
Mutter	95-100%	75%	50%	25%	0-5%
Vater	95-100%	75%	50%	25%	0-5%
Alter Kind zwischen 2-4 Jahre					
Mutter	95-100%	75%	50%	25%	0-5%
Vater	95-100%	75%	50%	25%	0-5%
Alter Kind zwischen 4-6 Jahre (Kindergarten)					
Mutter	95-100%	75%	50%	25%	0-5%
Vater	95-100%	75%	50%	25%	0-5%

Table A1: Overview of the questions used in this study in the two questionnaires



# Die Bedeutung des mentalen Lexikons beim Wortlesen und -schreiben: Einsichten anhand der semantisch-lexikalischen Verarbeitungskompetenzen von Kindern mit und ohne Schriftspracherwerbsstörungen

**Julia WINKES**

Universität Freiburg

Departement für Sonderpädagogik

Petrus-Kanisius-Gasse 21, CH-1700 Freiburg

julia.winkes@unifr.ch

Models of the reading and spelling process at the word level focus primarily on phonological and orthographic processing. The fact that the mental lexicon is the central processing unit, which is both the starting point (writing) and the goal (reading) of the cognitive processes involved, is easily forgotten. In this study, the semantic-lexical processing skills of sixth graders with average reading and spelling skills (N=30), with reading and spelling disorders (N=18) and with isolated spelling disorders (N=19) are compared. These processing skills were assessed on vocabulary, morphemic awareness, recall fluency and rapid naming ability (RAN). Differences between the groups are illustrated using single factor analysis of variance (ANOVA) and post-hoc analyses. Correlations between word reading and spelling performance and semantic-lexical processing skills in students without literacy disorders complete the emerging. The results underline the impact of rapid naming on the reading process and of word-specific orthographic knowledge, vocabulary and morphemic awareness on spelling. They provide evidence for the importance of the quality of lexical representations and the automated retrieval of these in written language processing.

## **Keywords:**

dyslexia, isolated spelling disorder, semantic-lexical processing, lexical quality.

## **Stichwörter:**

Lese-Rechtschreibstörung, Isolierte Rechtschreibstörung, semantisch-lexikalische Verarbeitung, lexikalische Qualität.

## **1. Einführung**

Die beiden sprachlichen Modalitäten Lesen und Schreiben sind bei vielen erwachsenen Personen hoch automatisiert. Gelingt der Schriftspracherwerb in der Schule nicht mühelos, sondern ist mit erheblichen Schwierigkeiten verbunden, dann wird die Komplexität der beteiligten Sprachverarbeitungsprozesse offensichtlich. Wissen über die Lautstruktur (phonologisch), die Graphemfolge (orthographisch) und die Bedeutung (semantisch) von Wörtern muss gespeichert, abgerufen, kurzfristig aufrechterhalten und ggf. manipuliert werden, wenn Wörter gelesen oder geschrieben werden (Coltheart 2006). Ziel des vorliegenden Beitrages ist es, die Relevanz lexikalischer Verarbeitungsprozesse beim Wortlesen und -schreiben aufzuzeigen. Der Zugang zur Thematik erfolgt über den Vergleich dieser Kompetenzen bei Schülerinnen und



Schülern mit Lese- Rechtschreibstörung, mit Isolierter Rechtschreibstörung und von Gleichaltrigen mit unauffälligen Lese- Rechtschreibleistungen. Mögliche Unterschiede zwischen diesen drei Gruppen mit verschiedenen Profilen im Lesen und Rechtschreiben tragen zum Verständnis des Stellenwertes von Umfang, Struktur und Abrufqualität des mentalen Lexikons bei der Verarbeitung von geschriebener Sprache bei.

## **2. Lexikalische Verarbeitungsprozesse beim Lesen und Schreiben**

Herzstück aller Modellvorstellungen zum Wortlesen und -schreiben – allen voran der Zwei-Wege-Modelle – sind drei Teilbereiche des mentalen Lexikons: das modalitätsunabhängige semantische System, in dem Informationen über die Bedeutungen von Wörtern gespeichert sind, das phonologische Lexikon mit Wissen über die lautliche Struktur von Wörtern und das orthographische Lexikon mit Wissen über die Schreibweisen von Wörtern (Miceli & Capasso 2006). Somit stellt das mentale Lexikon mit seinen verschiedenen Subkomponenten sowohl Ausgangspunkt (Schreiben) als auch Ziel (Lesen) der involvierten kognitiven Prozesse dar. Letztere werden häufig in die Oberbegriffe der phonologischen, orthographischen und semantisch-lexikalischen Verarbeitung unterteilt, worunter nach Wagner & Torgesen (1987) ganz allgemein die Nutzung von phonologischen bzw. orthographischen bzw. semantisch lexikalischen Informationen bei der Verarbeitung von geschriebener und gesprochener Sprache verstanden wird. Viele (wenn auch nicht alle) dieser Verarbeitungsprozesse lassen sich auch unabhängig vom Lesen und Schreiben durch mündliche Aufgabenstellungen operationalisieren, was sich insbesondere für Studien mit Kindern mit Schriftspracherwerbsschwierigkeiten eignet. Auf diese Weise lässt sich weitgehend ausschliessen, dass Defizite in den untersuchten Verarbeitungsprozessen lediglich die Lese- und Rechtschreibprobleme der Schülerinnen und Schüler widerspiegeln. Im Folgenden stehen die Verarbeitungsprozesse im Vordergrund, welche in einem direkten Zusammenhang zum mentalen Lexikon stehen, andere ebenfalls bedeutsame Prozesse (z.B. phonologische Bewusstheit, phonologisches Arbeitsgedächtnis, implizites orthographisches Wissen) werden an dieser Stelle bewusst ausgeklammert (siehe aber Winkes 2014 für eine vollständige Analyse).

### **2.1 Wortschatz**

Die Relevanz des Wortschatzes für das Leseverständnis ist unmittelbar einsichtig, es kann schliesslich ein Verständnis eines Textes nur dann erreicht werden, wenn die Bedeutung der einzelnen Wörter bekannt ist. Aber auch beim Erwerb von Wortlese- und Schreibfähigkeiten findet eine gegenseitige Beeinflussung mit dem Wortschatzumfang und -erwerb statt: zum einen hilft Wissen über die Schreibweisen von Wörtern dabei, diese Wörter neu zu erlernen und zum anderen ist Wissen über Wortbedeutungen und Wortformen hilfreich beim Erwerb von orthographischen Repräsentationen (Ouellette 2010).

Beispielsweise erwiesen sich verbal-semantische Kompetenzen (u.a. Wortschatz, semantische Wortflüssigkeit) in der Untersuchung von Nation & Snowling (2004) als bedeutsame Prädiktoren für die Leistung im Worterkennen bei achtjährigen Kindern, selbst unter statistischer Kontrolle der Dekodierfähigkeit. Verschiedene empirische Hinweise sprechen für einen grösseren Einfluss des Wortschatzes in den früheren Phasen des Leseerwerbs und eine Abnahme des Zusammenhangs zwischen semantisch-lexikalischen Kompetenzen und der Worterkennung mit zunehmendem Alter (z.B. Berendes et al. 2010; Nagy et al. 2003). In Bezug auf die Rechtschreibung fanden Berendes et al. (2010) bei Drittklässlern einen positiven Zusammenhang mit dem rezeptiven Wortschatz ( $r=.51$ ), welcher in der vierten Klassenstufe nicht auftrat ( $r=-.05$ ). Entsprechend stellte der Wortschatz bei den jüngeren Kindern eine Prädiktorvariable für das Rechtschreiben dar, bei den älteren hingegen nicht. In englischsprachigen Studien wurde in verschiedenen Altersstufen ein robuster Zusammenhang zwischen Rechtschreibung und Wortschatz dokumentiert (Burt & Fury 2000).

## *2.2 Orthographische Repräsentationen / wortspezifisches orthographisches Wissen*

Eine wichtige Komponente in allen Modellen zum Wortlesen und -schreiben ist das orthographische Lexikon, in welchem die Graphemstrukturen von spezifischen Wörtern gespeichert sind. Die Nutzung dieser orthographischen Repräsentationen versetzt geübte Leser in die Lage, Wörter hoch automatisiert zu erkennen. Die rapide und mühelose Worterkennung beim Lesen setzt kognitive Ressourcen frei, welche für das Textverständnis genutzt werden können. Für das korrekte Rechtschreiben sind gut spezifizierte orthographische Repräsentationen besonders wichtig, weil in der komplexen deutschen Orthographie häufig verschiedene potentielle Möglichkeiten existieren, ein Wort lautgetreu zu verschriften (z.B. Fuchs vs. Fux vs. Fuks) und die richtige Schreibweise auch durch Rückgriff auf Rechtschreibregeln oder morphematische Ableitungsstrategien nicht immer mit Sicherheit korrekt generiert werden kann. Es kann davon ausgegangen werden, dass geübte Leser und Schreiber sich zum grössten Teil auf lexikalische Verarbeitungsstrategien unter Rückgriff auf das orthographische Lexikon stützen (Moll & Landerl 2009) und dass andere Wissensquellen (z.B. Anwendung von Phonem-Graphem-Korrespondenzen, Nutzung von Analogien oder von Rechtschreibregeln) nur dann zur Generierung einer Schreibweise genutzt werden, wenn unbekannte Wörter verschriftet werden oder sich der Schreibende unsicher ist (Burt & Fury 2000).

## *2.3 Morphematische Bewusstheit*

Die deutsche Orthographie wird neben dem alphabetischen Prinzip sehr wesentlich durch das morphematische Prinzip geprägt, welches besagt, dass Morpheme immer gleich verschriftet werden, auch in Fällen, in denen dies dem

Lautprinzip widerspricht. Beispiele für morphematisch bedingte Schreibweisen sind die Auslautverhärtung (<Hund> statt <Hunt>), die Vokableitung (<Bäcker> statt <Becker> wegen /backen/) oder die Schreibung von Komposita (<Fahrrad> statt <Fahrad>). Morphematisches Wissen beeinflusst den Lese- und Rechtschreibprozess in mehrfacher Hinsicht. Eine gut ausgebildete morphematische Bewusstheit fördert das Verständnis für die Strukturen und Regelmäßigkeiten des orthographischen Systems (Bahr et al. 2020; Nagy et al. 2003). Insbesondere bei der Verarbeitung von längeren, morphematisch komplexen Wörtern unterstützt eine morphematische Segmentierung den Lesefluss und die korrekte Rechtschreibung. "Perceiving morphologically defined word parts (e.g., thanks in thanksgiving) rather than unanalyzed wholes (e.g., thanksgiving) may create bridges among multiple word forms in the lexicon, resulting in improved accuracy and rate of word reading, spelling, and reading comprehension" (Nagy et al. 2003: 731). Die Zuordnung zu Wortfamilien erleichtert zudem den Aufbau von orthographischen Repräsentationen (Burt & Fury 2000). Eine Vielzahl von Untersuchungen aus dem englischsprachigen Raum belegt erwartungsgemäss die Bedeutung von morphematischen Kompetenzen als Wissensquelle im Rechtschreibprozess, unabhängig von phonologischen Fähigkeiten (Larkin & Snowling 2008). Die Entwicklung von morphematischen Rechtschreibstrategien erstreckt sich über einen längeren Zeitraum und es dauert trotz expliziter Instruktion häufig bis in die Adoleszenz an, bis morphologisches Wissen systematisch rekonstruiert, assimiliert und angewendet werden kann (Pacton & Deacon 2008).

#### *2.4 Schnellbenennen: Rapid Automated Naming (RAN)*

Unter dem Begriff der Schnellbenennung werden Aufgabenstellungen bezeichnet, bei denen eine Reihe von Bildern hochfrequenter Items (z.B. Farben, Objekte, Zahlen) so schnell wie möglich benannt werden sollen. Die meisten Autoren fassen RAN als Komponente der phonologischen Verarbeitung auf und gehen davon aus, dass durch RAN-Tests die Geschwindigkeit gemessen wird, mit der eine Person auf Basis eines visuellen Reizes den entsprechenden phonologischen Code aus dem mentalen Lexikon abrufen kann. Diese Zugriffsgeschwindigkeit auf phonologische Informationen gilt – neben der Speicherqualität und der Verarbeitungsgenauigkeit – als wichtige Bedingung für die Automatisierung der Schriftsprachverarbeitung (Hagiliassis et al. 2006; Vukovic & Siegel 2006).

In der deutschen Orthographie ist RAN der bedeutsamste Prädiktor für die Leseflüssigkeit von Wörtern und Pseudowörtern und es zeigte sich, dass seine Erklärungskraft unabhängig ist von der phonologischen Bewusstheit (z.B. Moll & Landerl 2009). Zum Zusammenhang zwischen RAN und Rechtschreibung kristallisiert sich über verschiedene Studien hinweg heraus, dass das Schnellbenennen zwar einen kleinen Teil der Varianz im Rechtschreiben erklärt,

die phonologische Bewusstheit aber ein deutlich stärkerer Prädiktor für die Rechtschreibleistung ist (z.B. Furnes & Samuelsson 2011; Moll et al. 2009).

## 2.5 Abrufflüssigkeit

In der kognitiven Psychologie und in der neuropsychologischen Diagnostik werden häufig Aufgaben zur verbalen Flüssigkeit eingesetzt. Die Anweisung lautet dabei, innerhalb eines bestimmten Zeitraumes (meist 1-2 Minuten) so viele Wörter wie möglich zu einem bestimmten semantischen (z.B. Tiere, Lebensmittel) oder phonologischen (z.B. Wörter mit dem Anfangslaut /m/) Kriterium zu nennen. Viele Verfahren enthalten auch Aufgaben zum Kategorienwechsel. Diese Aufgabenstellungen evozieren strategische Suchmechanismen und erfordern den Einsatz exekutiver Funktionen (kognitive Flexibilität, Shifting) (da Silva et al. 2004; Sauzéon et al. 2004). Abrufflüssigkeit und RAN beziehen sich beide auf Abrufprozesse aus dem mentalen Lexikon. Die Art der Verarbeitung unterscheidet sich jedoch in wesentlicher Hinsicht: Während beim RAN (ebenso wie beim Lesen) ein hochgradig automatisierter Abruf auf Basis eines visuellen Hinweisreizes erfolgt, beinhalten Aufgaben zur verbalen Flüssigkeit einen freien Gedächtnisabruf, der nur unter Einsatz strategischer Such- und Abrufprozesse gelingen kann.

Tabelle 1 fasst zusammen, welche Verarbeitungsprozesse mit Zusammenhang zum mentalen Lexikon beim (Wort-) Lesen und -schreiben involviert sind.

Fachliche Bezeichnung	Art der Verarbeitung	Kognitive Erklärung	Operationalisierung (Verhaltensebene)
RAN/RAS Schnellbenennung	phonologisch	Geschwindigkeit des Zugriffs auf phonologische Codes im mentalen Lexikon	Schnelles Benennen von Zahlen, Objekten, Buchstaben, Farben
Orthographische Repräsentationen	orthographisch	Gespeichertes wortspezifisches Wissen über Graphemfolgen im mentalen Lexikon	Differenzierung von Homophonen oder Pseudohomophonen
Wortschatz	semantisch-lexikalisch	Quantität der Einträge im mentalen Lexikon	Wortschatztests
Morphematische Bewusstheit	semantisch-lexikalisch	Bewusstheit über die morphematische Struktur der Sprache; Operieren mit morphematischen Einheiten	z.B. Wortverwandtschaften erkennen; Aufgaben zur Komposition, Flexion, Derivation
Abrufflüssigkeit	semantisch-lexikalisch	Qualität des Zugriffs auf das mentale Lexikon	Wortabruf nach semantischen oder phonologischen Kriterien

Tabelle 1: beim Wortlesen/-schreiben beteiligte Verarbeitungsprozesse mit Bezug zum mentalen Lexikon

Ein Ansatz, welcher die Bedeutung von Speicherung, innerer Strukturierung des mentalen Lexikons und flexiblen, automatisierten Abrufprozessen für das Lesen und Schreiben modelliert, ist die lexikalische Qualitätshypothese (LQH) von Perfetti (Perfetti 2007; Perfetti & Hart 2002). Diese besagt, dass lexikalische Repräsentationen aus einem Triplet von Wortbedeutung, phonologischen und orthographischen Merkmalen bestehen. Repräsentationen mit einer hohen lexikalischen Qualität sind voll spezifiziert und redundant. Sie beinhalten somit die vollständige Buchstabensequenz des Wortes und weisen enge Verbindungen zwischen orthographischen und phonologischen Merkmalen auf verschiedenen Ebenen auf (z.B. Phoneme, Silben, Onset/Reim-Einheiten, Morpheme, Ganzwörter). Eine kohärente Repräsentation von hoher lexikalischer Qualität wird stabil, zuverlässig und synchron aus dem mentalen Lexikon abgerufen. Perfetti & Hart (2002) gehen davon aus, dass Menschen sich hinsichtlich der durchschnittlichen Qualität ihrer lexikalischen Repräsentationen unterscheiden, dass es aber auch intraindividuell grosse Differenzen zwischen der Qualität der Repräsentationen verschiedener Wörter gibt. Dies kann beispielsweise bedeuten, dass nicht alle Grapheme eines Wortes sicher gespeichert sind. Die lexikalische Qualität von Wörtern bedingt die Genauigkeit und Geschwindigkeit des Wortabrufs und beeinflusst, wie gross die notwendigen kognitiven Ressourcen für die Verarbeitung sind. Dieser Effekt wird als Effizienz bezeichnet (Perfetti 2007). Kompetente Leser und Schreiber verfügen über eine grosse Anzahl von lexikalischen Repräsentationen von hoher Qualität, die auch das weitere lexikalische Lernen unterstützen.

### **3. Lese- Rechtschreibstörung (LRS) und Isolierte Rechtschreibstörung (IR)**

Störungen des Schriftspracherwerbs gehören zu den häufigsten Lernstörungen und betreffen ca. 7-8% der Kinder eines Jahrgangs (Hasselhorn & Schuchardt 2006). Die Diskussionen über Definition, Ursachen und die diagnostischen Kriterien der LRS halten unvermindert an. Als internationaler Konsens kann derzeit am ehesten die Definition der International Dyslexia Association betrachtet werden:

Dyslexia is a specific learning disability that is neurobiological in origin. It is characterized by difficulties with accurate and/or fluent word recognition and by poor spelling and decoding abilities. These difficulties typically result from a deficit in the phonological component of language that is often unexpected in relation to other cognitive abilities and the provision of effective classroom instruction. Secondary consequences may include problems in reading comprehension and reduced reading experience that can impede growth of vocabulary and background knowledge (Lyon et al. 2003: 3).

Im deutschen Sprachraum orientiert sich die diagnostische Praxis in der Regel an den Kriterien der ICD-10 der Weltgesundheitsorganisation (WHO), welche ein doppeltes (bzw. dreifaches) Diskrepanzkriterium beinhalten (Dilling & Schulte-Markwort 2014). Dabei muss die Lese-Rechtschreibleistung erwartungswidrig sein in Bezug auf das Alter, die allgemeine Intelligenz und die

Beschulung. In der Zwischenzeit hat eine Vielzahl von Forschungsarbeiten aufzeigen können, dass Schriftspracherwerbsstörungen unabhängig vom IQ operationalisiert werden sollten (für eine Übersicht siehe Snowling & Hulme 2012). Neben der Lese-Rechtschreibstörung führt die ICD-10 auch das Störungsbild der Isolierten Rechtschreibstörung (IR) auf. Deren Hauptmerkmal besteht in einer umschriebenen und eindeutigen Beeinträchtigung in der Entwicklung von Rechtschreibfertigkeiten, ohne Vorgeschichte einer umschriebenen Lesestörung (Dilling & Schulte-Markwort 2014). In der Tat konnte in verschiedenen Studien, insbesondere im deutschsprachigen Raum, gezeigt werden, dass Lesen und Schreiben auch unabhängig voneinander von einer Entwicklungsstörung betroffen sein können (Moll & Landerl 2009).

#### **4. Lexikalische Verarbeitungsprozesse bei Kindern mit Schriftspracherwerbsstörungen**

LRS lässt sich nicht als Entwicklungsverzögerung erklären, sondern ist durch inhomogene Entwicklungen in den zugrundeliegenden Verarbeitungsprozessen gekennzeichnet (Romonath & Gregg 2003). Die Eigenschaften des orthographischen Systems beeinflussen weiterhin, wie sich ein bestimmtes Verarbeitungsprofil auf die Symptomatik einer Schriftspracherwerbsstörung auswirkt. Im Zusammenspiel mit der komplexen englischen Orthographie gilt ein phonologisches Sprachverarbeitungsdefizit als zentrale Ursache von LRS, welches sich insbesondere durch Schwierigkeiten in der phonologischen Bewusstheit manifestiert (Pennington et al. 2001). Im deutschsprachigen Raum ist ein solches bei älteren Kindern mit LRS oft kaum noch nachweisbar (Landerl 2001). Probleme im RAN scheinen dafür ein Hauptcharakteristikum von Schülerinnen und Schülern mit Leseschwierigkeiten in transparenten Orthographien darzustellen (Moll & Landerl 2011). Im Bereich der semantisch-lexikalischen Verarbeitung inklusive des orthographischen Lexikons stellt sich die Forschungslage uneindeutig dar. Im Vergleich zu Altersgenossen weisen Kinder mit LRS hier ebenfalls deutliche Defizite auf. Besteht die Vergleichsgruppe aber aus jüngeren Kindern mit demselben Lesealter, so zeigen sich in einigen Untersuchungen Unterschiede (z.B. Hauerwas & Walker 2003), in anderen nicht (z.B. Bourassa & Treiman 2008). Mündliche Sprachschwierigkeiten sind ein wesentlicher Risikofaktor für die Entwicklung einer LRS und entsprechend ist der Wortschatzumfang bei Kindergartenkindern ein guter Prädiktor für die spätere Lesefähigkeit (Scarborough 1990). Auf der anderen Seite können schlechtere lexikalische Kompetenzen bei Schülerinnen und Schülern mit LRS auch eine Folge von weniger Leseerfahrung darstellen, da das selbständige Lesen ab der dritten Klasse als wichtigstes Mittel zur Wortschatzerweiterung gilt (McGregor 2004).

In Bezug auf die IR ist der derzeitige Wissensstand nicht nur uneindeutig, sondern gar widersprüchlich. Während einige Studien IR ebenfalls als Folge

eines phonologischen Verarbeitungsdefizits ansehen (z.B. Moll & Landerl 2009, 2011), weisen andere auf ein spezifisches orthographisches Verarbeitungsdefizit hin, welches sich in fehlendem wortspezifischen orthographischen Wissen äussert (z.B. Frith 1980; Harrison 2004). Schülerinnen und Schülern mit IR werden häufig sehr gute sprachliche Fähigkeiten attestiert (Romani et al. 1999), was beispielsweise durch den Vergleich von Handlungs-IQ und verbalem IQ oder durch Wortschatz- und Hörverständnistests erhoben wurde. Es bleibt zu bilanzieren, dass die kognitiven Verarbeitungskompetenzen, die im Zusammenspiel mit der Tiefe des orthographischen Systems zu einer spezifischen Symptomatik von beeinträchtigten Lese- und/oder Rechtschreibprozessen führen, noch nicht vollständig verstanden sind. Daher fokussiert die folgend dargestellte Untersuchung auf die Bedeutung semantisch-lexikalischer Kompetenzen bei Kindern mit und ohne Schriftspracherwerbsstörungen.

## 5. Empirische Studie

### 5.1 Fragestellungen

Vor dem skizzierten theoretischen Hintergrund können nun folgende empirische Fragestellungen konkretisiert werden:

Unterscheiden sich Kinder mit IR, LRS und schriftsprachunauffällige Gleichaltrige hinsichtlich Wortschatz, wortspezifischem orthographischem Wissen, morphematischer Bewusstheit, RAN und Abrufflüssigkeit?

Welche Zusammenhänge gibt es zwischen dem Lesen und dem Rechtschreiben und den genannten Basisprozessen bei Kindern ohne Schriftspracherwerbsstörungen?

### 5.2 Methode

#### 5.2.1 Design und Stichprobe

Die Forschungsfragen werden durch eine gruppenvergleichende Querschnittsstudie mit quasi-experimentellem Design adressiert. Die Datenerhebung umfasste zwei Untersuchungszeitpunkte. In der ersten Phase wurde eine umfangreichere Stichprobe (N = 855) von Sechstklässlern aus fünf verschiedenen Schweizer Kantonen (51 Klassen) hinsichtlich ihrer Lese-Rechtschreibleistung und ihrer Intelligenz untersucht, um anschliessend die Untersuchungsgruppen zu bilden. Am zweiten Erhebungszeitpunkt nahmen nur noch die ausgewählten Gruppen (IR, LRS, KG) teil und bearbeiteten Testverfahren zu phonologischen, orthographischen und semantisch-lexikalischen Verarbeitungsprozessen.

In der Ausgangsstichprobe zu T1 (N=855) betrug das Durchschnittsalter der Probanden 12;2 Jahre (SD: 6,3 Monate). Es nahmen 434 Jungen (50,8%) und 421 Mädchen (49,2%) an der ersten Phase der Studie teil. Insgesamt 254 (29,8%) der Kinder gaben an, mehrsprachig zu sein.



Die Bildung der Untersuchungsgruppen für T2 erfolgte nach den hier dargestellten Kriterien:

Lese-Rechtschreibstörung: Lese- UND Rechtschreibleistung  $\leq -1,5$  Standardabweichungen (SD) vom Mittelwert der Gesamtstichprobe

Isolierte Rechtschreibstörung: Rechtschreibleistung  $\leq -1,5$  SD vom Mittelwert der Gesamtstichprobe bei einer Leseleistung  $\geq -1$  SD

Um das Vorliegen einer bedeutsamen Diskrepanz zwischen Lese- und Rechtschreibleistung bei der IR-Gruppe sicherzustellen, wurden mittels linearer Regression (Schätzung der Rechtschreibleistung ausgehend von der Leseleistung) Abweichungskoeffizienten errechnet. Die Diskrepanz zwischen vorhergesagter und tatsächlicher Rechtschreibleistung musste für die Kinder mit IR  $\geq 1,5$  SD sein (zu dieser Methodik siehe auch Klicpera et al. 1993).

Als Ausschlusskriterium für alle Untersuchungsgruppen galt ein nonverbaler IQ unter 70. Die Kontrollgruppe (KG) bestand aus einer Zufallsstichprobe von 30 schriftsprachunauffälligen Kindern.

	Geschlecht Männl./weib.	Alter in Monaten	Rechtschreibung (T-Wert)	Lesen (Z-Wert)	IQ	Mehrsprachigkeit (ein- /mehrsprachig)
IR (N=19)	12 / 7	149.42 (7.66)	32.0 (4.26)	-.25 (.40)	92.95 (14.05)	10 / 9
LRS (N=18)	12 / 6	148.89 (6.93)	32.5 (3.01)	-1.92 (.38)	94.94 (12.80)	11 / 7
KG (N=30)	15 / 15	146.27 (6.80)	48.6 (8.15)	.26 (.42)	104.00 (13.35)	18 / 12

Tabelle 2: Merkmale der drei Untersuchungsgruppen

Die Geschlechterverteilung stellt sich in den beiden Gruppen mit schriftsprachlichen Schwierigkeiten etwa im Verhältnis 2:1 zuungunsten der Buben dar, was mit Angaben aus der Literatur übereinstimmt (Hasselhorn & Schuchardt 2006). Der Faktor Gruppenzugehörigkeit hat einen signifikanten Einfluss auf den IQ ( $F(2, 64)=4,78$ ;  $p<.05$ ;  $\eta^2=0.13$ ), welcher jedoch bei den Einzelvergleichen nur zwischen IR und der KG statistische Signifikanz erreicht.

### 5.2.2 Instrumente

Rechtschreibung: Die Rechtschreibkompetenzen der Teilnehmenden wurden durch die Hamburger Schreibprobe (May et al. 2012) erhoben. Dieser standardisierte Test umfasst die Verschriftung von Wörtern und Sätzen und

beinhaltet sowohl quantitative, als auch qualitative Möglichkeiten der Auswertung. Als Mass für die Rechtschreibfähigkeit wurde der im Testmanual ersichtliche T-Wert für die Anzahl der Graphemtreffer der Kinder verwendet (alle Schulformen).

Lesen: Das Lesen wurde über den 1-Minute Wort- und Pseudowortlesetest des Salzburger Lese- und Rechtschreibtests II (Moll & Landerl 2010) erfasst. Dieser zielt auf die automatisierte Worterkennung und die alphabetische Strategie. Die statistischen Analysen der vorliegenden Studie bezogen sich auf z-transformierte Rohwerte, da der SLRT-II lediglich Prozenträge, aber keine T-Werte zur Verfügung stellt.

Wortschatz: Der Wortschatztest des CFT-20 R (Weiß 2007) misst laut Manual den über den Grundwortschatz der deutschen Sprache hinausgehenden Wortschatz in der Umgebungssprache. Er besteht aus 30 Mehrfachwahlaufgaben und differenziert optimal zwischen der dritten und sechsten Klasse. Für die statistischen Vergleiche und Analysen wurden die T-Werte der entsprechenden Altersstufe genutzt.

Wortspezifisches orthographisches Wissen: Die Operationalisierung des wortspezifischen orthographischen Wissens erfolgte durch zwei Subtests zur Differenzierung von Homophonen und von Pseudohomophonen. Diese beiden Aufgabenformate gelten als Standardmasse zur Erhebung von wortspezifischem orthographischem Wissen (Apel 2011). Beide Untertests wurden im Wesentlichen aus der "Universität Georgia Testbatterie phonologischer und orthographischer Fähigkeiten" übernommen (Romonath & Gregg 2003) jedoch leicht an die Altersstufe der Sechstklässler adaptiert. Bei der Differenzierung von Homophonen wurde schriftlich eine Frage präsentiert und die Schülerinnen und Schüler sollten aus zwei Alternativen die korrekte Antwort kennzeichnen (z.B.: Wer überbringt Nachrichten? Boote/Bote). Der Test besteht aus 30 Items und zeigte in einer Vorstudie zufriedenstellende Reliabilitätswerte (interne Konsistenz  $r=.79$ ; Testhalbierungskoeffizient  $r^{tt} = .77$  (Guttman)). Der Untertest zur orthographischen Differenzierung von Pseudohomophonen besteht aus 46 Items und fordert die Teilnehmenden ebenfalls auf, die korrekte Schreibweise aus zwei Antwortmöglichkeiten zu markieren (z.B. Chef – Schef; schwer – schwehr). Die interne Konsistenz betrug in der erwähnten Vorstudie mit 151 Schülerinnen und Schülern  $r=.78$  (Cronbachs  $\alpha$ ) und der Testhalbierungskoeffizient nach Guttman lag bei  $r^{tt} = .77$ .

Morphematische Bewusstheit: Zur Erhebung der morphematischen Bewusstheit wurde ein Untertest erstellt, welcher von den Teilnehmenden in einem rein mündlichen Format zu lösen war. Es wurden Nomen bzw. Adjektive präsentiert, zu welchen ein passendes Verb mit gleichem Wortstamm genannt werden sollte, z.B.: Die Schale (schälen); Das Kleid (bekleiden). Insgesamt bestand der Test aus 39 Items und der Testwert setzte sich aus der Summe der

korrekt genannten Verben zusammen. In der Vorstudie lag die durchschnittliche Schwierigkeit der Items bei  $P = 61$ . Die interne Konsistenz betrug  $\alpha = .83$  und der Testhalbierungskoeffizient war  $r^{tt} = .82$ .

**Schnellbenennung:** Das Verfahren "Rapid Automated Naming and Rapid Alternating Stimulus Test" (RAN/RAS) von Wolf & Denckla (2005) wurde in die deutsche Sprache übersetzt. In einer Vorstudie erreichten die Werte zur kriterialen Validität eine hohe Übereinstimmung mit den Angaben aus dem Manual des Originaltests. Das Verfahren setzt sich zusammen aus sechs Subtests zum Benennen hochfrequenter, einfacher Stimuli. Der Testwert für die vorliegende Untersuchung besteht aus der Zeit, die zusammenfassend benötigt wird, um alle sechs Subtests zu lösen.

**Abrufflüssigkeit:** Es wurden vier Untertests aus dem Regensburger Wortflüssigkeitstest (RWT; Aschenbrenner et al. 2000) ausgewählt (M-Wörter; Wechsel G-Wörter / R-Wörter; Tiere; Wechsel Sportarten / Früchte) und eine Zeitbegrenzung von 2 Minuten vorgegeben. Für die statistischen Auswertungen wurden die Rohwerte verwendet.

**Intelligenz:** Die Intelligenz wurde mittels der Kurzversion des CFT-20-R (Weiß 2006) erhoben und als Ausschlusskriterium verwendet ( $IQ < 70$ ).

### 5.2.3 Untersuchungsdurchführung

Die erste Datenerhebung fand zu Schuljahresbeginn statt. Alle teilnehmenden Schülerinnen und Schüler wurden im Hinblick auf ihre Lese-, Rechtschreib- und Intelligenzleistungen hin untersucht. Die zweite Untersuchungsphase folgte im Frühjahr desselben Schuljahres. Zu diesem Zeitpunkt nahmen nur noch die Kinder aus den Untersuchungsgruppen teil. Die Testdurchführungen und -auswertungen erfolgten an beiden Messzeitpunkten durch die Autorin und durch geschulte Studierende der Logopädie. Durch die Nutzung eines standardisierten Vorgehens (Testreihenfolge, Instruktionen, Testauswertungen) wurde die Vergleichbarkeit der Datenerhebung gesichert.

### 5.2.4 Statistische Analysen

Zur Beantwortung der Frage nach den Unterschieden in den lexikalischen Verarbeitungskompetenzen von Kindern mit IR, mit LRS und von schriftsprachunauffälligen Schülerinnen und Schülern werden univariate Varianzanalysen (ANOVA) eingesetzt. Als Maß für die Effektstärke wird der Determinationskoeffizient  $\eta^2$  verwendet. Multipliziert mit dem Faktor 100 gibt dieser an, wie viel Varianz der Messwerte durch den Faktor Gruppenzugehörigkeit erklärt wird. Für die Post-Hoc-Analysen wird entweder der Tukey HSD-Test (gleiche Varianzen) oder der Games-Howell-Test (ungleiche Varianzen) verwendet.

Die zweite Fragestellung wird explorativ durch Korrelationsanalysen adressiert. Diese Methode wurde aufgrund des geringen Stichprobenumfangs der

multiplen hierarchischen Regressionsanalyse vorgezogen. Hinzu kommt, dass die Verarbeitung von Schriftsprache neben den lexikalischen Prozessen weitere phonologische und orthographische Verarbeitungskompetenzen beinhaltet, die hier nicht berücksichtigt werden. Regressionsanalysen würden daher möglicherweise ein verzerrtes Bild der tatsächlichen Erklärungskraft der beteiligten Verarbeitungskompetenzen zeichnen. Die Korrelationskoeffizienten nach Pearson werden nach den Konventionen von Cohen (1992) interpretiert ( $r=.30$  = mittlerer Effekt;  $r=.50$  = starker Effekt). Aufgrund der geringen Stichprobengröße der Gruppen IR und LRS wird Frage 2 nur für die schriftsprachunauffälligen Kinder aufgeworfen.

### 5.3 Ergebnisse

#### 5.3.1 Unterschiede in den lexikalischen Verarbeitungskompetenzen von Kindern mit IR, mit LRS und schriftsprachunauffälligen Kindern

In Tabelle 3 sind die deskriptiven Statistiken (M, SD) zu den lexikalischen Verarbeitungskompetenzen in den drei Untersuchungsgruppen ersichtlich. Die sich dort bereits abzeichnenden Mittelwertsunterschiede wurden durch ANOVAs statistisch abgesichert, deren Resultate in Tabelle 4 aufgeführt sind.

	Wortschatz	Wortspezifisches ortho. Wissen	Morphematische Bewusstheit	Benennungsgeschwindigkeit RAN	Abrufflüssigkeit
IR (N = 19)	42.84 (4.99)	54.63 (6.64)	15.00 (6.28)	174.17 (19.68)	71.63 (12.82)
LRS (N = 18)	42.56 (9.09)	52.55 (9.05)	13.50 (7.38)	206.84 (32.05)	67.16 (13.71)
KG (N = 30)	48.57 (7.53)	65.96 (6.10)	20.83 (7.23)	172.74 (20.15)	73.70 (18.53)

Tabelle 3: Mittelwerte (M) und Standardabweichungen (SD) in den lexikalischen Verarbeitungsprozessen

	Quadratsumme	Mittel der Quadrate	F (2, 64)	p	$\eta^2$
<b>Wortschatz</b>					
Varianz zwischen	570.43	285.21	5.21	.008	.14
Varianz innerhalb	3500.33	54.69			
<b>Wortspez. orth. Wissen</b>					
Varianz zwischen	2564.67	1282.33	25.09	<.001	.44
Varianz innerhalb	3269.83	51.09			

<b>Morphematische Bew.</b>					
Varianz zwischen	734.40	367.20	7.45	.001	.18
Varianz innerhalb	3154.66	49.29			
<b>Schnellbenennen RAN</b>					
Varianz zwischen	14841.10	7420.55	13.11	<.001	.29
Varianz innerhalb	36225.89	566.03			
<b>Abrufflüssigkeit</b>					
Varianz zwischen	482.18	241.09	.95	.389	.02
Varianz innerhalb	16121.22	251.89			

Tabelle 4: Einfaktorielle ANOVAs für die Effekte der Gruppenzugehörigkeit auf semantisch-lexikalische Verarbeitungskompetenzen

Es zeigen sich signifikante Effekte der Gruppenzugehörigkeit auf vier der fünf untersuchten lexikalischen Verarbeitungsprozesse. Lediglich in der strategischen Abrufflüssigkeit können keine signifikanten Differenzen zwischen den Gruppen festgestellt werden ( $F=.95$ ;  $p=.389$ ). Wie die Effektstärken  $\eta^2$  ausweisen, wird durch den Faktor "Gruppenzugehörigkeit" zwischen 14% und 29% der Varianz im Wortschatz, wortspezifischen orthographischen Wissen, der morphematischen Bewusstheit und im Schnellbenennen erklärt, was als grosse Effekte bezeichnet werden können.

Einen genaueren Einblick in die Art der vorliegenden Gruppenunterschiede bieten die in Tabelle 5 aufgeführten Post-Hoc-Einzelvergleiche. In den drei Untertests zum Wortschatz, zum wortspezifischen orthographischen Wissen und zur morphematischen Bewusstheit schneiden die beiden Gruppen mit Schriftspracherwerbsstörungen (IR, LRS) jeweils auf vergleichbarem Niveau ab, während sie sich gleichzeitig signifikant von der Kontrollgruppe unterscheiden. Die Effektstärken  $d$  weisen durchwegs auf grosse Effekte hin. Der einzige statistisch signifikante Gruppenunterschied zwischen Schülerinnen und Schülern mit IR und mit LRS ergibt sich im Bereich des Schnellbenennens ( $d=1.22$ ;  $p=.003$ ). Die Kinder in den Gruppen IR und KG benötigen für die Lösung aller Subtests des RAN/RAS durchschnittlich knapp drei Minuten (173 Sekunden), die LRS-Gruppe hingegen brauchte für diese Aufgabe im Schnitt 30 Sekunden länger.

	Gruppen	Mittlere Differenz	d	p
Wortschatz	IR vs. LRS	.28	.03	.992
	<b>IR vs. KG</b>	<b>-5.72</b>	<b>.89</b>	<b>.028</b>
	<b>LRS vs. KG</b>	<b>-6.01</b>	<b>.72</b>	<b>.022</b>
Wortspez. orth. Wissen	IR vs. LRS	2.07	.26	.653
	<b>IR vs. KG</b>	<b>-11.33</b>	<b>1.79</b>	<b>&lt;.001</b>
	<b>LRS vs. KG</b>	<b>-13.41</b>	<b>1.83</b>	<b>&lt;.001</b>
Morphem.Bewusstheit	IR vs. LRS	1.50	.21	.793
	<b>IR vs. KG</b>	<b>-5.83</b>	<b>.86</b>	<b>.017</b>
	<b>LRS vs. KG</b>	<b>-7.33</b>	<b>1.00</b>	<b>.002</b>
Schnellbenennen RAN	<b>IR vs. LRS</b>	<b>-32.67</b>	<b>1.22</b>	<b>.003</b>
	IR vs. KG	1.42	.07	.967
	<b>LRS vs. KG</b>	<b>34.10</b>	<b>1.27</b>	<b>.001</b>
Abrufflüssigkeit	IR vs. LRS	4.46	.33	.670
	IR vs. KG	-2.06	.13	.897
	LRS vs. KG	-6.53	.40	.357

Tabelle 5: Einzelvergleiche der drei Gruppen bezüglich ihrer semantischen Verarbeitungs-kompetenzen

### 5.3.2 Zusammenhänge zwischen dem Wortlesen und Rechtschreiben und der lexikalischen Verarbeitung bei Kindern ohne Schriftspracherwerbsstörungen

Tabelle 6 beinhaltet die Korrelationen zwischen dem Rechtschreiben, dem Wort- und Pseudowortlesen und den einzelnen semantisch-lexikalischen Verarbeitungskompetenzen von Sechstklässlern ohne schriftsprachliche Schwierigkeiten. Die Rechtschreibleistung korreliert moderat bis hoch mit dem wortspezifischen orthographischen Wissen ( $r=.833^{**}$ ), dem Wortschatz ( $r=.464^{**}$ ) und der morphematischen Bewusstheit ( $r=.548^{**}$ ). Während das Schnellbenennen keinen signifikanten Zusammenhang mit der Rechtschreibung aufweist, ergibt sich ein mittlerer Zusammenhang zwischen Rechtschreibung und Abrufflüssigkeit ( $r=.421^*$ ). Wort- und Pseudowortlesen korrelieren jeweils mit dem wortspezifischen orthographischen Wissen ( $r=.490^{**}$  bzw.  $.586^{**}$ ) und mit dem Schnellbenennen ( $r=-.449^{**}$  bzw.  $-.542^{**}$ ), es ergeben sich jedoch keine signifikanten Zusammenhänge mit den weiteren lexikalischen Kompetenzen. Zwischen den einzelnen Komponenten der lexikalischen Verarbeitung erweisen sich Wortschatz, morphematische Bewusstheit, wortspezifisches orthographisches Wissen und Abrufflüssigkeit als eng

miteinander assoziiert ( $r$  zwischen  $.398^{**}$  und  $.713^{**}$ ), während das Schnellbenennen mit keiner dieser Kompetenzen signifikant korreliert.

	Recht-schreiben	Wort-lesen	Pseudo-wortlesen	Wort-schatz	Wortspez orth. Wissen	Morphem Bewussth	Schnell-Benennen RAN	Abruf-flüssigk
Recht-schreiben	-							
Wort-lesen	.537**	-						
Pseudo-wortlesen	.361*	.703**	-					
Wort-schatz	.464**	.319	.112	-				
Wortspez orth. Wissen	.833**	.586**	.490**	.561**	-			
Morphem Bew.	.548**	.299	.245	.694**	.600**	-		
RAN/RAS	-.244	-.449*	-.542**	-.070	-.141	-.134	-	
Abruf-flüssig-keit	.421*	.269	.108	.713**	.398*	.631**	-.296	-

\*\*= die Korrelation ist auf dem Niveau von 0.01 (2-seitig) signifikant

\* = die Korrelation ist auf dem Niveau von 0.05 (2-seitig) signifikant

Tabelle 6: Korrelationen zwischen dem Lesen und Rechtschreiben und den semantisch-lexikalischen Basisprozessen in der schriftsprachunauffälligen Gruppe (N=30)

## 6. Diskussion

Zusammenfassend kann festgehalten werden, dass Schülerinnen und Schüler mit LRS und IR gegenüber einer Alterskontrollgruppe signifikant schlechtere Leistungen im semantisch-lexikalischen Bereich aufweisen. Dies betrifft den Wortschatz, das wortspezifische orthographische Wissen und die morphematische Bewusstheit, nicht jedoch die strategische Abrufflüssigkeit. Besondere Aufmerksamkeit verdient das RAN, in dem die LRS-Gruppe ein deutliches Leistungsdefizit demonstriert, nicht jedoch die Kinder mit IR. Das Vorliegen von Leseschwierigkeiten (zusätzlich zu Rechtschreibproblemen) scheint folglich mit dem Defizit der Benennungsgeschwindigkeit assoziiert zu sein, was sehr gut übereinstimmt mit vorhandenen Forschungsbefunden, welche einen engen Zusammenhang zwischen Benennungsgeschwindigkeit und Leseflüssigkeit nahelegen (Furnes & Samuelsson 2011; Moll & Landerl 2009). Die explorativen Korrelationsanalysen weisen weiter darauf hin, dass bei schriftsprachunauffälligen Kindern gute Rechtschreibleistungen einhergehen

mit hohen Kompetenzen in allen untersuchten lexikalischen Verarbeitungsprozessen, ausser dem RAN. Das Wort- und Pseudowortlesen ist mit dem Schnellbenennen und dem wortspezifischen orthographischen Wissen assoziiert, was die Bedeutung des Schnellbenennens für das Lesen weiter unterstreicht.

Interpretiert man die oben dargestellten Resultate im Licht der Lexikalischen Qualitätshypothese (Perfetti 2007), so weist dies auf eine geringe Quantität und/oder Qualität der lexikalischen Repräsentationen bei LRS und bei IR hin. Diese offenbart sich in den schlechteren semantisch-lexikalischen Leistungen gegenüber den Peers. Gleichzeitig geben die Korrelationsanalysen in der schriftsprachunauffälligen Gruppe Hinweise auf die enge Verbindung zwischen dem für die Schriftsprachverarbeitung so bedeutsamen wortspezifischen orthographischen Wissen mit dem Wortschatz und der morphematischen Bewusstheit. Im Sinne der LQH gründen diese Aspekte der Verarbeitung auf einer gemeinsamen Basis: je mehr Wissen über Wörter vorhanden ist, desto enger sind die Verknüpfungen zwischen den verschiedenen Ebenen (Phonologie, Orthographie, Morphologie, Wortbedeutung) und desto besser kann neues Wissen in die vorhandenen Strukturen integriert werden, was auch den automatisierten Abruf dieser Informationen positiv beeinflussen sollte.

Besonders aussagekräftig ist auch der Unterschied hinsichtlich RAN und der strategischen Abrufflüssigkeit. Bei Kindern mit LRS kann ein bedeutsames Defizit in der Schnellbenennung festgestellt werden, welches bei Gleichaltrigen mit IR nicht auftritt. Der schnelle und hoch automatisierte Zugriff auf lexikalische Repräsentationen scheint bei Kindern mit LRS also beeinträchtigt zu sein, während strategische Suchmechanismen unter Einsatz exekutiver Funktionen bei LRS und bei IR keine Schwierigkeit darstellen. Den Schülerinnen und Schülern mit IR gelingt trotz der vermutlich schlechten lexikalischen Qualität ihrer Repräsentationen der automatisierte Zugriff beim Lesen mühelos. Dabei kommen ihnen modalitätsspezifische Unterschiede zwischen Lesen und Schreiben, aber auch die Besonderheiten der deutschen Orthographie entgegen. Das Lesen stellt einen Wiedererkennungsprozess dar, für welchen es ausreicht, das zu entziffernde Wort von anderen potentiellen Kandidaten abzugrenzen. Dafür ist es nicht notwendig, Identität und Reihenfolge aller Grapheme eines Wortes zu verarbeiten (Frith 1980). Sicheres, korrektes Rechtschreiben hingegen kann nur unter Rückgriff auf vollständig repräsentiertes orthographisches Wissen gelingen. In der deutschen Orthographie wird dieser Effekt verstärkt, da sie eine Asymmetrie zwischen regelhaften Graphem-Phonem-Korrespondenzen in der Leserichtung und weit weniger eindeutigen Phonem-Graphem-Korrespondenzen in der Schreibrichtung aufweist (Moll et al. 2009). Gleichwohl bleibt die besondere Stärke (bzw. ausbleibende Schwäche) der Schülerinnen und Schüler mit IR im RAN unerwartet im Vergleich zu ihrem sonstigen Profil in den semantisch-lexikalischen und schriftsprachlichen Kompetenzen.



Zu den Limitationen der vorliegenden Untersuchung gehört neben der geringen Stichprobengrösse der Subgruppen auch die Verwendung informeller Tests. Da der Schriftspracherwerb als Entwicklungsprozess zu verstehen ist, der in starkem Masse auch von soziokulturellen Faktoren beeinflusst wird, kann durch ein querschnittlich angelegtes Design nur eine Momentaufnahme entstehen. Es lässt keine Aussagen darüber zu, inwiefern die beobachteten semantisch-lexikalischen Schwierigkeiten von Kindern mit Schriftspracherwerbsstörungen Ursache oder Folge der erschwerten Lese- und Rechtschreibpraxis darstellen oder ob die geringeren Zusammenhänge zwischen den Basisprozessen und dem Lesen und Schreiben Ausdruck von Kompensationsstrategien sein könnten. Zukünftige Studien könnten diese Limitationen aufgreifen und durch den Einbezug von Kindern mit isolierten Lesestörungen das entstandene Bild vervollständigen.

## LITERATUR

- Apel, K. (2011). What is orthographic knowledge? *Language, Speech and Hearing Services in Schools*, 42, 592-603.
- Aschenbrenner, S., Tucha, O. & Lange, K. W. (2000). *Regensburger Wortflüssigkeitstest*. Bern: Hogrefe.
- Bahr, R. H., Silliman, E. R. & Berninger, V. W. (2020). Derivational morphology bridges phonology and orthography: Insights into the development of word-specific spellings by superior, average, and poor spellers. *Language, Speech and Hearing Services in Schools*, 51, 640-654.
- Berendes, K., Schnitzler, C. D., Willmes, K. & Huber, W. (2010). Die Bedeutung von Phonembewusstheit und semantisch-lexikalischen Fähigkeiten für Schriftsprachleistungen in der Grundschule. *Sprache · Stimme · Gehör*, 34, 33-41.
- Bourassa, D. C. & Treiman, R. (2008). Morphological constancy in spelling: A comparison of children with dyslexia and typically developing children. *Dyslexia*, 14, 155-169.
- Burt, J. S. & Fury, M. B. (2000). Spelling in adults: The role of reading skills and experience. *Reading and Writing*, 13, 1-30.
- Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science*, 1, 98-101.
- Coltheart, M. (2006). Dual route and connectionist models of reading: An overview. *London Review of Education*, 4, 5-17.
- Da Silva, C. G., Petersson, K. M., Faísca, L., Ingvar, M. & Dias Reis, A. I. (2004). The effects of literacy and education on the quantitative and qualitative aspects of semantic verbal fluency. *Journal of Clinical and Experimental Neuropsychology*, 26, 266-277.
- Dilling, H. & Schulte-Markwort, E. (2014). *Internationale Klassifikation psychischer Störungen: ICD-10. 9. Aufl.*. Bern: Huber.
- Frith, U. (1980). Unexpected spelling problems. In U. Frith (Hg.), *Cognitive processes in spelling* (S. 495-515). London: Academic Press.
- Furnes, B. & Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences*, 21, 85-95.
- Hagiliassis, N., Pratt, C. & Johnston, M. (2006). Orthographic and phonological processes in reading. *Reading and Writing*, 19, 235-263.

- Harrison, G. (2004). Differential use of orthographic and phonological coding strategies by students with symmetrical and asymmetrical reading and spelling skills. *Canadian Journal of School Psychology, 19*, 33-57.
- Hasselhorn, M. & Schuchardt, K. (2006). Lernstörungen. *Kindheit und Entwicklung, 15*, 208-215.
- Hauerwas, L. B. & Walker, J. (2003). Spelling of inflected verb morphology in children with spelling deficits. *Learning Disabilities Research and Practice, 18*, 25-35.
- Klicpera, C., Gasteiger-Klicpera, B. & Schabmann, A. (1993). *Lesen und Schreiben: Entwicklung und Schwierigkeiten*. Bern: Huber.
- Landerl, K. (2001). Word recognition deficits in German: More evidence from a representative sample. *Dyslexia, 7*, 183-196.
- Larkin, R. F. & Snowling, M. J. (2008). Morphological spelling development. *Reading & Writing Quarterly, 24*, 363-376.
- Lyon, G. R., Shaywitz, S. E. & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia, 53*, 1-14.
- May, P., Vieluf, U. & Malitzky, V. (2012). *Hamburger Schreib-Probe*. Stuttgart: VPM.
- Miceli, G. & Capasso, R. (2006). Spelling and dysgraphia. *Cognitive Neuropsychology, 23*, 110-134.
- Moll, K., Fussenegger, B., Willburger, E. & Landerl, K. (2009). RAN is not a measure of orthographic processing. Evidence from the asymmetric German orthography. *Scientific Studies of Reading, 13*, 1-25.
- Moll, K. & Landerl, K. (2009). Double dissociation between reading and spelling deficits. *Scientific Studies of Reading, 13*, 359-382.
- Moll, K. & Landerl, K. (2010). *SLRT-II. Lese- und Rechtschreibtest*. Bern: Huber.
- Moll, K. & Landerl, K. (2011). Assoziationen und Dissoziationen von Störungen des Lesens und Rechtschreibens. *Spektrum Patholinguistik, 4*, 47-74.
- Nagy, W., Berninger, V., Abbott, R., Vaughan, K. & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second-grade readers and at-risk fourth-grade writers. *Journal of Educational Psychology, 95*, 730-742.
- Nation, K. & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading, 27*, 342-356.
- Ouellette, G. (2010). Orthographic learning in learning to spell: The roles of semantics and type of practice. *Journal of Experimental Child Psychology, 107*, 50-58.
- Pacton, S. & Deacon, S. H. (2008). The timing and mechanisms of children's use of morphological information in spelling: A review of evidence from English and French. *Cognitive Development, 23*, 339-359.
- Pennington, B. F., Cardoso-Martins, C., Green, P. A. & Lefly, D. L. (2001). Comparing the phonological and double deficit hypotheses for developmental dyslexia. *Reading and Writing, 14*, 707-755.
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*, 357-383.
- Perfetti, C. & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro & P. Reitsma (Hgg.), *Studies in written language and literacy: Precursors of functional literacy* (S. 189-213). Amsterdam: Benjamins.
- Romani, C., Ward, J. & Olson, A. (1999). Developmental surface dysgraphia: What is the underlying cognitive impairment? *The Quarterly Journal of Experimental Psychology, 52*, 97-128.
- Romonath, R. & Gregg, N. (2003). Universität Georgia Testbatterie phonologischer und orthographischer Fähigkeiten (deutsche Version). In Ministerium für Bildung, Wissenschaft und Kultur des Landes Mecklenburg Vorpommern (Hg.), *Optimierung von Lese-Rechtschreibfähigkeiten bei Legasthenikern im Jugendalter*. Greifswald: Druckhaus Panzig.

- Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B. & Claverie, B. (2004). Verbal fluency output in children aged 7-16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language*, 89, 192-202.
- Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development*, 61, 17-28.
- Snowling, M. J. & Hulme, C. (2012). Annual research review: The nature and classification of reading disorders - a commentary on proposals for DSM-5. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 53, 593-607.
- Vukovic, R. K. & Siegel, L. S. (2006). The double-deficit hypothesis: A comprehensive analysis of the evidence. *Journal of Learning Disabilities*, 39, 25-47.
- Wagner, R. K. & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101, 192-212.
- Weiß, R. H. (2007). *WS/ZF-R. Wortschatztest und Zahlenfolgentest - Revision*. Ergänzungstest zum CFT 20-R. Bern: Hogrefe.
- Winkes, J. (2014). *Isolierte Rechtschreibstörung. Eigenständiges Störungsbild oder leichte Form der Lese-Rechtschreibstörung? Eine Untersuchung der kognitiv-linguistischen Informationsverarbeitungs-kompetenzen von Kindern mit Schriftspracherwerbsstörungen*. Thèse de doctorat: Université de Fribourg, 2014. <http://doc.rero.ch/record/210344>
- Wolf, M. & Denckla, M. B. (2005). *RAN/RAS. Rapid automatized naming and alternating stimulus test*. Austin: pro-ed.



# Standarddeutsche Interferenzen im Dialektwortschatz Schweizer Jugendlicher: Lexikalische und lautliche Entlehnungen

**Marie-Anne MORAND, Sandra SCHWAB & Stephan SCHMID**

Universität Zürich

Phonetisches Laboratorium, Institut für Computerlinguistik

Rämistrasse 71, CH-8006 Zürich

marie-anne.morand@uzh.ch, sandra.schwab@uzh.ch, stephan.schmid@uzh.ch

At least since the turn of the millennium, adolescents growing up in linguistically and culturally diverse neighborhoods of European cities have developed innovative ways of speaking, which are often referred to as "multiethnolectal". To investigate the phonetic features of multiethnolectal Swiss German, a research project is currently carried out in the city of Zurich. Forty-eight adolescents were recorded in different speaking styles, and in a subsequent perception experiment, these speakers were rated by other adolescents on how multiethnolectal they sound ("Screening Score"). In the present contribution, we analyze the transcripts of a picture description task, focusing on lexical and phonetic innovations with respect to traditional Zurich German. We assume that adolescents who are perceived as speaking rather multiethnolectal Zurich German show more Standard German interferences both on the lexical and the phonetic level. Indeed, the Screening Score is a statistically significant predictor for the number of interferences. In addition, the probability of purely phonetic interferences as opposed to lexical interferences increases with higher Screening Scores.

## **Keywords**

multiethnolects, Zurich German, diglossia, lexical interference, dialectal vocabulary.

## **Stichwörter**

Multiethnolekte, Zürichdeutsch, Diglossie, lexikalische Interferenz, Dialektwortschatz.

## **1. Einleitung**

Der vorliegende Beitrag stellt eine empirische Studie zu lexikalischen und lautlichen standarddeutschen Interferenzen im multiethnolektalen Schweizerdeutschen vor. Unter dem Begriff "Multiethnolekt" versteht Clyne (2000: 87) einen besonderen Typ von Ethnolekt, der von verschiedenen Minderheitsgruppen gesprochen wird und somit auch eine identitätsstiftende Funktion ausübt. Als Datenmaterial für unsere Untersuchung dienen Bildbeschreibungen, welche von ein- und mehrsprachigen Jugendlichen an zwei Zürcher Sekundarschulen aufgenommen wurden. Dieselben Sprecher\_innen wurden von Jugendlichen einer dritten Zürcher Sekundarschule bezüglich ihrer jeweiligen "Multiethnolektalität" bewertet.

Im Folgenden wird zuerst die bisherige Forschung zum multiethnolektalen Schweizerdeutschen umrissen, wobei sowohl auf phonetische als auch auf lexikalische Merkmale des multiethnolektalen Schweizerdeutschen eingegangen wird. Anschliessend werden lexikalische und lautliche Unterschiede zwischen (Schweizer) Standarddeutsch und traditionellem



Zürichdeutsch erläutert. Ein kontinuierlicher Übergang zwischen den beiden Varietäten wird für die Deutschschweizer Diglossie in der Regel nicht angenommen (im Gegensatz zum regionalsprachlichen Kontinuum zwischen Dialekt und Standard in weiten Teilen Österreichs und Deutschlands, vgl. Ammon 2003; Kehrein 2019); dementsprechend wird ein Wechsel innerhalb einer Äusserung normalerweise als *Code-Switching* interpretiert (vgl. Petkova 2016). Bei Mischformen auf der Wortebene kann man hingegen von lexikalischen Interferenzen sprechen. Aufgrund bisheriger Beobachtungen (vgl. Schmid 2020) kann angenommen werden, dass Jugendliche, welche eher multiethnolektales Schweizerdeutsch sprechen, vermehrt standarddeutsche Interferenzen im Dialekt aufweisen (vgl. auch Berthele 2004; Petkova 2011).

### 1.1 Multiethnolektales Schweizerdeutsch

Spätestens seit dem Jahrtausendwechsel wurde in verschiedenen europäischen Städten beobachtet, dass Jugendliche in kulturell und sprachlich vielfältigen Quartieren neue Sprechweisen entwickeln, welche in der Forschung unter anderem als "Multiethnolekte" bezeichnet werden (vgl. Clyne 2000; Quist 2000). Verschiedene sprachlich saliente Phänomene des multiethnolektalen Schweizerdeutschen wurden erstmals von Tissot et al. (2011) beschrieben, während eine systematische Untersuchung der phonetischen Merkmale nun in einem SNF-Projekt (Nr. 165798) vorgenommen wird, innerhalb dessen auch dieser Beitrag entstanden ist. In den nächsten beiden Abschnitten stellen wir zunächst phonetische Merkmale vor und gehen anschliessend auf einzelne Beobachtungen zu lexikalischen Merkmalen ein.

#### 1.1.1 Phonetische Merkmale des multiethnolektalen Zürichdeutschen

Tissot et al. (2011) führen für schweizerdeutsche Ethnolekte unter anderem die in den Beispielen (1)-(8) illustrierten phonetischen Merkmale auf: stimmhafte Plosive /b d g/ (1) anstatt der Lenis-Plosive /b̥ d̥ ġ/ (2), wortinitiale Fortis-Frikative /f s ʃ x/ (3) anstelle der Lenis-Frikative /f̥ s̥ ʃ̥ x̥/ (4), eine längere Dauer labiodentaler Approximanten (also /v/ (5) anstatt /v̥/ (6)), sowie das Ausbleiben von Sandhi-Prozessen ((7) anstatt (8)).

- (1) *Das isch en herte* [ˈbodə]<sup>1</sup> ("Das ist ein harter Boden")
- (2) *Das isch en herte* [ˈb̥o̥d̥ə]
- (3) *Ich gsee de* [fuxs] ("Ich sehe den Fuchs")
- (4) *Ich gsee de* [f̥uxs]
- (5) *Mir sueched e* [ˈv̥oniġ] ("Wir suchen eine Wohnung")
- (6) *Mir sueched e* [ˈv̥oniġ]
- (7) *Das chan i* [nød̥ ˈbru:xə] ("Das kann ich nicht brauchen")
- (8) *Das chan i* [nø̥ ˈpru:xə]

<sup>1</sup> Bei der orthographischen Wiedergabe des Zürichdeutschen orientieren wir uns an der von Dieth (1938/1986) vorgeschlagenen Schreibweise.

Diese Beispiele aus unserem Lesesprache-Korpus zeigen, dass bei Zürcher Jugendlichen sowohl eher multiethnolektale Realisierungen ((1), (3), (5), (7)) als auch eher traditionelle Aussprachevarianten ((2), (4), (6), (8)) vorkommen. Quantitative Analysen ergaben, dass multiethnolektal sprechende Jugendliche Lenis-Plosive signifikant stimmhafter realisieren als Studentinnen, die ein eher traditionelles Zürichdeutsch sprechen (Morand et al. 2019a). Jugendliche, deren Dialekt als eher multiethnolektal eingestuft wurde (siehe 2.4), produzierten ausserdem längere wortinitiale Frikative (Morand et al. 2019b) und realisierten weniger Sandhi-Phänomene (Morand et al. 2020) als Jugendliche, welche als eher traditionell zürichdeutschsprechend wahrgenommen wurden. Auf weitere phonetische Merkmale (Aspiration und Realisierung des velaren Frikativs als Fortis-Plosiv) wird in 2.5.2 eingegangen.

### 1.1.2 Lexikalische Merkmale des multiethnolektalen Zürichdeutschen

Systematische Untersuchungen zu lexikalischen Merkmalen des multiethnolektalen Zürichdeutschen liegen bisher nicht vor. Tissot et al. (2011) verweisen lediglich auf einige ausgewählte Besonderheiten: So beobachtet man z.B. Formen des *Crossings* (vgl. Rampton 2005) aus dem Albanischen und Standarddeutschen auch bei Jugendlichen ohne albanischen oder deutschen Sprachhintergrund. Für verschiedene multiethnolektale Varietäten Europas werden Interferenzen auf lexikalischer Ebene erwähnt; so findet man etwa Lehnwörter aus dem Arabischen und Türkischen in Multiethnolekten germanischer Sprachen (vgl. Dirim & Auer 2004; Quist 2010). Daneben erscheinen aber auch typische Diskursmarker und Interjektionen wie *Mann* oder *(h)ey Alter* (vgl. Auer 2003; Dirim & Auer 2004) sowie eine erweiterte Verwendung gewisser Bewegungsverben (vgl. Goschler et al. 2013).

Die Analyse eines Sprechers des multiethnolektalen Zürichdeutschen ergab, dass dieser einzelne lexikalische Formen verwendet, welche eher dem (Schweizer) Standarddeutschen als dem lokalen (traditionellen) Dialekt zuzuordnen sind, so dass seine Sprechweise zum Teil als Mischung zwischen Schweizer- und Standarddeutsch bezeichnet werden kann (Schmid 2020: 7).<sup>2</sup> Eine solche partielle Hybridisierung beider Varietäten, die in Kombination mit weiteren phonetischen und morphosyntaktischen Merkmalen zu einer strukturellen "Dialekt-Transformation" führt (vgl. Schmid 2020: 6-10), würde somit der traditionellen diglossischen Sprachsituation (vgl. Siebenhaar & Wyler 1997) zuwiderlaufen, da sich die beiden Sprachebenen nicht mehr klar unterscheiden lassen.

<sup>2</sup> Dieser Sprecher trat in einer 2001 ausgestrahlten Talkshow des Privatsenders TV3 auf.

## 1.2 Lexikalische und lautliche Unterschiede zwischen traditionellem Zürichdeutsch und (Schweizer) Standarddeutsch

Bekanntlich wird die Sprachsituation in der Deutschschweiz als klassischer Fall von Diglossie (Ferguson 1959) aufgefasst, in welcher sich Sprecher\_innen und Hörer\_innen immer bewusst darüber sind, welche Varietät verwendet wird (vgl. Christen 2000; Hove 2008). Für Aussenstehende ist diese Unterscheidung offenbar aber nicht immer eindeutig, da es zwischen den beiden Varietäten eine Schnittmenge gibt, welche bestimmte Ausdrücke für Rezipienten allenfalls ambig macht (Hove 2008: 65). Daneben gibt es aber auch eindeutig dialektale lexikalische Varianten und solche, die klar der Standardvarietät zugeordnet werden können. Im Folgenden illustrieren wir zuerst einige lautliche Differenzen zwischen traditionellem Zürichdeutsch und (Schweizer) Standarddeutsch, um anschliessend auf typische lexikalische Unterschiede einzugehen.

### 1.2.1 Lautliche Unterschiede zwischen den beiden Varietäten

Für die eindeutige Zuweisung bestimmter Wörter ist gemäss Hove (2008) insbesondere die phonologische Form ausschlaggebend. Dialektale Formen weisen zum Beispiel im Gegensatz zu den standarddeutschen Varianten keine Diphthongierung gewisser mittelhochdeutscher Monophthonge auf (*Huus* anstatt *Haus*). Andererseits gibt es aber auch mittelhochdeutsche Diphthonge, welche im Dialekt erhalten geblieben sind und standarddeutschen Monophthongen gegenüberstehen (*müed* anstatt *müde* oder /liəb/ anstatt /li:b/; vgl. Schirmunski 1962: 29, zitiert in Barbour & Stevenson 1990: 88). Hove (2008: 66) zählt hingegen für die meisten Dialektregionen das aspirierte *k* zu den Varianten, "deren Realisierung eindeutig auf die Standardsprache schliessen lässt [...]" ([k<sup>h</sup>unst] anstatt [k̄xunʃt] oder [k<sup>h</sup>ɪnt] anstatt [ʏɪnd]); traditionell kommt die Aspiration nur im Basler und Bündner Dialekt vor. Da die untersuchten Jugendlichen aber Zürichdeutsch sprechen, wurde der velare Fortis-Plosiv in den entsprechenden Lexemen als "lautliche Interferenz" codiert (siehe 2.5.2).

Für die lautliche Analyse des dialektalen Wortschatzes haben wir uns an der phonetischen Beschreibung des Zürichdeutschen von Fleischer & Schmid (2006) orientiert und eindeutig auf die Phonologie des Standarddeutschen zurückzuführende Abweichungen als "lautliche Interferenz" codiert (siehe 2.5.2). Im Gegensatz zum Standarddeutschen werden z.B. Fortis-Plosive vor betonten Vokalen im traditionellen Zürichdeutschen meist nicht aspiriert (siehe Bsp. 9 und 10); die Aspiration ist vielmehr lexikalisch bestimmt und insbesondere typisch für Entlehnungen und Eigennamen (Fleischer & Schmid 2006: 244; vgl. Bsp. 11).

- (9) Standarddeutsch: *Tag* [t<sup>h</sup>a:k], *Post* [p<sup>h</sup>ɔst]  
 (10) Zürichdeutsch: *Tag* [tɔ:ŋ], *Post* [pɔʃt]  
 (11) Zürichdeutsch: *Tee* [t<sup>h</sup>e:], *Paul* [p<sup>h</sup>æʊl]



Ladd & Schmid (2018: 232) nehmen aber an, dass die Aspiration bei der jüngeren Generation zunimmt (vgl. auch Schifferle 2010; Leemann et al. 2020).

### 1.2.2 Lexikalische Unterschiede zwischen den beiden Varietäten

Hove (2008: 77) hat allgemein festgestellt, dass nur wenige lexikalische Varianten eindeutig der Standardsprache zugeordnet werden können, da es in vielen Fällen möglich sei, "ursprünglich standardsprachliche Varianten durch lautliche und morphologische Anpassung in den Dialekt zu integrieren, z.B. *arbeite* [anstatt z.B. *schaffe* oder *wärche*"]". Solche Anpassungen werden aber meist als auffällig wahrgenommen und von uns deshalb als "lexikalische Interferenz" codiert (siehe 2.5.2). So geht Graf (2015: 69) z.B. davon aus, dass dem Sprecherkollektiv "vor allem Umschichtungen von einer dialektalen Bezeichnungsvielfalt hin zu einer einheitlichen Bezeichnungsweise" auffallen; da die Standardsprache als Leitvarietät über die grösste kommunikative Reichweite verfügt, würden standarddeutsche Bezeichnungen im Dialekt zwar übernommen, blieben aber trotzdem auffällig.

Aufgrund der Feststellung, dass viele Deutschschweizer standarddeutsche Ausdrücke mittels phonologischer und morphologischer Anpassungen ins Schweizerdeutsche integrieren, postulieren Häcki Buhofer & Burger (1998: 25-26) drei Lexika: Dialekt-Lexikon, Standard-Lexikon und gemeinsames Lexikon. Diese Einteilung in drei Lexika ist aber durchlässig und dem Wandel unterworfen. Beim vermutlich umfangreichsten "gemeinsamen Lexikon" handelt es sich um Wörter, welche in den beiden Varietäten identisch oder ähnlich sind. Hingegen kann davon ausgegangen werden, dass die Verwendung von Ausdrücken aus dem Standard-Lexikon im Dialekt und insbesondere umgekehrt die Verwendung von Ausdrücken aus dem Dialekt-Lexikon im Standard markiert sind und in der Regel als *Code-Switching* interpretiert werden (vgl. dazu auch Häcki Buhofer & Burger 1998: 26).

Der Einfluss der diglossischen Sprachsituation im Zusammenhang mit Mehrsprachigkeit wurde bereits aus unterschiedlichen Perspektiven beleuchtet. Häcki Buhofer et al. (2007) stellten zum Beispiel fest, dass die Komplexität der Spracherwerbssituation im Zusammenhang mit Diglossie und Mehrsprachigkeit keinen negativen Einfluss auf die standarddeutsche Schreib- und Lesekompetenz der Jugendlichen hatte. Ender et al. (2007) untersuchten im Projekt "Deutsch als Zweitsprache in Dialektumgebung" unter anderem die Sprachfähigkeiten und den Stellenwert der beiden Varietäten bei Kindern mit Migrationshintergrund. Die beiden deutschen Varietäten und die Erstsprache erfüllen dabei verschiedene Funktionen: so seien dialektale Fähigkeiten und die Erstsprache der Schlüssel zur sozialen Integration im Freundes- bzw. Familienkreis, während die Beherrschung des Standarddeutschen sich als notwendig für die Gleichberechtigung im akademischen Kontext erweist (Ender & Straßl 2009).

### 1.3 Forschungsfragen

In der vorliegenden Studie konzentrieren wir uns nicht auf standardsprachliche Kompetenzen, sondern untersuchen den Dialektgebrauch. Dabei gehen wir zwei zusammenhängenden Forschungsfragen nach:

- Lässt sich aufgrund der wahrgenommenen Multiethnolektalität der Sprecher\_innen die Wahrscheinlichkeit einer Interferenz vorhersagen?
- Lässt sich aufgrund der wahrgenommenen Multiethnolektalität der Sprecher\_innen die Wahrscheinlichkeit der Art einer Interferenz vorhersagen?

Die zweite Frage ist nur relevant, falls die erste Frage bejaht wird. Zur Erörterung dieser Forschungsfragen stellen wir die Anzahl der Interferenzen einer bestimmten Person dem sogenannten "Screening Score" derselben Person gegenüber; darunter verstehen wir die Bewertungen von Jugendlichen (Bewerter\_innen) in Bezug auf die jeweilige Multiethnolektalität der Sprecher\_innen (siehe 2.4).

Zur Beantwortung dieser beiden Fragen werden einerseits quantitative Analysen durchgeführt; andererseits wird auch eine qualitative Beschreibung einzelner lexikalischer und lautlicher Interferenzen vorgenommen. Somit soll systematisch untersucht werden, inwiefern die Sprechweise der Jugendlichen mit der Verwendung von lexikalischen Formen, welche eher dem Standarddeutschen zuzuordnen sind, einhergeht.

## 2. Methode

### 2.1 Bildbeschreibung

Im Rahmen unseres Forschungsprojekts wurden von Jugendlichen an zwei Zürcher Sekundarschulen verschiedene mündliche Sprachdaten erhoben. Unter anderem beschrieben die Schüler\_innen das sogenannte "Cookie Theft Bild" der *Boston Diagnostic Aphasia Examination* (BDAE, Goodglass & Kaplan 1983), eine Schwarzweisszeichnung, welche zwei Handlungen darstellt: Ein Junge und ein Mädchen versuchen, in der Küche Kekse zu stehlen, während gleichzeitig eine Frau das Geschirr spült. Der Junge steht auf einem Hocker, der zu kippen droht; bei der Frau läuft das Wasser über den Rand des Spültrogs. Die Jugendlichen wurden aufgefordert, alles zu beschreiben, was sie auf diesem Bild sehen. Für die Analyse im vorliegenden Beitrag wurde jeweils ein Sprachausschnitt am Anfang der Erzählung gewählt (bis zur ersten Intervention der Interviewerin).

### 2.2 Sprecher\_innen

Die 48 Jugendlichen, deren Sprachaufnahmen analysiert wurden, waren zum Zeitpunkt der Datenerhebung durchschnittlich 14.3 Jahre alt (Standardabweichung = 0.74); mehr als die Hälfte war weiblich (28 vs. 20).

Neun Personen gaben an, vor dem Eintritt in den Kindergarten nur Schweizerdeutsch gesprochen zu haben. Von den restlichen 39 gaben 21 an, vor dem Kindergarten eine andere oder mehrere andere Sprachen gesprochen zu haben; 18 sprachen vor dem Kindergarten Schweizerdeutsch und eine oder mehrere andere Sprachen. Das Durchschnittsalter für den Erwerbsbeginn des Schweizerdeutschen lag bei 3.7 Jahren (Standardabweichung = 3.10). Gemäss unserer Einschätzung stellte der Zürcher Dialekt bei allen untersuchten Personen eine ihrer Erstsprachen dar; Schüler\_innen, welche sich noch am Anfang des L2-Erwerbsprozesses befanden und somit eine typische Lernervarietät aufwiesen, wurden von der Analyse ausgeschlossen.

Schweizerdeutsch und Standarddeutsch wurden im soziolinguistischen Fragebogen und für die Wortschatzanalyse als unterschiedliche Sprachen kategorisiert. Neben Deutschkenntnissen verfügen die Jugendlichen aber – nebst den in der Schule gelernten Fremdsprachen – auch noch über muttersprachliche Kompetenzen in zahlreichen weiteren Sprachen (in alphabetischer Reihenfolge): Albanisch, Arabisch, Bosnisch, Englisch, Französisch, Italienisch, Kroatisch, Krobo, Paschtunisch, Portugiesisch, Punjabi, Rumänisch, Singhalesisch, Serbisch, Somalisch, Spanisch, Thai, Türkisch, Twi und Urdu.

### *2.3 Sprachaufnahmen*

Die Jugendlichen haben die beiden Situationen auf dem Bild unterschiedlich ausführlich beschrieben. Einige lieferten eher detailliertere Beschreibungen, zum Beispiel der Kleidung oder des im Fenster sichtbaren Hintergrunds. Ein Sprecher nahm eher eine Interpretation der Szene vor, die er als "chaotisch" bezeichnete und dabei nur am Rande beschrieb, welche Handlungen auf dem Bild dargestellt sind.

Insgesamt wurden rund 30 Minuten an spontansprachlichen Aufnahmen analysiert, wobei die Dauer der Bildbeschreibung zwischen den einzelnen Jugendlichen zum Teil stark variierte (Mittelwert = 36.6 Sekunden, Standardabweichung = 13.5 Sekunden, Min. = 12 Sekunden, Max. = 66 Sekunden).

### *2.4 Bewerter\_innen*

Kurze Ausschnitte dieser Bildbeschreibungen (5-7 Sekunden) wurden 40 Jugendlichen einer dritten Zürcher Schule vorgespielt.<sup>3</sup> Diese mussten einstufen, wie stark multiethnolektal diese Personen klingen; die Skala reichte von 1 (gar nicht multiethnolektal) bis 7 (komplett multiethnolektal). In der Skala erschien nicht der Begriff "multiethnolektal", sondern folgende Formulierung, die

---

<sup>3</sup> Für die Auswahl des Ausschnittes von 5-7 Sekunden wurde ein Zeitpunkt bestimmt, der exakt in der Mitte der ganzen Bildbeschreibung liegt. Vereinzelt wurden dabei längere Pausen zwischen zwei sehr kurzen Äusserungseinheiten gekürzt.

aufgrund der Resultate eines terminologischen Pilot-Tests gewählt worden war (vgl. Morand et al. akzeptiert):

(12) Inwiefern spricht diese Person Slang ("Ausländerdeutsch")?

Den Jugendlichen wurde zu Beginn erklärt, dass gewisse Erwachsene (z.B. Lehrpersonen oder Eltern) eine bestimmte Sprechweise als Slang oder eben "Ausländerdeutsch" bezeichnen. Den Schüler\_innen war sofort klar, was damit gemeint ist; sie bewerteten anschliessend alle Sprachausschnitte in einer randomisierten Reihenfolge. Aus dieser Einstufung wurden anschliessend für alle 48 Sprecher\_innen die individuellen *Screening Scores* berechnet. Diese ergeben sich jeweils aus dem Mittelwert aller Bewertungen pro Sprecher\_in und könnten theoretisch zwischen 1 (traditionelles Zürichdeutsch) und 7 (multiethnolektales Zürichdeutsch) liegen.

Die berechneten Mittelwerte erstrecken sich über einen grossen Teil der Skala und reichen von einem Minimum von 1.45 bis zum Maximum von 6.01. Im Durchschnitt wurden die Sprecher\_innen mit 3.78 eingestuft (Standardabweichung = 1.26). Wenn man die Sprecher\_innen abhängig von den *Screening Scores* sortiert, ergibt sich ein graduelles Kontinuum von eher traditionell zürichdeutschsprechend bis eher multiethnolektal zürichdeutschsprechend (vgl. Morand et al. akzeptiert).

## 2.5 Datenanalyse

### 2.5.1 Transkription und Annotation

Die Sprachausschnitte wurden gemäss den in der "Dialektschrift" von Dieth (1938/1986) ausgearbeiteten orthographischen Richtlinien transkribiert und mit dem Online-Tool *webMAUS* (Kisler et al. 2017) automatisch segmentiert. Der Output der automatischen Segmentation wurde anschliessend manuell kontrolliert und gegebenenfalls korrigiert.

### 2.5.2 Codierung

Für die Wortschatzanalyse wurden die Inhaltswörter (Nomen, Verben, Adjektive und Adverbien) der transkribierten Äusserungen in der Phonetik-Software *Praat* (Boersma & Weenink 2020) annotiert. Die restlichen Elemente der Äusserungen (Funktionswörter, Partikel, Häsitationen oder gefüllte Pausen) wurden nicht berücksichtigt. Die Inhaltswörter wurden anschliessend anhand einer varietätenlinguistischen Klassifikation mit fünf verschiedenen Codierungen versehen (siehe Tabelle 1).

Inhaltswörter, welche entweder klar Zürichdeutsch oder allgemein Schweizerdeutsch sind (*Bueb* anstatt *Junge*, *Fänschter* anstatt *Fenster*, *gsee* anstatt *sehe*) sowie solche, welche auf Zürichdeutsch und Standarddeutsch (quasi-)homophon sind (*Bild*, *Frau*, *Wasser*), wurden mit 0 codiert. Darunter fielen teilweise auch standarddeutsche Varianten, welche mithilfe phonologischer Anpassungen ins Zürichdeutsche integriert wurden (z.B.

*beschaffe*), aber vermutlich dem Sprecherkollektiv (vgl. Graf 2015: 69) weniger auffallen würden im Vergleich zur Codierung 2. Ebenfalls mit 0 codiert wurden einzelne zürichdeutsche Varianten, die analog zum Standarddeutschen gebildet wurden (z.B. *Chinder* analog zu *Kinder* anstatt *Chind* obwohl das traditionelle Zürichdeutsche für dieses Substantiv kein Pluralsuffix vorsieht).

Code	Beschreibung	Beispiele <sup>4</sup>
0	Klar Zürichdeutsch oder (quasi-)homophon auf Zürich- und Standarddeutsch	<i>Bueb, Fänschter, gsee</i> <i>Bild, Frau, Wasser</i>
1	Standarddeutsche lexikalische Interferenz ohne lautliche Anpassung	<i>Junge, Kekse, trägt</i>
2	Standarddeutsche lexikalische Interferenz mit lautlicher Anpassung	<i>arbäite, Küchi, Waschbeki</i>
3	Standarddeutsche lautliche Interferenz ohne lexikalische Interferenz	[nɔ <sup>h</sup> ʊr], [p <sup>h</sup> ʊtsə], [t <sup>h</sup> oxtər] [k <sup>h</sup> ɔs], [k <sup>h</sup> ɪnd], [klæuə]
4	unklar / Rest	Unverständliches oder englische Entlehnungen

Tabelle 1: Codierung der analysierten Wörter mit Beschreibung und Beispielen.

Mit 1 wurden lexikalische Interferenzen aus dem Standarddeutschen codiert, die lautlich nicht ans Zürichdeutsche angepasst wurden. Darunter fällt z.B. *Junge* anstatt *Bueb*, wobei die standarddeutsche Variante sehr auffällig scheint und deshalb als Interferenz codiert wurde. Bei *trägt* handelt es sich klar um eine lexikalische Interferenz, die lautlich nicht an den Dialekt angepasst wurde (z.B. *träit*); der typisch zürichdeutsche Ausdruck wäre *hätt aa* ("hat an").

Mit 2 wurden hingegen lexikalische Interferenzen aus dem Standarddeutschen codiert, die lautliche Anpassungen ans Zürichdeutsche aufweisen (z.B. *arbäite* anstatt *arbeiten*). Allerdings handelt es sich auch hier um Bedeutungen, für welche man im traditionellen Zürcher Dialekt klar einen anderen lexikalischen Typus verwendet (z.B. *schaffe*).

Neben den oben erwähnten untersuchten phonetischen Merkmalen des multiethnolektalen Zürichdeutschen (vgl. 1.1.1) fielen noch weitere Besonderheiten im Vergleich zum traditionellen Dialekt auf. So produzierten gewisse Jugendliche (stark) aspirierte Fortis-Plosive, welche im traditionellen Zürichdeutschen nicht erwartbar wären ([nɔ<sup>h</sup>ʊr] anstatt [nɔ<sup>h</sup>tʊr] oder [p<sup>h</sup>ʊtsə] anstatt [pʊtsə]). Zusätzlich wurden (wortinitiale) velare Frikative teilweise auch

<sup>4</sup> Die Beispiele für Codierung 0 und 2 werden in Dieth-Schrift (1938/1986) aufgeführt; für die Codierung 1 werden die standarddeutsche Orthographie und für die Codierung 3 phonetische Zeichen (IPA).

als velare Fortis-Plosive realisiert ([k<sup>h</sup>ind] anstatt [ɣind]). Solche Phänomene sind mit grosser Wahrscheinlichkeit auf standarddeutsche Interferenzen zurückzuführen und wurden deshalb in der vorliegenden Analyse als lautliche Interferenz mit 3 codiert.

Wie aus den Beispielen hervorgeht, kommen für die gleiche Bedeutung manchmal mehrere Varianten vor. So wird der auf der Zeichnung dargestellte Raum von den Jugendlichen z.T. traditionell Zürichdeutsch als *Chuchi* (0) bezeichnet, z.T. aber auch mit hybriden Formen wie *Küchi* (2) oder *Kuchi* (3) (siehe 3.2).

### 2.5.3 Datenauswertung

Für die Analyse wurden 1487 lexikalische Ausdrücke mit den Codierungen 0-3 betrachtet (Verlaufsformen wie z.B. *am Gschir spüele* wurden als einzelne lexikalische Ausdrücke kategorisiert). Ausgeschlossen wurden 52 Wörter (3.37%) mit der Codierung 4 (unklar / Rest); darunter fallen unverständliche Realisierungen und Versprecher (z.B. *verschiiinlich* als dialektale Kreuzung zwischen *warschiinlich* und *velicht*) sowie Entlehnungen aus dem Englischen (wie z.B. *Cookie*).

Die Codierung 0 (klar Zürichdeutsch oder homophon auf Zürich- und Standarddeutsch) kommt mit Abstand am häufigsten vor: 93% aller Inhaltswörter weisen weder lexikalische noch lautliche standarddeutsche Interferenzen auf. Ausserdem bestehen grosse Unterschiede bezüglich der Anzahl Wörter in den jeweiligen Wortartengruppen, weshalb diese nicht einzeln analysiert wurden.

Um die beiden Forschungsfragen (siehe 1.3) zu beantworten, wurden zwei gemischte logistische Regressionsmodelle mit R berechnet (Bates et al. 2015; R Development Core Team 2019). *Interferenz* wurde dabei als binäre abhängige Variable und *Screening Score* als kontinuierliche unabhängige Variable behandelt. In der ersten Analyse (siehe 3.1.1) wurde die Wahrscheinlichkeit einer Interferenz (Codierungen 1-3) mit der Wahrscheinlichkeit verglichen, dass es keine standarddeutsche Interferenz gab (Codierung 0). Bei der zweiten Analyse (siehe 3.1.2) wurde die Wahrscheinlichkeit lexikalischer Interferenzen (Codierung 1 und 2; ohne und mit lautlicher Anpassung) mit der Wahrscheinlichkeit rein lautlicher Interferenzen (Codierung 3) verglichen. Für beide Analysen wurde ein gemischtes Modell mit Sprecher\_innen als Zufallsvariable erstellt. In der zweiten Analyse wurden Sprecher\_innen ausgeschlossen, die entweder überhaupt keine Interferenzen produzierten (n = 12) oder keine Variabilität zwischen den beiden Arten von standarddeutschen Interferenzen aufwiesen (n = 7).



Die meisten Sprecher\_innen, die eher als multiethnolektal wahrgenommen wurden, haben einen grossen Anteil von lautlichen Interferenzen. Diese Beobachtung wird durch eine logistische Regressionsanalyse bestätigt ( $\beta = 0.42$ ,  $z = 2.02$ ,  $p = .04$ ). Aufgrund des *Screening Scores* ist die Wahrscheinlichkeit des Vorkommens von lautlichen Interferenzen vorhersagbar: Je multiethnolektaler die Sprecher\_innen wahrgenommen wurden, desto häufiger weisen sie lautliche statt lexikalische Interferenzen auf. Ausserdem erscheinen bei einzelnen Sprecher\_innen nur lexikalische Interferenzen (0% phonetische Interferenzen in Abbildung 2) oder nur lautliche Interferenzen (100% phonetische Interferenzen in Abbildung 2).

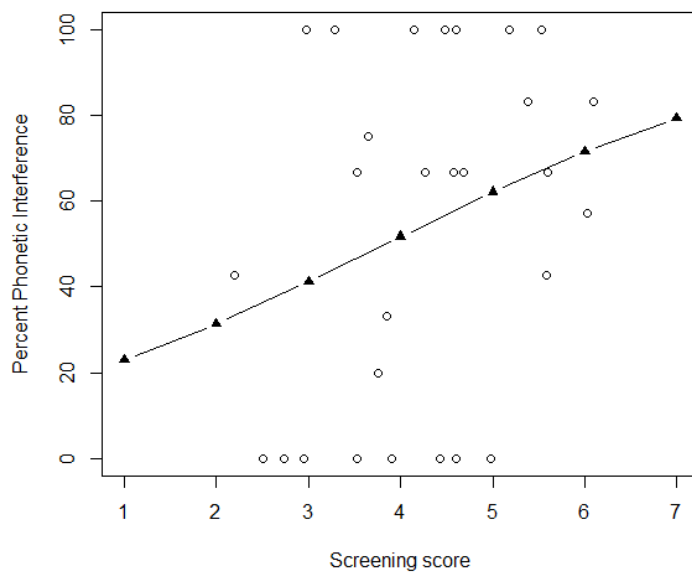


Abbildung 2: Prozentualer Anteil von beobachteten und vorhergesagten lautlichen Interferenzen (Punkte bzw. Dreiecke) im Vergleich zu den lexikalischen Interferenzen in Abhängigkeit des Screening Scores (jeder Punkt stellt eine\_n Sprecher\_in dar,  $n = 29$ ).

### 3.2 Weitere Beobachtungen

Neben den prüfstatischen Analysen sollen an dieser Stelle einige weitere, z.T. eher qualitative Beobachtungen ausgeführt werden. Bereits vor der Analyse fielen bei mehreren Jugendlichen standarddeutsche (und englische) Entlehnungen auf. So wurde das Gebäck, das vom Jungen im Bild gestohlen wird, in den 41 wörtlichen Benennungen im Datensatz als *Keks(e)*, *Guetzli*, *Cookies* oder *Süesses* bezeichnet (in absteigender Reihenfolge). Wie bereits erwähnt, wurden englische Lehnwörter für die Analyse ausgeschlossen (Codierung 4); in diesem Fall ist der Anglizismus sicher auch damit erklärbar, dass das Gefäss im Bild mit "Cookie Jar" beschriftet ist. Dies könnte vielleicht auch die Bezeichnung *Keks(e)* ausgelöst haben, da beide Bezeichnungen mit dem velaren Fortis-Plosiv /k/ beginnen. Wenn wir die eindeutig zürichdeutsche Variante *Guetzli* mit der standarddeutschen Interferenz *Keks(e)* vergleichen, fällt auf, dass Letztere etwa doppelt so häufig verwendet wurde (68% im



Vergleich zu 32%, siehe Tabelle 2). Ähnlich starke Interferenzen lassen sich auch bei der Bezeichnung des Raums feststellen (52% im Vergleich zu 48%, siehe Tabelle 2).

Objekt	Codierung 0	Codierung 1	Codierung 2	Codierung 3
<i>Gebäck</i>	13 (32%)	26 (63%)	2 (5%)	0 (0%)
<i>Raum</i>	10 (48%)	0 (0%)	8 (38%)	3 (14%)

Tabelle 2: Wörtliche Bezeichnungen des Gebäcks und des Raums aufgeteilt nach den Codierungen. Absolute Werte (Zeilenprozente in Klammern).

Bei den 21 Bezeichnungen des Raumes finden wir zur Hälfte standarddeutsche Interferenzen, die restlichen Bezeichnungen entsprechen dem traditionellen zürichdeutschen Lexem *Chuchi*. Die Interferenzen sind entweder rein lautlicher Natur (Codierung 3), wobei der wortinitiale velare Frikativ in drei Fällen durch einen velaren Fortis-Plosiv ersetzt wird. Bei den restlichen acht Fällen finden wir eine hybride Form, d.h. eine standarddeutsche lexikalische Interferenz mit lautlicher Anpassung (*Küchi*), welche mit 2 codiert wurde.

Standarddeutsche Interferenzen treten auch bei anderen Wortarten auf, wobei aber keine einzelnen Bezeichnungen hervorzuheben sind, sondern eher grundsätzliche Tendenzen bezüglich lautlicher Interferenzen der wortinitialen (aspirierten) Fortis-Plosive. Allerdings ist unklar, inwieweit dieses Phänomen Teil des Multiethnolekts ist oder vielmehr auf einen allgemeinen beginnenden Lautwandel hinweist, wonach innerhalb schweizerdeutscher Dialekte bilabiale und alveolare Fortis-Plosive vermehrt aspiriert werden (vgl. 1.2.1). Bei den velaren Fortis-Plosiven kann eine solche Entwicklung nicht eintreten, da diese im traditionellen Zürichdeutschen als Affrikaten oder velare Frikative realisiert werden.

#### 4. Diskussion

Wie eingangs erwähnt, wird die Sprachsituation der Deutschschweiz in der Vergangenheit meist als Diglossie beschrieben, weshalb gemäss Hove (2008: 65) eine Äusserung für Deutschschweizer Sprecher\_innen und Hörer\_innen leicht einer der beiden verwendeten Varietäten zugeordnet werden kann; für aussenstehende Hörer\_innen können bestimmte Äusserungen manchmal aber auch ambig sein. Die Verwendung von Wortformen aus der jeweils anderen Varietät wird innerhalb einer Äusserung normalerweise als *Code-Switching* interpretiert, da kein Kontinuum zwischen den beiden Varietäten besteht. Falls es sich aber nicht um mehr oder weniger absichtliches *Code-Switching* handelt, kann man von Entlehnungen oder lexikalischen Interferenzen sprechen.

Im vorliegenden Beitrag wurden standarddeutsche Interferenzen bei ein- und mehrsprachigen Jugendlichen zweier Zürcher Sekundarschulen in einer

Wortschatzstudie empirisch untersucht. Zu diesem Zweck wurde ein zuvor erhobener *Screening Score* herangezogen, welcher angibt, wie multiethnolektal die Sprechweise der Jugendlichen im Vergleich zum traditionellen Dialekt wahrgenommen wird, anstatt die Jugendlichen in vordefinierte Gruppen einzuteilen. Dabei wurden die beiden Hypothesen aufgestellt, dass (a) im eher multiethnolektalen Zürichdeutsch mehr standarddeutsche Interferenzen auftreten als im traditionellen Dialekt, und dass (b) bei eher multiethnolektal sprechenden Jugendlichen rein lautliche Interferenzen im Vergleich häufiger sind als lexikalische Interferenzen.

In den untersuchten zürichdeutschen Daten finden wir zwar standarddeutsche Interferenzen, allerdings sind diese nicht sehr zahlreich. Nur gut 7% der analysierten Wörter wurden als offensichtliche standarddeutsche Interferenzen kategorisiert. Die restlichen knapp 93% der analysierten Wörter sind entweder im traditionellen Zürichdeutschen und (Schweizer) Standarddeutschen homophon oder können klar dem Zürich- bzw. allgemeinen Schweizerdeutschen zugewiesen werden.

Der *Screening Score* hat aber in den statistischen Modellen einen signifikanten Effekt auf die Wahrscheinlichkeit einer Interferenz – insbesondere in Bezug auf lexikalische Interferenzen mit lautlicher Anpassung (z.B. *Küchi*) und auf rein lautliche Interferenzen (z.B. aspirierte Fortis-Plosive). Je multiethnolektaler bestimmte Sprecher\_innen wahrgenommen werden, desto häufiger sind bei ihnen standarddeutsche Interferenzen zu beobachten.

Ähnlich signifikante Effekte des *Screening Scores* stellt man im statistischen Modell fest beim Vergleich von lexikalischen und lautlichen Interferenzen. So steigt die Wahrscheinlichkeit einer rein lautlichen Interferenz im Gegensatz zu lexikalischen Interferenzen mit steigendem *Screening Score*. Je multiethnolektaler Sprecher\_innen wahrgenommen wurden, desto eher zeigten sie lautliche anstatt lexikalische Interferenzen.

Bei den lautlichen Interferenzen handelt es sich insbesondere um (wortinitiale) Fortis-Plosive, welche entweder aspiriert werden ( $[p^h t^h]$  anstatt  $[p t]$ ) oder um velare Frikative, die als Plosive realisiert werden ( $[k]$  anstatt  $[x \text{ ʔ}]$ ). In Bezug auf die Aspiration der bilabialen und alveolaren Fortis-Plosive gibt es Hinweise auf einen Lautwandel, bei welchem neben den typischen Entlehnungen wie  $[t^h e:]$  und  $[p^h a'kxet]$  auch bei Sprecher\_innen, die eigentlich nicht als multiethnolektal wahrgenommen werden, vermehrt Behauchung zu finden ist. So bemerken Leemann et al. (2020) in ihrer Pilotstudie einen Alterseffekt (mehr Aspiration in der jüngeren Kohorte), den sie als mögliche Anpassung an das Standarddeutsche interpretieren. In unserer zukünftigen Forschung soll die Hypothese des beginnenden Lautwandels (und die allfällige Rolle des Multiethnolekts) näher und systematischer verfolgt werden, da in den vorliegenden Bildbeschreibungen nur einzelne wenige Wörter mehrfach mit anlautendem Fortis-Plosiv erscheinen (z.B. *Tochter*, *Täller* oder *putze*).

Bezüglich der für Jugendliche mit Migrationshintergrund postulierten (strukturellen) Dialekt-Transformation (Schmid 2020) liefern die Ergebnisse der vorliegenden Wortschatzstudie erste Hinweise dafür, dass als multiethnolektal wahrgenommene Sprecher\_innen in ihrem Wortschatz z.T. standarddeutsche Interferenzen aufweisen. Für Kinder, die erst im Schulalter Schweizerdeutsch simultan mit Standarddeutsch erwerben, kann sich die Zuordnung bestimmter Lexeme zu den beiden Varietäten insofern als schwierig erweisen, als zahlreiche Wörter homophon in beiden Varietäten sind. So könnten sich bei diesen Sprecher\_innen allenfalls Ausdrücke im gemeinsamen Lexikon befinden (vgl. Häcki Buhofer & Burger 1998), welche traditionell eher dem Standard-Lexikon zugeordnet würden.

## 5. Abschliessende Bemerkungen

Die Ergebnisse der vorliegenden Studie zeigen, dass im dialektalen Wortschatz jugendlicher Sprecher\_innen der Stadt Zürich vermehrt standarddeutsche Interferenzen zu finden sind, welche aufgrund der Diglossiesituation im traditionellen Zürichdeutschen nicht erwartbar wären. Bei Jugendlichen mit Migrationshintergrund scheint zumindest ansatzweise eine Vermischung von Dialekt und Standardvarietät zu entstehen.

Zwar zeigt die quantitative Analyse, dass standarddeutsche Interferenzen im dialektalen Wortschatz nach wie vor die Ausnahme bilden; die Resultate zeigen aber auch, dass bei einem höheren *Screening Score* im Perzeptionsexperiment (Sprecher\_innen werden als stärker multiethnolektal zürichdeutschsprechend wahrgenommen) einerseits im Dialekt standarddeutsche Interferenzen mit grösserer Wahrscheinlichkeit auftreten und dass andererseits auch die Wahrscheinlichkeit von lautlichen im Gegensatz zu lexikalischen Interferenzen zunimmt. Zukünftige Forschungsarbeiten werden untersuchen können, welche Zusammenhänge zwischen Multiethnolekt und Jugendsprache bestehen und inwieweit in Bezug auf bestimmte Phänomene (etwa die Aspiration von Fortis-Plosiven) Anzeichen auf einen beginnenden Lautwandel festzustellen sind.

## Danksagung

Diese Forschungsarbeit wurde durch den Schweizerischen Nationalfonds (Nr. 165798) finanziert. Bedanken möchten wir uns insbesondere bei den Schüler\_innen und ihren Lehrkräften für die Teilnahme am Experiment. Ausserdem sind wir den studentischen Hilfskräften Melissa Bruno, Nora Julmi und Seraina Nadig für ihre Mitarbeit bei der Datenerhebung und -aufbereitung zu Dank verpflichtet. Dankbar sind wir auch den beiden anonymen Gutachter\_innen für ihre äusserst hilfreichen Kommentare. Für alle Unzulänglichkeiten dieses Beitrags bleiben wir allein verantwortlich.

## LITERATUR

- Ammon, U. (2003). Dialektschwund, Dialekt-Standard-Kontinuum, Diglossie: Drei Typen des Verhältnisses Dialekt – Standardvarietät im deutschen Sprachgebiet. In J. K. Androutsopoulos & E. Ziegler (Hgg.), *"Standardfragen": Soziolinguistische Perspektiven auf Sprachgeschichte, Sprachkontakt und Sprachvariation* (S. 163-171). Bern: Peter Lang.
- Auer, P. (2003). 'Türkenslang': Ein jugendsprachlicher Ethnolekt des Deutschen und seine Transformationen. In A. Häcki Buhofer (Hg.), *Spracherwerb und Lebensalter* (S. 255-264). Tübingen/Basel: Francke.
- Barbour, S. & Stevenson, P. (1990). *Variation in German: A critical approach to German sociolinguistics*. Cambridge: Cambridge University Press.
- Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Berthele, R. (2004). Vor lauter Linguisten die Sprache nicht mehr sehen: Diglossie und Ideologie in der deutschsprachigen Schweiz. In H. Christen (Hg.), *Dialekt, Regiolekt und Standardsprache im sozialen und zeitlichen Raum* (S. 111-136). Wien: Praesens.
- Boersma, P. & Weenink, D. (2020). *Praat: Doing phonetics by computer* [Computerprogramm]. 6.1.38.
- Christen, H. (2000). Standardsprachliche Varianten als stilistische Dialektvarianten? In A. Häcki Buhofer (Hg.), *Vom Umgang mit sprachlicher Variation. Soziolinguistik, Dialektologie, Methoden und Wissenschaftsgeschichte* (S. 245-260). Tübingen/Basel: Francke.
- Clyne, M. (2000). Lingua franca and ethnolects in Europe and beyond. *Sociolinguistica*, 14, 83-89.
- Dieth, E. (1938/1986). *Schwyzertütschi Dialäktschrift: Dieth-Schreibung* (2. bearbeitete Version herausgegeben von C. Schmid-Cadalbert). Aarau: Sauerländer.
- Dirim, İ. & Auer, P. (2004). *Türkisch sprechen nicht nur die Türken: Über die Unschärfebezeichnung zwischen Sprache und Ethnie in Deutschland*. Berlin: De Gruyter.
- Ender, A., Li, W. & Straßl, K. (2007). Das Projekt "Deutsch als Zweitsprache in Dialektumgebung". *Linguistik online*, 32(3), 25-36.
- Ender, A. & Straßl, K. (2009). The acquisition and use of German in a dialect-speaking environment: Facets of inclusion and exclusion of immigrant children in Switzerland. *International Journal of Applied Linguistics*, 19(2), 173-187.
- Ferguson, C. (1959). Diglossia. *Word*, 15(2), 325-340.
- Fleischer, J. & Schmid, S. (2006). Zurich German. *Journal of the International Phonetic Association*, 36, 243-255.
- Goodglass, H. & Kaplan, E. (1983). *Boston Diagnostic Aphasia Examination (BDAE)*. Distributed by Psychological Assessment Resources, Odessa, FL.
- Goschler, J., Woerfel, T., Stefanowitsch, A., Wiese, H. & Schroeder, C. (2013). Beyond conflation patterns: The encoding of motion events in Kiezdeutsch. *Yearbook of the German Cognitive Linguistics Association*, 1(1), 237-254.
- Graf, M. H. (2015). Frisst das Pfärd jetzt Kartoffle? *Sprachspiegel*, 3, 66-77.
- Häcki Buhofer, A. & Burger, H. (1998). *Wie Deutschschweizer Kinder Hochdeutsch lernen: Der ungesteuerte Erwerb des gesprochenen Hochdeutschen durch Deutschschweizer Kinder zwischen sechs und acht Jahren*. Stuttgart: Franz Steiner.
- Häcki Buhofer, A., Schneider, H. & Beckert, C. (2007). Mehrsprachige Jugendliche im Umgang mit Dialekt und Hochsprache in der Deutschen Schweiz. *Linguistik online*, 32(3), 49-70.

- Hove, I. (2008). Zur Unterscheidung des Schweizerdeutschen und der (schweizerischen) Standardsprache. In H. Christen & E. Ziegler (Hgg.), *Sprechen, Schreiben, Hören: Zur Produktion und Perzeption von Dialekt und Standardsprache zu Beginn des 21. Jahrhunderts* (S. 63-82). Wien: Praesens.
- Kehrein, R. (2019). Areale Variation Im Deutschen "Vertikal". In J. Herrgen & J. E. Schmidt (Hgg.), *Sprache und Raum: Ein internationales Handbuch der Sprachvariation* (S. 121-158). Berlin, Boston: De Gruyter Mouton.
- Kisler, T., Reichel, U. D. & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, 45, 326-347.
- Ladd, D. R. & Schmid, S. (2018). Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops. *Journal of Phonetics*, 71, 229-248.
- Leemann, A., Jeszenszky, P., Steiner, C. Messerli, J. & Studerus, M. (2020). Behauchung von Plosiven im Schweizerdeutschen: Ein Merkmal junger DialektsprecherInnen? Vortrag an der Tagung *Phonetik & Phonologie*, 16. Trier, 11.-12. September 2020 (online).
- Morand, M.-A., Bruno, M., Julmi, N., Schwab, S. & Schmid, S. (2019a). The voicing of lenis plosives in Zurich German: A sociophonetic marker of (multi-)ethnolectal speech. In S. Calhoun, P. Escudero, M. Tabain & P. Warren (Hgg.), *Proceedings of the 19th International Congress of Phonetic Sciences* (S. 1600-1604). Canberra: Australasian Speech Science and Technology Association Inc.
- Morand, M.-A., Bruno, M., Julmi, N., Schwab, S. & Schmid, S. (2019b). Duration of word-initial fricatives in Zurich German: A sociophonetic marker of (multi-)ethnolectal speech. Vortrag an der Tagung *Phonetik & Phonologie 15*. Düsseldorf, 26.-27. September 2019.
- Morand, M.-A., Bruno, M., Julmi, N., Schwab, S. & Schmid, S. (2020). Sandhi im multiethnolectalen Zürichdeutschen. Vortrag an der Tagung *Phonetik & Phonologie 16*. Trier, 11.-12. September 2020 (online).
- Morand, M.-A., Schwab, S. & Schmid, S. (akzeptiert). The perception of multiethnolectal Zurich German: A continuum rather than clear-cut categories. *Loquens*.
- Petkova, M. (2011). Zwischen Dialekt und Standardsprache: Code-Hybridisierung in der Deutschschweiz. In E. Glaser, J. E. Schmidt & N. Frey (Hgg.), *Dynamik des Dialekts: Wandel und Variation* (S. 241-266). Stuttgart: Steiner Verlag.
- Petkova, M. (2016). "Es Hochtitsch, won\_ich mues säge: *Hey, was soll das!*": Code-Switching als interaktionale Ressource im Deutschschweizer Kontext. *Gesprächsforschung: Online-Zeitschrift zur verbalen Interaktion*, 17, 231-249.
- Quist, P. (2000). Ny københavnsk "multietnolekt": Om sprogbrug blandt unge i sprogligt og kulturelt heterogene miljøer. *Danske Talesprog*, 1, 143-212.
- Quist, P. (2010). The sociolinguistic study of youth and multilingual practices in Denmark. In P. Quist & B. A. Svendsen (Hgg.), *Multilingual urban Scandinavia: New linguistic practices* (S. 6-11). Bristol: Multilingual Matters.
- R Core Team (2019). *R: A language and environment for statistical computing* [Computerprogramm]. 4.0.3.
- Rampton, B. (2005). *Crossing: Language and ethnicity among adolescents*. Manchester: St. Jerome.
- Schifferle, H.-P. (2010). Zunehmende Behauchung: Aspirierte Plosive im modernen Schweizerdeutsch. In H. Christen, S. Germann, W. Haas, N. Montefiori & H. Ruedi (Hgg.), *Alemannische Dialektologie: Wege in die Zukunft* (S. 43-55). Stuttgart: Steiner.
- Schirmunski, V. M. (1962). *Deutsche Mundartkunde*. Berlin: Akademie-Verlag.
- Schmid, S. (2020). Swiss German dialects spoken by second-generation immigrants: Bilingual speech and dialect transformation. *Journal of Multilingual and Multicultural Development*, <https://doi.org/10.1080/01434632.2020.1730386>

- Siebenhaar, B. & Wyler, A. (1997). *Dialekt und Hochsprache in der deutschsprachigen Schweiz*. Zürich: Pro Helvetia.
- Tissot, F., Schmid, S. & Galliker, E. (2011). Ethnolektales Schweizerdeutsch: soziophonetische und morphosyntaktische Merkmale sowie ihre dynamische Verwendung in ethnolektalen Sprechweisen. In E. Glaser, J. Schmidt & N. Frey (Hgg.), *Dynamik des Dialekts: Wandel und Variation* (S. 319-344). Stuttgart: Steiner.

# Der Wortschatz im Kindesalter: Entwicklung und Intervention

**Christina KAUSCHKE**

Universität Marburg  
Institut für Germanistische Sprachwissenschaft  
Pilgrimstein 16, D-35037 Marburg  
kauschke@uni-marburg.de

This paper gives an overview of developmental sequences in the acquisition of the mental lexicon and presents intervention approaches for children with lexical impairments. Vocabulary acquisition is a lifelong and multidimensional process. In its course, children learn to segment word forms from the input, link them to concepts, enrich concrete and abstract words with semantic information, build up a repertoire of parts of speech, apply word formation patterns, and retrieve the appropriate word forms in interaction contexts. Some children need special support because either their ambient language is insufficient, or internal mechanisms that enable effective lexical acquisition or word retrieval are limited. Parent- and child-centered intervention approaches have been proposed to enhance lexical breadth and depth and to reduce potential negative consequences for further learning and development. Essential components of successful lexical interventions comprise the selection of target words controlled for linguistic parameters, frequent presentation of these target words, elaboration of the semantic and phonological characteristics, multiple opportunities for word production, and strategies to cope with word finding difficulties. Intervention should be tailored to the individual needs, and combine direct and indirect methods.

## **Keywords**

vocabulary growth, fast and slow mapping, parts of speech, lexical disorders, intervention.

## **Stichwörter**

Wortschatzwachstum, fast und slow mapping, Wortarten, lexikalische Störungen, Intervention.

## **1. Entwicklungssequenzen im Lexikonerwerb**

Der Aufbau des Wortschatzes in einer oder mehreren Sprachen ist ein Prozess, der im Säuglingsalter beginnt und sich lebenslang fortsetzt. Wörter zu lernen ist eine vielschichtige und komplexe Aufgabe, bei der akustisch wahrnehmbaren Lautketten (den "Wortformen") Bedeutung zugewiesen wird (He & Arunachalam 2017). Wenn im Folgenden von "Wörtern" gesprochen wird, ist damit immer eine Verbindung zwischen einer Wortform und der zugehörigen Bedeutung gemeint. Wörter werden als Lexikoneinträge im mentalen Lexikon (siehe Dietrich & Gerwien 2017) gespeichert. Dieses meint einen Teil des Langzeitgedächtnisses, das ein wachsendes Repertoire unterschiedlichster Wörter umfasst und Wissen über die Form, Bedeutung und Verwendung von Wörtern sowie über die Verbindungen zwischen Wörtern enthält. Im Laufe der Zeit entwickelt sich das mentale Lexikon in die Breite und in die Tiefe, d.h. das Vokabular wird diverser und umfangreicher (Breite) und das Wissen über die jeweiligen Wörter wird detaillierter und spezifischer (Tiefe). Wie das mentale Lexikon im Laufe der kindlichen Entwicklung angelegt, erweitert und strukturiert

wird, welche Schwierigkeiten dabei auftreten können und wie diesen begegnet werden kann, ist Thema des vorliegenden Beitrags<sup>1</sup>. Dabei liegt der Schwerpunkt auf dem monolingualen Spracherwerb sowie auf der Entwicklung bis zur Einschulung, mit Ausblicken auf das Schulalter.

## 1.1 Lexikonerwerb in der frühen Phase (bis 3 Jahre)

### 1.1.1 Segmentation

Wortformen müssen zunächst aus dem kontinuierlichen Sprachstrom, dem Kinder von Geburt an ausgesetzt sind, heraus isoliert werden, bevor ihnen Bedeutung zugewiesen werden kann. Das Kind muss also erkennen, welche lautlichen Abfolgen wiederkehrende Einheiten bilden, die die Wörter und damit die Grundbausteine der zu erwerbenden Sprache darstellen. Säuglinge wenden bereits effiziente Mechanismen an, um ihre Umgebungssprache(n) zu segmentieren (Höhle 2015). Um festzustellen, wo die Grenzen zwischen Wörtern verlaufen, bedienen sich Kinder ab einem Alter von etwa sechs Monaten verschiedener Strategien, d.h. sie richten ihre Aufmerksamkeit auf bestimmte Eigenschaften des Sprachangebots und leiten daraus Annahmen über Wortgrenzen ab. Ein wichtiger früher Hinweisreiz ist die Wortbetonung: so nimmt ein Deutsch lernendes Kind ausgehend vom vertrauten Betonungsmuster seiner Muttersprache an, dass Wörter in der Regel mit betonten Silben beginnen und eine Wortgrenze daher vor einer betonten Silbe liegen muss. Mit Hilfe dieses so genannten "prosodischen *bootstrappings*" finden Kinder den Einstieg in die Wortidentifikation und damit in die Lexikonentwicklung (Schröder & Höhle 2011). Ab etwa acht Monaten können sie zuvor gehörte Inhalts- und Funktionswörter wiedererkennen. Das Erkennen eines häufigen, schon vertrauten Wortes, auch des eigenen Namens, stellt einen Ankerpunkt im Sprachstrom dar, so dass vor und nach diesem Wort eine Wortgrenze angenommen werden kann (Höhle 2015).

Nicht allen Kindern gelingt mit Hilfe derartiger Strategien ein effektiver und müheloser Einstieg in den Lexikonerwerb. Retrospektive Gruppenstudien fanden Hinweise darauf, dass Kinder, die später von Sprachauffälligkeiten betroffen sind, bereits in dieser frühen Phase schwach ausgeprägte Segmentierungsfähigkeiten zeigen. So schneiden Kinder, die in den ersten Lebensmonaten Betonungsmuster schlecht voneinander unterscheiden bzw. Wörter schlecht segmentieren können, während des Vorschulalters in Bezug auf ihren Vokabularumfang und/oder ihre grammatischen Fähigkeiten schlechter ab als Kinder mit typischer Sprachentwicklung (TSE; Friedrich et al. 2004; Höhle et al. 2014; Weber et al. 2005). Offenbar kann sich eine verzögerte

---

<sup>1</sup> Einige Inhalte dieses Beitrags werden in ausführlicherer Form in Kauschke (2012) und (2017b) dargestellt. Dort finden sich umfangreiche Literaturangaben und Hintergründe. Im vorliegenden Beitrag wird aus Platzgründen vor allem auf ausgewählte, neuere Referenzen zum Lexikonerwerb im Deutschen verwiesen. Zum Kontext der Mehrsprachigkeit siehe Till in diesem Band.



oder problematische Sprachentwicklung bereits im ersten Lebensjahr durch eingeschränkte Sprachverarbeitungsfähigkeiten ankündigen, auch wenn dies ohne experimentelle Untersuchungsmethoden noch nicht im Einzelfall nachweisbar ist.

### 1.1.2 Mapping

Nachdem Wortformen im Input erkannt und gespeichert wurden, können diese mit Konzepten verbunden werden. Konzepte entstehen durch Erfahrungen des Kindes in und mit seiner Umgebung und deren anschließende mentale Repräsentation. Werden Konzepte (bzw. Teile derselben) mit einer Wortform verknüpft, entstehen Wörter als Einträge im mentalen Lexikon mit einer Form- und einer Inhaltsseite. Die Kopplung der beiden Komponenten Wortform und Wortbedeutung wird als "*mapping*" bezeichnet. Nun entstehen Wörter als bedeutungsvolle Lautketten, die verwendet werden, um etwas Außersprachliches zu repräsentieren oder um kommunikative Funktionen zu erfüllen. Anfangs sind sowohl die gespeicherten Wortformen als auch die mit ihnen verbundenen Bedeutungsmerkmale noch rudimentär und vorläufig (Samuelson & McMurray 2017).

### 1.1.3 Wortschatzwachstum

Anfänge des Wortverstehens lassen sich unter Berücksichtigung von Blickbewegungsdaten bereits mit 6 Monaten nachweisen (Rohlfing 2019). Ab etwa zehn Monaten ist das Verständnis für Wörter als Kopplungen von Wortformen mit Bedeutung stabil nachweisbar. Darauf folgt ein rasantes Anwachsen des rezeptiven Wortschatzes, das von der beachtlichen Fähigkeit von Kindern zeugt, schnell eine wachsende Anzahl von Wörtern in das mentale Lexikon aufzunehmen ("*fast mapping*"). Dabei geht das Wortverständnis nicht nur zeitlich der Wortproduktion voraus, der Umfang des rezeptiven Wortschatzes übersteigt außerdem den des expressiven (Kauschke 2007).

Die Wortproduktion beginnt gegen Ende des ersten Lebensjahres mit kontextgebundenen Protowörtern, die in spezifische Handlungs- und Situationskontexte eingebettet sind (z.B. 'Ball' wird nur geäußert, während sich Mutter und Kind einen Ball zurollen). Ab etwa 13 Monaten folgen die ersten "echten" Wörter, die sich durch eine Loslösung vom Situationskontext auszeichnen. Im Gegensatz zu den Protowörtern werden echte Wörter kontextflexibel und symbolisch verwendet. Die Inhalte dieser frühen Wörter beziehen sich auf das unmittelbare Erleben der Kinder und ihr direktes Umfeld, d.h. auf Menschen, Tätigkeiten und Gegenstände der kindlichen Lebenswelt. Typische semantische Phänomene sind Verengungen oder Erweiterungen des Wortgebrauchs im Vergleich zur zielsprachlichen Verwendung, die vor allem im zweiten und dritten Lebensjahr auftreten: Bei der Übergeneralisierung (siehe Ambridge et al. 2013) wird ein Wort auf nicht passende Vertreter ausgedehnt ('Ball' als Bezeichnung für alle runden Gegenstände), während es bei einer

Unterdehnung nicht auf alle möglichen Bezugsobjekte bezogen wird ('Ball' nur für Fußbälle, nicht für Tennisbälle). Die produzierten Wortformen sind noch nicht stabil, sondern von Inkonsistenzen gekennzeichnet und gegenüber der zielsprachlichen Aussprache vereinfacht (z.B. /ba/ für 'Ball'; zur phonologischen Entwicklung im Deutschen siehe Fox-Boyer & Schäfer 2015).

In dieser frühen Phase, die etwa ein halbes Jahr andauern kann, wächst der Wortschatz zunächst langsam an. Ein Meilenstein ist erreicht, wenn das expressive Vokabular etwa 50 Wörter umfasst, was typischerweise mit circa 18 Monaten der Fall ist, wobei hinsichtlich des Zeitpunkts eine starke individuelle Variation zu berücksichtigen ist. Der durchschnittliche produktive Wortschatz von zweijährigen Kindern mit TSE übersteigt nach den Ergebnissen zahlreicher Studien in verschiedenen Sprachen 200 Wörter, wenn Elternangaben aus Vokabularchecklisten zugrunde gelegt werden<sup>2</sup>. Vielfach wurde nachgewiesen, dass der Wortschatz im zweiten und dritten Lebensjahr substantiell anwächst. Diese merkliche Beschleunigung der Wachstumsgeschwindigkeit wird als "Vokabularspurt" bezeichnet. Obwohl über die genauen Muster und Hintergründe des Wortschatzspurts keine Einigkeit in der Forschung besteht (Samuelson & McMurray 2017), stimmen die meisten Modellierungen dahingehend überein, dass eine Phase mit einem supralinearen bzw. exponentiellen Wachstum angenommen wird, d.h. der Wortschatz wächst in einem bestimmten Zeitraum nicht konstant, sondern beschleunigt an. Nachdem der Aufbau des Vokabulars auf diese Weise in Gang gekommen ist, findet weiterhin eine Zunahme und Ausdifferenzierung statt (siehe Abschnitte 1.2 und 1.3).

Nachgewiesene Assoziationen zwischen dem Wortschatzumfang und späteren Sprachkompetenzen machen die frühen lexikalischen Fähigkeiten zu einem wichtigen prognostischen Faktor für den weiteren Verlauf der Sprachentwicklung. Zweijährige Kinder mit einer verzögerten und verlangsamten Lexikonentwicklung werden als "*Late Talkers*" bezeichnet; sie tragen ein deutlich erhöhtes Risiko für spätere Auffälligkeiten in der Sprachentwicklung (Kauschke 2017a; Sachse 2015). Bei diesen Kindern bleibt der Wortschatzspurt im zweiten Lebensjahr aus, so dass ihr Wortschatz, dessen Umfang meist mit Hilfe von Elternfragebögen eingeschätzt wird, mit 24 Monaten deutlich kleiner ist als derjenige gleichaltriger Kinder mit TSE. Auch wenn es etwa einem Drittel der *Late Talkers* gelingt ihren Rückstand bis zum dritten Geburtstag aufzuholen, zeigen viele von ihnen später Sprachentwicklungsstörungen (SES) oder behalten bis ins Jugendalter sprachliche Schwächen bei (Rescorla 2015). Laut Ergebnissen einer Längsschnittstudie (Kühn et al. 2016) wurden 19% ehemaliger *Late Talkers* kurz vor der Einschulung als sprachschwach eingestuft, weitere 16% als

<sup>2</sup> Daten zum Vokabularumfang in unterschiedlichen Sprachen lassen sich in der "Wordbank"-Datenbank (<http://wordbank.stanford.edu>) einsehen, siehe dazu Frank et al. (2016).

sprachentwicklungsgestört. In einer Studie von Armstrong et al. (2017) zeigte sogar die Hälfte der ehemaligen *Late Talkers* noch mit 10 Jahren unterdurchschnittliche Leistungen in Sprachtests. Obwohl Armstrong und Kolleg\_innen betonen, dass die Datenlage zur langfristigen Entwicklung uneinheitlich ist, gehen sie davon aus, dass 20 bis 50% ehemaliger *Late Talkers* längerfristige Probleme beibehalten. Daher können Frühinterventionsmaßnahmen sinnvoll sein (Abschnitt 2.1), um die weitere Entwicklung frühzeitig positiv zu beeinflussen.

#### 1.1.4 Wortarten

Über die bloße Menge an Wörtern hinaus ist auch die Zusammensetzung des Lexikons in Bezug auf verschiedene Wortarten maßgeblich. Charakteristische Wortarten für die frühe Phase sind interaktive Wörter ('hallo', 'ja'), relationale Wörter ('da'), Verbpartikel ('auf'), Lautmalereien ('brumm') sowie Eigennamen ('Mama'). Die Zusammensetzung des Wortschatzes verändert sich im weiteren Verlauf dynamisch: die im zweiten Lebensjahr vorherrschenden Kategorien der relationalen und sozialen Wörter nehmen ab, während die Komposition des Lexikons durch weitere Wortarten ergänzt und vervollständigt wird (Kauschke & Hofmeister 2002). Für verschiedene Sprachen wurde ein ähnlicher, wellenartiger Verlauf der Wortartenentwicklung beschrieben: nach den anfänglich dominierenden Wortarten zeigt sich zunächst eine Expansion von Nomen als Gattungsbegriffen, der ein linearer Anstieg von Verben folgt. Gegen Ende des dritten Lebensjahres markiert eine Zunahme der Funktionswörter die letzte Welle in der Wortartenentwicklung. Eine nicht altersgemäße Komposition des Lexikons mit 18 Monaten (weniger Nomen und Verben, mehr interaktive und relationale Wörter) ist ein frühes Kennzeichen für eine verzögerte Sprachentwicklung (Lüke et al. 2017).

#### 1.1.5 Wortgebrauch

Wörter sind konventionell festgelegte sprachliche Symbole, mit denen Sprechende auf außersprachliche Inhalte verweisen und diese mit anderen Mitgliedern ihrer Sprachgemeinschaft teilen können. Insofern ist Worterwerb nicht von der sozial-kommunikativen und der kognitiven Entwicklung abzukoppeln. Bevor Kinder mit Wörtern kommunizieren, durchlaufen sie in den ersten beiden Lebensjahren wichtige Vorläuferprozesse dazu. Zentral ist die Fähigkeit, Aufmerksamkeit zu koordinieren und somit gemeinsame Aufmerksamkeitsbezüge ("*joint attention*") zu schaffen (Rohlfing 2019). Die Herstellung gemeinsamer Aufmerksamkeit erfordert die Fähigkeit, einen dyadischen Blickkontakt zur Bezugsperson herzustellen (ab ca. 2 Monaten möglich) und auch trianguläres Blickverhalten zu zeigen (ab ca. 12 Monaten), so dass eine triadische Interaktion entsteht, in der sich zwei Beteiligte auf ein Objekt ausrichten, das im Fokus des gemeinsamen Interesses steht. In solchen Situationen geteilter Aufmerksamkeit werden typischerweise Wörter angeboten und gelernt. Für die Aufmerksamkeitskoordination und den Bezug auf

anwesende Objekte nutzen Kinder im zweiten Lebensjahr verstärkt Gesten. Insbesondere der Zeigegeste kommt eine wichtige Rolle für die sozial-kommunikative Entwicklung, aber auch eine Vorläuferfunktion für den Worterwerb zu (Liszkowski 2015). Eine ausbleibende oder verspätete Produktion von Zeigegesten erweist sich als Vorbote für Sprachentwicklungsverzögerungen (Lüke 2019).

Wörter als sprachliche Zeichen erlauben eine Ablösung vom gegenwärtigen Kontext (vom "Hier und Jetzt"), da mit ihnen auch auf abwesende, weit entfernte, vergangene, zukünftige oder nur vorgestellte Inhalte Bezug genommen werden kann. Als kognitive Vorläufer für den Wortgebrauch sind daher die ab ca. 6 Monaten entstehende Objektpermanenz, das Symbolverständnis und die Kategorisierungsfähigkeit zu nennen. Zwischen sozial-kommunikativen und kognitiven Entwicklungsschritten und dem Wortschatzerwerb bestehen wechselseitige Zusammenhänge (Weinert 2020).

### 1.2 Lexikonerwerb im Vorschulalter (ab drei Jahren)

Die bisherigen Ausführungen haben gezeigt, wie wichtig die ersten drei Lebensjahre für den Aufbau des mentalen Lexikons und den Wortgebrauch sind. Die wesentlichen Mechanismen des Wortlernens werden in der frühen Phase angelegt. In der Altersspanne ab drei Jahren bis zur Einschulung erweitern und vertiefen sich die lexikalischen Fähigkeiten. Der Umfang des Vokabulars vergrößert sich auf etwa 300 bis 500 Wörter im Alter von drei Jahren und steigt in den folgenden Jahren stetig und eher gleichmäßig an. Segbers & Schroeder (2017) schätzen den Wortschatzumfang anhand von datenbankbasierten Simulationsmodellen in der ersten Klasse auf 6000 Wörter. Für den Erwerb und die Speicherung neuer Wörter ist das phonologische Arbeitsgedächtnis maßgeblich: Bei Kindern mit guten Arbeitsgedächtnisleistungen ist der Wortschatzumfang größer als bei Kindern mit geringeren Fähigkeiten in diesem Bereich (Weinert 2020). In Bezug auf die verwendeten Wortarten ist festzustellen, dass die Komposition des Lexikons nach der anfänglich dynamischen Entwicklung im Zeitraum zwischen drei und sechs Jahren recht konstant und stabil bleibt (Kauschke et al. 2015).

Damit der Wortschatz nicht nur in die Breite wächst, sondern auch an Tiefe gewinnt, muss das anfänglich noch rudimentäre Wortwissen detaillierter und spezifischer werden. Dies erfordert vielfältige Gelegenheiten, Wörter zu hören und zu verwenden. Über den Gebrauch in unterschiedlichen Situationen hinweg werden allmählich Informationen über Wörter gesammelt und in das mentale Lexikon integriert. Diese längere Konsolidierungsphase wird als "*slow mapping*" bezeichnet und schließt sich an das anfängliche "*fast mapping*" an (He & Arunachalam 2017; Rohlfing 2019).

Konkrete Inhalte überwiegen im Lexikon weiterhin, Kinder beginnen aber auch über innere Zustände zu sprechen (Kauschke 2012), d.h. über Empfindungen,

Wünsche, Bedürfnisse, Gefühle und Gedanken. Insbesondere Emotionswörter wie 'Angst' oder 'neidisch' stoßen die Loslösung vom unmittelbar Wahrnehmbaren an, womit ihnen eine Brückenfunktion zum Erwerb von Abstrakta zukommt (Vigliocco et al. 2013). Weitere semantische Veränderungen vollziehen sich hinsichtlich der Strukturierung des Lexikons, indem Kinder lernen vielfältige semantische Relationen auszudrücken. Anfänglich werden Begriffe auf der Basisebene ("*basic level*", z.B. 'Hund') bevorzugt, im Laufe des Vorschulalters treten Oberbegriffe ('Tier') und spezifische Unterbegriffe ('Pudel') hinzu.

*Late Talkers*, die ihre frühe Verzögerung nicht überwinden konnten, laufen Gefahr eine SES auszubilden, die ab drei Jahren festgestellt werden kann. Insgesamt wird die Prävalenz von Sprachentwicklungsstörungen mit ca. 7,5% eines Jahrgangs angegeben (Norbury et al. 2016). Liegt eine SES vor, ist oft die lexikalische Ebene betroffen (Rupp 2013; van Weerdenburg et al. 2006). Die Kinder verwenden dann weniger verschiedene Wörter in der Spontansprache, d.h. ihr Wortschatz ist weniger ausdifferenziert, der Wortschatzumfang begrenzt. Hinsichtlich der Komposition des Lexikons können bestimmte Wortarten, z.B. Verben, unterrepräsentiert sein. Bei vierjährigen Kindern mit SES stellten Kauschke et al. (2010) außerdem eine unzureichende semantische Organisation fest, da kaum Begriffe ober- und unterhalb der Basisebene verwendet wurden. In Wortschatztests schneiden Kinder mit lexikalischen Störungen unterdurchschnittlich ab; sowohl der rezeptive als auch der expressive Wortschatz kann eingeschränkt sein. Da die passenden Wörter fehlen oder nicht zugänglich sind, entstehen typische Benennfehler, z.B. die Verwendung von Kohyponymen ('Pullover' statt 'Hemd'), Hyperonymen ('Vogel' statt 'Pfau'), Umschreibungen ('Gläser zum Sehen' statt 'Brille'), Wortneuschöpfungen ('Luftpuppe' statt 'Drachen') oder auch Wortformveränderungen ('Vitalator' statt 'Ventilator'; Beispiele aus Kauschke 2012). Lexikalische Störungen erfordern eine detaillierte Diagnostik, bei der zwischen verschiedenen Störungsschwerpunkten differenziert wird (siehe bspw. Glück in diesem Band), sowie eine gezielte sprachtherapeutische Intervention (siehe Abschnitt 2.2).

### 1.3 Ausblick auf das Schulalter

Da der Schwerpunkt dieses Beitrags auf dem Vorschulalter liegt, erfolgt hier ein kürzerer Ausblick auf die weiteren Veränderungen, die mit dem Schuleintritt einhergehen. Im Zuge der Lesefähigkeit und des Fachunterrichts begegnen Kinder neuartigen Wissensinhalten und damit neuen Wörtern. McGregor & Duff (2015) schätzen, dass Schulkinder etwa 1000 neue Wortstämme pro Jahr lernen. Segbers & Schroeder (2017) ermittelten einen rasanten Wortschatzzuwachs ab Beginn der zweiten Klasse über das Schulalter hinweg, wobei 10 bis 20 Wörter pro Tag gelernt werden können, bis der (rezeptive) Wortschatz von jungen Erwachsenen einen Umfang von etwa 73000 Wörtern

erreicht. Je größer das Vokabular ist, d.h. je mehr Einheiten im mentalen Lexikon gespeichert sind, umso schwieriger wird es, die jeweils passenden Wortformen unter einer Vielzahl von Wortkandidaten zu aktivieren und gezielt abzurufen. Daher steigt der Anspruch an die Verarbeitungsfähigkeiten bei der Wortproduktion. Hein & Kauschke (2020) zeigten mit Experimenten zum lexikalischen Entscheiden und Schnellbenennen, dass die produktive und rezeptive Verarbeitung von Wortformen im Laufe des Grundschulalters sicherer und schneller wird.

Hinsichtlich der Lexikonkomposition ist – mit Ausnahme eines weiteren Anstiegs des Nomenanteils – während des Schulalters wenig Bewegung zu beobachten (Segbers & Schroeder 2017).

Kennzeichnend für lexikalisch-semantiche Weiterentwicklungen im Schulalter ist der Zuwachs abstrakter Begriffe. Ponari et al. (2018) stellten fest, dass der Anteil abstrakter Wörter im kindlichen Lexikon von 10% mit vier Jahren auf über 40% mit 12 Jahren wächst, wobei im Alter von achteinhalb Jahren ein besonders steiler Anstieg zu beobachten war. Die Verarbeitung von Emotionsbegriffen wird im Laufe des Grundschulalters stabiler: Neunjährige Kinder konnten Wörter für Emotionen signifikant besser und schneller erkennen sowie besser verstehen und produzieren als sechsjährige (Kauschke et al. 2017).

Außerdem steigt die morphologische Komplexität von Wörtern. Zu einfachen Wortstämmen treten vermehrt Wörter hinzu, die durch Wortbildungsprozesse entstehen. Obwohl im dritten Lebensjahr bereits einige Kompositionen (z.B. 'Flugzeug') und Derivationen auftauchen (z.B. 'Häschen', siehe Schipke & Kauschke 2011), werden abstraktere und seltene Wortbildungsmuster und Affixe (z.B. 'transportieren') erst im Laufe des Schulalters allmählich beherrscht (Mattes 2018). Auch Segbers & Schroeder (2017) stellten einen zunehmenden Anteil morphologisch komplexer Wörter im Lexikon fest, wobei insbesondere die Komposita anstiegen.

Lexikalisch-semantiche Störungen können sich im Schulalter mit den oben beschriebenen Symptomen fortsetzen. Wortschatzfähigkeiten und schriftsprachliche Leistungen stehen dabei in einem engen wechselseitigen Zusammenhang; insbesondere korreliert die Vokabulargröße mit den Lesefähigkeiten (McGregor & Duff 2015). Fällt es Kindern besonders schwer, auf eine adäquate Wortform zuzugreifen, obwohl die Bedeutung des Wortes bekannt ist, wird von Wortfindungsstörungen gesprochen (Beier & Siegmüller 2013), die besonders im Schulalter im Zuge des sich vergrößernden Lexikons hervortreten und therapeutischer Unterstützung bedürfen. Etwa 25% der Kinder mit SES zeigen diese Art von lexikalischen Auffälligkeiten (Messer & Dockrell 2006).

## 2. Intervenieren: wann und wie?

Wortlernen basiert immer darauf, dass Wörter im Sprachangebot häufig genug auftauchen, um erkannt, verstanden und ins mentale Lexikon integriert zu werden. Insofern ist es eine allgemeine Bildungsaufgabe, durch ein qualitativ gutes Sprachangebot und verbale Wissensvermittlung allen Kindern die Voraussetzungen für einen erfolgreichen Auf- und Ausbau ihres Wortschatzes zu bieten. Bei Kindern mit lexikalischen Problemen im Rahmen von Sprachentwicklungsverzögerungen (bis 3 Jahre) oder SES (ab 3 Jahren) reicht das familiäre oder in Bildungseinrichtungen vermittelte Sprachangebot nicht aus, sie benötigen zusätzliche Unterstützung. Im Vergleich zu Kindern mit typischer Sprachentwicklung müssen sprachauffällige Kinder ein Wort doppelt so oft hören, um es zu verstehen, und doppelt so viele Produktionsgelegenheiten erhalten, um es sicher zu produzieren (Gray 2003). Je nach Alter und Phase des Lexikonerwerbs eignen sich dazu unterschiedliche Interventionsmaßnahmen.

### 2.1 Frühintervention

Die Zielgruppe von Frühinterventionsmaßnahmen sind *Late Talkers* (vgl. Abschnitt 1.1.3 und Kauschke 2017a), also zweijährige Kinder mit auffallend geringer Wortproduktion. Da einerseits etwa ein Drittel dieser Kinder bis zu einem Alter von drei Jahren noch aufholt, sich andererseits aber die meisten späteren Sprachentwicklungsstörungen durch eine *Late Talker*-Symptomatik ankündigen, wird kontrovers diskutiert, ob eine Intervention vor dem dritten Geburtstag notwendig und sinnvoll ist. Eine Entscheidungshilfe bieten Prädiktoren, d.h. Faktoren, die einen positiven oder negativen weiteren Verlauf erwarten lassen. Die Befunde dazu sind uneinheitlich (für eine aktuelle Metaanalyse siehe Fisher 2017). Gesichert ist, dass ein eingeschränktes Sprachverständnis prognostisch besonders ungünstig ist, weitere mögliche negative Faktoren sind familiäre Vorbelastungen für Sprachstörungen, ein sehr geringer früher Lexikonumfang und ein niedriges elterliches Bildungsniveau. Um der Ausbildung von späteren Sprachentwicklungsstörungen und ihren möglichen negativen Folgen präventiv zu begegnen, wurden Frühinterventionsmaßnahmen erprobt, bei denen zwei grundsätzliche Verfahrensweisen unterschieden werden (Kauschke & de Langen-Müller 2020): In kindzentrierten Ansätzen arbeitet eine Fachkraft direkt mit dem Kind, während sich elternzentrierte Ansätze an die Eltern richten, die ihrerseits das erlernte, optimierte Kommunikationsverhalten an ihre Kinder weitergeben. In beiden Formen ist das Hauptziel die Sprachentwicklung anzuregen und den Wortschatz zu erweitern.

### 2.1.1 Elternbasierte Interventionsansätze

Elternzentrierte Ansätze<sup>3</sup> gehen davon aus, dass den Eltern als wichtigsten Bezugspersonen und hauptsächlichen Interaktionspartnern eine Schlüsselrolle für die Sprachentwicklung ihrer Kinder zukommt. In Elterntrainings werden daher sprachfördernde Strategien wie das Herstellen eines gemeinsamen Aufmerksamkeitsfokus, das feinfühliges, responsive Eingehen auf die Kommunikationsangebote des Kindes oder der sinnvolle Einsatz von Modellierungs- und Stimulierungstechniken vermittelt und eingeübt.

Die Wirksamkeit von Frühinterventionsprogrammen, die auf eine gezielte Anleitung der Eltern zu sprachförderndem Verhalten ausgerichtet sind, wurde vielfach nachgewiesen (Metaanalysen: Daniels et al. 2020; Roberts & Kaiser 2011). Im deutschsprachigen Raum ist das "Heidelberger Elterntraining frühe Sprachförderung" (HET Late Talkers) verbreitet, ein systematisches und strukturiertes Interaktionstraining für Elterngruppen. Eine kontrollierte Interventionsstudie (Buschmann et al. 2009) belegte, dass sich Kinder von Eltern, die an diesem Training teilgenommen hatten, in Bezug auf Wortschatz und Satzbildung stärker verbesserten als Kinder von Eltern ohne Training. Insgesamt weist die Forschungslage darauf hin, dass Elterntrainings zumindest kurzfristig positive Effekte zeigen, zur langfristigen Wirkung sind die Befunde uneinheitlich. Darüber hinaus scheinen vor allem *Late Talkers* zu profitieren, die durch eine rein expressive Einschränkung auffallen. Liegen zusätzlich zum geringen expressiven Wortschatz auch rezeptive Defizite (im Sinne eines eingeschränkten Wort- und/oder Satzverständnisses) vor, sind die Erfolge durch elternbasierte Interventionsansätze geringer. Das Durchlaufen eines Elterntrainings reicht somit nicht bei allen Kindern aus, um die Sprachentwicklung deutlich und langfristig voranzubringen.

### 2.1.2 Kindzentrierte Interventionsansätze

Kindzentrierte Ansätze arbeiten mit spezifischen Interventionsmethoden, die im familiären Umfeld nicht uneingeschränkt umsetzbar sind. Anspruch einer kindzentrierten Frühintervention ist es, das Sprachsystem zu aktivieren und damit einen Anstoß zu einem effektiven Voranschreiten der Sprachentwicklung zu geben. Als mögliche Ziele und Gegenstände der Frühintervention nennt Kauschke (2017a):

- Aufbau grundlegender kommunikativer und kognitiver Fähigkeiten wie Blickkontakt, gemeinsame Aufmerksamkeit, Konzeptbildung, Symbolisierungsfähigkeiten
- Wahrnehmung und Realisierung des typischen Betonungsmusters von Wortformen

---

<sup>3</sup> Umfassende Überblicke finden sich in Buschmann (2015, 2020), eine tabellarische Auflistung verschiedener Formen von Elterntrainings in Kauschke & de Langen-Müller (2020: 345).



- Verbesserung von Wortverständnis und Wortproduktion zum Aufbau eines altersgerechten Vokabulars
- Wortschatzaufbau für verschiedene Wortarten, orientiert an der Dynamik der kindlichen Wortartenentwicklung (siehe Abschnitt 1.1.4)
- Einstieg in die Semantik: Aufbau von sprachlichen Kategorien (Gattungsbegriffe)
- Einstieg in die Grammatik: Erweiterung des Verblexikons und Aufbau von Wortkombinationen

Methodisch gesehen kommt insbesondere der Erhöhung der Menge, Qualität und Intensität des Sprachangebots eine besondere Rolle zu. Wortschatzerweiterung setzt voraus, dass das Kind zunächst im Input nachdrücklich mit entwicklungsadäquaten Wörtern konfrontiert wird. Damit es den Input gewinnbringend verarbeiten kann, muss dieser in besonderer Form aufbereitet werden. Für die Wortschatzarbeit in der Frühintervention bedeutet dies, dass konkrete Wörter auf Basisebene aus der kindlichen Lebenswelt hochfrequent präsentiert werden. Dazu stehen verschiedene Verfahren der Inputoptimierung zur Verfügung, z.B. das Vorspielen oder Vorlesen von Sequenzen, die ausgewählte Zielwörter in einer kontrollierten Dichte enthalten (Siegmüller & Kauschke 2016).

Auch für die Wirksamkeit der kindzentrierten Frühintervention gibt es Belege. In den fünf Studien, die in einen Review von Cable & Domsch (2011) einfließen, zeigten sich nach Durchlaufen einer individuellen oder in Gruppen durchgeführten Intervention nachweisbare sprachliche Fortschritte. Im deutschsprachigen Raum erbrachte die Studie von Schlesiger (2009) den Nachweis kurzfristiger und langfristiger Effekte einer kindzentrierten Frühintervention. Diese dauerte zwei bis sechs Monate und setzte in Abhängigkeit von der individuellen Symptomatik flexibel an den Bereichen Kommunikationsfähigkeit, Symbolspiel, Wortverständnis und Wortproduktion, Wortkombinationen sowie Satzverstehen an. *Late Talkers*, die diese Intervention erhielten, zeigten direkt und sechs Monate danach bessere lexikalische und grammatische Fähigkeiten als Kinder einer unbehandelten Kontrollgruppe und bildeten seltener eine SES aus. Positive Effekte wurden auch mit einer rein inputorientierten Interventionsform erzielt (Ringmann et al. 2010), durch die das Erreichen des Wortschatzspurtes und das Wortschatzwachstum befördert werden konnte, insbesondere bei frühem Interventionsbeginn.

## 2.2 Lexikontherapie im Vorschulalter

Halten Wortschatzeinschränkungen im Vorschulalter weiterhin an, kann dies unterschiedliche Hintergründe haben. Zum einen kann es sein, dass das Sprachangebot an ein Kind nicht ausreichend war, um den Aufbau eines altersgerecht breiten und tiefen Wortschatzes zu ermöglichen. Hierbei handelt

es sich um umgebungsbedingte Sprachauffälligkeiten (AWMF-Leitlinie 2012), die auch Fälle des mehrsprachigen Aufwachsens umfassen, in denen ein zu geringes Sprachangebot in der Umgebungssprache zu einem entsprechend kleinen oder undifferenzierten Lexikon in dieser Sprache führte. Kinder mit umgebungsbedingten Sprachauffälligkeiten benötigen sprachliche Förderung in den Bildungseinrichtungen und in ihrem täglichen Umfeld, mit dem Ziel neue Wörter und ihre Bedeutungen kennenzulernen. Fördermöglichkeiten bestehen beispielsweise mit dem Ansatz des dialogischen Lesens (Marulis & Neumann 2010). Kinder mit SES dagegen zeigen sprachliche Einschränkungen trotz eines ausreichenden Sprachangebots. Ist die lexikalische Ebene betroffen, besteht hier Bedarf an therapeutischer Intervention. Die folgenden Ausführungen beziehen sich auf spezifische kindzentrierte Therapiemaßnahmen für Kinder mit lexikalischen Störungen im Rahmen einer SES.

Kernelemente der Therapie lexikalischer Störungen sind die Absicherung von relevanten Basisfähigkeiten, die semantische und phonologische Elaboration von Wörtern und die Stabilisierung des Wortzugriffs. Diese Elemente finden sich in unterschiedlichen Therapiekonzepten wieder (für eine Übersicht siehe Glück & Elsing 2013), von denen hier exemplarisch für den deutschsprachigen Raum der "Wortschatzsammler" sowie der patholinguistische Ansatz PLAN beschrieben werden.

Das pädagogisch orientierte Konzept des "Wortschatzsammlers" (Motsch et al. 2015) sieht das aktive Auffinden bislang unbekannter Wörter vor, mit denen lexikalische Lücken entdeckt und gefüllt werden können ("Schatztruhe"). Die neuen Wörter (Nomen und Verben) werden daraufhin mit Hilfe einer Handpuppe und eingebettet in ansprechende Rahmenthemen hinsichtlich ihrer semantischen Eigenschaften besprochen. Auch das wiederholte Nachsprechen der Wortform, Abrufspiele und das Kategorisieren der Wörter sind Bestandteile der Intervention. Randomisierte und kontrollierte Studien (Motsch & Ulrich 2012a, 2012b) mit vierjährigen Kindern weisen auf Vorteile dieses Ansatzes im Vergleich zu allgemeiner Wortschatzförderung hin.

Im linguistisch orientierten PLAN (Kauschke & Siegmüller 2017; Siegmüller & Kauschke 2016) wird besonderer Wert auf eine kriteriengeleitete, kontrollierte Auswahl von Zielwörtern gelegt. Je nach Alter des Kindes und nach Therapiephase werden die Zielwörter systematisch variiert bzw. in Bezug auf linguistische Parameter gesteigert. Steigerungsmöglichkeiten bestehen in abnehmender Wortfrequenz<sup>4</sup>, steigendem Erwerbssalter<sup>5</sup>, ansteigender phonologischer oder morphologischer Komplexität, verschiedenen hierarchischen Ebenen (ausgehend von der Basisebene Ausweitung auf Hyper-

<sup>4</sup> ermittelt über die childLex-Frequenzdatenbank, siehe Schroeder et al. (2015)

<sup>5</sup> ermittelt über Befragungen / Erwerbssalter-Ratings

und Hyponyme) sowie im Einbezug von Abstrakta. Das Therapiekonzept wird individuell erstellt und an die jeweilige Symptomatik angepasst. Mögliche Komponenten sind:

- Begriffsbildung: mentale Repräsentation von Erfahrungen mit Aufmerksamkeitslenkung auf die zugehörigen Wörter (anhand realer Gegenstände und Aktivitäten)
- Objektkategorisierung zum Aufbau von Gattungsbegriffen (unterschiedliche Vertreter einer Objektkategorie zusammenfassen)
- Auslösung des *fast mapping*-Prozesses (mit Pseudowörtern, neue Wortformen auf unbekannte Objekte beziehen)
- Erweiterung und Festigung des rezeptiven und expressiven Wortschatzes für ausgewählte und kontrollierte Zielwörter, die aus unterschiedlichen semantischen Feldern (Kleidung, Fahrzeuge, Lebensmittel, Tiere) und Wortarten (Nomen, Verben, Adjektive) stammen
- Merkmalsarbeit zur semantischen Elaboration
- Verbesserung der Strukturierung und Vernetzung des Lexikons
- Elaboration und Durchdringung der Wortform (Identifizieren, Merken, Analysieren und Synthetisieren von Wortformen)
- Stabilisierung und Automatisierung des Wortabrufs (schnelles Benennen und freie Produktion).

Im PLAN wird eine Kombination expliziter und impliziter Methoden eingesetzt. Zunächst werden neue Zielwörter hochfrequent und gezielt in Inputsequenzen präsentiert. In rezeptiven Übungen wird am Wortverständnis (z.B. Wort-Bild-Zuordnung), an der Vernetzung (Klassifikationsübungen) und an der Wortformverarbeitung (z.B. lexikalisches Entscheiden) gearbeitet. Expressive Übungen (Benennen) bieten Produktionsgelegenheiten unter verschiedenen Bedingungen und in verschiedenen Kontexten. Auch metasprachliche Methoden können eingesetzt werden, um semantische Merkmale zu thematisieren, Taxonomien bewusst zu machen oder die Auseinandersetzung mit Wortformen zu fördern.

Eine Metaanalyse zu Wortschatzinterventionen im Vorschulalter (Marulis & Neuman 2010) kommt zu dem Schluss, dass insbesondere mit einer Kombination von expliziten und impliziten Vermittlungsformen positive Effekte erzielt werden können. Vogt & Kauschke (2017a, 2017b) zeigten, dass Vorschulkinder mit Sprachentwicklungsstörungen darüber hinaus beim Wortlernen vom Einsatz ikonischer Gesten profitieren: Neue Wörter, die mit einer passenden bildhaften Geste verbunden wurden, wurden besser verstanden, produziert und definiert als Wörter, die nur mit einer allgemein hinweisenden Geste eingeführt wurden.

### 2.3 Ausblick: Wortschatzförderung im Schulalter

Setzen sich lexikalische Einschränkungen im Schulalter weiter fort, ist auch eine Fortführung von Interventionsmaßnahmen notwendig, die bei umgebungsbedingten Auffälligkeiten als Fördermaßnahme und bei SES als therapeutische Intervention angeboten wird. Therapiemaßnahmen sind ähnlich aufgebaut wie zuvor für das Vorschulalter beschrieben. Allerdings werden die Zielwörter dem steigenden Alter, den Veränderungen in der Lebenswelt und dem wachsenden schulischen Fachwissen angepasst, d.h. sie werden länger, abstrakter und morphologisch komplexer. Komposita, Derivationen und auch nicht native Wortstämme (Fremdwörter) werden einbezogen. Im Rahmen der semantischen Elaboration wird auf eine vertiefte Bedeutungsanreicherung und auf die Herstellung vielfältiger Vernetzungen zwischen Wörtern abgezielt. Mit phonologischer Elaboration wird die differenzierte und bewusste Auseinandersetzung mit Wortformen gefördert. Auch an Wörter gekoppelte grammatische Informationen (z.B. Wortart, morphologische Wortstruktur) oder orthografische Merkmale der geschriebenen Wortform können thematisiert werden.<sup>6</sup>

Ein stärkeres Gewicht kommt im Schulalter außerdem der Vermittlung von Strategien zu (Glück & Elsing 2014), mit deren Hilfe von Wortschatzeinschränkungen und Wortfindungsstörungen betroffene Kinder größere sprachliche Handlungsfähigkeit gewinnen können. Aktiv und selbstständig eingesetzte Strategien sollen zu mehr Selbstmanagement, Eigeninitiative und Eigenverantwortung befähigen. Mögliche Strategien sind:

- Worterwerbsstrategien: eigenständiges Suchen von Informationen über Wortbedeutungen, Wörter memorieren und visualisieren, Gedächtnisanker setzen
- Abrufstrategien, z.B. Anlauthilfen
- Kompensationsstrategien bei Wortschatzlücken, z.B. informative Umschreibung, Synonyme, metasprachliche Kommentare.

Eine Überblicksarbeit zur Effektivität von Wortschatzinterventionen für sprachentwicklungsgestörte Jugendliche zwischen 11 und 17 Jahren (Lowe et al. 2018) führte zu einem gemischten Bild, das vor allem für die Wirksamkeit einer kombinierten semantisch-phonologischen Wortschatzarbeit spricht, die unabhängig vom Setting (individuell oder in der Schulklasse) positive Effekte erbrachte. Aus den 13 Studien, die in die Überblicksarbeit von Lowe und Kolleg\_innen eingingen, sollen abschließend zwei Einzelstudien zur Effektivität von Wortschatzintervention herausgegriffen werden. Das unter 2.2 beschriebene Konzept "Wortschatzsammler" wurde auch für die Anwendung im schulischen Kontext ausgearbeitet. Motsch & Marks (2015) untersuchten

<sup>6</sup> Eine tabellarische Zusammenstellung von Elaborationsstrategien für das Grundschulalter findet sich in Kauschke (2019: 11).

dessen Wirksamkeit an 157 Drittklässlern in deutschen Sprachheilschulen. Kinder, die mit dem Wortschatzsammler-Konzept gefördert wurden, zeigten nach 20 Sitzungen in Bezug auf den Wortschatz und auf das Satzverständnis bessere Testleistungen als eine Kontrollgruppe, die nur die reguläre schulische Förderung erhielt. Wright et al. (2018) führten eine Lexikontherapie mit 9- bis 16-jährigen Kindern und Jugendlichen mit SES durch. Abgestimmt auf das Alter und die lexikalischen Fähigkeiten wurden Zielwortsets ausgewählt, die auch Wörter mit niedriger Frequenz und höherem Abstraktheitsgrad (z.B. 'zögern', 'Weisheit') beinhalteten. In den Interventionssitzungen wurden vielfältige Methoden eingesetzt, darunter Einführung und häufiges Angebot der Zielwörter, Besprechen der phonologischen und semantischen Eigenschaften, metasprachliche Bestimmung der Wortart, Nachsprechen, Übungen zur Wortproduktion, Lesen, Bilder zu den Wörtern zeichnen, Satzproduktion mit den Zielwörtern. Nach einer siebenwöchigen Interventionsphase (30 Minuten pro Woche) war der Umgang mit geübten Wörtern signifikant besser als mit ungeübten, wobei Nomen besser gelernt wurden als Verben. Die Teilnehmenden verbesserten sich in Bezug auf lexikalisches Entscheiden, Definieren und Wortproduktion im Satz.

Aus der derzeitigen Studienlage zu wortschatzbezogener Intervention im Schulalter lässt sich folgern, dass durch gezielte Lexikontherapie mit einem relativ geringen Zeitaufwand kurzfristig nachweisbare Effekte erzielt werden können, allerdings wurden mittel- oder langfristige Effekte bislang nicht ausreichend untersucht.

### 3. Fazit

Wortschatzaufbau ist ein lebenslanger und mehrdimensionaler Prozess, der im vorliegenden Beitrag mit dem Fokus auf dem Kindesalter betrachtet wurde. Im Zuge dieses Prozesses lernen Kinder Wortformen zu identifizieren und mit Bedeutung zu verknüpfen, konkrete und abstrakte Wörter mit Informationen anzureichern, ein Wortartenrepertoire aufzubauen, Wortbildungsmuster anzuwenden und im jeweiligen Situationskontext die passenden Wortformen abzurufen. Dabei brauchen manche Kinder besondere Unterstützung, da entweder das Sprachangebot ihrer Umgebung nicht ausreicht oder die internen Lern- und Verarbeitungsmechanismen eingeschränkt sind, die einen effektiven Aufbau des mentalen Lexikons und/oder einen mühelosen Wortabruf ermöglichen. Um diesen Problemen und möglichen negativen Folgen für den weiteren Bildungs- und Entwicklungsweg entgegenzuwirken, eignen sich gezielte Interventionsmaßnahmen, die auf die Breite und die Tiefe des Lexikons ausgerichtet sind. Übereinstimmend werden wesentliche Komponenten Erfolg versprechender kindzentrierter Intervention beschrieben (z.B. Glück & Elsing 2013; McGregor & Duff 2015; Steele & Mills 2011):

- Gezielte Auswahl, häufiges Angebot und systematische Steigerung von Zielwörtern mit Kontrolle linguistischer Variablen wie Frequenz, Wortart, Erwerbssalter, Konkretheit, phonologische und morphologische Komplexität
- Erarbeitung und Elaboration der phonologischen und semantischen Eigenschaften von Wörtern
- Methodenvielfalt unter Einsatz direkter Methoden
- Schaffung von häufigen und vielfältigen Produktionsgelegenheiten.

Werden diese Komponenten berücksichtigt, erscheinen Rahmenbedingungen wie das Setting (individuell versus Kleingruppen), die Interventionsdauer und die Gruppengröße für den Therapieerfolg eher unerheblich zu sein (Lowe et al. 2018). Der beste Ansatz ist somit "a comprehensive approach that provides children with multiple, rich opportunities to learn words, to learn how to learn words, and to learn the power of words" (McGregor & Duff 2015: 271).

## LITERATUR

- Ambridge, B., Pine, J.M., Rowland, C.F., Chang, F. & Bidgood, A. (2013). The retreat from overgeneralization in child language acquisition: Word learning, morphology, and verb argument structure. *WIREs Cognitive Science*, 4, 47-62.
- Armstrong, R., Scott J.G., Whitehouse, A.J.O., Copland D.A., McMahon K.L. & Arnott, W. (2017). Late talkers and later language outcomes: Predicting the different language trajectories. *International Journal of Speech-Language Pathology*, 19(3), 237-250.
- AWMF-Leitlinie zur Diagnostik von Sprachentwicklungsstörungen (SES) unter Berücksichtigung umschriebener Sprachentwicklungsstörungen (USES), Registernummer 049-006. <http://www.awmf.org/leitlinien/detail/II/049-006.html>
- Beier, J. & Siegmüller, J. (2013). Kindliche Wortfindungsstörungen. In S. Ringmann & J. Siegmüller (Hgg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen – Schuleingangsphase* (S. 79-99). München: Elsevier.
- Buschmann, A. (2015). Gezielte Anleitung von Bezugspersonen zur sprachförderlichen Alltagsinteraktion. In S. Sachse (Hg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen - Kleinkindphase* (1. Aufl.) (S. 185-205). München: Elsevier.
- Buschmann, A. (2020). Einbezug der Eltern in die Sprachförderung. In S. Sachse, A. K. Bockmann & A. Buschmann (Hgg.), *Lehrbuch Sprachentwicklung. Entwicklung – Diagnostik – Förderung im Kleinkind- und Vorschulalter* (S. 283-308). Berlin: Springer.
- Buschmann, A., Jooss, B., Rupp, A., Feldhusen, F., Pietz, J. & Philippi, H. (2009). Parent based language intervention for 2-year-old children with specific expressive language delay: A randomised controlled trial. *Archives of Disease in Childhood*, 94(2), 110-116.
- Cable, A. L. & Domsch, C. (2011). Systematic review of the literature on the treatment of children with late language emergence. *International Journal of Language & Communication Disorders*, 46(2), 138-154.
- Daniels, R., Fanselow, A. & Egert, F. (2020). Zur Wirksamkeit sprachfördernder Frühinterventionen bei Late Talkers: Eine Metaanalyse. *Sprache Stimme Gehör*. DOI: 10.1055/a-1060-6181.
- Dietrich, R. & Gerwien, J. (2017). *Psycholinguistik. Eine Einführung*. Stuttgart: Metzler.

- Fisher, E. L. (2017). A systematic review and meta-analysis of predictors of expressive-language outcomes among late talkers. *Journal of Speech, Language, and Hearing Research*, 60(10), 2935-2948.
- Fox-Boyer, A. & Schäfer, B. (2015). Die phonetisch-phonologische Entwicklung von Kleinkindern. In S. Sachse (Hg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen – Kleinkindphase* (S. 39-62). München: Elsevier.
- Frank, M., Braginsky, M., Yurovsky, D. & Marchman, V. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, 44(3), 677-694.
- Friedrich, M., Weber, C. & Friederici, A. D. (2004). Electrophysiological evidence for delayed mismatch response in infants at-risk for specific language impairment. *Psychophysiology*, 41, 772-782.
- Glück, C. W. (2021). Der WWT 6-10 und seine Testitems. *Bulletin suisse de linguistique appliquée*, 113, 29-47.
- Glück, C. W. & Elsing, C. (2013). Lexikontherapie. In A. Fox-Boyer (Hg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen – Kindergartenphase* (S. 173-186). München: Elsevier.
- Glück, C. W. & Elsing, C. (2014). Gestörte Lexikonentwicklung. In A. Fox-Boyer (Hg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen – Kindergartenphase* (S. 73-85). München: Elsevier.
- Gray, S. (2003). Word-learning by pre-schoolers with specific language impairment: What predicts success? *Journal of Speech, Language, and Hearing Research*, 46, 56-67.
- He, A. X. & Arunachalam, S. (2017). Word learning mechanisms. *Wiley Interdisciplinary Reviews: Cognitive Science* 2017, 8(4), e1435.
- Hein, K. & Kauschke, C. (2020). Word form processing in primary school children: A psycholinguistic perspective. *Journal of Speech, Language, and Hearing Research*, 63(11), 3685-3699.
- Höhle, B. (2015). Crosslinguistic perspectives on segmentation and categorization in early language acquisition. In E. Bavin & L. Naigles (Hgg.), *The Cambridge handbook of child language* (S. 159-182). Cambridge: University Press.
- Höhle, B., Pauen, S., Hesse, V. & Weissenborn, J. (2014). Discrimination of rhythmic pattern at 4 months and language performance at 5 years: A longitudinal analysis of data from German-learning children. *Language Learning*, 64 (Suppl. 2), 141-164.
- Kauschke, C. (2007). *Erwerb und Verarbeitung von Nomen und Verben*. Tübingen: Niemeyer.
- Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen. Verläufe, Forschungsmethoden, Erklärungsansätze*. Berlin: De Gruyter.
- Kauschke, C. (2017a). Frühe Sprachauffälligkeiten. In J. Siegmüller & H. Bartels (Hgg.), *Leitfaden Sprache Sprechen Stimme Schlucken* (5. Aufl.) (S. 62-72). München: Elsevier.
- Kauschke, C. (2017b). Psycho- und Patholinguistik 1: Normaler und gestörter Spracherwerb im Kindesalter. In M. Grohnfeldt (Hg.), *Kompendium der akademischen Sprachtherapie und Logopädie, Band 2: Interdisziplinäre Grundlagen* (S. 111-131). Stuttgart: Kohlhammer.
- Kauschke, C. (2019). Wortschatzerwerb bei Kindern – Erkenntnisse zur Entwicklung im Grundschulalter und davor. *Grundschule Deutsch*, 62(2), 8-11.
- Kauschke, C., Bahn, D., Vesker, M. & Schwarzer, G. (2017). Die semantische Repräsentation von Emotionsbegriffen bei Kindern im Grundschulalter. *Kindheit und Entwicklung*, 26(4), 251-260.
- Kauschke C., Bartl-Pokorny, K. D., Marschik, P. B. & Vollmann, R. (2015). Wortschatz in Erzählungen – eine Untersuchung lexikalischer Aspekte narrativer Fähigkeiten bei Vorschulkindern mit und ohne Sprachauffälligkeiten. *Sprache Stimme Gehör*, 39(3), 149-155.
- Kauschke, C. & de Langen-Müller, U. (2020). Sprachtherapie mit Kindern. In S. Sachse, A. K. Bockmann & A. Buschmann (Hgg.), *Lehrbuch Sprachentwicklung. Entwicklung – Diagnostik – Förderung im Kleinkind- und Vorschulalter* (S. 331-357). Berlin: Springer.

- Kauschke, C., Fauck, A. & Nachbarschulte, A. (2010). Zur hierarchischen Organisation des mentalen Lexikons bei Kindern mit spezifischer Sprachentwicklungsstörung. *Sprache Stimme Gehör*, 34, 228-236.
- Kauschke, C. & Hofmeister, C. (2002). Early lexical development in German: A study on vocabulary growth and vocabulary composition during the second and third year of life. *Journal of Child Language*, 29, 735-757.
- Kauschke, C. & Siegmüller, J. (2017). Der patholinguistische Ansatz in der Therapie von Sprachentwicklungsstörungen im Überblick. *Logos*, 25(4), 264-275.
- Kühn, P., Sachse, St. & v. Suchodoletz, W. (2016). Sprachentwicklung bei Late Talkern. *Logos*, 24(4), 256-264.
- Lizkowski, U. (2015). Kommunikative und sozial-kognitive Voraussetzungen des Spracherwerbs. In S. Sachse (Hg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen – Kleinkindphase* (S. 27-38). München: Elsevier.
- Lowe, H., Henry, L., Müller, L. M. & Joffe, V. L. (2018). Vocabulary intervention for adolescents with language disorder: A systematic review. *International Journal of Language & Communication Disorders*, 53(2), 199-217.
- Lüke, C. (2019). Verspätete gestische Kommunikation als Vorbote von Sprachentwicklungsverzögerungen. *Sprache Stimme Gehör*, 43(2), 77-85.
- Lüke, C., Gremplewski, K. & Ritterfeld, U. (2017). Analyse der lexikalischen Entwicklungen im zweiten Lebensjahr zur Identifikation von Sprachentwicklungsstörungen. *Sprache Stimme Gehör*, 41(1), 44-51.
- Marulis, L. M. & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of educational research*, 80(3), 300-335.
- Mattes, V. (2018). *Derivationsmorphologie und Wortarten im Erstspracherwerb des Deutschen. Mit einem Fokus auf dem Schuleintrittsalter*. Habilitationsschrift, Universität Graz 2018.
- McGregor, K. K. & Duff, D. (2015). Promoting diverse and deep vocabulary development. In T. A. Ukrainetz (Hg.), *School-age language intervention: Evidence-based practices* (S. 247-277). Austin, Texas: PRO-ED Inc. 8700 Shoal Creek Boulevard.
- Messer, D., & Dockrell, J. E. (2006). Children's naming and word-finding difficulties: Descriptions and explanations. *Journal of Speech, Language and Hearing Research*, 49(2), 309-324.
- Motsch, H. J. & Marks, D. K. (2015). Efficacy of the Lexicon Pirate strategy therapy for improving lexical learning in school-age children: A randomized controlled trial. *Child Language Teaching and Therapy*, 31(2), 237-255.
- Motsch, H. J., Marks, D. K. & Ulrich, T. (2015). *Wortschatzsammler. Evidenzbasierte Strategitherapie lexikalischer Störungen im Kindesalter*. München: Ernst Reinhardt Verlag.
- Motsch, H. J. & Ulrich, T. (2012a). Effects of the strategy therapy "lexicon pirate" on lexical deficits in preschool age: A randomized controlled trial. *Child Language Teaching and Therapy*, 28(2), 159-175.
- Motsch, H. J. & Ulrich, T. (2012b). "Wortschatzsammler" und "Wortschatzfinder" - Effektivität neuer Therapieformate bei lexikalischen Störungen im Vorschulalter. *Die Sprachheilarbeit*, 57(2), 70-78.
- Norbury, C.F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G. & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry*, 57(11), 1247-1257.
- Ponari, M., Norbury, C. F. & Vigliocco, G. (2018). Acquisition of abstract concepts is influenced by emotional valence. *Developmental science*, 21(2), e12549.



- Rescorla, L. (2015). Entwicklungswege von Late-Talkern. In S. Sachse (Hg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen - Kleinkindphase* (S. 101-119). München: Urban & Fischer.
- Ringmann, S., Dähn, S., Neumann, C., Lehnhoff, A., Rohdenburg, W., Schröders, C. & Siegmüller, J. (2010). Frühe inputorientierte Lexikontherapie - Ein Vergleich zwei- und dreijähriger Kinder im Late-Talker-Stadium. *LOGOS interdisziplinär*, 18(5), 358-369.
- Roberts, M. Y. & Kaiser, A. P. (2011). The effectiveness of parent-implemented language interventions: A meta-analysis. *American Journal of Speech-Language Pathology*, 20(3), 180-199.
- Rohlfing, K. J. (2019). *Frühe Sprachentwicklung*. Tübingen: Narr Francke.
- Rupp, S. (2013). *Semantisch-lexikalische Störungen bei Kindern*. Berlin, Heidelberg: Springer.
- Sachse, S. (2015). Vorhersage der weiteren Entwicklung von Late-Talkern mit besonderer Berücksichtigung von deutschsprachigen Studien. In S. Sachse (Hg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen - Kleinkindphase* (S. 119-128). München: Urban & Fischer.
- Samuelson, L. K. & McMurray, B. (2017). What does it take to learn a word? *Wiley Interdisciplinary Reviews: Cognitive Science*, 8, e1421.
- Schipke, C. & Kauschke, C. (2011). Early word formation in German language acquisition. *First Language*, 31(1), 67-82.
- Schlesiger, C. (2009). *Sprachtherapeutische Frühintervention für Late Talkers: Eine randomisierte und kontrollierte Studie zur Effektivität eines direkten und kindzentrierten Konzeptes* (1. Aufl.). Idstein: Schulz-Kirchner.
- Schröder, C. & Höhle, B. (2011). Prosodische Wahrnehmung im frühen Spracherwerb. *Sprache Stimme Gehör*, 35(3), e91-e98.
- Schroeder, S., Würzner, K. M., Heister, J., Geyken, A. & Kliegl, R. (2015). childLex: A lexical database of German read by children. *Behavior research methods*, 47(4), 1085-1094.
- Segbers, J. & Schroeder, S. (2017). How many words do children know? A corpus-based estimation of children's total vocabulary size. *Language Testing*, 34(3), 297-320.
- Siegmüller, J. & Kauschke, C. (2016). *Materialien zur Therapie nach dem Patholinguistischen Ansatz – Lexikon und Semantik*. München: Elsevier.
- Steele, S. C. & Mills, M. T. (2011). Vocabulary intervention for school-age children with language impairment: A review of evidence and good practice. *Child Language Teaching and Therapy*, 27(3), 354-370.
- Till, C. (2021). Semantisch-lexikalische Diagnostik bei sukzessiv-bilingualen Kindern. *Bulletin suisse de linguistique appliquée*, 113, 175-191.
- van Weerdeburg, M., Verhoeven, L. & van Balkom, H. (2006). Towards a typology of specific language impairment. *Journal of Child Psychology and Psychiatry*, 47(2), 176-189.
- Vigliocco, G., Kousta, S. T., Della Rosa, P. A., Vinson, D. P., Tettamanti, M., Devlin, J. T. & Cappa, S. F. (2013). The neural representation of abstract words: The role of emotion. *Cerebral Cortex*, 24(7), 1767-1777.
- Vogt, S. & Kauschke, C. (2017a). Observing iconic gestures enhances word learning in typically developing children and children with specific language impairment. *Journal of child language*, 44(6), 1458-1484.
- Vogt, S. & Kauschke, C. (2017b). With some help from others' hands: Iconic gesture helps semantic learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 60(11), 3213-3225.
- Weber, C., Hahne, A., Friedrich, M. & Friederici, A.D. (2005). Reduced stress pattern discrimination in 5-month-olds as a marker of risk for later language impairment: Neurophysiological evidence. *Cognitive Brain Research*, 25, 180-187.

- Weinert, S. (2020). Sprachentwicklung im Kontext anderer Entwicklungsbereiche. In S. Sachse, A. K. Bockmann & A. Buschmann (Hgg.), *Lehrbuch Sprachentwicklung. Entwicklung – Diagnostik – Förderung im Kleinkind- und Vorschulalter* (S. 131-162). Berlin: Springer.
- Wright, L., Pring, T. & Ebbels, S. (2018). Effectiveness of vocabulary intervention for older children with (developmental) language disorder. *International journal of language & communication disorders*, 53(3), 480-494.

# Semantisch-lexikalische Diagnostik bei sukzessiv-bilingualen Kindern

## Christoph TILL

Pädagogische Hochschule Bern  
Institut für Heilpädagogik  
Fabrikstrasse 8, CH-3012 Bern  
christoph.till@phbern.ch

The acquisition of a second language in childhood is affected by several factors, including the age of onset, quantity and quality of language input, context of language use, and many more. The impact of these variables is observed in the size and quality of the language-specific lexicons. Therefore, successive bilingual children may show similar vocabulary deficits in their second language as children with lexical developmental deficits. It is necessary to differentiate between these two groups of children, because the latter group needs customized word finding therapy in both their languages, while for the former group, additional education in the second language might be sufficient. Various procedures for diagnosing word finding disorders were developed to fit the needs of bilingual children. In general, these developments are an improvement to standard procedures, but they still fail to accurately identify word finding disorders in bilingual children. However, assessing vocabulary capabilities is a worthwhile endeavor. Strengths and weaknesses can be identified and incorporated into additional language education.

### Keywords:

successive bilingualism, vocabulary deficits, diagnostic procedures, diagnostic accuracy, dynamic assessment.

### Stichwörter:

Sukzessiver Zweitspracherwerb, semantisch-lexikalische Störung, diagnostische Methoden, diagnostische Genauigkeit, Dynamic Assessment.

## 1. Einleitung

Semantisch-lexikalische Störungen (SLS) werden als "erhebliche, nicht altersgemäße, häufige und anhaltende Schwierigkeiten, die eigenen Äußerungsintentionen in angemessen lexikalisch besetzten Äußerungen auszudrücken bzw. Äußerungen anderer wörtlich zu verstehen" (Glück 2011: 6) definiert. Die Ursachen werden in eingeschränkten Erwerbsmechanismen, Speicherkapazitäten und/oder Abruf Routinen vermutet. Als Erwerbsstörung äußern sich SLS als quantitatives Wortschatzdefizit und Wörter können weder expressiv noch rezeptiv verarbeitet werden. Als Speicherstörung machen sich SLS ggf. als Wortfindungsstörungen bemerkbar, wenn Informationen zur Wortbedeutung, zur Wortform und/oder zur semantischen Organisation nicht vollständig abgespeichert sind, weswegen es zu Ungenauigkeiten und Verwechslungen v.a. in der Wortproduktion kommt. Als Wortabrufstörungen äußern sich SLS dadurch, dass Wortformen nicht in angemessener Zeit abgerufen werden können, obwohl die Informationen erworben wurden und abgespeichert sind (Glück & Elsing 2014). Kinder mit SLS haben häufig alle der



beschriebenen Schwierigkeiten, prägen in der Regel aber ein individuelles Störungsprofil aus. Dieses muss differentialdiagnostisch erfasst werden, damit die im Einzelfall am besten passenden Unterstützungsmöglichkeiten ergriffen werden können (Ulrich 2017).

Darüber hinaus wirken sich SLS auf weitere Lern- und Entwicklungsbereiche aus. So wird der Lese-Rechtschreiberwerb stark von den semantisch-lexikalischen Fähigkeiten eines Kindes beeinflusst (siehe Beitrag von Winkes 2021). Auch ist der schulische Lernerfolg durch persistierende SLS gefährdet. Wie Glück & Spreer (2015) aufzeigen konnten, entsprechen die schulischen Leistungen der Kinder mit einem besonderen Förderbedarf im Bereich Sprache, darunter solche mit SLS, zu 75% dem Niveau der 25% leistungsschwächsten Kinder ohne diesen Förderbedarf. 40% der betroffenen Kinder haben sogar einen Leistungsdurchschnitt, der den Leistungen der schwächsten 5% der Kinder ohne Förderbedarf entspricht.

Sukzessiv-bilinguale Kinder, welche als solche definiert werden, die mit drei Jahren oder später eine Zweitsprache (L2) erwerben (Rothweiler & Ruberg 2011: 11), können ebenso SLS aufweisen. Auch bei mehrsprachigen Kindern äußern sich SLS wie oben beschrieben (Bedore & Peña 2008; Peña & Bedore 2009). Doch selbst zweitspracherwerbende Kinder, die keine genuinen Wortschatzstörungen haben, können an der Oberfläche ähnliche Schwierigkeiten in der Wortproduktion, im Wortverständnis, beim Wortabruf und damit verbunden auch vergleichbare Folgeprobleme demonstrieren wie Kinder mit SLS (Sheng et al. 2006; Ehl et al. 2014; Wilkens et al. 2018). Dieser Umstand ist auf die besonderen Bedingungen zurückzuführen, die im mehrsprachigen Erwerb vorherrschen (Golberg et al. 2008). Besonders großen Einfluss auf die Wortschatzentwicklung üben die Inputquantität und -qualität in den verschiedenen Sprachen aus (Thordardottir 2011; Hoff & Core 2013; Prevoo et al. 2014).

Während ungünstige Erwerbsbedingungen durch systematische Sprachförderung ausgeglichen werden können, können Kinder mit einer SLS nicht in ausreichendem Maß von pädagogischen Sprachfördermaßnahmen profitieren, sondern benötigen einen individuellen sprachtherapeutischen/ logopädischen Zugang (Scharff Rethfeldt 2018). In der Diagnostik ist folglich die Differenzierung einer genuinen SLS von ungünstigen Erwerbsbedingungen von zentraler Bedeutung.

Wie Rothweiler (2001: 123) deutlich macht, ist das Unterfangen, Wortschatzleistungen angemessen zu erheben, mit vielen Schwierigkeiten verbunden. Dies gilt erst recht für die Diagnostik von Kindern im Zweitspracherwerb, da sich der Zweitspracherwerb variabler gestaltet als der monolinguale Spracherwerb.

Im Folgenden sollen zwei Fragen untersucht werden:

1. Therapieindikation: Kann bei zweitspracherwerbenden Kindern das Vorliegen einer SLS zuverlässig abgeklärt werden und wenn ja, wie?
2. Differentialdiagnostik: Kann bei zweitspracherwerbenden Kindern das Störungsprofil zuverlässig bestimmt werden?

Um diese Fragen zu beantworten, wird zunächst ein Modell der semantisch-lexikalischen Diagnostik vorgestellt, welches sich für den Einsatz bei einsprachigen Kindern gut eignet.

## 2. Modell der semantisch-lexikalischen Diagnostik

Ulrich (2017) hat ein Modell der semantisch-lexikalischen Diagnostik entworfen, das ein effizientes und ökonomisches Vorgehen ermöglicht (s. Abb. 1). Nachdem ein Kind durch semantisch-lexikalische Fehlleistungen in der Spontansprache aufgefallen ist, sollte als nächstes ermittelt werden, ob eine SLS vorliegt (Therapieindikation). Diese äußert sich in der Regel durch unterdurchschnittliche Leistungen in einem aktiven Wortschatztest.

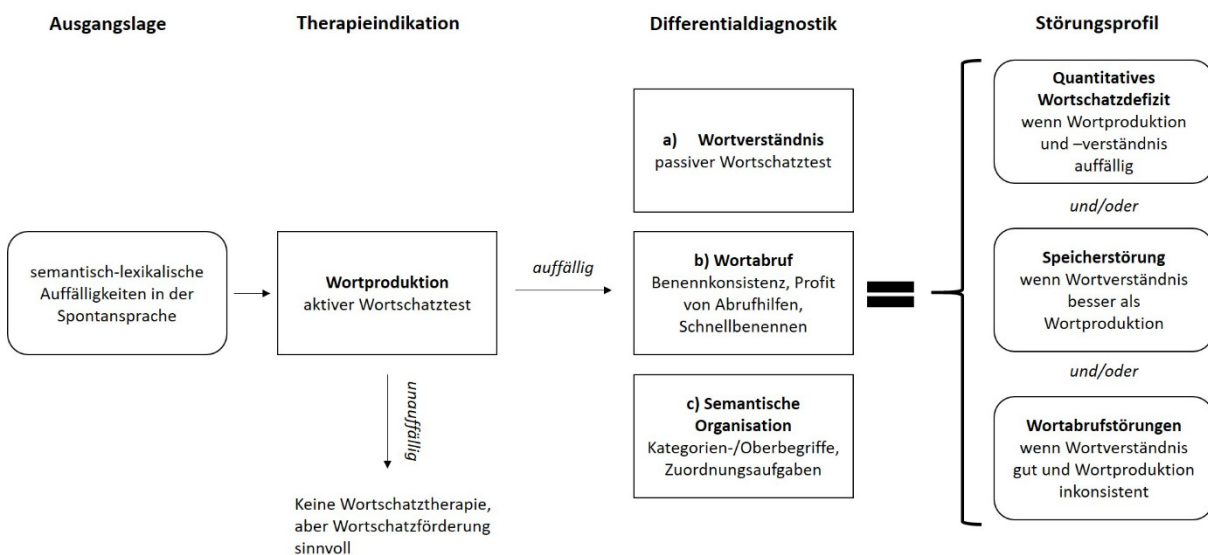


Abbildung 1: Modell der semantisch-lexikalischen Diagnostik (in Anlehnung an Ulrich 2017)

Sind die Leistungen im aktiven Wortschatztest unauffällig, kann die Überprüfung an dieser Stelle abgebrochen werden. Wird der Verdacht auf SLS jedoch erhärtet, sollte in einem nächsten Schritt das Störungsprofil ermittelt werden (Differentialdiagnostik). Ulrich (2017) empfiehlt in diesem Sinne, als nächstes den passiven Wortschatz zu überprüfen. Aus einer Auswahl von Bildern und/oder Objekten, die außer dem Zielwort semantische, phonologische oder nicht-relationierte Ablenker darstellen, muss das Kind ein vorgedachtes Wort wiedererkennen und auf das Bild/Objekt des entsprechenden Wortes zeigen. Sind sowohl die Leistungen im aktiven als auch im passiven Wortschatz unterdurchschnittlich, verweist dies auf ein generelles, quantitatives

Wortschatzdefizit. Sind die passiven Wortschatzleistungen deutlich besser als die aktiven, verweist dies auf ein eher qualitativ ausgeprägtes Wortschatzdefizit (Speicherstörung oder Wortabrufstörung). Das vorhandene Wortwissen ist ungenügend elaboriert, sodass zum Zielwort semantisch und/oder phonologisch ähnliche Begriffe produziert werden. Im Wortverständnis werden qualitative lexikalische Defizite auf der Wortbedeutungs- bzw. der Wortformebene durch die Wahl der semantischen oder eben phonologischen Ablenker ersichtlich.

Die Diskrepanz zwischen guten Wortverständnis- und schlechten Wortproduktionsleistungen kann auch durch reine Wortabrufprobleme erklärt werden. In diesem Fall werden häufig kommentierte Nullantworten gegeben ("Das müsste ich doch wissen" oder Ähnliches). Um die Abrufleistungen zu erheben, können die Benennkonsistenz, der Profit von Abrufhilfen und die Benennzeiten überprüft werden. Wenn sich noch keine Abrufrouninen etabliert haben, kann es sein, dass es dem Kind zu einem Zeitpunkt gelingt, ein Wort abzurufen, zu einem anderen Zeitpunkt jedoch nicht. Dies kann durch die zweimalige Durchführung ein und desselben aktiven Wortschatztests aufgedeckt werden. Während manche Begriffe in keiner Durchführung und manche in beiden Durchführungen benannt werden können, gibt es auch solche, die nur in einer der beiden Durchführungen, also inkonsistent benannt werden können. Beträgt der Anteil der inkonsistenten Benennungen mehr als 10%, darf von einer Wortabrufstörung ausgegangen werden (Ulrich 2017).

Beim Profit von Abrufhilfen wird ermittelt, welche Hinweise ein Kind benötigt, um ein Wort schließlich erfolgreich abrufen zu können. Beim *Wortschatz- und Wortfindungstest für 6- bis 10-Jährige* (WWT 6-10; Glück 2011; siehe auch Beitrag von Glück 2021) gibt es motivierende Hinweise ("Denk noch einmal in Ruhe darüber nach!"), die dem Kind mehr Zeit für den Wortabruf gewähren. Darauf folgt der semantische Hinweis, der zusätzliche Informationen über die Wortbedeutung bietet. Der dritte, phonologische Hinweis besteht in der Anlautvorgabe und stellt einen konkreten Verweis auf die gesuchte Wortform dar. Wenn keine der Abrufhilfen wirksam wird, kann man davon ausgehen, dass ein entsprechender Lexikoneintrag nicht existiert.

Zur Überprüfung des Schnellbenennens (*rapid automatized naming* oder RAN; Winkes 2014: 58f) werden in der Regel einige wenige hochfrequente Wörter (Objekte des Alltags, Zahlen, Farben u.Ä.) wiederholt dargestellt, die das Kind nacheinander so schnell wie möglich benennen soll. Gelingt der Abruf dieser Wörter in keiner angemessenen Geschwindigkeit, verweist dies auf gestörte bzw. nicht etablierte Abrufrouninen. Das Lexikon könnte dementsprechend einen altersgemäßen Umfang haben, der Zugriff auf die Lexikoneinträge ist aber erschwert.

Zusätzlich empfiehlt Ulrich (2017) das Überprüfen der semantischen Organisation. Bei unvollständigem oder nicht vernetztem Wortwissen können

Gemeinsamkeiten und Unterschiede verschiedener Begriffe nicht zur Organisation und zum Ausbau des mentalen Lexikons genutzt werden, was sich u.a. in der fehlenden Verwendungssicherheit von Kategorien bzw. Oberbegriffen zeigt. Dies kann einerseits durch das Benennen von Kategorien geprüft werden, indem zu mehreren Abbildungen ein passender Oberbegriff genannt werden muss. Andererseits kann das Vorliegen von Kategorien auch über Zuordnungs- bzw. Sortieraufgaben überprüft werden. Zu einer verbal und/oder visuell vorgegebenen Kategorie müssen passende Bilder zugeordnet und unpassende Bilder aussortiert werden.

Es stellt sich nun die Frage, ob das beschriebene Vorgehen auch bei Kindern im L2-Erwerb zielführend ist. Vor dem Hintergrund von Forschungsergebnissen zum Worterwerb sukzessiv-bilingualer Kinder sollen die eingangs gestellten Fragen nun diskutiert werden.

### **3. Forschungsergebnisse zu Methoden der Wortschatzdiagnostik im Kontext Mehrsprachigkeit**

#### *3.1 Ermitteln der Therapieindikation*

Wie in Abschnitt 2 erläutert, ist bei einem einsprachigen Kind eine SLS gegeben, wenn die Leistung in einem aktiven Wortschatztest unterdurchschnittlich ausfällt. Wenn man einen aktiven Wortschatztest in der L1 *oder* der L2 eines mehrsprachigen Kindes durchführt, fallen die entsprechenden Leistungen in der Regel geringer aus als bei einem einsprachigen Kind. Das liegt v.a. daran, dass ein mehrsprachiges Kind sein Wortwissen auf mehrere Sprachen verteilen muss (Bialystok et al. 2010; Hoff et al. 2012; Peña et al. 2016; Anaya et al. 2018). Ein Aufholen zu den produktiven Wortschatzleistungen monolingualer Sprecher derselben Sprachen gilt als nahezu unmöglich (Golberg et al. 2008; Thordardottir 2011; Hoff et al. 2012; Hoff & Core 2013). Im Gegensatz zum Erwerb der Grammatikregeln, der gesamthaft geschieht, gilt der Wortschatzerwerb nie als abgeschlossen. So bauen auch monolinguale Menschen ihren Wortschatz permanent aus, während zweitspracherwerbende Personen an diese Leistungen anzuknüpfen versuchen, um weiterhin kommunikativ teilhaben zu können (Golberg et al. 2008). Aufgrund dieses sog. "vocabulary gap" (Thordardottir 2011) werden die semantisch-lexikalischen Leistungen zweitspracherwerbender Kinder häufig als auffällig beurteilt, während alle anderen sprachlichen Entwicklungsbereiche unauffällig sein können (Lüke 2011; Prevoo et al. 2014; Bonifacci et al. 2020).

Wegen dieser offenkundigen Diskrepanzen der produktiven Wortschatzleistungen in den Einzelsprachen mehrsprachiger Kinder gegenüber einsprachigen Kindern wird generell empfohlen, das Wortwissen in L1 *und* L2 zu erheben und kombiniert zu bewerten (Hansen et al. 2019; Bonifacci et al. 2020). Eine Möglichkeit besteht in der Überprüfung des sog. *konzeptuellen* Wortschatzes (Peña et al. 2016). Dies kann in Anlehnung an das *revised*

*hierarchical model of lexical and conceptual representation in bilingual memory* (s. Abb. 2) von Kroll & Stewart (1994) geschehen. Hierzu ist es dem Kind erlaubt, bei der Durchführung eines Benenntests in der L1 *oder* in der L2 zu antworten. Kann das Kind ein Bild oder ein Objekt in einer Sprache nicht benennen, darf es dies in der anderen tun. Gelingt dies, weiß man, dass das Kind das entsprechende Konzept erworben hat, die zugehörige Wortform aber nur in einer seiner Sprachen kennt. Gelingt die Benennung gar nicht, kann man davon ausgehen, dass das entsprechende Konzept (noch) nicht erworben worden oder nicht abrufbar ist.

Anaya et al. (2018) werfen jedoch ein, dass der Wechsel zwischen den Sprachen innerhalb der Testsituation die Leistungen des Kindes zusätzlich schwächt. So muss der Wortabruf in einer der beiden Sprachen aktiv unterdrückt werden, was v.a. dann kognitive Ressourcen beansprucht, wenn es sich dabei um die dominante Sprache des Kindes handelt. Sie empfehlen daher, zunächst eine Sprache zu überprüfen, und dann - an einem anderen Tag - die andere Sprache (möglich z.B. mit dem WWT 6-10, der auf Deutsch und Türkisch durchgeführt werden kann).

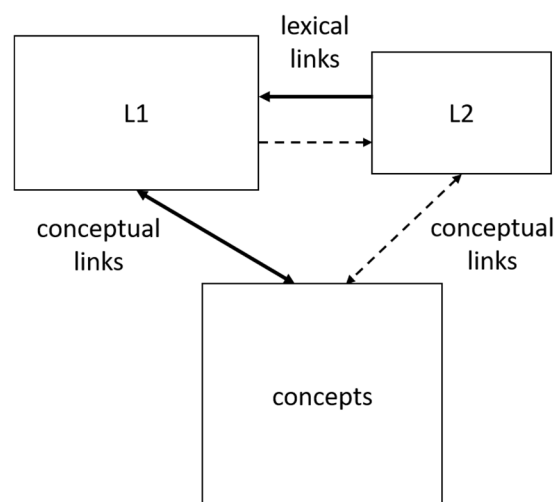


Abbildung 2: "revised hierarchical model of lexical and conceptual representation in bilingual memory" von Kroll & Stewart (1994: 158)

Aber auch auf diese Weise werden die aktiven Wortschatzleistungen nur unvollständig erfasst, da die Kinder für einen Teil der Konzepte Wortformen aus zwei Sprachen zur Verfügung haben - gemäß Peña et al. (2002) trifft dies auf rund 28% der erworbenen Konzepte zu, wobei dies vor dem Hintergrund der Erwerbskontexte betrachtet werden muss, die individuell stark variieren können. Hoff et al. (2012), Hoff & Core (2013) und Peña et al. (2016) empfehlen daher die Überprüfung des *totalen* Wortschatzes. Wiederum werden beide Sprachen überprüft, diesmal werden alle Wortformen gezählt, die in L1 *und/oder* L2 genannt wurden. Dieses Vorgehen würdige die Wortschatzleistungen eines



mehrsprachigen Kindes eher und führe zu einem angemesseneren Vergleich mit monolingualen Kindern. Anaya et al. (2018) entgegnen jedoch, dass dieses Vorgehen zu einer Überschätzung der Wortschatzleistungen führt. So haben mehrsprachige Kinder zwar für einen Teil der Konzepte Wörter aus zwei Sprachen erworben, aber andere Wortschatzfelder können durchaus unterrepräsentiert sein. Der Einwand von Ayala et al. (2018) bezüglich des Sprachenwechsels innerhalb der Testsituation muss natürlich auch hier geltend gemacht werden.

Auch unter Berücksichtigung dieser alternativen Vorgehensweisen wird in der Fachliteratur zur Vorsicht bei der Überprüfung des aktiven Wortschatzes geraten. Aufgrund ungenügender Werte für Sensitivität (korrekte Erfassung einer vorliegenden Störung) und Spezifität (korrekte Ermittlung der Abwesenheit einer Störung) (je unter 80%) sind falsch positive wie falsch negative Diagnosen bei allen beschriebenen Vorgehensweisen nach wie vor sehr wahrscheinlich. Die Überprüfung des aktiven Wortschatzes sollte somit auch bei angepasster Vorgehensweise nicht zur Ermittlung einer Therapieindikation genutzt werden (Bedore & Peña 2008; Kapantzoglou et al. 2012; Anaya et al. 2018; Wilkens et al. 2018).

### 3.2 *Ermitteln des Störungsprofils*

Auch bei der Überprüfung des passiven Wortschatzes fallen die Wortverständnisleistungen sukzessiv-bilingualer Kinder in den Einzelsprachen geringer aus als bei monolingualen Kindern. Bezüglich der dominanten Sprache bewegt sich die Abweichung (gerade noch) im Normbereich (Golberg et al. 2008; Bialystok et al. 2010), in der schwachen Sprache fällt die Diskrepanz deutlicher aus. Bei der Interpretation der Ergebnisse eines passiven Wortschatztests in der L2 sollten neben dem Alter des Kindes, der Erstsprache, der Begabung, dem Bildungsniveau der Mutter (vgl. Paradis 2016) unbedingt die Inputquantität und die Erwerbsdauer berücksichtigt werden. Aus ihrer Untersuchung mit simultan-bilingualen Kindern konnte Thordardottir (2011) schließen, dass ein L2-Input von mindestens 40-60% nötig ist, um vergleichbare Wortverständnisleistungen wie ein monolinguales Kind zu erreichen. Diese Zahl dürfte bei sukzessiv-bilingualen Kindern noch größer sein. Außerdem benötigen zweitspracherwerbende Kinder gemäß Golberg et al. (2008) drei Jahre Spracherfahrung mit der L2, um in den rezeptiven Leistungen zu ihren monolingualen Peers aufzuschließen. Wenn man diese Umstände nicht berücksichtigt, führt die Überprüfung des Wortverständnisses in der L2 zu vielen falsch positiven Urteilen (Peña et al. 2016; Wilkens et al. 2018; Bonifacci et al. 2020). Stattdessen sollte das Wortverständnis v.a. in der L1 untersucht werden, da andere klassische Vorgehensweisen nur bedingt aussagekräftig sind (Bonifacci et al. 2020). Die Ermittlung eines Therapiebedarfs ist also auch mit der Wortverständnisüberprüfung nicht sicher möglich. Allerdings lässt sich differentialdiagnostisch ermitteln, ob ein quantitatives Defizit vorliegt, nämlich

wenn aktive und passive Wortschatzleistungen beide auffällig sind, oder ob ein eher qualitatives Defizit vorliegt, wenn die passiven Leistungen deutlich besser ausfallen als die aktiven.

Wie der Hinweis von Anaya et al. (2018) bezüglich des Sprachenwechsels in der Testsituation bereits erahnen lässt, ist auch die Untersuchung der Wortabrufleistungen im mehrsprachigen Kontext erschwert. Groot (2011) identifiziert zwei Faktoren, die den Wortabruf mehrsprachiger Kinder maßgeblich beeinflussen, die Sprachdominanz und der Sprachkontext. Wird der Wortabruf in der L2 überprüft, die L1 ist aber die dominante Sprache, kann es zu Verzögerungen im Wortabruf, evtl. sogar zu unwillkürlichen Sprachwechseln kommen. Ist die L2 hingegen die dominante Sprache, gelingt der Abruf schneller und Sprachwechsel sind unwahrscheinlich. Wenn sich das Kind unmittelbar vor der Wortabrufuntersuchung in der L2 in einem L1-Kontext befunden hat, ist das L1-Lexikon vorrangig aktiviert und der Abruf in der L2 entsprechend erschwert. Befand sich das Kind vor der Überprüfung aber schon in einem L2-Kontext, ist auch das L2-Lexikon bereits aktiviert, wodurch die Abrufleistungen in der L2 gesteigert werden.

Für die Überprüfung der Benennkonsistenz bedeutet das, dass sich die Abrufbedingungen bei der zweiten Durchführung des aktiven Wortschatztests günstiger gestalten, da die untersuchte Sprache in der ersten Durchführung schon aktiviert wurde. Je nach Ausprägung von Sprachdominanz und -kontext wird auch das Schnellbenennen positiv oder negativ beeinflusst (Kohnert & Bates 2002). Es lässt sich leicht vorstellen, dass auch der Profit von Abrufhilfen von diesen Faktoren beeinflusst wird: Ein Kind, dessen dominante Sprache Spanisch ist und das einen Affen benennen soll, könnte durch die Anlauthilfe "Es fängt mit 'A...' an" zusätzlich verwirrt werden, da es tendenziell eher die Wortform *mono* sucht. Um einen realistischen Eindruck der Abrufleistungen zu erhalten, sollte man diese idealerweise in der dominanten Sprache überprüfen. Wenn dies nicht möglich ist, sollten die Sprachdominanz und der Sprachkontext bei der Interpretation der Ergebnisse unbedingt berücksichtigt werden.

Auch die Leistungen der semantischen Organisation werden durch einen mehrsprachigen Erwerbskontext beeinflusst. So kennen bilinguale Kinder in der Regel zwar die gleichen Kategorien wie einsprachige, die Kategorien können aber unterschiedlich besetzt sein. Peña et al. (2002) befragten englisch-spanischsprachige Kinder in beiden Sprachen nach Lebensmitteln und Tieren. Während in der Kategorie Lebensmittel die drei häufigsten Antworten (in dieser Reihenfolge) auf Englisch *apple*, *banana* und *hamburger* lauteten, waren die drei häufigsten spanischen Nennungen *hamburguesa* (Hamburger), *manzana* (Apfel) und *sopa* (Suppe). In der Kategorie Tiere lauteten die häufigsten Antworten auf Englisch *elephant*, *lion* und *dog*, auf Spanisch *caballo* (Pferd), *elefante* (Elefant) und *tigre* (Tiger).

Wenn zu Hause die L1 und in der Schule die L2 gesprochen wird, dürfte sich die Organisation der sprachspezifischen Lexika ebenfalls unterscheiden. So kommen Begriffe aus dem Familienalltag wohl eher in der L1 und Begriffe aus dem Schulalltag eher in der L2 vor (Hoff & Core 2013). Diese Annahme konnten Bialystok et al. (2010) nur teilweise bestätigen. So konnten die von ihnen untersuchten Kinder Wörter aus dem Kontext Schule in ihren beiden Sprachen etwa gleich gut verstehen, Wörter aus dem Familienalltag wurden in der L1 aber deutlich besser verstanden als in der L2.

In der Besetzung von Kategorien und in der Verwendungshäufigkeit kategorieller Begriffe können also sprachspezifische Unterschiede bestehen. Bezüglich der Benennung von Oberbegriffen heißt das, dass eventuell kein Oberbegriff für eine bestimmte Gruppe von Gegenständen genannt werden kann, weil diese Gegenstände aus Sicht des mehrsprachigen Kindes keine typischen Vertreter dieser Kategorie sind. Im Rahmen von nonverbalen Sortieraufgaben kann ermittelt werden, ob ein Kind über verschiedene Kategorien verfügt und ob diese angemessen besetzt sind. Ist dies der Fall, kann man davon ausgehen, dass das Kind über ein System zur Organisation des mentalen Lexikons verfügt. Auf der anderen Seite könnten aufgrund der oben beschriebenen Unterschiede zum monolingualen Erwerb von Kategorien mehrsprachige Kinder möglicherweise andere Zuordnungen vornehmen, als man es von monolingualen Kindern kennt.

Die Gegenüberstellung der verschiedenen Testresultate zum aktiven und passiven Wortschatz, zur Abrufqualität und der semantischen Organisation erlaubt es, ein semantisch-lexikalisches Fähigkeitsprofil zu ermitteln, auch wenn diverse Besonderheiten bei der Interpretation der Ergebnisse berücksichtigt werden müssen. So sollte der aktive Wortschatz eines mehrsprachigen Kindes in allen seinen Sprachen, als konzeptueller oder als totaler Wortschatz erhoben werden, um einen besseren Eindruck der produktiven Wortschatzleistungen zu erhalten. Der passive Wortschatz sollte hingegen in der L1 untersucht werden, da hier die besten Leistungen zu erwarten sind. Bei der Untersuchung der Abrufleistungen muss der Sprachkontext der Testsituation bewusst gestaltet werden, um die möglichen Einflüsse korrekt interpretieren zu können. Im Bereich der semantischen Organisation ist v.a. die Durchführung von nonverbalen Sortieraufgaben aussichtsreich, die einen qualitativen Einblick in die Kategorisierungsleistungen der Kinder ermöglicht. Da die Sprachdominanz und der Sprachkontext die Leistung von sukzessiv-bilingualen Kindern in den verschiedenen Tests beeinflussen, kann trotz eines semantisch-lexikalischen Fähigkeitsprofils keine Therapieindikation festgestellt werden. Im Gegensatz zu den hier diskutierten statischen Testverfahren könnten dynamische Vorgehensweisen (*Dynamic Assessment*) gute Alternativen darstellen.

#### 4. Dynamische Wortschatzdiagnostik

In den letzten drei Jahrzehnten sind Vorgehensweisen, die als *Dynamic Assessment* (DA) bezeichnet werden, in den Fokus der Sprachentwicklungsdiagnostik gerückt (Orellana et al. 2019). DA hat den Anspruch, nicht das Sprachwissen, sondern das Sprachlernpotential zu messen. Das Sprachwissen ist nämlich – wie oben erläutert – von den Bedingungen, unter denen es erworben wird, abhängig. Gestalten sich die Bedingungen eher ungünstig, wird das Sprachwissen eine Diskrepanz zum Wissensstand derjenigen Kinder aufzeigen, die unter besseren Bedingungen lernen (Peña et al. 2007). Untersucht man jedoch das Sprachlernpotential, wird offensichtlich, was ein Kind leisten könnte, wenn es bessere Bedingungen hätte.

Um herauszufinden, ob und unter welchen Bedingungen sich die Testleistungen verbessern, werden dem Kind während des Testens verschiedene Hilfestellungen angeboten. Verbessern sich die Leistungen durch die Hilfestellungen unmittelbar, ist das Lernpotential groß und eine intensive Sprachtherapie ist nicht nötig. Verbessern sich die Leistungen jedoch nicht oder nur in geringem Ausmaß, bedeutet das, dass das Sprachlernpotential beeinträchtigt ist und längerfristig intensivere Maßnahmen notwendig sind, um angemessene Fortschritte erreichen zu können (Orellana et al. 2019). Es gibt verschiedene Umsetzungsmöglichkeiten des DA. Mit Bezug auf die Diagnostik semantisch-lexikalischer Fähigkeiten soll jeweils ein Beispiel aus der Forschungsliteratur aufgezeigt werden.

In einem *test-teach-retest*-Design haben Kapantzoglou et al. (2012) eine Gruppe spanischsprachiger Kinder mit Englisch als Zweitsprache, die sich ca. zur Hälfte aus Kindern mit SLS zusammensetzte, untersucht. Die Umsetzungen der *test-teach-retest*-Prozedur erfolgte auf Spanisch und dauerte zwischen 30 und 40 Minuten. In der *test*-Phase wurden alle Kinder dazu aufgefordert, drei Bilder von hochfrequenten Objekten (Blume, Pizza, Sonnenbrille) und drei Bilder von niederfrequenten Objekten (ein nicht identifizierbares Tier, Samenkörner, Wasserwaage) zu benennen. Die niederfrequenten Objekte konnten von den Kindern erwartungsgemäß nicht benannt werden und wurden von der Testleitung mit Pseudowörtern bezeichnet (*fote, depa, kina*). In der anschließenden *teach*-Phase wurden den Kindern die Kategorie, die Funktion und die Eigenschaften der Objekte erläutert. Außerdem wurden die Kinder dreimal dazu aufgefordert, den Namen des Objekts zu wiederholen. Erst wenn diese drei Imitationen korrekt erfolgten, wurde die *retest*-Phase eingeleitet. Wiederum wurden die sechs Bilder mit den hoch- und niederfrequenten Objekten vorgelegt, die die Kinder benennen sollten. Im Vergleich der *test*- und *retest*-Ergebnisse offenbarte sich, ob und wie viel die Kinder ihre Leistungen unter gleichen *teach*-Bedingungen steigern konnten. Mit diesem Vorgehen konnten 78% der Kinder korrekt identifiziert werden, was dem geforderten Minimum für Spezifität und Sensitivität (80%) nur knapp nicht genügt. Die

Autorinnen zeigen sich jedoch zuversichtlich, dass mit gezielten Überarbeitungen des geschilderten Vorgehens bessere Werte erreicht werden können.

Beim *graduated prompting* werden dem Kind während des Tests zunehmend stärker werdende Hinweise zur Lösung der Aufgabe angeboten, woraufhin die Anzahl benötigter Hilfestellungen ausgewertet wird. Ein solches Vorgehen haben Camilleri & Botting (2013) zur Wortlernendiagnostik umgesetzt. Sie legten 15 Kindern (ein- und mehrsprachig, mit und ohne SLS) Bilder vor, die Alltagsszenen darstellten. Es wurde für jedes Bild ein Zielwort definiert, das auf jedem Bild zweimal abgebildet war. Die Kinder wurden im ersten Schritt dazu aufgefordert, zu beschreiben, was sie alles auf dem Bild sehen. Wenn das Kind von sich aus das Zielwort gebraucht hat, wurde die volle Punktzahl gegeben. War dies nicht der Fall, wurden nach und nach stärker werdende Hinweise auf das Zielwort gegeben: zunächst längere, semantisch unspezifische Hinweise, dann immer spezifischere und kürzer werdende Hinweise. Zuletzt wurde auf eine Darstellung des Zielobjekts gezeigt, der Begriff genannt und nach der zweiten Darstellung des Begriffs gefragt. Je weniger Hinweise ein Kind benötigte, um das Zielwort korrekt zu identifizieren, desto mehr Punkte bekam es. Braucht ein Kind viele Hinweise, weist dies darauf hin, dass das Kind über ungeeignete Wortlernstrategien verfügt und sehr klare, strukturierte Anweisungen benötigt, um Wortformen und ihre Referenten zu erwerben. Camilleri & Botting (2013) decken hohe Korrelationen zwischen den Testergebnissen und den Spracherwerbsbedingungen auf, verweisen aber gleichzeitig darauf, dass die Testergebnisse nicht den Anforderungen an Sensitivität und Spezifität genügen. Weiterhin lässt sich kritisch anmerken, dass das beschriebene Vorgehen nach wie vor Sprachwissen (und eben nicht Sprachlernpotential) belohnt, da die sofortige Benennung des Zielworts die höchste Punktzahl bringt. Das Vorgehen kann den diagnostischen Prozess jedoch unterstützen und nützliche Informationen zu den Wortlernkapazitäten eines Kindes liefern.

Als dritte Alternative des DA soll noch der Einsatz von sog. *modifiability scales* erwähnt werden. Hierbei handelt es sich um Beobachtungsbögen, die das Verhalten von Kindern in Lernsituationen beschreiben. Peña et al. (2007) haben die *Mediated Learning Observation* (MLO) entwickelt, die Beobachtungen zu Affekten, zur kognitiven Aktivierung, zur kognitiven Elaboration und zum sozio-emotionalen Verhalten auf einer fünfstufigen Skala abbildet. Auf der Skala zur kognitiven Elaboration beispielsweise findet sich das Item *Problemlösen*. Von *Stufe 1 = planlos, unsystematisch, raten* bis *Stufe 5 = systematisch und effizient, vorausschauend, reflektiert* kann das Arbeitsverhalten des Kindes differenziert eingeschätzt werden. Die Auswertung der Daten von 40 Kindern mit Englisch als Zweitsprache (davon 15 mit Sprachentwicklungsstörungen), die bei der Bearbeitung einer Wortlernaufgabe beobachtet wurden, hat ergeben, dass die Kinder mit 93% Genauigkeit korrekt zugeordnet wurden.

Die hier exemplarisch ausgewählten Studienergebnisse zu verschiedenen Vorgehensweisen des DA spiegeln die Ergebnisse der Metaanalyse von Orellana et al. (2019) wider. In allen untersuchten Studien zum Einsatz der *modifiability scales* konnten sprachentwicklungsauffällige und -unauffällige Kinder gut voneinander differenziert werden. *Test-teach-retest*-Verfahren hingegen lieferten nur in einer von sechs Studien ähnlich zuverlässige Nachweise.

Insgesamt sind die Spezifitätswerte der DA-Verfahren höher als die Sensitivitätswerte. Das steht im Gegensatz zu einer Metaanalyse statischer Testverfahren, wo sich das Verhältnis der Werte umkehrt (Dollaghan & Horner 2011). Daraus ziehen Orellana et al. (2019) den Schluss, dass weder statische noch dynamische Diagnostikverfahren allein zuverlässige Aussagen ermöglichen, sondern dass die aussagekräftigsten Ergebnisse aus der Kombination beider Vorgehensweisen resultieren. Insofern stellt DA keine Alternative, sondern eine wertvolle Ergänzung zu üblichen Vorgehensweisen in der Wortschatzdiagnostik dar. So können mit den in Abschnitt 3 beschriebenen Vorgehensweisen zumindest qualitative Aussagen zu den Wortschatzleistungen sukzessiv-bilingualer Kinder gewonnen werden, die dann durch Erkenntnisse zum Sprachlernpotential ergänzt werden.

## 5. Diskussion

Die geschilderten Zusammenhänge lassen nachvollziehbar erkennen, warum die Identifikation einer vorliegenden SLS und damit die Ermittlung der Therapieindikation bei Kindern im Zweitspracherwerb erschwert ist. Für einsprachige Kinder konzipierte diagnostische Vorgehensweisen zur Identifikation einer SLS an sich und zur Identifikation des Störungsschwerpunkts können nicht ohne Weiteres auf mehrsprachige Kinder übertragen werden. Die Bedingungen, unter denen Kinder eine zweite Sprache erwerben, unterscheiden sich so stark voneinander, dass klassische, standardisierte Vorgehensweisen in der Diagnostik zu vielen falschen Diagnosen führen können (Kapantzoglou et al. 2012).

Wie Tabelle 1 zeigt, reicht für eine Therapieindikation die Untersuchung des L2-Wortschatzes nicht aus. Die L1 muss unbedingt miteinbezogen werden (Hansen et al. 2019; Bonifacci et al. 2020). Die Untersuchung des konzeptuellen bzw. totalen Wortschatzes führt zu Ergebnissen, die vergleichbar mit monolingualen Leistungen sind. Dennoch sind Leistungsdiskrepanzen feststellbar, die zu Fehlinterpretationen führen können, weswegen Peña et al. (2016) und Thordardottir (2015) empfehlen, bei der Interpretation der Wortschatzleistungen die Sprachdominanz zu berücksichtigen.

Aktiver Wortschatz	Überprüfung der L1 und L2 als konzeptueller oder totaler Wortschatztest
Passiver Wortschatz	Vorzugsweise Überprüfung der L1, L2 nur unter Berücksichtigung von Inputquantität und Erwerbsdauer (z.B. gemäß PABIQ; Tuller 2015)
Wortabruf	Vorzugsweise Überprüfung der L1, L2 nur unter Berücksichtigung von Sprachdominanz und -kontext (z.B. gemäß PABIQ; Tuller 2015)
Semantische Organisation	In beiden Sprachen möglich, aber nur unter Berücksichtigung sprachspezifischer Unterschiede

Tabelle 1: Zusammenfassung der Empfehlungen zur bilingualen Wortschatzdiagnostik

Welches die dominante Sprache ist, lässt sich z.B. mithilfe des speziell für die Eltern mehrsprachiger Kinder entwickelten "Questionnaire for Parents of Bilingual Children" (PABIQ; Tuller 2015) ermitteln. Das PABIQ bietet die Möglichkeit, zu erfragen, wie häufig, in welchem Kontext und mit wem welche Sprachen gesprochen werden, was Aufschluss über die Sprachdominanz des Kindes gibt. Da sich die Erwerbsbedingungen für die Sprachen des Kindes aber kontinuierlich ändern, sollten diese Erhebungen in regelmäßigen Abständen wiederholt werden. So kann sich z.B. das Dominanzverhältnis von L1 und L2 nach ca. sechs bis sieben Jahren Sprachkontakt zugunsten der Zweitsprache ändern, da diese als Umgebungs- bzw. Mehrheitssprache allgegenwärtig ist (Kohnert & Bates 2002; Hoff et al. 2012; Hoff & Core 2013; Sheng 2014; Dijkstra et al. 2016).

Um die L1 der Kinder in der Diagnostik berücksichtigen zu können, braucht es entsprechende Tests, die für viele Erstsprachen nicht existieren. Selbst wenn es solche L1-Wortschatztests gibt, sind diese nicht per se geeignet. Der Wortgebrauch in den Erstsprachen entwickelt sich in Migrationssettings häufig anders als in den Herkunftssettings (Hamann & Abed Ibrahim 2017) und entsprechende Wortschatztests sind folglich unpassend konzipiert. Tatsächlich müssten passende Wortschatztests mithilfe der Zielgruppe (sukzessiv-bilinguale Kinder) entwickelt und normiert werden. Dabei gilt jedoch, dass alle relevanten Einflussfaktoren (Erstsprache, Inputquantität und -qualität, Erwerbsdauer u.v.m.) berücksichtigt werden müssen, was in angemessener Art und Weise kaum zu leisten ist, wie auch Gagarina (2013) bei der Untersuchung des Sprachstandstests Russisch für mehrsprachige Kinder (SRUK) feststellen musste.

Auch Übersetzungen von L2-Wortschatztests in die L1 der Kinder sind mit Problemen behaftet. Xuan & Dollaghan (2013) haben die Wortschatzzusammensetzung von Kindern, die Englisch und Mandarin erwerben, untersucht. Während Englisch eher als nomenlastig gilt, zeigt Mandarin eine Dominanz der Verben. Entsprechend setzte sich der Englisch-Wortschatz zu 54% aus Nomen zusammen, der Mandarin-Wortschatz nur zu 38%. Würde man einen (eher nomenlastigen) englischen Wortschatztest in Mandarin übersetzen,

würde man die Wortartenverteilung übernehmen, was zum Nachteil in der Bewertung der Mandarin-Fähigkeiten führen würde. Auch die Untersuchung von Peña et al. (2002) hat aufgezeigt, dass die verschiedensprachigen Wortformen desselben Konzepts unterschiedlich häufig gebraucht werden: Während *hot dogs* im englischsprachigen Wortgebrauch hochfrequent ist, trifft das für das spanischsprachige Pendant (*perrito caliente*) nicht gleichermaßen zu. Dies führt letzten Endes dazu, dass die Items in den Testübersetzungen unterschiedliche Schwierigkeiten haben (Hansen et al. 2019).

Im Projekt COST Action IS0804 versucht man, dieses Problem zu umgehen. Aus einem Pool von insgesamt 158 Nomen und 141 Verben werden anhand der Kriterien Erwerbsalter und linguistische Komplexität des Wortes Testwörter zusammengestellt, die zu sprachspezifischen, vergleichbar schwierigen Wortschatztests führen sollen (Cross-linguistical Lexical Tasks, kurz: CLT; Haman et al. 2015). Der Pool umfasst Wörter aus 34 Sprachen, bislang wurden 24 Sprachversionen der CLT entwickelt. Noch konnte nicht gezeigt werden, dass die CLT einen auffälligen von einem unauffälligen Worterwerb zuverlässig differenzieren kann. Ein Problem besteht in den nach wie vor unterschiedlichen Schwierigkeitsindizes der verschiedenen Sprachversionen (Haman et al. 2017).

Das Ziel der logopädisch-therapeutisch orientierten Diagnostik, eine vorliegende Störung zu identifizieren (vgl. Scharff Rethfeldt 2018), ist nach aktuellem Forschungsstand mit Wortschatzdiagnostik allein nicht zufriedenstellend zu erreichen.

Orientiert man sich am Vorgehen von Ulrich (2017), lassen sich semantisch-lexikalische Fähigkeiten eines Kindes jedoch umfassend beschreiben: Gibt es bedeutende Diskrepanzen zwischen den aktiven und passiven Wortschatzleistungen oder fallen diese ähnlich aus? Wenn ein Kind Fehler beim aktiven Wortgebrauch oder beim Wortverstehen macht, sind diese eher semantisch oder phonologisch orientiert? Über welche semantischen Kategorien verfügt ein Kind, welche sind nicht oder schwach besetzt? Gelingt es dem Kind, den ihm zur Verfügung stehenden Wortschatz gezielt und angemessen schnell abzurufen oder muss es jeweils sehr lange überlegen?

Im Sinne einer pädagogisch-diagnostischen Zielsetzung (vgl. Scharff Rethfeldt 2018) können diese Informationen gut zur Planung einer passenden Wortschatzförderung genutzt werden. Abbildung 3 zeigt auf, aus welchen Leistungsprofilen welche Förderziele abgeleitet werden können, damit das Kind im Wortschatzaufbau und beim Erwerb des Lesens und Schreibens sowie beim schulischen Lernen optimal unterstützt wird. Zu diesem Zweck kann auch der Einsatz von für die Zweitsprache der Kinder entwickelten Testverfahren gerechtfertigt werden (Smolander et al. 2020). Auch die verschiedenen Methoden des Dynamic Assessment können in diesem Kontext einen wertvollen Beitrag leisten. So zeigen diese nicht nur auf, wie groß das



Lernpotential der Kinder ist, sondern auch, welche Hilfestellungen besonders wirksam sind, damit Fortschritte erreicht werden können.

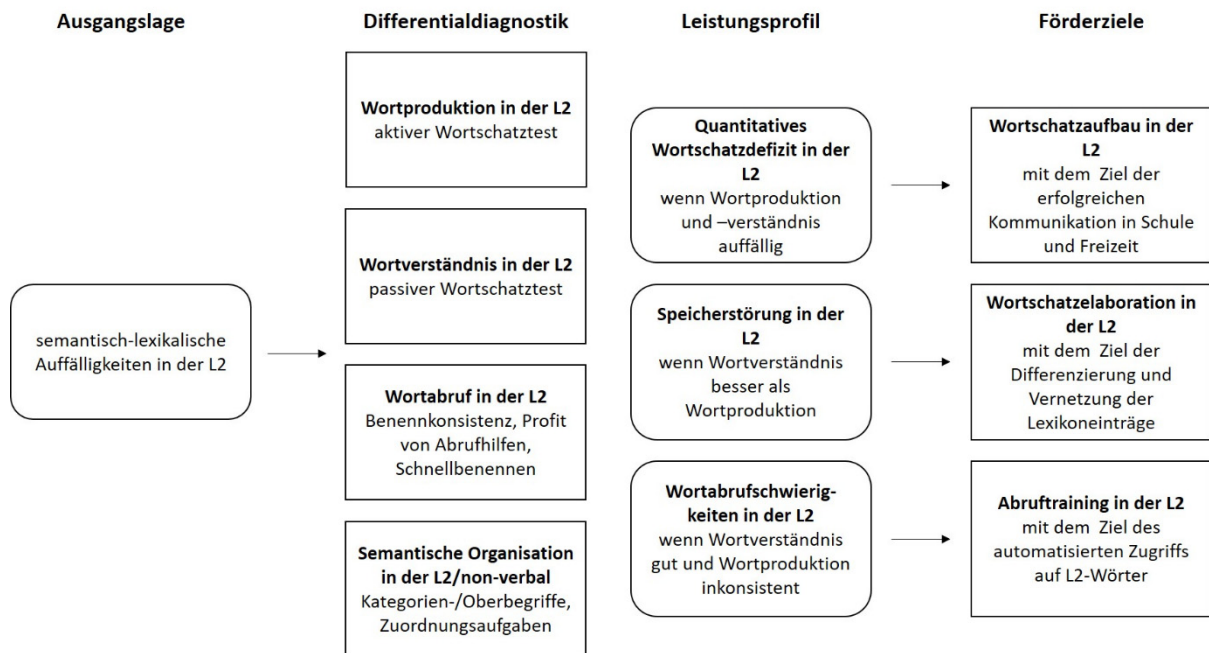


Abbildung 3: Pädagogisch-orientierte L2-Wortschatzdiagnostik in Anlehnung an Ulrich (2017)

Obwohl semantisch-lexikalische Leistungen diagnostisch nur schwer zu erfassen sind, lohnt es sich trotzdem, diese gezielt und umsichtig zu erheben, um im Sinne der pädagogischen Diagnostik ein optimales Förderangebot ableiten zu können. Nur so kann Problemen im Erwerb des Lesens und Schreibens sowie bezüglich der schulischen Leistungen präventiv entgegengetreten werden. Dies gilt für alle Kinder gleichermaßen, ob ein- oder mehrsprachig.

## LITERATUR

- Anaya, J. B., Peña, E. D. & Bedore, L. M. (2018). Conceptual scoring and classification accuracy of vocabulary testing in bilingual children. *Language, Speech, and Hearing Services in Schools*, 49(1), 85-97.
- Bedore, L. M. & Peña, E. D. (2008). Assessment of bilingual children for identification of language impairment. Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 11(1), 1-29.
- Bialystok, E., Luk, G., Peets, K. F. & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, 13(4), 525-531.
- Bonifacci, P., Atti, E., Casamenti, M., Piani, B., Porrelli, M. & Mari, R. (2020). Which measures better discriminate language minority bilingual children with and without developmental language disorder? A study testing a combined protocol of first and second language assessment. *Journal of Speech, Language, and Hearing Research*, 63(6), 1898-1915.
- Camilleri, B. & Botting, N. (2013). Beyond static assessment of children's receptive vocabulary: The dynamic assessment of word learning (DAWL). *International Journal of Language & Communication Disorders*, 48(5), 565-581.

- Dijkstra, J., Kuiken, F., Jorna, R. J. & Klinkenberg, E. L. (2016). The role of majority and minority language input in the early development of a bilingual vocabulary. *Bilingualism: Language and Cognition*, 19(1), 191-205.
- Dollaghan, C. A. & Horner, E. A. (2011). Bilingual language assessment: A meta-analysis of diagnostic accuracy. *Journal of Speech, Language and Hearing Research*, 54(4), 1077-1088.
- Ehl, B., Schrey-Dern, D. & Willmes, K. (2014). Der AWST-R bei sukzessiv mehrsprachigen Kindern. Eignung und Anpassung der Auswertung bei sukzessiven Erwerbsbedingungen. *Forum Logopädie*, 28(1), 30-34.
- Gagarina, N. (2013). Sprachdiagnostik in der Erstsprache mehrsprachiger Kinder (am Beispiel des Russischen). *Sprache · Stimme · Gehör*, 37(4), 196-200.
- Glück, C. W. (2011). *Wortschatz- und Wortfindungstest für 6- bis 10-Jährige (WWT 6-10)*. Handbuch. München, Jena: Elsevier, Urban & Fischer.
- Glück, C. W. (2021). Der WWT 6-10 und seine Testitems. *Bulletin suisse de linguistique appliquée*, 113, 29-47.
- Glück, C. W. & Elsing, C. (2014). Gestörte Lexikonenwicklung. In A. V. Fox-Boyer (Hg.), *Handbuch Spracherwerb und Sprachentwicklungsstörungen* (S. 73-86). München: Elsevier, Urban & Fischer.
- Glück, C. W. & Spreer, M. (2015). Zur Bildungsrelevanz semantisch-lexikalischer Störungen. *Sprache · Stimme · Gehör*, 39(2), 81-85.
- Golberg, H., Paradis, J. & Crago, M. B. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied Psycholinguistics*, 29(1), 41-65.
- Groot, A. M. B. d. (2011). *Language and cognition in bilinguals and multilinguals. An introduction*. New York, Hove: Psychology Press.
- Haman, E., Łuniewska, M., Hansen, P., Simonsen, H. G., Chiat, S., Bjekić, J., Blažienė, A., Chyl, K., Dabašinskienė, I., Engel de Abreu, P., Gagarina, N., Gavarró, A., Håkansson, G., Harel, E., Holm, E., Kapalková, S., Kunnari, S., Levorato, C., Lindgren, J., Mieszkowska, K., Montes Salarich, L., Potgieter, A., Ribu, I., Ringblom, N., Rinker, T., Roch, M., Slančová, D., Southwood, F., Tedeschi, R., Müge Tuncer, A., Ünal-Logacev, Ö., Vuksanović, J. & Armon-Lotem, S. (2017). Noun and verb knowledge in monolingual preschool children across 17 languages. Data from cross-linguistic lexical tasks (LITMUS-CLT). *Clinical Linguistics & Phonetics*, 31(11-12), 818-843.
- Haman, E., Łuniewska, M. & Pomiechowska, B. (2015). Designing cross-linguistic lexical tasks (CLTs) for bilingual preschool children. In S. Armon-Lotem (Hg.), *Assessing multilingual children. disentangling bilingualism from language impairment* (S. 196-240). Bristol: Multilingual Matters.
- Hamann, C. & Abed Ibrahim, L. (2017). Methods for identifying specific language impairment in bilingual populations in Germany. *Frontiers in Communication*, 2, 1-19.
- Hansen, P., Łuniewska, M., Simonsen, H. G., Haman, E., Mieszkowska, K., Kołak, J. & Wodniecka, Z. (2019). Picture-based vocabulary assessment versus parental questionnaires: A cross-linguistic study of bilingual assessment methods. *International Journal of Bilingualism*, 23(2), 437-456.
- Hoff, E. & Core, C. (2013). Input and language development in bilingually developing children. *Seminars in Speech and Language*, 34(4), 215-226.
- Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M. & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39(1), 1-27.
- Kapantzoglou, M., Restrepo, M. A. & Thompson, M. S. (2012). Dynamic assessment of word learning skills: Identifying language impairment in bilingual children. *Language, Speech, and Hearing Services in Schools*, 43(1), 81-96.
- Kohnert, K. J. & Bates, E. (2002). Balancing bilinguals II: Lexical comprehension and cognitive processing in children learning Spanish and English. *Journal of Speech, Language, and Hearing Research*, 45(2), 347-359.

- Kroll, J. F. & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149-174.
- Orellana, C. I., Wada, R. & Gillam, R. B. (2019). The use of dynamic assessment for the diagnosis of language disorders in bilingual children: A meta-analysis. *American Journal of Speech-Language Pathology*, 28(3), 1298-1317.
- Peña, E. D. & Bedore, L. M. (2009). Bilingualism in child language disorders. In R. G. Schwartz (Hg.), *Handbook of child language disorders* (S. 281-307). New York: Psychology Press.
- Peña, E. D., Bedore, L. M. & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one. *International Journal of Language & Communication Disorders*, 51(2), 192-202.
- Peña, E. D., Bedore, L. M. & Zlatic-Giunta, R. (2002). Category-generation performance of bilingual children: The influence of condition, category, and language. *Journal of Speech, Language, and Hearing Research*, 45, 938-947.
- Peña, E. D., Reséndiz, M. & Gillam, R. B. (2007). The role of clinical judgements of modifiability in the diagnosis of language impairment. *International Journal of Speech-Language Pathology*, 9(4), 332-345.
- Prevo, M. J. L., Malda, M., Mesman, J., Emmen, R. A. G., Yeniad, N., Van Ijzendoorn, M. H. & Linting, M. (2014). Predicting ethnic minority children's vocabulary from socioeconomic status, maternal language and home reading input: Different pathways for host and ethnic language. *Journal of Child Language*, 41(5), 963-984.
- Rothweiler, M. (2001). *Wortschatz und Störungen des lexikalischen Erwerbs bei spezifisch sprachentwicklungsgestörten Kindern*. Heidelberg: Universitätsverlag C. Winter.
- Rothweiler, M. & Ruberg, T. (2011). *Der Erwerb des Deutschen bei Kindern mit nichtdeutscher Erstsprache. Sprachliche und außersprachliche Einflussfaktoren*. München: Dt. Jugendinst.
- Scharff Rethfeldt, W. (2018). Evidenzbasierte logopädische Diagnostik bei mehrsprachigen Kindern mit Verdacht auf Sprachentwicklungsstörung. *Praxis Sprache*, 63(1), 12-17.
- Sheng, L. (2014). Lexical-semantic skills in bilingual children who are becoming English-dominant: A longitudinal study. *Bilingualism: Language and Cognition*, 17(3), 556-571.
- Sheng, L., McGregor, K. K. & Marian, V. (2006). Lexical-semantic organization in bilingual children. Evidence from a repeated word association task. *Journal of Speech Language and Hearing Research*, 49(3), 572.
- Smolander, S., Laasonen, M., Arkkila, E., Lahti-Nuutila, P. & Kunnari, S. (2020). L2 vocabulary acquisition of early sequentially bilingual children with TD and DLD affected differently by exposure and age of onset. *International Journal of Language & Communication Disorders*, 56(1), 72-89.
- Thordardottir, E. T. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, 15(4), 426-445.
- Thordardottir, E. T. (2015). Proposed diagnostic procedures for use in bilingual and cross-linguistic contexts. In S. Armon-Lotem (Hg.), *Assessing multilingual children. Disentangling bilingualism from language impairment* (S. 331-358). Bristol: Multilingual Matters.
- Tuller, L. (2015). Clinical use of parental questionnaires in multilingual contexts. In S. Armon-Lotem (Hg.), *Assessing multilingual children. Disentangling bilingualism from language impairment* (S. 301-330). Bristol: Multilingual Matters.
- Ulrich, T. (2017). Diagnostik lexikalischer Störungen bei spracherwerbsgestörten Kindern. *Sprachförderung und Sprachtherapie in Schule und Praxis*, 6(3), 140-147.
- Wilkens, R., Lein, T. & Rothweiler, M. (2018). Sprachdiagnostik bei zweisprachigen Kindern: Phonologische Verarbeitung und Wortschatzleistungen. *Praxis Sprache*, 63(1), 24-30.

- Winkes, J. (2014). *Isolierte Rechtschreibung*. Dissertation. Universität Freiburg, Freiburg/CH.
- Winkes, J. (2021). Die Bedeutung des mentalen Lexikons beim Wortlesen und -schreiben: Einsichten anhand der semantisch-lexikalischen Verarbeitungskompetenzen von Kindern mit und ohne Schriftspracherwerbsstörungen. *Bulletin suisse de linguistique appliquée*, 113, 117-135.
- Xuan, L. & Dollaghan, C. A. (2013). Language-specific noun bias: Evidence from bilingual children. *Journal of Child Language*, 40(5), 1057-1075.

# Wortschatzkompetenz erheben – eine synoptische Zusammenstellung wortschatzdiagnostischer Verfahren

**Kerstin ALBER**

PH Ludwigsburg  
Reuteallee 46, D-71634 Ludwigsburg  
alber@ph-ludwigsburg.de

In the article, a set of vocabulary diagnostic tools for children aged 0 to 12 were analyzed to determine to what extent specific areas and skills of lexical competence are being probed. Therefore, the three dimensions of the lexical space, "vocabulary breadth", "vocabulary depth", and "fluency" according to Daller et al. (2010) have been integrated in the distinction of representational knowledge and procedural abilities. Based on the classifications used for second language acquisition (Nation 2013; Schmitt 2014) and for language development disorders (Kannengieser 2019), a new characterization of lexical competence is derived. The sub-components of the complex construct vocabulary depth are aligned according to their hypothesized developmental order. This new categorization provides criteria for the synopsis that includes different survey methods such as parent questionnaires, language observation sheets, screenings and language tests. The compilation shows which sub-competencies of the lexical competence are not yet sufficiently taken into account in the language diagnostic procedure.

**Keywords:**

lexical competence, vocabulary breadth, vocabulary depth, fluency, procedural abilities.

**Stichwörter:**

Lexikalische Kompetenz, Wortschatzumfang, Wortschatztiefe, Wortzugriffsgeschwindigkeit, prozedurale Fähigkeiten.

## 1. "Wörter im Kopf"

### 1.1 Repräsentationale und prozedurale Aspekte der Wortschatzkompetenz

Das Ziel von wortschatzdiagnostischen Verfahren ist die Erfassung der Wortschatzkompetenz einer Person. Der Wortschatz einer Person ist im sogenannten mentalen Lexikon gespeichert (Aitchison 1997). Bei der mündlichen oder schriftlichen Sprachproduktion und Sprachrezeption wird auf das mentale Lexikon zugegriffen. Wortschatzkompetenz umfasst demnach eine repräsentationale Komponente, die sich auf die Speicherung der Wörter bzw. lexikalischen Einträge im Langzeitgedächtnis bezieht, und die prozedurale Fähigkeit, auf dieses Wissen zuzugreifen. Eine geeignete Grundlage für die Operationalisierung dieser beiden Aspekte stellt das Modell des lexikalischen Raums von Daller et al. (2007) dar. In diesem Modell entspricht die Dimension Wortschatzumfang ("vocabulary breadth") der repräsentationalen Komponente. Erfasst wird, wie viele lexikalische Einträge, also Kombinationen aus Wortform und Wortbedeutung, im mentalen Lexikon gespeichert sind. Die zweite Dimension im Modell des lexikalischen Raums, die Wortschatztiefe ("vocabulary

depth"), umfasst zum einen das verfügbare Wissen zu einem lexikalischen Eintrag als repräsentationale Komponente, das sich aus verschiedenen Wissenskomponenten zusammensetzt, und zum anderen die prozedurale Fähigkeit des Zugriffs auf die einzelnen repräsentierten Wissenskomponenten. In einer konkreten Interaktionssituation, beispielsweise ausgelöst durch die Instruktion "Was ist das Gegenteil von 'hell'?" im WWT 6-10 (Glück 2011; zum WWT siehe auch Beitrag von Glück 2021), wird die Verbindung von einem lexikalischen Eintrag, wie "hell", zu einer Wissenskomponente des lexikalischen Eintrags, zu "dunkel", hergestellt. Das repräsentierte Wissen stellt die Grundlage für die prozedurale Fähigkeit dar. Die prozedurale Fähigkeit ist ein mentales Werkzeug, das auch bei noch nicht voll ausgebildetem repräsentationalem Wissen eingesetzt werden kann. Das ist beispielsweise im frühen Spracherwerb von Kindern in deren Wortneuschöpfungen, z.B. "unhell", beobachtbar. Die Wortform "dunkel" ist nicht verfügbar und eine prozedurale Fähigkeit wird genutzt. Die dritte Dimension im Modell des lexikalischen Raums, die Wortzugriffsgeschwindigkeit ("fluency") ist eine prozedurale Fähigkeit auf Worteinträge zuzugreifen.

Soll Wortschatzkompetenz umfassend erhoben werden, dann müssen in diagnostischen Verfahren sowohl repräsentationale als auch prozedurale Aspekte berücksichtigt werden.

## *1.2 Repräsentationales Wissen und prozedurale Fähigkeiten in wortschatzdiagnostischen Verfahren*

### *1.2.1 Repräsentationales Wissen: Sprachliche Konzepte*

Im Gedächtnis ist mehr repräsentiert als das, was versprachlicht werden kann. Das zeigt sich zum einen individuell, wenn uns die Worte für Dinge, die wir ausdrücken möchten, fehlen oder wenn ein Kind im AWST-R (Kiese-Himmel 2005) die zu der bildlichen Darstellung von "schälen" passende Wortform nicht kennt. Zum anderen ist es möglich, dass in der betreffenden Sprache zu einer bestimmten Vorstellung keine Wortform existiert. Das mentale Lexikon ist somit nur ein Ausschnitt des gesamten konzeptuellen Repräsentationssystems "unter der Brille derjenigen Konzepte, die versprachlicht werden können" (Plieger 2006: 47). Für wortschatzdiagnostische Verfahren bedeutet das, dass nur Konzepte, für die es in der jeweiligen Sprachgemeinschaft per Konvention eine Wortform gibt, als Items verwendet werden können. Für Kinder mit Deutsch als Zweitsprache (DaZ) kann die Orientierung an der deutschsprachig-konzeptuellen Gliederung der Welt eine zusätzliche Schwierigkeit im Rahmen der Wortschatzdiagnose darstellen. So war es für einen Schüler mit Lingala als Erstsprache im WWTexpressiv (Glück 2011) nicht möglich, zur Instruktion "Was ist das alles?" das kategoriale Nomen "Pflanzen" abzurufen, weil in seiner Erstsprache Pflanzen in kleine und große Pflanzen kategorisiert und entsprechend versprachlicht werden (Schleher 2012). In seiner Erstsprache konnte er problemlos die Wortform "Baum" für große Pflanzen und die Wortform

"Blume" bzw. "Gras" für kleine Pflanzen abrufen, jedoch fehlte der Oberbegriff "Pflanzen" in der Zweitsprache. Das Beispiel stützt die Forderung, dass in zukünftigen wortschatzdiagnostischen Verfahren für die Ermittlung des Wortschatzumfangs insgesamt gezählt werden sollte, ob zu einem Konzept eine Wortform vorhanden ist, unabhängig davon, ob sie in der Erstsprache (L1) oder der Zweitsprache (L2) abgerufen werden kann (vgl. dazu Cross-linguistic lexical tasks (CLT), Rinker & Gagarin 2014; adaptives Testen im Subtest WWW\_D\_TR\_Standard\_adaptives\_Testen\_Lang-/Kurzform, Glück 2011). Diese Zählung wird vor allem Lernenden mit noch gering ausgeprägter Wortschatzkompetenz in L2 gerecht. Erst bei sehr fortgeschrittenen Lernern kann man von der konzeptuellen Reorganisation des L2-Lexikons ausgehen, was bedeutet, dass für jede Wortform in L1 und L2 ein eigenes Konzept aufgerufen werden kann (Flecken & von Stutterheim 2018).

Die Repräsentationen der sprachlichen Konzepte bestehen aus Wissenskomponenten unterschiedlicher Modalitäten wie imaginal, strukturell-abstrakt und emotiv-bewertend (Graf et al. 1996). Die Bedeutung eines Wortes ist Teil eines viel reichhaltigeren Konzepts (Löbner 2015; Aitchison 1997).

In Wortschatztests wie dem AWST-R (Kiese-Himmel 2005) und WWT 6-10 (Glück 2011) wird im Rahmen der qualitativen Auswertung auf die repräsentationale Komponente der Wortschatztiefe indirekt geschlossen. Dazu werden die fehlerhaften Antworten mit Hilfe der semantisch-lexikalischen Kriterien (semlexkrit) klassifiziert, indem die Beziehung zwischen der geforderten und der tatsächlich produzierten Benennung analysiert wird. Ruft ein Kind beispielsweise im WWTexpressiv zum Bildimpuls "Ellenbogen" die Bezeichnung "Arm" auf, so besteht zwischen der geforderten und der tatsächlich geäußerten Bezeichnung die Teil-Ganzes-Relation. Diese Form der Ersetzung wird als fortgeschrittene semantische Ersetzung bezeichnet (Glück 2011) und gilt als höherwertiger, als wenn das Kind auf seinen eigenen Ellenbogen zeigt. Ebenso erlaubt die fehlerhafte Bildauswahl (Wahl des Distraktors) im Subtest WWTrezeptiv einen Rückschluss darauf, über welche repräsentationalen Wissenskomponenten auf Wortform- und Bedeutungsseite ein Kind bereits verfügt. Die Klassifikation fehlerhafter Antworten erlaubt somit Rückschlüsse auf die Qualität des repräsentationalen Wissens.

### 1.2.2 Prozedurale Fähigkeiten in wortschatzdiagnostischen Verfahren: Zugriff auf sprachliche Konzepte

Die Fähigkeit, auf die Wissenskomponenten einer konzeptuellen Repräsentation eines sprachlichen Konzepts zuzugreifen, stellt eine prozedurale Fähigkeit dar, die der Wortschatztiefe zuzuordnen ist.

In wortschatzdiagnostischen Verfahren wird die prozedurale Fähigkeit beispielsweise im englischsprachigen WORD Test 2 Elementary (Bowers et al. 2004) überprüft, indem Kinder aufgefordert werden, ein Synonym zu einem

vorgegebenen Wort zu nennen. Die Notwendigkeit der Unterscheidung in repräsentationale und prozedurale Komponenten der Wortschatztiefe lässt sich mit dem Subtest WWTexpressiv-Abrufhilfen verdeutlichen. Wenn ein Kind die korrekte Benennung nicht nennen kann, dann wird zunächst eine allgemeine Abrufhilfe (z.B. "das kennst du bestimmt"), danach eine semantische Abrufhilfe mit Nennung von semantischen Relationen (z.B. dem Ort "das ist zwischen dem Unter- und dem Oberarm") und schließlich eine phonologische Abrufhilfe gegeben, bei welcher der Anfangslaut der Wortform genannt wird. Über das Abrufverhalten des Kindes kann differenziert diagnostiziert werden, ob das repräsentationale Wissen nicht ausreichend vorhanden ist (Nichtwirksamkeit der Abrufhilfe) oder ob repräsentationales Wissen zwar vorhanden ist, jedoch der Zugriff auf dieses Wissen das Problem darstellt (Wirksamkeit der Abrufhilfe).

Als weitere prozedurale Fähigkeit wird in aktiven Wortschatztests die Fähigkeit zum schnellen Abruf einer Wortform überprüft. Im AWST-R soll bei der Durchführung notiert werden, wenn der Abruf sich um mehr als 10 Sekunden verzögert. Am präzisesten wird die Abrufgeschwindigkeit im WWT 6-10 in der Computerversion erfasst. Mit einem Tastendruck wird die Zeit, bis das Kind mit der Benennung beginnt, automatisch festgehalten.

Auf sprachliche Konzepte kann mittels eines Bildimpulses zugegriffen werden und anschließend soll ein akustischer Output, eine Benennung, produziert werden. Dieser Verarbeitungsweg wird in den meisten Verfahren zur Überprüfung des aktiven Wortschatzes zugrunde gelegt. Der Zugriff erfolgt in diesem Fall über das nonverbale Eingangssystem, über das Bildmarkensystem (Engelkamp & Rummer 1999). Dieses aktiviert das konzeptuelle System und anschließend wird im verbalen Ausgangssystem ein motorisches Sprechprogramm aktiviert. Bei sehr jungen Kindern werden als Input zusätzlich zu den Abbildungen von Objekten auf Bildkarten echte Objekte verwendet wie im SETK-2 (Grimm 2016) und als Output wird die Produktion von Geräuschen oder ikonischen Wortformen erfasst wie im ELFRA-1 (Grimm et al. 2019). Beim Wortlesen bildet die graphematische Wortform den Input. Der Zugriff auf das konzeptuelle System erfolgt über das visuelle Wortmarkensystem (Engelkamp & Rummer 1999). Beim Wortlesen in lesediagnostischen Verfahren mit wortschatzdiagnostischen Anteilen muss beispielsweise in ELFE II (Lenhard et al. 2018) zu einem Bild die korrespondierende graphematische Wortform aus vier Auswahlmöglichkeiten ausgewählt werden. Wortschatzdiagnostische Verfahren variieren hinsichtlich der Modalität des Inputs, der verbal (akustisch oder graphematisch) oder nonverbal (Abbildung oder reales Objekt) sein kann sowie in der Modalität des geforderten Outputs, der verbal (Sprech- oder Schreibprogramm) oder nonverbal (motorisches Programm) gefordert sein kann. Da modalitätsspezifische Repräsentationssysteme für Bilder und Wörter angenommen werden (Kiefer 1999), stellen diese modalitätsspezifischen Zugriffsformen auf das mentale Lexikon ebenso wie die Prozessrichtungen Produktion oder Rezeption unterschiedliche, voneinander unabhängige



Verarbeitungswege dar. Das belegen ungewöhnliche Sprachstörungsbilder, wenn z.B. Personen zu einem gelesenen Wort keine Bedeutung aufrufen können, ihnen der Zugriff auf das konzeptuelle System von einem akustisch dargebotenen Wort jedoch möglich ist oder umgekehrt (Apeltauer 2006). In LEMO 2.0 (Stadie et al. 2013), einem individual-wortschatzdiagnostischen Verfahren für Aphasiepatienten, können diese Verarbeitungswege getrennt voneinander überprüft werden. Erkenntnisse aus verschiedenen wortschatzdiagnostischen Verfahren sind somit nur bedingt vergleichbar, wenn sich die Modalitäten von In- und Output und die berücksichtigten Verarbeitungswege unterscheiden, weil jeweils unterschiedliche prozedurale Fähigkeiten untersucht werden.

## 2. "What does it mean to know a word?"

Für das Konstrukt Wortschatzkompetenz und deren Teilkomponenten Wortschatzumfang, Wortschatztiefe und Wortzugriffsgeschwindigkeit existieren verschiedene Klassifikationsvorschläge. Die umfassendste, von Schmitt und Schmitt als "the best specification to date" (Schmitt & Schmitt 2020: 32) bezeichnete Aufgliederung der Wortwissensaspekte ist die Klassifikation von Nation (2013) aus dem Bereich der Zweitspracherwerbsforschung des Englischen. Die Klassifikation umfasst 18 Wissenskomponenten und bezieht sich auf das repräsentative Wissen ("word knowledge"). Es wird also nicht der tatsächliche Gebrauch, die prozedurale Fähigkeit, erfasst, sondern das Wissen über den Gebrauch, wie beispielsweise das rezeptive Wissen zur Auftretenshäufigkeit eines Wortes. Repräsentative Wissenskomponenten und prozedurale Fähigkeiten wie die Beherrschung der rezeptiven und produktiven Verarbeitungswege ("receptive versus productive mastery") und die Verarbeitungsgeschwindigkeit ("the ability to use lexical items fluently") finden sich in der von Schmitt (2014) vorgestellten Klassifikation.

Für den sonderpädagogischen Bereich stellt Kannengieser (2019) eine umfassende Klassifikation von Wortwissenskomponenten vor, die sich an der psycholinguistischen Modellkonzeption von Levelt (1989) mit der Unterscheidung von Lemma- und Lexeminformation für einen Eintrag im mentalen Lexikon orientiert. Auf der Lemmaebene wird in semantisches und syntaktisches Wortwissen und auf der Lexemebene in morphologisches und phonologisches Wortwissen klassifiziert. Zusätzlich kommt pragmatisches Wortwissen (z.B. "in welcher kommunikativen Situation ist der Gebrauch des Wortes passend?") dazu. Die Klassifikation enthält außerdem die schriftsprachliche Wortwissenskomponente (z.B. "aus welcher Buchstabenfolge besteht das geschriebene Wort?"). Das repräsentative Wortwissen bildet die Grundlage für die darauf fußenden Verarbeitungsprozesse. Die Auftrennung von repräsentativem Wissen und prozeduraler Fähigkeit innerhalb der Wortschatztiefe wird gestützt durch Störungsbilder von Kindern mit einer

spezifischen Sprachentwicklungsstörung, bei denen entweder die Speicherung des repräsentationalen Wortwissens beeinträchtigt ist (Strukturstörung) oder der Zugriff auf das gespeicherte Wissen, das differenziert ausgeprägt sein kann, erschwert ist (Prozessstörung).

### 3. Klassifikation der Wortschatzkompetenz

Der im Abschnitt 4 vorgestellten Synopse der wortschatzdiagnostischen Verfahren liegt die folgende Kategorisierung der Wortschatzkompetenz zugrunde.

- a) Alle drei Dimensionen des lexikalischen Raums mit der Unterscheidung in repräsentationales Wissen und prozedurale Fähigkeiten werden berücksichtigt.
- b) Modalitätsspezifische Zugriffswege auf sprachliche Konzepte, die von nicht-sprachlichen Inputs ausgehen oder auf nicht-sprachliche Outputs abzielen, werden berücksichtigt.
- c) Die prozeduralen Fähigkeiten werden entsprechend der angenommenen Entwicklungsabfolge der Wortschatzaneignung angeordnet.

Im Folgenden werden entlang der drei Dimensionen des lexikalischen Raums die potenziell messbaren Komponenten der Wortschatzkompetenz erläutert.

#### 3.1 Wortschatzumfang

Ein Eintrag im mentalen Lexikon besteht aus Wissen zur Wortform (Lexem) und zur Wortbedeutung (Lemma). Dieses Wissen bildet in wortschatzdiagnostischen Verfahren die Grundlage zur Ermittlung des Wortschatzumfangs. Dabei wird unterschieden, ob der lexikalische Eintrag im rezeptiven oder im produktiven Wortschatz vorhanden ist. Zum rezeptiven Wortschatzumfang zählt ein Eintrag dann, wenn zu einer Wortform eine Bedeutung aufgerufen werden kann (Rezeptiv [1]; die in eckiger Klammer angegebene Ziffer bezieht sich jeweils auf die Spalte der Synopse in Kapitel 4). Im produktiven Wortschatzumfang ist ein Eintrag vorhanden, wenn zu einer Bedeutung, beispielsweise zu einem Bild, eine akustische Wortform abgerufen werden kann (Produktiv [2]). Wenn ein Eintrag im mentalen Lexikon entweder rezeptiv oder produktiv vorhanden ist, dann geht mit diesem Wissen bereits eine minimale Wortschatztiefe einher (vgl. Schmitt 2014 zur Abgrenzung von Wortschatzumfang und Wortschatztiefe). Zum einen muss zumindest rudimentäres repräsentationales Wissen verfügbar sein und zum anderen die prozedurale Fähigkeit zum Zugriff auf dieses Wissen vorhanden sein. Da in den meisten Verfahren zur Erfassung des Wortschatzumfangs die Qualität des repräsentationalen Wissens nicht erhoben wird bzw. unspezifiziert bleibt, werden die Spalten zu Lexem [3] oder Lemma [4] in der Synopse in Kapitel 4 im Falle rein quantitativer Verfahren nicht markiert. Die jeweilige prozedurale Fähigkeit, rezeptiv [14] oder produktiv [15], wird markiert.

## 3.2 Wortschatztiefe

Wortschatztiefe umfasst repräsentationale Wissenskomponenten und prozedurale Fähigkeiten.

### 3.2.1 Repräsentationale Wissenskomponenten

Die repräsentationalen Komponenten der Wortschatztiefe beinhalten zum einen das Wissen zu einem lexikalischen Eintrag. Dieser umfasst phonologisches, graphematisches und morphologisches Wissen auf der Wortformseite (Lexem [3]) und semantisches und syntaktisches Wissen auf der Bedeutungsseite (Lemma [4]). Zum anderen zählt hierzu das kategoriale Wissen zu lexikalischen Einträgen. Dazu gehört Wissen zu den Wortarten wie Inhaltswörtern (Autosemantika [5]), Strukturwörtern (Synsemantika [6]) sowie lexikalischen Einheiten, die zu einer festen Form verwachsen sind ("formulaic language" (Schmitt 2010), d.h. formelhafte Mehrwortverbindungen (Formulaic [7]). Zum kategorialen Wissen zählt außerdem das Wissen zu verschiedenen Wortschatzebenen wie Alltagswortschatz, Fachwortschatz und akademischer Wortschatz. Dieser Aspekt ist in der Synopse nicht aufgeführt, da in den ausgewählten wortschatzdiagnostischen Verfahren aufgrund des Alters der Zielgruppe nur der Alltagswortschatz berücksichtigt wird. Und schließlich zählt zum repräsentationalen Wissen metasprachliches Wissen, wie das Wissen über die Auftretenshäufigkeit von Wörtern oder deren Gebrauchsbedingungen wie stilistische Varianten oder Unterschiede zwischen mündlicher und schriftlicher Sprache. Diese Wissenskomponente wird in der Synopse nicht dargestellt, weil sie in den aktuellen sprachdiagnostischen Verfahren bisher keine Rolle spielt.

### 3.2.2 Prozedurale Fähigkeiten

Die prozeduralen Fähigkeiten sind Katalysatoren für den Aufbau und die Nutzung von repräsentationalem Wissen. In Alber (2014) werden die prozeduralen Fähigkeiten in kognitive Grundfähigkeiten, implizite und explizite Strategien sowie metakognitive Fähigkeiten aufgegliedert. Die Entwicklungsabfolge zur "Wortform" ([8] bis 12]) beinhaltet die folgenden Schritte.

Erstes Wissen über die Wortform wird schon sehr früh, auch bereits schon vor der Geburt aufgebaut (Kauschke 2012). Im Alter von 6 Monaten nutzen Kinder rhythmische und etwas später phonotaktische Merkmale zum Segmentieren des kontinuierlichen Lautstroms, um einzelne Wortformen zu isolieren. Auch beim Erwerb einer neuen Sprache müssen die Wortformen zunächst aus dem Lautstrom isoliert werden. Vorschulkinder mit DaZ benötigen zum Dekodieren von Wörtern aus dem Grundwortschatz viele Wochen bis Monate (Apeltauer 2006: 22). Die isolierten Wortformen werden in der phonologischen Schleife des Arbeitsgedächtnisses in einer vorläufigen Form kurzfristig gespeichert und eine phonologische Repräsentation der Wortform wird erstellt, sodass diese bei erneuter Präsentation wiedererkannt werden kann (Rothweiler 2001). Im Alter

von 6 Monaten können Kinder ihren eigenen Namen wiedererkennen. Die Fähigkeit, isolierte Wortformen zu memorieren, kann als Eintrittspforte eines Wortes auf seinem Weg in das mentale Lexikon angesehen werden. In den testdiagnostischen Verfahren der Einschulungsuntersuchung wie HASE (Schöler & Brunner 2008) oder SSV (Grimm 2017) wird diese Teilkompetenz der Wortschatztiefe überprüft, indem Pseudowörter wie beispielsweise "wunore", also reine Wortformen ohne Bedeutung, nachgesprochen werden müssen (Prälinguistische Fähigkeiten [8]).

Die frühen Produktionen von Wortformen weichen zunächst noch von den korrekten zielsprachlichen Realisierungen ab. Die Wortformen für die Artikulation werden vereinfacht, indem Laute oder Silben ausgelassen oder ersetzt werden (Kauschke 2012: 35). Schließlich kann die Wortform korrekt artikuliert werden (Produktion [9]).

Mit zunehmender Verfestigung der Wortformen unterstützt die Fähigkeit, Verbindungen zu phonologisch ähnlichen und morphologisch verwandten Wortformen herzustellen, den Aufbau der repräsentationalen Struktur des mentalen Lexikons. Die Fähigkeit zur Herstellung von phonologischen Verbindungen wird in wortschatzdiagnostischen Verfahren in Form von Reimen überprüft wie im BISC (Jansen et al. 2002) oder wie im RWT (Aschenbrenner et al. 2000), indem zu einem Anfangslaut möglichst viele Wörter mit demselben Anfangslaut abgerufen werden sollen (Phonologische Verbindungen [10]). Die Fähigkeit zur Herstellung von morphologischen Verbindungen zeigt sich im kreativen Umgang mit Sprache, beispielsweise in den Wortneuschöpfungen von Kindern. Überprüft wird die morphologische Fähigkeit in wortschatzdiagnostischen Verfahren so gut wie nicht, bis auf den Subtest AD in HSET (Grimm & Schöler 2001) (Morphologische Verbindungen [11]). Mit der Aneignung der Schriftsprache durch das Lesen kommt eine visuelle Modalität hinzu, die Wahrnehmung des geschriebenen Wortes. Die graphematische Wortform wird in eine phonologische Repräsentation überführt, indem den einzelnen Buchstaben die passenden Laute zugeordnet werden. Die Abfolge der Laute der Wortform wird in der phonologischen Schleife des Arbeitsgedächtnisses repräsentiert und mit einer existierenden phonologischen Repräsentation abgeglichen. Diese wiederum aktiviert den dazu passenden Artikulationsplan (Levelt 1989). Wird innerhalb dieses Transformationsprozesses eine Lautfolge einer falschen phonologischen Repräsentation zugeordnet, dann stellt sich ein Lesefehler ein. Ein Beispiel hierfür wäre, wenn ein Kind in der WLLP-R zur graphematischen Wortform <Ei> die Abbildung "Eis" wählt (Graphematisch Rezeptiv [12]). Beim Schreiben bildet die graphematische Wortform den Output. Diese Fähigkeit wird in der Synopse nicht berücksichtigt. Hierfür wäre beispielsweise die HSP (May et al. 2018) ein geeignetes Instrument zur Erfassung der graphematisch produktiven Fähigkeit.

Die Entwicklungsabfolge zur "Wortbedeutung" ([13] bis [17]) umfasst die folgenden Schritte.

Lange bevor die ersten Einträge im mentalen Lexikon gespeichert werden, kategorisieren Kinder die sie umgebende Welt und bilden Konzepte aus (Poulin-Dubois & Pauen 2017, Motsch et al. 2018). Kategorisierung und Konzeptualisierung werden als sprachunabhängige kognitive Leistungen betrachtet (Rothweiler 2001, Löbner 2015). Eine zentrale Vorläuferfertigkeit für die Repräsentation von Wortbedeutungswissen im mentalen Lexikon ist der trianguläre Blickkontakt. Über diesen verfügt ein Kind in der Regel ab zwölf Monaten und dieser ermöglicht es dem Kind zu erkennen, welches Objekt im Zentrum der gemeinsamen Aufmerksamkeit steht. Da zu diesem Zeitpunkt bereits die ersten rudimentären Informationen zu Wortformen verfügbar sind, erkennt das Kind, dass Bedeutungen und Wortformen aufeinander bezogen sind. Unterstützend setzen Kinder Zeigegesten ein. Da es einen Zusammenhang gibt zwischen dem Einsatz von Gesten und dem späteren Wortschatzumfang (Grimminger et al. 2016) wird in wortschatzdiagnostischen Verfahren, die sich wie ELFRA-1 an sehr junge Kinder richten, der Einsatz von Zeigegesten erhoben. Die Fähigkeit zur Nutzung der Constraints, Ganzheits-, Taxonomie- und Disjunktionsannahme, vereinfachen den Zuordnungsprozess (Mapping) von Bedeutung zur Wortform ebenfalls (Kauschke 2012: 67). Die ersten Wortkonzepte von Kindern unterscheiden sich noch stark von den Wortbedeutungen der Erwachsenen. Häufig kommen bei Kindern Bedeutungsverengungen und Übergeneralisierungen vor (Szagun 2019: 133). Selbst bei Schuleintritt unterscheiden sich die singulären, also individuellen, Wortkonzepte von Kindern noch stark von den regulären Wortkonzepten von Erwachsenen (Osburg 2002) (Prälinguistische Fähigkeiten [13]).

Wenn zu einer akustischen Wortform ein Konzept aufgerufen werden kann, dann zählt ein Eintrag im mentalen Lexikon zum rezeptiven Wortschatzumfang. Entsprechend kreuzt beispielsweise die Bezugsperson in ELFRA die Antwortmöglichkeit "versteht" an, wenn das Kind das Wort versteht, es aber selber noch nicht sagt (Rezeptiv [14]). In einem folgenden Schritt kann zu einem Wortkonzept eine Wortform aufgerufen werden (Produktiv [15]).

Sehr früh werden inhaltlich zusammengehörende oder häufig im Satzkontext miteinander auftretende Wörter verbunden. Diese Fähigkeit, Kollokationen zu bilden, wird in wortschatzdiagnostischen Verfahren noch kaum berücksichtigt (Syntaktische Verbindungen [16]). In der weiteren Entwicklung kommt die Fähigkeit dazu, zu Wortkonzepten weitere Wortkonzepte abzurufen, die zueinander in einer bestimmten semantischen Beziehung stehen. Relativ früh können Ober- und Unterbegriffsrelationen sowie Beziehungen zwischen Kohyponymen gebildet werden, später folgen die Fähigkeiten zur Herstellung von Troponymie-, Antonymie-, Meronymie-, Synonymie- und Homonymiebeziehungen (Motsch et al. 2018: 24). Eine besonders produktive

Fähigkeit zur Erweiterung von Wortkonzepten ist die Fähigkeit zur Polysemie. Sie zeigt sich darin, dass die Bedeutung eines Konzepts auf weitere Bereiche übertragen oder ausgeweitet werden kann. Kinder sind schon sehr früh in der Lage Bedeutungsübertragungen vorzunehmen, wenn beispielsweise ein Globus als "Murmel" bezeichnet wird. Die früh auftretenden Übergeneralisierungen ähneln dem Prinzip der Polysemie, d.h. die Fähigkeit, Bedeutungsverschiebungen vorzunehmen, ist schon früh vorhanden, nur dass die Übertragungen nicht gebräuchlich sind. Lexikalisierte Polysemie mit ihren unterschiedlichen Bedeutungen können Kinder im Alter von 5 bis 6 Jahren produzieren (Semantische Verbindungen [17]).

Eine weitere Fähigkeit besteht darin, den Wortschatzaneignungsprozess eigenaktiv zu steuern. Hierzu zählen Kompensationsstrategien, Strategien zur Schließung lexikalischer Lücken, die Fähigkeit zum Nachfragen oder die Selbsteinschätzung der eigenen Wortschatzkompetenz (Selbstmanagement [18]).

### *3.3 Wortzugriffsgeschwindigkeit*

Erfolgreiche Sprachproduktion und -rezeption erfordert, dass auf repräsentiertes Wissen in angemessener Zeit zugegriffen werden kann. Die Verarbeitungsprozesse sind äußerst komplex und beinhalten kognitive Prozesse, die mit sprachdiagnostischen Verfahren nicht messbar sind. Sprachdiagnostische Verfahren können nur die wahrnehmbaren Komponenten der Wortzugriffsgeschwindigkeit erfassen. In aktiven Wortschatztests wird die Wortzugriffsgeschwindigkeit durch die Schnelligkeit des Abrufs einer Wortform gemessen. In Lesetests ist die Lesegeschwindigkeit Teil der Leseflüssigkeit (Wortzugriffsgeschwindigkeit (WZG) [19]).

### 4. Synopse wortschatzdiagnostischer Verfahren

Wortschatzkompetenz	Anzahl Wortschatzitems (%)	Wortschatzumfang	Wortschatztiefe												WZG					
			Repräsentationales Wissen				Prozedurale Fähigkeiten													
			Lexikal. Eintrag	Wortarten			Lexem (Wortform)				Lemma (Wortbedeutung)									
AWST-R 3-5 Jahre	100	Rezeptiv [1]	Lexem [3]	Lemma [4]	Autosemantika [5]	Synsemantika [6]	Formulaic [7]	Prälinguistische Fähigkeiten [8]	Produktion [9]	Phonologische Verbindungen [10]	Morphologische Verbindungen [11]	Graphematisch Rezeptiv [12]	Prälinguistische Fähigkeiten [13]	Rezeptiv [14]	Produktiv [15]	Syntaktische Verbindungen [16]	Semantische Verbindungen [17]	Selbstmanagement [18]	[19]	
		[2]: 82 Items, Instruktion: "Was ist das? Was macht die?" (bei Vorlage von Bild) [8]-[11], [13], [17]: qualitative Auswertung [18]: Auswertung: Nonverbale und sonstige Antworten: z.B. verbales Suchverhalten (Wie heißt das noch mal?)																		
WWT 6-10 5;6 bis 10;11 Jahre	100																			
		[1], [2]: <u>WWT</u> expressiv/rezeptiv: 95 (Langform) bzw. 40 (Kurzform) Items (Vorlage von Bild/Bildern) [3], [4]: qualitative Auswertung <u>SemLex</u> Krit: <u>WWT</u> Abstraktilien (Phonologie, Semantik) [8] - [11], [13]: qualitative Auswertung <u>SemLex</u> Krit: z.B. [11]: "fortgeschrittenes Bewältigungsverhalten: z.B. Wortneuschöpfung" [17]: Kategoriale Nomen: "Was ist/sind das alles?" (Hyperonymie); "Was ist das Gegenteil von...?" (Antonymie) [18]: Auswertung <u>SemLex</u> Krit: fortgeschrittenes Bewältigungsverhalten: Nachfrage um Bedeutung/Wortform																		
PPTV 3;0 bis 16;11 Jahre	100																			
		[1]: 228 Items (19 <u>Itemsets</u> à 12 Items; Vorschulalter (8-9 <u>Itemsets</u> = 96-108 Items), Grundschulalter (10 <u>Itemsets</u> = 120 Items), (Vorlage von Bildern) [5]: mit zunehmendem Schwierigkeitsgrad auch Fachwortschatz																		



Wortschatzkompetenz	Wortschatzumfang	Wortschatztiefe										WZG								
		Repräsentationales Wissen				Prozedurale Fähigkeiten														
		Lexikal. Eintrag		Wortarten		Lexem (Wortform)					Lemma (Wortbedeutung)									
	Anzahl Wortschatzitems (%)	Rezeptiv [1]	Produktiv [2]	Lexem [3]	Lemma [4]	Autosemantika [5]	Synsemantika [6]	Formulaic [7]	Prälinguistische Fähigkeiten [8]	Produktion [9]	Phonologische Verbindungen [10]	Morphologische Verbindungen [11]	Graphematisch Rezeptiv [12]	Prälinguistische Fähigkeiten [13]	Rezeptiv [14]	Produktiv [15]	Syntaktische Verbindungen [16]	Semantische Verbindungen [17]	Selbstmanagement [18]	[19]
Wortschatzkompetenz	100			RWT																
Sprachtests mit Wortschatzanteil	50			SETK-2 2;0 bis 2;11 Jahre																
	50			SETK-3-5 3;0 bis 3;11 Jahre																
	25			4;0 bis 5;11 Jahre																
	30			SET 5-10 5 bis 10 Jahre																
	38			HSET 3 bis 9 Jahre																







Wortschatzkompetenz	Anzahl Wortschatzitems (%)	Wortschatzumfang	Wortschatztiefe												WZG						
			Repräsentationales Wissen				Prozedurale Fähigkeiten														
			Lexikal. Eintrag		Wortarten		Lexem (Wortform)				Lemma (Wortbedeutung)										
Elternfragebögen	100	Rezeptiv [1] Produktiv [2]	Lexem [3]	Lemma [4]	Autosemantika [5]	Synsemantika [6]	Formulaic [7]	Prälinguistische Fähigkeiten [8]	Produktion [9]	Phonologische Verbindungen [10]	Morphologische Verbindungen [11]	Graphematisch Rezeptiv [12]	Prälinguistische Fähigkeiten [13]	Rezeptiv [14]	Produktiv [15]	Syntaktische Verbindungen [16]	Semantische Verbindungen [17]	Selbstmanagement [18]	[19]		
									[2]: 319 Wörter (verschiedene semantische Bereiche)												
	80																				
	33																				
	90																				

Wortschatzkompetenz	Wortschatzumfang	Wortschatztiefe														WZG					
		Repräsentationales Wissen				Prozedurale Fähigkeiten															
		Lexikal. Eintrag		Wortarten		Lexem (Wortform)					Lemma (Wortbedeutung)										
Wortschatzkompetenz	Anzahl Wortschatzitems (%)	Rezeptiv [1]	Produktiv [2]	Lexem [3]	Lemma [4]	Autosemantika [5]	Sysemantika [6]	Formulaic [7]	Pällinguische Fähigkeiten [8]	Produktion [9]	Phonologische Verbindungen [10]	Morphologische Verbindungen [11]	Graphematisch Rezeptiv [12]	Pällinguische Fähigkeiten [13]	Rezeptiv [14]	Produktiv [15]	Syntaktische Verbindungen [16]	Semantische Verbindungen [17]	Selbstmanagement [18]	[19]	
		(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)
		[2]: nur Grobeinschätzung: M3 "kann Gegenstände (z.B. 'Apfel', 'Stift', 'Schere', 'Auto') benennen: kann keine benennen/kann einige benennen/kann die meisten benennen" [17]: nur Grobeinschätzung: M4 "kann Gegenstände beschreiben (z.B. Apfel...ist rot,...ist rund,...ist rund,...kann man essen): gar nicht/Beschreibung eher grob/Beschreibung sehr differenziert")																			
Sprachbeobachtung	10	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	
		[2]: nur Grobeinschätzung: J5 "Alltagswortschatz des Kindes ist: reichhaltig/ausreichend/ingeschränkt/sehr eingeschränkt" [17]: nur Grobeinschätzung: J2 "kann Oberbegriffe finden, z.B. 'Kleider', 'Obst', 'Tiere', 'Pflanzen', 'Musikinstrumente' "																			
		[2]: nur Grobeinschätzung: L7 "Alltagswortschatz des Kindes ist reichhaltig/ausreichend/ingeschränkt, sehr eingeschränkt" [17]: nur Grobeinschätzung: L8: "kann Synonyme finden: mühelos/mit etwas Mühe/mit großer Mühe/gar nicht" [7]: L12: "versteht Redewendungen: mühelos/mit etwas Mühe/mit großer Mühe/gar nicht) [8]: nur Grobeinschätzung L3: "kann sich neue Wörter/Ausdrücke merken: mühelos/mit etwas Mühe/mit großer Mühe/gar nicht" [18]: L11: "umschreibt sinnvoll und nachvollziehbar, was es sagen möchte, wenn ihm ein Wort nicht einfällt: mühelos/mit etwas Mühe/mit großer Mühe/gar nicht") [19]: L9 "sucht beim Sprechen nach Wörtern, die es eigentlich kennt																			
Sprachbeobachtung	40	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	(✓)	
		[2]: Verbaler Wortschatz: Erfassung der produzierten Verben [5]: Verben [6]: Verbindung von Sätzen: Konjunktionen, Subjunktionen und Relativpronomen [19]: Indirekt/Grobeinschätzung: Bewältigung der Gesprächssituation: Flüssigkeit des Sprechens																			
		[19]: Indirekt/Grobeinschätzung: Bewältigung der Gesprächssituation: Flüssigkeit des Sprechens																			



Wortschatzkompetenz	Anzahl Wortschatzitems (%)	Wortschatzumfang	Wortschatztiefe												WZG
			Repräsentationales Wissen				Prozedurale Fähigkeiten								
			Lexikal. Eintrag		Wortarten		Lexem (Wortform)				Lemma (Wortbedeutung)				
Dile-D 1. bis 3. Klasse	100	Rezeptiv [1] Produktiv [2]	Lexem [3] Lemma [4]	Autosemantika [5] Synsemantika [6] Formuläre [7]	Prälinguistische Fähigkeiten [8] Produktion [9]	Phonologische Verbindungen [10] Morphologische Verbindungen [11]	Graphematisch Rezeptiv [12]	Prälinguistische Fähigkeiten [13]	Rezeptiv [14]	Produktiv [15]	Syntaktische Verbindungen [16]	Semantische Verbindungen [17]	Selbstmanagement [18]	[19] (✓)	
					↙		↘								
Lesetests	100													(✓)	
ELFE II 1. bis 7. Klasse	30													(✓)	

## 5. Fazit

Im Artikel wurde eine umfassende Kategorisierung der Komponenten der Wortschatzkompetenz vorgestellt, auf deren Grundlage die kriteriengestützte Analyse verschiedener wortschatzdiagnostischer Verfahren erfolgte. Die Synopse zeigt, dass die drei Dimensionen des lexikalischen Raums in den verschiedenen wortschatzdiagnostischen Verfahren unterschiedlich gut erfasst werden.

Für die Erfassung des Wortschatzumfangs existieren Tests, die eine lückenlose Diagnose des Wortschatzumfangs für die gesamte Altersgruppe von 0 bis 12 Jahren ermöglichen.

Die Wortschatztiefe hingegen wird in den Verfahren nur teilweise erfasst. Hier punktet neben dem WWT 6-10 (Glück 2011) interessanterweise das Urgestein der Sprachentwicklungstests, die HSET (Grimm & Schöler 2001). Sprachbeobachtungsverfahren wie selsa (Mayr et al. 2012) berücksichtigen einige der Teilkomponenten, sind jedoch aufgrund der grobkörnigen Einschätzungen auf uneinheitlichen Likert-Skalen für die Wortschatzanalyse nicht zu empfehlen.

Mit Blick auf die Entwicklung zukünftiger wortschatzdiagnostischer Verfahren sollten die prozeduralen Fähigkeiten stärker berücksichtigt werden, da die Fähigkeiten, phonologische, morphologische, semantische und syntaktische Verbindungen zwischen Einträgen im mentalen Lexikon herstellen zu können, positiv mit dem Wortschatzumfang korrelieren (Alber 2016). Insbesondere wird in keinem der vorgestellten wortschatzdiagnostischen Verfahren die Fähigkeit zur Herstellung von syntaktischen Verbindungen überprüft. Diese Fähigkeit ist grundlegend für die Wortschatzaneignung und zeigt sich schon in einem sehr frühen Stadium mit der Verwendung der sogenannten "chunks" und auch später bei der Ausbildung des bildungssprachlichen Wortschatzes, wenn beispielsweise Funktionsverbgefüge verwendet werden. Im anglo-amerikanischen Sprachraum wird diese Fähigkeit, syntaktische Verbindungen herzustellen, bei der Erfassung von Wortschatzkompetenzen bereits berücksichtigt (Lewis 1997, 2000, Schmitt & Schmitt 2005).

Zukünftige wortschatzdiagnostische Verfahren sollten zudem die Fähigkeit zur eigenaktiven Steuerung des Wortschatzerwerbs mitberücksichtigen.

## LITERATUR

- Aitchison, J. (1997). *Wörter im Kopf. Eine Einführung in das mentale Lexikon*. Tübingen: Niemeyer.
- Alber, K. (2014). Strategiebasierte Wortschatzaneignung. *dgs Praxis Sprache*, 59, 27-32.
- Alber, K. (2016). Wortschatzumfang, Wortschatztiefe und Verarbeitungsgeschwindigkeit – Analyse der lexikalischen Kompetenz. *Zeitschrift für angewandte Linguistik*, 65, 107-128.

- Apeltauer, E. (2006). Förderprogramme, Modellvorstellungen und empirische Befunde. Zur Wortschatz- und Bedeutungsentwicklung bei türkischen Vorschulkindern. In B. Ahrenholz (Hg.), *Kinder mit Migrationshintergrund. Spracherwerb und Fördermöglichkeiten* (S. 11-33). Freiburg im Breisgau: Fillibach.
- Aschenbrenner, S., Tucha, O. & Lange, K.W. (2000). *RWT. Regensburger Wortflüssigkeits-Test*. Göttingen: Hogrefe.
- Bockmann, A. & Kiese-Himmel, C. (2012). *ELAN. Eltern Antworten – Elternfragebogen zur Wortschatzentwicklung im frühen Kindesalter – Revision*. Göttingen: Hogrefe.
- Bowers, L., Huisingh, R., LoGiudice, C. & Oman, J. (2004). *The WORD Test 2 Elementary. East Moline. IL: LinguiSystems, Inc.*
- Daller, H., Milton, J. & Trefers-Daller, J. (2007). Editor's introduction. In H. Daller, J. Milton & J. Treffers-Daller (Hgg.), *Modelling and assessing vocabulary knowledge* (S. 1-32). Cambridge: University Press.
- Engelkamp, J. & Rummer, R. (1999). Die Architektur des mentalen Lexikons. In A. Friederici (Hg.), *Sprachrezeption. Enzyklopaedie für Psychologie* (S. 155-201). Göttingen: Hogrefe.
- Flecken, M. & von Stutterheim, C. (2018). Sprache und Kognition: Sprachvergleichende und lernersprachliche Untersuchungen zur Ereigniskonzeptualisierung. In S. Schimke & H. Hopp (Hgg.), *Sprachverarbeitung im Zweitspracherwerb* (S. 325-355). Berlin: de Gruyter.
- Glück, C. W. (2011). *Wortschatz- und Wortfindungstest für 6-bis 10-jährige. Handbuch WWT 6-10*. München: Elsevier.
- Glück, C. W. (2021). Der WWT 6-10 und seine Testitems. *Bulletin suisse de linguistique appliquée*, 113, 29-47.
- Graf, R., Herrmann, T., Grabowski, J. & Schweizer, K. (1996). Grundriß eines Modells der Aktivierung von Konzepten Wörtern und Figuren. In J. Grabowski, G. Harras & T. Herrmann (Hgg.), *Bedeutung. Konzepte. Bedeutungskonzepte* (S. 154-210). Opladen: Westdeutscher Verlag.
- Grimm, H. (2015). *SETK-3-5. Sprachstandserhebungstest für Kinder im Alter zwischen 3 und 5 Jahren*. Göttingen: Hogrefe.
- Grimm, H. (2016). *SETK-2. Sprachentwicklungstest für zweijährige Kinder (2;0 bis 2;11 Jahre)*. Göttingen: Hogrefe.
- Grimm, H. (2017). *Sprachscreening für das Vorschulalter - SSV. Kurzform des SETK 3-5*. Göttingen: Hogrefe.
- Grimm, H. & Schöler, H. (2001). *Heidelberger Sprachentwicklungstest - HSET*. Göttingen: Hogrefe.
- Grimm, H., Doil, H., Aktas, M. & Frevert, S. (2019). *ELFRA. Elternfragebögen für die Früherkennung von Risikokindern*. Göttingen: Hogrefe.
- Grimminger, A., Lüke, C., Ritterfeld, U., Liszkowski, U. & Rohlfing, K.J. (2016). Effekte von Objekt-Familiarisierung auf die frühe gestische Kommunikation. Individuelle Unterschiede in Hinblick auf den späteren Wortschatz. *Frühe Bildung*, 5, 91-98.
- Jansen, H., Mannhaupt, G., Marx, H. & Skowronek, H. (2002). *BISC. Bielefelder Screening zur Früherkennung von Lese-Rechtschreibschwierigkeiten*. Göttingen: Hogrefe.
- Kannengieser, S. (2019). *Sprachentwicklungsstörungen. Grundlagen, Diagnostik und Therapie*. München: Urban & Fischer.
- Kauschke, C. (2012). *Kindlicher Spracherwerb im Deutschen. Verläufe, Forschungsmethoden, Erklärungsansätze*. Berlin: de Gruyter.
- Kiefer, M. (1999). *Die Organisation des semantischen Gedächtnisses. Ereigniskorrelierte Potentiale bei der Kategorisierung von Bildern und Wörtern*. Bern: Hans Huber.
- Kiese-Himmel, C. (2005). *AWST-R. Aktiver Wortschatztest für 3- bis 5-jährige Kinder*. Göttingen: Hogrefe.

- Lenhard, A., Lenhard, W., Segerer, R. & Suggate, S. (2015). *Peabody Picture Vocabulary Test. 4th Edition*. Deutsche Adaption. Frankfurt: Pearson.
- Lenhard, W., Lenhard, A. & Schneider, W. (2018). *ELFE II. Ein Leseverständnistest für Erst- bis Siebtklässler*. Göttingen: Hogrefe.
- Levelt, W. (1989). *Speaking. From intention to articulation*. Cambridge, Mass: MIT Press.
- Lewis, M. (1997). *Implementing the lexical approach. putting theory into practice*. Hove: Language Teaching Publications.
- Lewis, M. (2000). Language in the lexical approach. In M. Lewis (Hg.), *Teaching collocations. Further developments in the lexical approach* (S. 126-154). Boston: Thomson Heinle.
- Löbner, S. (2015). *Semantik. Eine Einführung*. Berlin: de Gruyter.
- May, P., Malitzky, V. & Vieluf, U. (2018). *HSP 1-10. Hamburger Schreib-Probe 1-10*. Göttingen: Hogrefe.
- Mayr, T., Hofbauer, C., Simic, M. & Ulich, M. (2012). *Selsa. Sprachentwicklung und Literacy bei Kindern im Schulalter (1. bis 4. Klasse)*. Freiburg: Herder.
- Motsch, H.-J., Marks, D.-J. & Ulrich, T. (2018). *Wortschatzsammler. Evidenzbasierte Strategitherapie lexikalischer Störungen im Kindesalter*. München: Reinhardt.
- Nation, I.S.P. (2013). *Learning vocabulary in another language*. Amsterdam: John Benjamins.
- Osburg, C. (2002). *Begriffliches Wissen am Schulanfang: Schulalltag konstruktivistisch analysiert*. Freiburg im Breisgau: Fillibach.
- Paleczek, L., Seifert, S., Obendrauf, T., Schwab, S. & Gasteiger-Klicpera, B. (2017). *DiLe-D. Differenzierter Lesetest – Dekodieren*. Göttingen: Hogrefe.
- Petermann, F. (2018). *Sprachstandserhebungstest für Fünf- bis Zehnjährige (SET 5-10)*. Göttingen: Hogrefe.
- Plieger, P. (2006). *Struktur und Erwerb des bilingualen Lexikons*. Berlin: LIT Verlag.
- Poulin-Dubois, D. & Pauen, S. (2017). The development of categories. What? When? How? In H. Cohen & C. Lefebvre (Hgg.), *Handbook of categorization in cognitive science* (S. 653-666). Elsevier: Amsterdam.
- Rinker, T. & Gagarina, N. (2014). *CLT – Crosslinguistic Lexical Task - Deutsche Version*. Universität Konstanz & ZAS Berlin.
- Rothweiler, M. (2001). *Wortschatz und Störungen des lexikalischen Erwerbs bei spezifisch sprachentwicklungsgestörten Kindern*. Heidelberg: Winter.
- Schleher, L. (2012). *Der Einfluss des kulturellen Hintergrunds auf den Wortschatzaufbau von mehrsprachig aufwachsenden Kindern. Eine Einzelfallstudie*. (Unveröffentlichte Magisterarbeit). PH Ludwigsburg.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan UK.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64, 913-951.
- Schmitt, N. (2020). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schmitt, D. & Schmitt, N. (2005). *Focus on vocabulary. Mastering the academic word list*. White Plains, NY: Pearson Education.
- Schmitt, N. & Schmitt, D. (2020). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schneider, W., Blanke, I., Faust, V. & Küspert, P. (2011). *WLLP-R. Würzburger Leise Leseprobe – Revision*. Göttingen: Hogrefe.
- Schöler, H. & Brunner, M. (2008). *HASE. Heidelberger Auditives Screening in der Einschulungsuntersuchung*. Göttingen: Hogrefe.

- Schulz, P. & Tracy, R. (2011). *LiSe-DaZ. Linguistische Sprachstandserhebung – Deutsch als Zweitsprache*. Göttingen: Hogrefe.
- Stadie, N., Cholewa, I. & De Bleser, R. (2013). *LEMO 2.0 Lexikon modellorientiert. Diagnostik für Aphasie, Dyslexie und Dysgraphie*. Hofheim: NAT-Verlag.
- Szagun, G. (2019). *Sprachentwicklung beim Kind*. Weinheim: Beltz.
- Ulich, M. & Mayr, T. (2006a). *SISMIK. Sprachverhalten und Interesse an Sprache bei Migrantenkindern in Kindertageseinrichtungen*. Freiburg im Breisgau: Herder.
- Ulich, M. & Mayr, T. (2006b). *SELDAK. Sprachentwicklung und Literacy bei deutschsprachig aufwachsenden Kindern*. Freiburg im Breisgau: Herder.



# Assessing primary grade children's lexical inferencing strategies while reading – A review

**Ladina BRUGGER**

PHBern, Institut für Forschung, Entwicklung und Evaluation  
Fabrikstrasse 2a, CH-3012 Bern  
ladina.brugger@phbern.ch

**Britta JUSKA-BACHER**

PHBern, Institut Primarstufe  
Fabrikstrasse 8, CH-3012 Bern  
britta.juska@phbern.ch

Die Forschung zu lexikalischen Inferenzstrategien und -prozessen bei der Bedeutungserschliessung und damit beim Erwerb unbekannter Wörter konzentriert sich mehrheitlich auf zweitsprachliche Kontexte und interlinguale Hinweise. Zu lexikalischen Inferenzprozessen in der Erstsprache hingegen gibt es nur wenige und methodisch unterschiedliche Studien – insbesondere in der frühen Schulzeit. Ziel dieses Beitrags ist es, die Erkenntnisse zu Inferenzstrategien von Unterstufenkindern beim Lesen in ihrer Erstsprache in dieser begrenzten Zahl von Studien zusammenzutragen, zu untersuchen und unter methodischem Fokus zu diskutieren. Verschiedene Faktoren, besonders das Alter und die Lesefähigkeit, aber auch die Wahl der zu erschliessenden Wörter und des sie beinhaltenden Lesetexts scheinen nicht nur für den Erfolg der korrekten Identifikation eines unbekanntes Wortes entscheidend zu sein, sondern auch in Bezug auf die für die Bedeutungserschliessung verwendeten Hinweise und die Art der Strategien im Umgang mit den unbekanntes Wörtern. Es werden schliesslich Grenzen und Möglichkeiten dieser Methoden zur Erfassung lexikalischer Inferenzstrategien auf der Zielstufe diskutiert und weitere Forschungsperspektiven vorgeschlagen.

**Stichwörter:**

Inferenzstrategien, Erschliessungsprozesse, lexikalisches Inferieren, Leseerwerb, Wortschatzerweiterung beim Lesen.

**Keywords:**

inferencing strategies, deducing processes, lexical inferencing, reading acquisition, vocabulary acquisition while reading.

## 1. Introduction

In their first years of life, children learn new words through oral contexts. In mapping meaning onto unknown word forms, they take advantage of any clues available to them (Clark 2017: 396). When children learn to read, the main source of new vocabulary seems to gradually shift from oral to written contexts (Nagy et al. 1987; Anglin 2002; Bloom 2002). While the number and variety of unfamiliar words in oral contexts can dwindle, the variety of unfamiliar words encountered in written contexts increases<sup>1</sup>.

---

<sup>1</sup> Although vocabulary learning through reading becomes important when children learn to read, we need to keep in mind that the oral modality may still be an important source when children frequently interact with older peers and adults. In the absence of empirical data, it cannot be guaranteed that written contexts become the most important source of new vocabulary and it



Nagy et al. (1987) assume that English-speaking children learn up to 3000 new words annually between the third and eleventh grade. These words are acquired incidentally through reading – without direct instruction or conscious efforts of memorisation. This observation can at least be traced back to Gray & Holmes (1938: 28), who note that "practically all pupils acquire many meanings from context with little or no help from teachers".

Processes of lexical inferencing are thus crucial not only for reading comprehension, but also for incidental vocabulary acquisition (Wesche & Paribakht 2009). Empirical investigations suggest that acquisition processes of new vocabulary and reading skills are reciprocal: larger vocabulary favours reading skills and better reading abilities favour vocabulary acquisition through inferencing (Wagner & Meros 2010; Verhoeven et al. 2011).

Most children have acquired the technical aspects of reading in the first and second grade. Consequently, cognitive capacities are set free to engage in reading comprehension and vocabulary plays an increasingly important role (Biemiller 2012). An investigation of inference skills in reading thus becomes particularly interesting from the third school year onwards. During the primary school years, the ability to infer meaning increases more and more (Fukkink et al. 2001). This ability mediates the relation between reading comprehension and vocabulary knowledge (Cain et al. 2003).

Research on lexical inferencing has become prominent in the early 1970s with a focus on English as a Second Language (ESL) (Wesche & Paribakht 2009). Building up on these early investigations, different lines of research focussed on L1 and L2 vocabulary acquisition through reading – lines of research that were independent, but still at least indirectly influenced by each other. The bulk of research on lexical inferencing, however, is rooted in a Second Language Acquisition (SLA) framework and has mainly focussed on English (ESL). Methodologies in studies on lexical inferencing vary and cover authentic and manipulated text materials, case studies and cross-sectional group comparisons (e.g., skilled vs. less skilled readers), and intro- and retrospective verbal reporting (Wesche & Paribakht 2009). In both, L1 and L2 contexts, inferencing and the use of cues is important for reading comprehension. The process of lexical inferencing is similar in L1 and L2 (Wesche & Paribakht 2009) and empirical approaches are comparable (Haastrup 2008). Many theoretical approaches as well as empirical findings in the domain of foreign language learning are comparable to first language contexts. However, to better understand the reciprocal relations between vocabulary and reading acquisition and to answer specific questions on what information children use and how they use it to infer the meaning of unfamiliar words in written L1

---

cannot be generalised beyond societies where literacy and schooling are paramount. However, when children enter school, they will encounter new and more sophisticated words in written texts (Schleppegrell, 2001) and eventually learn them through this modality.

contexts, empirical investigations in the specific domain of reading in a first language in early school years are required. This contribution aims to provide a basis for such empirical investigations. A variety of methods used to assess inferencing strategies in a first language are reported and the potentials and questions that arise from these methods, as well as the transferability to early school years and other languages than English, are discussed.

## 2. Reading as a source for vocabulary acquisition

For several decades in reading research, it has been well established that reading – especially extensive reading – is an important source for vocabulary acquisition (Nagy et al. 1985, 1987; Krashen 1989; Parry 1991; Pigada & Schmitt 2006; Verhoeven et al. 2011). However, neither the precise nature of the relation between vocabulary knowledge and reading ability, let alone the mechanisms of vocabulary acquisition through reading, can clearly be explained by research evidence (Paribakht & Wesche 1997).

Some researchers have argued that learning words that are embedded in a given context can foster the learning of their referential, syntactic, pragmatic, or even emotional information (Gu & Johnson 1996) and lead to better retention than when learnt in isolation (Nation 1982; Bialystok 1983; Nation & Coady 1988). The latter assumes that inferring or inducing the solution of a problem implies an increased mental effort than when the solution of a problem is given. This increase in mental effort leads to a better retention of the information than when learned with less mental effort (Hulstijn 1992; Laufer & Hulstijn 2001).

While Nation (2001: 233), similarly to Schmidt (1993), acknowledges that "all learning implies conscious attention", he nevertheless distinguishes intentional from incidental vocabulary learning. In general terms, intentional learning involves the intention of learning and committing to one's memory, while incidental vocabulary learning refers to vocabulary learned from a certain context – without the intention of doing so or the learning of a particular feature or word while intending to learn another (Richards & Schmidt 2002).<sup>2</sup> One of the most frequently cited examples of incidental learning is vocabulary learning being a by-product of reading (Krashen 1989). Goodman (1967: 127) describes the process of reading as "a psycholinguistic guessing game". What defines efficient reading is not the precise identification of all elements, but the skill in selecting the most relevant cues to produce guesses (Goodman 1967).

---

<sup>2</sup> The discussion on the distinction between incidental and intentional learning is characterised by a terminological fuzziness: while some scholars use the terms synonymously to implicit and explicit learning, others maintain a difference (Hulstijn 2003). According to Ellis (1994: 1), the most characteristic features distinguishing implicit from explicit learning are the presence or absence of "conscious operations". While incidental learning is always implicit, implicit learning entails more than incidental learning (Hulstijn 2003).

Several scholars (e.g., Nation & Coady 1988; Parry 1993; Huckin & Coady 1999) make a distinction between guessing the meaning of a word in a context by means of several cues and the actual retention of the word's meaning. Hence, it needs to be borne in mind that inferencing the meaning of a word in a text does not imply the actual acquisition of that same word (e.g., Laufer & Hulstijn 2001). What a reader actually does – particularly a young reader in early school years – i.e., which strategies she or he applies to infer the meaning of an unknown word in a text and which strategies eventually lead to a better retention of a word, remains a question that merits empirical attention.

### 3. Lexical inferencing

Inferencing has been described as guessing the meaning of an unfamiliar word (Haastrup 2008; Wesche & Paribakht 2009). While reading, inference can, on the one hand, refer to text comprehension or "sense creation" where meaning is created based on linguistic and situational cues from the text (Haastrup 2008). According to Haastrup (1991), it is presupposed that the two types of inference depend on each other and the understanding of the procedures of one type explains the procedures of the other type, i.e., the acquisition of a word is prototypic of accessing the meaning of a text and vice versa. The type of inference referring to word identification has become a prominent concept in SLA and is referred to as lexical inference and is defined as follows by Haastrup (1991: 13):

The process of lexical inferencing involves making informed guesses as to the meaning of a word in the light of all available linguistic cues in combination with the learner's general knowledge of the world, her awareness of the co-text and her relevant linguistic knowledge.

Following this definition, which applies to all age groups including primary school children, lexical inferencing refers to much more than merely making random guesses. The process of "making informed guesses" comprises different cognitive processes such as analysing, extracting and integrating linguistic knowledge from context in combination with the existing knowledge of the learner. Hence, familiar attributes are used in recognising unfamiliar ones (Carton 1971). In Schmitt's terms (2010), lexical inferencing is thus best described as "qualified guessing of the meaning of lexical items in context, rather than guessing from context, as contextual cues are only one of several knowledge sources".

Lexical inferencing can thus be seen as a guessing procedure, promoted by a variety of cues, supported by different strategies and influenced by several factors.

### 3.1 Cues for lexical inferencing

Examining the types of cues to infer the meaning of a word in an L2, Carton (1971) establishes a categorisation of main cue-types: extra-lingual cues, intra-lingual cues and inter-lingual cues. This three-way distinction has largely been adopted and refined by scholars such as Haastrup (1991; 2008), Paribakht (2005) and Wesche & Paribakht (2009). In principle, this classification<sup>3</sup> can also be used for L1-speakers (interlingual cues<sup>4</sup>, however, only if the speakers know other languages, which is often not the case with younger children).

Within the top-down category of extra-lingual or contextual cues, Haastrup (1991) distinguishes between cues from the co-text and knowledge of the world. The co-text may refer to only one or two words surrounding the test word; to the immediate co-text, i.e., the sentence the test word is embedded in; to a specific part of the co-text beyond the sentence containing the test word; or to an unspecific, more general part of the text. Knowledge of the world refers to factual knowledge, attitudes, beliefs, prejudices – knowledge that is not taken from the text containing the test word.

Intra-lingual cues comprise cues on the level of phonology, orthography, morphology, lexis, semantics, as well as syntax (Haastrup 1991)<sup>5</sup>. As for phonological and orthographic cues, a learner may, for instance, search for similarities to a familiar word and does not consider meaning. Word association studies have shown that younger children tend to give more phonologically based associations before shifting to syntagmatic and paradigmatic associations as they get older (Namei 2004). This suggests that the mental lexicon is primarily phonologically based in earlier stages of development. When it comes to lexical inferencing, it can therefore be supposed that younger children tend to use more phonological cues than older ones. Although an analysis at the phonological or orthographic level alone can be sufficient for successful inferencing for both younger and older children, it represents usually only a first step of analysis before switching to a lexical or semantic level for more advanced language users. Word inference by means of morphological cues involves the decomposition of words by morphological

<sup>3</sup> The cues described and referred to in this paper are delimited to verbal context and do not take into account visual aids.

<sup>4</sup> Inter-lingual cues are relevant if a reader disposes of languages or language varieties in his or her repertoire that are similar to or remind her/him of the L1. In this case, as with intra-lingual cues, analyses at the level of phonology, orthography, morphology, lexis, collocation, semantics as well as syntax can be applied. This category is not further discussed since it goes beyond the scope of the present article where we focus exclusively on first language acquisition processes.

<sup>5</sup> At a more advanced stage, intra-lingual cues such as part of speech or collocations can also play an important role for inferencing. In early primary grades however, these cues are less used.

rules. At the lexical level, the meaning is taken into consideration. Usually, this category is a starting point to the semantic level of analyses, where meaning is explicitly reflected on. As for syntax, the learner focusses on the sentence structure to infer the meaning of a given test word. These intra-lingual cues are used differently according to age and language skills. While Werner & Kaplan (1950) show that children at age nine rather ignore syntactical cues, the findings of McKeown (1985) suggest that fifth graders with lower verbal ability often misused or did not use contextual cues at all and findings of Bangel (2018) show for fifth-graders that morphological cues were more often used by pupils showing higher verbal skills. Following these results, younger children and poor readers often seem to stop somewhere in the process of this stepwise analysis and take the meaning of a phonologically similar sounding word as the solution even though it would not fit in the syntactic or semantic context of the sentence.

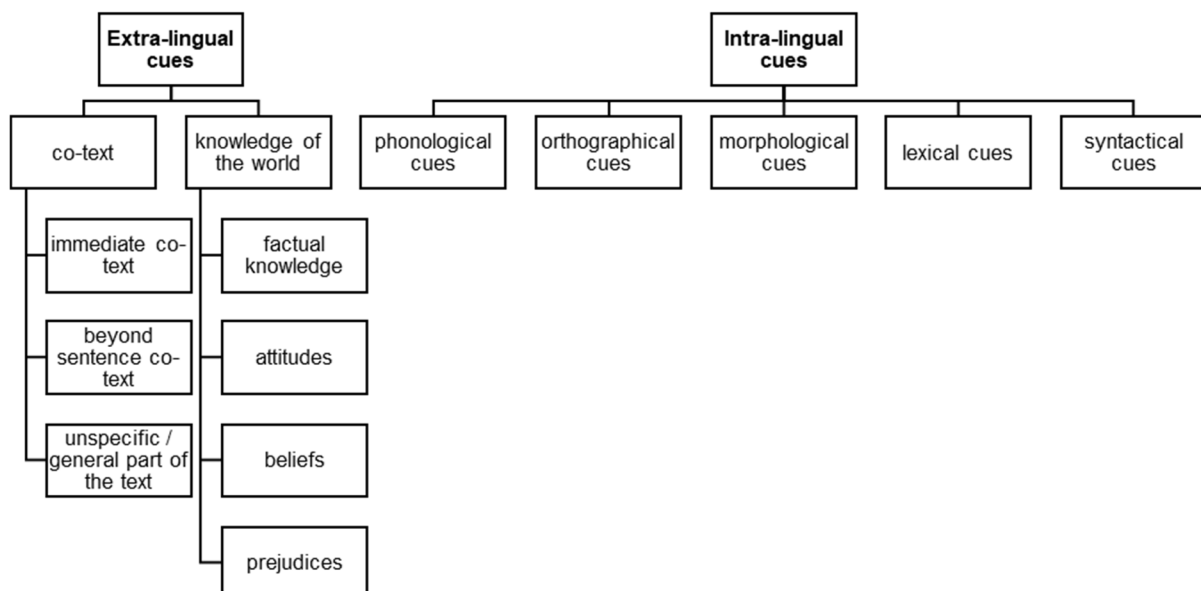


Figure 1: Cues for lexical inferencing according to Haastруп (1991). Figure created by the authors.

### 3.2 Lexical inferencing strategies

In the process of inferencing, a reader needs both different strategies and his or her background knowledge (Hu & Nassaji 2014). According to Kintsch's psycholinguistic model of text comprehension (Kintsch 2004), the reader builds a situation model in the process of inferencing, in which the information is given by the text and the schemata. The meaning selection can occur in a bottom-up manner and in a top-down manner where the schema acts as a filter, i.e., the context suppresses irrelevant information. Empirical research has shown support for both the top-down (Bensoussan & Laufer 1984) and the bottom-up manner (Schouten-van Parreren 1989) of meaning selection.

Refining the sources and strategies discussed in Nassaji (2003) and further addressing the question in how successful and less successful inferencers differ, Hu & Nassaji (2014) identify twelve types of inferencing strategies. These strategies are divided into four categories. The first category comprises form-focussed strategies such as analysing, associating and repeating. The second category refers to meaning-focussed strategies including the use of textual cues, prior knowledge and paraphrasing. The next category, termed evaluating strategies, refers to inquiry making, confirming or disconfirming and commenting. Finally, the fourth category refers to monitoring strategies including stating the failure or difficulty, suspending judgement, i.e. postponing the inference making, and reattempting.

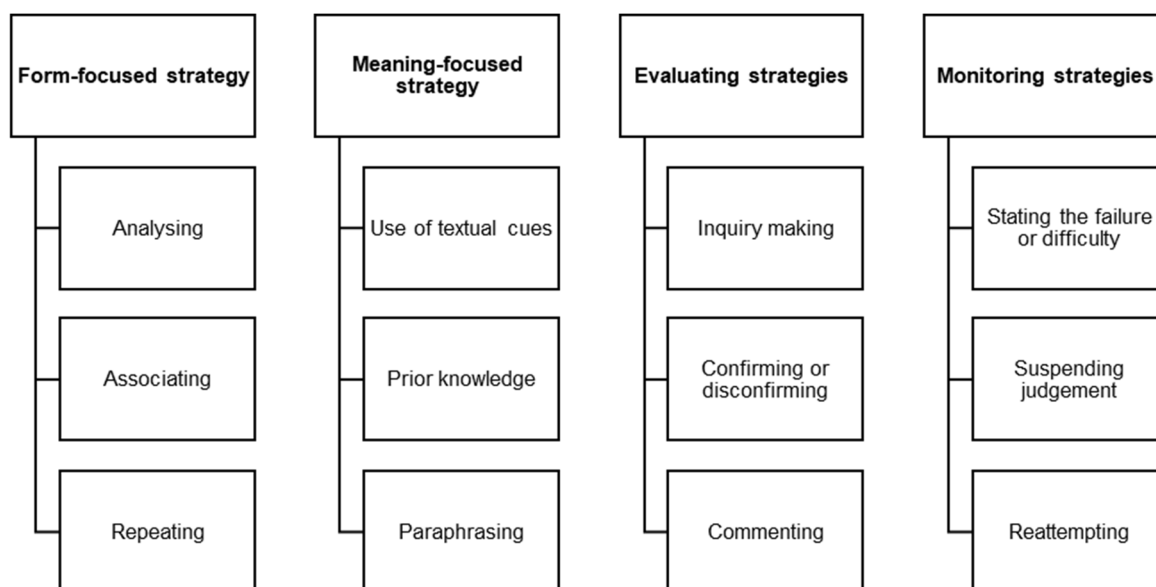


Figure 2: Lexical inferencing strategies according to Hu & Nassaji (2014: 68). Figure created by the authors.

Comparing the strategies used by successful and less successful inferencers, Hu & Nassaji (2014: 35) observe that it is not the quantity but rather the quality of strategies applied that differs:

There appeared to be no one way of combining different strategy types by the successful inferencers. They used a wide range of strategies and used them flexibly, depending on the word and the context. They not only used multiple strategies, but they also relied on other knowledge sources. They did not judge words individually but always attempted to relate them to the broader context. These qualities were not observed as much in less successful inferencers.

Hence, successful inferencing seems to depend on an appropriate choice of strategies, which is shaped by several factors.

### 3.3 Factors influencing lexical inferencing

Focussing on the question of what is required for successful inferencing, some scholars have identified several factors influencing the process of lexical inferencing. Among these factors are sufficient linguistic knowledge (Haastrup 1991; Nassaji 2006; Wesche & Paribakht 2009), background knowledge (Pulido 2009), motivation (Laufer & Hulstijn 2001; Hu & Nassaji 2012) and constant cognitive effort (Fraser 1999; Laufer & Hulstijn 2001; Hu & Nassaji 2012) during the process of lexical inferencing. These factors can be categorised as learner factors, language factors and task factors (Haastrup 1991). Among the learner-related factors are age (Nagy et al. 1985; Fukkink 2005), intelligence (Bloom 2002), working memory capacity (Cain et al. 2004), breadth and depth of vocabulary (Cain et al. 2004; Geva et al. 2017) and reading skills (Costa 2010). Concerning the latter factor, Costa (2010) states that skilled readers more easily make use of the context, activate previous knowledge, note and connect different parts of the text to infer meaning, while poor readers more frequently tend to ignore gaps of knowledge. This observation is in line with prominent research on reading development stating that novice readers or poor readers read in a slow, laborious manner which tends to impair comprehension. In other words, a process of automatisisation of basal reading skills is a prerequisite for additional cognitive demands induced by the process of deriving word meaning (Schwanenflugel et al. 2006). This automatisisation of decoding abilities or 'fluent reading' is usually developed from the third grade onwards when decoding skills are confirmed through practice (cf. Biemiller 2012).

In terms of the task factors or context, the variability of context, the relevance and the frequency of unknown words may also play a crucial role. As for the quantity of unknown words, the learner needs to recognise most of the surrounding words at first sight (Huckin & Coady 1999) to successfully infer the meaning. The findings of several studies – albeit from SLA research – suggest that the quantity of unknown words should not be more than 5%, ideally about 2% (e.g., Hu & Nation 2000). If the quantity of unknown words exceeds 5%, it may hamper comprehension of the overall text (Hu & Nation 2000).

## 4. Methods to assess lexical inferencing in first language contexts

In first language contexts, methods to assess lexical inferencing strategies differ first and foremost in terms of participants, target words, the context they are embedded in, as well as methods of asking which strategies participants used and how (see subsections 4.1 to 4.3).



#### 4.1 Participants

Most of the studies reviewed have been focusing on high school students (e.g., Quealy 1969; Sternberg & Powell 1983; Sternberg 1987; Albrechtsen et al. 2008) or PhD students (e.g., Ames 1966). Studies focussing on primary school children, are usually on upper primary grades such as fifth grade or above (McKeown 1985; Nagy et al. 1985; Neumann 1989; Van Daalen-Kapteijns & Elshout-Mohr 2001; Bangel 2018). We only found a few studies investigating lexical inferencing strategies of younger children (Werner & Kaplan 1950 and Nagy et al. 1987 with third-graders, Fukkink et al. 2001 and Fukkink 2005 with second- and fourth-graders). In studies with young participants from primary school, however, some methodological issues need to be considered. First, the quantity and quality of the text to assess inferencing strategies need to be adapted to novice readers. Secondly, the kinds of questions to be asked in the verbal protocol need to be adapted for this age group, given that even for advanced readers verbalising the thought processes is not an easy task.

Furthermore, there is also variation in terms of the number of participants and the profile of participants, such as linguistic background and literacy skills. The sample size varies considerably across studies. It is not uncommon to have a rather small sample size when verbal protocols in individual interviews are analysed (e.g., Van Daalen-Kapteijns & Elshout-Mohr 2001 or Bangel 2018, both with 16 participants). However, there are also studies with larger sample sizes (e.g., Werner & Kaplan 1950 with 125 participants or Nagy et al. 1987 with 352 participants).

Most studies on lexical inferencing focus on the English language with few exceptions (e.g., Van Daalen-Kapteijns & Elshout-Mohr 2001 on Dutch or Bangel 2018 on German). Several of the studies reviewed also emphasise basic differences between skilled and poor readers (McKeown 1985; Van Daalen-Kapteijns et al. 2001; Bangel 2018). These studies suggest that not only success or failure in deriving the meaning of a word, but also the type of vocabulary inferencing strategies depend on reading skills. Some scholars have explicitly taken the variable of "verbal ability" into account. This variable is operationalised differently: Van Daalen-Kapteijns & Elshout-Mohr (2001), for instance, used a Dutch vocabulary knowledge test called "Woordentoets Nederlands" consisting of lexical decision tasks; McKeown (1985) used a vocabulary subtest of the Stanford Achievement Test (Madden et al. 1973). Reading skills, however, have usually not been considered in the studies reviewed. An exception is the work of Bangel (2018) where a battery of different reading tests is used, such as ELFE 1-6 (Lenhard & Schneider 2006) to measure decoding ability and SLS 5-8 (Auer et al. 2005) to measure reading speed. Children who scored best on the tests showed a preference for an orientation on morphological cues, whereas children who scored less used

both context and orientation on morphological cues to the same degree to derive the meaning of a word.

#### 4.2 Target words and contexts

The methods applied in studies exploring lexical inferencing strategies in L1 vary greatly. A basic distinction can be made between "naturalistic" studies in which readers recorded or recounted how they dealt with unknown words in a text and how they inferred meaning and situations where researchers deliberately manipulated inferencing contexts (Wesche & Paribakht 2009).

A manipulated situation, for instance, would be the use of a pseudoword in a given context. Pseudowords ensure that the target words are unknown to all participants. The method is used in a handful of studies with different approaches of embedding a pseudoword in a context (Werner & Kaplan 1950; Quealy 1969; McKeown 1985). Werner and Kaplan (1950), for instance, embedded 12 pseudowords for their third graders referring to objects or actions in six different sentences each. Pseudowords are constructed differently with regards to the manner they match an existing word they are intended to replace. In Werner & Kaplan (1950), for instance, length of the pseudoword was not the same as the word that was replaced and the pseudoword could be translated in multiple ways.

In some studies, target words are real, usually low-frequency words (Nagy et al., 1987; Fukkink et al. 2001; Van Daalen-Kapteijns et al. 2001; Fukkink 2005; Albrechtsen et al. 2008; Bangel 2018). Real words may have the advantage that they represent a more natural inferencing context as the new label does not necessarily match a known concept, but simultaneously involves the learning of a new label. This, however, does not guarantee that in an experimental situation, participants do not simply try to replace unknown words by known words.

Target words – pseudowords or real words – are usually content words, mostly concrete nouns. In some studies, however, other word types<sup>6</sup> such as verbs or adjectives (McKeown 1985; Fukkink et al. 2001) are investigated. While target words are mostly embedded in sentences (Werner & Kaplan 1950) or texts, some scholars have also presented target words in isolation (e.g., Bangel 2018) – usually in addition to a task where target words are presented in context. In contrast to context-bound inferencing, this approach allows to focus more closely on morphological inference strategies (Bangel 2018).

The type and length of texts that the target words are embedded in depend on participants' age and profile. While texts for adult speakers usually consist of authentic material, texts for primary grade children are often constructed for

<sup>6</sup> Function words such as prepositions and discourse connectors usually lack clear semantic referents, are more difficult to infer meaning and are therefore usually not used in studies focussing on lexical inferencing (Wesche & Paribakht 2009).

the purpose of the study and adapted to the world of children (e.g., McKeown 1985). Given that the density of unknown words can influence vocabulary understanding and learning, it is important to consider how many test words should be embedded in a given text. As mentioned above, it is common practice to adopt the numbers of <5% in terms of density of test words. While these numbers have been taken up in contexts other than ESL, it is open to doubt whether they are directly transferable to other languages. Moreover, it is questionable whether the quota of <5% can also be applied in FLA contexts or if FLA contexts allow for more unknown words. Hypothetically, however, cognitive capacity and thus age are factors that need to be considered when considering the quota of pseudowords.

### *4.3 Methods of questioning*

There is not only a great deal of variation in how target words are presented, but also in how the derivation of meaning is asked and identified. The most prominent methodological approach to assess word inference strategies of older students or adults is the use of verbal protocols (e.g., Haastrup 1987, 1991; Paribakht & Wesche 1997; Nassaji 2003; Hu & Nassaji 2012, 2014). Verbal protocols can be used simultaneously or retrospectively and differ in terms of degree of metacognitivity (Heine 2005). Thus, verbal protocols showing a lower degree of metacognitivity – usually referred to as think-aloud protocols – comprise all kinds of mental processes including emotions and perceptions. The advantage of an introspective think-aloud procedure is that there is less interpretation or justification on the action as there is less time to reflect upon it. An obvious challenge, however, is to formulate questions that animate children to actually verbalise their thoughts. Verbal protocols showing a high degree of metacognitivity are characterised only by observations and reflections on thinking and doing (Heine 2005). Usually, these interview questions are conducted individually.

Depending on the particular aim of the study, the participants, the setting and the type of verbal protocols, interview questions can be more or less structured. As stated above, however, it is usually cognitively too demanding for primary school children to describe their thinking processes and alternative methods of questioning are required. Werner & Kaplan (1950) conducted individual interviews with 125 children (age range: 8.5 to 13.5 years). They were asked to provide meanings for the words and tell how and why the word fits into the sentence. In a second step, the same word was presented in another context and children were asked if and how the meaning could be applied to the second.

Further approaches such as multiple-choice format or definitions are also frequently employed. However, while these techniques may show whether a word is known and how well it is known, the results usually show little about the actual cues and strategies of inferencing the meaning of a word (e.g.,

Fukkink et al. 2001; Fukkink 2005). Nagy et al. (1987), for instance, have looked on how readers infer meaning from context by highlighting words and asking participants to provide definitions for these words. This method, however, does not show what cues and strategies a reader might use in a normal reading situation (*Id.*) and instead of learning from context, this scenario would rather be described as learning from definitions with examples (Gipe 1978).

## 5. Avenues for further research

In order to assess the vocabulary learning strategies in reading and to contribute to a deeper understanding of the processes underlying lexical inferencing strategies in early school years, there are several possible agendas following the literature review for further research to fill the gaps in the field. In the present section, the most important points are taken up and discussed: investigating young readers in first language contexts, assessing reading skills and vocabulary knowledge and creating adapted testing materials.

### 5.1 Studying young readers

As stated in the introductory part, the ability to inferencing is a crucial link between reading and vocabulary. In order to diagnose and eliminate difficulties at an early stage, it is important to investigate this ability. In the literature cited above, there are only a few studies investigating children in their early school years. The youngest children were in the second or third grade (Werner & Kaplan 1950; Nagy et al. 1987; Fukkink et al. 2001; Fukkink 2005). Provided that the critical age of the automatising of decoding abilities or fluent reading is usually developed from the third grade onwards (Biemiller 2012) when decoding skills are confirmed through practice (see section 3.3) (Biemiller 2012), this particular age group is interesting to investigate. However, investigating children needs careful methodological reflections. First, texts need to be related to their age, interests and the topics of reading they are used to (Bailey 2017). This involves ruling out certain endeavours or materials such as long texts and therefore many target words. Second, children only start to develop a metalinguistic knowledge at this age so that introspective think-aloud procedures are often too demanding and interview questions to assess word deriving strategies should not be too complex (McKay 2006).

In all studies reviewed, these interview sessions are conducted individually. Pair and group interviews would be an interesting method for further research for several reasons. First, pair interviews have the potential to show consensus, may generate richer responses by allowing participants to challenge one another's opinions and may allow interesting insights in co-construction of knowledge (Lewis 2012). Second, talking to peers can help to stimulate a more natural interview setting and help to remove formal or

intimidating barriers that may exist between an unknown experimenter and primary school children. It is advisable that pairs or groups are carefully selected (Lewis 2012) which can, for example, be arranged with the help of the classroom teacher who knows how the children work in groups.

### *5.2 Skilled and less skilled readers*

While in the majority of studies, the focus lies on skilled readers, some studies also underline the importance of having both skilled and less skilled readers in their sample given that they differ first and foremost in terms of word recognition (Bangel 2018). Reading skills, however, are hardly ever assessed in these studies. It is therefore important for further research to measure the different dimensions of reading skills systematically and comprehensively in order to explore children's inferencing strategies and drawing conclusions on their processes underlying the derivation of unknown words while reading. As for reading skills, it is also advisable to systematically assess vocabulary skills as both variables interact and influence each other. While this variable has been considered in certain studies (cf. previous section), there is still a lack of a multidimensional assessment of vocabulary knowledge which comprises both vocabulary depth and vocabulary breadth (see e.g., Binder et al. 2017 for a discussion of the relation between these variables and reading skills).

Furthermore, a lower level of reading skills or vocabulary knowledge often goes hand in hand with lower language skills in general. According to Namei (2004) then, it may be expected that children with a lower language proficiency level may significantly differ from children with a higher language proficiency level in terms of their strategies of lexical inferencing. Less skilled readers or readers with a lower vocabulary may more frequently turn to phonological cues, while their peers with better vocabulary knowledge or advanced readers may more frequently focus on semantic and syntactic relations of the target words (cf. intra-lingual cues in chapter 3.1).

Thus, the field of lexical inferencing research would benefit from large-scale studies connecting strategies with different dimensions of reading skills and vocabulary knowledge.

### *5.3 Target words and their density*

As seen in the previous section, pseudowords vary considerably in terms of how they are presented and embedded in a context, how they are created and how many pseudowords in relation to real words are inserted. When creating pseudowords, it is important that the combination of syllables conforms to the language's phonotactic rules. Programs such as Wuggy (Keuleers & Brysbaert 2010) or WordGen (Duyck et al. 2004) help to simplify this task by automatically controlling for variables such as number of letters, orthographic relatedness and bigram frequency. When it comes to the question of how many pseudowords can be inserted in a text, the quota of <5% adopted from the SLA

literature is often taken as a point of reference. The quantity of pseudowords depends without doubt very much on the reader's linguistic profile. Given that text length is an important criterion when assessing inferencing strategies of children, it needs to be considered that keeping the quota of <5% pseudowords inevitably means that only few pseudowords can be inserted in a given text. Furthermore, there is no reason why the same quota should be adopted to FLA contexts or other languages than English. When studying primary grade children, including less skilled readers, however, too many pseudowords can also be overwhelming and result in a decrease of motivation.

#### *5.4 Language*

Finally, given that the vast majority of studies are on lexical inferencing strategies in English, research would benefit from insights in other languages than English. Languages such as French have a rich morphological system which may facilitate lexical inference (Saidane et al. 2020). German, for instance, also has many complex compound words (Bangel 2018). Studying inferencing strategies in these languages would give insights about how parts of words are recognised and how morphological cues are considered. In addition, nouns are marked by capitalisation. Therefore, the question whether parts of speech are recognised and considered while inferring unknown words can be adopted when studying inferencing strategies in German texts.

### **6. Conclusions & Outlook**

In this paper, we aimed to show that several lines of inquiry are interesting to be explored in the domain of lexical inferencing strategies. For at least two decades, the trend in the field seemed to focus on lexical inferencing strategies in the domain of SLA. This domain is certainly worth being investigated and unquestionably yields interesting and important insights in how foreign language texts can be understood and words can be learned while reading. However, there are still some additional questions to be explored in FLA contexts in early school years. Lexical inferencing processes form the interface between vocabulary acquisition and reading comprehension and therefore serve as the basis for learning in all disciplines. They have an impact on all learners, be they L1 or L2. This has mainly been studied in L2 children. It is crucial to investigate these fundamental processes also for L1 children as soon as their focus is less on reading technique and more on reading comprehension such as in the third grade, when there is usually still an independence of the influence of other languages for many children.

### **Acknowledgements**

This article was funded by the SNSF (Project-173245). We thank Martina Röthlisberger and two anonymous reviewers for their insightful comments on an earlier version of this article.

## REFERENCES

- Albrechtsen, D., Haastrup, K. & Henriksen, B. (2008). Lexical knowledge, lexical inferencing and writing. In D. Albrechtsen, K. Haastrup & B. Henriksen (eds.), *Vocabulary and writing in a first and second language: Processes and development* (pp. 160-194). London: Palgrave Macmillan.
- Ames, W. S. (1966). The development of a classification scheme of contextual aids. *Reading Research Quarterly*, 2(1), 57.
- Anglin, J. M. (2002). The acquisition of word meaning II: Later lexical and semantic development. In D. A. Cruse et al. (eds.), *Lexikologie: Ein internationales handbuch zur natur und struktur von wörtern und wortschätzen* (pp. 1789-1800). Berlin: de Gruyter.
- Auer, M., Gruber, G., Mayringer, H. & Wimmer, H. (2005). *Salzburger Lese-Screening für die Klassenstufen 5–8*. Göttingen: Hogrefe.
- Bailey, A. L. (2017). Assessing the language of young learners. In E. Shohamy, I. G. Or & S. May (eds.), *Language testing and assessment* (pp. 323-342). Cham: Springer International Publishing.
- Bangel, M. (2018). *Wortbildungsstrukturen als Wegweiser beim Dekodieren*. Wiesbaden: Springer Fachmedien.
- Bensoussan, M. & Laufer, B. (1984). Lexical guessing in context in EFL reading comprehension. *Journal of Research in Reading*, 7(1), 15-31.
- Bialystok, E. (1983). Inferencing: Testing the "Hypothesis-Testing" hypothesis. In M. H. Long & H. W. Seliger (eds.), *Classroom oriented research in second language acquisition* (pp. 104-123). Cambridge Mass: Newbury House.
- Biemiller, A. (2012). Teaching vocabulary in the primary grades. In E. J. Kame'enui & J. F. Baumann (eds.), *Vocabulary instruction: Research to practice* (pp. 34-50). New York: Guilford Press.
- Binder, K. S., Cote, N. G., Lee, C., Bessette, E. & Vu, H. (2017). Beyond breadth: The contributions of vocabulary depth to reading comprehension among skilled readers. *Journal of research in reading*, 40(3), 333-343.
- Bloom, P. (2002). *How children learn the meanings of words* (Reprint). Cambridge Mass etc.: MIT Press.
- Cain, K., Oakhill, J.V. & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties. *Journal of Child Language*, 30, 681-694.
- Cain, K., Oakhill, J. & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge and memory capacity. *Journal of Educational Psychology*, 96(4), 671-681.
- Carton, A. (1971). Inferencing: A process in using and learning language. In E. Paul Pimsleur & E. Terence Quinn (eds.), *The psychology of second language learning* (pp. 671-681). Cambridge: University Press.
- Clark, E. V. (2017). Later lexical development and word formation. In P. Fletcher & B. MacWhinney, *The handbook of child language* (pp. 393-412). Malden MA: Blackwell.
- Costa, S. (2010). *Fremde Texte – fremde Wörter* (Ed. 8). Frankfurt etc.: Peter Lang.
- Duyck, W., Desmet, T., Verbeke, L. P. C. & Brysbaert, M. (2004). WordGen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments, & Computers*, 36(3), 488-499.
- Ellis, N. (1994). *Implicit and explicit learning of languages*. London: Academic Press.
- Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition*, 21(2), 225-241.



- Fukkink, R. G. (2005). Deriving word meaning from written context: A process analysis. *Learning and Instruction, 15*(1), 23-43.
- Fukkink, R. G., Blok, H. & De Glopper, K. (2001). Deriving word meaning from written context: A multicomponential skill. *Language Learning, 51*(3), 477-496.
- Geva, E., Galili, K., Katzir, T. & Shany, M. (2017). Learning novel words by ear or by eye? An advantage for lexical inferencing in listening versus reading narratives in fourth grade. *Reading and Writing, 30*(9), 1917-1944.
- Gipe, J. P. (1978). Investigating techniques for teaching word meanings. *Reading Research Quarterly, 14*(4), 624-644.
- Goodman, K. S. (1967). Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist, 6*(4), 126-135.
- Gray, W. S. & Holmes, E. (1938). *The development of meaning vocabularies in reading: An experimental study*. Chicago: University of Chicago.
- Gu, Y. & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning, 46*(4), 643-679.
- Haastrup, K. (1987). Using thinking aloud and retrospection to uncover learners' lexical inferencing procedures. In C. Faerch (ed.), *Introspection in second language research* (pp. 197-212). Clevedon, Avon, England: Multilingual Matters.
- Haastrup, K. (1991). *Lexical inferencing procedures or, talking about words: Receptive procedures in foreign language learning with special reference to English*. Tübingen: Gunter Narr.
- Haastrup, K. (2008). Lexical inferencing procedures in two languages. In D. Albrechtsen, K. Haastrup, & B. Henriksen (eds.), *Vocabulary and writing in a first and second language: Processes and development* (pp. 67-111). New York NY: Palgrave Macmillan.
- Heine, L. (2005). Lautes Denken als Forschungsinstrument in der Fremdsprachenforschung. *Zeitschrift für Fremdsprachenforschung: ZFF, 16*(2), 163-186.
- Hu, H. M. & Nassaji, H. (2012). Ease of inferencing, learner inferential strategies, and their relationship with the retention of word meanings inferred from context. *The Canadian Modern Language Review / La revue canadienne des langues vivantes, 68*, 54-77.
- Hu, H. M. & Nassaji, H. (2014). Lexical inferencing strategies: The case of successful versus less successful inferencers. *System, 45*, 27-38.
- Hu, H. M. & Nation, P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403-430.
- Huckin, T. & Coady, J. (1999). Incidental vocabulary acquisition in a second language: A review. *Studies in Second Language Acquisition, 21*(2), 181-193.
- Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. L. Arnaud & H. Béjoint (eds.), *Vocabulary and applied linguistics* (pp. 113-125). Basingstoke, UK: Macmillan.
- Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (eds.), *The handbook of second language acquisition* (pp. 349-381). Malden MA: Blackwell.
- Keuleers, E. & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*(3), 627-633.
- Kintsch, W. (2004). Revisiting the construction–integration model of text comprehension and its implications for instruction. In D. E. Alvermann, N. Unrau & R. B. Ruddell (eds.), *Theoretical models and processes of reading* (pp. 1270-1328). Newark DE: International Reading Association.
- Krashen, S. (1989). We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis. *The Modern Language Journal, 73*(4), 440-464.



- Laufer, B. & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26.
- Lenhard, W. & Schneider, W. (2006). *ELFE 1-6. Ein Leseverständnistest für Erst- bis Sechstklässler*. Göttingen: Hogrefe Verlag.
- Lewis, T. N. (2012). The effect of context on the L2 thinking for speaking development of path gestures. *L2 Journal*, 4(2), 247-268.
- Madden, R., Gardner, E. F., Rudman, H. C., Karlsen, B. & Merwin, J. C. (1973). *Stanford achievement test*. New York NY: Harcourt Brace Jovanovic.
- McKay, P. (2006). *Assessing young language learners*. Cambridge: University Press.
- McKeown, M. G. (1985). The acquisition of word meaning from the context by children of high and low ability. *Reading Research Quarterly*, 20(4), 482-496.
- Nagy, W. E., Anderson, R. C. & Herman, P. A. (1987). Learning word meanings from context during normal reading. *American Educational Research Journal*, 24(2), 237.
- Nagy, W. E., Herman, P. A. & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20(2), 233-253.
- Namei, S. (2004). Bilingual lexical development: A Persian-Swedish word association study. *International Journal of Applied Linguistics*, 14(3), 363-388.
- Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly*, 37(4), 645-670.
- Nassaji, H. (2006). The relationship between depth of vocabulary knowledge and L2 learners' lexical inferencing strategy use and success. *The Modern Language Journal*, 90(3), 387-401.
- Nation, I. S. P. (1982). Beginning to learn foreign vocabulary: A review of the research. *RELC Journal*, 13(1), 14-36.
- Nation, I. S. P. & Coady, J. (1988). Vocabulary and reading. In M. Carter & M. McCarthy (eds.), *Vocabulary and language teaching* (pp. 97-110). London: Longman.
- Neuman, S. (1989). Assessing children's inferencing strategies. *Paper presented at the annual meeting of the national reading conference*, 39<sup>th</sup>. Austin, Texas.
- Paribakht, T. S. (2005). The influence of first language lexicalization on second language lexical inferencing: A study of farsi-speaking learners of English as a foreign language. *Language Learning*, 55(4), 701-748.
- Paribakht, T. S. & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174-200). Cambridge: University Press.
- Parry, K. (1991). Building a vocabulary through academic reading. *TESOL Quarterly*, 25(4), 629-653.
- Parry, K. (1993). Too many words: Learning the vocabulary of an academic subject. In T. Huckin, M. Haynes & J. Coady (eds.), *Second language reading and vocabulary learning* (pp. 109-129). Norwood: Praeger.
- Pigada, M. & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, 18, 1-28
- Pulido, D. (2009). How involved are American L2 learners of Spanish in lexical input processing tasks during reading? *Studies in Second Language Acquisition*, 31(1), 31-58.
- Quealy, R. J. (1969). Senior high school students use of contextual aids in reading. *Reading Research Quarterly*, 4(4), 512-533.
- Richards, J. C. & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics* (3rd ed.). London: Pearson Longman.

- Saidane, R., Fejzo, A. & Whissell-Turner, K. (2020). La relation entre les connaissances morphologiques et l'acquisition des mots chez des élèves francophones de 9 ans. *Repères*, 61, 57-74.
- Schlepppegrell, M. J. (2001). Linguistic features of the language of schooling. *Linguistics and Education*, 12(4), 431-459.
- Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13, 206-226.
- Schouten-van Parreren, C. (1989). Reading foreign languages in the first phase of secondary education: Why, what and how? *European Journal of Teacher Education*, 12(2), 131-140.
- Schwanenflugel, P. J., Meisinger, E. B., Wisenbaker, J. M., Kuhn, M. R., Strauss, G. & Morris, R. D. (2006). Becoming a fluent and automatic reader in the early elementary school years. *Reading research quarterly*, 41(4), 496-522.
- Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (eds.), *The nature of vocabulary acquisition* (pp. 89-105). Hillsdale, NJ: Lawrence Erlbaum.
- Sternberg, R. J. & Powell, J. S. (1983). Comprehending verbal comprehension. *American Psychologist*, 38(8), 878-893.
- Van Daalen-Kapteijns, M. van, Elshout-Mohr, M. & Glopper, K. de. (2001). deriving the meaning of unknown words from multiple contexts. *Language Learning*, 51(1), 145-181.
- Verhoeven, L., Van Leeuwe, J. & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading*, 15(1), 8-25.
- Wagner, R. K. & Meros, D. (2010). Vocabulary and reading comprehension: Direct, indirect, and reciprocal influences. *Focus on Exceptional Children*, 43, 1-17.
- Werner, H. & Kaplan, E. (1950). The acquisition of word meaning: A developmental study. *Monographs of the Society of Research in Child Development*, 25(1), (Serial No. 51).
- Wesche, P. M. B. & Paribakht, P. T. S. (2009). *Lexical inferencing in a first and second language: Cross-linguistic dimensions*. Bristol: Multilingual Matters.

## Appendix

## Research on primary school children's lexical inferencing strategies while reading

Study	Method	Participants
<i>Bangel 2018</i>	<p>Materials: Texts &amp; isolated real words.</p> <p>Task &amp; Procedure: Children were asked to read aloud texts and words, summarise them and verbalise their thoughts in deriving the meaning.</p>	16 skilled readers, 11 poor readers (2 reading tests helped to categorize the readers), 5 <sup>th</sup> grade
<p><i>Fukkink, 2005</i></p> <p><i>And also:</i></p> <p><i>Fukkink et al. 2001</i></p>	<p>Materials: Short narratives for 12 target words (6 abstract &amp; 6 concrete real words of low frequency)</p> <p>Task &amp; Procedure: Children were asked to read the texts aloud and to provide definitions of target words in interview questions.</p>	30 primary school children in 2 <sup>nd</sup> , 4 <sup>th</sup> and 6 <sup>th</sup> grade at four different schools in Amsterdam with varying cultural and SES backgrounds.
<i>McKeown 1985</i>	<p>Material: 6 pseudowords (of different word types: 2 nouns, 2 verbs &amp; 2 adjectives) in different contexts</p> <p>Task &amp; Procedure: Different meanings for each word were provided and children were asked if the meaning was correct and why or why not.</p>	30 5 <sup>th</sup> graders with high and low verbal ability operationalized via the vocabulary subtest of the Stanford Achievement Test (Madden et al. 1973)
<i>Nagy et al. 1987</i>	<p>Materials: an expository or narrative text with difficult but real words (evaluated by different raters before the experiment)</p> <p>Task &amp; Procedure: vocabulary assessment tasks on 15 target words from each passage, an individual interview (defining target words and using them in a sentence) and a multiple-choice test, both designed to tap partial knowledge of word meanings</p>	352 3 <sup>rd</sup> , 5 <sup>th</sup> and 7 <sup>th</sup> graders
<i>Neumann 1989</i>	<p>Materials: 2 mystery stories divided into 6 episodes each ending with a clue for resolving the case.</p> <p>Task &amp; procedure: in individual think-aloud protocols children were asked to verbalize their thoughts about resolving the cases.</p>	42 5 <sup>th</sup> graders from 11 classes

<i>Werner &amp; Kaplan 1950</i>	<p>Material: 12 pseudowords embedded in 6 different contexts</p> <p>Task &amp; Procedure: In individual interviews, children were asked to provide meanings for the words and tell how and why the word fit into the sentence. In a second step, the same word was presented in another context and children were asked if and how the meaning could be applied to the second</p>	125 children in five age groups from 8.5 to 13.5 years
<i>Van Daalen-Kapteijns, Elshout-Mohr &amp; De Glopper, 2001</i>	<p>Material: 10 unknown words (Dutch low-frequency words) embedded in 3 different contexts.</p> <p>Task &amp; Procedure: Think-aloud protocols on: decontextualizing, inferring meaning with subsequent contextual aid, providing definitions</p>	16 children (11-12 years old), half of them with high, the other half with low verbal ability. Verbal ability was operationalized via a test for Dutch vocabulary knowledge "Woordentoets Nederlands" (p. 159).