Zeitschrift: Bulletin suisse de linguistique appliquée / VALS-ASLA

Herausgeber: Vereinigung für Angewandte Linguistik in der Schweiz = Association suisse de linguistique appliquée

Band: - (2021)

Heft: 113: Wortschatzkompetenzen definieren, erheben und fördern = Defining, assessing and fostering vocabulary skills

Artikel: Der WWT 6-10 und seine Testitems

Autor: Glück, Christian W.

DOI: https://doi.org/10.5169/seals-1030124

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 10.12.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

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, Erwerbsalter, 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 therapieleitende Aussagen zum expressiv und rezeptiv aktivierbaren Lexem-Bestand und über Detailanalysen korrekter und nicht-korrekter 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 kontrastierenden 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 abrufbare 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 therapieleitenden 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 Itemsubsets (Wortarten) und die Analyse der Falschantworten.

Damit bilden die Aufgaben isoliert Prozesse aus psycholinguistischen inkrementellen (Levelt 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 Erwerbsalter, 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 Erwerbsalter 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 ≥ .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 Erwerbsalter, 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 Werte-Niveau 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. 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 Erwerbsalter, Imageability, Familiarität und Konkretheit

Als einer der stärksten Einflussfaktoren auf die Geschwindigkeit des Abrufs (für eine Übersicht Perret & Bonin 2019) gilt das Erwerbsalter (*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 Erwerbsalter und Imageability korreliert. Als deutschsprachige Datenbasis dieser Itemmerkmale kommen in Betracht die Arbeiten von Schröder et al. (2003), Schröder et al. (2012) bezugnehmend 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 Erwerbsalter und Imageability ableitbar.

Für 258 deutsche Adjektive liegen Normen zum Erwerbsalter 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 Erwerbsalter angegeben werden (MW 5.2 Jahre, SD 1.7). Die Übereinstimmung des Erwerbsalters 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 Erwerbsalters 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 Grundschulkinder 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 Benennaufgabe 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 Repräsentaten 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.

Altersbereich / Itemkategorie	5;6-6;11	7:0-7;11	8;0-8;11	9;0-9;11	10;0-10;11
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 dichterer 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 relatiert. Insbesondere die phonologisch (Anlaut) relatierten Items mussten entlang der türkischen Aussprache gefunden werden. Wenn semantische oder nicht-relatierte 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. Rasch-skalierte 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

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 ¾ 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

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. Dabóczi & S. Hartlmaier (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

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

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

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 Grundschülern 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

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

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

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.