



<https://bulletin.vals-asla.ch> · e-ISSN: 2813-7973 · p-ISSN: 1023-2044

Till, C. (2021). Semantisch-lexikalische Diagnostik bei sukzessivbilingualen Kindern. *Bulletin suisse de linguistique appliquée*, 113, 175-192. <https://doi.org/10.5169/seals-1030131>

Dieser Artikel ist unter der *Creative Commons Attribution 4.0 International* Lizenz veröffentlicht (CC BY):
<https://creativecommons.org/licenses/by/4.0>



© Christoph Till, 2021

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.

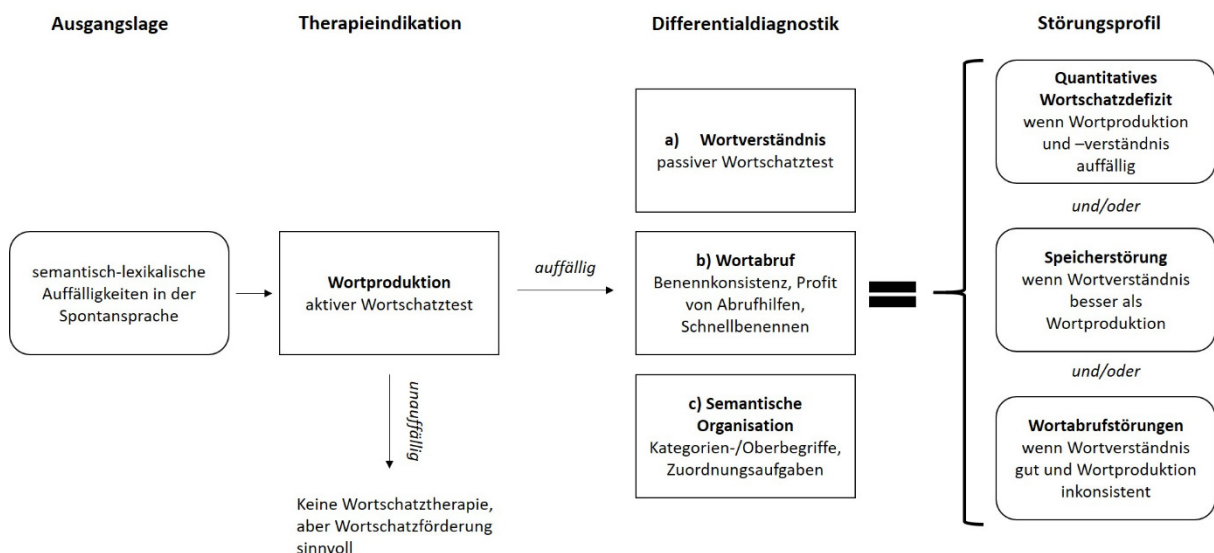


Abbildung 1: Modell der semantisch-lexikalischen Diagnostik (in Anlehnung an Ulrich 2017)

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 vorgesprochenes 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 Wortverständnistest 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 Abrufrouinen 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 Abrufrouinen. 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).

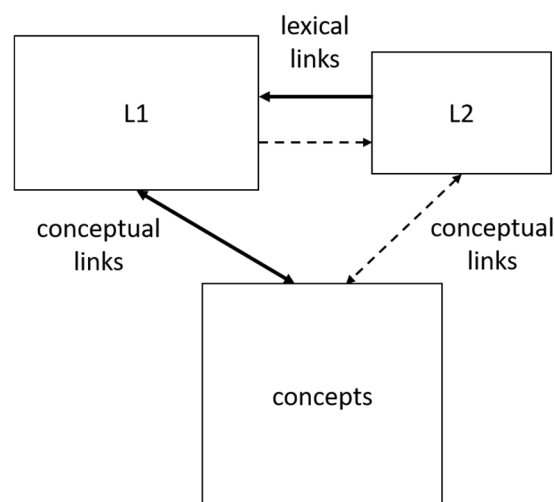


Abbildung 2: "revised hierarchical model of lexical and conceptual representation in bilingual memory" von Kroll & Stewart (1994: 158)

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 Wortverständnisseleistungen 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 Wortverständnisseleistungen 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 Wortlern Diagnostik 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.

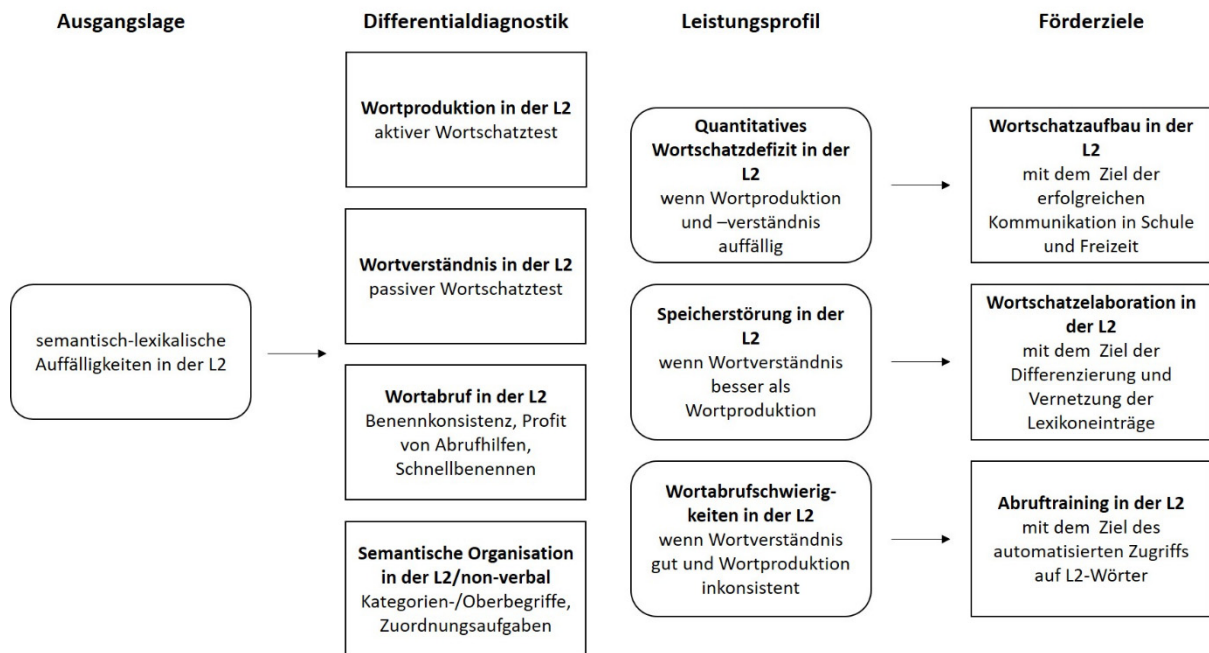


Abbildung 3: Pädagogisch-orientierte L2-Wortschatzdiagnostik in Anlehnung an Ulrich (2017)

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 Lexikonentwicklung. 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-Nuuttila, 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.