

# Ausagieren oder Satz-Bild-Zuordnung?

## Zwei Methoden zur Untersuchung des Grammatikverständnisses im Vergleich

Act-out or picture selection?

Comparing two methods for the assessment of grammatical comprehension

**Schlüsselwörter:** Testmethoden, Ausagieren, Satz-Bild-Zuordnung, Grammatikentwicklung, Satzverständnis

**Keywords:** test methods, act-out, picture selection, grammatical development, sentence comprehension

**Zusammenfassung:** Sowohl in der Spracherwerbsforschung als auch in der Diagnostik von Sprachentwicklungsstörungen spielt die Untersuchung des Grammatikverständnisses eine wichtige Rolle. Zwei dazu häufig eingesetzte Methoden sind das Ausagieren (Aag) und die Satz-Bild-Zuordnung (SBZ). Sie stellen unterschiedliche Anforderungen an das Kind und sind jeweils mit spezifischen Vor- und Nachteilen bezüglich der Testkonstruktion, -durchführung und -auswertung verbunden. Hieraus ergibt sich die Frage, inwiefern die Wahl der Methode das Testergebnis beeinflusst bzw. ob eine der Methoden zur Erfassung von Sprachverständnisleistungen besser geeignet ist als die andere, auch im Hinblick auf das Alter der zu untersuchenden Kinder.

In der vorliegenden Studie wurden die Aag- und die SBZ-Methode anhand einer Satzverständnisaufgabe miteinander verglichen. ProbandInnen waren 48 monolingual deutschsprachige, entwicklungsunauffällige Kinder in zwei Altersgruppen (3;0 bis 3;5 und 5;0 bis 5;5 Jahre).

Die Ergebnisse zeigten keine statistisch signifikanten Auswirkungen der Methode auf die Testleistung, d. h. die Unterschiede zwischen den Verfahren hatten keinen wesentlichen Einfluss auf die Testergebnisse der Kinder. Zudem erwiesen sich beide Methoden als entwicklungssensitiv, da sich mit beiden altersbedingte Verbesserungen im Satzverstehen aufzeigen ließen. Bezogen auf die Dreijährigen ergaben sich bei genauerer Betrachtung der Ergebnisse in Zusammenhang mit Verhaltensbeobachtungen jedoch Hinweise auf eine höhere Zuverlässigkeit der Aag-Methode, während die SBZ-Methode hier häufig zu Rateraktionen führte. Bei den Fünfjährigen erwiesen sich beide Methoden auch in dieser Hinsicht als gleichwertig. Für junge Kinder könnte die Aag-Methode also im Hinblick auf die Verlässlichkeit der Ergebnisse geeigneter sein.

Die Ergebnisse der Studie bieten einen Ausgangspunkt für weitere Untersuchungen mit sprachentwicklungsgestörten Kindern.

**Abstract:** The assessment of children's grammatical comprehension plays an essential role in both, experimental research on language acquisition and the assessment of developmental language disorders. Two prevalent methods used for this purpose are the act-out task and the picture selection task. They differ in their requirements on the child and are characterised by specific advantages and disadvantages concerning the construction, administration, and interpretation of the tests in which they are applied.

This leads to the question as to how far the adopted method influences children's performance, and accordingly, whether one of the methods is better suited to assessing sentence comprehension than the other. An additional question is whether task-specific effects change with age.

In the study reported here, act-out and picture selection were, by means of a sentence comprehension task, compared to each other. Subjects were 48 monolingual German-speaking children with normal language development, divided into two age groups (3;0 to 3;5 and 5;0 to 5;5 years).

The results showed no statistically significant effects of the method on test performance, i. e. differences between the procedures had no considerable impact on the children's test results. In addition, both methods proved to be age-sensitive, since age-related improvement of sentence comprehension abilities could be demonstrated with both methods.

Regarding the three-year-olds, however, a detailed analysis of the results and behavioural observations during the test indicated a higher reliability of the act-out method, whereas picture selection seemed to increase the risk of guessing behaviour in this age group. With the five-year-olds, both methods proved to be equally reliable. A tentative conclusion is that the act-out method might lead to more stable results in young children.

The results of the study offer a starting point for further research in language impaired children.

## Einleitung

Die Erfassung kindlicher Sprachverständnisseleistungen ist ein wesentlicher Bestandteil sowohl der Spracherwerbsforschung als auch der klinischen Diagnostik von Sprachentwicklungsstörungen. In beiden Bereichen kommt den rezeptiven Fähigkeiten ein wichtiger Stellenwert bei der Beurteilung der Sprachkompetenz eines Kindes zu.

Im ungestörten Spracherwerb interagieren Sprachverständnis und Sprachproduktion in ihrer Entwicklung zwar, entwickeln sich jedoch keineswegs immer völlig synchron (Miller & Paul, 1995). So zeigen die Ergebnisse verschiedener Studien, dass v. a. in der Grammatikentwicklung nach dem dritten Lebensjahr Asymmetrien zwischen rezeptiven und produktiven Fähigkeiten auftreten können. Nicht selten verwenden Kinder in diesem Alter bestimmte grammatische Strukturen bereits, bevor sie diese vollständig zielsprachgemäß verstehen können (z. B. Chapman & Miller, 1975; Chipman & Dannenbauer, 1988; Clark, 1982; Schrey-Dern, 2006; Schulz & Grimm, 2012). Daher kann von der Sprachproduktion eines Kindes nicht ohne Weiteres auf dessen Sprachverständnis geschlossen werden. Die gezielte Erfassung rezeptiver Sprachleistungen ist also unabdingbar, um die Mechanismen des kindlichen Spracherwerbs umfassend verstehen zu können (Elben & Lohaus, 2000; Grimm, Aktaş, & Frevert, 2001).

Auch bei Störungen des Spracherwerbs spielt die rezeptive Modalität eine nicht zu unterschätzende Rolle. In Bezug auf

die Spezifische Sprachentwicklungsstörung (SSES) ist belegt, dass Sprachverständnisdefizite, v. a. im Bereich der Grammatik, zum Erscheinungsbild der Störung dazugehören (z. B. Bishop, 1997; de Langen-Müller, Kauschke, Kiese-Himmel, Neumann, & Noterdaeme, 2012; Gebhard, Dames, & Baur, 1994; Gebhard, 2008; Schlesiger, 2001; Schulz & Grimm, 2012). Zudem gilt besonders bei jungen Kindern mit Sprachentwicklungsverzögerungen (Late Talkern) das Ausmaß der rezeptiven Beeinträchtigungen als ein wichtiger prognostischer Faktor für die weitere Sprachentwicklung: Je ausgeprägter die Sprachverständnisdefizite, desto schlechter die Prognose (Buschmann & Neubauer, 2012; Sachse & v. Suchodoletz, 2009; Schmitz & Fox, 2007). Eine verlässliche Einschätzung der Sprachverständnissfähigkeiten – ggf. mit therapeutischen Konsequenzen – ist zudem Voraussetzung für den Schulerfolg eines Kindes, der wesentlich vom Verstehen verbal vermittelter Lerninhalte und Anweisungen abhängt (Elben & Lohaus, 2000; Mathieu, 2008). Eine umfassende Diagnostik von Sprachentwicklungsstörungen sollte daher immer auch eine Beurteilung der rezeptiven Sprachkompetenz beinhalten (Schrey-Dern, 2006; de Langen-Müller et al., 2012).

Tatsächlich ist die Erfassung von Sprachverständnisseleistungen jedoch nicht ganz unproblematisch, da die rezeptive Sprachverarbeitung nicht direkt beobachtet, sondern nur indirekt aus den Reaktionen des Kindes erschlossen werden kann (Gebhard, 2008). Aufgrund der vielfältigen Faktoren, die diese Reaktionen beeinflussen, bspw. Motivation und Aufmerksamkeit des Kindes, aber auch der verbale und nonverbale Untersuchungskontext, können freie Gesprächs- oder Spielsituationen hier leicht zu Fehleinschätzungen führen. Formelle Testverfahren dagegen erlauben eine stärkere Kontrolle vieler dieser Faktoren, insbesondere des Kontexteinflusses, und damit eine gezieltere Erfassung des Sprachverständnisses im Sinne des linguistischen Dekodierens (Endres & Baur, 2000; Kauschke & Siegmüller, 2009; Schlesiger, 2001). Umgesetzt wird dies meist mit Hilfe von Ausagierungs- oder Satz-Bild-Zuordnungsaufgaben, wobei jede Methode spezifische Anforderungen

an das Kind stellt und bestimmte Vor- und Nachteile hinsichtlich Testkonstruktion, -durchführung und -auswertung aufweist (Goodluck, 1996; Gerken & Shady, 1996). Daraus ergibt sich die Frage, ob die Wahl der Methode das Untersuchungsergebnis beeinflusst – evtl. auch in Abhängigkeit vom Alter der zu untersuchenden Kinder. In der vorliegenden Studie wird daher ein direkter Vergleich der beiden Methoden mit parallelisiertem Material vorgenommen, wobei sprachunauffällige Kinder zweier Altersgruppen untersucht werden (siehe auch Watermeyer, 2010; Watermeyer, Höhle, & Kauschke, 2011). Im Folgenden werden die Methoden des Ausagierens und der Satz-Bild-Zuordnung zunächst beschrieben und mit Blick auf ihre Vor- und Nachteile einander gegenübergestellt. Die Ergebnisse bisheriger Vergleichsstudien werden berichtet. Daraus werden die Fragestellungen und Hypothesen der Studie abgeleitet. Nach der Präsentation und Diskussion der Ergebnisse wird ein Ausblick auf Fragestellungen für die zukünftige Forschung gegeben.

## Darstellung und Vergleich der Methoden

### Ausagieren

Beim Ausagieren (Aag) besteht die Aufgabe des Kindes darin, die Inhalte auditiv präsentierter Stimuli mithilfe von Spielzeugfiguren und/oder -gegenständen nachzuspielen.

Die Erfassung des Grammatikverständnisses mit der Aag-Methode setzt voraus, dass sich die Inhalte der Stimulussätze figürlich darstellen lassen, was Sätze über mentale Zustände oder Gefühle eher ausschließt (Goodluck, 1996).

In der Spracherwerbsforschung wird die Aag-Methode seit den 1970er Jahren v. a. in Studien zur Syntaxentwicklung eingesetzt. Untersuchungsgegenstände sind z. B.:

- das Verständnis komplexer Sätze (Clark, 1971; Grimm & Schöler, 1975; Schöler, 1975; Stevenson & Pollit, 1987)
- das Verständnis von Relativsatzstrukturen (Fragman, Goodluck, & Heggie, 2007; Grimm & Wintermantel, 1975b; Sheldon, 1974)
- das Verständnis von Nullelementen/

## KURZBIOGRAFIE

**Melanie Watermeyer** ist Logopädin und Patholinguistin. Nach dem Abschluss ihrer Logopädieausbildung 2002 in Göttingen arbeitete sie in verschiedenen Praxen in Hamburg, Berlin und Potsdam. Berufsbegleitend absolvierte sie bis 2011 den Diplom-Studiengang Patholinguistik an der Universität Potsdam. Ihre Arbeitsschwerpunkte sind Sprachentwicklungsstörungen und Störungen des Schriftspracherwerbs.

- Pronomina (Chomsky, 1969; Goodluck & Behne, 1992; Hsu, Smith Cairns, Eisenberg, & Schlüsselberg, 1991)
- das Verständnis von Passivkonstruktionen (de Villiers & de Villiers, 1973; Grimm, 1975)
  - das Verständnis von Verb-Argument-Strukturen (van der Lely, 1994)
  - das Verständnis ambiger Sätze (Kidd & Bavin, 2005).

Im Bereich der klinischen Diagnostik kommt die Aag-Methode sowohl in standardisierten und normierten Tests als auch in informellen Verfahren zur Untersuchung des Grammatikverständnisses zum Einsatz. Tabelle 1 gibt einen Überblick über Diagnostikverfahren aus dem deutschen Sprachraum, die – gänzlich oder in einzelnen Subtests – mit der Aag-Methode arbeiten.

Zu einem Mindestalter, ab dem die Aag-Methode sinnvoll eingesetzt werden kann, finden sich in der Fachliteratur leicht variierende Angaben: So halten Miller und Paul (1995) Aag-Tests schon bei Kindern ab zwei Jahren für einsetzbar. Tatsächlich werden in einigen Spracherwerbsstudien schon zweijährigen oder noch jüngeren Kindern Aag-Aufgaben präsentiert (z. B. de Villier & de Villiers, 1973; Thal & Flores, 2001). Goodluck (1996) dagegen plädiert aufgrund der kognitiven Komplexität der Aufgabe (Wahrnehmung, Speicherung und Dekodierung des Testsatzes, Aufbau eines Bezuges zu den verfügbaren Materialien, Umsetzen in eine Handlung und schließlich deren Ausführung) für ein Mindestalter von drei Jahren. Die meisten diagnostischen Verfahren sind dementsprechend für Kinder ab einem Alter von drei oder auch vier Jahren ausgelegt (vgl. Tab. 1).

Durch die vergleichsweise hohen kognitiven Anforderungen ist die Menge der am Stück zu bewältigenden Items in Aag-Untersuchungen begrenzt – laut Goodluck (1996) auf ca. 20 bis 30 für Kinder ab vier Jahren. In Experimenten wird diese Anzahl teilweise durchaus überschritten (z. B. Hsu et al., 1991; Stevenson & Pollit, 1987). Die diagnostischen Verfahren dagegen sehen meist weniger Items vor, insbesondere, wenn sie zu einem Gesamtverfahren gehören, das mehrere Untertests enthält (vgl. Tab. 1).

Als Materialien für Aag-Verfahren kom-

men je nach Untersuchungsgegenstand, Zielgruppe und Itemkonstruktion verschiedene handelsübliche Spielzeugfiguren und -gegenstände in Frage. Wichtig ist in jedem Fall, dass die Materialien pragmatisch neutral sind, d. h. mit ihnen verschiedene thematische Rollen besetzt bzw. Handlungen ausgeführt werden können (Goodluck, 1996). Das in Experimenten verwendete Material reicht von lediglich zwei Schachteln (Grieve, Hoogenraad, & Murray, 1977) über vier Mitglieder einer Puppenfamilie (Goodluck, Terzi, & Díaz, 2001) bis hin zu 16 verschiedenen Tierfiguren (Miolo, Chapman, & Sindberg, 2005). In den diagnostischen Tests kommt meist eine Mischung aus Spielzeugfiguren und -gegenständen zum Einsatz, wobei letztere als Objekte oder Instrumente von auszuführenden Handlungen, teilweise aber auch zum Aufbau einer Art Szenerie dienen (vgl. Tab. 1).

Die Testdurchführung wird je nach Studienegegenstand bzw. Testverfahren sowie Alter der ProbandInnen/PatientInnen etwas unterschiedlich gehandhabt: Meist ist das Kind angehalten, die Testsätze so auszuagieren, dass jeweils eine der Spielfiguren die Rolle des Agens übernimmt. Einige Experimente bzw. Tests sind aber auch so angelegt, dass das Kind selbst der/die Handelnde ist (z. B. Stevenson & Pollit, 1987; Chien & Wexler, 1990; vgl. Tab. 1).

Eine weitere Variationsmöglichkeit hinsichtlich der Durchführung besteht in der Anzahl der Figuren und/oder Gegenstände, die dem Kind pro Item zur Verfügung stehen: In einigen Experimenten (z. B. Naigles & Lehrer, 2002; Miolo et al., 2005) sowie in vielen diagnostischen Verfahren (vgl. Tab. 1) wird für alle Items dieselbe Menge von Materialien bereitgestellt, so dass das Kind für jedes Item zunächst die benötigten Figuren und/oder Gegenstände auswählen muss. Dies setzt eine lexikalische Verarbeitungsleistung voraus. In anderen Experimenten (z. B. Thal & Flores, 2001; Kidd & Bavin, 2005) und vereinzelt auch in diagnostischen Verfahren (vgl. Tab. 1) werden dem Kind jeweils nur die Materialien zur Verfügung gestellt, die es zum Ausagieren des betreffenden Items braucht, so dass vom Kind keine lexikalische Auswahl mehr zu treffen ist. In jedem Fall ist bei der Platzierung

der Materialien eine möglichst gleichwertige räumliche Relation zum Kind wichtig, um eine Bevorzugung aufgrund besserer Erreichbarkeit zu vermeiden. Abgesehen von diesem Risiko verleitet die Aag-Methode mit ihrem freien Antwortformat kaum zu situations-, material- oder untersucherbedingten Reaktionen, die das Untersuchungsergebnis verfälschen können. Zudem ist der spielerische Charakter dieser Methode für die meisten Kinder sehr attraktiv und daher fast ein Garant für eine hohe kindliche Motivation (Grimm & Wintermantel, 1975a). Für den Untersucher/die Untersucherin ist die Durchführung allerdings recht aufwendig.

Die Bewertung der kindlichen Reaktionen bei der Aag-Methode gilt aufgrund der freien Antwortmöglichkeiten als nicht ganz einfach (Goodluck, 1996; Schmitz & Fox, 2007; Lienert & Raatz, 1998). So kann z. B. durch eine ungenaue kindliche Darstellung der Handlungen oder durch nicht eindeutig erkennbare Vergabe der thematischen Rollen die Einschätzung der Korrektheit der Reaktionen erschwert werden. Auch Nullreaktionen oder nur partielles Ausagieren von Handlungen können in dieser Hinsicht problematisch sein (Mathieu, 2008). Durch eine sorgfältige Untersuchungskonstruktion und genau festgelegte Auswertungskriterien können diese Probleme jedoch weitestgehend vermieden werden (Kubinger, 2009). Die Auswertung der Ergebnisse erfolgt durch die Summierung der korrekten Reaktionen, die daraufhin statistisch weiterverarbeitet oder mit Normdaten abgeglichen werden können. Außerdem kann eine Analyse der Reaktionsmuster im Hinblick auf Erwerbslevel, Sprachverständnisstrategien (Miller & Paul, 1995; Schlesiger, 2001) oder Fehlerarten erfolgen. Das freie Antwortformat in Aag-Verfahren, das u. U. zu vielfältigen individuellen Fehlreaktionen führt, kann diese Reaktionsmusteranalyse zwar einerseits erschweren, andererseits aber auch aufschlussreiche Einblicke in die kindliche Grammatikverarbeitung gewähren (Goodluck, 1996).

### Satz-Bild-Zuordnung

Bei der Satz-Bild-Zuordnung (SBZ) besteht die Aufgabe des Kindes darin, aus einer Auswahlmenge von Bildern das

Test (-Subtest)	Gegenstand	Alter	Items	Material	Durchführung	Auswertung
HSET-VS (Grimm & Schöler, 1978)	Verständnis verschiedener Satzstrukturen, u. a. mit - Passiv - kausativierenden Verben - Temporalsätzen - Relativsätzen	3–9 J.	10–17	4 Puppen 10 Tierfiguren 1 Holzklotz 1 Waschlappen	<u>Agens</u> : Figur <u>Material</u> : Vorauswahl jeweils durch UntersucherInnen, Anordnung standardisiert	Abgleich mit Normdaten
PSST-A (Wettstein, 1997)	Verständnis verschieden langer Sätze, u. a. mit - Präpositionalphrasen - Kausalsätzen - Temporalsätzen	4–9 J.	16	4 Puppen 2 Tierfiguren 1 kleiner Ball + <u>Szenerie</u> : Haus, Turm und Garten	<u>Agens</u> : Figur <u>Material</u> : für alle Items Gesamtmenge verfügbar, Anordnung in etwa vorgegeben	Abgleich mit Normdaten, Reaktionsmusteranalyse (Satzstrukturen, Sprachverständnis-Strategien)
ESGRAF-Anteile (Motsch, 2000)	Verständnis komplexer Satzstrukturen, u. a. mit - Temporalsätzen - Kausalsätzen - Finalsätzen - Relativsätzen - indirekten Fragen	4–10 J.	bis 10	7 Puppen 1 Puppenstuhl 1 Flasche 2 Körbe 2 Einkaufszettel	<u>Agens</u> : teils Kind, teils Figur <u>Material</u> : für alle Items Gesamtmenge verfügbar, Anordnung in etwa vorgegeben	Reaktionsmusteranalyse (Satzstrukturen)
IVÜS (Endres & Baur, 2000)	Verständnis verschieden langer Sätze, u. a. mit - Negation - Temporalsätzen - Kausalsätzen - Konditionalsätzen	4–8 J.	45–70	4 Puppen 4 Tierfiguren 1 kleiner Ball 1 Auto + <u>Szenerie</u> : Haus, Turm und Garten	<u>Agens</u> : Figur <u>Material</u> : für einige Items Gesamtmenge verfügbar, für einige Vorauswahl durch UntersucherInnen, Anordnung vorgegeben	Reaktionsmusteranalyse (Satzstrukturen, SV-Strategien)
SETK3–5-VS (Grimm et al., 2001)	Verständnis verschieden langer Sätze, u. a. mit - Temporalsätzen - Kausalsätzen - Relativsätzen - Präpositionalphrasen	3–5 J.	10–15	2 Malstifte 1 Stoffsack 4 Knöpfe 1 Kiste 2 kleine Bälle 1 Bilderbuch 1 Teddy	<u>Agens</u> : Kind (3-Jährige), Figur (4–5-Jährige) <u>Material</u> : Vorauswahl jeweils durch Untersucher, Anordnung standardisiert	Abgleich mit Normdaten
PDSS-16 (Kauschke & Siegmüller, 2009)	Verständnis verschiedener Satzstrukturen, u. a. mit - Objektvoranstellung - Passiv - Präpositionalphrasen - Temporalsätzen - Relativsätzen	2–6 J.	12	2 Puppen 2 Tierfiguren 2 Autos 1 Haus 1 Turm 1 Baum	<u>Agens</u> : Kind (die 3 ersten Items), Figur (Rest) <u>Material</u> : für alle Items Gesamtmenge verfügbar, Anordnung beliebig	Reaktionsmusteranalyse (Satzstrukturen)

Tabelle 1 Deutschsprachige Diagnostikverfahren, die mit der Aag-Methode arbeiten

zum auditiv präsentierten Stimulus passende herauszufinden und zu zeigen.

Die SBZ-Methode wird in erster Linie zur Untersuchung des Grammatikverständnisses eingesetzt, eignet sich aber auch zur Beurteilung des Wortverstehens – jeweils vorausgesetzt, die Inhalte der zu untersuchenden linguistischen Strukturen lassen sich genügend gut bildlich darstellen, was z. B. bei Handlungen aufgrund der darzustellenden Dynamik problematisch sein kann (Gerken & Shady, 1996).

In der Spracherwerbsforschung findet die SBZ-Methode breite Anwendung in vielfältigen Studien zum Syntaxerwerb, die unterschiedliche linguistische Strukturen untersuchen, z. B.:

- das Verständnis von Relativsatzstruktu-

ren (Friedmann & Novogrodsky, 2004; Hakansson & Hansson, 2000)

- das Verständnis von Passivkonstruktionen (van der Lely, 1996)

- das Verständnis von Verbzeitformen (Weist, Wysocka, & Lyytinen, 1991)

- das Verständnis ambiger Sätze (Kidd & Bavin, 2005)

- das Verständnis von Pronominalkonstruktionen (Oshima-Takane, 1992)

- den Zusammenhang von Arbeitsgedächtniskapazität und Satzverständniskapazitäten (Montgomery, 2000; Montgomery & Evans, 2009)

- den Einfluss von Grammatikalität und Prosodie auf das Verstehen von Funktionswörtern (Gerken & McIntosh, 1993).

Im klinisch-diagnostischen Rahmen wird die SBZ-Methode in vielen standardisierten und normierten Sprachver-

ständnistests eingesetzt. In Tabelle 2 sind Diagnostikverfahren aus dem deutschen Sprachraum zusammengestellt, die – gänzlich oder in einzelnen Subtests – die SBZ-Methode verwenden.

Über eine Altersgrenze, ab der die SBZ-Methode anwendbar ist, gehen die Meinungen recht weit auseinander: So mutmaßen Gerken und Shady (1996), dass die SBZ-Methode für schüchterne oder zum Ausagieren noch zu unreife Kinder besonders gut geeignet sein könne, da diese weniger Eigeninitiative vom Kind verlange. In der Tat sind die ProbandInnen einiger Studien, die mit der SBZ-Methode arbeiten z. T. jünger als zwei Jahre (z. B. Oshima-Takane, 1992; Gerken & McIntosh, 1993). Nach Einschätzung von Mathieu (2008) dagegen

Test (-Subtest)	Gegenstand	Alter	Items	Material	Durchführung	Auswertung
<b>IDIS-MM</b> (Schöler, Schakib-Ekbatan, Spohn, & Spohn, 1999)	Verständnis verschiedener Satzstrukturen, u. a. mit - Perfekt - Passiv - Objektvoranstellung - Pronomina	4–6 J.	8	Zielbild + 3 Ablenker	Reihenfolge der Präsentation: Bilder > Stimulus	Abgleich mit Risikowerten
<b>MSVK</b> (Elben & Lohaus, 2000) <sup>1</sup>	Wortverständnis - passiver Wortschatz - Ober-/Unterbegriffe Verständnis verschiedener Satzstrukturen, u. a. mit - Plural - Perfekt - Passiv Pragmatikverständnis - personenbezogene - situationsbezogene Sprechakte	5–7 J.	72	Zielbild + 2-4 Ablenker	Reihenfolge der Präsentation: Bilder > Stimulus	Abgleich mit Normdaten, Reaktionsmusteranalyse (Satzstrukturen, Fehlerarten)
<b>SETK2-V I-II</b> (Grimm, Aktaş, & Frevert, 2000)	Wortverständnis - passiver Wortschatz Verständnis verschiedener Satzstrukturen, u. a. mit - Negation - Präpositionalphrasen	2 J.	17	Zielbild + 3 Ablenker	Reihenfolge der Präsentation: Bilder > Stimulus	Abgleich mit Normdaten
<b>TROG-D</b> (Fox, 2006)	Wortverständnis - passiver Wortschatz - Ober-/Unterbegriffe Verständnis verschiedener Satzstrukturen, u. a. mit - Negation - Präpositionalphrasen - Perfekt - Plural - Passiv - Pronomina - Temporalsätzen - Relativsätzen - Doppelobjektsätzen - Objektvoranstellung	3–10 J.	72–84	Zielbild + 3 Ablenker	Reihenfolge der Präsentation: Bilder > Stimulus	Abgleich mit Normdaten (Blockwertung, d. h. Punktvergabe nur bei vollständig korrekt gelöstem Block von 4 Items pro Struktur), Reaktionsmusteranalyse (Satzstrukturen, Fehlerarten)
<b>TSVK</b> (Sieg Müller, Kauschke, Bartke, & Bittner, 2011) <sup>2</sup>	Verständnis verschiedener Satzstrukturen, u. a. mit - Perfekt - Objektvoranstellung - Passiv - Reflexivbindungen - Objektbindungen - Relativsätzen	2–8 J.	12–90	Zielbild + 2 Ablenker	Reihenfolge der Präsentation: Bilder > Stimulus	Abgleich mit Normdaten, Reaktionsmusteranalyse (Satzstrukturen, Fehlerarten)

Tabelle 2 Deutschsprachige Diagnostikverfahren, die mit der SBZ-Methode arbeiten

<sup>1</sup> Der MSVK beinhaltet auch einen Untertest zum Instruktionsverständnis, der jedoch keine SBZ-Aufgaben enthält und deshalb hier nicht mit aufgeführt ist. Entsprechend wurden die Items für diesen Untertest nicht mitgezählt.

<sup>2</sup> Der TSVK enthält neben der hier beschriebenen Langversion auch noch eine Kurzversion (TSVK Screen), der aus Platzgründen hier nicht mit aufgeführt werden konnte.

verfügen Kinder erst ab vier Jahren über die notwendige Abstraktionsfähigkeit für die ausschließliche Arbeit mit Bildmaterial, wie die SBZ-Methode sie erfordert (Wahrnehmung, Speicherung und Dekodierung des Testsatzes, Verarbeitung der Auswahlbilder und Abgleich mit dem Inhalt des Testsatzes, schließlich Auswahl des passenden Bildes). Schlesiger (2001) setzt das Mindestalter für SBZ-Aufgaben sogar noch höher an, da diese sich auch bei drei- bis vierjährigen Kindern als wenig reliabel erwiesen hätten (S. 91). In der klinischen Praxis werden SBZ-Aufgaben z. T. schon mit Zweijährigen durchgeführt, die Mehrzahl der Tests ist

jedoch für etwas ältere Kinder konzipiert (vgl. Tab. 2).

Die Anzahl der Items, die mit der SBZ-Methode am Stück getestet werden, ist oft relativ hoch: In Experimenten werden vierjährigen Kindern meist 30 bis 40 Items oder mehr präsentiert (z. B. Bishop, 1982; Veit, 1986; Montgomery, 2000). Auch die diagnostischen Verfahren, und dabei insbesondere die reinen Sprachverständnistests, enthalten meist eine große Anzahl von Items (vgl. Tab. 2). Das Material für SBZ-Verfahren muss sorgfältig konstruiert werden, um das bei dieser Methode bestehende Raterisiko (siehe Exkurs zur Ratewahrscheinlich-

keit) möglichst gering zu halten. Neben einem dem Stimulussatz entsprechenden Zielbild werden ein oder mehrere Ablenker benötigt, die einen bestimmten linguistischen Kontrast zum Stimulussatz abbilden und dabei untereinander möglichst gleichwertig sein, d. h. für die TestkandidatInnen ähnlich plausibel erscheinen sollten (vgl. Lienert & Raatz, 1998).

Bei der Testdurchführung sind in Bezug auf die Abfolge der Präsentation von Bildern und Stimulus unterschiedliche Vorgehensweisen möglich: Meistens werden dem Kind zuerst die Bilder und anschließend der Stimulussatz präsentiert – so

auch in sämtlichen deutschsprachigen Sprachverständnistests (vgl. Tab. 2). Eine Alternative stellt das umgekehrte Vorgehen dar, also zuerst den Stimulus und danach die Bilder zu präsentieren (z. B. Veit, 1986; Montgomery, 2000). In jedem Fall muss während der Bild- und Stimuluspräsentation darauf geachtet werden, dass das Kind nicht durch das Verhalten des Untersuchers/der Untersucherin, insbesondere durch evtl. unbewusstes Blicken in Richtung des Zielbildes, in seiner Reaktion beeinflusst wird (Schmitz & Fox, 2007). Auch für andere Einflüsse auf die kindliche Reaktion ist die SBZ-Methode anfällig: Zum einen besteht aufgrund des Multiple-Choice-Formats immer ein gewisses Risiko, dass das Kind, statt gezielt zu reagieren, lediglich rät (siehe Exkurs zur Ratewahrscheinlichkeit). Zum anderen sind auch nicht-linguistisch determinierte Antwortneigungen, z. B. zu einer bestimmten Bildposition oder zu dem Bild mit den interessantesten Details denkbar. Für den Untersucher/die Untersucherin ist die Durchführung von SBZ-Tests dennoch einfach, kaum aufwendig und daher außerordentlich praktikabel (Gebhard, 2008).

Die Bewertung der kindlichen Reaktionen in SBZ-Verfahren ist unkompliziert, da lediglich die Zeigegeste des Kindes interpretiert werden und daraufhin die Summe korrekter Reaktionen ermittelt werden muss. Schwierigkeiten kann allenfalls der Umgang mit Nullreaktionen oder mit dem Zeigen mehrerer Bilder zum selben Stimulus bereiten – dieser sollte daher bei der Planung des Experiments bzw. vor Einsatz des Diagnostikverfahrens genau festgelegt werden (Gerken & Shady, 1996).

Die Weiterverarbeitung der Ergebnisse kann wiederum mittels statistischer Verfahren bzw. eines Normdatenabgleichs und/oder durch eine Reaktionsmusteranalyse erfolgen. Die Analyse der Reaktionsmuster hinsichtlich des Erwerbslevels der getesteten Strukturen, kindlicher Sprachverständnisstrategien oder Fehlerarten ist in SBZ-Tests bei entsprechender Konstruktion der Ablenker meist unproblematisch und eindeutig möglich, ihr diagnostischer Aufschlussreichtum bleibt allerdings auf die durch die Ablenker vorgegebenen Fehlermöglichkeiten beschränkt.

## Exkurs zur Ratewahrscheinlichkeit

Die Ratewahrscheinlichkeit bezieht sich auf die Wahrscheinlichkeit, allein durch eine zufallsgeleitete Reaktion zur richtigen Lösung zu kommen. So könnte ein Kind bei SBZ-Aufgaben einfach auf irgendein beliebiges Bild zeigen. Nur durch die Verwendung von sorgfältig konstruierten (untereinander gleichwertigen, d. h. ähnlich plausiblen) Ablenkerbildern ist die Ratewahrscheinlichkeit pro Item überhaupt kalkulierbar. Sie beträgt bei einem Ablenker zusätzlich zum Zielbild, also insgesamt zwei gleichzeitig präsentierten Bildern, 50 Prozent, bei zwei Ablenkern zusätzlich zum Zielbild, also drei gleichzeitig präsentierten Bildern, rund 33 Prozent, bei drei Ablenkern zusätzlich zum Zielbild, also vier gleichzeitig präsentierten Bildern, 25 Prozent usw. (vgl. Abb. 1).

Um dem Problem des bestehenden Raterisikos zu begegnen, muss für jede zu untersuchende linguistische Struktur eine ausreichende Menge von Items verwendet werden. Daraufhin kann die Mindestanzahl korrekter Reaktionen auf diese Items ermittelt werden, ab der davon ausgegangen werden kann, dass das Kind nicht geraten, sondern die Stimuli tatsächlich verarbeitet und bewusst korrekt reagiert hat. Diese Mindestanzahl korrekter Reaktionen lässt sich mittels der Binomialverteilung berechnen, wie sie in der Testpsychologie üblicherweise für Multiple-Choice-Verfahren verwendet wird. Die Formel lautet:

$P(k) = (n^k) \cdot (p)^k \cdot (1-p)^{n-k}$ , wobei  $n^k$  (sprich „n über k“) =  $(n!) : (k! \cdot (n-k)!)$  ist

Die maßgeblichen Variablen sind hierbei die Anzahl der Items (n) und die Ratewahrscheinlichkeit pro Item (p). Daraus lässt sich ableiten, wie viele Items mindestens korrekt gelöst werden müssen (k), damit die Gesamtratewahrscheinlichkeit (P) für das Experiment bzw. den Test ein akzeptables Niveau erreicht. Analog zur Wahrscheinlichkeitsgrenze, an der bei statistischen Signifikanztests die Nullhypothese zurückgewiesen wird, sollte dieses bei unter fünf Prozent liegen (Bortz & Lienert, 2003). Tabelle 3 zeigt die Interaktion zwischen Itemzahl, Ratewahrscheinlichkeit pro Item und Gesamtratewahrscheinlichkeit anhand von Beispielwerten.

In den meisten SBZ-Untersuchungen spielen Überlegungen zur Ratewahrscheinlichkeit erstaunlicherweise keine Rolle. Nur in wenigen Spracherwerbsstudien wird als Auswertungskriterium eine Mindestmenge korrekter Reaktionen auf Basis der Binomialverteilung angegeben (z. B. van der Lely, 1990; Hakansson & Hansson, 2000) oder zumindest zur Ratewahrscheinlichkeit pro Item Stellung bezogen (z. B. Fraser, Bellugi, & Brown, 1963). Auch in den meisten diagnostischen SBZ-Verfahren wird das Raterisiko für die verschiedenen Ergebnisrohre nicht berücksichtigt. In einigen Tests bzw. Subtests ist nicht sicher auszuschließen, dass das getestete Kind allein durch Raten zu einem unauffälligen Testergebnis

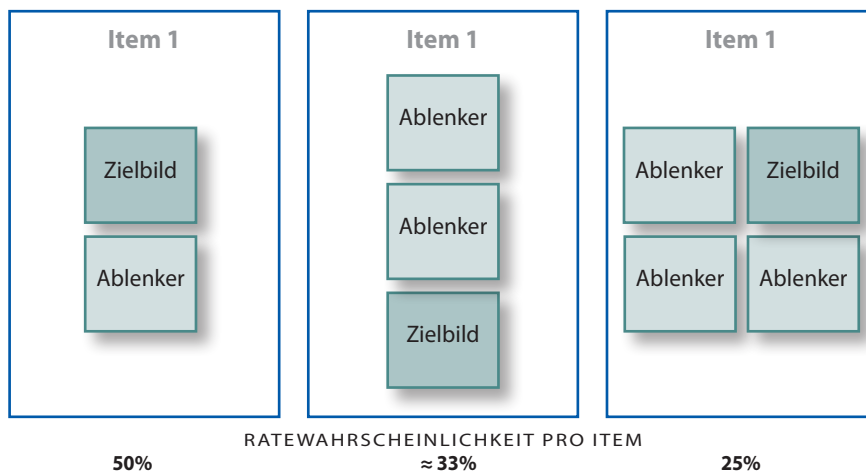


Abbildung 1 Ratewahrscheinlichkeit pro Item in Abhängigkeit von der Ablenkerzahl

kommt. Wenn jedoch auch mit einer Zufallsreaktion ein unauffälliger Testwert erzielt werden kann, d. h. wenn selbst Kinder aus der Normierungsstichprobe möglicherweise nur geraten haben, ist keine sichere Differenzierung zwischen unauffälligen und auffälligen Leistungen möglich. Für eine sinnvolle normdatenbezogene Auswertung der Ergebnisse von SBZ-Tests wäre es also wünschenswert, nicht nur die Ratewahrscheinlichkeit pro Item zu beachten, sondern auch die Ratewahrscheinlichkeit für die Anzahl korrekter Reaktionen zu ermitteln, ab der ein Kind als unauffällig eingestuft wird.

## Vergleich der Methoden

Wie aus den vorangegangenen Abschnitten hervorgeht, weisen die Aag- und die SBZ-Methode einige Ähnlichkeiten, aber auch große Unterschiede auf. Daraus ergeben sich spezifische Vor- und Nachteile, die für ihren Einsatz sowohl in der Spracherwerbsforschung als auch in der klinischen Diagnostik relevant sein können. Tabelle 4 fasst die Vor- und Nachteile in einer Gegenüberstellung zusammen. Zum direkten Vergleich des Ausagierens mit der Satz-Bild-Zuordnung finden sich in der Literatur lediglich einige wenige Studien: Cocking und McHale (1981) analysierten bild- und objektbasierte Untersuchungsmethoden und erhoben dabei u. a. die Sprachverständnisleistungen von Kindern im Alter von vier bis fünf Jahren anhand von 100 aus verschiedenen syntaktischen Strukturen bestehenden (englischen) Stimulussätzen. Je zwölf Kinder wurden entweder mit einer

Aag- oder einer SBZ-Aufgabe getestet – mit der Erwartung, dass das Ausagieren schwieriger sein würde: „In comprehension / pictures [entspricht SBZ], the child need only recognize a situation, whereas in comprehension / objects [entspricht Aag] the child had to construct it; one would expect the latter to be harder“ (S. 5). Tatsächlich erzielten die mit der Aag-Methode getesteten Kinder signifikant schlechtere Ergebnisse als die mit der SBZ-Methode untersuchten – allerdings nur, wenn sie die benötigten Figuren vor dem Ausagieren selbst aus einer Menge auswählen mussten. Die Aag-Methode mit Materialauswahl durch das Kind erwies sich für Vier- bis Fünfjährige also im Vergleich zur SBZ-Methode als schwieriger.

Da in dieser Studie unabhängige ProbandInnengruppen mit jeweils einer Methode konfrontiert wurden, könnten die Ergebnisse jedoch auch durch unterschiedliche Fähigkeitsniveaus der Kinder in den beiden relativ kleinen Gruppen bedingt sein.

Eine etwas anderes Vorgehen liegt der Studie von Gebhard (2008) zugrunde: Hier wurden die Leistungen einer Gruppe von 50 Vor- und Grundschulkindern im HSET-VS (Grimm & Schöler, 1978) und im TROG-D (Fox, 2006) miteinander verglichen, wobei ersterer mit der Aag-Methode arbeitet – und zwar in der Variante mit Vorauswahl des Materials durch den Untersucher/die Untersucherin (vgl. Tab. 1), während letzterer die SBZ-Methode einsetzt (vgl. Tab. 2). Es zeigte sich, dass „Kinder, die im HSET-VS [mit der Aag-Methode] den Normbereich verfehlten,

ihn im TROG-D [mit der SBZ-Methode] öfter noch erreichten“ (Gebhard, 2008, S. 159). Insgesamt ergaben sich signifikante Unterschiede zwischen den Leistungen der Kinder in den verschiedenen Tests zugunsten der SBZ-Methode, was nach Ansicht des Autors auf „die unterschiedliche Herangehensweise der beiden Verfahren (Bilder vs. Figuren) zurückzuführen ist“ (S. 160). Die Aag-Methode erwies sich also auch hier als schwieriger. Eine Vergleichbarkeit der Ergebnisse aus den beiden Tests ist allerdings nur bedingt gegeben, da die verwendeten Testverfahren sich sowohl bezüglich der Anzahl und des Aufbaus der enthaltenen Items als auch hinsichtlich der Bewertung der kindlichen Reaktionen stark unterscheiden (HSET-VS: zehn Items, TROG-D: 72 bis 84 Items; syntaktisch und lexikalisch unterschiedlich aufgebaute Stimulussätze).

Interessant ist schließlich auch die Studie von Dittmar, Abbot-Smith, Lieven und Tomasello (2008) zur Entwicklung des Verstehens von transitiven Aussagesätzen im Deutschen (prototypische SVO-Sätze, ambige SVO-Sätze und OVS-Sätze). Untersucht wurden die Verständnisleistungen von jeweils 16 zweijährigen und vier- bis fünfjährigen Kindern beim Ausagieren von sieben semantisch reversiblen Sätzen, davon sechs mit Kunstverben. Im Ergebnis zeigten sich überraschend späte Erwerbszeitpunkte für die getesteten Strukturen: Die Zweijährigen verstanden nur den einen Satz mit dem bekannten Verb korrekt, die Vier- bis Fünfjährigen darüber hinaus nur die SVO- Sätze mit Kunstverben (z. B. „Der Hund wief den

Items (n)	bei p = 50% (1 Ablenker)	bei p ≈ 33% (2 Ablenker)	bei p = 25% (3 Ablenker)
2	-	-	-
3	-	P (k = 3) < 5%	P (k = 3) < 5%
4	P (k = 4) < 5%	P (k = 4) < 5%	P (k = 4) < 5%
5	P (k = 5) < 5%	P (k ≥ 4) < 5%	P (k ≥ 4) < 5%
6	P (k = 6) < 5%	P (k ≥ 5) < 5%	P (k ≥ 4) < 5%
7	P (k = 7) < 5%	P (k ≥ 5) < 5%	P (k ≥ 5) < 5%
8	P (k ≥ 7) < 5%	P (k ≥ 6) < 5%	P (k ≥ 5) < 5%
9*	P (k ≥ 8) < 5%	P (k ≥ 6) < 5%	P (k ≥ 5) < 5%
10	P (k ≥ 9) < 5%	P (k ≥ 7) < 5%	P (k ≥ 6) < 5%
11	P (k ≥ 9) < 5%	P (k ≥ 7) < 5%	P (k ≥ 6) < 5%
12	P (k ≥ 10) < 5%	P (k ≥ 8) < 5%	P (k ≥ 7) < 5%

n = Anzahl der Items, p = Ratewahrscheinlichkeit pro Item,  
k = Anzahl korrekter Reaktionen,  
P(k) = Gesamtratewahrscheinlichkeit für den Test bei k korrekten Reaktionen

### \*Erläuterung:

Hat man bspw. neun Items (n = 9) mit je zwei Ablenkern (und damit p ≈ 33%), müssen mindestens sechs dieser Items korrekt gelöst werden (k ≥ 6), damit die Gesamtratewahrscheinlichkeit ein akzeptables Niveau erreicht (P < 5%). Bei derselben Anzahl von Items mit nur einem Ablenker (und damit p = 50%) müssten mindestens acht korrekte Antworten erzielt werden (k ≥ 8), um eine akzeptable Gesamtratewahrscheinlichkeit zu erreichen. Bei drei Ablenkern (also p = 25%) wäre dies dagegen schon bei mindestens fünf korrekt gelösten von insgesamt neun Items (k ≥ 5) der Fall. Ebenso verschiebt sich die Mindestanzahl korrekt zu lösender Items, wenn sich die Gesamtzahl der Items verändert.

Tabelle 3 Mindestanzahl korrekter Reaktionen für eine Gesamtratewahrscheinlichkeit von < 5%

Ausagieren (Aag)		Satz-Bild-Zuordnung (SBZ)
zur Untersuchung v. a. des Grammatikverständnisses (sofern ausagierbar) bei Kindern ab 2;0 bzw. 3;0 Jahren	Einsatzbereich	zur Untersuchung des Wort-/Grammatikverständnisses (sofern bildlich darstellbar) bei Kindern ab 2;0 bzw. 4;0 / 5;0 Jahren
👉 Materialbeschaffung einfach, günstig (handelsübliches Spielmaterial)	Testkonstruktion	👉 Materialherstellung aufwendig, evtl. teuer (mehrere Bilder zu jedem Item)
<ul style="list-style-type: none"> <li>👉 besonders motivierend für das Kind (da starker Spielcharakter)</li> <li>👉 am Stück testbare Items begrenzt (da anstrengend durch Eigeninitiative)</li> <li>👉 Rateverhalten / Antwortneigungen eher unwahrscheinlich (da freie Reaktionen)</li> <li>👉 relativ aufwendig für den/die UntersucherIn (durch wiederholte Materialanordnung)</li> </ul>	Durchführung	<ul style="list-style-type: none"> <li>👉 etwas weniger motivierend für das Kind (da eher Testcharakter)</li> <li>👉 besonders viele Items am Stück testbar (da Zeigen von Bildern kaum ermüdend)</li> <li>👉 anfällig für Rateverhalten / Antwortneigungen (da multiple-choice-Format)</li> <li>👉 sehr praktikabel für den/die UntersucherIn (da lediglich Umblättern von Bildseiten)</li> </ul>
<ul style="list-style-type: none"> <li>👉 Reaktionen u.U. schwierig zu bewerten (durch freie Reaktionsmöglichkeiten)</li> <li>👉 eher unverfälschte Ergebnisse (da Rate- / Antwortneigungsrisiko gering)</li> <li>👉 Reaktionsmuster u.U. schwierig zu interpretieren (da viele verschiedene Fehlerarten möglich)</li> <li>👉 u.U. diagnostisch sehr aufschlussreich (da individuelle, evtl. unerwartete Fehler)</li> </ul>	Auswertung	<ul style="list-style-type: none"> <li>👉 Reaktionen eindeutig zu bewerten (da nur Zeigegeste zu deuten)</li> <li>👉 Risiko der Ergebnisverfälschung (da evtl. Rate- / Antwortneigungseffekte)</li> <li>👉 Reaktionsmuster meist eindeutig interpretierbar (bei entsprechend konstruierten Ablenkern)</li> <li>👉 u.U. etwas weniger aufschlussreich (da nur vorgegebene Fehler möglich)</li> </ul>

Tabelle 4 Vorteile (👉) und Nachteile (👉) der Aag- und der SBZ-Methode

Löwen.“), nicht aber die OVS-Sätze (z. B. „Den Bären weist der Tiger.“). Die Autoren führten dies auf den hohen Schwierigkeitsgrad der Aag-Methode infolge ihrer Anforderungen an das Arbeitsgedächtnis und an die Exekutivfunktionen zurück (S. 1159). Das Experiment wurde daraufhin noch einmal mit der SBZ-Methode durchgeführt, diesmal mit neun Sätzen, die allesamt Kunstverben enthielten. Tatsächlich zeigten die Zweijährigen nun bessere Sprachverständnisleistungen. Sie erzielten deutliche höhere Korrekturraten bei den prototypischen SVO-Sätzen mit Kunstverben als mit der Aag-Methode. Bei den Vier- bis Fünfjährigen hingegen veränderte sich das Ergebnis nicht maßgeblich, d. h. sie hatten weiterhin Probleme mit den OVS-Sätzen. Die Autoren schlussfolgerten aus diesen Befunden, dass sehr junge Kinder ihre Sprachverständniskompetenz in SBZ-Aufgaben besser zeigen können als in Aag-Tests, was ihrer Ansicht nach v. a. der „memory-burdensome nature of the act-out method per se with small children“ geschuldet ist (S. 1163). Zusammenfassend bleibt festzuhalten, dass in allen drei beschriebenen Studien

Vorteile für die SBZ-Methode gefunden wurden, während die Aag-Methode für Kinder anspruchsvoller erschien und zu schlechteren Ergebnissen führte. Jedoch wurden die beiden Verfahren nur in einer der Studien mit dem gleichen Stimulusmaterial durchgeführt (Coking & McHale, 1981). In dieser Studie schnitten Kinder beim Ausagieren nur dann schlechter ab, wenn sie die Figuren selbst auswählen mussten. Der bisherige Forschungsstand zur Aussagekraft der beiden Methoden ist somit noch uneindeutig. Studien, die durch parallelisiertes Stimulusmaterial einen direkten Vergleich beider Verfahren erlauben, Altersaspekte berücksichtigen und die Ratewahrscheinlichkeit gering halten, stellen damit ein Desiderat dar.

### Empirische Studie Fragestellungen und Hypothesen

Um die Frage nach dem Einfluss der Methoden auf das Testergebnis und damit nach ihrer Eignung zur Erfassung des kindlichen Grammatikverständnisses unter Ausschluss der oben genannten

Einschränkungen zu klären, wurde eine Untersuchung konzipiert, die einen direkten Vergleich der beiden Aag- mit der SBZ-Methode ermöglicht – mit jeweils denselben ProbandInnen und denselben Items für beide methodischen Varianten des Experiments sowie einem ausreichend geringen Raterisiko bei den SBZ-Aufgaben. Zudem wurden zwei Altersgruppen einbezogen, um eine mögliche Abhängigkeit der Aussagekraft der Methoden vom Alter des Kindes zu ermitteln. Als Teststrukturen wurden semantisch reversible Sätze mit Objektvoranstellung verwendet. Diese Auswahl erfolgte zum einen aus Gründen der Umsetzbarkeit, zum anderen aufgrund der spracherwerbtheoretischen und diagnostischen Relevanz dieser Strukturen: Die Beherrschung objektinitialer Sätze gilt als Beleg für den vollständigen Erwerb der Verbzweitstellung, umgekehrt können Schwierigkeiten mit diesen Strukturen noch im späteren Vorschul- bzw. Grundschulalter als Hinweis auf eine syntaktische Störung verstanden werden (Clahsen, 1986; Penner, 1999; Siegmüller, 2003; siehe auch Watermeyer, 2010). Die Testsätze wurden drei- und fünfjährigen



Kategorie	Interpretationshinweise	Beispiel
I (leicht)	transparenter AKK, eindeutige SVK im Plural	Den Ritter kitzeln die Gespenster.
II (mittel)	transparenter AKK, nicht-eindeutige SVK	Den Drachen jagt der Zauberer.
III (schwer)	nicht-transparenter AKK, nicht-eindeutige SVK	Das Einhorn beißt der Drache.

Tabelle 6 Aufbau der OVS-Items mit Abstufung nach Schwierigkeitsgrad

AKK = Akkusativ, SVK = Subjekt-Verb-Kongruenz

Kindern jeweils mit beiden Methoden präsentiert und die erzielten Ergebnisse miteinander verglichen. Im Einzelnen ergaben sich dazu folgende Fragestellungen und Hypothesen:

1) Beeinflusst die Wahl der Methode das Testergebnis?

Aus den beschriebenen vergleichenden Studien lässt sich zunächst die Erwartung ableiten, dass die SBZ-Methode möglicherweise zu besseren Testergebnissen führt als die Aag-Methode. Allerdings bleibt es eine offene Frage, ob sich diese Methodendifferenz auch bei einer strikt parallelisierten Vorgehensweise zeigt.

2) Welche Rolle spielt das Alter der getesteten Kinder, sind die Methoden für verschiedene Altersgruppen unterschiedlich gut geeignet?

Dazu lässt sich aufgrund der sehr unterschiedlichen Einschätzungen verschiedener AutorInnen (Miller & Paul, 1995 vs. Goodluck, 1996; Gerken & Shady, 1996 vs. Mathieu, 2008 und Schlesiger, 2001) keine eindeutige Erwartung formulieren. Einerseits erscheint die Aag-Methode im Hinblick auf Gedächtniskapazität und Eigeninitiative gerade für sehr junge Kinder anspruchsvoller als die SBZ-Methode, andererseits birgt die SBZ-Methode ein höheres Rate- und Antwortneigungsrisiko und ist daher insbesondere bei jungen Kindern möglicherweise weniger reliabel als die Aag-Methode.

## ProbandInnen

An der Studie nahmen insgesamt 48 Kinder, verteilt auf zwei Altersgruppen, teil. Alle wuchsen monolingual deutschsprachig auf und waren in ihrer Entwicklung unauffällig. Ein altersgemäßes Grammatikverständnis wurde bei allen Kindern mittels des TROG-D (Fox, 2006) sichergestellt. Tabelle 5 gibt Auskunft über die Alters- und Geschlechtsstruktur der ProbandInnengruppen.

Altersgruppe	Altersschnitt	Jungen	Mädchen
3;0 – 3;5 Jahre	3;25 Jahre	n = 12	n = 12
5;0 – 5;5 Jahre	5;18 Jahre	n = 12	n = 12

Tabelle 5 Alter und Geschlecht der ProbandInnen  
n = Anzahl

## Material

Für die Studie wurden insgesamt zwölf Items, bestehend aus einfachen Hauptsätzen mit transitivem Verb, Subjekt und Objekt, entwickelt. Die Verben wurden so gewählt, dass die entsprechenden Handlungen gleichermaßen ausagierbar und bildlich darstellbar waren. Als Argumente wurden Märchenfiguren verwendet, da diese für Kinder besonders ansprechend und zugleich in verschiedenen thematischen Rollen einsetzbar sind. Letzteres war entscheidend, um die Stimulussätze so konstruieren zu können, dass sie semantisch reversibel waren, d. h. die

Handlungsbeteiligten jeweils die Agens- oder Patiensrolle einnehmen konnten, ohne dass eine bestimmte Interpretation durch das Weltwissen favorisiert werden konnte.

Von den zwölf Testsätzen bestanden neun aus objektinitialen Strukturen (OVS). Durch unterschiedlich viele morphosyntaktische Interpretationshinweise (Transparenz der Akkusativ-Markierung, Eindeutigkeit der Subjekt-Verb-Kongruenz) wurde eine dreifache Abstufung des Schwierigkeitsgrades der Sätze erreicht (vgl. Tab. 6).

Die übrigen drei Testsätze bestanden aus subjektinitialen Strukturen (SVO), um eine Interpretationsstrategie im Sinne der Regel ‚die erste Konstituente ist immer Patiens‘ zu verhindern.

Für die Untersuchung wurden die zwölf Items in eine syntaktisch und lexikalisch abwechslungsreiche Reihenfolge gebracht, um das Risiko möglicher Antwortneigungen zu minimieren. Ein SVO-Satz diente als Übungssitem.

Für den Aag-Test wurden Biegepuppen und Kunststofffiguren verwendet (s. Abb. 2).

Für den SBZ-Test wurden zu jedem Item ein Ziel- und zwei Ablenkerbilder erstellt. Der syntaktische Ablenker basierte auf der Vertauschung der thematischen Rollen, der syntaktisch-morphologische Ablenker entstand durch zusätzliche Veränderung des Numerus bei einem der Verbargumente. Pro Item standen also immer drei Bilder zur Auswahl, wobei sich Ziel- und Ablenkerbilder gleich häufig im oberen, mittleren und unteren Teil der dem Kind vorgelegten DIN-A4-Seite befanden (s. Abb. 3).



Abbildung 2 Material zum Aag-Test (Das Material wurde eigens für die Studie zusammengestellt und fotografiert.)



Abbildung 3 Material zum SBZ-Test (Das Bildmaterial wurde von der Erstautorin selbst erstellt.)

Item 5: „Den Drachen beißen die Gespenster.“



Item 6: „Den König kitzelt der Zauberer.“

## Durchführung

Die Kinder wurden einzeln in einem separaten Raum ihres Kindergartens untersucht und zwar an je zwei verschiedenen Terminen im Abstand von etwa einer Woche. In Tabelle 7 ist der Gesamttablauf der Untersuchung dargestellt.

Sitzung	Inhalte	Dauer
1	<ul style="list-style-type: none"> <li>• Spiel zum Kennenlernen</li> <li>• TROG-D (Fox, 2006)</li> <li>• Spiel zum Abschluss</li> </ul>	20–30 Min.
2	<ul style="list-style-type: none"> <li>• Spiel zum Einfinden</li> <li>• Aag-Test</li> <li>• Spiel als Pause</li> <li>• SBZ-Test</li> <li>• Spiel zum Abschluss</li> <li>• Dankeschön aussuchen</li> </ul>	30–45 Min.

Tabelle 7 Untersuchungsablauf

Für den Aag-Test wurden alle Figuren in einer festen Anordnung (s. Abb. 2) vor dem Kind aufgebaut und dabei die Figurenbezeichnungen erfragt oder zumindest rezeptiv abgesichert. Anschließend

wurde dem Kind das Ausagieren der im Test enthaltenen Verben demonstriert. Danach wurde die Aufgabe erklärt und mit dem Übungsitem begonnen. Hierbei wurden korrekte Reaktionen inhaltlich bestätigt und auf Fehler ausdrücklich aufmerksam gemacht, bis sichergestellt war, dass das Kind die Aufgabe verstanden hatte. Schließlich wurden die zwölf Testsätze mit neutraler Betonung vorgelesen, wobei jedes Item falls nötig (bei Unsicherheit, langem Zögern oder Nachfragen des Kindes) bis zu zweimal wiederholt wurde. Um Lerneffekte zu vermeiden, wurde dem Kind nun – abgesehen von lobenden Äußerungen zu seiner Mitarbeit – keine inhaltliche Rückmeldung mehr zu seiner Reaktion gegeben. Nach jeder Ausagierungsaktion wurden die Figuren wieder an ihren ursprünglichen Platz zurückgestellt, so dass bei jedem Item dieselbe Ausgangssituation bestand.

Da im SBZ-Test dieselben Items verwen-

det wurden, war hier keine erneute Einführung der Handlungen und Figuren notwendig. Es wurde sofort mit der Erklärung der Aufgabe begonnen, darauf folgte das Übungsitem – auch hier mit inhaltlicher Bestätigung bzw. Fehlerkorrektur und ggf. Erklärung der korrekten Antwort. Schließlich wurden die Testsätze vorgelesen – ebenfalls mit neutraler Betonung, maximal zwei Wiederholungen und ohne inhaltliche Rückmeldung zur Korrektheit der Reaktion des Kindes. Zu jedem Item wurde zuerst die Bildseite aufgedeckt, so dass das Kind sich alle Bilder genau anschauen konnte, bevor der entsprechende Testsatz präsentiert wurde. Dabei wurde darauf geachtet, das Kind nicht durch den eigenen Blick zum Zielbild zu führen.

## Auswertungskriterien

Die Auswertung bezog sich zum einen auf die Anzahl korrekter Reaktionen, zum anderen wurde unter Verwendung eines

Erwerbskriteriums betrachtet, wie viele Kinder der jeweiligen Altersgruppe die relevante Struktur (OVS) rezeptiv erworben hatten.

### a) Korrektheit

Im Aag-Test wurden alle Reaktionen als korrekt klassifiziert, die den Inhalt des Stimulussatzes in syntaktischer Hinsicht korrekt wiedergaben. Syntaktische Abweichungen (Vertauschung der thematischen Rollen Agens und Patiens), Nullreaktionen (keine Handlung) und nicht klassifizierbare Reaktionen (unklare Handlungen) wurden als inkorrekt klassifiziert. Reaktionen, die auf einer inkorrekten lexikalischen oder morphologischen Verarbeitung beruhten, syntaktisch jedoch der Zielreaktion entsprachen (z. B. falsche Figurenauswahl bei ansonsten korrekten thematischen

Rollen) führten nicht zu einer Bewertung der Reaktion als inkorrekt, da in der Untersuchung das Syntaxverständnis im Mittelpunkt stand.

Im SBZ-Test wurde nur ein eindeutiges Zeigen auf das Zielbild als korrekte Reaktion bewertet, während das Zeigen auf eines der Ablenkerbilder sowie Nullreaktionen und nicht klassifizierbare Reaktionen (z. B. Zeigen mehrerer Bilder) als inkorrekt bewertet wurden.

Anschließend wurde die Anzahl korrekter Reaktionen aus jedem der beiden Tests über alle Kinder der jeweiligen Altersgruppe gemittelt – sowohl für die Gesamtitems als auch separat für die SVO- und OVS-Items.

In der statistischen Auswertung wurden zunächst die Unterschiede zwischen den durchschnittlichen Ergebnissen aus den beiden Verfahren innerhalb einer Alters-

gruppe geprüft. Aufgrund der geringen Stichprobengröße wurde dafür ein nicht-parametrisches Verfahren gewählt (Wilcoxon-Test). Anschließend wurde berechnet, ob sich die durchschnittlichen Leistungen der Kinder innerhalb eines Verfahrens in Abhängigkeit vom Alter unterscheiden. Dazu wurde ebenfalls ein nicht-parametrisches Verfahren verwendet (Mann-Whitney-U-Test).

### b) Beherrschung der Zielstruktur (Erwerbskriterium)

Darüber hinaus sollte individuell festgestellt werden, ob die in Frage stehende Struktur (OVS-Sätze) von dem jeweiligen Kind sicher verstanden wird.

Für den SBZ-Test, bei dem methodenbedingt mit Rateeffekten zu rechnen war, wurde das Kriterium der ausreichend geringen Gesamtratewahrscheinlichkeit ( $P < 5\%$ , s. Exkurs zur Ratewahrscheinlichkeit) angesetzt. Danach mussten mindestens sechs OVS-Items korrekt gelöst werden. Bei einer Anzahl von insgesamt neun Testsätzen beginnt erst hier der Bereich, in dem sicher mit einer überzufälligen Leistung zu rechnen ist ( $P(k \geq 6) = 0,042$ , vgl. Tab. 3).

Analog wurde auch für den Aag-Test ein Level von sechs von neun korrekten Reaktionen angesetzt (auch wenn das Raterisiko hier nicht wie für den SBZ-Test berechnet werden kann). Diese Mindestmenge korrekter Reaktionen entspricht einem Erwerbskriterium von ca. 67 Prozent, wie es z. B. Fox (2004) unabhängig von Ratewahrscheinlichkeiten in ihrer Studie zum Phonologieerwerb verwendet hat.

In der Auswertung wurde dann pro Test ermittelt, wie viele Kinder der jeweiligen Altersgruppe bei den OVS-Sätzen mindestens 67 Prozent (sechs von neun) korrekte Reaktionen zeigten. Mit Hilfe des nicht-parametrischen McNemar-Tests wurde ermittelt, ob die Anzahl der Kinder, deren Leistungen für eine Beherrschung der Zielstruktur sprechen, in Abhängigkeit von der jeweiligen Methode variiert.

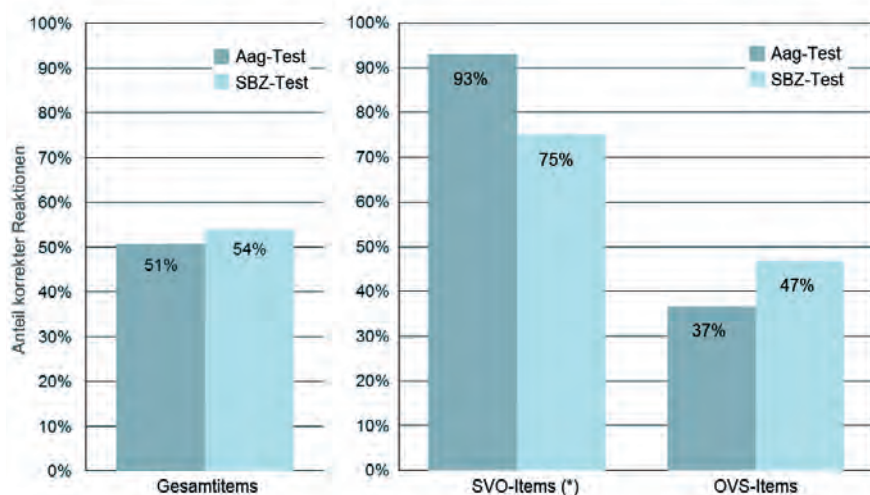


Abbildung 4 Durchschnittliche Leistungen der 3;0- bis 3;5-jährigen Kinder

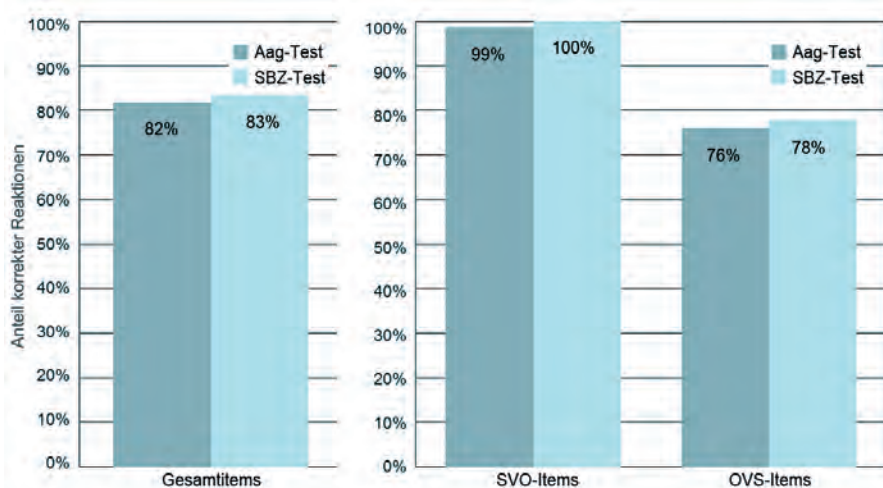


Abbildung 5 Durchschnittliche Leistungen der 5;0- bis 5;5-jährigen Kinder

## Ergebnisse

### a) Korrektheit

Die 3;0- bis 3;5-jährigen Kinder erzielten bezogen auf die zwölf Gesamtitems ähnliche Leistungen mit beiden Metho-

den (Aag: Mittelwert 6,08, SD 2,99; SBZ: Mittelwert 6,46, SD 1,98; vgl. Abb. 4). Die Mittelwerte unterscheiden sich nicht signifikant.

Bei separater Betrachtung der einzelnen Itemkategorien, zeigt sich folgendes Bild: Bei den neun OVS-Items wurden im SBZ-Test mehr korrekte Reaktionen erreicht als im Aag-Test (Aag: Mittelwert 3,29, SD 2,81; SBZ: Mittelwert 4,21, SD 1,84; vgl. Abb. 4). Dieser leichte Vorteil für die SBZ-Methode ist tendenziell signifikant (Wilcoxon:  $p = 0,074$ ). Da die durchschnittliche Leistung im SBZ-Test jedoch im Ratebereich (weniger als sechs von neun Reaktionen korrekt) liegt, d. h. durchaus auf Zufallsreaktionen zurückzuführen sein könnte, ist der leichte Methodenunterschied bei den OVS-Sätzen nur vorsichtig zu bewerten.

Bei den drei SVO-Items zeigten die Dreijährigen im Aag-Test bessere Leistungen als im SBZ-Test (Aag: Mittelwert 2,79, SD 0,59; SBZ: Mittelwert 2,25, SD 0,9; vgl. Abb. 4). Diese Mittelwertdifferenz ist signifikant (Wilcoxon:  $p = 0,006$ ). Da bei dreijährigen Kindern eigentlich davon auszugehen ist, dass sie SVO-Sätze sicher verarbeiten können, deutet dies auf ein konsequenteres Antwortverhalten der jüngeren Kinder bei der Aag-Methode hin.

Die 5;0- bis 5;5-jährigen Kinder erzielten, sowohl bezogen auf die zwölf Gesamtitems (Aag: Mittelwert 9,79, SD 1,82; SBZ: Mittelwert 10,0, SD 2,15) als auch auf die neun OVS-Items (Aag: Mittelwert 6,83, SD 1,81; SBZ: Mittelwert 7,0, SD 2,15) sowie die drei SVO-Items (Aag: Mittelwert 2,96, SD 0,2; SBZ: Mittelwert 3,0, SD 0,0) vergleichbare Leistungen mit beiden Methoden (vgl. Abb. 5). Die Mittelwerte unterscheiden sich in keinem Fall signifikant. In dieser Altersgruppe führten die verschiedenen Methoden also nicht zu einem Leistungsunterschied.

Wie aus den o. g. Mittelwerten bereits ersichtlich wird, waren die Leistungen der älteren Kinder in beiden Tests besser als die der jüngeren. Statistische Gruppenvergleiche bestätigen diese Altersdifferenz für die Gesamtitems sowie für die OVS-Items: Die Unterschiede in der durchschnittlichen Anzahl korrekter Reaktionen zwischen den älteren und jüngeren Kindern sind in beiden Testverfahren hoch signifikant (Mann-

Whitney-U-Test: jeweils  $p < 0,001$ ). Im Falle der SVO-Items war bereits von den Dreijährigen ein sicheres Verständnis zu erwarten. Im Aag-Test ergab sich hier dementsprechend kein signifikanter Unterschied zu den Leistungen der Fünfjährigen – im SBZ-Test allerdings durchaus (Mann-Whitney-U-Test:  $p = 0,000$ ), was wiederum auf ein systematischeres Antwortverhalten der jüngeren Kinder im Aag-Test hinweist.

#### b) Beherrschung der Zielstruktur

Innerhalb der jüngeren Altersgruppe erreichten mit der Aag-Methode fünf Kinder (21% der Stichprobe) und mit der SBZ-Methode vier Kinder (16%) Werte, die auf ein sicheres Verständnis der Zielstruktur (OVS) hindeuten (mindestens 67%, d. h. sechs von neun Reaktionen korrekt). Im McNemar-Test ergab sich diesbezüglich kein signifikanter Unterschied. Die Mehrzahl der Dreijährigen hatte also – unabhängig von der Testmethode – noch massive Schwierigkeiten mit der Verarbeitung von OVS-Sätzen.

Von den Fünfjährigen erreichten mit der Aag-Methode 18 Kinder (75%) Werte über dem festgelegten Level, mit der SBZ-Methode waren es 22 Kinder (92%). Wiederum ergab sich kein signifikanter Unterschied zwischen diesen Werten. Der überwiegende Teil der älteren Kinder konnte die Zielstruktur demnach – ebenfalls unabhängig von der Testmethode – sicher verstehen.

## KURZBIOGRAFIE

**Prof. Dr. Christina Kauschke** ist Professorin für Germanistische Linguistik mit dem Schwerpunkt Klinische Linguistik an der Philipps-Universität Marburg. Nach einer Ausbildung als Logopädin und dem Studium der Sprachwissenschaft war sie an den Universitäten Berlin (FU) und Potsdam tätig. Ihre Schwerpunkte sind der normale und gestörte Spracherwerb, Wortverarbeitung bei Kindern, Sprachentwicklungsstörungen, Entwicklung von Diagnostik- und Therapieverfahren sowie Sprache und Emotion.

## Diskussion und Schlussfolgerungen

In der vorliegenden Studie wurde das Verstehen syntaktischer Strukturen bei Kindern mit zwei unterschiedlichen Testmethoden, dem Ausagieren und der Satz-Bild-Zuordnung untersucht. Ziel war es festzustellen, ob die beiden Methoden zu vergleichbaren Ergebnissen führen und ob mögliche Methodeneffekte abhängig vom Alter der Kinder sind. Um einen direkten Vergleich zu ermöglichen, wurde in beiden Verfahren exakt parallelisiertes Material verwendet und die Ergebnisse innerhalb jeweils einer Stichprobe von drei- bzw. fünfjährigen Kindern miteinander verglichen. Ein solcher Methodenvergleich wurde in der bisherigen Forschung noch nicht durchgeführt.

Als Teststimuli wurden Sätze mit Objektvoranstellung verwendet. Bei dieser Satzstruktur war zu erwarten, dass die dreijährigen Kinder noch beträchtliche Probleme mit der korrekten Interpretation haben würden, während bei den fünfjährigen Kindern von deutlich fortgeschritteneren Verständnisfähigkeiten auszugehen war.

Zunächst wurde die durchschnittliche Anzahl korrekter Reaktionen in den beiden Testvarianten innerhalb der jeweiligen Altersgruppe miteinander verglichen: In Bezug auf die Gesamtitems führten beide Methoden zu vergleichbaren Testergebnissen – sowohl bei den Dreijährigen als auch bei den Fünfjährigen. Auch bei separater Betrachtung der im Fokus der Untersuchung stehenden OVS-Sätze ergab sich in keiner der beiden Altersgruppen ein bedeutsamer Unterschied zwischen den durchschnittlichen Leistungen, die mit der jeweiligen Methode erzielt wurden.

Außerdem wurde untersucht, wie viele Kinder aus den jeweiligen Altersgruppen bei den OVS-Sätzen Reaktionen zeigten, die auf ein sicheres Verständnis dieser Struktur hinweisen. Dazu wurde ein Kriterium von mindestens 67 Prozent korrekter Reaktionen angesetzt. Auch hierbei konnte in keiner der beiden Altersgruppen eine methodenbedingte Differenz festgestellt werden: Von den Dreijährigen erreichten sowohl mit der Aag- als auch mit der SBZ-Methode weniger als ein Viertel der Kinder das Er-

werkskriterium, bei den Fünfjährigen waren es mit beiden Methoden mindestens drei Viertel der Kinder.

Der Vergleich der durchschnittlichen Leistungen der beiden Altersgruppen innerhalb der jeweiligen Testvariante ergab zudem, dass die älteren Kinder bezüglich der Gesamt- und der OVS-Items mit beiden Methoden deutlich besser abschnitten als die jüngeren.

Insgesamt zeigen die Ergebnisse, dass die Aag- und die SBZ-Methode im Wesentlichen zu gleichwertigen Testergebnissen führen, sowohl in Bezug auf durchschnittliche Gruppenwerte als auch hinsichtlich der individuellen Beherrschung der Zielstruktur. Darüber hinaus erwiesen sich beide Methoden als entwicklungs sensitiv, insofern als sie altersbedingten Zuwachs an Verständnisfähigkeiten gleichermaßen deutlich abbildeten. Der in der Literatur oft geäußerte Vorteil für die SBZ- bzw. Nachteil für die Aag-Methode hinsichtlich des Schwierigkeitsgrades (Cocking & McHale, 1981; Gebhard, 2008; Dittmar et al., 2008) lässt sich somit nicht bestätigen.

#### Fazit 1:

Bei der Untersuchung des Grammatikverständnisses hat die Wahl der Methode keinen wesentlichen Einfluss auf die Testergebnisse. Bestehende Unterschiede hinsichtlich der Anforderungen an das Kind wirken sich nicht bedeutsam auf die Testleistung aus. Beide Methoden sind entwicklungs sensitiv.

Da das verwendete Stimulusset nicht nur die im Fokus stehenden OVS-Sätze, sondern auch einige kanonische Sätze mit SVO-Stellung enthielt, wurden auch diese separat betrachtet. Dabei fiel auf, dass die dreijährigen Kinder hier mit der Aag-Methode bessere Leistungen erzielten als mit der SBZ-Methode. Vor dem Hintergrund des im Alter von drei Jahren anzunehmenden sicheren Verständnisses von SVO-Strukturen deutet dies auf ein konsequenteres Antwortverhalten der jüngeren Kinder bei der Aag-Methode hin.

Im Altersvergleich zeigte sich zudem, dass die Leistungen der Drei- und Fünfjährigen bei den SVO-Items im Aag-Test vergleichbar waren, im SBZ-Test jedoch ein deutlicher Unterschied zugunsten der älteren Kinder bestand. Angesichts

der auch von den Dreijährigen zu erwartenden Verständnisfähigkeiten für SVO-Sätze erscheint ein vergleichbares Ergebnis in beiden Altersgruppen hier schlüssiger, was wiederum für verlässlichere Reaktionen der Dreijährigen bei der Aag-Methode spricht.

Insgesamt deuten diese Befunde darauf hin, dass die Aag-Methode bei jüngeren Kindern möglicherweise zu konsequenterem Antwortverhalten führt und damit verlässlichere Ergebnisse liefert als die SBZ-Methode, bei der aufgrund des Multiple-Choice-Formats die Möglichkeit zu willkürlichen Reaktionen besteht.

Beobachtungen zum Verhalten der Kinder während der Tests untermauern diese Vermutung: Etwa die Hälfte der Dreijährigen wirkte im SBZ-Test eher unkonzentriert: Die Kinder schauten sich trotz Aufforderung nicht alle Bilder gleich intensiv an, zeigten auch bei wiederholter Stimuluspräsentation offensichtliche Präferenzen für eine bestimmte Bildposition oder das Bild mit den interessantesten Details. Im Aag-Test dagegen zeigte nur eines der jüngeren Kinder ähnlich unkonzentriertes Verhalten. Die Fünfjährigen dagegen verhielten sich in keinem der beiden Tests derart auffällig.

Die Beobachtungen legen die Vermutung nahe, dass die SBZ-Methode bei jüngeren Kindern mit einem besonders hohen Raterisiko behaftet ist, das beim Einsatz dieser Methode unbedingt berücksichtigt werden sollte. Diese Schlussfolgerung deckt sich zum einen mit der Einschätzung von Schlesiger (2001), dass die SBZ-Methode bei drei- und vierjährigen Kindern wenig reliabel sei. Zum anderen spiegelt sich darin möglicherweise auch die von Mathieu (2008) geäußerte Einschätzung, dass Kinder unter vier Jahren noch nicht über die notwendige Abstraktionsfähigkeit für die Arbeit mit reinem Bildmaterial verfügen.

#### Fazit 2:

Für jüngere Kinder könnte die Aag-Methode im Hinblick auf die Verlässlichkeit der Testergebnisse geeigneter sein als die SBZ-Methode. Bei älteren Kindern hingegen ist das Grammatikverständnis mit beiden Methoden gleichermaßen zuverlässig zu erfassen.

Die geringe Anzahl der in dieser Studie verwendeten SVO-Items lässt noch

keine endgültigen Schlüsse in der Frage nach der Eignung der Methoden für verschiedene Altersgruppen zu – hier besteht für die zukünftige Forschung noch Klärungsbedarf, z. B. in Form einer Studie mit deutlich mehr Testsätzen, deren Struktur von Dreijährigen sicher verarbeitet werden kann.

Um die Anwendbarkeit und Aussagekraft der in dieser Studie verglichenen Methoden außerdem auch im Hinblick auf die Diagnostik von Störungen des Sprachverständnisses beurteilen zu können, sind im nächsten Schritt entsprechende Untersuchungen mit sprachauffälligen Kindern notwendig. Interessant wäre es herauszufinden, ob sich SSES-Kinder diesbezüglich eher wie gleichaltrige sprachunauffällige Kinder verhalten, die sich kognitiv auf einem ähnlichen Entwicklungsstand befinden, oder ob sie eher wie jüngere Kinder mit vergleichbarem Sprachentwicklungsniveau reagieren.

Abschließend ist zu bedenken, dass beide hier verglichenen Methoden auf Verhaltensbeobachtungen beruhen. Welche Verarbeitungsprozesse beim Kind tatsächlich stattfinden, kann aus dem beobachteten Verhalten nicht unmittelbar erkannt werden. Experimentelle Methoden dagegen erlauben es, den tatsächlichen kognitiven Prozessen und neuronalen Mechanismen bei der Satzverarbeitung näher zu kommen. So untersuchte Schipke (2012) das Verstehen von Sätzen mit Objektvoranstellung sowohl mit einer Satz-Bild-Zuordnungsaufgabe als auch mit einem eye-tracking-Verfahren sowie mit der Ableitung ereigniskorrelierter Potentiale (EKP). Die Ergebnisse zeigen, dass auf einer neurophysiologischen Ebene durchaus eine Sensitivität für die besondere Struktur objektinitialer Sätze nachgewiesen werden kann, obwohl die Verhaltensdaten noch sehr unsichere Reaktionen zeigen. Experimentelle Methoden können also tiefere Einblicke in die Satzverarbeitungsfähigkeiten von Kindern bieten.

#### Literatur

- Bishop, D. V. M. (1982). Comprehension of Spoken, Written and Signed Sentences in Childhood Language Disorders. *Journal of Child Psychology and Psychiatry*, 1, 1-20.
- Bishop, D. V. M. (1997). *Uncommon Understanding – Development and Disorders of Language Comprehension in Children*. Hove: Psychology Press.
- Bortz, J., & Lienert, G. (2003). *Kurzgefasste Statistik für die klinische Forschung – Leitfaden für*

- die verteilungsfreie Analyse kleiner Stichproben. Berlin, Heidelberg: Springer.
- Buschmann, A., & Neubauer, M. (2012). Prädikatoren für den Entwicklungsverlauf spät sprechender Kinder. *Sprache Stimme Gehör*, 3, 135-141.
- Chapman, R. S., & Miller, J. F. (1975). Word Order in Early Two and Three Word Utterances: Does Production Precede Comprehension? *Journal of Speech and Hearing Research*, 18, 355-371.
- Chien, Y.-C., & Wexler, K. (1990). Children's Acquisition of the Locality Conditions in Binding as Evidence for the Modularity of Syntax and Pragmatics. *Language Acquisition*, 1, 225-295.
- Chipman, H. H., & Dannenbauer, F. M. (1988). Kleine Kinder verstehen anders: Zum Problem der Erfassung kindlicher Strategien des Sprachverständnisses. In K. B. Günther (Hrsg.), *Sprachstörungen – Probleme ihrer Diagnostik bei mentalen Retardierungen, Entwicklungsdysphasien und Aphasien*. Heidelberg: Heidelberger Verlagsanstalt und Druckerei GmbH, Edition Schindele.
- Chomsky, N. (1969). *The Acquisition of Syntax in Children from 5 to 10*. Cambridge, Mass.: MIT Press.
- Clahsen, H. (1986). *Die Profilanalyse. Ein linguistisches Verfahren für die Sprachdiagnose im Vorschulalter*. Berlin: Marhold.
- Clark, R. (1982). Theory and Method in Child-Language Research: Are we Assuming too much? In S. A. Kuczaj II (Hrsg.), *Language Development, Volume 1: Syntax and Semantics*. Hillsdale, N. J.: Earlbaum.
- Clark, E. V. (1971). On the Acquisition of the Meaning of Before and After. *Journal of Verbal Learning and Verbal Behavior*, 10, 266-275.
- Cocking, R., & McHale, S. (1981). A Comparative Study of the Use of Pictures and Objects in Assessing Children's Receptive and Productive Language. *Journal of Child Language*, 8, 1-13.
- Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German Children's Comprehension of Word Order and Case Marking in Causative Sentences. *Child Development*, 4, 1152-1167.
- de Langen-Müller, U., Kauschke, C., Kiese-Himmel, C., Neumann, K., & Noterdaeme, M. (2012). *Interdisziplinäre S2k-Leitlinie: Diagnostik von Sprachentwicklungsstörungen (SES) unter Berücksichtigung umschriebener Sprachentwicklungsstörungen*. Frankfurt/Main: Peter Lang. Verfügbar unter <http://www.awmf.org/leitlinien/aktuelle-leitlinien.html>.
- de Villiers, J. P., & de Villiers, P. A. (1973). Development of the Use of Word Order in Comprehension. *Journal of Psycholinguistic Research*, 4, 331-341.
- Elben, C. E., & Lohaus, A. (2000). *MSVK – Marburger Sprachverständnistest für Kinder*. Göttingen: Hogrefe.
- Endres, R., & Baur, S. (2000). Informelles Verfahren zur Überprüfung von Sprachverständnisseleistungen (IVÜS) – Diskussion und Materialien. *Die Sprachheilarbeit*, 2, 64-71.
- Fox, A. V. (2004). *Kindliche Aussprachestörungen: Phonologischer Erwerb · Differenzialdiagnostik · Therapie*. Idstein: Schulz-Kirchner.
- Fox, A. V. (2006). *TROG-D – Test zur Überprüfung des Grammatikverständnisses*. Idstein: Schulz-Kirchner.
- Fragman, C., Goodluck, H., & Heggie, L. (2007). Child and Adult Construal of Restrictive Relative Clauses: Knowledge of Grammar and Differential Effects of Syntactic Context. *Journal of Child Language*, 34, 345-380.
- Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behavior*, 2, 121-135.
- Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: a study of SLI and normal development. *Journal of Child Language*, 31, 661-681.
- Gebhard, W. (2008). *Entwicklungsbedingte Sprachverständnisstörungen bei Kindern im Grundschulalter – Status und Diagnostik im klinischen Kontext*. München: Utz.
- Gebhard, W., Dames, K., & Baur, S. (1994). Diagnostik und Therapie schwerer Sprachentwicklungsstörungen in einer Klinik. In J. Martinius, & H. Amorosa (Hrsg.), *Teilleistungsstörungen*. Berlin, München: Quintessenz.
- Gerken, L. A., & McIntosh, B. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, 29, 448-457.
- Gerken, L. A., & Shady, M. E. (1996). The Picture Selection Task. In D. McDaniel, C. McKee, & H. Smith Cairns (Hrsg.), *Methods for Assessing Children's Syntax*. Cambridge, Mass.: MIT Press.
- Goodluck, H., & Behne, D. (1992). Development in Control and Extraction. In J. Weißenborn, H. Goodluck, & T. Roeper (Hrsg.), *Theoretical Issues in Language Acquisition: Continuity and Change in Development*. Hillsdale, N.J.: Earlbaum.
- Goodluck, H., Terzi, A., & Díaz, G. C. (2001). The Acquisition of Control Crosslinguistically: Structural and Lexical Factors in Learning to Licence PRO. *Journal of Child Language*, 28, 153-172.
- Grieve, R., Hoogenraad, R., & Murray, D. (1977). On the Young Child's Use of Lexis and Syntax in Understanding Locative Instructions. *Cognition*, 5, 235-250.
- Grimm, H. (1975). Verstehen, Imitation und Produktion von Passivsätzen. In H. Grimm, H. Schöler, & M. Wintermantel (Hrsg.), *Zur Entwicklung sprachlicher Strukturformen bei Kindern*. Weinheim, Basel: Beltz.
- Grimm, H., Aktaş, M., & Frevert, S. (2000). *SETK2 – Sprachentwicklungstest für zweijährige Kinder – Diagnose rezeptiver und produktiver Sprachverarbeitungsfähigkeiten*. Göttingen: Hogrefe.
- Grimm, H., Aktaş, M., & Frevert, S. (2001). *SETK 3-5 – Sprachentwicklungstest für drei- bis fünfjährige Kinder – Diagnose von Sprachverarbeitungsfähigkeiten und auditiven Gedächtnisleistungen*. Göttingen: Hogrefe.
- Grimm, H., & Schöler, M. (1975). Erlauben – Befehlen – Lassen: Wie gut verstehen kleine Kinder kausativierende Beziehungen? In H. Grimm, H. Schöler, & M. Wintermantel (Hrsg.), *Zur Entwicklung sprachlicher Strukturformen bei Kindern*. Weinheim, Basel: Beltz.
- Grimm, H., & Schöler, M. (1978). *HSET – Heidelberger Sprachentwicklungstest*. Göttingen: Hogrefe.
- Grimm, H., & Wintermantel, M. (1975a). Zur Entwicklung von Sprachtests. In H. Grimm, H. Schöler, & M. Wintermantel (Hrsg.), *Zur Entwicklung sprachlicher Strukturformen bei Kindern*. Weinheim, Basel: Beltz.
- Grimm, H., & Wintermantel, M. (1975b). Über das Verstehen von Relativsatzstrukturen. In H. Grimm, H. Schöler, & M. Wintermantel (Hrsg.), *Zur Entwicklung sprachlicher Strukturformen bei Kindern*. Weinheim, Basel: Beltz.
- Hakansson, G., & Hansson, K. (2000). Comprehension and production of relative clauses: a comparison between Swedish impaired and unimpaired children. *Journal of Child Language*, 27, 313-333.
- Hsu, J. R., Smith Cairns, H., Eisenberg, S., & Schlüsselberg, G. (1991). When Do Children Avoid Backward Coreference? *Journal of Child Language*, 18, 339-353.
- Kauschke, C., & Siegmüller, J. (2009). *Pathologische Diagnostik bei Sprachentwicklungsstörungen*. München: Elsevier.
- Kidd, E., & Bavin, E. L. (2005). Lexical and Referential Cues to Sentence Interpretation: An Investigation of Children's Interpretations of Ambiguous Sentences. *Journal of Child Language*, 32, 855-876.
- Kubinger, K. D. (2009). *Psychologische Diagnostik – Theorie und Praxis psychologischen Diagnostizierens*. Göttingen: Hogrefe.
- Lienert, G. A., & Raatz, U. (1998). *Testaufbau und Testanalyse*. Weinheim: Psychologie Verlagsunion, Beltz.
- Mathieu, S. (2008). Entwicklung und Abklärung des Sprachverständnisses. In B. Zollinger (Hrsg.), *Kinder im Vorschulalter – Erkenntnisse, Beobachtungen und Ideen zur Welt der Drei- bis Siebenjährigen*. Bern: Haupt.
- Miller, J. F., & Paul, R. (1995). *The Clinical Assessment of Language Comprehension*. Baltimore: Brookes.
- Miolo, G., Chapman, R. S., & Sindberg, H. A. (2005). Sentence Comprehension in Adolescents With Down Syndrome and Typically Developing Children: Role of Sentence Voice, Visual Context, and Auditory Short-Term Memory. *Journal of Speech, Language, and Hearing Research*, 48, 172-188.
- Montgomery, J. W. (2000). Verbal Working Memory and Sentence Comprehension in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, 43, 293-308.
- Montgomery, J. W., & Evans, J. L. (2009). Complex Sentence Comprehension and Working Memory in Children With Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, 52, 269-288.
- Motsch, H.-J. (2000). *ESGRAF – Evozierte Sprachdiagnose grammatischer Fähigkeiten*. München: Reinhardt.
- Naigles, L. R., & Lehrer, N. (2002). Language-General and Language-Specific Influences on Children's Acquisition of Argument Structure: Comparison of French and English. *Journal of Child Language*, 29, 545-566.
- Oshima-Takane, Y. (1992). Analysis of pronominal errors: a case study. *Journal of Child Language*, 19, 111-131.
- Penner, Z. (1999). *Screeningverfahren zur Feststellung von Störungen in der Grammatikentwicklung*. Luzern, Biel: Edition der schweizerischen Zentralstelle für Heilpädagogik, SHZ.
- Sachse, S., & v. Suchodoletz, W. (2009). Prognose und Möglichkeiten der Vorhersage der Sprachentwicklung bei Kindern mit verzögertem Sprechbeginn (Late Talkers). *Kinderärztliche Praxis*, 5, 318-328.
- Schipke, C. S. (2012). *Processing Mechanisms of Argument Structure and Case Marking in Child Development: Neural Correlates and Behavioral Evidence*. Leipzig: MPI Series in human cognitive and brain sciences.
- Schlesiger, C. (2001). *Sprachverstehen bei spezifischer Sprachentwicklungsstörung – Grundlagen und Diagnostik*. Frankfurt/Main: Lang.

- Schmitz, P., & Fox, A. V. (2007). Sprachverstehenstests im Deutschen unter besonderer Berücksichtigung des TROG-D. *Forum Logopädie*, 4, 18-25.
- Schöler, H. (1975). Verstehen und Imitation temporaler Satzformen. In H. Grimm, H. Schöler, & M. Wintermantel (Hrsg.), *Zur Entwicklung sprachlicher Strukturformen bei Kindern*. Weinheim, Basel: Beltz.
- Schöler, H., Schakib-Ekbatan, K., Spohn, B., & Spohn, S. (1999). *IDIS – Inventar diagnostischer Informationen bei Sprachentwicklungsauffälligkeiten*. Heidelberg: Winter.
- Schrey-Dern, D. (2006). *Sprachentwicklungsstörungen – Logopädische Diagnostik und Therapieplanung*. Stuttgart: Thieme.
- Schulz, P., & Grimm, A. (2012). Spracherwerb. In H. Drügh, S. Komfort-Hein, A. Kraß, C. Meier, G. Rohowski, R. Seidel, & H. Weiß (Hrsg.), *Germanistik. Sprachwissenschaft · Literaturwissenschaft · Schlüsselkompetenzen*. Stuttgart, Weimar: Metzler.
- Sheldon, A. (1974). The Role of Parallel Function in the Acquisition of Relative Clauses in English. *Journal of Verbal Learning and Verbal Behavior*, 13, 272-281.
- Siegmüller, J., Kauschke, C., Bartke, S., & Bittner, D. (2011). *Test zum Satzverstehen von Kindern*. München: Elsevier.
- Siegmüller, J. (2003). Sprachtherapeutische Intervention bei Störungen auf der grammatischen Ebene bei Schulkindern. *L.O.G.O.S. INTERDISZIPLINÄR*, 1, 36-41.
- Stevenson, R. J., & Pollit, C. (1987): The Acquisition of Temporal Terms. *Journal of Child Language*, 14, 533-545.
- Thal, D. J., & Flores, M. (2001). Development of Sentence Interpretation Strategies by Typically Developing and Late-Talking Toddlers. *Journal of Child Language*, 28, 173-193.
- van der Lely, H. K. (1990). Comprehension of reversible sentences in specifically language-impaired children. *Journal of Speech and Hearing Disorders*, 55, 101-117.
- van der Lely, H. K. (1994). Canonical Linking Rules: Forward Versus Reverse Linking in Normally Developing and Specifically Language-Impaired Children. *Cognition*, 51, 29-72.
- van der Lely, H. K. (1996). Specifically language-impaired and normally developing children: Verbal passive vs. adjectival passive sentence interpretation. *Lingua*, 98, 243-272.
- Veit, S. (1986). Das Verständnis von Plural- und Komparativformen bei (entwicklungs-)dysgrammatischen Kindern im Vorschulalter. In G. Kegel, T. Arnhold, K. Dahlmeier, & B. Tischer (Hrsg.), *Sprechwissenschaft und Psycholinguistik: Beiträge aus Forschung und Praxis*. Opladen: Westdeutscher Verlag.
- Watermeyer, M. (2010). *Ausagieren von Sätzen versus Satz-Bild-Zuordnung: Vergleich zweier Methoden zur Untersuchung des Sprachverständnisses anhand von semantisch reversiblen Sätzen mit Objektvoranstellung bei drei- und fünfjährigen Kindern*. Diplomarbeit im Studiengang Patholinguistik an der Universität Potsdam. Verfügbar unter <http://www.opus.kobv.de/ubp/volltexte/2011/5238/>.
- Watermeyer, M., Höhle, B., & Kauschke, C. (2011). Ausagieren von Sätzen versus Satz-Bild-Zuordnung: Vergleich zweier Methoden zur Untersuchung des Sprachverständnisses anhand von semantisch reversiblen Sätzen mit Objektvoranstellung bei drei- und fünfjährigen Kindern. In S. Hanne, T. Fritzsche, S. Ott, & A. Adelt (Hrsg.), *Spektrum Patholinguistik: Band 4*. Potsdam: Universitätsverlag.
- Weist, R. M., Wysocka, H., & Lyytinen, P. (1991). A cross-linguistic perspective on the development of temporal systems. *Journal of Child Language*, 18, 67-92.
- Wettstein, P. (1997). *Psycholinguistischer Sprachverständnis- und Sprachentwicklungstest*. Uster: BSSI, Wettstein.



#### Autorinnen

Melanie Watermeyer  
(Logopädin, Dipl.-Patholinguistin)  
Zülpicher Wall 40, 50674 Köln  
Melanie@Watermeyer.org

Prof. Dr. Christina Kauschke  
(Logopädin, klinische Linguistin)  
Wilhelm-Röpke-Straße 6, 35032 Marburg  
kauschke@staff.uni-marburg.de



DOI dieses Beitrags  
([www.doi.org](http://www.doi.org))  
10.7345/prolog-1304264