

N° 02  
2014

VOL. 37  
ISSN 1664-8595

# Aphasie

und verwandte Gebiete  
et domaines associés

Originalbeitrag

**Kognitive Sprach- und  
Kommunikationsstö-  
rungen – Diagnostik und  
Therapieimplikationen  
anhand des Screening-  
verfahrens MAKRO**

Originalbeitrag

**Investigating treatment  
fidelity in a conver-  
sation-based aphasia  
therapy**

Originalbeitrag

**Semantisches Wieder-  
holungspriming in der  
Therapie von Wortabru-  
fstörungen bei Aphasie  
– Eine Pilotstudie**



aphasiesuisse (...)

wenn Worte fehlen.  
quand les mots font défaut.  
quando le parole sfuggono.  
sch'ils plets mancan.

## Originalbeiträge | Articles originaux

---

**3 Kognitive Sprach- und Kommunikationsstörungen –  
Diagnostik und Therapieimplikationen anhand des Screening-  
verfahrens MAKRO**

Büttner, Julia

**14 Investigating treatment fidelity in a conversation-based  
aphasia therapy**

Heilemann, Claudia; Best, Wendy; Johnson, Fiona; Beckley, Firlie;  
Maxim, Jane; Beeke, Suzanne

**27 Semantisches Wiederholungspriming in der Therapie von  
Wortabrufstörungen bei Aphasie – Eine Pilotstudie**

Schröder, Astrid; Hausmann, Nadin; Stadie, Nicole

Originalbeitrag

# Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO

Büttner, Julia<sup>1 2</sup>

## DE | Zusammenfassung

Nicht-aphasische Sprach- und Kommunikationsstörungen sind in unterschiedlicher Ausprägung eine häufige Folge von neurologischen Erkrankungen. Die Symptome dieser kognitiven Sprach- und Kommunikationsstörungen können in ihrem Schweregrad je nach begleitenden neuropsychologischen Defiziten deutlich variieren. In der Forschungsliteratur werden in diesem Zusammenhang auch Dysfunktionen an der Schnittstelle zwischen exekutiven und linguistischen Fähigkeiten diskutiert (Ardila 2012, Lê et al. 2012, Ferstl et al. 2005). Während für die Diagnostik und Therapie der aphasisch bedingten Sprachstörungen empirisch überprüfte und modelltheoretisch fundierte Verfahren zur Verfügung stehen, ist der Themenbereich der kognitiven Sprach- und Kommunikationsstörungen im deutschsprachigen Raum bislang eher wenig untersucht. Für das systematische Erfassen von makrostrukturellen Fähigkeiten bei gestörten Exekutivfunktionen wurde daher im Rahmen eines Forschungsprojekts das empirisch überprüfte Screeningverfahren MAKRO (Büttner 2012, 2013, im Druck) für die Diskurs- und Textverarbeitung entwickelt, das auch Implikationen für Therapieansätze bietet. Im Folgenden wird auf die Grundzüge des Screeningverfahrens, die Ergebnisse einer Gruppenstudie und die daraus ableitbaren therapeutischen Möglichkeiten eingegangen.

## EN | Abstract

Non-aphasic language disorders are to a varying degree a common, but not yet well examined result of neurological disorders. The severity of the symptoms of this cognitive language and communication disorder can vary according to concomitant neuropsychological deficits. In the current research those problems are linked to dysfunctions in the interface between executive and linguistic abilities (Ardila 2012, Lê et al. 2012, Ferstl et al. 2005). While there are model-based and empirically verified assessments and therapy programs for aphasic disorders available, the field of 'cognitive communication impairments' has not yet been studied very well in the german-speaking world. In the context of a research project the screening MAKRO (Büttner 2012, 2013, in print), which evaluates the text and discourse level and provides indications for treatment approaches, was developed. In the following, the main features of the screening are outlined and the outcomes of a group study are discussed as well as the resulting implications for possible treatments.

<sup>1</sup> Schön Klinik Bad Aibling, Klinik für Neurologie und neurologische Rehabilitation

<sup>2</sup> LMU München (Dt. Philologie, Studiengang Sprachtherapie); Julia.Buettner@lmu.de

## 1. Einführung: Diskurs- und Textverarbeitung bei kognitiven Kommunikationsstörungen

Menschen mit kognitiven Sprach- und Kommunikationsstörungen können strukturelle Auffälligkeiten im Diskursverhalten und Schwierigkeiten im Verstehen und Produzieren von Texten aufweisen. In der Forschungsliteratur werden auch Auffälligkeiten in der Wortflüssigkeit und Schwierigkeiten bei Aufgaben, die inferentielle Prozesse erfordern, angeführt (z.B. Blyth et al. 2012, Dressler et al. 2004, Zalla et al. 2002). Der Schweregrad der kommunikativen Beeinträchtigung kann in Abhängigkeit vom Läsionsort und den begleitenden neuropsychologischen Defiziten unterschiedlich stark ausgeprägt sein. Auch das Erstellen einer Differentialdiagnose wird durch die Überlagerung von sprachsystematischen und kognitiven Defiziten (wie z.B. nach Schädelhirntrauma) deutlich erschwert (Drechsler 1997, Ferstl et al. 1999). Im Gegensatz zu den Errungenschaften der Aphasologie, die sich in der aktuellen Forschungsliteratur auf psychometrisch gesicherte und modellgeleitete Diagnostikverfahren und Therapieansätze beziehen kann, wurde jedoch bislang kein standardisiertes Verfahren veröffentlicht, das eine differenzierte Beurteilung von kognitiven Sprach- und Kommunikationsstörungen ermöglicht und Implikationen für Therapieansätze bietet (vgl. dazu Ziegler et al. 2008; Spitzer 2009; Hielscher-Fastabend & Jaecks 2010; Heidler 2006:57). Bereits entwickelte Materialien für Textstörungen nach neurologischer Erkrankung sind für die Diagnostik nur bedingt geeignet, da bei diesen keine Normwerte bzw. Vergleichswerte einer Kontrollgruppe vorliegen oder eine differenzierte Beurteilung nur für bestimmte Modalitäten oder einzelne Textsorten gegeben ist.

In einem Forschungsüberblick zu Studien der Verarbeitung von Texten (Büttner, im Druck: 64ff.) zeigte sich, dass sich durchaus Parameter und Prozesse beschreiben lassen, die für die Diagnostik und Therapie von Störungen der Diskurs- und Textverarbeitung als sensitiv erachtet werden können. Obwohl die genannten Forschungsarbeiten im methodischen Aufbau oder der Wahl der Aufgabenstellung nicht immer einheitlich sind, lassen sich Gemeinsamkeiten in der Zielsetzung ausmachen. Viele Arbeiten gehen dem Vorliegen von Leistungsdissoziationen auf Textebene nach. Durch Variation des Texttyps, der Modalität und der jeweiligen makrostrukturellen Anforderung wird versucht, den Verarbeitungsprozess auf Textebene möglichst genau zu erfassen (z.B. Blyth et al. 2012, Body/Perkins 2004, Coelho 2002, Coelho et al. 1995, Ferstl et al. 2005, Lê et al. 2012). Die jeweiligen Fragestellungen beziehen sich dabei auf Unterschiede in der Produktion und Rezeption, im Verste-

hen von impliziter und expliziter Textinformation, Dissoziationen in Abhängigkeit von hochfrequentem und niederfrequentem Skriptwissen oder offenen und geschlossenen Themen (z.B. Boelen et al. 2011, Ausflug planen vs. Zimmer streichen). Wesentliche Überschneidungen lassen sich auch in der Beurteilung der Informativität und der Kohärenz eines Textes ausmachen. In der Textproduktion wird die Anzahl von generierten thematischen Einheiten als zentral erachtet, ebenso wie das Verhältnis dieser essentiellen Sinneinheiten zu irrelevanten Sinneinheiten. In der Analyse der Texte wird dabei meist von zentralen bzw. obligatorischen Propositionen und peripheren Propositionen gesprochen (Dressler et al. 2004, Lojek-Osiejuk 1996, Lê et al. 2012). Vielen Arbeiten ist auch gemeinsam, dass sie in der Beurteilung von expressiven Fähigkeiten nicht nur die propositionale Dichte (*propositional density*, Anzahl der Prädikat-Argument-Struktur pro Satz, Coelho et al. 2005) messen, sondern auch die Verarbeitung von semantischen Inferenzen prüfen. Den verschiedenen Typen von Inferenz (u.a. emotionale, kausale Inferenzen) kommt in der Etablierung von Kohärenz ein besonderer Stellenwert zu, da sie kognitive Prozesse erfordern, welche die propositionale Information der Textbasis mit dem Weltwissen des Rezipienten bzw. Produzenten verknüpfen (z.B. Graesser et al. 2007, Singer 1994). Die Methoden der Textbeurteilung bei neurologischen Erkrankungen unterscheiden zwischen einer mikrolinguistischen und makrolinguistischen Analyse (Übersicht in Büttner, im Druck: 69ff.). Für die Beurteilung der Text- und Diskursebene wird in Zusammenhang mit kognitiven Kommunikationsstörungen die makrolinguistische Analyse bevorzugt, da hierbei die Analyse der lokalen und globalen Kohärenz, die Verwendung von kohäsiven Mitteln und die thematische Strukturierung von besonderem Interesse sind (Chapman et al. Coelho et al. 2005, Lê et al. 2012). Die Erhebung der makrostrukturellen Fähigkeiten wird auf Diskursebene durch Beurteilungsbögen und Ratingverfahren in strukturierten Interviews, aber auch im monologischen Diskurs durchgeführt. Beide Diskurstypen haben ihre Vorzüge. Im elizitierten Monolog zu einem umgrenzten Thema kann der geforderte Informationsgehalt objektiv erfasst werden, während im interaktiven Diskurs soziopragmatische Beeinträchtigungen deutlicher sichtbar werden. Hohe diagnostische Relevanz wird der Realisierung der thematischen Organisation und der Vollständigkeit der Informationsstruktur der unterschiedlichen Texttypen zugesprochen. Bei narrativen Texten bezieht sich dies auf die Anzahl an obligatorischen Propositionen sowie auf die episodische Strukturierung nach Situierung, Komplikation und Resolution. Bei prozeduralen Texten (z.B. Handlungsabläufe wie *Kaffee kochen, Reifen wechseln, Reise planen*)

## Originalbeitrag: Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO

bestimmt schwerpunktmäßig die Auswahl und die Sequenzbildung von relevanten Handlungsschritten die interne Struktur (Lê et al. 2012). Bei narrativen Texten fällt Patienten mit kognitiven Kommunikationsstörungen nicht nur das Generieren der wesentlichen Sinneinheiten schwer, auch die Zuordnung dieser obligatorischen Propositionen zu den Komponenten der narrativen Superstruktur (Situierung – Komplikation – Resolution) ist problematisch (für SHT siehe z.B. Blyth 2012, Drechsler 1997). Narrative Texte sind zudem deshalb anspruchsvoll, da sie das Verknüpfen von aufeinanderfolgenden Ereignissen und die Einnahme einer Perspektive erfordern, um die Motivation von Handlungsträgern zu erkennen und zu evaluieren. Gerade dieser flexible Wechsel der Perspektive im Textgenerierungsprozess fällt Patienten mit kognitiven Kommunikationsstörungen besonders schwer (Büttner, im Druck; 192f., Dressler et al 2004, Ferstl et al. 2005). Im Textverstehen wird die Lesedauer, vermehrtes «Hängenbleiben» im Text oder «Zurückspringen» zu bereits gelesenen Passagen sowie die Anzahl an Metakomentaren (z.B. assoziative Äußerungen) während des Lesens beurteilt. In der Rezeptionsleistung wird darauf Bezug genommen, ob Fragen zum Textinhalt korrekt und vollständig beantwortet werden können. Dabei wird auch erhoben, ob sich die Fragen auf implizite oder explizite Textinformation, Detailwissen (Mikroproposition) oder zu inferierende Informationen beziehen (wie die Pointe oder Motivationen von Handlungsträgern).

### 2. Diskurs- und Textverarbeitung bei Exekutivstörungen

Als zentrales Merkmal wird in vielen Arbeiten in Zusammenhang mit kognitiv bedingten Diskurs- und Textstörungen die fehlende Kohärenz und Einschränkungen in der intentionalen Sprachplanung an den situativen Kontext thematisiert (z.B. Chapman et al. 1992, Dressler et al. 2004, Sirigu *et al.* 1998, Zalla et al. 2002). Als mögliche Ursache werden Störungen von kognitiven Teilleistungen diskutiert bzw. im Übergangsbereich zwischen kognitiven und sprachsystematischen Fähigkeiten verortet, die dann zu einer defizitären Textorganisation führen können. Eine gestörte makrostrukturelle Planung wird als Folge einer unzureichenden inhibitorischen Kontrolle (Lojek-Osiejuk 1996: 73), als eingeschränkte Kapazität des Arbeitsgedächtnis, als gestörte Figur-Grund-Wahrnehmung (Dressler et al. 2004) oder auch als Resultat eines fragmentarischen internen Schemas (Marini et al. 2005) erklärt. Gestützt werden die Annahmen zum Zusammenhang von kognitiven und sprachlichen Fähigkeiten auf makrostruktureller Ebene durch Untersuchungen, die statistisch bedeutsame Bezüge zwischen gestörten Exe-

kutivfunktionen und Beeinträchtigungen in der Verarbeitung von Texten nachweisen konnten (Ferstl et al. 2005). Die enge Verbindung von exekutiven und makrostrukturellen Fähigkeiten zeigt sich auch in einer konzeptuellen Nähe.

Als dysexekutive Symptome werden in der einschlägigen Literatur ein eingeschränktes Monitoring und Selbstkorrekturverhalten, Störungen in der Handlungsplanung und im Sozialverhalten, eine reduzierte kognitive Flexibilität sowie verminderte Aufmerksamkeits- und Arbeitsgedächtnisleistungen (Boelen et al. 2011, Coelho et al. 1995, Ferstl et al. 2005, z.B. Müller & Münte 2009) angeführt. Als charakteristisches Merkmal von gestörten Text- und Diskursstörungen gilt das Fehlen von zielgerichteten Prozessen in der kontextspezifischen Sinnkonstruktion, die zu Störungen in der Organisation der Informationsstruktur führen können. Ein hoher Anteil von irrelevanten Sinneinheiten und Metakomentaren (periphere Propositionen), Fehler im temporalen und kausalen Anordnen von Propositionen und ein fragmentarisches Skriptwissen können somit als eine Störung der Sprachplanung aufgefasst werden. In neueren Arbeiten wird vermehrt der Fokus auf funktionale Beeinträchtigungen gelegt. So findet sich beispielsweise bei Boelen und Kollegen (2011) als zentrales Einschlusskriterium das Vorliegen von Exekutivstörungen, während die Ätiologie und Lokalisation der Hirnschädigung weniger zentral für die Untersuchung der Textebene und des Skriptwissens angesehen wurden.

### 3. Evaluation des Screeningverfahrens MAKRO zur Text- und Diskursverarbeitung

Für die Einordnung von kommunikativen Beeinträchtigungen auf Text- und Diskursebene ist es erforderlich, einzelne Teilprozesse in der makrostrukturellen Verarbeitung erfassen zu können. Im Forschungsüberblick wurde auch deutlich, dass sich die Entwicklung von geeignetem Diagnostikmaterial sowohl auf textlinguistische als auch neuropsychologische Modellvorstellungen beziehen muss, um funktionale Störungen, aber auch erhaltene Komponenten in der Textverarbeitung greifbar zu machen. Daher wurden für die Entwicklung von MAKRO relevante textlinguistische Parameter und psycholinguistische Modellvorstellungen zur Verarbeitung von Makrostrukturen berücksichtigt. Im Rahmen einer Gruppenstudie (n=41) wurden erste Erfahrungen mit MAKRO bei Patienten mit Exekutivstörungen (n=21) gesammelt und diese einer gematchten Kontrollgruppe (n=20) ohne neurologische Erkrankungen gegenüber gestellt (Büttner, im Druck). Nachfolgend werden die

## Originalbeitrag: Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO

Grundzüge des Verfahrens vorgestellt und auf einige zentrale Ergebnisse der Gruppenstudie eingegangen.

### 3.1 Psycholinguistische Implikationen der Materialentwicklung

Eine differenzierte Beurteilung der Textverarbeitung in verschiedenen Modalitäten ermöglicht es zudem, spezifische und individuelle Ziele für die Behandlung von Text- und Diskursstörungen anhand der Diagnostikergebnisse abzuleiten. In linguistischen Modellvorstellungen wird von einer semantisch-thematischen Textbasis ausgegangen, die sich aus propositional codierten Themen und Subthemen zusammensetzt und als hierarchisches Gebilde spezifische Planungsprozesse erfordert (z.B. Brinker 2005, Quaestio Ansatz von Stutterheim/Klein 1987, 2008). Makrostrukturelle Planungsprozesse beinhalten auch die wichtigsten konzeptuellen Domänen (Angaben zu Entitäten, Handlungen oder Eigenschaften, Angaben zu Ort, Zeit und Modalität) und bedingen zudem auch eine mikrostrukturelle Umsetzung auf propositionaler Ebene (von Stutterheim/Klein 2008). Die Generierung eines Textes kann daher als eine kommunikative Aufgabe beschrieben werden, die sowohl die Auswahl als auch die Linearisierung von Sachverhaltswissen umfasst. Zur erfolgreichen Bewältigung der Diskurs- und Textplanung gehören darüber hinaus auch Prozesse der Perspektivierung und die Anpassung des «Granularitätsniveaus» des Sachverhaltswissens an den situativen Kontext (vgl. Klein/von Stutterheim 1992, von Stutterheim/Klein 2008). Bereits in den früheren psycholinguistischen Modellvorstellung (van Dijk 1980, Situationsmodell van Dijk & Kintsch 1983) werden kognitive Prozesse als Voraussetzung für das Generieren von kohärenten Texten vorgeschlagen. Zu diesen sogenannten *Makroregeln* werden das *Selektieren*, *Generalisieren* und *Konstruieren* von propositionalen Einheiten gezählt. Als *Generalisieren* fasst van Dijk reduktive Prozesse auf, die z.B. notwendig sind, um Zwischenüberschriften oder Titel zu einer Textvorlage zu produzieren. *Konstruieren* zielt dabei auf die spezifische Superstruktur des jeweiligen Texttyps ab, was sich beispielsweise bei narrativen Texten auf die episodische Strukturierung (Situierung, Komplikation, Resolution, Evaluation) bezieht. Im Situationsmodell (van Dijk & Kintsch 1983) werden auch Inferenzen als integrale Prozesse für die Etablierung der Textkohärenz angesehen. Eine Inferenz kann dabei als die inferierte Information, aber auch als Prozess aufgefasst werden (Guthke 1992: 31), der während der Textverarbeitung automatisch oder strategisch abläuft. Mit inferentiellen Propositionen kann der explizite Textinhalt durch zusätzliche Information aus dem semantischen oder episodischen Wissen angereichert werden (Linke/Nussbaumer 2001: 439, Rickheit/Strohner 2003: 566, Singer 1994: 480).

Kohärenz kann dann entstehen, wenn die Anwendung

von Selektions-, Integrations- und Konstruktionsprozessen situationsspezifisch und nach *Relevanzaspekten* (van Dijk 1977, Sperber/ Wilson 1981) ausgerichtet ist. Die entwickelte Testkonstruktion orientiert sich folglich an psycholinguistischen Texttheorien (u.a. der Konzeption der Makroregeln, van Dijk 1980), die von selektiven Teilfunktionen im Aufbau einer kohärenten Textbasis ausgehen. Die einzelnen Testbereiche (*Textrezeption*, *Textproduktion*, *Prozedurale Sequenzen*, *Inferenzen*) sind daher in unterschiedlicher Gewichtung nach der Selektion von *obligatorischen Propositionen*, deren *temporaler* und *kausaler Anordnung*, und der Integration von *semantischen Inferenzen* ausgerichtet. In den einzelnen Untertests wird überprüft, ob Propositionen nach Relevanzaspekten ausgewählt und in Bezug zu einem übergeordneten Thema nach temporalen und kausalen Aspekten organisiert werden können. Aufgrund der strukturellen Unterschiede von prozeduralen und narrativen Texten ist davon auszugehen, dass dabei auch unterschiedliche kognitive Prozesse gefordert werden. Prozedurale Texte beziehen sich vorrangig auf das Selektieren und Sequenzieren von obligatorischen Propositionen, während sich narrative Texten durch eine episodische Strukturierung und inferentielle Prozesse (Graesser et al. 1994, 2007) ausweisen. Das Diagnostikmaterial wurde daher danach differenziert, welche makrostrukturellen Teilfunktionen in den jeweiligen Untertests des Verfahrens schwerpunktmäßig gefordert sind. Es prüft das *Selektieren*, *Sequenzieren* und *Inferieren* von Propositionen sowie mit der Fähigkeit des Konstruierens die texttypspezifische Organisation von Propositionen.

### 3.2 MAKRO – ein Screeningverfahren für die Text- und Diskursverarbeitung

Untertest	Aufgabe	Bewertung	Punkte
<b>Textrezeption</b>	Single-Choice-Fragen	Thematisch richtige Proposition	30
<b>Textproduktion</b>	Mündlich expressiv, (4- und 8-stufige Bildsequenzen)	Anzahl obligatorischer Propositionen und Inferenzen	31 (A) 30 (B)
<b>Inferenzen</b>	Schriftliches Generieren von Schlussfolgerungen im Kontext	Generierung einer inferentiellen Proposition	30
<b>Prozedurale Sequenzen</b>	Handlungsabfolgen sequenzieren	Kausales + temporales Sequenzieren von Propositionen	30

Abbildung 1: Übersicht über die Testbereiche von MAKRO (Büttner, im Druck)

## Originalbeitrag: Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO

Das Verfahren testet makrostrukturelle Fähigkeiten für verschiedene Textsorten (narrative, prozedurale Texte) in unterschiedlichen Modalitäten und zielt darauf ab, eine möglichst differenzierte Einordnung von Text- und Diskursstörungen zu erreichen. Mit diesem methodischen Vorgehen ist es daher möglich, selektive Störungen von Textverarbeitungsprozessen zu prüfen. Auch quantitative können qualitativen Abweichungen in den produzierten Texten gegenübergestellt werden. Ebenso lassen sich in einem gewissen Umfang auch modalitätsspezifische Abweichungen (Bild vs. Text, Rezeption vs. Produktion) empirisch prüfen.

Der Untertest *Textrezeption* besteht aus einem narrativen Text, zu dem 10 Single-Choice-Fragen entwickelt wurden. Diese beziehen sich auf das Selektieren, Inferieren und Generalisieren von propositionaler Information beim Textverstehen. Im Untertest *Inferenzen* wird bei 10 Kurztexten die Fähigkeit geprüft, kausale Inferenzen (Brückeninferenzen) zu generieren. Bei den *prozeduralen Sequenzen* wird vorrangig die Fähigkeit der Sequenzbildung von Handlungsabfolgen geprüft. Anhand von zwei unterschiedlich langen Bildgeschichten wird im Untertest *Textproduktion* das *Konstruieren* geprüft: In diesem Fall die Produktion von narrativen Texten in Bezug zu einer visuellen Vorlage.

Das Verfahren besteht aus zwei parallelisierten Versionen, um die Fähigkeit in der Eingangsdiagnostik mit der Leistung nach einer Therapiephase von 15 Einheiten vergleichen zu können. Für die Vergleichbarkeit der Leistungsprofile zwischen den Patienten mit kognitiven Sprach- und Kommunikationsstörungen und gesunden Probanden sowie zwischen unterschiedlichen Testzeitpunkten wurde ein Bewertungssystem nach Punkten entwickelt. Dieses Bewertungssystem zielt darauf ab, für jede zu verarbeitende Proposition in der Bearbeitung einer Aufgabenstellung einen Punkt zu vergeben (siehe Abb.2)

Das entwickelte Material wurde für alle Untertests sowohl für die Version A (*Eingangsdiagnostik*) und Version B (*Abschlussdiagnostik*) – unter Berücksichtigung der Natürlichkeit der Texte – nach Möglichkeit parallelisiert.

### 3.2.1 Beschreibung des Untertests Textrezeption

Die beiden narrativen Texte der Textverständnisaufgaben wurden in Bezug auf die Anzahl der Wörter, die Verteilung der Wortarten und die Anzahl der Propositionen parallelisiert (vgl. Vorgehen bei Klingenberg 1997: 89f. zur Textverarbeitung bei Aphasie). Sie zeigen jeweils auch eine vergleichbare thematische Progression und beschreiben ein Missgeschick im Rahmen einer alltäglichen Situation. Die thematische Struktur setzt sich aus

mehreren Episoden zusammen, die sich jeweils aus Situierung, Komplikation, Auflösung und Evaluierung zusammensetzen. Der Text wird zunächst – ohne zeitliche Beschränkung – zum Lesen vorgelegt. Im Anschluss daran werden den Probanden Fragen ausgehändigt, die sich explizit oder implizit auf den Textinhalt beziehen und ohne erneute Konsultation der Textvorlage beantwortet werden sollen.

### Textauszug aus dem Untertest Textrezeption

So ein Mist! Warum musste Luzia das heute nur passieren? Gerade wo sie heute um 09.00 Uhr in München ein wichtiges Bewerbungsgespräch in einer beliebten Firma hatte! Nach nur 50 Metern war die Fahrt durch den dichten Novembernebel mit ihrem Auto zu Ende gewesen. Sie hatte einen Platten! Ein Reifenwechsel würde zu viel Zeit in Anspruch nehmen und sie käme viel zu spät. Jetzt half nur noch die Bahn! (...).

	Warum hatte Luzia heute Glück im Unglück?
	a) weil heute kein Stau auf der Autobahn war.
X	b) weil sie gerade noch den Zug erwischt hatte.
	c) weil sie bei einer beliebten Firma arbeitet.

Abbildung 2: Beispiel aus dem Untertest **Textrezeption**.

### 3.2.2. Textproduktion

Für den narrativen Texttyp wurden zwei Bildgeschichten verwendet, die eigens für die Studie entwickelt und gezeichnet wurden.<sup>1</sup>

Die vier- und achtstufige Bildgeschichten sind hinsichtlich der zu verarbeitenden Informationsmenge und -komplexität unterschiedlich konzipiert.

Mit den Ergebnissen der gesunden Kontrollgruppe wurden für die zu produzierenden Bildgeschichten Propositionenlisten verifiziert, die nach der Frequenz der Nennung in obligatorische Propositionen und optionale Propositionen unterteilt wurden. Als Proposition werden semantisch-thematische Sinneinheiten aufgefasst, die durch das finite Verb der Satzstruktur konstituiert werden. Die Analyse zur Ermittlung von obligatorischen Propositionen wurde analog zu etablierten Verfahren (Dressler *et al.* 2004, Regenbrecht *et al.* 1992) durchgeführt, in denen eine Unterscheidung nach für die Textkonstitution relevanten (obligatorischen, zentralen) im Gegensatz zu irrelevanten (peripheren) Sinneinheiten getroffen wird.

Während obligatorische Propositionen durch eine Nennungshäufigkeit von über 80% pro Proposition gekennzeichnet sind und die thematische Progression vor-

<sup>1</sup> Ein großer Dank geht an T. Dauenhauer für die Zeichnung der Bildsequenzen.

## Originalbeitrag: Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO

anbringen, werden mit peripheren Propositionen thematisch falsche Propositionen und Konfabulationen beschrieben. Periphere Propositionen werden mit Metakommentaren zu den nicht-thematischen Propositionen gezählt, da bei beiden eine thematische Abweichung vorliegt und somit die Textwelt verlassen wird. In der Instruktion wird ausdrücklich darauf hingewiesen, eine Geschichte zu erzählen und nicht nur die Bilder zu beschreiben. Die Probanden sollen so erzählen, dass auch jemand, der die Bilder nicht kennt, die Geschichte versteht. Für die Bildergeschichten wurde anhand des oben beschriebenen Verfahrens die Anzahl der obligatorischen Propositionen ermittelt. Die kürzere Bildgeschichte A1 «Katze und Fisch» (mit 4 Bildern und einer Episode) kommt auf 12 obligatorische Propositionen, die längere Bildgeschichte («Geldbeutel», bestehend aus 8 Bildern, 2 Episoden) hat 19 obligatorische Propositionen. In der Parallelversion wurde für die längere Geschichte nur bei 18 Propositionen eine Nennungshäufigkeit von > 80% ermittelt, daher liegt hier eine leichte Abweichung zur Testversion A vor (vgl. Abb. 1).

### 3.2.3 Beschreibung des Untertests Inferenzen

Im Untertest Inferenzen wird anhand von Kurztexten mit kausaler Diskontinuität geprüft, ob Inferenzen zu kausaler Antezedenz produziert werden können. Die Kurztexte orientierten sich in Bezug auf die Makrostruktur an den Übungen aus der Aphasietherapie zur Inferenzbildung von Engl und Kollegen (1996).

Anhand von 10 Kurztexten wird gefordert, kausale Inferenzen zu generieren, um das Vorkommen eines Ereignisses am Ende des Textes zu erklären. In der Konstruktion der Texte wurde darauf geachtet, dass diese kausale Diskontinuität anzeigen. Die Bewertung der Antworten richtet sich danach, ob eine kausale Inferenz generiert wird. Nur wenn eine Brückeninferenz generiert wird, erhält der Proband die volle Punktzahl. Für die Durchführung der Aufgabe werden die Probanden instruiert, die Kohärenzlücke durch das Schreiben von Sätzen zu schließen. Damit sollte das bloße Nennen von Schlüsselbegriffen vermieden werden, die keine eindeutigen referentiellen Bezüge erkennen lassen.

#### Instruktion zum Untertest Inferenzen:

Nachfolgend lesen Sie kurze Texte. Schreiben Sie in die leere Zeile in 1-2 Sätzen, was in der Zwischenzeit passiert ist.

Herr Weizel beisst in einen knackigen Apfel.

**z.B.:** Dabei fiel ihm eine Plombe aus einem Backenzahn.

Am nächsten Tag muss er zum Zahnarzt.

Abbildung 3: Beispiel aus dem Untertest Inferenzen.

### 3.2.4 Beschreibung des Untertests Prozedurale Sequenzen

Der Test prüft schwerpunktmäßig die Fähigkeit, die einzelnen Propositionen in Bezug auf ihre zeitliche Abfolge zu sequenzieren. Im Untertest Prozedurale Sequenzen sollen auch thematisch relatierte Propositionen von für eine Handlungssequenz irrelevanten Propositionen getrennt werden. Dabei wird die Aktivierung von Skriptwissen zu Alltagstätigkeiten gefordert. Zu Alltagshandlungen werden Propositionenlisten vorgelegt (z.B. *Einkaufen im Supermarkt, Reklamation eines Elektrogeräts*), die in die richtige Reihenfolge gebracht werden sollten. Bei manchen Aufgaben sollen auch unpassende Handlungen herausgefunden und passende Überschriften produziert werden.

#### Titel: Reklamation, Umtausch eines Radios

6	Ersatzgerät erhalten oder wahlweise Geld erstattet bekommen
2	Radio in Originalverpackung einpacken
5	Technischen Schaden zeigen/bemängeln
4	Den Verkäufer von Vortag suchen
3	Zu Elektrogeschäft fahren
1	Originalverpackung und Rechnung holen

Abbildung 4: Beispiel aus dem Untertest Prozedurale Sequenzen.

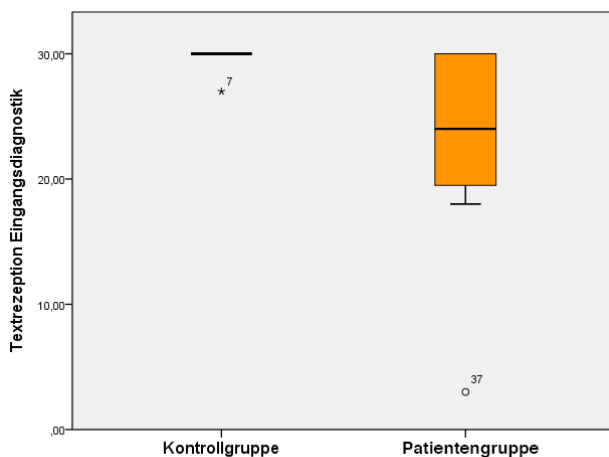
Um die Handlungen in eine passende Reihenfolge zu bringen, sollen Zahlen von 1-6 bzw. 1-8 verwendet werden. Bei einigen Aufgaben (z.B. Planen einer Zugreise) gibt es mehrere Lösungsmöglichkeiten.

### 3.3 Ergebnisse einer Gruppenstudie

Zielgruppe der entwickelten Textdiagnostik sind Patienten mit kognitiven Kommunikationsstörungen. Daher wurden in der Gruppenstudie auch Patienten mit aphasischen Sprachstörungen (AAT: Token Test, Huber et al. 1983) ausgeschlossen. Ausgehend von einem funktionalen Ansatz wurden Patienten mit kognitiven Sprach- und Kommunikationsstörungen bei unterschiedlicher Ätiologie in die Studie eingeschlossen (vgl. Ansatz bei Boelen et al. 2011). Die gruppenbildenden Merkmale waren Einschränkungen in den kommunikativen Fähigkeiten und Beeinträchtigungen in Exekutivfunktionen. Dazu wurden für jeden Patienten sowohl Experteneinschätzungen von sprachtherapeutischer als auch neuropsychologischer Perspektive als auch Werte von Bewertungsskalen (FIM-Wert < 7) erhoben. Bei den neurologischen Probanden wurden die Leistungen im Arbeitsgedächtnis (HAWIE-R; Zahlenspanne vorwärts und rückwärts, Tewes 1991), im Planerischen Denken (Tower of London, Dt. Version, Tucha und Lange 2004)



## Originalbeitrag: Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO



**Abbildung 5:** Ergebnisse des Gruppenvergleichs im Untertest Textrezeption.

und in der Wortflüssigkeit (Regensburger Wortflüssigkeitstest, Aschenbrenner et al. 2000) erfasst, um mögliche Korrelationen zwischen Exekutivfunktionen und den Ergebnissen der Textdiagnostik transparent zu machen. Als Indikatoren und damit als Einschlusskriterium für beeinträchtigte Exekutivfunktionen gelten für die Gruppe der Patienten in mindestens einem der Testbereiche unterdurchschnittliche Werte in den genannten neuropsychologischen Testverfahren.

Im Rahmen der empirischen Studie wurde auch der Frage nachgegangen, ob sich divergierende Leistungen in Abhängigkeit der Komplexität des Materials zeigen und ein Bezug zu einzelnen exekutiven Funktionen (Arbeitsgedächtnis, Wortflüssigkeit, Planerisches Denken) möglich ist. Mit dem Verfahren sollte zunächst eine Trennung zwischen einer beeinträchtigten und unbeeinträchtigten Textverarbeitungsleistung erreicht werden.

Nachfolgend wird auf einige zentrale Fragestellungen der Studie eingegangen.

Der Schwerpunkt wird hierbei auf die Untertests *Textrezeption* und *Textproduktion* gelegt. Eine ausführliche Darstellung und Diskussion der Testergebnisse findet sich bei Büttner (2013, im Druck). Im Rahmen der Testkonstruktion wurde auch durch anschließende inferenzstatistische Verfahren der Grad der Ausprägung der Beeinträchtigung im makrostrukturellen Verarbeiten erfasst. Mit struktorentdeckenden Verfahren wurde geprüft, inwiefern die Variablen der Exekutivfunktionen mit den

Variablen der Textdiagnostik Gemeinsamkeiten aufweisen (Büttner, im Druck: 202ff.).

*Gibt es Unterschiede in der Textrezeptionsleistung zwischen Patienten mit dysexekutiven Symptomen und einer gesunden*

Im Untertest Textrezeption hat die Patientengruppe im Vergleich zu den anderen Untertests von MAKRO das beste Ergebnis ( $M = 23,473$ ;  $sd = 6,874$ ), sie liegt jedoch auch in diesem Testbereich signifikant unter dem Ergebnis der Kontrollgruppe (Mann-Whitney-U-Test;  $z = -4,115$ ;  $p < 0,001$ , einseitig).

Die Patientengruppe zeigt Fehler bei Fragen, die das Verarbeiten von Detailinformationen, also ein Verarbeiten auf mikropropositionaler Ebene, erfordern, aber auch Fehler, die über die Selektion von Propositionen hinausreichen und Inferenzprozesse verlangen. Den Patienten fiel das Abstrahieren und Reduzieren von Textinformation schwer, was sich z.B. durch Schwierigkeiten in der Ableitung des Textthemas oder beim Finden von Titeln und Überschriften zeigte.

In den Korrelationsberechnungen zwischen den Ergebnissen der Textrezeption und den geprüften Exekutivfunktionen lassen sich interessante Dissoziationen ausmachen. Statistisch bedeutsame Bezüge zeigen sich in Messungen zum Wortflüssigkeitsparadigma (semantisch-lexikalisch,  $r = .491$ ,  $p < 0,01$  phonologisch,  $r = .438$ ,  $p < 0,03$ ). Ebenfalls positiv und signifikant korrelieren die Punktwerte in der Textrezeption mit der Arbeitsgedächtnisleistung ( $r = .431^*$ ), jedoch nicht mit der Handlungsplanung ( $r = .372$ ).

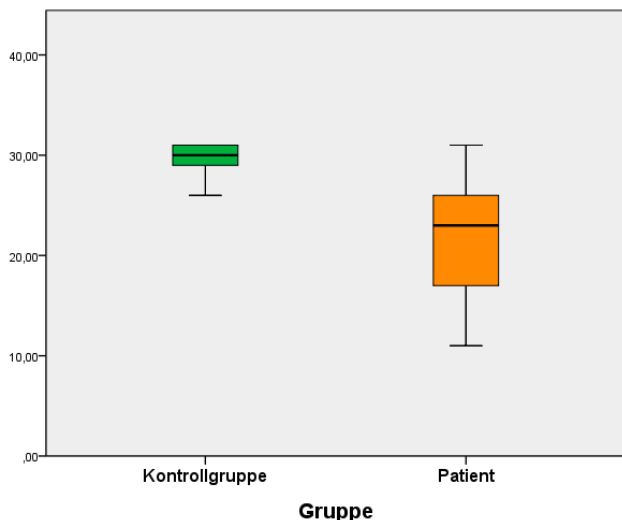
*Durch welche makrostrukturellen Parameter (Anzahl der Teilsätze, obligatorische oder nicht-thematische Propositionen) in der mündlichen Textproduktion unterscheiden sich die Probandengruppen?*

Die Unterschiede zwischen der sprachgesunden Gruppe und der Patientengruppe werden nicht nur anhand quantitativer Merkmale (Anzahl der obligatorischen Propositionen) deutlich, sondern zeigen sich auch durch qualitative Unterschiede. Bei der Patientengruppe werden fehlgeleitete Selektions- und Integrationsprozesse während der Sprachplanung deutlich, wenn das Verhältnis von nicht-thematischen zu obligatorischen Propositionen in Betracht gezogen wird (siehe Abb. 6). Die Ermittlung der Gesamtanzahl von geäußerten Propositionen (Anzahl an Teilsätzen) orientiert sich für eine quantitative Analyse an den produzierten Teilsät-

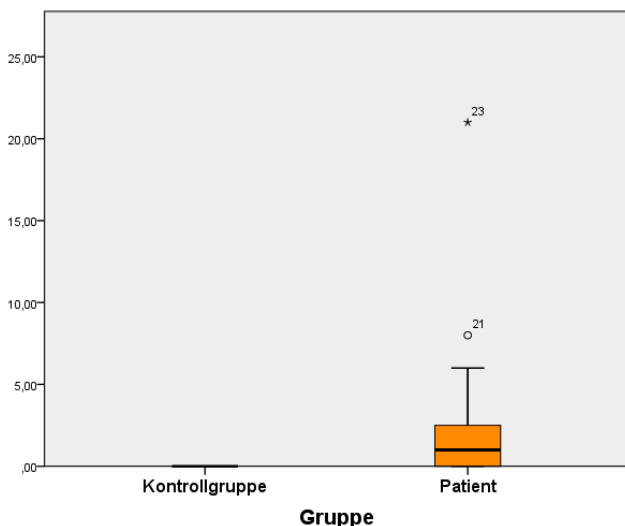
FIM: Functional Independence Measure. Der FIM ist ein Instrument, um die funktionale Selbstständigkeit im Alltag zu erheben. Ein Wert von «7» bedeutet, dass der Patient die bewertete Aktivität, wie Sprechen oder Verstehen, in gewohnter Weise ausführen kann. Ein Wert von «6» verweist auf Einschränkungen, die entweder mit Hilfsmitteln oder einem erhöhten Zeitbedarf kompensiert werden können. Ab einem Wert von «5» ist der Betroffene zur Verrichtung der Alltagsaktivität auf eine Hilfsperson angewiesen. Der Wert «1» zeigt an, dass der Betroffene eine totale Hilfestellung benötigt. Die deutsche Version wurde von der Internationalen Vereinigung für Assessment in der Rehabilitation (IVAR) (1999) herausgegeben.

## Originalbeitrag: Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO

Anzahl von obligatorischen Propositionen



Anzahl von nicht-thematischen Propositionen



**Abbildung 6:** Gruppenvergleich zwischen obligatorischen und nicht-thematischen Propositionen in der Textproduktion.

zen. Für die qualitative Analyse wird zwischen obligatorischen und nicht-thematischen Propositionen unterschieden. In der mündlichen Textproduktion (4- und 8-stufige Bildgeschichten) führt die Messung von obligatorischen Propositionen und Inferenzen zu signifikanten Gruppenunterschieden ( $z = -4,44$ ;  $p < 0,01$ ), während die Gesamtanzahl von produzierten Teilsätzen nicht zwischen den Gruppen trennt. Im Vergleich der Mittelwerte lag kein signifikanter Unterschied vor (Text kurz,  $t(37) = 0,70$ ,  $p < 0,48$ ; Text lang,  $t(37) = 0,42$ ,  $p < 0,67$ ). Periphere Propositionen, wie Perseverationen oder thematische Fehler sowie Metakommentare kommen während der Aufgabestellung nur bei den Patienten vor. Dies bedeutet, dass sich die Patienten und die Kontrollgruppe nicht hinsichtlich der Länge der Texte unterscheiden, sondern qualitative Differenzen zeigen. Auch für den Un-

tertest Textproduktion lassen sich eindeutige Bezüge zwischen den gemessenen Exekutivfunktionen und den makrostrukturellen Parametern festmachen. Hoch signifikante Korrelation lagen – vergleichbar mit der Erhebung des Textverständnisses – mit dem Wortflüssigkeitsparadigma vor (semantisch-lexikalische Wortflüssigkeit  $r = .688^{**}$ ). In Bezug auf die Handlungsplanung (Tower of London) und die Messung der Arbeitsgedächtnisleistung (Zahlenspanne rückwärts) liegen jedoch deutliche Abweichungen zwischen den Untertests vor. Die Anzahl der gemessenen obligatorischen Propositionen korreliert in der narrativen Textproduktion bei einer komplexen Bildgeschichte stark (8-stufig) und signifikant mit der Handlungsplanung ( $r = .611^{**}$ ), jedoch nicht signifikant mit dem Arbeitsgedächtnis ( $r = .391$ ).

#### 4. Diskussion:

Das Diagnostikverfahren erweist sich als geeignet, um Störungen im Verarbeiten von Texten bei dysexekutiven Symptomen differenziert zu erkennen. Dabei wurden auch Leistungsdissoziationen deutlich, die für das Ableiten von Therapiezielen eine wichtige Rolle spielen. Anhand des entwickelten Screenings konnte auch die beeinträchtigte Leistung der Patienten mit Exekutivstörungen von der Leistung gesunder gemachter Probanden getrennt werden. Dabei zeigte sich im Untertest Textproduktion nicht die Länge der produzierten Texte als ausschlaggebend, sondern die Selektion von relevanten Sinneinheiten und deren propositionale Anordnung. Durch die Leistungsbewertung nach einem parallelisierten Punktsystem können Störungen auf Text- und Diskursebene auch in Schweregrad und Modalität unterteilt werden. Die Unterschiede zwischen der Kontrollgruppe und den Patienten in der Textverarbeitung können durch die Parameter obligatorische Propositionen und Inferenzen sensitiv gemessen werden. Dysexekutive Symptome beeinflussen die Verarbeitung von Makrostrukturen in unterschiedlicher Ausprägung. Für das Produzieren eines narrativen Textes scheinen auf Grundlage der bisherigen Daten spezifische sprachliche Planungsprozesse stärker gefordert zu sein als für das Textverstehen. In der Textrezeption wirkten sich Beeinträchtigungen im Arbeitsgedächtnis mehr auf das Decodieren von obligatorischen Propositionen und Inferenzen aus.

Eine Fortführung der Datenerhebung an Patienten mit dysexekutiven Symptomen und einer Normierungstichprobe erscheint vielversprechend und ist für die weitere Zukunft geplant. Damit wird es auch möglich sein, die bisherigen Gütekriterien des Screeningverfahrens weiter auszubauen. Da das Testverfahren auch in

---

**Originalbeitrag: Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO**

den einzelnen Untertests Schwerpunkte in Bezug auf die verschiedenen Prozesse der der Textverarbeitung nimmt, ist es in einer sich anschließenden Therapiephase möglich, diese gezielt zu beüben.

Dabei sollten in die Arbeit mit etablierten Ansätzen der Texttherapie (Claros-Salinas 1993, Freudenberg et al. 1997) auch Behandlungsabsätze integriert werden, die eine Verbesserung der Selbstwahrnehmung sowie das systematische Üben von Strategien fokussieren (z.B. TEACH-M von Ehlhard et al 2005, Ansätze zur Selbstregulation von Kennedy & Coelho 2005).

Der Nutzen von strukturierten Interventionsmethoden konnte bereits durch eine Verlaufsbeobachtung bei einer Gruppe von Patienten mit Exekutivstörungen aufgezeigt werden (Büttner, im Druck: 227f.). Die Fähigkeiten auf Text- und Diskursebene verbessern sich bei dieser Patientengruppe vor allem in den Untertests, die zu Beginn der Behandlung am stärksten betroffen waren und daher am intensivsten trainiert wurden.

**Danksagung:**

Das Screeningverfahren MAKRO entstand im Rahmen meiner Dissertation im Linguistik – Internationalen Promotionsprogrammes (LIPP/Class of Language) an der LMU München. Ein herzlicher Dank geht an Herrn Dr. Ralf Glindemann (Ltg. Neurolinguistik/ Bogenhausen), Frau Prof. Dr. Elisabeth Leiss (LMU München), Herrn Prof. Dr. Wolfram Ziegler (EKN, München) und Frau Dr. Gudrun Klingenberg (Ltg. Neurolinguistik / Schön Klinik Bad Aibling) für wichtige Impulse und Diskussionen während der Arbeitsphase. Tobias Dauenhauer danke ich für eine kritische Durchsicht des Textentwurfs.

---

**Kontakt** | Büttner, Julia; <sup>1</sup>Schön Klinik Bad Aibling, Klinik für Neurologie und neurologische Rehabilitation, Abt. Sprachtherapie <sup>2</sup>LMU München (Dt. Philologie, Studiengang Sprachtherapie); Julia.Buettner@lmu.de

---

**Literatur**

- Ardila, A. (2012): The Executive Functions in Language and Communication. In: R.K. Peach & L.P. Shapiro (Hrsg.) (2012): Cognition and Acquired Language Disorders. An Information Processing Approach. St. Louis: Elsevier Mosby, S. 147-166.
- Aschenbrenner, S., Tucha, O., Lange, K. (2000): Regensburger Wortflüssigkeitstest – RWT. Göttingen – Testzentrale.
- Blyth, T., Scott, A., Bond, A., & Paul, E. (2012): A Comparison of Two Assessments of High Level Cognitive Communication Disorders in Mild Traumatic Brain Injury. *Brain Injury* 26 (3), S. 234-240.
- Body, R. & Perkins, M.R. (2004): Validation of Linguistic Analyses in Narrative Discourse after Traumatic Brain Injury. *Brain Injury* 18 (7), S. 707-724.
- Boelen D.H., Allain P., Spikman, J.M. & Fasotti, L. (2011): Script Generation and the Dysexecutive Syndrome in Patients with Brain Injury. *Brain Injury* 25 (11), S. 1091-1100.
- Brinker, K. (2005<sup>6</sup>): Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden (= Grundlagen der Germanistik 29). Berlin: Erich Schmidt.
- Büttner, J. (2012): Wenn der Sinn fehlt – erste Ergebnisse einer Pilotstudie zum Verarbeiten von Texten bei dysexekutiven Symptomen. In: Schnittstellen. Diskussionsforum Linguistik in Bayern. *Bavarian Working Papers in Linguistics* 1, S. 5-27.
- Büttner, J. (2013): Executive Dysfunction and Text Processing Measurement – Insights into Verbal Planning. *Procedia Social and Behavioral Sciences* 94, S. 88-89.
- Büttner, J. (im Druck): Sprache und Kognition. Diskurspragmatik und Textverarbeitung bei Exekutivstörungen. Reihe Neurokognition (Hrsg. H.M. Müller). Tübingen: Stauffenburg. [erscheint voraussichtlich im Frühjahr 2014].
- Chapman, S.B., Culhane, K.A., Levin, H.S., Harward, H., Mendelsohn, D., Ewing-Cobbs, L., Fletcher, J.M. & Bruce, D. (1992): Narrative Discourse After Closed Head Injury in Children and Adolescents. *Brain and Language* 43, S. 42-65.
- Claros-Salinas, D. (1993): Text Verstehen. Dortmund: Borgmann.
- Coelho, C.A. (1995): Discourse Production Deficits Following Traumatic Brain Injury: A Critical Review of the Recent Literature. *Aphasiology* 9 (5), S. 409-429.
- Coelho, C.A. (2002): Story Narratives of Adults with Closed Head Injury and Non-Brain-Injured Adults: Influence of Socioeconomic Status, Elicitation Task, and Executive Functioning. *Journal of Speech, Language, and Hearing Research* 45 (6), S. 1232-1248.

## Originalbeitrag: Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO

- Coelho, C.A., Grela, B., Corso, M., Gamble, A. & Feinn, R. (2005): Microlinguistic Deficits in the Narrative Discourse of Adults with Traumatic Brain Injury. *Brain Injury* 19 (13), S. 1139-1145.
- Coelho, C.A., Liles, B.Z. & Duffy, R.J. (1995): Impairments of Discourse Abilities and Executive Functions in Traumatically Brain-Injured Adults. *Brain Injury* 9 (5), S. 471-477.
- Dijk, T.A van (1977): Semantic Macrostructures and Knowledge Frames in Discourse Comprehension. In: M.A. Just & P.A. Carpenter (Hrsg.) (1977): *Cognitive Processes in Comprehension*. Hillsdale, New Jersey: John Wiley & Sons, S. 3-32.
- Dijk, T.A van (1980): *Textwissenschaft*. München: dtv.
- Dijk, T.A. van & Kintsch, W. (1983): *Strategies of Discourse Comprehension*. New York: Academic Press.
- Drechsler, R. (1997): Sprachstörungen nach Schädelhirntrauma. *Diskursanalytische Untersuchungen aus textlinguistischer und neuropsychologischer Sicht*. Tübingen: Gunther Narr.
- Drechsler, R. (2007): Exekutive Funktionen: Übersicht und Taxonomie. *Zeitschrift für Neuropsychologie* 18 (3), S. 233-248.
- Dressler, W.U., Stark, H.K., Vassilakou, M., Rauchensteiner, D., Tomic, J., Weitzenauer, S.M., Wasner, P., Pons, C., Stark, J. & Brunner, G. (2004): Textpragmatic Impairments of Figure-Ground Distinction in Right-Brain-Damaged Stroke Patients Compared With Aphasics and Healthy Controls. *Journal of Pragmatics* 36 (2), S. 207-215.
- Engl, E.M., Kotten, A., Ohlendorf, I. & Poser, E. (1996): Sprachübungen zur Aphasiebehandlung. Ein linguistisches Übungsprogramm mit Bildern (Teil 1-2). (= Logotherapia. Materialien und Methoden der Sprachtherapie, Bd.1), Berlin: Marhold.
- Ehhardt, L., Sohlberg, M.M., Glang, A. & Albin, R. (2005): TEACH-M: A Pilot Study Evaluating an Instructional Sequence for Persons with Impaired Memory and Executive Functions. *Brain Injury* 19 (8), S. 569-583.
- Ferstl, E.C., Guthke, T. & Cramon, D.Y. von (1999): Change of Perspective in Discourse Comprehension: Encoding & Retrieval Processes After Brain Injury. *Brain and Language* 70, S. 385-420.
- Ferstl, E.C., Walther, K., Guthke, T., & Cramon, D.Y. von (2005): Assessment of Story Comprehension Deficits After Brain Damage. *Journal of Clinical and Experimental Neuropsychology* 27 (4), S. 367-384.
- Freudenberg, M., Honekamp, A., Mende, M. & Zückner, H. (1997): *Etwas vom Kurs abgekommen. Zur Behandlung von Textstörungen bei Aphasie*. Köln: Steiner.
- Graesser, A.C, Singer, M., & Trabasso, T. (1994): Constructing Inferences During Narrative Text Comprehension. *Psychological Review*, 101 (3), S. 371-395.
- Graesser, A.C., Louwerse, M.M., McNamara, D.S., Olney, A.M., Cai, Z., & Mitchell, H.H. (2007): Inference Generation and Cohesion in the Construction of Situation Models: Some Connections with Computational Linguistics. In: F. Schmalhofer & C. A. Perfetti (Hrsg.) (2007), S. 289-310.
- Guthke, T. (1992): *Psychologische Untersuchungen zu Inferenzen beim Satz- und Textverstehen* (= Dissertation. Mathematisch-Naturwissenschaftliche Fakultät. Humboldt-Universität Berlin). Berlin: Hochschulschrift.
- Heidler, M.-D. (2006): *Kognitive Dysphasien. Differenzialdiagnostik aphasischer und nichtaphasischer zentraler Sprachstörungen sowie therapeutische Konsequenzen*. Frankfurt am Main: Peter Lang.
- Hielscher-Fastabend, M. & Jaecks P. (2010): Textverstehen und Textproduktion in der klinischen Linguistik. In: Blanken, G. & Ziegler, W.: *Klinische Linguistik. Ein Lehrbuch für die Diagnose und Behandlung von erworbenen Sprach- und Sprechstörungen im Erwachsenenalter* (= Mentale Sprachverarbeitung 6. Psycho- und Neurolinguistische Studien. Hrsg. von Gerhard Blanken & Ria De Bleser), Freiburg: Hochschulverlag, S. 329-352.
- Huber, W., Poeck, K., Weniger, D. & Willmes, K. (1983): *Aachener Aphasie Test*. Göttingen: Hogrefe.
- Internationale Vereinigung für Assessment in der Rehabilitation (IVAR) (Hrsg.) (1999): *FIM-Funktionale Selbstständigkeitsmessung. Manual. Version 1.1*. Straubing: o.V.
- Kennedy, M.R.T. & Coelho, C. (2005): Self -Regulation after Traumatic Brain Injury: A Framework For Intervention Of Memory And Problem Solving. In: *Seminars In Speech And Language* 26 (4), S. 244-255.
- Klein, W. & Stutterheim, C. von (1992): Textstruktur und referentielle Bewegung. *Zeitschrift für Literaturwissenschaft und Linguistik* 22, S. 67-93.
- Lê, K., Mozeiko, J. & Coelho, C. (2011): Discourse Analyses: Characterizing Cognitive-Communication Disorders Following TBI. The ASHA Leader. Online unter: <http://www.asha.org/Publications/leader/2011/110215/Discourse-Analyses.html> (Zuletzt abgerufen am 01.04.2014).
- Lê, K., Coelho, C., Mozeiko, J., Krueger, F. & Grafman, J. (2012): Predicting Story Goodness Performance from Cognitive Measures following Traumatic Brain Injury. *American Journal of Speech & Language Pathology* 21 (Suppl.), S. 115-125.

Originalbeitrag: Kognitive Sprach- und Kommunikationsstörungen – Diagnostik und Therapieimplikationen anhand des Screeningverfahrens MAKRO

- Linke, A. & Nussbaumer, M. (2001): Konzepte des Impliziten. Präsuppositionen und Implikaturen. In: K. Brinker et al. (Hrsg.) (2001): Text- und Gesprächslinguistik Linguistics of Text and Conversation. Ein internationales Handbuch zeitgenössischer Forschung. An International Handbook of Contemporary Research (= Handbücher zur Sprach- und Kommunikationswissenschaft 16, 2. Halbband: Gesprächslinguistik). Berlin – New York: de Gruyter, S. 435-448.
- Marini, A., Carlomagno, S., Caltagirone, C. & Nocenti, U. (2005): The Role Played by the Right Hemisphere in the Organization of Complex Textual Structures. *Brain and Language* 93, S. 46-54.
- Marini, A., Galetto, V., Zampieri, E., Vorano, L., Zettin, M., Carlomagno, S. (2011): Narrative Language in Traumatic Brain Injury. *Neuropsychologia* 49, S. 2904-2910.
- Müller, S.V. & Münte, T.F. (2009): Exekutivfunktionen. In: W. Sturm, M. Herrmann & T.F. Münte (Hrsg.) (2009): Lehrbuch der klinischen Neuropsychologie. Heidelberg: Spektrum Akademischer Verlag, S. 480-499.
- Regenbrecht, F., Huber, W. & Glindemann, R. (1992): Zum Verhältnis von Kohärenz und Kohäsion bei Aphasie. In: G. Rickheit, R. Mellies & A. Winnecken (Hrsg.) (1992): Linguistische Aspekte der Sprachtherapie. Forschung und Intervention bei Sprachstörungen. Opladen: Westdeutscher Verlag, S. 111-137.
- Rickheit, G. & Strohner, H. (2003): Inferenzen. In: G. Rickheit, T. Herrmann & W. Deutsch (2003) (Hrsg.), Psycholinguistik - Psycholinguistics. Ein internationales Handbuch. An international Handbook. (= Handbücher zur Sprach- und Kommunikationswissenschaft/ Handbooks of Linguistics and Communication Science/ HSK Band 24) Berlin - New York: de Gruyter, S. 566-577.
- Singer, M. (1994): Discourse Inference Processes. In: M.A. Gernsbacher (1994) (Hrsg.): Handbook of Psycholinguistics. San Diego: Academic Press, S. 479-515.
- Sirigu, A., Cohen, L., Zalla, T., Pradat-Diehl, P., Eeckhout, P. van, Grafman, J., & Agid, Y. (1998): Distinct Frontal Regions for Processing Sentence Syntax and Story Grammar. *Cortex* 34, S. 771-778.
- Sperber, W. & Wilson, D. (1981): Pragmatics. *Cognition* 10, S. 281-286.
- Spitzer, L., Beushausen, U. & Grötzbach, H. (2009): Exekutivfunktionen in der Aphasietherapie – Auswirkungen der Exekutivfunktionen und der Umstellungsfähigkeit auf den Alltag von Patienten mit einer Aphasie. *Forum Logopädie* 3, S. 6-9.
- Stutterheim, C. von & Klein, W. (2008): Mündliche Textproduktion. Informationsorganisation in Texten. In N. Janich (Hrsg.) (2008): Textlinguistik. 15 Einführungen. Tübingen: Narr, S. 200-236.
- Stutterheim, C. von (1987): Einige Prinzipien des Textaufbaus. Empirische Untersuchung zur Produktion mündlicher Texte. Tübingen: Niemeyer.
- Tewes, U. (1991): Hamburg -Wechsler-Intelligenztest für Erwachsene – Revision 1991 (HAWIE-R). Bern – Stuttgart – Toronto: Huber.
- Tucha, O. & Lange, K.W. (2004): Turm von London – Deutsche Version: TL-D. Hogrefe, Göttingen, S. 7-49.
- Zalla, T., Phipps, M. & Grafman, J. (2002): Story Processing in Patients with Damage to the Prefrontal Cortex. *Cortex* 38 (2), S. 215-231.
- Ziegler, W., Ackermann, H., Goldenberg, G., Huber, W., Sedlmeier, C., Schupp, W., Willmes, K., Auff, E., Benke, T. & Annoni, J.M. (2008): Rehabilitation aphasischer Störungen nach Schlaganfall. In: H.C. Diener & N. Putzki (Hrsg.) (2008): Leitlinien für die Diagnostik nach Schlaganfall. Stuttgart – New York: Thieme, S. 920-928.

Originalbeitrag

## Investigating treatment fidelity in a conversation-based aphasia therapy

Heilemann, Claudia<sup>1</sup>; Best, Wendy<sup>1</sup>; Johnson, Fiona<sup>1,2</sup>; Beckley, Firle<sup>1</sup>; Edwards, Susan<sup>1,3</sup>; Maxim, Jane<sup>1</sup>; Beeke, Suzanne<sup>1</sup>

### EN | Abstract

When conducting a speech and language therapy intervention study, one essential focus is the evaluation of outcomes. The therapy itself and its actual delivery, often discussed in the context of treatment fidelity (TF), are aspects which are mostly carried out in the background. According to Cherney et al. (2013), however, TF is an important component of any behavioural treatment study and should therefore be investigated.

This article presents a study of TF embedded in a wider research project that evaluates a new conversation-based therapy for people with agrammatic aphasia and their conversation partners (Beckley et al., 2013; Beeke et al., 2011; Beeke et al., 2014). The therapy is called Better Conversations with Aphasia (BCA). Using the concept of TF, the degree to which BCA was delivered as planned can be measured. A pilot version of a BCA-specific observational fidelity tool was developed, based on a conceptual model of TF (Carroll et al., 2007), on practices reported in the TF literature and on the generic therapy session plans. The results indicate that, in terms of adherence to the therapy content, a high fidelity level (91.9 %) was reached for BCA. This article will also report on the degree to which the therapist showed desired behaviour associated with the delivery of BCA, and findings from an inter-rater reliability investigation of the fidelity tool. It concludes with reflections on the importance and value of TF investigations in speech and language therapy intervention.

<sup>1</sup> Division of Psychology and Language Sciences, University College London, London, UK

<sup>2</sup> Guys & St Thomas' NHS Foundation Trust, London, UK.

<sup>3</sup> School of Psychology and Clinical Language Sciences, University of Reading, UK.

## DE | Zusammenfassung

Die Evaluierung einer Therapiestudie besteht typischerweise aus der Messung des tatsächlichen Therapieeffekts. Die Frage, inwieweit hierbei die ursprünglich vorgesehene Therapie angewendet wurde (z.B. wie in einem Therapiemanual beschrieben), wird hingegen oft vernachlässigt. In der englischsprachigen Literatur wird in diesem Zusammenhang häufig von dem Konzept der treatment fidelity (TF) gesprochen. Das Konzept der TF wird nach Cherney et al. (2013) als eine wichtige Komponente einer verhaltenstherapeutisch orientierten (Sprach-)Therapiestudie angesehen.

In diesem Artikel wird eine Studie vorgestellt, die in ein übergreifendes Forschungsprojekt (Beckley et al., 2013; Beeke et al., 2011; Beeke et al., 2014) eingebettet ist. Das Konzept der TF wird auf eine konversationsorientierte Therapie (Better Conversations with Aphasia; BCA) für Menschen mit Agrammatismus und deren Konversationspartner angewendet. Ziel ist es, den Grad der Konformität mit der ursprünglich vorgesehenen Therapie zu ermitteln. Hierzu wurde eine Pilotversion eines BCA-spezifischen Beobachtungsinstruments (fidelity tool) entwickelt, basierend auf einem speziellen TF-Modell (Carroll et al., 2007), weiteren Methoden aus der TF-Literatur sowie BCA-spezifischen Therapieplänen. Die Analyse der Videoaufnahmen von Therapiesitzungen mithilfe des Beobachtungsinstruments zeigt einen fidelity score von 91,9 %, was laut Literatur auf eine hohe Therapiekonformität hindeutet. Zudem werden Ergebnisse zu der Qualität der Therapieübermittlung vorgestellt sowie die Inter-rater Reliabilität des fidelity tools diskutiert. Insgesamt demonstriert diese Studie die Wichtigkeit und das Potential einer fidelity Evaluation anhand der Anwendung des TF Konzepts auf die BCA Therapie.

## 1. Conversation-based therapy approaches

Aphasia is traditionally defined as an acquired language impairment resulting from a brain lesion in the language dominant hemisphere (Papathanasiou, Coppens, & Potagas, 2013). Generally, aphasia can be divided into several subtypes characterised by typical symptom complexes (Huber, Poeck, & Springer, 2006). One of these symptom complexes consists of the use of simplified syntactic structures, omission of morphology and few if any verbs. Characterised by non-fluent speech production, this symptom complex is known as agrammatism. It will be the focus of this article.

Beeke et al. (2011) suggest that targeting the grammar of everyday conversations may be more effective in improving the spontaneous speech of people with agrammatism than targeting decontextualised language (e.g. the language used when describing pictures). Beeke, Wilkinson and Maxim (2007) demonstrated that the language

Partner Scheme) (McVicker, Parr, Pound, & Duchan, 2009), developed at Connect, the communication disability network, in London, to train volunteers to converse with PWA who experience social isolation (see also Swinburn, McVicker, & Pearce-Willis, 2011). Another is 'Supported Conversation for adults with Aphasia' (SCA; Kagan, 1998a, 1998b; Kagan & Gailey, 1993), which also focuses on conversation partner training. Table 1 shows an overview of selected conversation-based therapy approaches for aphasia.

It should be noted that the examples included in Table 1 do not exclusively target agrammatic speech and that they are not intended to be a complete enumeration of all existing conversation-based therapy approaches. For the most part, studies using these programmes focus on training the CP rather than the PWA (for a systematic review, see Simmons-Mackie, Raymer, Armstrong, Holland, & Cherney, 2010).

**Table 1:** Conversation-based therapy approaches for aphasia (chronologically ordered)

Therapy approach	Abbreviation	Authors (Years)
«Conversation Coaching»	-	Holland (1991, 1997)
«Communication Partners»	-	Lyon et al. (1997)
«Supported Conversation for adults with aphasia»	SCA	Kagan (1998a, 1998b), Kagan & Gailey (1993)
«Supporting Partners of People with Aphasia in Relationships and Conversation»	SPPARC	Lock et al. (2001a)
«Conversation Partner Scheme»	-	McVicker et al. (2009)
«Partners van Afasiepatiënten Conversatie Training» (Dutch adaptation of SPPARC)	PACT	Wielert & Wilkinson (2012)
Swedish adaptation of SPPARC (conversation partner training)	-	Saldert, Backman, & Hartelius (2013)

output produced by speakers with agrammatism in everyday conversations is different to speech elicited by a task-based language assessment, and they therefore suggest that focusing therapy on elicited speech may be limiting generalisation.

Since the 1990s, researchers have developed and tested various approaches that target aphasic conversation directly. So-called conversation-based (sometimes referred to as interaction-focused or pragmatic-oriented) therapy approaches focus on the conversation behaviour of conversation partners (CPs) of people with aphasia (PWA), or the CP and the PWA as a couple, with the objective of enhancing daily, natural conversations, by practising the use of conversation strategies (for a review, see Wilkinson & Wielert, 2012). One example is the «Conversation

### The Better Conversations with Aphasia therapy programme

Beeke and colleagues (Beckley et al., 2013; Beeke et al., 2011; Beeke et al., 2014) have designed a conversation-based therapy approach, called BCA, based on SPPARC (Lock et al., 2001a). SPPARC is a conversation training that applies a qualitative research method, called Conversation Analysis. It consists of three progressive main steps: (1) to raise awareness of the broad idea of conversation and conversation behaviour (such as «the aim of a turn» or «overlapping talk») in relation to aphasia, (2) to raise awareness of a CP's own conversation behaviour, and (3) to facilitate the identification and use of strategies for change (e.g. waiting until the PWA's turn is finished before speaking). The key techniques included in



the therapy (video feedback, role-play and conversation activities) are based on a model of experiential learning (Kolb, 1984; see Beckley et al., 2013 for an overview). Several studies found mostly qualitative indicators for SPPARC's efficacy (Burch, Wilkinson, & Lock, 2002; Lock et al., 2001a, 2001b; Wilkinson, Bryan, Lock, & Sage, 2010) and showed that changes in a CP's behaviour could lead to indirect changes in a PWA's conversation. SPPARC is designed to be delivered either in a group setting (that is, a group of CPs) or with the PWA and the CP together as a couple; either way, it focuses on directly changing the CP's conversation behaviour.

**Table 2:** Session goals of the BCA therapy programme

Session #	The overall goal(s) of the session is/are...
1	...to raise the dyad's overall awareness of conversation.
2	...to raise the dyad's awareness of different aims of turns.
3	...to raise the dyad's awareness of repair in general AND to help the dyad to identify their own patterns of repair.
4	...to identify patterns of turn building in the PWA's own conversation AND for the PWA to select strategies for change and to experience them within a structured task.
5	...to identify patterns of turn building in the CP's own conversation AND for the CP to select and practice strategies for change.
6	...to facilitate the identification and implementation of strategies for change in relation to topic.
7	...to facilitate the implementation of strategies for change.
8	...to support the dyad to implement strategies for change.

BCA uses Kolb's adult learning model (1984) as its basis, as does SPPARC. However, Beeke and colleagues created two significant changes from SPPARC in their BCA therapy programme. First, they focus on conversational difficulties resulting from agrammatism instead of aphasia in general (Beeke et al., 2011). Secondly, in BCA, the PWA is an active participant throughout all sessions. Instead of solely targeting the CP (as is common in most of the previous research in this area), BCA is designed to train a PWA as well as their CP to use conversation strategies. This was done in order to discover whether a PWA can learn from conversation therapy (see Beckley et al., 2013). This adaptation led to the creation of a therapy session focused solely on strategies for the PWA (e.g. use of gesture, key words, writing and drawing in a conversational

turn), plus linked therapy handouts and activities. BCA is designed as a programme of 8 weekly sessions of around 1.5 hours in length taking place at the clients' home (Beckley et al., 2013). The overall aim of BCA is to educate the participants about the effects of agrammatism on conversation, and teach conversation strategies to allow a PWA to produce more successful turns, which should increase mutual understanding between a dyad (i.e. a PWA and their frequent CP). The specific goals for each BCA therapy session are illustrated in Table 2. The therapy programme itself is freely available as part of an e-learning package on UCLeXtend (<https://extend.ucl.ac.uk>). The change in conversations associated with the therapy is expected to be both quantitative and qualitative in nature (Beeke et al., 2011).

In summary, conversation-based therapy programmes such as CP training have become popular clinically. Yet they can be regarded as complex interventions (Beeke et al., 2014; Cherney et al., 2013), with interacting components and individually variable outcomes. Thus, there is a need to explore the delivery of therapy (see also Medical Research Council-guidelines: Craig et al., 2008), especially with regard to the research quality of an intervention study. Cherney et al., 2013, in their paper about the methodological quality of studies on CP training in aphasia, speak of treatment fidelity (TF) as an important element of a behavioural treatment study. This concept will be explored next.

## 2. The concept of treatment fidelity (TF)

When building an evidence base for a new and complex behavioural treatment – like BCA – researchers need to report on the treatment itself (Craig et al., 2008) as well as the therapy outcomes. In the context of designing, implementing and evaluating treatments, one key element is TF (Hennessey & Rumrill, 2003), a measure of the reliability of the provision of a treatment (Hinckley & Douglas, 2013). Literature on the concept of TF has been published in different research areas, such as psychology, education and medicine. Particularly in the last 20 to 30 years many researchers have developed fidelity assessments (e.g. Bellg et al., 2004; Carroll et al., 2007; Waltz, Addis, Koerner, & Jacobson, 1993; Yeaton & Sechrest, 1981). The literature discussed in this article is predominantly from studies that investigate aspects of TF from the field of psychology, where the concept originated (Di Rezze, Law, Gorter, Eva, & Pollock, 2012), because this field has most in common with conversation-based speech and language therapy. Compared to this rich body of research, little literature on TF exists in the area of speech and language therapy studies. This may

be because TF issues are usually reported in the context of higher-level study designs such as randomised control trials (see e.g. Godfrey, Chalder, Ridsdale, Seed, & Ogden, 2007). Such trials are difficult to design and execute in speech and language therapy and are therefore relatively rare (Cherney et al., 2013). Some papers, however, do report on some aspects of TF (e.g. monitoring and documenting fidelity) especially when assessing complex speech and language therapy interventions (see e.g. Adams, Lockton, Gaille, Earl, & Freed, 2012). Last year, Hinckley & Douglas (2013) published the first review of the importance of TF and the frequency with which it is reported in studies related to aphasia treatment. This underlines the growing attention to TF in the field.

## 2.1 Definition

TF is a term that encapsulates a concept originally known as treatment integrity. It assesses whether a treatment or therapy was delivered as intended (Hennessey & Rumrill, 2003). Other terms to describe TF are procedural reliability, intervention fidelity, implementation fidelity, program(me) fidelity, treatment adherence process research and therapist's or clinician's adherence or competence. This range of terminology reflects variability in definitions of TF and leads to confusion in terms of what exactly should be measured when it is assessed (Nelson, Cordray, Hulleman, Darrow, & Sommer, 2012). In this article, the term 'TF' will be used consistently.

The specific term treatment fidelity was first employed by Moncher and Prinz (1991). Lichstein, Riedel, and Grieve (1994) added two aspects to previous definitions: receipt of treatment (i.e. that the client understands and uses the skills that are delivered) and enactment (i.e. that the client implements these skills into daily life). This addition of client-focused behaviour enlarges the traditional understanding of treatment integrity related to adherence (Were all components of the therapy delivered?) and competence (In what way has the therapy been delivered?).

## 2.2 Measurement

Methods of measuring TF vary according to intervention type (see e.g. Chan et al., 2004; Kiran & Thompson, 2003; Lichstein et al., 1994). In general, TF measurement can take place either once (e.g. at an initial stage of a new therapy or at any point after implementation), or repeatedly (Bond, Evans, Salyers, Williams, & Kim, 2000).

Fidelity measures also vary in their aims. For example, researchers might aim to document adherence, or to discriminate treatments which share similarities, to synthesise a body of research, or to identify core ingredients of a therapy (e.g. specific treatment targets or therapeutic techniques).

Common fidelity tools include direct observation, observer and self-report check-lists, indirect observation via video and ratings by experts based on documentation data (Kaderavek & Justice, 2010; Mowbray et al., 2003; Resnick et al., 2005). Fidelity tools are easier to develop if «detailed practice manuals» for a therapy exist (Bond et al., 2000: p. 78).

In order to investigate TF as precisely as possible, many researchers use a multidimensional approach to fidelity, including structural elements (adherence to the therapy programme) and procedural aspects (quality of the delivery, emotional climate in a session, therapy principles) (e.g. Hasson, Blomberg, & Dunér, 2012; Odom et al., 2010).

The study presented in this article is underpinned by benchmarks in relation to fidelity procedures, synthesised from the TF literature (see Table 3).

## 2.3 Conceptual model of TF

The framework created by Carroll et al. (2007) is useful for integrating TF into the research process and describing the ingredients and potential influencing aspects of TF (Hasson, 2010). This so-called Implementation Fidelity Framework (IFF), presented in Figure 1, includes traditional components of TF (e.g. adherence to therapy content) as well as so-called moderating factors, which are expected

**Table 3:** Summary of suggested benchmarks in relation to fidelity procedures (based on a synthesis of TF literature)

Question regarding fidelity procedures	Suggestion derived from the TF literature
Amount of therapy sessions to be checked?	15% - 40%, randomly selected
Type of rating scale used in order to conduct a fidelity check?	Likert-type scale or occurrence/non-occurrence
Characteristics of rater(s)?	Trained, independent, familiar with the intervention
«High» fidelity level?	80% and above
Amount of sessions to be checked by a second rater?	10-30%
Acceptable level for inter-rater percentage agreement?	70% and above
Computing inter-rater reliability (IRR) of a fidelity tool?	Intra-class correlation, Kappa

ted to influence the degree of overall TF. It served as a foundation for the present study. The IFF components will now be outlined.

According to the IFF, the core of TF is adherence, with subcategories content, coverage, frequency and duration. The assessment of adherence involves ascertaining whether the main elements of an intervention have been implemented as intended and if the participants received these elements as often and for as long as designed by the developers of a therapy. Adherence may be affected by moderating factors. Potential moderating factors include the complexity of the intervention, facilitation strategies such as training of therapists or preparation of a manual, the quality of therapy delivery and participant responsiveness.

The link between TF and therapy outcomes is visualised with the help of a broken line (see Figure 1). This indicates that the relationship between a treatment and its outcomes is external to TF, although, according to Carroll et al. (2007), the degree of TF can in theory affect the outcomes.

The notation «identification of «essential» components» (see Figure 1) implies that the aspect of programme differentiation in the model, other than suggested in the literature (e.g. Moncher & Prinz, 1991), is regarded as separate from TF.

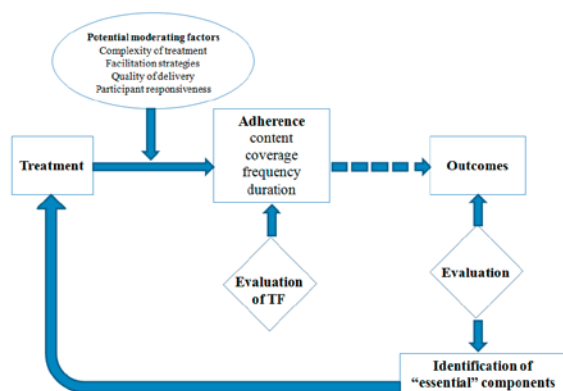


Figure 1: IFF (adapted from Carroll et al., 2007: p.4)

### 3. Research questions

As outlined above, the evaluation of TF is an important part of the methodological quality of a study targeting behaviour change. With regard to future investigations of conversation-based therapy approaches, it is therefore necessary to develop adequate tools to measure TF.

A multidimensional fidelity tool was designed, based on the first author's review of the TF literature, with the intention of capturing as complete a representation of TF as possible for the research project evaluating BCA. The

IFF (Carroll et al., 2007) served as a conceptual model in developing the following research questions:

1. To what degree were the planned components of the BCA therapy programme delivered to the participants of the main research project? [adherence]
2. To what degree did the behaviour of the therapist reflect desired BCA intervention principles? [therapy delivery]

## 4. Method

The data for this study stem from a research project (Beeke et al., 2011) executed at University College London (UCL) in which the effectiveness of the BCA therapy programme was tested. The study design of that project comprised a case series evaluation (for details see Beckley et al., 2013).

### 4.1 Participants

Participants were recruited for the main project from speech and language therapists working in the National Health Service, aphasia support groups, private speech and language therapists and university aphasia clinics. Seven dyads from the main project, in each case one person classified as having agrammatic aphasia and his or her significant other (spouse or family member), were analysed in the present study. All people with aphasia (PWA) had had a left hemisphere stroke at least 6 months prior to involvement in the project. They were classified as agrammatic according to their spoken output in conversation with the research speech and language therapist (SLT) and from their verbal description of the Cookie Theft picture from the Boston Diagnostic Aphasia Examination (BDAAE, Goodglass & Kaplan, 2001). All participants were English native speakers. Each dyad was treated by the same research SLT, who had more than 10 years of clinical experience. PWA consisted of 3 females and 4 males (time since stroke  $\bar{O} 40 \pm 16$  months, age  $\bar{O} 56 \pm 10$  years, age left education  $\bar{O} 18 \pm 2$  years) and their 7 CPs (5 females, 2 males).

### 4.2 Description of the data

During the main research project, each therapy session was videotaped using a Panasonic digital video camera. The dyad and the SLT are clearly visible on screen. Data analysed here consist of a randomly selected sample of 25% of all videotaped therapy sessions (N=14; two sessions out of 8 for each of 7 dyads), equating to a total of 17.3 hours of data. In addition, written notes by the research SLT were consulted for information on the actual content of individual sessions. Data were collated retrospectively allowing TF to be studied independently.

The generic BCA session plans were used to construct a fidelity tool since a BCA manual has yet to be developed.

### 4.3 Fidelity tool

A pilot fidelity tool was devised, covering the IFF (Carroll et al., 2007) components of adherence, with subcomponents content and dose, and potential moderating factors. As the essential components of BCA, i.e. those that are expected to create therapeutic change, are not known yet, fidelity to each component of the therapy programme, was examined (Carroll et al., 2007). The tool consisted of a 'procedural' section, created to carry out the main fidelity check, a 'qualitative' section, to analyse therapy delivery, and a 'client-focused' section, to capture participant responsiveness (i.e. opinions of the clients expressed during a therapy session). The procedural and qualitative sections will now be described in more detail, because the present article focuses on adherence to therapy content and therapy delivery. Therefore, the client-focused section stays in the background of the present article.

#### 4.3.1 Procedural section

The items included in the procedural section of the tool were based on the a priori constructed generic session plans that exist for each session, i.e. they were specific to the activities and materials of BCA therapy (e.g. 'The therapist showed a video example of a successful conversation repair'). They were therapist-oriented (following the recommendations by Hogue et al., 1996 to rate only therapist behaviour when assessing adherence). The section was designed to enable an external rater to fill it in on observation of therapy videos. Almost all of the items identified adherence, but some items were expected to indicate both adherence and quality of delivery (e.g. when the therapist leads a discussion with the dyad). The three-point Likert-type rating scale that was applied to the procedural section distinguished between therapy content that was fully, partly or not delivered as planned (1=fully delivered; 0.5=partly delivered; 0=not delivered). This permitted a rater to record items that were not fully delivered (e.g. because of a certain participant reaction), and so the interactive nature of the therapy was captured.

#### 4.3.2 Qualitative section

With regard to the present study, it is important to know which therapist skills are associated with the delivery of BCA. A list of principles reflecting desired therapist behaviour was created from the SPPARC manual and the adaptations made for BCA. Thus, ideally, the therapist:

- supplies individualised advice (based on analyses of conversation between the PWA and the CP)

- guides the dyad to make their own choices
- focuses both on the PWA and on the CP, so that the PWA and the CP have equal roles during the session
- avoids making judgments about what conversation patterns the dyad should retain or change
- uses active listening skills
- gives skilful summaries of what has been said
- expresses warmth and empathy towards the dyad
- affirms and encourages the dyad

These fundamental principles were listed for rating in the qualitative section of the tool. Following Chan et al. (2004), the rating scale used distinguishes between therapist behaviour which the rater observes not at all (0), occasionally (0.5) or most of the time (1).

Another aspect that was included in the qualitative part of the fidelity tool was whether the overall aim of the session was judged to be met. A statement for each session (e.g. session 2: 'The overall aim of the session, to raise the dyad's awareness of different aims of turns, was achieved') was listed in order to be rated with 0 (I don't agree), 0.5 (I partly agree) or 1 (I fully agree).

### 4.4 Procedures

A mixed-methods approach was conducted in order to answer the research questions. Guided by the IFF (Carroll et al., 2007), data concerning adherence and potential moderating factors were obtained via observation of the video-recorded therapy sessions by the first author of the present article. Additionally, secondary data sources (an email survey and document analyses such as notes on individual session plans or dates on video tapes) were conducted in order to examine aspects of TF. A second rater (a speech and language therapy graduate of UCL already familiar with the main research project) observed and rated 20% of the sessions coded by the first author (N=3) in order to investigate inter-rater reliability (IRR) of the procedural and the qualitative sections of the fidelity tool. The aim was to uncover whether the fidelity tool can be reliably implemented by different independent raters in the same way using the same video samples.

## 5. Results

The IRR analysis revealed that the overall percentage agreement of the procedural and qualitative section was similar (86.8% and 87.5% respectively), but higher variability was evident in the qualitative section of the fidelity tool. More detailed results regarding IRR are summarised in Appendix 1.

### 5.1 Adherence

«To what degree were the planned components of the BCA therapy programme delivered to the participants of the main research project?»

Here the focus is what Carroll et al. (2007) in their model refer to as «content». Across the sample of 14 sessions, the overall number of observations (i.e. items / therapy activities) was 232. Numbers of planned activities (e.g. talking through a specific handout) varied across each of the sessions. It was not possible to carry out the fidelity check on session N° 8 for dyad N° 3, because there were technical problems with the video recording, and the handwritten session notes by the SLT did not include enough information on the missing items to complete the ratings. The data for this session were therefore removed from further analysis.

The results of the ratings for the procedural section of the fidelity tool are shown in Table 4. The mean number of observations per dyad was 32 (SD=9.7) with a range between 18 for dyad N° 3 (session 5) and 41 for dyad N° 1 (sessions 4 and 6). Per session across all dyads, the mean number of observations was 18 (SD=5.2; range: 5-25).

The final data include a total of 227 observations over 13 sessions which would allow a maximum overall fidelity score of 227. The achieved score given by the rater is 208. This represents 91.9% (SD=3.9), i.e. the overall fidelity score across all dyads related to the IFF subcategory «con-

tent» is 91.9%, based on an analysis of 23% of the 56 therapy sessions.

A closer inspection of the dyad-specific fidelity scores reveals four of these are above 90% (for dyads N° 1, 2, 3 and 6). For two dyads, the scores lie just below 90% (dyads N° 4 and 5). The fidelity score for dyad N° 7 is the lowest at 86.1%. It must be kept in mind that the scores of the dyads are based on different numbers of observations, as these varied between sessions.

Averaged across sessions and dyads, 88.5% of the items were given a rating of 1 (fully delivered, N=201), 6.2% of the ratings indicated partial adherence (corresponding to a rating of 0.5, N=14) and 5.3% of the items were given a rating of 0 (not delivered, N=12).

### 5.2 Therapy delivery

«To what degree did the behaviour of the therapist reflect desired BCA intervention principles?»

The results for the qualitative section of the fidelity tool are reported in Table 5. Here the focus is what Carroll et al. (2007) refer to as «quality of delivery».

Again, session N° 8 for dyad N° 3 was removed from the analysis (see section 5.1). The results of the TF check for the qualitative section, based on the remaining 13 sessions, show scores near ceiling: They range between 90% (dyad N° 7) and 100% (dyads N° 3, 4 and 5). Overall, the degree to which desired therapist behaviour was present was 96.7% (SD=4.1).

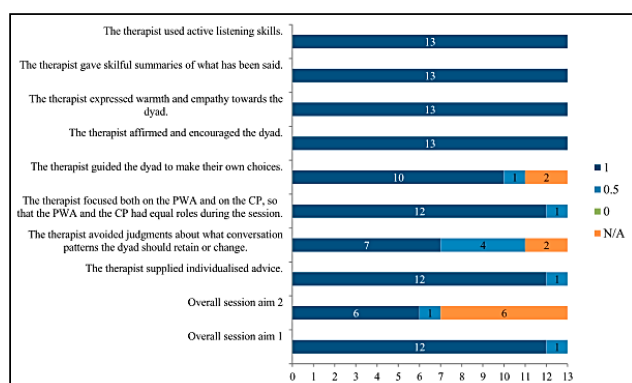
**Table 4:** Dyad-specific fidelity scores and overall fidelity score

Dyad #	Session # checked	Activities planned, Maximum score	Activities delivered, Actual score	Percentage $\frac{\text{Actual score}}{\text{Maximum score}} \times 100$
1	4 6	41	37.5	<b>91.5</b>
2	2 3	38	35.5	<b>93.4</b>
3	5	18	17.5	<b>97.2</b>
4	1 6	39	35	<b>89.7</b>
5	7 8	19	17	<b>89.5</b>
6	3 4	36	34.5	<b>95.8</b>
7	3 4	36	31	<b>86.1</b>
<b>Total</b>		<b>227</b>	<b>208</b>	
<b>Mean</b>		<b>32.4</b>	<b>29.7</b>	<b>91.9</b>
Standard deviation		9.7	8.7	3.9

**Table 5:** Dyad-specific quality scores and overall quality score

Dyad #	Session # checked	Maximum score of desired behaviour	Actual score	Percentage $\frac{\text{Actual score}}{\text{Maximum score}} \times 100$
1	4 6	19	17.5	<b>92.1</b>
2	2 3	19	18.5	<b>97.4</b>
3	5	10	10	<b>100</b>
4	1 6	16	16	<b>100</b>
5	7 8	16	16	<b>100</b>
6	3 4	20	19.5	<b>97.5</b>
7	3 4	20	18	<b>90.0</b>
<b>Mean</b>		<b>17.1</b>	<b>16.5</b>	<b>96.7</b>
Standard deviation		3.6	3.1	4.1

The frequency with which each item was rated with 0 (not at all), 0.5 (occasionally) or 1 (most of the time) can be found in Figure 2. The item «The therapist avoided judgements about what conversation patterns the dyad should retain or change» was most frequently rated at 0.5 compared to other items. No single item in the fidelity check, however, was rated at 0. Thus, for this section of the TF tool, Likert-scaling turned into a binary rating system. The most consistently rated qualitative items were «The therapist used active listening skills», «The therapist gave skilful summaries of what has been said», «The therapist expressed warmth and empathy towards the dyad» and «The therapist affirmed and encouraged the dyad», each of which reached the maximum score of 13 across all observations.



**Figure 2:** Frequency of ratings (0=not at all; 0.5=occasionally; 1=most of the time) of therapist behaviour across the observed sessions (N=13); N/A=not applicable

## 6. Discussion

### 6.1 Adherence

Across the observed therapy sessions, fidelity to the session plans (score: 91.9%) can be regarded as high, according to the published TF literature. This indicates that BCA therapy delivery was consistent with the prototype therapy as designed by the main project team. It is therefore likely that the so-called active ingredients of BCA therapy have been delivered with high fidelity, although further research is necessary in order to identify these. Averaged across all observed sessions, the percentages of items that were given a rating of 1 (88.5%), 0.5 (6.2%) and 0 (5.3%) are similar to or better than the results reported in previous literature. For example, Lewinsohn et al. (1990) stated that 78% of items were given a rating of 2 (corresponding to a rating of 1 in the present study), 17% a rating of 1 (corresponding to 0.5) and 5% of the ratings indicated no compliance. However, it is important to keep in mind that the design of the study by Lewinsohn et al. (ibid.), the nature of the intervention and the small

number of items in the fidelity tool (11) differ from the current investigation.

Dyad N° 7 showed a lower individual fidelity score (86.1%) compared to the other six dyads (89.5% and above). This score can be explained by the relatively high proportion of 0 and 0.5 ratings given to items observed. This shows that even when only one experienced therapist is delivering an intervention (which can be described as an «ideal situation», according to Hennessey & Rumrill, 2003:p. 124), there appears to be a certain amount of variety in therapy delivery. Moreover, it shows that a fidelity evaluation can serve as a tool to uncover certain participants for which therapy was delivered in a different way compared to the majority. It also points to the importance of measuring fidelity when there is more than one therapist delivering an intervention. In this regard, it is important to provide appropriate training to all therapists and use strategies such as supervision or a detailed manual in order to ensure that each therapist delivers the planned therapy.

One issue worth noting in this study is the selected sample. Interpretations have to be handled with caution because it could be argued that the final sample (23% of all sessions) is not representative of the whole therapeutic process, although guidelines from the TF literature concerning the amount of therapy sessions to check were followed. Also, each dyad-specific fidelity score is based on a different number of observations, depending on the individual sessions checked.

The BCA therapy can be described as highly interactive, which is why some constituents of the procedural section of the fidelity tool are worded in a rather open way (e.g. «The therapist had a discussion with the dyad on how aphasia affects conversations»). The data reveal that a relatively high percentage of the 0.5-ratings given by rater 1 (corresponding to «partly delivered») consisted of items belonging to the major domain of having a discussion. This could be a sign of problems for the rater with the definition of these items (the issue of specificity of treatment components is also discussed in Whyte & Hart, 2003). This issue could be further examined in the future, for example, by assessing the degree of complexity of the BCA therapy (e.g. with the help of a survey for experts in the specific therapy approach, see also Carroll et al., 2007) or by rewording or concretising these items. A general issue in the context of the procedural section of the fidelity tool is the influence which each of the procedural items has on the overall fidelity score. As each single element from the generic session plans was included, there is no variation in their weight related to the overall fidelity score (i.e. each item is equally weighted). By including an estimate of the therapeutic potency of each item, hypothesised active ingredients could be in-

vestigated for their potential to influence fidelity more than elements which might be regarded as less important. In general, future research should aim to develop a shorter fidelity tool reflecting the most important aspects of the BCA therapy approach.

## 6.2 Therapy delivery

Items in the qualitative section of the fidelity tool are an attempt to describe desired therapist behaviour associated with the delivery of the BCA therapy. However, it is important to acknowledge the complexity of this aspect of TF. High quality scores (ranging from 90% to 100%) were found for the seven dyads using a 3-point Likert scale. This ceiling effect could theoretically reflect rater bias. Consequently, it could be advantageous to use a 5-point Likert scale. This may lead to a higher degree of differentiation in the rating of desired therapist behaviour. Another possibility would be to create clear definitions and examples that correspond to the 0, 0.5 or 1-ratings for each qualitative item, to clarify how to use the scale. Furthermore, the challenge for future investigations is to consider whether there are additional behaviours (not currently listed in the TF tool) that are key to competent BCA therapy delivery. One possible skill relates to the ability to select appropriate positive and negative samples of conversation to show a dyad during therapy, as this underpins a vital component of the intervention - the identification of facilitator and barrier conversation behaviours to target in therapy. As yet unidentified aspects of therapy delivery such as this might also have an impact on future studies where more than one therapist is delivering BCA therapy. This highlights the need for comprehensive SLT training when preparing for a reliable delivery of the BCA therapy. One resource that begins to answer such a need can be found at <https://extend.ucl.ac.uk>, a free e-learning resource based around BCA to help SLTs to plan, carry out and evaluate conversation therapy.

The qualitative item which was most frequently rated with a score of 0.5 was «The therapist avoided making judgments about what conversation patterns the dyad should retain or change». This suggests that the therapist showed the desired behaviour only «occasionally». On the other hand, the coding of this item might reflect the suitability of the clients for the intervention, since it is important in the therapeutic exchange for the therapist to react to client behaviour. For example, if a client is not able to choose what conversation strategy to work on, as a consequence the therapist is more likely to have to judge herself which patterns or strategies should be changed or retained.

## 7. Conclusion

The fidelity and therapy delivery levels for BCA therapy reported here can be regarded as high, whilst at the same time reflecting a certain amount of flexibility in therapy delivery, which is likely to be desirable for such an interactive therapy approach. These findings contribute to the group outcomes of the main research project by providing evidence that each of the seven dyads received the BCA therapy as originally planned. This study underlines the growing prominence of TF in the field of speech and language therapy, especially with regard to ensuring the methodological quality of research reports. However, valid fidelity tools need to be created for speech and language interventions, of which conversation-based therapy approaches represent one example. In conclusion, this study has demonstrated the multifaceted nature of TF, and its importance and value for the evaluation of a complex speech and language intervention for aphasia.

## Acknowledgements

This article presents selected findings from the first author's Master's thesis in the programme «Sprachtherapie» (Speech and Language Therapy) at the University of Munich. The thesis is available at <http://nbn-resolving.de/urn/resolver.pl?urn=nbn:de:bvb:19-epub-17910-9>. The Master's project was embedded in a research project at University College London which was funded by a Stroke Association project grant (TSA 2007/05, 2008-2011) and approved by Cambridgeshire 1 Research Ethics Committee (project-ID: 08/H0304/40). We want to thank all the dyads who participated in the main project. Furthermore, special thanks go to Jo Barclay, who volunteered to be rater 2 for the investigation of inter-rater reliability of the fidelity tool, and to Prof Wolfram Ziegler for his advice and supervision.

## Appendix 1

Inter-rater reliability of the independent ratings for the procedural and the qualitative section of the fidelity tool averaged across the three sessions observed by both raters

Section of the fidelity tool	Percentage agreement	Range	ICC
Procedural section (adherence/content)	86.8%	80.0-92.9%	.674 (good)
Qualitative section (therapy delivery)	87.5%	62.5-100%	.258 (poor) <sup>a</sup>

<sup>a</sup>This ICC value is not an average but reflects the ICC of one session, session N° 4 for dyad N° 7 (there was a lack of variability in the raw data of the other two sessions so that ICC could not be calculated for all of the three sessions).

## Originalbeitrag: Investigating treatment fidelity in a conversation-based aphasia therapy

IRR was examined for the pairs of dual-rated video-recorded therapy sessions. Both percentage agreement and an intra-class correlation (ICC; two-way random, consistency, single-measures; Shrout & Fleiss, 1979) were performed, the latter to correct for chance agreement (Hallgren, 2012). The table below summarises the results of the IRR investigation. The overall percentage agreement of both sections of the fidelity tool is similar, but the range

indicates higher variability in the qualitative section. Averaged across rated sessions, ICC values indicate good agreement across raters for the procedural section, but poor agreement for the qualitative section. The latter could be due to the fact that percentage agreement for the specific session which built the foundation for this ICC value of .258 was only 62.5 %.

**Contact |** Heilemann, Claudia; UCL Division of Psychology and Language Sciences Chandler House 2 Wakefield Street London WC1N 1PF UK; c.heilemann@ucl.ac.uk

### References

- Adams, C., Lockton, E., Gaile, J., Earl, G., & Freed, J. (2012). Implementation of a manualized communication intervention for school-aged children with pragmatic and social communication needs in a randomized controlled trial: the Social Communication Intervention Project. *International Journal of Communication Disorders*, 47 (3), 245-256.
- Beckley, F., Best, W., Johnson, F., Edwards, S., Maxim, J., & Beeke, S. (2013). Conversation therapy for agrammatism: exploring the therapeutic process of engagement and learning by a person with aphasia. *International Journal of Language and Communication Disorders*, 48 (2), 220-239.
- Beeke, S., Wilkinson, R., & Maxim, J. (2007). Grammar without sentence structure: A conversation analytic investigation of agrammatism. *Aphasiology*, 21 (3/4), 256-282.
- Beeke, S., Maxim, J., Best, W., & Cooper, F. (2011). Redesigning therapy for agrammatism: initial findings from a conversation-based approach. *Journal of Neurolinguistics*, 2, 222- 236.
- Beeke, S., Beckley, F., Johnson, F., Heilemann, C., Edwards, S., Maxim, J., & Best, W. (2014). Conversation focused aphasia therapy: investigating the adoption of strategies by people with agrammatism. *Aphasiology*. DOI: 10.1080/02687038.2014.881459.
- Bellg, A.J., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D.S., Ory, M., Ogedegbe, G., et al. (2004). Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology*, 23 (5), 443-451.
- Bond, G.R., Evans, L., Salyers, M.P., Williams, J., & Kim, H.-W. (2000). Measurement of Fidelity in Psychiatric Rehabilitation. *Mental Health Services Research*, 2 (2), 75-87.
- Burch, K., Wilkinson, R., & Lock, S. (2002). A single case study of conversation-focused therapy for a couple where one partner has aphasia. *British Aphasiology Society Therapy Symposium Proceedings*. London: British Aphasiology Society, 1-9.
- Carroll, C., Patterson, M., Wood, S. Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science*, 40 (2). DOI:10.1186/1748-5908-2-40
- Chan, E.K.H., O'Neill, I., McKenzie, M., Love, A., & Kissane, D.W. (2004). What works for therapists conducting family meetings: treatment integrity in family-focused grief therapy during palliative care and bereavement. *Journal of Pain and Symptom Management*, 27 (6), 502-512.
- Cherney, L.R., Simmons-Mackie, N., Raymer, A., Armstrong, E., & Holland, A. (2013). Systematic review of communication partner training in aphasia: methodological quality. *International Journal of Speech-Language Pathology*, 15 (5), 535-545.
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). Developing and evaluating complex interventions: the new Medical Research Council guidance. *British Medical Journal* (337:a1655), 979-983. DOI: <http://dx.doi.org/10.1136/bmj.a1655>
- Di Rezze, B., Law, M., Gorter, J.W., Eva, K., & Pollock, N. (2012). A Narrative Review of Generic Intervention Fidelity Measures. *Physical & Occupational Therapy in Pediatrics*, 32 (4), 430-446.
- Godfrey, E., Chalder, T., Ridsdale, L., Seed, P., & Ogden, J. (2007). Investigating the 'active ingredients' of cognitive behaviour therapy and counselling for patients with chronic fatigue in primary care: Developing a new process measure to assess treatment fidelity and predict outcome. *British Journal of Clinical Psychology*, 46, 253-272.
- Goodglass, H., & Kaplan, E. (2001). *Boston Diagnostic Aphasia Examination*. Austin Texas: Pro-Ed.



## Originalbeitrag: Investigating treatment fidelity in a conversation-based aphasia therapy

- Hallgren, K.A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, 8 (1), 23-34.
- Hasson, H. (2010). Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implementation Science*, 67 (5). DOI:10.1186/1748-5908-5-67
- Hasson, H., Blomberg, S., & Dunér, A. (2012). Fidelity and moderating factors in complex interventions: a case study of a continuum of care program for frail elderly people in health and social care. *Implementation Science*, 23 (7).
- Hennessey, M.L., & Rumrill, P.D. (2003). Treatment fidelity in rehabilitation research. *Journal of Vocational Rehabilitation*, 19, 123-126.
- Hinckley, J.J., & Douglas, N.F. (2013). Treatment Fidelity: Its Importance and Reported Frequency in Aphasia Treatment Studies. *American Journal of Speech-Language Pathology*, 22, 279-284.
- Hogue, A., Liddle, H.A., & Rowe, C. (1996). Treatment adherence process research in family therapy: a rationale and some practical guidelines. *Psychotherapy*, 33 (2), 332-345.
- Holland, A.L. (1991). Pragmatic Aspects of Intervention in Aphasia. *Journal of Neurolinguistics*, 6 (2), 197-211.
- Holland, A.L. (1997). Pragmatic assessment and treatment for aphasia. In G.L. Wallace (Ed.), *Adult Aphasia Rehabilitation* (pp.161-174). Boston: Butterworth-Heinemann.
- Huber, W., Poeck, K., & Springer, L. (2006). *Klinik und Rehabilitation der Aphasie*. Stuttgart: Thieme.
- Kaderavek, J.N., & Justice, L.M. (2010). Fidelity: An Essential Component of Evidence-Based Practice in Speech-Language Pathology. *American Journal of Speech-Language Pathology*, 19, 369-379.
- Kagan, A. (1998a). Supported conversation for adults with aphasia: Methods and resources for training conversation partners. *Aphasiology*, 12 (9), 816-830.
- Kagan, A. (1998b). Philosophical, practical and evaluative issues associated with «Supported Conversation for Adults with Aphasia»: a reply. *Aphasiology*, 12 (9), 851-864.
- Kagan, A., & Gailey, G. (1993). Functional is not enough: training conversation partners of aphasic adults. In A. Holland & M.M. Forbes (Eds.), *Aphasia treatment: world perspectives* (pp.199-226). London: Chapman & Hall.
- Kiran, S., & Thompson, C.K. (2003). The role of semantic complexity in treatment of naming deficits: Training semantic categories in fluent aphasia by controlling exemplar atypicality. *Journal of Speech, Language and Hearing Research*, 46, 608-622.
- Kolb, D.A. (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice Hall.
- Lewinsohn, P.M., Clarke, G.N., Hops, H., & Andrews, J. (1990). Cognitive-Behavioral Treatment for Depressed Adolescents. *Behavior Therapy*, 21, 385-401.
- Lichstein, K.L., Riedel, B.W., & Grieve, R. (1994). Fair tests of clinical trials: A treatment implementation model. *Advances in Behaviour Research and Therapy*, 16 (1), 1-29.
- Lock, S., Wilkinson, R., & Bryan, K. (2001a). *Supporting Partners of People with Aphasia in Relationships and Conversation (SPPARC): A Resource Pack*. Bicester, UK: Speechmark.
- Lock, S., Wilkinson, R., Bryan, K., Maxim, J., Edmundson, A., Bruce, C., & Moir, D. (2001b). *Supporting Partners of People with Aphasia in Relationships and Conversation (SPPARC)*. *International Journal of Language & Communication Disorders*, 36 (1), 25-30.
- Lyon, J.G., Cariski, D., Keisler, L., Rosenbek, J., Levine, R., Kumpula, J., Ryff, C., Coyne, S., & Blanc, M. (1997). Communication partners: Enhancing participation in life and communication for adults with aphasia in natural settings. *Aphasiology*, 11 (7), 693-708.
- McVicker, S., Parr, S., Pound, C., & Duchan, J. (2009). The Communication Partner Scheme: A project to develop long-term, low-cost access to conversation for people living with aphasia. *Aphasiology*, 23 (1), 52-71.
- Moncher, F.J., & Prinz, R.J. (1991). Treatment fidelity in outcome studies. *Clinical Psychological Review*, 11, 247-266.
- Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: development, measurement and validation. *American Journal of Evaluation*, 24, 315-340.
- Nelson, M.C., Cordray, D.S., Hulleman, C.S., Darrow, C.L., & Sommer, E.C. (2012). A Procedure for Assessing Intervention Fidelity in Experiments Testing Educational and Behavioral Interventions. *The Journal of Behavioral Health Services & Research*, 1-22.
- Odom, S.L., Fleming, K., Diamond, K., Lieber, J., Hanson, M., Butera, G., Horn, E., Palmer, S., & Marquis, J. (2010). Examining different forms of implementation and in early childhood curriculum research. *Early Childhood Research Quarterly*, 25, 314-328.

## Originalbeitrag: Investigating treatment fidelity in a conversation-based aphasia therapy

- Papathanasiou, I., Coppens, P., & Potagas, C. (2013). *Aphasia and Related Neurogenetic Communication Disorders*. Burlington: Jones & Bartlett Learning.
- Resnick, B., Bellg, A.J., Borrelli, B., De Francesco, C., Breger, R., Hecht, J., et al. (2005). Examples of implementation and evaluation of treatment fidelity in the BCC studies: Where we are and where we need to go. *Annals of Behavioral Medicine*, 29 (2), 46-54.
- Saldert, C., Backman, E., & Hartelius, L. (2013). Conversation partner training with spouses of persons with aphasia: A pilot study using a protocol to trace relevant characteristics. *Aphasiology*, 27 (3), 271-292.
- Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86 (2), 420-428.
- Simmons-Mackie, N., Raymer, A., Armstrong, E., Holland, A., & Cherney, L.R. (2010). Communication Partner Training in Aphasia: A Systematic Review. *Archives of Physical Medicine and Rehabilitation*, 91, 1814-1837.
- Swinburn, K., McVicker, S., & Pearce-Willis, E. (2011). Living with aphasia for life: What can community-based voluntary organizations offer? *Aphasie und verwandte Gebiete*, 2, 36-46.
- Waltz, J., Addis, M.E., Koerner, K., & Jacobson, N.S. (1993). Testing the Integrity of a Psychotherapy Protocol: Assessment of Adherence and Competence. *Journal of Consulting and Clinical Psychology*, 61 (4), 620-630.
- Whyte, J., & Hart, T. (2003). It's more than a black box; it's a Russian doll: Defining rehabilitation treatments. *American Journal of Physical Medicine & Rehabilitation*, 82 (8), 639-652.
- Wielaeert, S., & Wilkinson, R. (2012). Partners van Afasiepatiënten Conversatie Training (PACT). [Conversation Training for Partners of People with Aphasia] Houten: Bohn Stafleu van Loghum.
- Wilkinson, R., & Wielaeert, S. (2012). Rehabilitation Targeted at Everyday Communication: Can We Change the Talk of People With Aphasia and Their Significant Others Within Conversation? *Archives of Physical Medicine and Rehabilitation*, 93, Suppl 1, 70-76.
- Wilkinson, R., Bryan, K., Lock, S., & Sage, K. (2010). Implementing and evaluating aphasia therapy targeted at couples' conversations: A single case study. *Aphasiology*, 24 (6), 869-886.
- Yeaton, W.H., & Sechrest, L. (1981). Critical dimensions in the choice and maintenance of successful treatments: strength, integrity and effectiveness. *Journal of Consulting and Clinical Psychology*, 49 (2), 156-167.

Originalbeitrag

# Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie

Schröder, Astrid; Hausmann, Nadin; Stadie, Nicole

## DE | Zusammenfassung

Die vorliegende Studie beschreibt eine Überprüfung der Wirksamkeit des semantischen Wiederholungsprimings zur Behandlung des mündlichen Bildbenennens bei Aphasie (Laine & Martin, 1996; Martin, Fink, & Laine, 2004). Teilnehmer der Pilotstudie war ein deutschsprachiger Patient mit erhaltenen semantischen Fähigkeiten und einer post-semantisch bedingten Störung des mündlichen Bildbenennens. In insgesamt sieben Behandlungseinheiten wurden zehn Vertreter der Kategorien OBST und MUSIKINSTRUMENTE mit Aufgaben zum Benennen, Wort-Bild-Zuordnen und Nachsprechen der Zielwörter intensiv geübt. Nach der Behandlung zeigten sich nachhaltige Übungseffekte sowie Generalisierungen auf das Benennen von semantisch relatierten, ungeübten Wörtern. Die Ergebnisse zeigen, dass das semantische Wiederholungspriming eine effektive Therapiemethode zur Behandlung von post-semantisch bedingten Störungen des Bildbenennens darstellt. Erhaltene semantische Fähigkeiten können dabei als Ressource genutzt werden, um die mündliche Wortproduktion zu unterstützen, nachhaltige Übungseffekte zu erzielen und Generalisierungen auf semantisch ähnliche, ungeübte Wörter zu ermöglichen.

## EN | Summary

This study investigates the efficacy of semantic repetition priming in the treatment of picture naming in aphasia (Laine & Martin, 1996; Martin et al., 2004). Participant of this pilot study was a German-speaking patient with intact semantics and impaired post-semantic access to lexical word representations. Ten members of the semantic categories FRUITS and MUSICAL INSTRUMENTS were repeatedly trained with naming, word-picture matching and repetition tasks in seven therapy sessions. After treatment, significant, long-lasting practice effects as well as generalization to semantically related, untreated words were seen. The results show that semantic repetition priming is effective for the treatment of post-semantic picture naming deficits in aphasia. In these participants, intact access to semantics can be used as a resource to support spoken word production and to allow for long-term practice and generalization effects.

## 1. Einleitung

Der Begriff ‹Priming› (dt. Bahnung) beschreibt die Tatsache, dass die Verarbeitung eines Wortes durch ein zeitlich davor präsentiertes Wort beeinflusst wird, wobei sowohl unterstützende als auch hemmende Effekte auftreten können (vgl. z.B. Stadie, Drenhaus, Höhle, Spalek, & Wartenburger, 2012). Primingeffekte mit verkürzten Reaktionszeiten für das Verarbeiten von Wörtern nach der vorherigen Präsentation eines so genannten ‹Primes›, werden dabei als Faszilitierung bzw. Bahnung der Verarbeitung durch den zuvor präsentierten Prime interpretiert. Beim so genannten *Wiederholungspriming* zeigt sich, dass die wiederholte Präsentation eines Wortes den Abruf eines Wortes beim mündlichen Bildbenennen sowohl bei Sprachgesunden als auch bei Menschen mit Aphasie erleichtert (Breitenstein, Kamping, Jansen, Schomacher, & Knecht, 2004; Cave, 1997; Miceli, Amitrano, Capasso, & Caramazza, 1996; Nickels, 2002; Wheeldon & Monsell, 1992). Beim *semantischen Priming* zeigt sich dieser Effekt auch dann, wenn z.B. beim Bildbenennen vor der Präsentation des Zielbildes (z.B. *Löwe*) zwar nicht das Zielwort, aber ein semantisch verwandtes Wort oder eine semantische Eigenschaft (z.B. *TIER, Tiger, oder hat eine Mähne*) präsentiert wird (vgl. z.B. Hutchison et al., 2013; Moss & Tyler, 1995).

Die Möglichkeit der Faszilitierung der Wortverarbeitung durch semantische Voraktivierung machen sich zahlreiche Methoden zur Behandlung von aphasisch bedingten Wortabrufstörungen zu Nutze. So werden zum Beispiel in der Therapie des Bildbenennens Übungen durchgeführt, die eine erhöhte semantische Aktivierung für das Zielitem erwirken, um anschließend den nachfolgenden Wortabruf zu erleichtern (z.B. Aufgaben zum Wort-Bild-Zuordnen, Beantworten von Fragen zu semantischen Merkmalen, Vorgabe von semantischen Hilfen beim Bildbenennen (siehe Stadie & Schröder, 2009; Whitworth, Webster, & Howard, 2014, für einen Überblick). Howard, Hickin, Redmond, Clark, und Best (2006) postulieren, dass sich explizite sprachtherapeutische Methoden wie z.B. das bewusste Reflektieren über semantische Merkmale (Boyle & Coelho, 1995; Kiran, 2008; Kiran & Thompson, 2003) eher für die Therapie von Störungen des Bildbenennens bei semantischer Störungsursache eignen. Für Patienten mit erhaltener semantischer Verarbeitung und postsemantisch bedingten Störungen im lexikalischen Zugriff hingegen seien vielmehr implizite Methoden wirksam. Hierzu zählt u.a. das semantische Priming, bei dem unmittelbar vor dem Bildbenennen Aufgaben zum Wort-Bild-Zuordnen durchgeführt werden, um durch die erzielte semantische Voraktivierung den Wortabruf zu erleichtern (Howard, Patterson, Franklin, Orchard-Lisle, & Morton, 1985).

Martin und Laine (Laine & Martin, 1996; Martin & Laine, 2000) entwickelten die Methode des sogenannten *kontextuellen Wiederholungsprimings* für die Behandlung von Störungen des Wortabrufes bei Aphasie (vgl. auch Martin et al., 2004; Martin, Fink, Renvall, & Laine, 2006; Martin & Laine, 2000; Renvall, Laine, Laakso, & Martin, 2003; Renvall, Laine, & Martin, 2005; Renvall, Laine, & Martin, 2007). Bei dieser Methode werden in der Therapie die geübten abbildbaren Wörter stets gemeinsam in einem Bilderset (z.B. bestehend aus fünf bis sechs Übungselementen) präsentiert. Innerhalb einer Therapiesitzung werden dann für jedes Zielbild in diesem Set Aufgaben zum Benennen, Wort-Bild-Zuordnen und Nachsprechen der Übungselemente durchgeführt. Diese Übungen werden für jedes Zielitem mehrfach innerhalb der Behandlungseinheit wiederholt. Durch die Wiederholung wird die mündliche Wortproduktion während einer Therapiesitzung für jedes Zielitem hochfrequent geübt (je nach Durchführung in den erwähnten Publikationen zwischen 12 und 49 mal). Der dadurch erzielte Mechanismus des Wiederholungsprimings wird ergänzt durch die *Art des Kontextes* bei der Zusammenstellung der Bilder, die entweder phonologisch, semantisch, oder nicht related sind. Durch die gemeinsame Präsentation relateder Begriffe wird insgesamt eine erhöhte Aktivierung semantisch bzw. phonologisch ähnlicher Repräsentationen ermöglicht (Martin et al., 2004; Martin et al., 2006; Renvall et al., 2003; Renvall et al., 2005; Renvall et al., 2007). Generalisierungseffekte auf das Benennen ungeübter Wörter wurden bislang nur für einen Einzelfall nach der Behandlung mit der Methode des semantischen Wiederholungsprimings beschrieben (Renvall et al., 2003), so dass ausschließlich diese Methode für die vorliegende Studie ausgewählt wurde.

Für das Üben mit einem Set aus semantisch relateden Wörtern beim *semantischen Wiederholungspriming* sind die in der Literatur beschriebenen Therapieeffekte für Patienten mit semantischen und postsemantischen Störungen sowie für Patienten mit erhaltenen rezeptiven semantischen Fähigkeiten in Tabelle 1 zusammengefasst<sup>1</sup>. Zwei Ergebnisse werden in dieser Zusammenstellung deutlich: Nachhaltige Übungseffekte wurden bislang lediglich bei Patienten mit erhaltenen semantischen Fähigkeiten erzielt, und eine Generalisierung auf das Benennen von ungeübten, semantisch relateden Wörtern konnte nur für einen Einzelfall (Renvall et al., 2003), ebenfalls mit erhaltenen semantischen Fähigkeiten, beobachtet werden (vgl. auch Howard et al., 2006; Pring, Davis, & Marshall, 1993 für ähnliche Ergebnisse).

Wie lässt sich die Beobachtung, dass nachhaltige Übungseffekte ausschließlich bei Patienten mit *erhalte-*

<sup>1</sup> Für einen weiteren Patienten mit semantischen Defiziten wurden nach der Anwendung einer leicht modifizierten Therapiemethode ebenfalls nachhaltige Effekte erzielt (Renvall et al., 2007).

## Originalbeitrag: Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie

	Patient	Störungslokalisierung		Therapieeffekt			Quelle
		SEM	POST-SEM	Übungseffekt	Nachhaltigkeitseffekt	Generalisierungseffekt	
1	PH	✗	✗	Ja	Nein	Nein	Renvall et al., 2005
2	BM	✗	✗	Ja	Nein	Nein	Martin et al., 2006
3	BQ	✗	✗	Ja	Nein	Nein	
4	AS	✗	✗	Ja	Nein	Nein	Martin et al., 2004
5	LP	✓	✗	Ja	Ja (4 Monate)	Nein	
6	YK	✓	✗	Ja	Ja (6 Wochen)	Ja	Renvall et al., 2003
7	JP	✓	✗	Ja	Ja (6 Wochen)	Nein	Renvall et al., 2007

**Tabelle 1:** Evidenzen zur Wirksamkeit des semantischen Wiederholungsprimings. ✗ = beeinträchtigt, ✓ = erhalten, SEM = semantische Fähigkeiten, POST-SEM = postsemantische phonologisch-lexikalische und phonologisch-segmentale Fähigkeiten

nen semantischen Fähigkeiten beobachtet wurden, erklären? Martin et al. (2006) nehmen an, dass bei Patienten mit unbeeinträchtigten semantischen Fähigkeiten während der Therapie eine top-down Aktivierung semantischer und lexikalischer Repräsentationen erfolgen kann, während bei Patienten mit semantischen Störungen eine semantische Aktivierung beim Nachsprechen von Wörtern ausbleibt. Dadurch können bei Patienten mit erhaltener Semantik während der Therapie die auch bei Sprachgesunden beschriebenen Mechanismen des Wiederholungsprimings wirksam werden und es werden nachhaltige Übungseffekte durch die Therapie erzielt. Für die Erklärung des Generalisierungseffektes auf ungeübte, semantisch relatierte Items gehen Renvall und Kollegen (2003) davon aus, dass die erhaltenen semantischen Fähigkeiten des Patienten YK einen Mechanismus der sich ausbreitenden Aktivierung auf erhaltene semantische Repräsentationen innerhalb des semantischen Systems ermöglicht haben. Durch einen zusätzlichen «boost» (Renvall et al., 2003, S. 326) an Aktivierung wurde dadurch auch die post-semantische Aktivierung für die ungeübten Items aus den geübten semantischen Kategorien verstärkt und ein Generalisierungseffekt erzielt.

Vor dem Hintergrund der evidenzbasierten Literatur zur Behandlung von Wortabrufstörungen ist das Ausbleiben eines Generalisierungseffektes für das Benennen von ungeübten, semantisch relatierten Items nicht unbedingt überraschend. Für Patienten mit post-semantischen, lexikalischen (Zugriffs)defiziten werden Generalisierungseffekte (außer beim Erlernen einer expliziten Strategie) nach der Therapie des Bildbenennens in der Literatur nahezu nicht berichtet (vgl. z.B. Best et al., 2013; Wisenburn & Mahoney, 2009). Diese werden auch theoretisch nicht erwartet, da es sich beim Training des lexikalischen Zugriffs um item-spezifisches Lernen für

die arbiträre Verbindung zwischen einer semantischen Repräsentation und einem lexikalischen Eintrag handelt (Howard, 2000). Generalisierungseffekte mit besseren Leistungen für das Benennen ungeübter, semantisch ähnlicher Begriffe werden für Patienten mit semantischen Defiziten meist nach einem expliziten Training über das Wissen einzelner semantischer Merkmale und deren bedeutungsunterscheidender Funktion beobachtet (z.B. Boyle, 2004; Hillis, 1989; Kiran, 2007). Kiran und Kollegen konnten zeigen, dass der Parameter *semantische Typikalität* einen Einfluss auf die mögliche Generalisierung ausübt. So zeigten sich in einer Reihe von Studien Generalisierungseffekte lediglich nach dem Üben von semantisch atypischen Vertretern einer semantischen Kategorie mit besseren Leistungen für das Benennen ungeübter typischer Vertreter der geübten Kategorie. Umgekehrte Effekte nach dem Üben von eher typischen Kategorievertretern wurden hingegen nicht beobachtet. (Kiran, 2008; Kiran & Thompson, 2003; vgl. auch Stanczak, Waters, & Caplan, 2006). Ein Einfluss der semantischen Typikalität zeigt sich auch ohne therapeutische Intervention in einfachen Aufgaben zum Bildbenennen mit schnelleren Benennzeiten für typische als für weniger typische Vertreter einer semantischen Kategorie bei Sprachgesunden (Dell'Acqua, Lotto, & Job, 2000; Holmes & Ellis, 2006) sowie weniger Fehlern für typische als für atypische Kategorievertreter bei Probanden mit semantischer Demenz und Aphasie (Laiacona, Luzzatti, Zonca, Guarnaschelli, & Capitani, 2001; Rossiter & Best, 2013; Woollams, Cooper-Pye, Hodges, & Patterson, 2008). Diese Effekte werden im Rahmen von Modellen erklärt, in denen semantische Konzepte in Form von semantischen Merkmalen (z.B. *kann fliegen, hat Flügel*) repräsentiert sind (z.B. Plaut, 1996; Rogers et al., 2004; Rosch, 1975). Es wird angenommen, dass atypische Vertreter einer Kategorie (z.B. *Pinguin* für die Kategorie VÖGEL) eine

## Originalbeitrag: Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie

breitere Varianz an distinktiven semantischen Merkmalen aufweisen als typische Vertreter (z.B. *Spatz*), die mehr Merkmale untereinander teilen. Durch diese Interkorrelationen weisen typische Vertreter einer semantischen Kategorie stabilere Repräsentationen auf und sind auch bei einer semantischen Störung noch leichter abrufbar. Aufgrund der breiteren Varianz an unterschiedlichen, distinktiven semantischen Merkmalen bei atypischen Vertretern einer semantischen Kategorie kommt es dann zu Generalisierungen auf typische Vertreter einer Kategorie, deren Merkmale ein Subset der geübten Merkmale für die atypischen Kategorievertreter darstellen (z.B. Kiran, 2007; Kiran & Thompson, 2003). Ob es jedoch aufgrund der breiteren Varianz an semantischen Merkmalen bei atypischen Vertretern durch Ko-Aktivierung auch zu einer Generalisierung auf andere, ungeübte atypische Vertreter einer semantischen Kategorie kommen kann, wurde bislang noch nicht untersucht.

### Ziel und Fragestellungen der vorliegenden Studie

Ziel der vorliegenden Therapiestudie war die Replikation der bereits in der englischsprachigen Literatur beschriebenen Effekte zur Wirksamkeit des semantischen Wiederholungsprimings an einem deutschsprachigen Patienten mit Aphasie. Folgende Fragestellungen sollten beantwortet werden:

1. Führt ein hochfrequentes Benennen, Wort-Bild-Zuordnen und Nachsprechen der Zielbegriffe mit der Methode des semantischen Wiederholungsprimings zu einem Übungseffekt?
2. Treten nach dem Training von atypischen Kategorievertretern Generalisierungen auf das Benennen von anderen, ungeübten atypischen Kategorievertretern aus den geübten semantischen Kategorien auf?
3. Sind die beobachteten Therapieeffekte nachhaltig?

## 2. Methode

### Proband

Teilnehmer der Studie ist der 48jährige Patient HW mit einer chronischen Aphasie, verursacht durch einen Teilinfarkt im Mediastromgebiet links 12 Jahre vor Studienbeginn. Im Aachener Aphasie Test (AAT, Huber, Poeck, Weniger, & Willmes, 1983) zeigte Herr HW in der ALLOC-Klassifikation zu 100% eine amnestische Aphasie (vgl. Abb. 1 zur Illustration des AAT-Profiles). In der flüssigen Spontansprache traten sehr viele Wortfindungsstörungen auf, die sich durch Stocken im Sprachfluss, Pausen, Fehlen von Satzteilen sowie Redefloskeln äußerten (z.B. *und dann hab ich ne Umschulung gekriegt als . äh . hm . Kommunikationselektroniker . man kann auch sagen . ähm . . . oh Mensch, wie n . . Fernmeldetechniker oder so was in der*

*Art ; Ja da hat ich noch äh wurd ich von ner Firma . . hat ich äh bei jemandem in ner Firma äh war ich da . äh hab da wie sagt man ne Bewerbung gehabt).*

	Punktwerte	Prozentränge	Schweregrad	
			allgemein	Syndrom
Spontansprache (Profil)	4 5 5 4 4 4	----	----	----
Token Test (alterskor.)	20	65	M* L*	* S*
Nachsprechen	124	66	M* L*	* S*
Schriftsprache	75	80	* L*	S* M*
Benennen	95	72	M* L*	* S* M
Sprachverständnis	113	99	L* M*	* L*

Abbildung 1: AAT- Profil, Patient HW

Zur Ermittlung der semantischen Fähigkeiten wurde in einer Eingangsdiagnostik eine Auswahl von Untertests aus dem Diagnostikverfahren Lexikon modellorientiert (LEMO, De Bleser, Cholewa, Stadie, & Tabatabaie, 2004; Stadie, Cholewa, & De Bleser, 2013), der Bogenhausener Semantik Untersuchung (BOSU, Glindemann, Klintwort, Ziegler, & Goldenberg, 2002) und dem Pyramids and Palm Trees Test (Visch-Brinck & Denes, 1992, unveröffentlicht) durchgeführt (vgl. Tab. 2)

Test	Anzahl korrekt	% korrekt
Sortieren von Objekten /sem. Nebenmerkmale (BOSU)	9/10	90% (N)
Pyramids & Palm Trees Test	30/30	100% (N)
Wort-Bild-Zuordnen auditiv (LEMO)	20/20	100% (N)
Wort-Bild-Zuordnen visuell (LEMO)	20/20	100% (N)
Synonymie-Entscheiden mit sem. Ablenker auditiv (LEMO)	30/40	75% (B)
Synonymie-Entscheiden mit sem. Ablenker visuell (LEMO)	19/20	95% (N)
Benennen mündlich (LEMO)	19/20	95% (N)

Tabelle 2: Ergebnisse zur Überprüfung der semantischen Verarbeitung vor der Therapie. N= Normalbereich, B= beeinträchtigt, jeweils ermittelt in Form einer vorhandenen bzw. nicht vorhandenen signifikanten Abweichung zur Kontrollgruppe (Crawford & Howell, 1998).

In Tabelle 2 wird deutlich, dass im LEMO-Test mündliches Benennen, aber auch in non-verbalen (BOSU, Pyramids and Palm Trees Test) sowie verbalen Tests zur Überprüfung semantischer Fähigkeiten (auditives und visuelles Wort-Bild-Zuordnen, mündliches Benennen, visuelles Synonymie-Entscheiden) unbeeinträchtigte Leistungen vorlagen. Lediglich die Leistungen im auditiven Synonymie-Entscheiden waren beeinträchtigt. Die Leistungen in der visuellen Variante dieses Tests lagen jedoch im Normbereich, so dass insgesamt erhaltene se-

## Originalbeitrag: Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie

mentische Fähigkeiten angenommen werden können. Die modalitätsspezifischen Auffälligkeiten beim auditiven Synonymie-Entscheiden sind wahrscheinlich auf präsemantische Beeinträchtigungen zum Beispiel im auditiv-verbale Arbeitsgedächtnis oder in einer Störung des modalitätsspezifischen Zugriffs auf semantische Repräsentationen zurück zu führen.

Da die Auffälligkeiten in der Spontansprache auf Schwierigkeiten des Wortabrufes hinwiesen, wurde mit Herrn HW eine vertiefende Diagnostik zum Benennen und Wort-Bild-Zuordnen von 180 Bildern mit variierender semantischer Typikalität durchgeführt (vgl. Tab. 3). Hierzu wurde das mündliche Benennen sowie das auditive und visuelle Wort-Bild-Zuordnen (ein Zielbild, drei semantisch relationale Ablenker) für ein Set aus 180 Bildern mit je 30 Items aus sechs alltagsrelevanten semantischen Kategorien (TIERE, OBST, GEMÜSE, KLEIDUNG, TRANSPORT, MUSIKINSTRUMENTE) überprüft (Schröder, Gemballa, Ruppig, & Wartenburger, 2012; Schröder & Stadie, in Vorbereitung). Alle Kategorievertreter in diesem Bilderset variieren systematisch nach semantischer Typikalität, wobei in jedem Set je zehn typische (z.B. GEMÜSE - *Gurke*), mitteltypische (z.B. GEMÜSE - *Bohne*) und atypische (z.B. GEMÜSE - *Ingwer*) Kategorievertreter enthalten sind. Beim Benennen von Bildern wurden alle Reaktionen als korrekt gewertet, die innerhalb von 20 Sekunden korrekt erfolgten (inklusive Selbstkorrekturen).

Test	gesamt	in Prozent
Wort-Bild-Zuordnen auditiv, 180-Bilderset *	174/180	97%
Wort-Bild-Zuordnen visuell, 180-Bilderset *	179/180	99%
Benennen mündlich, 180-Bilderset*	115/180	63%

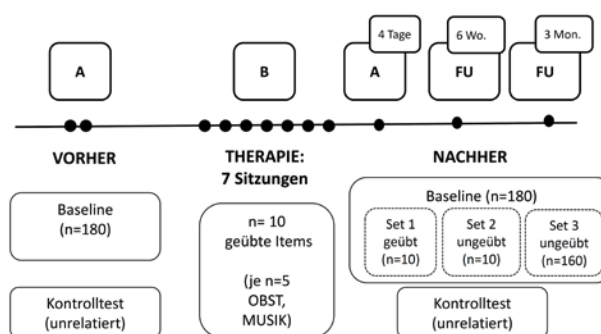
**Tabelle 3:** Ergebnisse der vertiefenden Diagnostik zum mündlichen Benennen, auditiven und visuellen Wort-Bild-Zuordnen von 180 Bildern mit variierender semantischer Typikalität. \* = Schröder und Stadie, in Vorbereitung.

Tabelle 3 zeigt, dass beim Benennen der Bilder Einschränkungen vorlagen, während für das gleiche Bilderset die rezeptiven Leistungen im auditiven und visuellen Wort-Bild-Zuordnen unbeeinträchtigt waren. Der Unterschied zwischen den beeinträchtigten Leistungen beim mündlichen Benennen und unbeeinträchtigten Leistungen beim visuellen und auditiven Wort-Bild-Zuordnen ist statistisch signifikant (Fisher's exakt Test, beide  $p < .001$ ). Beim Benennen von Bildern war ein Typikalitätseffekt mit signifikant besseren Leistungen beim Benennen von Bildern mit hoher semantischer Typikalität (51/60 korrekt) im Vergleich zu Bildern mit mitteltypi-

scher (37/60 korrekt) und atypischer Typikalität (27/60 korrekt) beobachtbar (Fisher's exakt Test, beide  $p < .01$ ). Eine qualitative Fehleranalyse ergab, dass Herr HW beim Bildbenennen überwiegend semantische Paraphrasen produzierte (insg. 54%, z.B. Zielbild: Tiger, Reaktion: *Löwe*). Weiterhin lagen Nullreaktionen (insg. 18% der Fehler), semantische Umschreibungen (insg. 11%, z.B. Zielbild: Balalaika, Reaktion: *'ne russische Mandoline*) und einige phonematische Paraphrasen vor (insg. 6%). Semantische Paraphrasen traten signifikant häufiger als alle anderen Fehlertypen auf (Fisher's exakt, zweiseitig, alle  $p < .001$ ). Das zahlreiche Auftreten semantischer Paraphrasen bei gleichzeitig diagnostizierten erhaltenen semantischen Fähigkeiten spricht für die Lokalisation der Störung im Zugriff von semantischen Repräsentationen auf phonologisch-lexikalische Wortformen (Caramazza & Hillis, 1990). Die wenigen phonematischen Paraphrasen könnten auf ein zusätzliches Defizit im phonologischen Arbeitsspeicher hinweisen, was jedoch im Rahmen der vorliegenden Studie nicht vertiefend untersucht wurde. Zusammenfassend zeigen die Ergebnisse, dass Herr HW das für die Teilnahme an der Studie notwendige Inklusionskriterium «erhaltene semantische Fähigkeiten» bei Vorliegen eines post-semantisch bedingten Defizites im Zugriff auf lexikalische Wortformen erfüllte.

### Therapieplan

Ziel der sprachtherapeutischen Intervention war die Verbesserung des mündlichen Wortabrufes. Da der Patient erhaltene semantische Fähigkeiten in Kombination mit Beeinträchtigungen beim Benennen von Bildern aufwies, wurde er für die vorliegende Pilotstudie zur Wirksamkeit des semantischen Wiederholungsprimings bei Störungen des Bildbenennens ausgewählt. Für die Behandlung wurde ein item-spezifisches ABA-Kontrolltest-Design mit zwei Nachhaltigkeitsuntersuchungen sechs Wochen und drei Monate nach Beendigung der Therapie ausgewählt (z.B. Franklin, 1997, vgl. Abb. 2).



**Abbildung 2:** Therapiedesign. A= Diagnostik, B= Therapie, FU = Nachhaltigkeitsuntersuchung (Follow-Up)

## Originalbeitrag: Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie

Als unrelatierter Kontrolltest wurde der LEMO-Test Schreiben nach Diktat von Neologismen ausgewählt (De Bleser et al., 2004; Stadie et al., 2013). Zur Ermittlung der Stabilität der Leistungen vor der Therapie wurden zwei Baseline-Untersuchungen mit zwei parallelisierten Bildersets im Abstand von neun Tagen durchgeführt. Hierzu wurde das 180-Bilderset in zwei Sets à n = 90 Items geteilt, die im Hinblick auf ihren Schwierigkeitsgrad (Anzahl an typischen, mitteltypischen und atypischen Vertretern der einzelnen semantischen Kategorien) angeglichen waren. Die Leistungen für die geteilten Bildersets waren über beide Testungen vor der Therapie vergleichbar (Sitzung 1: 60/90 korrekt, Sitzung 2: 55/90 korrekt, Fisher's exakt Test, zweiseitig, alle  $p > .10$ ).

Für die Durchführung der Therapie wurden nach der Baseline-Diagnostik jeweils zehn geübte Items (Set 1) und zehn ungeübte Items (Set 2) ausgesucht, die vor der Therapie beim Wort-Bild-Zuordnen korrekt zugeordnet werden konnten, aber inkorrekt benannt wurden. Aus jedem Set an geübten und ungeübten Items wurden je fünf Vertreter der Kategorien OBST und MUSIKINSTRU-

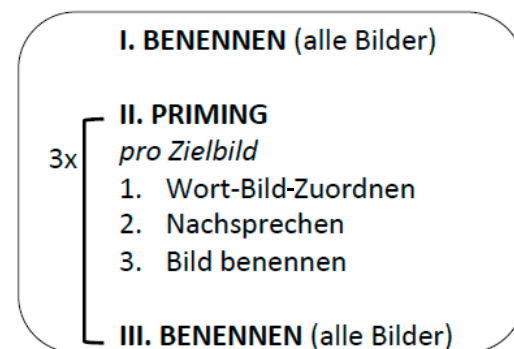
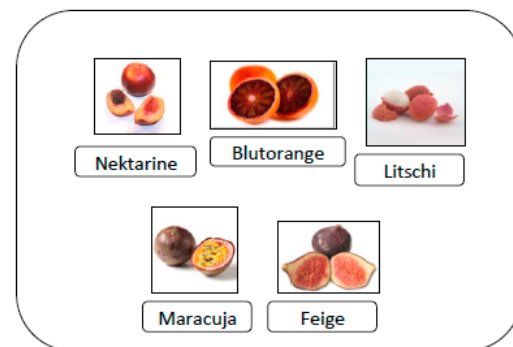
MENTE ausgewählt. Set 1 und Set 2 mit den je zehn geübten und ungeübten Items waren im Hinblick auf die semantische Typikalität, sowie weitere Variablen (Erwerbsalter, Familiarität, Wortfrequenz, Wortlänge; Schröder et al., 2012) angeglichen und somit unmittelbar miteinander vergleichbar (vgl. Tabelle 4). Set 3 enthielt die weiteren n=160 Items des 180er-Bildersets. Vier Tage nach Beendigung der Therapie wurden die Benennleistungen für die 180 Bilder aus der Eingangsbaseline erneut überprüft. Die Nachhaltigkeit der Ergebnisse wurde in zwei Follow-Up Untersuchungen sechs Wochen und drei Monate nach Beendigung der Therapie erfasst.

### Therapiedurchführung

Die Therapie fand 2-3x wöchentlich über jeweils 60-90 Minuten statt. Abbildung 3 illustriert das Vorgehen in der Therapie am Beispiel des Bildersets der Kategorie OBST. Alle fünf Übungsbilder eines Übungssets wurden unabhängig von der Art der Reaktion (korrekt oder inkorrekt) während einer Therapiesitzung nach einem feststehenden Ablauf trainiert (Martin et al., 2004). Zunächst wurden alle Bilder des Übungssets mündlich benannt (vgl. I. in Abb. 3). Anschließend wurden für jedes Zielbild (z.B. Nektarine, Blutorange, Litschi, usw.) Aufgaben zum Wort-Bild-Zuordnen, Nachsprechen und Benennen des Übungsbildes durchgeführt (vgl. II. in Abb. 3). Daraufhin wurden noch einmal alle fünf Bilder münd-

Kontrollierte Variable	Set 1 (geübt, n=10)	Set 2 (ungeübt, n=10)
<b>Semantische Typikalität (7-Punkte-Skala)*</b>		
MW	3,08	3,42
s	0,74	0,95
<b>Erwerbsalter (7-Punkte Skala)*</b>		
MW	5,29	5,31
s	0,52	0,91
<b>Familiarität (5-Punkte Skala)*</b>		
MW	2,91	2,66
s	0,66	0,75
<b>Wortfrequenz (per Million)#</b>		
MW	0,62	0,60
s	0,56	0,64
<b>Anzahl Phoneme</b>		
MW	6,40	6,90
s	2,41	2,13
<b>Anzahl Silben</b>		
MW	2,90	3,1
s	0,88	1,10

**Tabelle 4:** Kontrollierte Variablen bei der Zusammenstellung für die Sets an geübten und ungeübten Items (je n= 10). MW = Mittelwert, s = Standardabweichung, \* = Werte aus Schröder et al., 2012, # = normalisierte Lemmafrequenz per Million (Heister et al., 2011)



**Abbildung 3:** Vorgehen in der Therapie am Beispiel für das Bilderset der Kategorie OBST (n= 5 geübte Items)

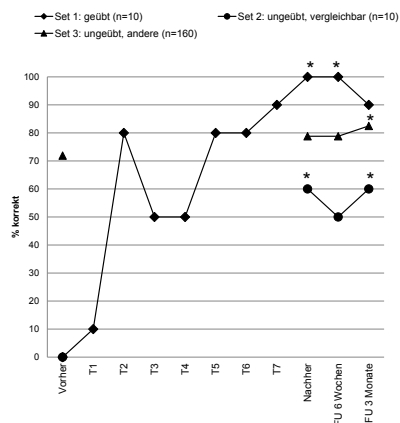


## Originalbeitrag: Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie

lich benannt (vgl. III. in Abb.3). Jeder Durchgang, bestehend aus Wort-Bild-Zuordnen, Nachsprechen und Benennen jedes einzelnen Items sowie das anschließende Benennen des gesamten Bildersets (vgl. II. und III. in Abb. 3) wurde 3x wiederholt, so dass pro Sitzung jedes Item mindestens 10x mündlich produziert wurde. Im Falle einer Fehlbenennung wurde das korrekte Item von der Therapeutin vorgegeben und vom Patienten noch einmal nachgesprochen. Die Abfolge der semantischen Kategorien OBST und MUSIKINSTRUMENTE wurde in den einzelnen Therapiesitzungen systematisch variiert. Nach Ablauf von sieben Therapiesitzungen war das vorher in Anlehnung an Martin et al. (2004) festgelegte cut-off-Kriterium von 80% korrekten Benennreaktionen für die geübten Items in drei aufeinanderfolgenden Sitzungen erreicht.

### 3. Ergebnisse

Abb. 4 zeigt eine Zusammenfassung der Ergebnisse für das Benennen der geübten Items im Therapieverlauf. Für die prozentuale Anzahl korrekter Benennungen während der Therapie sind jeweils die Reaktionen aus dem ersten Durchlauf (vgl. I.: Benennen, alle Bilder in Abb. 3) angegeben. Die Ergebnisse der Nachher-Diagnostik für die 180 Bilder wurden getrennt für die vorab definierten Sets (Set 1: geübt [n=10]; Set 2; ungeübt, vergleichbar [n=10]; Set 3: andere ungeübte [n=160]) ausgewertet. Signifikante Veränderungen für die wiederholte Testung identischer Itemsets wurden mit dem McNemar Test (bei einer erwarteten Häufigkeit von  $n < 5$  pro Zelle mit Kontinuitätskorrektur) überprüft. Für die Untersuchung der Nachhaltigkeit wurden die Leistungen aus den Follow-Up-Sitzungen sechs Wochen und drei Monate nach der Therapie mit den Leistungen vor Beginn der Therapie verglichen.



**Abbildung 4:** Darstellung der Ergebnisse: Therapieverlauf sowie Prozent korrekt beim Benennen geübter Items (Set 1) und ungeübter Items (Set 2 und Set 3) vor und nach der Therapie. T = Therapiesitzung, \* = signifikante Veränderung im Vergleich zur Vorher-Diagnostik (McNemar,  $p < .05$ )

#### Unrelatierter Kontrolltest

Im unrelatierten Kontrolltest (Schreiben nach Diktat von Neologismen) lagen vor und nach der Therapie keine Veränderungen in den Leistungen vor (29/40 bzw. 31/40 korrekt, McNemar,  $p > .10$ ).

#### Geübte Items

Abb. 4 zeigt, dass für geübte Items die Benennleistungen bereits nach der ersten Sitzung anstiegen. Nach der siebten Sitzung war das cut-off-Kriterium von 80% Benennleistungen in drei aufeinanderfolgenden Sitzungen erreicht, so dass die Therapie beendet wurde. In der Nachher-Diagnostik zeigten sich signifikante Verbesserungen beim Benennen der geübten Bilder (0/10 korrekt vs. 10/10 korrekt, McNemar,  $p < .01$ ). Dieser Anstieg war in beiden Follow-Up-Messungen stabil (McNemar, beide  $p < .01$ ).

#### Ungeübte Items

Nach der Therapie zeigten sich für die zehn unmittelbar vergleichbaren ungeübten Items aus den geübten Kategorien OBST und MUSIKINSTRUMENTE ein signifikanter Leistungsanstieg von 0 auf 60% (0/10 korrekt vs. 6/10 korrekt, McNemar,  $p < .05$ ). Dieser fiel in der ersten Follow-Up Sitzung sechs Wochen nach Beendigung der Therapie leicht ab, so dass lediglich ein statistischer Trend für den Anstieg der Leistungen im Vergleich zur Messung vor der Therapie (0/10 korrekt vs. 5/10 korrekt, McNemar,  $p = .063$ ) beobachtbar war. In der zweiten Follow-Up Sitzung drei Monate nach Beendigung der Therapie lagen wieder statistisch signifikante Verbesserungen vor (0/10 korrekt vs. 6/10 korrekt, McNemar,  $p < .05$ ). Für die weiteren 160 ungeübten Bilder aus Set 3 zeigten sich in der Nachher-Diagnostik und in der Follow-Up Untersuchung sechs Wochen nach Beendigung der Therapie nur leichte, statistisch nicht signifikante Veränderungen (jeweils 115/160 vs. 126/160 korrekt, McNemar,  $p = .082$  bzw.  $p = .109$ ). In der zweiten Follow-Up Untersuchung drei Monate nach Beendigung der Therapie war jedoch erneut eine leichte Verbesserung im Benennen der ungeübten Items aus ungeübten semantischen Kategorien beobachtbar, die statistische Signifikanz erreichte (115/160 versus 132/160 korrekt, McNemar,  $p < .01$ ).

### 4. Diskussion

Ziel der vorliegenden Pilotstudie war die Überprüfung der Wirksamkeit des semantischen Wiederholungsprimings für die Behandlung des mündlichen Wortabrufes bei einem deutschsprachigen Patienten mit erhaltenen semantischen Fähigkeiten und einer post-semantisch lokalisierten Störungen des lexikalischen Zugriffs. Teil-

## Originalbeitrag: Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie

nehmer der Studie war der 48jährige Patient HW mit einer chronischen Aphasie (12 Jahre post-onset) und einer Restsymptomatik von starken Wortfindungsstörungen in der Spontansprache. Erhaltene semantische Fähigkeiten wurden mit einer Reihe von Tests zur Überprüfung der semantischen Verarbeitung belegt. Beim mündlichen Bildbenennen wurden deutliche Einschränkungen beobachtet, die sich signifikant von den unauffälligen Leistungen beim auditiven und visuellen Wort-Bild-Zuordnen für ein Set von 180 Bildern mit variierender semantischer Typikalität (Schröder & Stadie, in Vorbereitung) unterschieden. Fehler beim mündlichen Benennen zeigten sich überwiegend in Form von semantischen Paraphrasen. Diese Beobachtung liefert weitere Evidenz dafür, dass semantische Paraphrasen nicht per se auf eine semantische Störung hinweisen, sondern auch bei einem Defizit im Zugriff auf phonologisch-lexikalische Wortformen bei erhaltener Semantik auftreten (Caramazza & Hillis, 1990). Der beobachtete Typikalitätseffekt beim Benennen mit besseren Leistungen für typische als für weniger typische Vertreter einer Kategorie bestätigt aktuelle Befunde darüber, dass neben anderen psycholinguistischen Variablen auch semantische Typikalität eines Items die Benennleistungen von Menschen mit Aphasie beeinflusst (Rossiter & Best, 2013).

Für die Therapie mit dem semantischen Wiederholungspriming wurden aus den in der Baseline benannten Bildern zehn atypische Vertreter aus zwei semantischen Kategorien (OBST und MUSIKINSTRUMENTE) ausgewählt, die vor der Therapie nicht benannt aber in beiden rezeptiven Tests zum auditiven und visuellen Wort-Bild-Zuordnen korrekt zugeordnet werden konnten. In insgesamt sieben Therapiesitzungen wurden die Begriffe in zwei Bildersets mit je fünf Bildern in einer festen Abfolge aus Benennen, Wort-Bild-Zuordnen, Nachsprechen und Benennen hochfrequent geübt (mindestens zehn Benennungen des Zielitems innerhalb einer Therapiesitzung bei 2 - 3 Behandlungseinheiten à 60-90 Minuten in der Woche). Ein zweites Set von vergleichbaren Items aus beiden behandelten semantischen Kategorien blieb ungeübt. Vor der Therapie konnten in zwei separat durchgeführten Baselineuntersuchungen stabile Leistungen in den parallelisierten Sets zum mündlichen Benennen festgestellt werden. Stabile Leistungen zeigten sich auch für den unrelatierten Kontrolltest (Schreiben nach Diktat) nach Durchführung der Therapie, so dass alle nachfolgend beschriebenen Effekte als therapiespezifisch zu bezeichnen sind.

Insgesamt wurden für das Set von geübten Items nach der Therapie hochsignifikante Übungseffekte erreicht, die auch sechs Wochen und drei Monate nach Beendigung der Therapie nachhaltig waren. Zusätzlich zu den geübten Items waren signifikante Verbesserungen für

das Set von ungeübten Items beobachtbar, die aus den geübten semantischen Kategorien stammten und im Hinblick auf die semantische Typikalität, Familiarität, Frequenz und Wortlänge mit den geübten Items vergleichbar waren. Für das restliche Set von n=160 ungeübten Bildern lagen in der Messung unmittelbar nach der Therapie sowie sechs Wochen nach der Therapie keine signifikanten Veränderungen vor. Drei Monate nach Beendigung der Therapie stiegen die Benennleistungen für dieses Set jedoch erneut leicht an, so dass signifikant bessere Leistungen als zu Beginn der Therapie für diese ungeübten Items beobachtbar waren. Neben den bereits beschriebenen Patienten aus Finnland (n= 2, Renvall et al., 2003, 2007) und den USA (n=1, Martin et al., 2004) konnte in der vorliegenden Studie nun auch für einen deutschsprachigen Patienten mit erhaltenen semantischen Fähigkeiten gezeigt werden, dass mit dem semantischen Wiederholungspriming post-semantisch bedingte Wortabrufstörungen nachhaltig verbessert werden können und Generalisierungen auf ungeübte, semantisch relationierte Items erzielt werden können. Im Rahmen der vorliegenden Pilotstudie wurde nur ein kleines Set von geübten Items (n=10) in einer relativ kurzen Therapiedauer (n=7 Sitzungen) behandelt. Weitere Replikationen mit einer größeren Stichprobe sind notwendig, um aussagekräftige Schlussfolgerungen aus den Ergebnissen ziehen zu können. Die Ergebnisse liefern jedoch erste Hinweise auf wichtige Implikationen, die wir im Folgenden diskutieren möchten.

### **Nachhaltige Verbesserungen beim Benennen der geübten Items**

Die Studie bestätigt den in der Literatur vielfach beschriebenen Effekt, dass das wiederholte Üben des Bildbenennens zu einer Verbesserung des Wortabrufes für die geübten Items führt. Diese itemspezifischen Lerneffekte werden sowohl für das Üben mit semantischen Methoden, als auch mit phonologischen, oder gemischten Methoden beschrieben (Wisburn & Mahoney, 2009). Der zugrundeliegende Wirkmechanismus hierbei ist laut Howard (2000) das Stärken der itemspezifischen Verbindungen zwischen der durch die Bildpräsentation aktivierten semantischen Repräsentation und durch das mündliche Produzieren des Wortes aktivierten lexikalischen Repräsentation des Wortes sowie deren Verknüpfung. In der vorliegenden Studie konnte dieser Effekt durch hochfrequentes Üben der Zielitems in einer relativ kurzen Behandlungszeit (sieben Therapiesitzungen à 60-90 Min. in einem Zeitraum über drei Wochen) erreicht werden. Innerhalb einer Therapiesitzung wurden die geübten Items sehr häufig, mindestens zehn Mal, vom Probanden mündlich produziert, so dass innerhalb der Behandlungszeit von drei Wochen jedes Übungsittem

## Originalbeitrag: Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie

mindestens 70mal in der Therapiesituation produziert wurde. Nach Martin et al. (2004) und Renvall et al. (2003) wird – insbesondere bei Patienten mit erhaltenen semantischen Fähigkeiten – auf diese Weise das Prinzip des Wiederholungsprimings wirksam, so dass durch die häufig wiederholte Aktivierung der lexikalischen Repräsentationen nachhaltig eine Faszilitierung des Wortabrufes für die geübten Items ermöglicht wird. In Bezug auf das häufige Wiederholen von geübtem Material während der Therapie (sog. *massed practice*) zeigt die vorliegende Studie, dass im Rahmen des viel diskutierten Faktors der Therapieintensität (z.B. Bhogal, Teasell, Speechley, & Albert, 2003; Cherney, 2012; Cherney, Patterson, Raymer, Frymark, & Schooling, 2008) nicht nur die *Dauer* und *Anzahl* der *Therapieeinheiten* pro Tag oder Woche, sondern auch die *Anzahl der Wiederholungen für geübte Items* innerhalb einer Sitzung berücksichtigt werden muss. Weitere Studien sind notwendig, um zu überprüfen, welchen Einfluss die Anzahl an Wiederholungen pro Item innerhalb einer Sitzung auf die Größe des Lerneffektes haben und ob möglicherweise dabei zusätzlich auch die Größe des behandelten Sets von geübten Items eine Rolle spielt (vgl. z.B. Boyle, 2004). Insgesamt sollte bei der Zusammenstellung des Übungsmaterials darauf geachtet werden, dass alltagsrelevante Sets für im Alltag häufig verwendete Wörter zusammen getragen werden (vgl. z.B. Renvall, Nickels, & Davidson, 2013a, b), sodass die Patienten einen direkten Nutzen auch aus den zu erwartenden itemspezifischen Lerneffekten für ein umschriebenes Set an Wörtern ziehen können.

### **Nachhaltige Verbesserungen beim Benennen ungeübter, vergleichbarer Items**

In der vorliegenden Studie wurden die zu übenden Bilder nach der Methode des *semantischen Wiederholungsprimings* zusammen gestellt, d.h. alle Übungselemente wurden stets im Kontext von vier weiteren, semantisch relatierten Begriffen aus der gleichen semantischen Kategorie geübt. Im Rahmen des verwendeten Therapiedesigns ist es nicht möglich zu entscheiden, ob nur einer der beiden Wirkmechanismen (semantisches Priming bzw. Wiederholungspriming) entscheidend zum Erfolg der Therapie beigetragen hat, oder ob von einem additiven Effekt beider Mechanismen ausgegangen werden muss. Allerdings zeigten sich auch drei Monate nach Beendigung der Therapie nachhaltige Verbesserungen auch für das Benennen von *ungeübten Vertretern* der geübten semantischen Kategorien, die im Hinblick auf die semantische Typikalität, Familiarität, Wortfrequenz und Wortlänge mit den geübten Items vergleichbar waren. Diese Beobachtung einer Generalisierung auf semantisch ähnliche Items unterstützt die Annahmen von Renvall et al. (2003), dass durch das Üben in einem semantischen Kontext eine

sich ausbreitende semantische Aktivierung auf ungeübte Vertreter der geübten semantischen Kategorien ermöglicht wurde. Unklar bleibt jedoch, warum dieser angenommene Mechanismus der Unterstützung der post-semantischen Verarbeitung durch sich ausbreitende Aktivierung auf semantischer Ebene nicht auch für andere in der Literatur beschriebene Patienten mit erhaltenen semantischen Fähigkeiten wirken konnte, bei denen ein Generalisierungseffekt aber ausblieb (z.B. Martin et al., 2004; Renvall et al., 2007), oder lediglich zu itemspezifischen Verbesserungen in einer ungeübten Modalität führte (wie z.B. bei dem Patient TRC in Best, 2000, Nickels & Best, 1996).

Die vorliegenden Ergebnisse mit Generalisierungen auf Vertreter der geübten semantischen Kategorien mit einer eher geringen, aber vergleichbaren semantischen Typikalität (z.B. geübt: *Feige*, ungeübt: *Datteln*) zeigen auch, dass die zu erwartenden Generalisierungen möglicherweise nicht auf das Spektrum «von atypisch zu typisch» beschränkt sind (Kiran, 2007). Weitere Studien sind notwendig, um zu erforschen, ob und bei welchen Patienten der bereits von Renvall et al. (2003) beschriebene Generalisierungseffekt repliziert werden kann und wie das Material strukturiert sein muss, um möglichst weit fassende Generalisierungen zu erhalten.

### **Verbesserungen für ungeübte, andere Items nach mehrmaliger Präsentation**

Für die weiteren 160 ungeübten Items aus Set 3 waren weder in der Nachher-Untersuchung unmittelbar nach der Therapie, noch in der Nachhaltigkeitsuntersuchung sechs Wochen nach Beendigung der Therapie Verbesserungen beim Benennen beobachtbar. Drei Monate nach Ablauf der Therapie stieg die Leistung im Benennen dieser Items jedoch an, so dass signifikant bessere Leistungen als vor Beginn der Therapie vorlagen. Wie ist dieser Leistungsanstieg zu erklären? In der Literatur wurde bereits mehrfach beschrieben, dass der einfache Mechanismus der häufigen Wiederholung in der Testsituation auch ohne therapeutisches Feedback bei wiederholtem Testen zu signifikanten Lernerfolgen führen kann (z.B. Breitenstein et al., 2004; Nickels, 2002). Für den Patienten HW konnten wir beobachten, dass das verwendete Bilderset bei wiederholter Testung zunehmend vertrauter für den Patienten wurde und dieser begann, seine Leistungen in der Diagnostiksituation auf vorherige Testungen zu beziehen und zu vergleichen. Wir gehen daher davon aus, dass Verbesserungen in der vierten Untersuchung einen itemspezifischen Lerneffekt, verursacht durch die wiederholte Präsentation der Bilder, widerspiegeln und nicht als eine genuine Generalisierung betrachtet werden können.

## 5. Fazit

Zusammenfassend liefern die Befunde dieser Pilotstudie Hinweise für eine weitere Bestätigung der Hypothese, dass das semantische Wiederholungspriming eine effektive Therapiemethode für die Behandlung von Wortabrufstörungen darstellt. Als Resultat werden bei Patienten mit erhaltenen semantischen Fähigkeiten nachhaltige Übungseffekte sowie möglicherweise Generalisierungen auf semantisch ähnliche Begriffe erwartet. Im klinischen Alltag sollten daher für diese Patienten Bildersets mit alltagsrelevanten Items aus alltagsrelevanten semantischen Kategorien zusammengestellt und

nach dem beschriebenen Vorgehen wiederholt und hochfrequent geübt werden. Weitere Forschung ist notwendig, um die Ergebnisse zu replizieren und weitere Einflussparameter auf den Erfolg der Behandlung (z.B. Art und Anzahl der geübten Items pro Sitzung, Anzahl von Wiederholungen pro Sitzung) zu erfassen.

## 6. Danksagung (Remerciement)

Wir danken dem Patienten HW für die Teilnahme an der Studie sowie Anne Adelt und Wiebke Bruchmüller für das Bereitstellen von Daten aus der Eingangsdiagnostik.

**Kontakt** | Schröder, Astrid; Universität Potsdam, Exzellenzbereich Kognitionswissenschaften, Neurokognition der Sprache, Department Linguistik, Karl-Liebknecht-Straße 24-25, 14476 Potsdam; astrid.schroeder@uni-potsdam.de; (+49) 0331-977-2387

## Bibliographie

- Best, W., Greenwood, A., Grassly, J., Herbert, R., Hickin, J., & Howard, D. (2013). Aphasia rehabilitation: Does generalisation from anomia therapy occur and is it predictable? A case series study. *Cortex*, 49, 2345-2357.
- Bhagal, S. K., Teasell, R., Speechley, M., & Albert, M. L. (2003). Intensity of aphasia therapy, impact on recovery. *Stroke*, 34, 987-993.
- Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology*, 13, 236-249.
- Boyle, M., & Coelho, C. A. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech Language Pathology*, 4, 94-98.
- Breitenstein, C., Kamping, S., Jansen, A., Schomacher, M., & Knecht, S. (2004). Word learning can be achieved without feedback: Implications for aphasia therapy. *Restorative Neurology and Neuroscience*, 22, 445-458.
- Caramazza, A., & Hillis, A. E. (1990). Where do semantic errors come from? *Cortex*, 26, 95-122.
- Cave, C. B. (1997). Very long-lasting priming in picture naming. *Psychological Science*, 8, 322-325.
- Cherney, L. R. (2012). Aphasia treatment: Intensity, dose parameters, and script training. *International Journal of Speech-Language Pathology*, 14, 424-431.
- Cherney, L. R., Patterson, J. P., Raymer, A., Frymark, T., & Schooling, T. (2008). Evidence-based systematic review: Effects of intensity of treatment and constraint-induced language therapy for individuals with stroke-induced aphasia. *Journal of Speech, Language, and Hearing Research*, 51, 1282-1299.
- Crawford, J. R., & Howell, D. C. (1998). Comparing an individual's test score against norms derived from small samples. *The Clinical Neuropsychologist*, 12, 482-486.
- De Bleser, R., Cholewa, J., Stadie, N., & Tabatabaie, S. (Eds.). (2004). LEMO. Lexikon modellorientiert. Einzelfalldiagnostik bei Aphasie, Dyslexie und Dysgraphie. München: Urban & Fischer, Elsevier.
- Dell'Acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, 32, 588-615.
- Franklin, S. (1997). Designing single case treatment studies for aphasic patients. *Neuropsychological Rehabilitation*, 7, 401-418.
- Glindemann, R., Klintwort, D., Ziegler, W., & Goldenberg, G. (Eds.). (2002). Bogenhausener Semantik-Untersuchung (BOSU). München: Urban & Fischer.
- Heister, J., Würzner, K.-M., Bubbenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexDB- eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, 61, 10-20.
- Hillis, A. E. (1989). Efficacy and generalization of treatment for aphasic naming errors. *Archives of Physical Medicine and Rehabilitation*, 70, 632-636.

## Originalbeitrag: Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie

- Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition*, 13, 884-910.
- Howard, D. (2000). Cognitive neuropsychology and aphasia therapy: the case of word retrieval. In I. Papathanasiou (Ed.), *Acquired Neurogenic Communication Disorders: A Clinical Perspective*. (pp. 76-99). London: Whurr.
- Howard, D., Hickin, J., Redmond, T., Clark, P., & Best, W. (2006). Re-visiting «semantic facilitation» of word retrieval for people with aphasia: facilitation yes but semantic no. *Cortex*, 42, 946-962.
- Howard, D., Patterson, K., Franklin, S., Orchard-Lisle, V., & Morton, J. (1985). The facilitation of picture naming in aphasia. *Cognitive Neuropsychology*, 2, 49-80.
- Huber, W., Poeck, K., Weniger, D., & Willmes, K. (1983). *Der Aachener Aphasie Test (AAT)*. Göttingen: Hogrefe.
- Hutchison, K., Balota, D., Neely, J., Cortese, M., Cohen-Shikora, E., Tse, C.-S., Yap, M., Bengson, J., Niemeier, D., & Buchanan, E. (2013). The semantic priming project. *Behavior Research Methods*, 45, 1099-1114.
- Kiran, S. (2007). Complexity in the treatment of naming deficits. *American Journal of Speech-Language Pathology*, 16, 18-29.
- Kiran, S. (2008). Typicality of inanimate category exemplars in aphasia treatment: further evidence for semantic complexity. *Journal of Speech, Language, and Hearing Research*, 51, 1550-1568.
- Kiran, S., & Thompson, C. K. (2003). The role of semantic complexity in treatment of naming deficits: training semantic categories in fluent aphasia by controlling exemplar typicality. *Journal of Speech, Language, and Hearing Research*, 46, 608-622.
- Laiacina, M., Luzzatti, C., Zonca, G., Guarnaschelli, C., & Capitani, E. (2001). Lexical and semantic factors influencing picture naming in aphasia. *Brain and Cognition*, 46, 184-187.
- Laine, M., & Martin, N. (1996). Lexical retrieval deficit in picture naming: implications for word production models. *Brain and Language*, 53, 283-314.
- Martin, N., Fink, R., & Laine, M. (2004). Treatment of word retrieval deficits with contextual priming. *Aphasiology*, 18, 457-471.
- Martin, N., Fink, R. B., Renvall, K., & Laine, M. (2006). Effectiveness of contextual repetition priming treatments for anomia depends on intact access to semantics. *Journal of the International Neuropsychological Society*, 12, 853-866.
- Martin, N., & Laine, M. (2000). Effects of contextual priming on impaired word retrieval. *Aphasiology*, 14, 53-70.
- Miceli, G., Amitrano, A., Capasso, R., & Caramazza, A. (1996). The treatment of anomia resulting from output lexical damage: Analysis of two cases. *Brain and Language*, 52, 150-174.
- Moss, H. E., & Tyler, L. K. (1995). Investigating semantic memory impairments: The contribution of semantic priming. *Memory*, 3, 359-395.
- Nickels, L. (2002). Improving word finding: Practice makes (closer to) perfect? *Aphasiology*, 16, 1047-1060.
- Plaut, D. C. (1996). Relearning after damage in connectionist networks: toward a theory of rehabilitation. *Brain and Language*, 52, 25-82.
- Pring, T., Davis, A., & Marshall, J. (1993). Therapy for word finding deficits: can experimental findings inform clinical work? In F. J. Stachowiak (Ed.), *Developments in the assessment and rehabilitation of brain-damaged patients: perspectives from a European concerted action*. Tübingen: Gunter Narr.
- Renvall, K., Laine, M., Laakso, M., & Martin, N. (2003). Anomia treatment with contextual priming: A case study. *Aphasiology*, 17, 305-328.
- Renvall, K., Laine, M., & Martin, N. (2005). Contextual priming in semantic anomia: A case study. *Brain and Language*, 95, 327-341.
- Renvall, K., Laine, M., & Martin, N. (2007). Treatment of anomia with contextual priming: exploration of a modified procedure with additional semantic and phonological tasks. *Aphasiology*, 21, 499-527.
- Renvall, K., Nickels, L., & Davidson, B. (2013a). Functionally relevant items in the treatment of aphasia (part I): Challenges for current practice. *Aphasiology*, 27, 636-650.
- Renvall, K., Nickels, L., & Davidson, B. (2013b). Functionally relevant items in the treatment of aphasia (part II): Further perspectives and specific tools. *Aphasiology*, 27, 651-677.
- Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). The structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, 111, 205-235.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, 104, 192-233.

---

**Originalbeitrag: Semantisches Wiederholungspriming in der Therapie von Wortabrufstörungen bei Aphasie – Eine Pilotstudie**

- Rossiter, C., & Best, W. (2013). «Penguins don't fly»: An investigation into the effect of typicality on picture naming in people with aphasia. *Aphasiology*, 27, 784-798.
- Schröder, A., Gemballa, T., Ruppin, S., & Wartenburger, I. (2012). German norms for semantic typicality, age of acquisition, and concept familiarity. *Behavior Research Methods*, 44, 380-394.
- Schröder, A., & Stadie, N. (in Vorbereitung). Semantische Aufgaben. Neurolinguistische Materialien zur Diagnose, Behandlung und Evaluation von semantischen Defiziten. Potsdam: Universität Potsdam.
- Stadie, N., Cholewa, J., & De Bleser, R. (2013). LEMO 2.0. Lexikon modellorientiert. Diagnostik für Aphasie, Dyslexie und Dysgraphie. Hofheim: NAT-Verlag.
- Stadie, N., Drenhaus, H., Höhle, B., Spalek, K., & Wartenburger, I. (2012). Forschungsmethoden der Psycholinguistik. In B. Höhle (Ed.), *Psycholinguistik*. Berlin: Wissenschaftsverlag, 2. unveränderte Auflage.
- Stadie, N., & Schröder, A. (2009). *Kognitiv orientierte Sprachtherapie. Methoden, Material und Evaluation für Aphasie, Dyslexie und Dysgraphie*. München: Elsevier Urban & Fischer.
- Stanczak, L., Waters, G., & Caplan, D. (2006). Typicality-based learning and generalisation in aphasia: Two case studies of anomia treatment. *Aphasiology*, 20, 374-383.
- Visch-Brinck, E., & Denes, G. (1992, unveröffentlicht). Semantic Association Test: A Modified Version of the «Pyramid and Palmtree Test». Rotterdam: Academic Hospital.
- Wheeldon, L. R., & Monsell, S. (1992). The locus of repetition priming of spoken word production. *Quarterly Journal of Experimental Psychology*, 44, 723-761.
- Whitworth, A., Webster, J., & Howard, D. (2014). *A cognitive neuropsychological approach to assessment and intervention in aphasia. A clinician's guide*. (2 ed.). Hove: Psychology Press.
- Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology*, 23, 1338-1352.
- Woolams, A. M., Cooper-Pye, E., Hodges, J. R., & Patterson, K. (2008). Anomia: A doubly typical signature of semantic dementia. *Neuropsychologia*, 46, 2503-2514.