



universität
wien

MASTERARBEIT

Titel der Masterarbeit

Das Arbeitsgedächtnis beim Simultandolmetschen.
Erfahrungs- und modalitätsspezifische
Kapazitätsunterschiede

Verfasser

Adrian Carl High, Bakk.phil.

angestrebter akademischer Grad

Master of Arts (MA)

Wien, im März 2012

Studienkennzahl lt. Studienblatt: A 065 342 345

Studienrichtung lt. Studienblatt: Masterstudium Dolmetschen Englisch Französisch

Betreuer: Ao. Univ.-Prof. Dr. Franz Pöchhacker

Dank

Das Thema „Arbeitsgedächtnis“ hat sich in den letzten 20 Jahren zu einem dynamischen Forschungsbereich der Dolmetschwissenschaft entwickelt. Die Einsicht, dass ein weiterer Beitrag zur Erkundung arbeitsgedächtnisrelevanter Aspekte beim Dolmetschen aufschlussreich wäre, verdanke ich den Anregungen meines fachlichen Betreuers Franz Pöchlhacker, der mir auch im weiteren Verlauf dieses Masterarbeitsprojekts vortrefflich in fachlichen und praktischen Belangen zu helfen wusste.

Ein Wort des Dankes sei an dieser Stelle an alle jene gerichtet, die sich Zeit genommen haben, die vorliegende Arbeit vollständig oder teilweise zu lesen. Für wertvolle Korrekturen, scharfsinnige Kommentare und nützliche Ratschläge bin ich insbesondere Theresia Pantzer und Yuuzuki Ripplinger zu großem Dank verpflichtet. Auch meine Schwester Karina A. High und mein Vater Carl J. High unterstützten mich in redaktioneller Hinsicht unermüdlich.

Selbstverständlich wäre das Experiment ohne die bereitwillige Teilnahme so vieler Vorversuchs- und Hauptversuchspersonen zum Scheitern verurteilt gewesen. Daher gebührt auch ihnen großer Dank.

Als äußerst aufschlussreich erwies sich die Korrespondenz mit einigen führenden Persönlichkeiten der dolmetschwissenschaftlichen Arbeitsgedächtnisforschung. Barbara Köpke, Minhua Liu und Šárka Timarová gilt ein aufrichtiges Dankeschön für ihre fachlichen Einschätzungen meiner Ergebnisse.

Zu guter Letzt sei auch dem Medienteam des Zentrums für Translationswissenschaft für die Bereitstellung der Versuchsräumlichkeiten und die technische Betreuung gedankt.

Inhalt

Abbildungsverzeichnis	13
Tabellenverzeichnis	13
Abkürzungsverzeichnis	15
0. Einleitung	17
1. Kognitionspsychologische Grundlagen	19
1.1. Allgemeines	19
1.2. AG als Forschungsgegenstand der kognitiven Psychologie	20
1.3. Zwei Arbeitgedächtnismodelle	22
1.3.1. AG-Modell nach Baddeley	22
1.3.2. AG-Modell nach Cowan	26
1.4. Vergleichende Diskussion der AG-Modelle	28
2. Das Arbeitsgedächtnis beim Dolmetschen	30
2.1. Der AG-Begriff in der Dolmetschwissenschaft	30
2.2. Experimentelle Methoden	32
2.3. Dolmetschwissenschaftliche Untersuchungen des AG	34
2.3.1. Freie Reproduktion mit Artikulatorischer Suppression	34
2.3.2. Reading Span Task	35
2.3.3. Listening Span Task	36
2.4. Kategorisierung der Erklärungsansätze	40
2.5. Kritische Bewertung der Untersuchungsmethoden	41
2.6. Methodische Anregungen	43
3. Der Reading Span Task	45
3.1. Entwicklung des interdisziplinären Austauschs	45

3.2.	Der RST in der Kognitionspsychologie	48
3.2.1.	<i>Ursprünge des RST</i>	48
3.2.2.	<i>Methodische Entwicklung</i>	52
3.2.3.	<i>Bearbeitungsaufgabe (Processing Task)</i>	54
3.2.4.	<i>Scoringmethoden</i>	55
3.2.5.	<i>Materialerzeugung</i>	58
3.2.6.	<i>Darbietungsreihenfolge der Satzebenen</i>	59
3.2.7.	<i>Vereinbarkeit mit gängigen AG-Modellen</i>	59
3.3.	Der RST in der Dolmetschwissenschaft	61
3.3.1.	<i>Problemlage</i>	61
3.3.2.	<i>Bisherige Ergebnisse</i>	62
3.3.3.	<i>Methodische Faktoren</i>	69
4.	Experiment	75
4.1.	Allgemeine Zielsetzung	75
4.2.	Forschungsfragen.....	76
4.3.	Methode	77
4.3.1.	<i>TeilnehmerInnen</i>	78
4.3.2.	<i>Material</i>	81
4.3.3.	<i>Versuchsaufbau und -ablauf</i>	84
4.3.4.	<i>Durchführung</i>	86
4.3.5.	<i>Scoring</i>	88
4.3.6.	<i>Versuchsplan und statistisches Analysekonzept</i>	89
4.4.	Ergebnisse	90
4.4.1.	<i>Deskriptivstatistische Auswertung</i>	90

4.4.2. Inferenzstatistische Auswertung.....	92
4.5. Diskussion	94
4.6. Schlussfolgerungen.....	99
5. Zusammenfassung	101
Bibliographie	103
Anhang	109
Anhang A: Instruktionen für TeilnehmerInnen.....	110
Anhang B: Stimulussätze des Reading Span Task (Auszug)	112
Anhang C: Stimulussätze des Listening Span Task (Auszug)	113
Anhang D: Protokolle	114
Anhang E: Screenshots	117
Anhang F: Statistische Datenauswertung.....	119
Abstract (Deutsch)	123
Abstract (Englisch)	124
Lebenslauf	125

Abbildungsverzeichnis

Abbildung 1: AG-Modell nach Baddeley	23
Abbildung 2: AG-Modell nach Cowan.....	26
Abbildung 3: Mittlere RST- und LST-Scores nach Gruppen.....	92
Abbildung 4: Interaktion zwischen Erfahrung und Modalität.....	94

Tabellenverzeichnis

Tabelle 1: Demographische Daten der Versuchsgruppen.....	80
Tabelle 2: RST- und LST-Ergebnisse nach Gruppen	91

Abkürzungsverzeichnis

AG	Arbeitsgedächtnis
FoA	Focus of Attention
KZG	Kurzzeitgedächtnis
LST	Listening Span Task
LZG	Langzeitgedächtnis
RST	Reading Span Task

0. Einleitung

Die Einführung des Arbeitsgedächtnisbegriffs im Sinne einer dynamischen und prozessorientierten Weiterentwicklung des Kurzzeitgedächtniskonzepts eröffnete in der Kognitionspsychologie neue Wege der Forschung. Von diesem Umdenken profitierte letztendlich auch die Dolmetschwissenschaft. Die Erkenntnis, dass gängige Arbeitsgedächtnismodelle und -versuchsanordnungen herangezogen werden können, um die kognitiven Prozesse beim Dolmetschen besser zu verstehen, bildete den Ausgangspunkt einer regen Auseinandersetzung mit diesem Thema innerhalb der Dolmetschwissenschaft.

Das Simultandolmetschen galt in der Wissenschaft schon lange als hochkomplexer Prozess. Mit dem Arbeitsgedächtnis (= AG) stand nun ein kognitives Konstrukt zur Verfügung, das zur Konzeptualisierung dieser Tätigkeit herangezogen werden konnte. Angesichts der vielschichtigen kognitiven Anforderungen beim Dolmetschen schien es offenkundig, dass professionelle DolmetscherInnen ein in gewisser Hinsicht besonders ausgeprägtes AG haben müssen. Um genauer zu erkunden, welche determinierende Eigenschaft des AG für erfolgreiches Dolmetschen verantwortlich gemacht werden kann, kam in den letzten zwei Jahrzehnten immer wieder ein von Daneman und Carpenter (1980) vorgestelltes Verfahren, der Lesespannentest (*reading span task* = RST), zum Einsatz. Von diesem Experiment erhoffte man sich Hinweise auf einen kausalen Zusammenhang zwischen Dolmetschkompetenz und AG-Kapazität. Bisher konnte diese Verbindung allerdings nicht eindeutig belegt werden. Sowohl die visuelle (RST) als auch die akustische Variante (*listening span task* = LST) dieses AG-Tests wurden in der Dolmetschwissenschaft mehrmals verwendet. Aus den Ergebnissen geht scheinbar kein klar interpretierbares Muster hervor.

Die vorliegende Masterarbeit stellt daher den Versuch dar, diesen rätselhaften empirischen Befund aus einer frischen Perspektive zu betrachten, um alternative Erklärungsansätze zur Beseitigung der Widersprüche zu liefern. Als Grundlage dient dabei das bereits erwähnte AG-Experiment von Daneman und Carpenter (1980), das im Rahmen dieser Untersuchung repliziert wurde. Durch ein angepasstes Versuchsdesign, bei dem die Modalität als zusätzliche unabhängige Variable realisiert wurde, war es mög-

lich, einer in der dolmetschwissenschaftlichen Literatur bisher wenig erforschten Facette auf den Grund zu gehen.

Die Abhandlung beginnt mit einer Erörterung der theoretischen Grundlagen. Zunächst erfolgt eine vergleichende Betrachtung zweier in der Dolmetschwissenschaft maßgeblicher AG-Modelle: Alan D. Baddeleys *multi-component model* und Nelson Cowans *embedded-processes model*. Anschließend wird ein Überblick über die AG-Forschung aus dolmetschwissenschaftlicher Sicht gegeben. Hier stehen Forschungsbeiträge im Mittelpunkt, die sich mit der Rolle des AG beim Simultandolmetschen befassen und mögliche AG-spezifische Besonderheiten bei ausgebildeten SimultandolmetscherInnen thematisieren.

Danach folgt eine Analyse der Entwicklung des interdisziplinären Austauschs zwischen Kognitionspsychologie und Dolmetschwissenschaft, die den Hintergrund für eine Beschreibung des Lesespannentests bildet. Dabei gilt es, der LeserIn durch eine detaillierte Besprechung der grundlegenden Konzepte des RST sowie damit im Zusammenhang stehender methodischer Überlegungen und Weiterentwicklungen ausreichend Hintergrundwissen zu bieten, um das darauf folgende Kapitel über die Verwendung dieses Verfahrens in der Dolmetschwissenschaft verstehen zu können.

Aus einer detaillierten Analyse der bisherigen RST/LST-Untersuchungen mit DolmetscherInnen geht hervor, dass der Faktor „Modalität“ der Schlüssel zur Entwirrung der auf den ersten Blick so heterogenen Forschungsergebnisse sein könnte. Um diesen Erklärungsvorschlag empirisch zu testen, wurde daher als Teil dieser Arbeit ein Experiment mit 18 BerufsdolmetscherInnen, 18 Dolmetschstudierenden und 18 Nicht-Dolmetschstudierenden durchgeführt, dessen entscheidendes Merkmal in der gekoppelten Abwicklung eines visuellen RST gemeinsam mit einem akustischen LST besteht. Dieser empirischen Unternehmung ist ein eigenes Kapitel gewidmet.

Abschließend wird die gesamte Arbeit zusammengefasst, wobei besonderes Augenmerk auf die aus dem Experiment gewonnenen Erkenntnisse gelegt wird. Die umfangreiche Dokumentation im Anhang soll zusätzlich einen Einblick in die methodische Vorgangsweise und die statistische Datenauswertung gewähren.

1. Kognitionspsychologische Grundlagen

1.1. Allgemeines

Die Gedächtnisforschung stellt einen Bereich der Kognitionspsychologie dar, dessen facettenreiche Wechselbeziehung mit der Dolmetschwissenschaft eine Vielzahl aufschlussreicher interdisziplinärer Untersuchungen hervorgebracht hat. An der Schnittstelle dieser Wissenschaftszweige wird das methodische Instrumentarium der Kognitionspsychologie verwendet, um einen Extremfall menschlicher Kognitionsleistungen, das Simultandolmetschen, unter besonderer Berücksichtigung der Rolle des Gedächtnisses näher zu beleuchten.

Die kognitionswissenschaftliche Forschung erkannte schon früh, dass das Gedächtnis keineswegs als einheitliches, unteilbares Ganzes konzeptualisiert werden kann. James (1890) postulierte in seiner grundlegenden Darstellung der Psychologie eine Zweiteilung in „primäres“ und „sekundäres Gedächtnis“. Dabei entspricht das primäre Gedächtnis dem moderneren Begriff des Kurzzeitgedächtnisses (= KZG), das begrenzte Informationsmengen für kurze Zeit im Bewusstsein aufrecht erhält; und das sekundäre Gedächtnis deckt sich weitgehend mit dem Langzeitgedächtnis (= LZG), das als dauerhafter Speicher einer nahezu unbegrenzten Menge an Gedächtnisinhalten charakterisiert wird.

Bis in die 1960er-Jahre herrschte die Ansicht vor, dass das KZG ausschließlich als passive Speicherkomponente innerhalb des Informationsverarbeitungsprozesses zu bewerten sei. Die Ergebnisse zahlreicher empirischer Untersuchungen gaben jedoch bald Anlass, diese Sichtweise zu überdenken. Erstmals begann man dem KZG in der Forschungsliteratur (Atkinson & Shiffrin 1968; Baddeley & Hitch 1974) eine größere Bedeutung als bisher zuzuschreiben, wobei für diese erweiterte Systemkomponente die Bezeichnung „Arbeitsgedächtnis“ (= AG; *working memory*) eingeführt wurde. Die wesentliche Neuerung besteht darin, dass das KZG nun nicht mehr nur die Funktion eines passiven Zwischenspeichers erfüllt, sondern als aktiver Bestandteil kognitiver Gedächtnisprozessmodelle aufgefasst wird. Das Arbeitsgedächtnis ist also in der Lage, Informationen vorübergehend zu behalten, diese entsprechend aktueller Zielsetzungen des Indi-

viduums zu bearbeiten und Zwischenergebnisse dieser Operationen wiederum zu speichern.

Die ersten zwei Kapitel dieser Arbeit haben zum Ziel, die vielschichtigen Verflechtungen dieses dynamischen Arbeitsgedächtnisbegriffs mit der Tätigkeit des Simultandolmetschens zu untersuchen. Insbesondere gilt es, auf der Grundlage aktueller Forschungsergebnisse näher darauf einzugehen, wie sich SimultandolmetscherInnen hinsichtlich der Funktionsweise des Arbeitsgedächtnisses von anderen Personengruppen unterscheiden könnten. Die Ausübung dieses Berufs setzt die Fähigkeit voraus, die durch simultanes Hören und Sprechen verursachte ständige Belastung des AG-Systems so weit in den Griff zu bekommen, dass eine zufriedenstellende Dolmetschleistung möglich ist. In der Forschungsliteratur werden verschiedene Faktoren genannt, um das Funktionieren dieser besonders komplexen kognitiven Tätigkeit verständlich zu machen. Es soll hier auf Erklärungsansätze eingegangen werden, die dem AG dabei eine zentrale Bedeutung einräumen. Die in der kognitiven Psychologie mit dem AG-Begriff eng verknüpften Themenbereiche der Aufmerksamkeit und des LZG werden im Rahmen der vorliegenden Arbeit nur in dem Ausmaß behandelt, als sie für das Verständnis darauf aufbauender AG-Aspekte nötig sind.

1.2. AG als Forschungsgegenstand der kognitiven Psychologie

Die von James (1890) postulierte duale Gliederung in ein primäres und ein sekundäres Gedächtnis konnte im Laufe des 20. Jahrhunderts durch empirische Untersuchungen und Erkenntnisse aus der Neuropsychologie bestätigt werden. Wesentlich erwiesen sich in diesem Zusammenhang neurophysiologische Befunde von gehirngeschädigten Patienten. Es wurden entgegengesetzte Gedächtnissyndrome beobachtet, die bei einem Teil der Fälle bewirkten, dass das LZG durch Läsionen in bestimmten Arealen beeinträchtigt war, während das KZG weiterhin normal funktionierte, und bei einem anderen Teil das umgekehrte Phänomen hervorriefen, d.h. ein intaktes LZG und ein dysfunktionales KZG. Es herrscht daher weitgehend Einigkeit darüber, dass man von zwei unabhängigen Gedächtnissystemen, dem LZG und dem KZG, die auch in verschiedenen kortikalen Bereichen verortet werden können, ausgehen muss (vgl. Darò 1997: 623; Gruber 2011: 16).

Als besonders richtungsweisend erwies sich bei der Erforschung des KZG die Feststellung Millers ([1956]/1994), dass die Anzahl der Einheiten, die ein Mensch kurzzeitig auf einmal bewusst im Gedächtnis zu behalten vermag, auf 7 ± 2 beschränkt ist. Die Gedächtnisspanne lässt sich laut Miller jedoch durch den Vorgang des *recoding* (oder *chunking*) erheblich erweitern. Dabei werden kleinere Informationseinheiten mit Hilfe sinnvoller Assoziationen zu immer größeren Einheiten, oder *chunks*, zusammengefasst. Diesem Ansatz folgend könnte man sich beispielsweise die Buchstabenkette N–G–O–U–N–O–N–A–T–O viel leichter merken, wenn man sie folgendermaßen gruppiert: NGO–UNO–NATO (vgl. Gruber 2011: 43; Miller [1956]/1994: 349f.).

Lag der Schwerpunkt dieser frühen Phase der kognitiven Gedächtnisforschung noch auf der empirischen Ermittlung der genauen Kapazität der Speicherkomponenten des Menschen, so setzte in den 1960er- und 1970er-Jahren eine grundlegende Neuorientierung ein, die der Forderung vieler Wissenschaftler nach einer feineren strukturellen Gliederung des Gedächtnissystems, die über die starre Zweiteilung in KZG und LZG hinausging, Rechnung trug. Die Modelle wurden zunehmend komplex, da man bemüht war, auf der Grundlage der aus zahlreichen Experimenten gewonnenen Erkenntnisse ein überprüfbares strukturelles Gefüge für gedächtnisrelevante Prozesse zu erkennen.

Baddeley und Hitch (1974) gehören zu den prominentesten Vertretern dieser stark strukturorientierten Denkrichtung. Als richtungsweisend für die gesamte Disziplin erwies sich ihre Auffassung des KZG als multimodulares AG, das nicht mehr nur der passiven Speicherung von Gedächtniseinträgen dient, sondern auch aktiv an der Informationsverarbeitung beteiligt ist. Die überwältigende Mehrheit der Untersuchungen, welche die Rolle des AG beim Simultandolmetschen zum Thema haben, verwendet das von Baddeley formulierte Modell als Ausgangspunkt, weshalb auch in der vorliegenden Arbeit eine Beschreibung hiervon sinnvoll ist.

In der neueren dolmetschwissenschaftlichen Forschung gewinnt neben der noch immer sehr dominanten Sichtweise Baddeleys auch zunehmend ein Ansatz an Einfluss, bei dem weniger die Gliederung des Systems in modalitätsspezifische Bestandteile mit klar definierten Aufgaben, sondern das zielorientierte Zusammenspiel verschiedener Prozesse innerhalb eines integrativen und modalitätsunabhängigen Gesamtsystems im Vordergrund steht. Maßgeblich an der Entwicklung dieser funktional-prozessualen Auffassung des AG beteiligt war Cowan (1988), der zu den prominentesten Vertretern die-

ser jüngeren Strömung in der kognitiven Gedächtnisforschung zählt. Nach einer Auseinandersetzung mit dem von Cowan vorgeschlagenen AG-Modell wird in diesem Kapitel abschließend eine vergleichende Betrachtung vorgenommen, die wesentliche Unterschiede, aber auch Überlappungen zwischen Baddeley und Cowan zum Vorschein bringen soll.

1.3. Zwei Arbeitgedächtnismodelle

1.3.1. AG-Modell nach Baddeley

Dem AG-Modell von Baddeley liegt ein multimodulares Konzept zu Grunde, dessen ursprüngliche Version (Baddeley & Hitch 1974) drei spezialisierte Einzelkomponenten umfasste: die zentrale Exekutive (*central executive*), die phonologische Schleife (*phonological loop*) und den räumlich-visuellen Notizblock (*visuo-spatial sketchpad*). Baddeley (2000a) erkannte jedoch, dass sein Modell in gewissen Punkten, vor allem betreffend die Verbindung des AG mit dem LZG und die Verschmelzung von Stimuli verschiedener Modalitäten (visuell und akustisch) zu neuen episodischen Gedächtniseinheiten, problematisch war. Durch die Einführung eines neuen Subsystems, des episodischen Puffers (*episodic buffer*), will er neuen Erkenntnissen aus empirischen Untersuchungen Rechnung tragen und bringt sein Modell so auf den aktuellen Stand (vgl. Timarová 2008: 4f.; s. Abb. 1).

Die phonologische Schleife und der räumlich-visuelle Notizblock erfüllen jeweils die Funktion eines modalitätsspezifischen Informationsspeichers, wobei in der phonologischen Schleife temporär akustische und sprachliche Einheiten und im räumlich-visuellen Notizblock vorübergehend Informationen behalten werden, die visuelle und räumliche Wahrnehmung voraussetzen. Diese zwei Speicherkomponenten sind der Kontrolle der zentralen Exekutive unterworfen, die für die Koordination der Prozesse, die Aufmerksamkeitszuwendung und die Aktivierung von LZG-Repräsentationen zuständig ist. Der erst viel später eingeführte episodische Puffer wird ebenfalls als der zentralen Kontrollinstanz untergeordnete Speichereinheit betrachtet. Die Informationen in diesem Subsystem stammen aus verschiedenen Kanälen und werden hier mit Hilfe eines multimodalen Codes zu kohärenten Episoden zusammengefügt. Eine wesentliche

Neuerung des überarbeiteten Modells nach Baddeley besteht außerdem darin, dass nun explizite Verbindungen zwischen dem LZG, den *crystallized systems* in Abbildung 1, und allen Subsystemen postuliert werden (Baddeley [2001]/2002; Baddeley & Logie 1999).

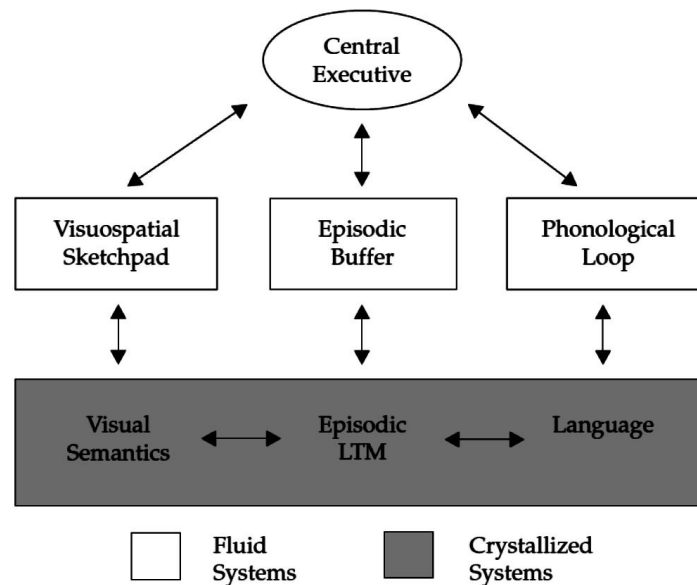


Abbildung 1

AG-Modell nach Baddeley ([2001]/2002: 93)

Der Untersuchung der phonologischen Schleife wurde in der Forschung bisher am meisten Aufmerksamkeit geschenkt. Da die Verarbeitung verbaler Reize in diesem Zusammenhang im Zentrum steht, ist es nachvollziehbar, dass die Dolmetschwissenschaft an einer Implementierung dieses AG-Modells in dolmetschrelevanten Experimenten äußerst interessiert ist. Die phonologische Schleife kann laut Baddeley und Logie (1999) in zwei weitere Module unterteilt werden, einen passiven phonologischen Speicher (*phonological store*) und einen aktiven artikulatorischen Kontrollprozess (*rehearsal process* oder *articulatory control process*).

Der phonologische Speicher ermöglicht das Aufrechterhalten akustisch dargebotener Reize. Die Kapazität dieses Subsystems ist aber begrenzt, da nur so viele Wörter kurzzeitig behalten werden können, wie eine Person in ungefähr zwei Sekunden zu artikulieren vermag. Um verbale Informationen länger als zwei Sekunden im Gedächtnis zu behalten, ist es nötig, den artikulatorischen Kontrollprozess einzuschalten, der einer inneren Stimme gleicht und durch den ein behaltendes Wiederholen (*subvocal rehearsal*) erfolgt.

sal) des verbalen Materials durchgeführt wird. Gedächtnisspuren, die nicht durch diesen Rehearsal-Prozess aufgefrischt werden, zerfallen und können nicht mehr abgerufen werden. Wird hingegen verbales Material visuell präsentiert, so kommt zuerst der artikulatorische Kontrollprozess zum Einsatz, um es zunächst in einen phonologischen Code umzuwandeln. Anschließend können die umgewandelten Gedächtniseinheiten in die phonologischen Speicher transferiert und durch den Rehearsal-Prozess beliebig lang behalten werden (vgl. Baddeley & Logie 1999: 32; Gruber 2011: 39; Timarová 2008: 4).

Ein für die Tätigkeit des Simultandolmetschens besonders relevantes Phänomen ist die sogenannte „artikulatorische Suppression“ (*articulatory suppression*), die durch die Funktionsweise der phonologischen Schleife erklärt werden kann. Dieses Phänomen wird dadurch charakterisiert, dass sich die phonologische Behaltensleistung einer Person drastisch verschlechtert, wenn gleichzeitiges Zuhören und Sprechen erforderlich sind. Bei Versuchen zur Messung der Gedächtnisspanne forderte man die TeilnehmerInnen auf, sich Zahlen- oder Wortfolgen einzuprägen und unmittelbar danach in der dargebotenen Reihenfolge wiederzugeben. Die Anzahl der Items wurde gesteigert, bis keine korrekte Reproduktion mehr gelang. Die Item-Anzahl der letzten korrekt abgerufenen Item-Folge entsprach der Gedächtnisspanne. Verlangte man von den ProbandInnen in der Lernphase, ständig irrelevante Laute zu vokalisieren (z.B. „Bah, bah, bah, ...“), so konnte beobachtet werden, dass der unmittelbare Abruf erheblich beeinträchtigt war. Auf der Grundlage des AG-Modells von Baddeley erscheint folgender Erklärungsansatz plausibel: Durch das gleichzeitige Artikulieren (*concurrent articulation*) von sinnlosen Silben wird der artikulatorische Kontrollprozess in Anspruch genommen. Daher steht der behaltende Wiederholungsmechanismus (*subvocal rehearsal*) nicht oder nur mehr sehr eingeschränkt für die Auffrischung akustischer Spuren im phonologischen Speicher zur Verfügung und die zu memorierenden Items unterliegen einem stärkeren Zerfall (vgl. Chincotta & Underwood 1998: 2-5; Gruber 2011: 40f.).

In der Dolmetschwissenschaft wurde vielfach darauf hingewiesen, dass die SimultandolmetscherIn in einer kognitiven Konfiguration arbeitet, die genau den Voraussetzungen für das Auftreten artikulatorischer Suppression entspricht (vgl. Mizuno 2005: 740f.). Die beim Dolmetschen zu beobachtenden Auswirkungen des Phänomens sollten diesem Ansatz zufolge sogar noch deutlicher ausfallen, da das Arbeitsgedächtnis hier

eine wesentlich komplexere Aufgabe zu bewältigen hat als bei Experimenten, die serielle Reproduktionsaufgaben mit gleichzeitiger Artikulation umfassen. Dass die VertreterInnen dieser Berufsgruppe trotzdem zufriedenstellende Leistungen erbringen können, bedarf einer Erklärung. Mögliche AG-spezifische Aspekte, die der SimultandolmetscherIn dabei helfen, die Beeinträchtigung der phonologischen Schleife möglichst gering zu halten oder vielleicht sogar durch andere Mechanismen zu umgehen, werden in den folgenden Kapiteln besprochen.

Weitaus weniger eingehend erforscht ist der räumlich-visuelle Notizblock, für den eine ähnliche duale Struktur und vergleichbare Mechanismen wie bei der phonologischen Schleife postuliert werden. Die meisten Versuchsparadigmen haben sich bisher jedoch auf die akustisch-verbale Modalität konzentriert. Baddeley und Logie (1999: 32) gehen von einem passiven Subsystem für die Speicherung visueller Reize (*visual cache*) und einem aktiven Subsystem für die Manipulation räumlicher Gedächtnisinhalte (*inner scribe*) aus. Weniger Klarheit herrscht darüber, wie sich im räumlich-visuellen Notizblock der Rehearsal-Prozess vollzieht. In dolmetschwissenschaftlichen Arbeiten findet diese Komponente kaum Beachtung.

Als viel grundlegender für die Erklärung gedächtnisrelevanter Phänomene beim Simultandolmetschen kann die zentrale Exekutive betrachtet werden, wobei auch hier einschränkend angemerkt werden muss, dass der Forschungsstand in diesem Bereich im Vergleich zur Untersuchung der phonologischen Schleife weniger weit fortgeschritten ist. Baddeley und Logie sind sich der Gefahr bewusst, dass diese Kontrollinstanz einfach als „(...) useful ragbag to contain all phenomena that cannot be readily accounted for otherwise“ (1999: 39) verwendet wird. Diese AG-Komponente wird im Allgemeinen mit Kontroll-, Steuerungs- und Koordinationsaufgaben in Verbindung gebracht. Zu ihren Aufgaben gehören also die Steuerung und Koordination der Subsysteme, die Aufmerksamkeitsverlagerung und -fokussierung sowie die Aktivierung von Repräsentationen im LZG. Der zentralen Exekutive daneben auch eine Speicherfunktion zuzuteilen, halten Baddeley und Logie (1999: 37) für nicht plausibel.

1.3.2. AG-Modell nach Cowan

Nelson Cowan führte in seinen Arbeiten (1988; 1995; 1999) ein AG-Modell ein, das die Ansichten der funktionalen Gedächtnisforschung widerspiegelt. Dabei steht die Rolle des AG bei der Erfüllung eines vom Bewusstsein bestimmten Zieles und nicht die Struktur des Systems wie bei Baddeley im Vordergrund. Entscheidend bei Cowans Ansatz ist die Einbettung des AG in das LZG und die enge Verknüpfung des Gedächtnissystems mit der Aufmerksamkeit.

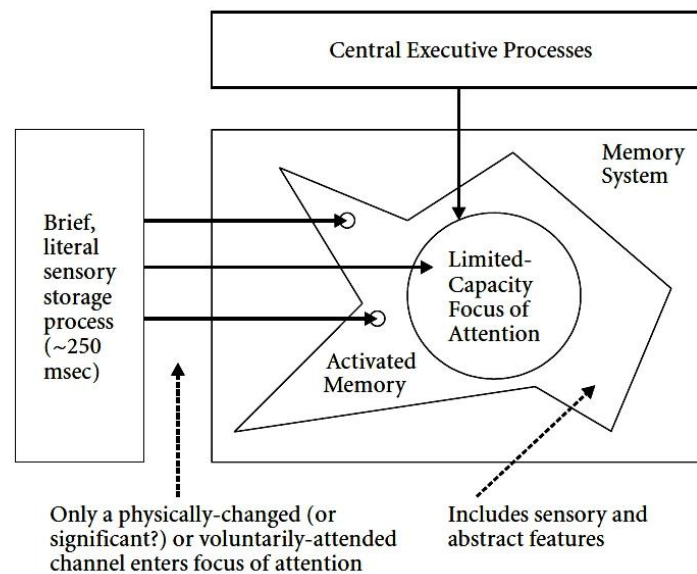


Abbildung 2

AG-Modell nach Cowan (2000: 134)

Die im AG verarbeiteten Inhalte stammen aus hierarchisch abgestuften Aktivierungszonen innerhalb des LZG. Das LZG wird als unterster Bereich aufgefasst. Die Gedächtnisinhalte sind hier grundsätzlich in einem passiven Zustand, können aber je nach kognitiver Zielsetzung durch die richtigen Hinweisreize aktiviert und zur Verarbeitung bereitgestellt werden. Darüber befindet sich in der Hierarchie ein aktivierter Teilbereich des LZG, das sogenannte „aktivierte Gedächtnis“ (*activated memory*). Diese Informationen befinden sich zwar außerhalb des Bewusstseins eines Individuums. Wegen ihres erhöhten Aktivierungsgrades sind sie jedoch wesentlich leichter abrufbar als Items im LZG. Auf der höchsten Aktivierungsstufe befinden sich schließlich solche

Informationselemente, die gerade vom Fokus der Aufmerksamkeit (*focus of attention* = FoA) verarbeitet werden (Cowan 1999; Gruber 2011: 44-49; s. Abb. 2).

Der Fokus der Aufmerksamkeit wird einerseits durch die zentrale Exekutive (*central executive system*) gesteuert, wenn es darum geht, bewusst und willentlich bestimmte Elemente innerhalb des Systems für momentane kognitive Prozesse abzurufen. Andererseits ist es auch möglich, dass eine plötzlich auftretende Veränderung eines Reizkanals (z.B. ein Aufschrei während einer Vorlesung) zu einer automatischen Aufmerksamkeitsverschiebung (*automatic attending*) führt. Um auf diese Weise in den Fokus der Aufmerksamkeit zu gelangen, muss der Reiz eine gewisse Aktivierungsschwelle überschreiten. Das AG-System ist nämlich nach Cowans Auffassung mit einer Vielzahl von eintreffenden Reizströmen konfrontiert. Durch einen Filtermechanismus werden jene Stimuli, die für die momentanen kognitiven Abläufe irrelevant sind, zu einem gewissen Grad ausgeblendet (*habituation*). Tritt ein neuer Reiz in Erscheinung, so kann dieser, wenn er stark genug ist, automatisch Aufmerksamkeit für sich rekrutieren (*dis-habituation*). Stimuli, die sich im Fokus der Aufmerksamkeit befinden, werden vollständiger verarbeitet (Cowan 1999). Dem AG ist außerdem ein kurzlebiger sensorischer Speicher (*sensory store*) vorgelagert, in dem die Reize vorübergehend in einer rohen, rein physikalischen Form behalten werden, bevor sie ins eigentliche Gedächtnissystem vordringen. Dort erfolgt die Umwandlung physikalischer Merkmale in konzeptuelle Repräsentationen (vgl. Gruber 2011: 45). Aktivierte Informationseinheiten, die gleichzeitig oder nacheinander in den Fokus der Aufmerksamkeit gelangen, werden oft zu neuen Kombinationen verknüpft und anschließend im LZG gespeichert (z.B. neue Vokabeln oder neue Erlebnisse; vgl. Cowan 1999: 63).

Zwei verschiedene Beschränkungen werden von Cowan (1999: 68) hinsichtlich seines AG-Modells postuliert: Das aktivierte Gedächtnis unterliegt einer empirisch ermittelten Zeitbegrenzung von zehn bis zwanzig Sekunden. Nach diesem Zeitabstand beginnt die Aktivierung in der betreffenden Zone nachzulassen. Der Fokus der Aufmerksamkeit hingegen hat ein begrenztes Fassungsvermögen. Die Anzahl nicht-assoziiierter Elemente, die auf einmal bewusst erfasst werden können, beträgt laut Cowan drei bis fünf, wobei diese Zahl mit Hilfe von Chunking- und Strukturierungsstrategien auch erhöht werden kann. Gedächtniseinträge werden aufrechterhalten, indem sie immer wieder durch den Fokus der Aufmerksamkeit zirkuliert werden (*rehearsal*).

1.4. Vergleichende Diskussion der AG-Modelle

Wie aus den bisherigen Ausführungen hervorgegangen ist, wählen Baddeley und Cowan für ihre AG-Modelle jeweils einen anderen Schwerpunkt. Baddeley und Logie (Baddeley & Logie 1999) betrachten das AG als Summe klar abgegrenzter Bausteine, während Cowan (1999) eine integrative Sichtweise vertritt, bei der alle Mechanismen zum AG gerechnet werden, die zur erfolgreichen Erfüllung einer Aufgabe beitragen. Die Schöpfer dieser Modelle weisen in ihren Arbeiten (Baddeley 2003; Cowan 1999) jedoch auch darauf hin, dass Parallelen existieren. Hier sei zunächst ein Überblick über grundlegende Unterschiede gegeben, bevor Gemeinsamkeiten ausfindig gemacht werden.

Das wesentliche Charakteristikum des AG-Modells von Cowan ist darin zu sehen, dass keine strukturelle Trennung zwischen LZG und AG vorgenommen wird. Für ihn ist das AG eine aktivierte Teilmenge des LZG. Dadurch sind zahlreiche kognitive Prozesse hoher Komplexität, die erwiesenermaßen auf LZG-Inhalte zurückgreifen, wie beispielsweise das Lesen, leichter erklärbar, als es mit Hilfe des Ansatzes von Baddeley möglich wäre, der mehr Wert auf eine saubere Gliederung in LZG und AG legt. Mit der Einführung des episodischen Puffers versuchte Baddeley (2000a) die Kritik zu entschärfen, die auf Grund dieser mit dem empirischen Befund unvereinbaren Entkoppelung zwischen LZG und AG aufgekommen war.

Der episodische Puffer stellt gleichzeitig den Versuch dar, eine weitere Schwäche in Baddeleys Konzept auszumerzen. Während das AG nach Cowan relativ flexibel mit der Verarbeitung von Reizen unabhängig von deren Modalität umgehen kann und mühelos die Verknüpfung von Reizen verschiedener Modalitäten zu neuen LZG-Einträgen konzeptualisiert, wirkte das AG-Modell von Baddeley vor der Einführung des episodischen Puffers vergleichsweise starr, weil es nur zwei Modalitäten, die visuelle und verbale, berücksichtigte und die eintreffenden Reize streng getrennt nach Modalität in unabhängigen Systemen gespeichert wurden. Der episodische Puffer ermöglicht es, Gedächtniseinheiten aus den modalitätsspezifischen Komponenten zu neuen Kombinationen, Episoden genannt, zu verschmelzen. Diese Ergänzung greift jedoch m. E. noch immer zu kurz, da bei weitem nicht das ganze Spektrum sensorischer Informationsquellen abgedeckt ist. Cowan (vgl. 1999: 71) gibt zu bedenken, dass haptische und nonver-

bale Reize nicht ausreichend zur Geltung kommen. Cowan (1999) betrachtet visuelle und verbale Reize als nur zwei von einer Vielzahl von Stimuli, die bestimmte Merkmale im LZG aktivieren.

Wie bereits angedeutet, überlappen sich die zwei Erklärungsansätze aber auch in einigen Punkten. Sie dürfen nicht als grundsätzlich inkompatibel verstanden werden (vgl. Baddeley 2003: 837). Beide Modelle postulieren eine zentrale Exekutive, die ähnliche Steuerungs- und Koordinationsaufgaben übernimmt und weitgehend synonym mit dem Begriff der Aufmerksamkeit verstanden wird. Gemeinsam ist den Auffassungen außerdem, dass gewisse Beschränkungen für die AG-Komponenten vorgesehen sind. Während bei Baddeley die phonologische Schleife einer zeitlichen Speicherungslimitierung von ca. zwei Sekunden unterworfen ist, sind bei Cowan zwei verschiedene Arten der Begrenzung definiert: eine zeitliche im Falle des aktivierten Gedächtnisses (zehn bis zwanzig Sekunden) und eine kapazitätsbezogene hinsichtlich des Fokus der Aufmerksamkeit (drei bis fünf Chunks). Eine weitere Ähnlichkeit ist in der Funktionsweise des Fokus der Aufmerksamkeit und der phonologischen Schleife festzustellen: Das Behalten von Informationseinheiten wird durch einen Rehearsal-Prozess bewerkstelligt, bei dem die Inhalte in der jeweiligen Komponente wiederholt zirkuliert werden.

Die Tatsache, dass das AG-Modell von Baddeley in der Dolmetschwissenschaft am weitesten verbreitet ist und am öftesten zitiert wird, ist wohl darauf zurückzuführen, dass es aus klar abgegrenzten Bestandteilen konstruiert ist und daher empirisch leichter untersucht werden kann. Gegen diese Konstruktion wäre einzuwenden, dass sie zu starr ist und daher der ganzen Komplexität der Tätigkeit des Simultandolmetschens nicht gerecht werden kann.

Die für die Dolmetschwissenschaft wahrscheinlich aussichtsreichere Konzeptualisierung des AG ist jene von Cowan. Sein Modell ist flexibler und prozessorientierter. Es ermöglicht eine bessere Einbindung pragmatischer Faktoren und stellt eine klare Verknüpfung zum LZG her. Gleichzeitig bedeutet die weniger klare Struktur, dass sich eine empirische Erforschung hier schwieriger gestaltet als beim AG-Modell von Baddeley.

2. Das Arbeitsgedächtnis beim Dolmetschen

2.1. Der AG-Begriff in der Dolmetschwissenschaft

Bereits wenige Jahre nach der Veröffentlichung der bahnbrechenden Arbeit von Baddeley und Hitch (1974), die das Konzept des AG in den Mittelpunkt der kognitionswissenschaftlichen Forschung gerückt hatten, wurde dieser Terminus auch in Verbindung mit dem Dolmetschen aufgegriffen. Gerver verwendete die Bezeichnung *short-term working memory* in einem wissenschaftlichen Beitrag, in dem er sein Modell des Informationsverarbeitungsprozesses beim Simultandolmetschen darlegte. Gerver hält dieses Kurzzeit-AG für notwendig „(...) in order to account for the ability of interpreters to translate while further information is received (...)“ (vgl. 1976: 191). Es geht aus Gervers Ausführungen hervor, dass er diese Komponente als einen von mehreren Puffermodulen im Informationsverarbeitungsprozess versteht. Zwar fehlt eine präzise Definition des Begriffs in Gervers Beschreibung und eine direkte Bezugnahme auf Baddeley und Hitch (1974) ist zu dieser Zeit noch unwahrscheinlich. Die Verwendung dieses Ausdrucks erwies sich jedoch als zukunftssträchtig.

Bald erkannte man in der Dolmetschwissenschaft die zentrale Bedeutung des von der kognitiven Psychologie beschriebenen AG beim Dolmetschen. Es war empirisch erwiesen, dass das AG bei anspruchsvollen kognitiven Aufgaben eine wichtige Rolle spielt. Da das Dolmetschen eine kognitive Tätigkeit der höchsten Komplexitätsstufe ist, lag die Schlussfolgerung nahe, dass auch hierbei das AG grundlegend sei. Das Dolmetschen stellt für KognitionspsychologInnen grundsätzlich eine sehr interessante Verarbeitungssituation dar, da das AG auf Grund der zahlreichen, gleichzeitig ablaufenden Prozesse auf das Äußerste belastet wird. Dieser gedächtnisstrapazierende Extremzustand steht im Mittelpunkt des Interesses vieler ForscherInnen aus Kognitionspsychologie und Dolmetschwissenschaft und zeitigte eine Fülle an wissenschaftlichen Publikationen.

Die kognitive Konfiguration beim Simultandolmetschen kann als ein für den Menschen ungewöhnlicher Informationsverarbeitungsmodus betrachtet werden, der aktuellen AG-Modellen zufolge zu erheblichen Kapazitätsengpässen im AG-System führen sollte. Dementsprechend wäre zu erwarten, dass das Funktionieren dieser Tätig-

keit kaum möglich bzw. stark beeinträchtigt ist. Dem widerspricht aber die Tatsache, dass das Simultandolmetschen sehr wohl möglich ist. Die Auswirkungen dieser extremen AG-Belastung sind zwar feststellbar, aber keineswegs katastrophal für die Dolmetschleistung einer ausgebildeten DolmetscherIn. Offensichtlich ist es irgendwie zu schaffen (vgl. Chincotta & Underwood 1998: 5; Mizuno 2005: 740). Ein Großteil der Untersuchungen zu diesem Thema basiert auf der Annahme, dass das AG von SimultandolmetscherInnen besondere Eigenschaften bezüglich Funktionsweise, Kapazität oder Struktur besitzen muss, die sie von Nicht-SimultandolmetscherInnen oder DolmetschstudentInnen unterscheiden und erfolgreiches Dolmetschen erlauben. Die konkreten Unterschiede, die dabei hinsichtlich des AG vorgeschlagen werden, sind Gegenstand weiterer Unterkapitel.

Die Einbindung des kognitionswissenschaftlichen AG-Begriffs in die Dolmetschwissenschaft hat sich als sehr aufschlussreich für beide Disziplinen erwiesen. Lange Zeit diente das AG-Modell von Baddeley als Grundlage dieses interdisziplinären Forschungsbereichs (z.B. Chincotta & Underwood 1998; Darò 1995; Darò 1997; Darò & Fabbro 1994; Isham 2000; Padilla *et al.* 2005; Padilla *et al.* 1995; Viezzi 1990), wobei oft AG-Kapazitätsunterschiede zwischen professionellen DolmetscherInnen und Nicht-DolmetscherInnen bzw. DolmetschstudentInnen im Vordergrund standen, weil man davon ausging, dass dieser Aspekt für das Funktionieren des Simultandolmetschens entscheidend sei. Eine relativ neue Tendenz in der Dolmetschwissenschaft ist die Verwendung des AG-Modells von Cowan (z.B. Mizuno 2005). Parallel dazu ist in den letzten Jahren in der AG-bezogenen Dolmetschforschung vermehrt ein Ansatz anzutreffen, der keine konkrete Zuordnung zu einem AG-Modell zulässt, sondern eine versöhnliche Position zwischen den Forschungstraditionen darstellt. Brauchbare Aspekte verschiedener Denkweisen werden unterstrichen und machen Überbrückungen möglich, die dem besseren Verständnis AG-relevanter Phänomene dienen sollen (z.B. Köpke & Nespoulous 2006; Liu *et al.* 2004; Shlesinger 2003). Zunächst erscheint es zielführend, eine kurze Beschreibung methodischer Vorgangsweisen bei der Erforschung dieser Phänomene einzuschieben. Anschließend kann eine beispielhafte Auswahl dolmetschwissenschaftlicher Studien im Bereich der AG-Forschung erfolgen.

2.2. Experimentelle Methoden

Im Bereich der Gedächtnisforschung trägt eine Vielzahl von Quellen dazu bei, dass wissenschaftliche Fortschritte erzielt werden. Neben der Entwicklungsforschung, den bildgebenden Verfahren und der Neuropsychologie bildet vor allem die experimentelle Kognitionspsychologie die empirische Grundlage für die Erforschung des AG. Von Anfang an stützte sich die Dolmetschwissenschaft auf die empirischen Methoden der kognitiven Psychologie. Erst in den letzten Jahren begannen sich experimentelle Methoden durchzusetzen, die die externe Validität der Untersuchungen durch die Einbindung dolmetschspezifischer Aufgaben bei Versuchen erhöhen sollen.

Die Experimentalanordnungen zur Erforschung der Rolle des AG beim Simultandolmetschen sind von einer großen Heterogenität gekennzeichnet. Die Basis für viele Versuche bilden die von KognitionspsychologInnen entwickelten Gedächtnisspannungsaufgaben (*memory span tasks*), mit denen versucht wird, die AG-Leistung der ProbandInnen zu quantifizieren. Man unterscheidet im Wesentlichen zwei Arten von experimentellen Vorgehensweisen, die es erlauben, die AG-Kapazität zu ermitteln.

Bei einfachen Gedächtnisspannungsaufgaben (*simple span tasks*) steht meist einzig die statische Speicherkapazität des AG im Mittelpunkt. In der einfachsten Konstellation wird der Versuchsperson eine Reihe von Wörtern oder Ziffern während einer Lernphase präsentiert. Diese Informationseinheiten müssen im Gedächtnis behalten werden (*retention*), bis die Aufforderung erfolgt, dieses Material wiederzugeben. Die Wiedergabe kann seriell (*serial recall*), d.h. in der dargebotenen Reihenfolge, oder frei (*free recall*) erfolgen. Die Anzahl der zu behaltenden Items wird innerhalb eines Durchgangs ständig erhöht, bis eine fehlerfreie Reproduktion nicht mehr möglich ist. Als Gedächtnisspanne gilt bei diesen Tests die längste Itemkette, die von der Versuchsperson richtig abgerufen werden kann. Es wäre allerdings ein Fehler, die so ermittelte Größe als Maß der gesamten AG-Leistung eines Menschen zu betrachten. Zwar wird die Kapazität der passiven Speicherkomponente des AG dadurch ermittelt. Allein die für das AG essentiellen exekutiven Aspekte, nämlich die Verarbeitung von Gedächtnisinhalten und die Prozesskoordination, werden dabei nicht berücksichtigt. Wir können daher festhalten, dass mit dieser Methode nur die Kapazität des passiven KZG gemessen wird.

Um die Leistung des AG vollständig zu erfassen, entwarf man komplexe Gedächtnisspannenaufgaben (*complex span tasks*). Diese Ausprägung von Gedächtnistests ermöglicht es, die kombinierte Leistung der beiden Dimensionen des AG, des passiven Speichers und des aktiven Verarbeitens von Informationen, zu messen und gilt als Standardtest für die AG-Spanne. Der klassische Versuchstyp ist in dieser Kategorie der von Daneman und Carpenter (1980) eingeführte Lesespannentest (*reading span test*). Die Versuchsperson muss hierbei einen Block von visuell präsentierten Sätzen laut vorlesen und sich das letzte Wort jedes Satzes einprägen. Nach jedem Block wird von der ProbandIn verlangt, dass sie die finalen Wörter in der dargebotenen Reihenfolge wiedergibt. Die Größe der Satzgruppen wird gesteigert, bis der fehlerfreie Abruf der letzten Wörter nicht mehr gelingt. Die AG-Spanne wird mit der Anzahl der Sätze in der letzten korrekt wiedergegebenen Serie gleichgesetzt. Es wurde beobachtet, dass eine sehr enge Beziehung zwischen komplexen Gedächtnisspannenaufgaben und kognitiv anspruchsvollen Tätigkeiten, wie dem Lesen oder logischen Denken, besteht (vgl. Daneman & Carpenter 1980; Gathercole 2008: 42f.). Da beim Simultandolmetschen ebenfalls sehr komplexe kognitive Prozesse ablaufen, kommt diese Art von Experiment in der Dolmetschwissenschaft oft zur Anwendung. Der Lesespannentest soll in weiteren Kapiteln noch ausführlicher behandelt werden, da er die Grundlage der im Rahmen dieser Arbeit durchgeführten empirischen Untersuchung bildet.

Die zwei besprochenen Versuchsparadigmen können auf verschiedene Weisen erweitert oder modifiziert werden, um bestimmte Aspekte besser zu operationalisieren. Zahlreiche Parameter stehen der VersuchsleiterIn zur Verfügung, um die für ihre Fragestellung sinnvollen Schwerpunkte zu setzen. So steht es ihr beispielsweise offen, den Darbietungsmodus (visuell oder akustisch) oder das Versuchsmaterial (Wort, Nichtwort, Pseudowort, Silben, Buchstaben, Ziffern etc.) zu variieren. Eine in der dolmetschwissenschaftlichen Forschung beliebte Variante des klassischen Lesespannentests ist der Hörspannentest (*listening span task*), bei dem statt des visuellen Darbietungsmodus eine akustische Präsentation der Satzsequenzen erfolgt. Diese Methode entspricht der Tätigkeit des Simultandolmetschens eher, da die Stimuli in beiden Fällen über den akustischen Kanal geliefert werden. Weitere Testvarianten integrieren eine Reihe verschiedener Doppelaufgaben und abgewandelter Versuchsbedingungen. Wird während der Lernphase verlangt, dass die VersuchsteilnehmerIn ständig sinnlose Silben vokalisiert,

spricht man von artikulatorischer Suppression. Ferner besteht die Möglichkeit, Verzögerungsintervalle zwischen Items oder das Sprechtempo bei akustisch dargebotenem Material zu verändern. Die Bandbreite der Anpassungsmöglichkeiten ist damit bei weitem nicht erschöpft. Ein wichtiger Aspekt ist auch die Zusammensetzung der Experimental- und Kontrollgruppen. Handelt es sich um einen Vergleich von professionellen DolmetscherInnen mit Nicht-DolmetscherInnen? Wie viel Berufserfahrung haben die professionellen DolmetscherInnen? Oft werden bei Experimenten auch DolmetschstudentInnen oder zweisprachige Nicht-DolmetscherInnen herangezogen.

2.3. Dolmetschwissenschaftliche Untersuchungen des AG

Ab dem Anfang der 1990er-Jahre ist eine zunehmende Publikationsdichte im Bereich der dolmetschwissenschaftlichen AG-Forschung zu beobachten. Griff man in dieser frühen Phase fast ausschließlich auf das von Baddeley vorgeschlagene AG-Modell zurück, um die empirischen Ergebnisse zu interpretieren, so rückt man in der neueren Forschung immer mehr von Baddeley ab. An dieser Stelle sollen beispielhaft ein paar nach experimenteller Vorgehensweise gegliederte Untersuchungen mit DolmetscherInnen präsentiert werden.

2.3.1. *Freie Reproduktion mit Artikulatorischer Suppression*

Darò und Fabbro (1994) führten eine Studie mit 24 fortgeschrittenen DolmetschstudentInnen durch, deren A-Sprache Italienisch war. Ausgangspunkt für die Untersuchung war die auf Baddeley basierende Hypothese, dass das Zuhören in einer Ausgangssprache und die gleichzeitig erfolgende Artikulation der Ausgangsbotschaft in der Zielsprache zu einer Beeinträchtigung des subvokalen Rehearsal-Prozesses, d.h. zu artikulatorischer Suppression, in der phonologischen Schleife des AG führen könnte (vgl. 1994: 369). Das Experiment umfasste zwei Aufgaben: die vollständige Wiedergabe von akustisch präsentierten Kurzgeschichten und einen akustischen Ziffernspannentest. Die Wiedergabe von Kurzgeschichten erfolgte unter zwei Bedingungen: Abruf nach Zuhören (Kontrollbedingung) und Abruf nach Simultandolmetschung des dargebotenen Texts. Den Ziffernspannentest konzipierte man mit vier Bedingungen: Abruf nach Zu-

hören (Kontrollbedingung), Abruf nach begleitendem Mitsprechen (*shadowing*), Abruf nach artikulatorischer Unterdrückung, Abruf nach Simultandolmetschung. Die TeilnehmerInnen wurden in zwei Gruppen unterteilt. Eine arbeitete von der A-Sprache (Italienisch) in die B-Sprache (Englisch), eine von der B-Sprache in die A-Sprache.

Die Ergebnisse bestätigten die Hypothese, die Darò und Fabbro aufgestellt hatten: Die Retentionsleistung war nach einer Simultandolmetschung des dargebotenen Materials bei beiden Aufgaben und bei beiden TeilnehmerInnengruppen signifikant niedriger als bei allen anderen Versuchsbedingungen. Die Ergebnisse des Ziffernspannentests zeigten, dass die Behaltensleistung bei artikulatorischer Suppression gegenüber der Zuhör- und *shadowing*-Bedingung deutlich abfällt, jedoch immer noch besser ausfällt als nach einer Simultandolmetschung. Darò und Fabbro (vgl. 1994: 375) führen den von ihnen beobachteten Leistungsabfall bei der Simultandolmetsch-Bedingung allerdings nur zum Teil auf die durch artikulatorische Suppression hervorgerufene phonologische Interferenz zurück. Weitere Faktoren sind ihrer Ansicht nach auch die geteilte Aufmerksamkeit und der strapazierende Übersetzungsprozess.

2.3.2. *Reading Span Task*

Padilla *et al.* (1995) konzipierten ein Experiment, an dem vier verschiedene ProbandInnengruppen teilnahmen: zehn professionelle DolmetscherInnen, von denen die Hälfte gerade die Dolmetscherabschlussprüfung absolviert hatte und die andere Hälfte ungefähr fünf Jahre Berufserfahrung hatte; zehn Nicht-DolmetscherInnen, die alle ein Studium in einer anderen Studienrichtung als Dolmetschen mit ausgezeichnetem Erfolg abgeschlossen hatten; zehn im zweiten Jahr studierende DolmetschstudentInnen und zehn im dritten Jahr studierende DolmetschstudentInnen. Mit dieser Untersuchung beabsichtigten Padilla *et al.* (vgl. 1995: 63) die Hypothese zu überprüfen, nach der professionelle DolmetscherInnen hinsichtlich AG-Kapazität den DolmetschstudentInnen und Nicht-DolmetscherInnen signifikant überlegen sind. Dies wäre ihrer Ansicht nach der empirische Nachweis dafür, dass das AG eine zentrale Rolle beim Simultandolmetschen spielt und ferner, dass die AG-Kapazität nicht unveränderlich ist, sondern eine Kapazitätssteigerung durch Ausbildung und Berufserfahrung erzielt werden kann. Es wurden zwei Experimente durchgeführt: Experiment 1 umfasste einen Ziffernspannentest und einen

Lesespannentest; Experiment 2 war ein freier Reproduktionstest, der einmal ohne artikulatorische Suppression (Kontrollbedingung) und einmal mit artikulatorischer Suppression abgewickelt wurde.

Professionelle DolmetscherInnen schnitten bei beiden Aufgaben im Rahmen von Experiment 1 signifikant besser ab als alle anderen Gruppen, woraus Padilla *et al.* (vgl. 1995: 70) schließen, dass professionelle DolmetscherInnen eine größere AG-Kapazität besitzen. Experiment 2 lieferte ein differenzierteres Ergebnis. Während in der Kontrollbedingung des freien Reproduktionstests keine signifikanten Gruppenunterschiede festgestellt werden konnten, ergab die Versuchsvariante mit artikulatorischer Suppression eine klare Überlegenheit der professionellen DolmetscherInnen. Padilla *et al.* zufolge können professionelle DolmetscherInnen den Ausfall einer AG-Komponente, in diesem Fall der artikulatorischen Schleife, besser verkraften, weil ihr AG effizienter arbeitet. Das bessere Abschneiden dieser Gruppe bei AG-Tests sei auf die regelmäßige Übung zuerst in der Ausbildung und später im Berufsleben zurückzuführen.

Padilla *et al.* (2005) differenzierten ihre Ansichten in einer späteren Studie. Die AG-Leistungen dreier Versuchsgruppen (professionelle DolmetscherInnen, Personen mit hoher Gedächtnisspanne und eine Kontrollgruppe) wurden in drei Experimenten verglichen. Die Resultate zeigten, dass professionelle DolmetscherInnen deshalb in der Lage sind, gleichzeitig zu verstehen und zu produzieren, weil sie schneller auf lexikalische und semantische Repräsentationen im LZG zugreifen können. Der in der ersten Untersuchung hervorgehobene Zusammenhang mit AG-Speicherkapazität und AG-Effizienz wird in dieser neuen Publikation abgeschwächt.

2.3.3. *Listening Span Task*

Besonders aufschlussreich im Hinblick auf die Rolle des AG beim Simultandolmetschen ist die Dissertation von Liu (2001). Die Grundzüge dieses Experiments werden überblicksmäßig von Liu *et al.* (2004) besprochen. Mit ihrer Studie leitete Liu eine methodische Neuorientierung in der Dolmetschwissenschaft ein. Ihr Ansatz wird in der Dolmetschwissenschaft als vielversprechend betrachtet (vgl. Timarová 2008: 22), da sie bei ihrem Versuch nicht ausschließlich Standardaufgaben aus der Kognitionspsychologie verwendet, sondern eine Vorgehensweise wählt, bei der neben einem klassischen

AG-Test auch eine nach festgelegten Kriterien bewertete Dolmetschaufgabe verwendet wird. Drei verschiedene Personengruppen nahmen an der Studie teil: elf professionelle DolmetscherInnen, die eine einjährige Ausbildung absolviert hatten und mindestens zwei Jahre Berufserfahrung vorweisen konnten; elf DolmetschstudentInnen, die das zweite Ausbildungsjahr fast hinter sich hatten; und elf DolmetschstudentInnen, die das erste Ausbildungsjahr fast abgeschlossen hatten. Die A-Sprache aller TeilnehmerInnen war Chinesisch, die B-Sprache Englisch.

Mit dieser Untersuchung möchte Liu feststellen, ob Personen mit ähnlich gut ausgeprägten kognitiven Fähigkeiten, aber unterschiedlichem Expertiseniveau hinsichtlich der Tätigkeit des Simultandolmetschens auch unterschiedlich gute Dolmetschleistungen erbringen. Mit dieser Zielsetzung wurden die Simultandolmetschleistungen professioneller DolmetscherInnen mit jenen von DolmetschstudentInnen verglichen.

Das Experiment bestand, wie erwähnt, aus zwei Aufgaben, nämlich einem Hörspannentest und einer Dolmetschaufgabe. Ausgangssprache war stets Englisch. Für die Dolmetschaufgabe wurden drei experimentelle Texte entworfen, die nach klar definierten Gesichtspunkten zusammengestellt waren. Es wurden sog. „kritische Sätze“ (*critical sentences*) eingebaut, deren Schwierigkeit man gezielt veränderte, um die Auswirkungen einer Veränderung dieses Parameters auf die Qualität der Dolmetschung des „Folgesatzes“ (*continuation sentence*) zu untersuchen. Darüber hinaus wurde der Ausgangstext in Bedeutungseinheiten segmentiert, deren Wichtigkeit von Muttersprachlern auf einer zweistufigen Skala eingeordnet wurde. Als zweite Aufgabe war ein Hörspannentest vorgesehen, der eine domänenunabhängige Messung der AG-Kapazität ermöglichen sollte.

Die Dolmetschaufgabe wurde von den professionellen DolmetscherInnen signifikant besser bewältigt als von den Dolmetschstudierenden. Beim Hörspannentest konnten hingegen keine signifikanten Gruppenunterschiede festgestellt werden. Die drei Gruppen hielten einander also hinsichtlich allgemeiner kognitiver Leistungsfähigkeit und AG-Kapazität die Waage. Die Überlegenheit der professionellen DolmetscherInnen bei der Dolmetschaufgabe steht daher nicht im Zusammenhang mit kapazitätsspezifischen Variablen, sondern ist Liu *et al.* (vgl. 2004: 36) zufolge der Tatsache zuzuschreiben, dass erfahrene DolmetscherInnen auf Strategien und Fertigkeiten zurückgreifen, die sie eigens für den Simultandolmetschprozess, vielleicht sogar für einzelne Spra-

chenpaare, entwickelt haben. Beim Hörspannentest würden sie dieses domänenspezifische Wissen nicht anwenden können und daher keine Vorteile gegenüber anderen Gruppen vorweisen. Für die DolmetscherInnenausbildung würde dies bedeuten, dass weniger die Steigerung allgemeiner AG-Kapazitäten, sondern der Erwerb dolmetschspezifischer Fertigkeiten im Mittelpunkt stehen sollte (vgl. 2004: 38). Die Untersuchung zeigte, dass professionelle DolmetscherInnen die vom AG auferlegten Kapazitätsbeschränkungen oft durch eine domänenspezifische Fertigkeit überwinden, die es ihnen erlaubt, weniger wichtige Informationseinheiten zu erkennen und im Notfall wegzulassen, ohne die primäre Botschaft zu beeinträchtigen.

Köpke und Nespoulous (2006) reihen sich mit ihrer sehr breit angelegten Studie in die Gruppe von ForscherInnen ein, deren Experimente sehr stark auf der von der experimentellen Kognitionspsychologie ausgearbeiteten Versuchspalette aufbauen. Die hohe TeilnehmerInnenzahl und die facettenreiche Experimentalanordnung verleihen dieser Untersuchung besondere Aussagekraft.

79 ProbandInnen nahmen an dieser Studie teil. Sie entfielen auf vier Gruppen: 21 professionelle DolmetscherInnen mit vier bis 35 Jahren Berufserfahrung (Durchschnittsalter: 44,4 Jahre), deren A-Sprache Französisch war und die Englisch als B- oder C-Sprache hatten; 18 im zweiten Jahr studierende DolmetschstudentInnen mit der gleichen Sprachenkombination wie die erste Gruppe (Durchschnittsalter: 26,2 Jahre); 20 französisch-englische zweisprachige AkademikerInnen, die sich mehrere Jahre in einem englischsprachigen Land aufgehalten hatten, aber keine DolmetscherInnen waren (Durchschnittsalter: 44,7 Jahre); 20 französischsprachige StudentInnen ohne ausgeprägte Fremdsprachenkenntnisse (Durchschnittsalter: 21,5 Jahre).

Die auszuführenden Aufgaben wurden in verschiedene Kategorien gegliedert: Aufgaben, die sich auf die Speicherung von Informationen beschränkten; solche, die sowohl Behaltensleistungen als auch Informationsverarbeitung erforderten; und einen sog. Stroop-Test, bei dem die Versuchspersonen aufgefordert werden, die Farbe dargebotener (Farb-)Wörter oder Items zu bestimmen, wobei bei bestimmten Varianten erhebliche AG- und Aufmerksamkeitskapazitäten nötig sind, um die Farbe von Schriftzügen richtig auszusprechen, wenn es sich um ein Farbwort handelt, das nicht mit der Farbe der Buchstaben übereinstimmt (z.B. beim Farbwort „grün“ in gelbem Schriftzug wäre die korrekte Antwort „gelb“).

Köpke und Nespoulous (vgl. 2006: 10f.) überprüften folgende Hypothesen: Erstens erwartete man sich empirische Anhaltspunkte für die Annahme, dass das AG ausgebaut und weiterentwickelt werden kann. Der Beweis dafür wäre ein besseres Abschneiden der professionellen DolmetscherInnen vor den DolmetschstudentInnen gefolgt von den zwei anderen Gruppen. Zweitens nahm man an, dass diese Unterschiede bei komplexen Aufgaben noch deutlicher zum Vorschein kommen würden, weil die zentrale Exekutive wesentlich mehr beteiligt ist. Und drittens ging man davon aus, dass die durch den Stroop-Test gemessene selektive Aufmerksamkeit bei professionellen Dolmetschern besser ausgeprägt sein würde.

Die erste Hypothese wurde durch die Ergebnisse bei den einfachen Gedächtnisaufgaben nicht gestützt: Es wurden keine signifikanten Unterschiede zwischen den Gruppen beobachtet. Hypothese zwei wurde dahingehend bestätigt, dass bei komplexen Aufgaben signifikante Leistungsunterschiede festgestellt wurden. Allerdings lagen entgegen den Erwartungen die DolmetschstudentInnen bei einer großen Anzahl von Aufgaben an der Spitze, gefolgt von den professionellen DolmetscherInnen und dann den Kontrollgruppen. Die dritte Hypothese wurde widerlegt.

Daraus ergibt sich die Feststellung, dass weder die Behaltensleistung noch die selektive Aufmerksamkeit bei DolmetscherInnen besser ausgeprägt ist. Bei komplexen Aufgaben entsteht aber ein anderes Bild: Hier schneiden DolmetscherInnen durchwegs deutlich besser ab, wenn es die Aufgabe erlaubt, die Reize semantisch tiefer zu verarbeiten. Die Ergebnisse scheinen also darauf hinzudeuten, dass DolmetscherInnen bei Ausfall der phonologischen Komponente des AG auf gut ausgeprägte semantische Speicherungs- und Verarbeitungsmechanismen zurückgreifen können. Köpke und Nespoulous (vgl. 2006: 18) betonen die Rolle der zentralen Exekutive (Baddeley & Logie 1999) bzw. des Fokus der Aufmerksamkeit (Cowan 1999) bei semantischen Verarbeitungsstrategien.

Für das unerwartet bessere Abschneiden der DolmetschstudentInnen bei manchen komplexen Aufgaben gegenüber professionellen DolmetscherInnen liefern die Versuchleiter die Erklärung, dass bei AnfängerInnen dieser Berufsgruppe noch ganz andere Prozesse ablaufen als bei erfahrenen DolmetscherInnen. Da AnfängerInnen mangels Erfahrung noch sehr oft kognitiv an ihre Grenzen stoßen, verwenden sie das KZG als Notfallmechanismus. Aus demselben Grund ist auch die semantische Verarbei-

tungsfunktion der zentralen Exekutive bei AnfängerInnen besser ausgebildet. Professionelle DolmetscherInnen sind offensichtlich nicht mehr in gleichem Maße auf ihr AG angewiesen, da sie bereits andere spezialisierte Strategien entwickelt haben, die unter anderem durch das LZG unterstützt werden (vgl. Köpke & Nespoulous 2006: 18).

2.4. Kategorisierung der Erklärungsansätze

In der Forschungsliteratur ist eine große Vielfalt an Vorschlägen zur Erklärung der AG-spezifischen Unterschiede zwischen SimultandolmetscherInnen und Nicht-SimultandolmetscherInnen anzutreffen. In einigen Arbeiten werden die Struktur und Kapazität des AG in den Vordergrund gestellt. Andere wissenschaftliche Beiträge betonen die Funktionsweise dieses Gedächtnissystems. In jedem Fall ist festzustellen, dass sich die Antworten oft ergänzen oder überlappen. Um einen besseren Überblick zu gewähren, erscheint es an dieser Stelle sinnvoll, eine Zusammenstellung der in der Literatur dargelegten Erklärungsansätze zu versuchen:

- AG-Kapazität: Ein erheblicher Teil der Forschung geht von einer größeren allgemeinen AG-Kapazität bei DolmetscherInnen aus (Darò 1995; Darò & Fabbro 1994; Padilla *et al.* 1995). Diese Ansicht wird allerdings von einigen Studien widerlegt oder relativiert (Köpke & Nespoulous 2006; Liu 2001; Liu 2008; Liu *et al.* 2004; Padilla *et al.* 2005).
- Expertenleistung: Erfahrene SimultandolmetscherInnen besitzen im Gegensatz zu EinsteigerInnen spezifische, hochspezialisierte Fertigkeiten, die das AG entlasten. Dieser Erklärungsansatz wird in der Literatur auch als *expert-novice paradigm* bezeichnet. Zu diesen Expertenstrategien gehören beispielsweise die effiziente Aufmerksamkeitsverschiebung oder die Automatisierung bestimmter Abläufe (vgl. Cowan 2000: 129; Ericsson 2000; Moser-Mercer *et al.* 2000). Auch die Fähigkeit, die Wichtigkeit von Informationseinheiten schnell einzuschätzen und Ressourcen effizient zu verteilen, wird als charakteristisch für eine Expertenleistung betrachtet (Liu 2001; Liu *et al.* 2004). Die Bedeutung des prozeduralen Gedächtnisses wird in diesem Zusammenhang ebenfalls hervorgehoben (vgl. Darò 1997: 626). Padilla *et al.* (2005) sprechen sich für besser ausgeprägte und daher

schneller abrufbare lexikalische und semantische Repräsentationen bei DolmetscherInnen aus.

- Steuerung des AG: Bei SimultandolmetscherInnen laufen die Prozesse im AG generell effizienter ab und werden besser koordiniert (Bajo *et al.* 2001; Chincotta & Underwood 1998; Darò 1995; Liu *et al.* 2004). Padilla *et al.* (vgl. 2005: 213) können diese Annahme mit ihrem Experiment aber nicht bestätigen.
- Verarbeitungs-/Aktivierungsgeschwindigkeit: Informationseinheiten werden im AG schneller verarbeitet und transferiert (Cowan 2000; Padilla *et al.* 2005).
- AG-Entleerung: Die Inhalte werden von SimultandolmetscherInnen schneller wieder aus dem AG gelöscht als bei Nicht-DolmetscherInnen (Padilla *et al.* 1995).
- Inhibition: Die Fähigkeit, irrelevante Stimuli zu unterdrücken, unterscheidet SimultandolmetscherInnen von anderen Gruppen. Mechanismen im AG erlauben es, die relevantere und wichtigere Information akustisch und semantisch besser herauszufiltern, was einer besseren Resistenz gegen Interferenz gleichkommt (Bajo *et al.* 2001; Cowan 2000; Darò 1995; Darò 1997; Liu 2008; Liu *et al.* 2004; Moser-Mercer *et al.* 2000).
- Robuste Gedächtniseinträge: Die Gedächtniseinträge sind bei SimultandolmetscherInnen robuster und daher für das AG besser zugänglich (vgl. Chincotta & Underwood 1998: 15).
- Semantisch-kontextuelle Enkodierung: Da beim Dolmetschen die phonologische Komponente ausgelastet ist, entwickeln DolmetscherInnen die Fähigkeit, Informationen verstärkt semantisch-kontextuell zu enkodieren statt phonologisch, wie üblicherweise bei anderen Personengruppen. Semantische Codes werden auch länger im Gedächtnis aufrechterhalten (Köpke & Nespoulous 2006; Liu 2008; MacWhinney 1997; Mizuno 2005; Shlesinger 2000a).

2.5. Kritische Bewertung der Untersuchungsmethoden

Eine große Schwäche fast aller bisher durchgeführten empirischen Studien im Bereich der dolmetschwissenschaftlichen AG-Forschung betrifft die Anzahl der Versuchspersonen. Liu *et al.* (vgl. 2004: 39) geben beispielsweise selbst zu bedenken, dass ihre Er-

kenntnisse auf Grund der kleinen TeilnehmerInnenzahl nur sehr beschränkt generalisierbar sind.

Außerdem führen oft unterschiedliche Einteilungskriterien hinsichtlich der TeilnehmerInnengruppen zu widersprüchlichen Ergebnissen und wirken sich nachteilig auf die Vergleichbarkeit von Studien untereinander aus. So ist die Gruppe der professionellen DolmetscherInnen beispielsweise bei Padilla *et al.* (1995) viel jünger (Durchschnittsalter: 30 Jahre) und hat viel weniger Berufserfahrung (null bis fünf Jahre) als die gleichnamige Gruppe bei Köpke und Nespoulous (2006), wo sie ein mittleres Alter von 44,4 Jahren und zwischen vier und 35 Jahren Berufserfahrung hat.

Mizuno (vgl. 2005: 741) gibt ferner zu bedenken, dass Retentionsaufgaben eigentlich wenig aussagekräftig für das Simultandolmetschen sind, da es DolmetscherInnen nach einem Einsatz nur in den seltensten Fällen nötig haben, verarbeitete Informationseinheiten wieder abzurufen. Chincotta und Underwood (vgl. 1998: 6) äußern Vorbehalte darüber, ob die von Padilla *et al.* (1995) verwendete freie Reproduktionsaufgabe tatsächlich über das KZG Aufschluss gibt. Sie vermuten, dass hier das LZG eine große Rolle spielt. Gleiches könnte man auch bei zahlreichen anderen AG-Spannentests behaupten, wo die Verzögerung bis zum Abruf der Items nicht den Gegebenheiten bei einer Dolmetschleistung entspricht.

Ein anderer methodischer Einwand betrifft die Darbietungsmodalität des Testmaterials. Der Lesespannentest ist im Zusammenhang mit dem Simultandolmetschen vielleicht weniger aussagekräftig, da das Material visuell präsentiert wird. Aufschlussreich wäre demzufolge eher der Hörspannentest, da hier die Stimuli akustisch geliefert werden. Ungeeignet für die Untersuchung der Rolle des AG beim Simultandolmetschen ist entsprechend dieser Ansicht auch der Stroop-Test, da hier ebenfalls visuelle Aspekte im Vordergrund stehen (Köpke & Nespoulous 2006).

Köpke und Nespoulous (2006: 18) und Hacker *et al.* (1994: 314f.) weisen schließlich darauf hin, dass bei der Lösung von komplexen Gedächtnisspannenaufgaben oft mehrere Strategien zur Auswahl stehen. Welche Strategie von den Versuchspersonen gewählt wird, könne man fast nie mit Sicherheit sagen (dazu mehr in Kap. 4.5).

2.6. Methodische Anregungen

In der dolmetschwissenschaftlichen Forschung ist in den letzten Jahren vermehrt die Forderung nach mehr dolmetschspezifischen Experimentalanordnungen bei der Untersuchung der Rolle des AG beim Simultandolmetschen laut geworden.

Timarová (vgl. 2008: 21f.) hält es für wichtig, bei empirischen Studien den eigentlichen Dolmetschaufgaben mehr Gewicht zu geben. Liu (2001) schlägt m. E. in dieser Hinsicht den richtigen Weg ein, da sie die Dolmetschleistung in den Mittelpunkt ihres Experiments rückt, aber gleichzeitig die domänenunabhängige AG-Leistung mit Hilfe eines Standardtests aus der Kognitionspsychologie prüft. Als durchaus aufschlussreich haben sich in diesem Zusammenhang auch Verfahren erwiesen, bei denen Dolmetschleistungen nicht nur wie üblich von DolmetscherInnen erbracht werden müssen, sondern auch von Nicht-DolmetscherInnen (Christoffels *et al.* 2003; Tzou *et al.* 2011). Shlesinger (vgl. 2003: 45) schlägt sogar vor, zukünftig korpusbasierte Studien durchzuführen, die auf Korpora bereits gedolmetschter Texte zurückgreifen und dazu dienen könnten, Ergebnisse empirischer Arbeiten durch Daten aus der Praxis zu bestätigen oder zu widerlegen.

Von Vorteil für die Aussagekraft der Versuche wäre natürlich auch eine größere Anzahl an TeilnehmerInnen. Die einzigen Forscher, die dieser Forderung bisher gerecht werden konnten, waren Köpke und Nespoulous (2006), die mit 79 ProbandInnen arbeiteten.

Ein letzter Forschungsansatz, der bisher in den Studien vollkommen vernachlässigt wurde, aber interessante Erkenntnisse verspricht, besteht in einer longitudinalen Herangehensweise, bei der man das AG einer DolmetscherIn vom Beginn der Ausbildung an bis ins Berufsleben hinein in regelmäßigen Abständen überprüft, um festzustellen, ob es sich mit zunehmender Erfahrung auf irgendeine Art und Weise verändert (vgl. Cowan 2000: 141; Padilla *et al.* 2005: 218).

Insgesamt geht aus diesem Kapitel deutlich hervor, dass in Zukunft eine noch engere Zusammenarbeit zwischen Kognitionspsychologie und Dolmetschwissenschaft wünschenswert wäre. Denn es wird eine Reihe interessanter Fragen bei kognitionspsychologischen Experimenten mit DolmetscherInnen aufgeworfen. Um schlüssige Antworten zu finden, müsste allerdings die Kooperation der zwei Disziplinen noch weiter

ausgebaut werden. Die große Palette an Erklärungsansätzen und Forschungsperspektiven zeigt, dass interdisziplinäres Synergiepotenzial durchaus vorhanden ist.

3. Der Reading Span Task

3.1. Entwicklung des interdisziplinären Austauschs

Rückblickend wäre es wohl im Sinne dolmetschwissenschaftlichen Erkenntnisgewinns gewesen, wenn sich VertreterInnen der Kognitionspsychologie bereits ab der Entstehungsphase ihrer Disziplin am Anfang der zweiten Hälfte des 20. Jahrhunderts mit Translationsprozessen auseinandergesetzt hätten. Tatsächlich wurde dieser vielschichtigen Sonderform menschlicher Kommunikation von kognitionswissenschaftlicher Seite bei der Entwicklung von theoretischen Modellen und Versuchsparadigmen wenig Aufmerksamkeit geschenkt.

Wenig überraschend erscheint daher die Feststellung, dass der durch gängige AG-Modelle gebotene theoretische Rahmen oft keine restlos zufriedenstellende Einbettung für die Untersuchung dolmetschrelevanter Prozesse bietet. In diesem Zusammenhang bereitet vor allem die Parallelität von Verstehen, Translation und Produktion sprachlicher Information, wie sie beim Simultandolmetschen anzutreffen ist, Erklärungsnot – beruhen doch die meisten theoretischen Überlegungen hinsichtlich des AG auf einsprachiger Kommunikation.

Mögliche Gründe für dieses mangelnde Interesse sieht De Groot (1997: 30) einerseits in der fehlenden Behandlung dieses Themas in einflussreichen Fachzeitschriften und Handbüchern der kognitiven Psychologie. Andererseits sei die große Komplexität der kognitiven Abläufe beim Simultandolmetschen ein Faktor, der kognitionspsychologische Forschungsbestrebungen in diese Richtung erheblich bremse. Es herrsche die Meinung vor, dass es zunächst erforderlich sei, ein gründliches Verständnis einfacher Prozesse wie Sprachverstehen, Sprachproduktion, Gedächtnis, Aufmerksamkeit, visuelle bzw. akustische Wahrnehmung und Entscheidungsprozesse zu erlangen, bevor man sich mit dem Dolmetschen befassen könne, bei dem all diese kognitiven Mechanismen in einem komplizierten Zusammenspiel gemeinsam zur Verwendung kommen und sich möglicherweise gegenseitig beeinflussen (De Groot 1997).

Natürlich lässt sich genauso argumentieren, dass das Simultandolmetschen gerade wegen seiner Komplexität für viele Teilbereiche der Psychologie und Linguistik einen äußerst fruchtbaren Forschungsgegenstand darstellen müsste. Denn nicht selten

treten bei dieser Tätigkeit auf Grund kognitiver Überlastung Fehlleistungen auf, die ausgezeichnet über die ablaufenden mentalen Prozesse Aufschluss geben und möglicherweise wesentliche Beiträge zur Aufstellung allgemeiner Theorien der Kognition und Sprachverarbeitung liefern könnten (MacWhinney 1997).

Erste Annäherungen zwischen Kognitionspsychologie und Dolmetschwissenschaft sind den Bemühungen einflussreicher ForscherInnen wie David Gerver und Barbara Moser zu verdanken. Gerver (1976) präsentierte ein Prozessmodell des Simultandolmetschens, das ein *short-term working memory* (1976: 191) umfasst. Es kann als Verdienst dieser WissenschaftlerInnen betrachtet werden, dass kognitionspsychologische Konzepte und Forschungsmethoden in der Dolmetschwissenschaft immer mehr in den Mittelpunkt rückten.

Allerdings blieb diese interdisziplinäre Befruchtung lange Zeit einseitig. Die in dieser Arbeit beschriebenen AG-Modelle von Baddeley (Baddeley & Hitch 1974) und Cowan (1988) entstanden ohne jegliche Berücksichtigung der ungewöhnlichen kognitiven Konstellation beim Dolmetschen. Erst als Nachgedanke wurden mögliche Berührungspunkte zwischen dem jeweiligen AG-Konzept und der Dolmetschwissenschaft thematisiert (Baddeley 2000b; Cowan 2000; Ericsson 2000). Der bilaterale Brückenschlag zwischen den zwei Disziplinen scheint erst Ende der 1990er-Jahre des letzten Jahrhunderts gelungen zu sein. Vereinzelt beschäftigen sich nun auch KognitionspsychologInnen mit dem Dolmetschen, wie die Veröffentlichung des Sammelbandes „Cognitive Processes in Translation and Interpretation“ (Danks *et al.* 1997) zeigt.

Eine Folge dieser schwachen interdisziplinären Vernetzung ist darin zu sehen, dass Fortschritte in der Kognitionspsychologie zumeist erst stark zeitversetzt von der Dolmetschwissenschaft rezipiert werden. Das multimodulare AG-Modell von Baddeley wurde beispielsweise 1974 in der Forschungsliteratur präsentiert. Es entwickelte sich rasch zum maßgeblichen Ansatz in der AG-Forschung und bildete den Ausgangspunkt vieler wissenschaftlicher Arbeiten, die strukturelle Verfeinerungen vorschlugen oder neue Versuchsparadigmen einführten. In der Dolmetschwissenschaft fanden Baddeleys theoretische Vorschläge aber erst wirklich Anfang der 1990er ihren Niederschlag, als Darò und Fabbro (1994) ein allgemeines Gedächtnismodell für den Simultandolmetschprozess beschrieben, das in wesentlichen Zügen auf Baddeley aufbaut.

Dieses Phänomen spiegelt sich in der Anwendung experimenteller Forschungsmethoden wider. Denn um die Tragfähigkeit verschiedener AG-Modelle, darunter an prominenter Stelle jenes, das von Baddeley postuliert wurde, zu prüfen, wurde eine ganze Reihe gedächtnisrelevanter Versuchsparadigmen entwickelt. Dabei ist im Rahmen dieser Arbeit der von Daneman und Carpenter im Jahre 1980 vorgestellte Lesespannentest (*reading span task*) von besonderer Relevanz. Es vergingen 15 Jahre, bis diese Experimentalanordnung von der Dolmetschwissenschaft aufgegriffen wurde (Padilla *et al.* 1995). Dieser zeitliche Abstand weist auf die Kluft hin, die das Verhältnis der zwei Disziplinen trotz bemerkenswerter Annäherungsversuche von prägenden Gestalten der Dolmetschwissenschaft wie Barbara Moser in dieser Phase charakterisierte. Erst vor etwas mehr als einem Jahrzehnt begann das Dolmetschen verstärkt ins Bewusstsein von VertreterInnen der kognitiven Psychologie zu rücken, sodass nun von einem einigermaßen wechselseitigen Austausch gesprochen werden kann.

Die beschriebene Situation entpuppte sich zwar einerseits als Nachteil, da bei der Entwicklung kognitionspsychologischer Tests dolmetschrelevante Aspekte nicht berücksichtigt und folglich auch keine standardisierten Versuchsparadigmen zur Erforschung der kognitiven Mechanismen beim Simultandolmetschen hervorgebracht wurden. Andererseits darf nicht vergessen werden, dass dolmetschwissenschaftlichen ForscherInnen trotz der dargelegten geschichtlichen Entwicklung ein vielfältiger Fundus an bewährten experimentellen Forschungsmethoden zur Verfügung steht, der es zwar nicht erlaubt, den Dolmetschprozess in seiner Gesamtheit kognitionspsychologisch zu erfassen, sich allerdings dazu eignet, die kognitiven Subprozesse und Teilaspekte, wie das AG, gründlich zu durchleuchten.

Eine weit verbreitete experimentelle Methode zur Untersuchung des AG ist der bereits angesprochene Reading Span Task (= RST) von Daneman und Carpenter (1980), auf dem auch die vorliegende empirische Studie fußt. Von seiner Reliabilität und Validität zeugt zum einen die lange Liste von Publikationen (s. Conway *et al.* 2005), in denen Experimente beschrieben werden, welche die ursprüngliche Versuchsanordnung einfach beibehalten oder mit gezielten methodischen Korrekturen übernehmen. Für die empirische Güte des Lesespannentests spricht darüber hinaus die Tatsache, dass er weit über die Grenzen der Kognitionspsychologie hinaus bekannt ist und beispielsweise in

der Entwicklungspsychologie und der klinischen Psychologie als verlässliches Messinstrument gilt (Conway *et al.* 2005; Van den Noort *et al.* 2008).

In einem allgemeinen Abschnitt gilt es nun zunächst, grundlegende Konzepte des RST zu erörtern und methodische Überlegungen und Weiterentwicklungen, die bei der Durchführung des hier beschriebenen Experiments maßgeblich waren, zu besprechen. Das folgende Kapitel wendet sich dann dem eigentlichen Versuch zu.

3.2. Der RST in der Kognitionspsychologie

3.2.1. *Ursprünge des RST*

Der Lesespannentest entsprang dem Bestreben, die allgemeine kognitive Kapazität des Menschen messbar zu machen. Die Forschung war sich weitgehend einig, dass allen kognitiv komplexen Aktivitäten, wie beispielsweise dem Verstehen, dem logischen Denken oder dem Problemlösen, eine gemeinsame kognitive Ressource zu Grunde liegen muss, die begrenzt ist und interindividuell variiert. Das unterschiedliche Abschneiden bei Intelligenztests und Lesetests sowie im weiteren Sinne auch der schulische Erfolg wurde mit diesem domänenunabhängigen Konstrukt in Verbindung gebracht (Gathercole 2008: 43). Mit der Einführung des AG-Begriffs gelang es, diese kognitive Grundkapazität genauer zu definieren, werden diese Konzepte doch als weitgehend gleichbedeutend aufgefasst.

Einfache Gedächtnisspannentests, wie der Ziffernspannen- oder Wortspannentest, erwiesen sich im Zusammenhang mit der AG-Kapazität als wenig zielführend. Sie sind in einer überholten Forschungstradition verankert, die der Idee eines passiven Kurzzeitspeichers zentrale Bedeutung einräumt. Dementsprechend schwach korrelieren diese einfachen Gedächtnisaufgaben auch mit Maßen der Intelligenz und des Leseverstehens (Daneman & Carpenter 1980: 451; Köpke & Nespoulous 2006: 18). Daneman und Carpenter erkannten die Notwendigkeit einer aussagekräftigeren Operationalisierung der AG-Kapazität, die ihnen mit dem RST auch gelang.

Wichtig bei der Entstehung der neuen Experimentgattung war die Erkenntnis, dass das AG ein aktiver Bestandteil des menschlichen Informationsverarbeitungssystems ist. Folglich muss für eine präzise Messung seiner Kapazität gewährleistet sein,

dass nicht nur seine Speicherfunktion, sondern vor allem auch seine Verarbeitungsfunktion eingebunden ist. Allen komplexen AG-Spannungsaufgaben, zu denen an erster Stelle der klassische Lesespannungstest und in weiterer Folge eine breite Palette an Varianten zählt, ist somit die Eigenschaft gemeinsam, dass die Versuchsperson aufgefordert wird, gleichzeitig zur Retentionsaufgabe eine strapazierende Verarbeitungstätigkeit zu bewältigen. Man kann von einem Doppelaufgabenparadigma (*dual task paradigm*) sprechen, bei dem das koordinierte Bewältigen von Behalten und Verarbeiten von Gedächtnisinhalten im Mittelpunkt steht (Hacker *et al.* 1994: 298).

Nimmt man an, dass diesen Aufgaben eine gemeinsame Gesamtressource, nämlich das AG, zur Verfügung steht, so ist die logische Folge, dass diese Versuchsanordnung Konkurrenzerscheinungen (*trade-off*) zwischen Behaltens- und Verarbeitungsleistungen hervorrufen muss. Konzentriert sich die ProbandIn zu sehr auf das Speichern von Elementen, so leidet darunter die Leistung bei der vom Experiment vorgegebenen Verarbeitungsaufgabe. Gleiches gilt für die entgegengesetzte Konstellation.

In der Wissenschaft werden zwei Perspektiven geboten, um unterschiedliches Abschneiden beim Lesespannungstest und verwandten AG-Versuchen zu erklären. Quantitativ betrachtet könnten einfach interindividuelle Differenzen im Umfang der rohen Gesamtressource postuliert werden. Just und Carpenter (1992) prägten in diesem Zusammenhang den Begriff Aktivierung (*activation*). Darunter verstehen sie die gesamte AG-Kapazität, auf die eine Person zurückgreifen kann, um gleichzeitige Behaltens- und Verarbeitungsanforderungen zu erfüllen.

Als Alternative wäre die Möglichkeit einer qualitativen Interpretation von Kapazitätsunterschieden zu erwähnen. Denn besseres Abschneiden beim RST könnte auch mit effizienteren mentalen Prozessen in Verbindung gebracht werden, was nicht unbedingt eine größere Kapazität im quantitativen Sinne voraussetzen würde. Lesespannungsvorteile wären demnach bei Personen zu beobachten, welche die vorgegebene Aufgabe mit weniger Verarbeitungsschritten, beispielsweise beim lexikalischen Zugriff oder bei Kodierungsprozessen, lösen. Daraus würde ein doppelter Nutzen erwachsen. Zum einen würde auf diese Weise mehr Platz für Speichervorgänge geschaffen. Zum anderen würden die Verarbeitungsprozesse auch grundsätzlich wegen der kleineren Anzahl an Zwischenschritten schneller ablaufen, wodurch sich der Zerfall der Gedächtnisinhalte langsamer vollziehen würde. Eine größere AG-Kapazität könnte auf diese Weise rein funk-

tional erklärt werden (Daneman & Carpenter 1980: 451). Selbstverständlich kann dieser Index kognitiver Leistungsfähigkeit auch als Ausdruck eines Zusammenwirkens sowohl qualitativer als auch quantitativer Kapazitätsfaktoren verstanden werden.

Ein AG-Test ist erst dann sinnvoll kalibriert, wenn die AG-Ressourcen auf das Äußerste erschöpft werden. Nur so können interindividuelle Unterschiede in der Kapazität verlässlich festgestellt werden. Daneman und Carpenter (1980) erreichten dies, indem sie das Grundprinzip des Wortspannentests, bei dem eine von Serie zu Serie anwachsende Anzahl an unzusammenhängenden Wörtern zu behalten und anschließend wiederzugeben ist, beibehielten, aber um einen Verarbeitungsaspekt erweiterten.

Der von ihnen entwickelte RST operierte mit sinnmäßig unzusammenhängenden Sätzen, die in Blöcke variabler Größe zusammengefasst waren. Es waren 2-Satz-, 3-Satz-, 4-Satz-, 5-Satz- und 6-Satz-Ebenen vorgesehen. Jede Ebene wurde dreimal durchlaufen. Die Satzstimuli wurden einzeln präsentiert, wobei nach jedem Block von der TeilnehmerIn erwartet wurde, dass sie die Endwörter der soeben dargebotenen Sätze in der richtigen Reihenfolge reproduzierte. Die 2-Satz-Ebene setzte sich folglich aus drei 2-Satz-Blöcken und den drei dazugehörigen Reproduktionsphasen zusammen. Dann wurde auf drei 3-Satz-Blöcke erhöht und die Anzahl wuchs schrittweise weiter, bis das Abrufen der Endwörter nicht mehr gelang, maximal aber bis sechs. Dieser Teil des Experiments gleicht weitgehend dem Wortspannentest.

Die Verarbeitungsaufgabe bestand in der ersten Version des Verfahrens darin, jeden Satz laut vorzulesen. Daneman und Carpenter (1980) nahmen an, dass mit dieser Methode eine ausreichende Auslastung des AG gewährleistet sei. Sie gingen davon aus, dass der durch das laute Vorlesen erzwungene Verstehensprozess gekoppelt mit einer raschen Präsentation der Satzstimuli den erwünschten Effekt einer AG-Kapazitätserschöpfung habe. Als Wert für die AG-Spanne galt die maximale Satzebene, auf der bei zwei von drei Blöcken noch alle Endwörter richtig abgerufen wurden.

Die Ergebnisse ihrer ersten Testreihe können als Bestätigung der aufgestellten Hypothesen interpretiert werden. Der Lesespannentest korrelierte anders als der herkömmliche Wortspannentest gut mit drei verschiedenen Maßen des Leseverstehens. Der Schluss, dass die starke Korrelation von RST-Studien und verschiedenen Tests des Leseverstehens und der Intelligenz durch eine gemeinsame zugrunde liegende Komponente, das AG, verursacht wird, erschien daher durchaus plausibel. Die AG-Kapazität wäre

damit als gemeinsamer interindividueller Differenzierungsfaktor zu bezeichnen. Allerdings konnten Daneman und Carpenter (1980) nicht gänzlich ausschließen, dass diese Unterschiede durch von Testperson zu Testperson variierende Lesegeschwindigkeiten hervorgerufen wurden. Bei schnellerem Lesen ist nämlich das Argument legitim, dass die Gedächtnisinhalte dem zeitlich bedingten Zerfall weniger anheimfallen. Diese Überlegung sowie die Absicht, modalitätsspezifische Faktoren zu untersuchen, veranlassten Daneman und Carpenter (1980), eine zweite Testreihe durchzuführen.

Sie erweiterten ihre Experimentpalette dabei um eine akustische Variante des RST, den Hörspannentest (*listening span task*), und bemühten sich gleichzeitig, die in der ersten Testreihe verwendeten methodischen Parameter zu verbessern. Der Listening Span Task (= LST) schloss nun einen Lesegeschwindigkeitseffekt aus, da die akustische Darbietung des Reizes gleiche zeitliche Wahrnehmungsbedingungen für alle ProbandInnen sicherstellte. Das Aufgabenkonzept blieb beim Hörspannentest weitgehend unverändert. Lediglich die Anzahl der Blöcke pro Satzebene wurde in der akustischen Version des Tests auf fünf erhöht.

Eine wesentliche Neuerung erfuhr die Verarbeitungskomponente des Verfahrens, um für alle Ausprägungen des Versuchs gleichermaßen anwendbar zu sein. Lautes Vorlesen war nun beim Listening Span Task keine praktikable Verarbeitungsaufgabe mehr. Die Versuchsleiterinnen fügten daher direkt nach der Darbietung jeden Reizes eine Satzverifikationskomponente ein, die eine Bewertung der gehörten oder gelesenen Aussage nach „richtig“ oder „falsch“ innerhalb von 1,5 Sekunden erforderte. Mittels dieser Vorgehensweise versuchte man zu gewährleisten, dass die VersuchsteilnehmerInnen den Satz vollständig verarbeiteten und die Bedeutung zu erfassen versuchten. Andernfalls wäre es durchaus vorstellbar, dass sich Testpersonen ausschließlich auf das Endwort jedes Satzes konzentrieren (Daneman & Carpenter 1980: 457). Wie die Evaluierungen der Sätze ausfielen, wurde bei der Analyse der Ergebnisse nicht berücksichtigt. Den Versuchspersonen wurde aber nichtsdestoweniger zu verstehen gegeben, dass dies ein wichtiger Bestandteil des Verfahrens sei.

Weitere methodische Anpassungen betrafen das Auswertungsverfahren und die Punktevergabe. Daneman und Carpenter (1980) rückten von der streng seriellen Wiedergabeanforderung ab und ließen alle korrekt reproduzierten Wörter ungeachtet der Reihenfolge in das Ergebnis einfließen. Bei der Punktevergabe wurde eine Präzisierung

dergestalt vorgenommen, dass nun auch halbe Punkte für teilweise erfolgreich abgeschlossene Satzebenen angerechnet wurden. Versuchspersonen, die beim RST einen von drei Blöcken und beim LST zwei von fünf Blöcken richtig beantworteten, erhielten einen halben Punkt.

Die Analyse der aus Experimentreihe 2 resultierenden Daten untermauerte die in der ersten Phase gezogenen Schlussfolgerungen. Wieder wurde eine enge Beziehung zwischen der Lesespanne und verschiedenen Maßen des Lese- und Hörverstehens beobachtet. Der nun hinzugekommene Listening Span Task unterstützt dieses Ergebnismuster zusätzlich. Er korreliert stark mit dem Reading Span Task und ebenfalls mit den von Daneman und Carpenter (1980) eingesetzten Verfahren zur Ermittlung des Lese- und Hörverstehens. In Summe lassen diese Korrelationswerte also durchaus den Schluss zu, dass eine einheitliche kognitive Gesamtressource, die AG-Kapazität, für interindividuelle Differenzen in der kognitiven Leistungsfähigkeit verantwortlich gemacht werden kann. Weiterhin schwer zu klären bleibt die Frage, ob ein besseres Abschneiden durch eine größere AG-Kapazität im Sinne eines größeren Gesamtfassungsvermögens oder aber im Sinne einer gesteigerten Verarbeitungseffizienz und der daraus erwachsenden funktionalen Kapazitätserweiterung bedingt ist (Daneman & Carpenter 1980: 464f.).

3.2.2. *Methodische Entwicklung*

Der Lesespannentest und der davon abgeleitete Hörspannentest sind in einer dynamischen 30-jährigen Forschungstradition verwurzelt. Naturgemäß machen neue Erkenntnisse und der stetig wachsende Wissensstand in der Kognitionspsychologie methodische Schärfungen und Anpassungen dieses klassischen Versuchsparadigmas unvermeidlich. Die rege Beschäftigung mit diesem experimentellen Verfahren gerade in den letzten Jahren zeigt, dass es auch weiterhin ein bewährtes Instrument zur Erkundung AG-relevanter Phänomene ist.

Abgesehen von den offensichtlichen Fortschritten, die der technologische Wandel der 1980er und 1990er mit dem Einzug computergesteuerter Experimente in den psychologischen Disziplinen brachte (Daneman und Carpenter verwendeten noch analoge Ausrüstung), ist auch zu beobachten, dass es in der Wissenschaft vermehrt Bemühungen gibt, mehr Einheit in die große Vielfalt methodischer Vorgehensweisen zu brin-

gen. Damit verfolgt man das Ziel, wissenschaftliche Untersuchungen, welche die gleiche Versuchsgattung verwenden, besser vergleichbar zu machen. In dieser Hinsicht wegweisend ist die Publikation von Conway *et al.* (2005). Eine Gruppe namhafter Psychologen um Andrew R. A. Conway verfasste gemeinsam einen Aufsatz, der eine Mischung aus methodologischer Metaanalyse der bisherigen Forschung zu komplexen Gedächtnisspannungsaufgaben und praktischem Leitfaden für die Abwicklung derartiger Experimente ist. Eine methodische Standardisierung steht auch bei Van den Noort (2008) im Vordergrund. Allerdings zielen seine Normierungsbestrebungen eher auf die Entwicklung einheitlicher RST für verschiedene Sprachen (Englisch, Deutsch, Niederländisch und Norwegisch) ab. Zur erfolgreichen Verbreitung dieser experimentellen Methode über Sprachgrenzen hinweg trugen bereits früher erschienene Arbeiten bei, beispielsweise Hacker *et al.* (1994) oder Desmette *et al.* (1995), die eine deutsche bzw. französische Variante des RST beschreiben.

Zwar stünden noch weitere Messinstrumente zur Ermittlung der domänenunabhängigen AG-Kapazität zur Auswahl, die am Grundprinzip einer aus Behaltens- und Bearbeitungskomponente bestehenden Doppelanforderung festhalten und sich ebenfalls als robuste Operationalisierungen kognitiver Leistungsfähigkeit herausgestellt haben (Conway *et al.* 2005). Es gäbe beispielsweise den *operation span task*, bei dem einfache Rechenaufgaben gelöst und die resultierenden Ergebnisse behalten werden, oder den *counting span task*, bei dem visuelle Elemente in verschiedenen Zusammenstellungen gezählt werden und die so ermittelte Summe der Objekte memoriert wird. Doch obwohl diese experimentellen Varianten zweifellos auch aufschlussreiche Ergebnisse liefern würden, kamen in der vorliegenden Untersuchung aus zwei Gründen die klassischeren Tests, nämlich der RST und der LST, zum Einsatz.

Einerseits ist die Vermutung naheliegend, dass bei diesen Methoden auf Grund ihrer Textbezogenheit eine größere externe Validität in Bezug auf das Dolmetschen, einer inhärent textverarbeitenden Tätigkeit, erzielt wird. Die beim Lesespannen- und Hörspannentest eingesetzten AG-Mechanismen entsprechen mit größerer Wahrscheinlichkeit dolmetschspezifischen Subprozessen, als es bei anderen Verfahren mit visuellen oder arithmetischen Komponenten der Fall wäre. Andererseits existieren für alternative AG-Spannenparadigmen einfach noch keine dolmetschwissenschaftlichen Anhaltspunkte, sodass eine derartige Untersuchung zwar als Pionierarbeit Anerkennung verdienen

würde, aber im Zusammenhang mit den in diesem Forschungsbereich etablierten Paradigmen, dem RST und dem LST, von noch ungewisser Aussagekraft ist.

3.2.3. *Bearbeitungsaufgabe (Processing Task)*

Während hinsichtlich der Retentionsaufgabe, bei der von Anfang an die Speicherung von Satzendwörtern vorgesehen war, kaum methodische Veränderungen zu berichten sind, vollzog sich seit der Vorstellung des Experiments als Reaktion auf berechtigte Kritik eine Auffächerung der Bearbeitungsfacetten des RST/LST in mehrere Varianten.

Wie oben beschrieben, begnügten sich Daneman und Carpenter (1980) in ihrer ersten Experimentreihe damit, die Versuchspersonen Stimulussätze laut vorlesen zu lassen, um die verarbeitende Funktion des AG zu beschäftigen. Mit der Einführung des akustischen LST im zweiten Versuchslauf wurde auch eine neue Bearbeitungsaufgabe notwendig. Die ProbandInnen wurden daher aufgefordert, nach der Darbietung jedes Satzitems eine allgemeine inhaltliche Beurteilung nach „richtig oder falsch“ abzugeben.

Neuere Versionen des Verfahrens beinhalten ebenfalls Verifikationskomponenten. Allerdings steht nun oft eine Einschätzung nach grammatikalischen, semantischen oder syntaktischen Gesichtspunkten auf der Satzebene im Vordergrund. Diese Satzevaluierungskomponente kann im Gegensatz zur Vorgangsweise von Daneman und Carpenter (s. oben) auch bei der Auswertung der Ergebnisse eingebunden werden. Dies trägt zur Güte des Tests bei, denn auffallend hohe Fehlerquoten bei der Satzbewertung könnten darauf hinweisen, dass die Aufmerksamkeit fast ausschließlich auf die Behaltensaufgabe gerichtet war und daher eher von einem Verhalten wie bei einem einfachen KZG-Experiment zu sprechen wäre. Fälle dieser Art können als ergebnisverzerrend betrachtet und folglich ausgeschlossen werden (Conway *et al.* 2005: 772).

Hacker *et al.* (1994: 298, 301) befürworten eine größere Gewichtung der inhaltlichen Sinnerfassung jeden Satzes ohne Verifikationsanforderung. Sie verlangten daher von den ProbandInnen, zusätzlich zu den Endwörtern auch die „sinngemäße[n] Kernaussage“ (Hacker *et al.* 1994: 301) der Sätze wiederzugeben. Auf diese Weise wird überprüft, ob sowohl Behaltens- als auch Bearbeitungsmechanismen des AG im Spiel waren, ob also das Doppelanforderungsparadigma (*dual task paradigm*) tatsächlich erfüllt wurde. Hacker *et al.* (1994: 298) stehen der Satzbeurteilung nach Sinnhaftig-

keitskriterien kritisch gegenüber. Sie geben zu bedenken, dass diese Art der Bearbeitungsaufgabe eine zusätzliche Erschwerung bei der Materialerzeugung bedeute, da ein Teil der Stimuli *per definitionem* keinen Sinn ergeben dürfe und Erinnerungsleistungen daher durch ungewöhnlich oder absurd anmutende Sätze gesteigert werden könnten. Als weiteres Argument für die von ihnen vorgeschlagene Sinnerfassungskomponente führen Hacker *et al.* (1994: 298) die geringere Wissensabhängigkeit ins Treffen. Denn die Verifikationsaufgabe setzt ein gewisses Wortwissen oder Weltwissen voraus, während bei der Sinnerfassungsaufgabe nur die Wiedergabe der allgemeinen Idee erforderlich ist. Eine ähnliche Verstehensanforderung ist auch bei Service *et al.* (2002) zu finden, allerdings erwarten sie von ihren Versuchspersonen, dass sie die Kernaussage von akustisch präsentierten Sätzen erfassen und mittels Tastendruck beurteilen, ob diese Aussagen auf ein gleichzeitig dargebotenes Bild zutreffen.

Van den Noort *et al.* (2008: 38) wählen in ihrer mehrsprachigen Variante des RST eine Mischform der bisher genannten Methoden. Wie im ursprünglichen RST von Daneman und Carpenter (1980) müssen die TeilnehmerInnen die Sätze laut vorlesen. Zusätzlich werden vor Beginn des Experiments inhaltliche Fragen angekündigt. Nach Abschluss des Tests sind dann zehn Fragen zu bestimmten Satzstimuli zu beantworten.

3.2.4. *Scoringmethoden*

Das in der Forschung herrschende Durcheinander bei der Berechnung von RST-Scores hat in jüngster Zeit vermehrt Zweifel an der Vergleichbarkeit von RST-Studien hervorgerufen. Die verschiedenen Punktevergabeverfahren sind Gegenstand einer eingehenden Analyse durch Conway *et al.* (2005: 774-776). An dieser Stelle erscheint eine kurze Charakterisierung der bisher in kognitionswissenschaftlichen Arbeiten angewandten Scoringmethoden sinnvoll. Eine Beschreibung der in der vorliegenden empirischen Studie gewählten Vorgangsweise, die maßgeblich von den Empfehlungen in Conway *et al.* (2005) beeinflusst wurde, ist im Kapitel „Scoring“ (s. Kap. 4.3.5) zu finden.

Fast allen Scoringmethoden ist gemeinsam, dass sie nur die Behaltensleistung einer Versuchsperson widerspiegeln. Meistens werden die Ergebnisse der Bearbeitungsaufgabe, soweit diese überhaupt registriert werden, bei der Ermittlung des Scores nicht

berücksichtigt, sondern dienen lediglich als Überprüfungsmechanismus, der es erlaubt, auffällig fehlerhafte Bearbeitungsleistungen auszusortieren (Conway *et al.* 2005: 774).

Wie oben besprochen (s. Kap. 3.2.1) galt beim ursprünglichen RST von Daneman und Carpenter (1980) jene Satzebene als Lesespanne (*reading span*) einer Versuchsperson, auf der die Endwörter gerade noch fehlerfrei und in richtiger Reihenfolge reproduziert werden konnten. Gelang diese Leistung nicht mehr, hatte man diesem Ansatz zufolge die Kapazitätsobergrenze erreicht und das Experiment war beendet. Die Lesespanne entsprach dann der letzten Satzebene, bei der entweder alle Blöcke, oder zumindest die Mehrheit der Blöcke (zwei von drei beim RST bzw. drei von fünf beim LST) einwandfrei bewältigt wurden. Bei teilweise richtig beantworteten Blöcken gab es halbe Punkte. Zur nächstgrößeren Satzebene gelangte man allerdings nur, wenn die vorhergehende ohne Fehler absolviert worden war.

Conway *et al.* (2005: 774) plädieren dafür, sich von dieser traditionellen Methode zu distanzieren. Sie bemängeln an der Berechnung einer absoluten AG-Spanne, wie sie bei den Schöpfern des RST zu finden ist, dass dieser Wert bei ein und derselben Person abhängig von der Schwierigkeit der Satzstimuli hinsichtlich verschiedener Parameter wie Länge oder semantischer Ähnlichkeit erheblich schwankt. Der als „Lesespanne“ bezeichnete kognitive Kapazitätsindex büßt somit beim Vergleich mehrerer RST-Studien mit unterschiedlichem Satzmaterial an Aussagekraft ein. Außerdem kritisieren Conway *et al.* (2005: 774) am herkömmlichen Scoringverfahren den Mangel an Empfindlichkeit, da die Skala meist nur zwei bis sechs Punkte umfasse und außerdem auf eine Abstufung nach Halbpunkten beschränkt sei.

Ein erster Schritt, um diesem offensichtlichen Eichungsdilemma entgegenzuwirken, ist eine in neueren RST-Untersuchungen weit verbreitete Vorgangsweise, bei der ein vollständiges Durchlaufen aller Satzebenen ohne vorzeitiges Abbrechen bei Fehlleistungen vorsieht. Damit schafft man ein differenzierteres und gründlicheres Messinstrument, da die kognitive Leistung auf allen Ebenen – auch auf solchen, wo das AG überfordert zu sein scheint – berücksichtigt wird und sich unerwartete Leistungsbeeinträchtigungen auf Grund bestimmter Satzkonstellationen weniger gravierend auf das Gesamtergebnis auswirken. Bei diesem Design besteht die einfachste Scoringmethode darin, die Gesamtzahl der erinnerten Endwörter bei allen Blöcken als Score zusammenzurechnen (s. z.B. Christoffels *et al.* 2006; Signorelli *et al.* 2011; Tzou *et al.* 2011).

Damit eng verbunden ist die Festlegung der Reihenfolge, in welcher die Endwörter reproduziert werden müssen. Fließen nur die in der ursprünglichen Reihenfolge wiedergegebenen Wörter in das Ergebnis ein oder ist jede beliebige Reihenfolge zulässig, solange die Reize erfolgreich behalten wurden? Die Verwendung der freien Reproduktion ist jedenfalls für das Untersuchen serieller Positionseffekte (Primäreffekt und Rezenzeffekt, s. Gathercole 2008: 37) wichtig, wie Van den Noort (2008: 38) feststellt. Köpke und Signorelli (2011: 11) merken darüber hinaus an, dass in diesem Modus das von Craik und Lockhart (1972) postulierte *elaborative rehearsal* begünstigt werde. Da die Informationen nicht einfach mechanisch wiedergegeben werden müssen (*maintenance rehearsal*), kann eine tiefere Verarbeitung der Elemente und damit eine Herstellung assoziativer Verbindungen stattfinden. Diese Überlegung gewinnt im Zusammenhang mit dem Simultandolmetschen zusätzlich an Bedeutung, da hier möglicherweise eben derartige Verhaltensmuster charakteristisch sind. Köpke und Signorelli (2011: 13) halten es für wahrscheinlich, dass BerufsdolmetscherInnen auf Grund ihrer Erfahrung besser ausgebildete elaborative Rehearsalfähigkeiten besitzen. Folglich sei auch ein besseres Abschneiden dieser Gruppe bei AG-Tests mit freier Wiedergabe zu erwarten.

Conway *et al.* (2005) tragen mit ihrer Besprechung wesentlich zur Verfeinerung der Scoringverfahren bei. Ein wichtiges Merkmal ihrer Arbeit ist die Hervorhebung standardisierter Auswertungsmethoden, die mit einer besseren Vergleichbarkeit von RST-Ergebnissen verbunden sind. Während der RST-Wert bei einer absoluten Berechnung stets von der Anzahl der eingesetzten Stimuli abhängig ist, erhält man bei einer relativen Kalkulation einen Betrag zwischen null und eins, der wahlweise auch in Prozenten angegeben werden kann und damit auch über den Rahmen des betreffenden Experiments hinaus aussagekräftig bleibt.

Das relative Scoringschema gliedert sich weiter in vier Unterarten auf, die ebenfalls von Conway *et al.* (2005: 774-776) ausführlich behandelt werden. Die Unterteilung erfolgt anhand von zwei Fragen: Erstens, werden bei der Auswertung nur solche Blöcke gezählt, bei denen alle Endwörter vollständig richtig wiedergegeben wurden, unvollständige Blöcke aber disqualifiziert (*all-or-nothing*)? Oder werden auch solche Satzgruppen berücksichtigt, bei denen nur ein Anteil behalten wurde (*partial-credit*)? Zweitens ist zu entscheiden, wie die Satzebenen gewichtet werden. Da höhere Satzebenen schwieriger zu bewältigen sind, wäre es denkbar, ihnen im Gesamtscore mehr Gewicht

zuzuteilen (*load scoring*). Alternativ dazu können alle Satzgruppen als ebenbürtig behandelt werden (*unit scoring*). Es stehen Conway *et al.* (2005: 775) zufolge also vier Scoringvarianten zur Auswahl: *partial-credit unit scoring* (PCU), *all-or-nothing unit scoring* (ANU), *partial-credit load scoring* (PCL), *all-or-nothing load scoring* (ANL).

Die vorliegende Arbeit setzt in der Scoring-Frage die Empfehlungen von Conway *et al.* (2005: 776) um. Dort wird auf der Grundlage empirischer Daten der Schluss gezogen, dass das Partial-Credit Unit Scoring (PCU) wohl als die zuverlässigste Lösung betrachtet werden darf. Eine detailliertere Beschreibung dieses Verfahrens erfolgt im Rahmen des Methodenteils (s. Kap. 4.3.5).

3.2.5. Materialerzeugung

Die Wahl des Satzmaterials war bei Daneman und Carpenter (1980) im Vergleich zu derzeitigen Standards in der AG-Forschung noch relativ bedenkenlos. Die einzigen Einschränkungen waren die Satzlänge, die zwischen neun und 16 Wörter schwanken durfte, und der Schwierigkeitsgrad, der an der subjektiven Einschätzung der Versuchsleiterinnen gemessen als „mäßig“ beschrieben wird (Daneman & Carpenter 1980: 458). Als Quelle verwendeten sie Quizbücher zu allgemeinen Themen.

Empirische Untersuchungen der letzten Jahre (s. z.B. Van den Noort *et al.* 2008) zeigen, dass die methodischen Regeln bei der Materialerzeugung viel präziser gestaltet werden können. Verschiedene Prüfverfahren und eine ganze Reihe linguistischer Werkzeuge erlauben es, bei der Herstellung der Sätze ein straffes Korsett an Kriterien anzulegen.

Eine Darstellung der Eigenschaften des in dieser Arbeit verwendeten und von Maurits van den Noort zur Verfügung gestellten Materials erfolgt an späterer Stelle im Rahmen eines eigenen Kapitels (s. Kap. 4.3.2 und Van den Noort *et al.* 2008). Einige der wichtigsten Aspekte können allerdings bereits hier angeschnitten werden. Zu den Maßnahmen zählt beispielsweise eine bessere Kontrolle der Satzlänge durch eine Festlegung nicht nur der zulässigen Anzahl an Wörtern, sondern auch an Silben und Buchstaben. Weitere homogenisierte Parameter umfassen die Frequenz und den Konkretheitsgrad der Endwörter. Zur Ermittlung der Frequenz wurde die CELEX-Datenbank herangezogen. Dabei handelt es sich um ein korpusbasiertes Sprachfor-

schungsinstrument, das den Wörtern verschiedener Sprachen relative Häufigkeitswerte zuordnet (Van den Noort *et al.* 2008: 36).

Zwei verschiedene empirische Eichungsverfahren mit jeweils 120 TeilnehmerInnen wurden durchgeführt. Eines zielte darauf ab, die Plausibilität der Sätze zu kontrollieren. Fiel die durchschnittliche Plausibilitätsbewertung unter eine bestimmte Schwelle, wurde der betreffende Stimulus durch einen anderen ersetzt. In einem zweiten Verfahren mussten die Versuchspersonen die Konkretheit der Endwörter auf einer sechsstufigen Skala beurteilen, was eine zusätzliche Vereinheitlichung des Materials erlaubte (Van den Noort *et al.* 2008: 38).

3.2.6. *Darbietungsreihenfolge der Satzebenen*

Eine sinnvolle methodische Neuerung, die bereits von Van den Noort *et al.* (2008) angeregt und in der vorliegenden Studie übernommen wurde, betrifft die Abfolge der verschiedenen Satzebenen. Das ursprüngliche Konzept beruht auf einer schrittweisen Erhöhung der Satzebene. Beispielsweise folgen nach der erfolgreichen Bewältigung von drei 2-Satz-Blöcken drei 3-Satz-Blöcke etc. Diese Gesetzmäßigkeit hinsichtlich der Darbietungsreihenfolge wird den TeilnehmerInnen in den Instruktionen meistens auch mitgeteilt, sodass nicht auszuschließen ist, dass dieser Faktor einen Einfluss auf die Behaltensleistung hat. Van den Noort *et al.* (2008: 38) beschließen angesichts dieser Tatsache, die Reihenfolge der Satz-Blöcke zu randomisieren. Dies verhindert, dass Versuchspersonen das Präsentationsschema antizipieren und Strategien entwickeln können. Satzebenen aus Blöcken einheitlicher Größe werden also durch Serien ersetzt, die durch eine randomisierte Abfolge von Blöcken unterschiedlichen Umfangs charakterisiert sind.

3.2.7. *Vereinbarkeit mit gängigen AG-Modellen*

Ein klarer Vorteil des hier gewählten experimentellen Paradigmas ist in der Vereinbarkeit mit allen gängigen AG-Modellen zu sehen. Da eine globale kognitive Kapazität gemessen wird und nicht die Leistungsfähigkeit einzelner postulierter Subkomponenten des AG im Mittelpunkt steht, sind die Ergebnisse des RST bzw. LST sowohl im Rah-

men integrativ-funktionaler AG-Ansätze, wie von Cowan befürwortet (s. Kap. 1.3.2), als auch stärker strukturell-modular orientierter Konzeptualisierungen, wie von Baddeley vertreten (s. Kap. 1.3.1), aufschlussreich.

Obwohl der Lesespannentest in der Forschungsliteratur oft mit dem AG-Modell von Baddeley in Verbindung gebracht wird (Conway *et al.* 2005: 771; Gathercole 2008: 43), diene der ursprünglichen Variante des Experiments als theoretische Grundlage einzig die Annahme, dass es eine kognitive Gesamtressource gebe, die durch in Konkurrenz stehende Behaltens- und Bearbeitungsaufgaben bis an ihre Grenzen mobilisiert wird. Aus der Perspektive von Baddeley fällt das Bearbeiten von Inhalten in den Aufgabenbereich der zentralen Exekutive. Das Speichern von verbalen Reizen hingegen bewältigt nach dieser Auffassung die phonologische Schleife. Analog dazu wäre die Retention von visuellen oder räumlichen Reizen, wie sie bei anderen komplexen AG-Experimenten erforderlich ist, dem Wirken des räumlich-visuellen Notizblocks zuzuschreiben. Das Doppelaufgabenparadigma ist diesem Ansatz zufolge als gleichzeitige Beanspruchung der zentralen Exekutive und eines weiteren domänenspezifischen Subsystems zu verstehen (Gathercole 2008: 43).

In Anlehnung an Cowan *et al.* (2005) ist es möglich, den Lesespannentest als Maß des Fassungsvermögens des Fokus der Aufmerksamkeit (*focus of attention* = FoA) zu betrachten. Zwar merken Cowan *et al.* (2005) an, dass sie dieses Verfahren aus zahlreichen Gründen für wenig zielführend halten und anderen Messinstrumenten den Vorzug geben. Allerdings gestehen sie ein, dass der RST trotz einiger Bedenken auch eine Interpretation im Sinne ihres AG-Modells zulässt. Vor dem Hintergrund des *embedded-processes model* ist die Lesespanne der Kapazität des FoA gleichzusetzen, die als *scope of attention* bezeichnet wird. Diese Kapazität wird als maximale Anzahl unverbundener Informationseinheiten (*chunks*) verstanden, die gleichzeitig im FoA gehalten werden können. Wesentlich ist dieser Auffassung des RST zufolge die Rolle der Bearbeitungsaufgabe. Diese verhindere nämlich, dass Rehearsal- und Chunkingprozesse zum Einsatz kommen. Durch derartige Strategien entstehen Verknüpfungen zwischen Gedächtnisinhalten, die zu gesteigerten Behaltensleistungen führen. Wenn allerdings mit dem Einsatz von gedächtnisstützenden Mechanismen zu rechnen ist, kann unmöglich die eigentliche Grundkapazität des FoA im Sinne von Cowan eruiert werden, weil man es nicht mehr mit einzelnen isolierten Einheiten zu tun hat. Durch die Bearbeitungskomponente des

RST wird diese Strategieanwendung unterbunden und ermöglicht somit eine Auslegung des Versuchs als Maß des *scope of attention* (Cowan *et al.* 2005).

3.3. Der RST in der Dolmetschwissenschaft

3.3.1. Problemlage

Mit der Erschließung kognitionspsychologischer Forschungswege in der Dolmetschwissenschaft wurde auch der RST als relevantes Verfahren zur Untersuchung dolmetschspezifischer Fragen entdeckt. ForscherInnen dieser Strömung beschäftigen sich intensiv mit der Frage der kognitiven Subprozesse, in die das Simultandolmetschen zerlegt werden könnte (Liu 2008: 168-173).

Darunter sind einerseits die einzelnen kognitiven Teilfertigkeiten zu verstehen, die eine DolmetscherIn beherrschen muss, um ihren Beruf erfolgreich auszuüben. Ein prominenter Vertreter dieses Forschungsansatzes ist Daniel Gile, der den Simultandolmetschprozess im Rahmen seines *effort model* in vier Teilkomponenten (*efforts*) gliedert (Gile 1997: 198): *listening and analysis effort*, *production effort*, *memory effort*, *coordination effort*. Auch Moser-Mercer *et al.* (2000: 107) beschreiben eine Unterteilung der Tätigkeit in zahlreiche Sprachverarbeitungsfacetten, sodass sich insgesamt das Bild einer hochkomplexen kognitiven Leistung ergibt.

Andererseits ist mit dieser rein dolmetschwissenschaftlichen Sicht noch keine Verbindung zu gängigen Konzepten und Modellen der Kognitionspsychologie hergestellt, die es erlauben würde, auf den für jene Disziplin bezeichnenden Reichtum bewährter Versuchsparadigmen zurückzugreifen. Um eine aufschlussreiche Verknüpfung zu ermöglichen, musste also jenes kognitionspsychologische Konstrukt ausfindig gemacht werden, das für die Bewältigung komplexer Kognitionsanforderungen verantwortlich gemacht wird. Diese Pionierleistung gelang Darò und Fabbro (1994), die in ihren Arbeiten erstmals Baddeleys AG-Modell als Grundlage für ihre Experimente heranzogen. Der Verwendung klassischer Experimentgattungen aus der Kognitionswissenschaft stand nun nichts mehr im Wege.

Es gilt in der Kognitionspsychologie als allgemein anerkannt, dass das AG als treibender Mechanismus bei der Lösung diverser komplexer Aufgaben, wie verbalen

Verstehensleistungen, dem Problemlösen und dem logischen Denken, aufgefasst werden kann. Ebenso weit verbreitet ist die Ansicht, dass dieses AG eine begrenzte Kapazität hat. Darin ist auch der Ausgangspunkt zahlreicher dolmetschwissenschaftlicher Untersuchungen zu sehen. Denn das Simultandolmetschen erfordert, wie oben bereits erwähnt, eine ungewöhnlich große Anzahl an parallel ablaufenden kognitiven Sprachverarbeitungsprozessen, sodass der von den gängigen AG-Konzeptualisierungen gebotene Kapazitätsrahmen eigentlich gesprengt werden müsste. Es ist in diesem Zusammenhang also die Frage von Interesse, wie simultane Sprachmittlerleistungen auf der Grundlage gängiger theoretischer AG-Modelle überhaupt zu erklären sind. Was ist das entscheidende Charakteristikum, das das AG einer SimultandolmetscherIn auszeichnet?

Zur Untersuchung des AG hat sich in der Dolmetschwissenschaft weitgehend die bereits oben besprochene experimentelle Methode von Daneman und Carpenter (1980), der RST bzw. der LST, durchgesetzt. Man erhoffte sich von Anfang an eindeutige Hinweise auf AG-Leistungszuwächse mit dem Fortschreiten der Dolmetscherausbildung und zunehmender Berufserfahrung, die man mit einer statischen oder funktionalen Kapazitätsvergrößerung durch die regelmäßige Ausübung einer kognitiv strapazierenden Tätigkeit begründen könnte. Alternativ dazu hielt man es für vorstellbar, dass VertreterInnen dieser Berufsgruppe einfach von Natur aus hinsichtlich kognitiver Kapazitäten großzügiger veranlagt seien. Auch in diesem Fall wäre dann der Lesespannentest ein geeignetes Messinstrument. Allgemein ist festzuhalten, dass der RST als bewährtes Werkzeug angesehen wird, um interindividuelle Unterschiede hinsichtlich der Belastbarkeit des AG zum Vorschein zu bringen.

So sehr man sich aber eine klare Überlegenheit professioneller DolmetscherInnen beim RST/LST erwarten würde, so unterschiedlich und teilweise widersprüchlich fielen bisher die Ergebnisse aus (Köpke & Signorelli 2011: 9; Liu 2008: 170).

3.3.2. *Bisherige Ergebnisse*

Dolmetschwissenschaftliche Untersuchungen, die sich die von Daneman und Carpenter (1980) eingeführte Experimentgattung zu Nutze machen, können in Anlehnung an Köpke und Signorelli (2011: 8) sinnvollerweise nach Präsentationsmodus unterteilt werden. Häufiger kam bisher die visuelle Variante, der RST, zum Einsatz: Padilla *et al.* (1995),

Bajo *et al.* (2000), Christoffels *et al.* (2006), Signorelli *et al.* (2011), Tzou *et al.* (2011). Für die akustische Ausprägung des Tests, den LST, entschieden sich Nordet und Voegtlin (1998, zit. n. Köpke & Signorelli 2011), Liu (2001; s. auch Liu *et al.* 2004) und Köpke und Nespoulous (2006).

Wegbereitend war unbestritten die Studie von Padilla *et al.* (1995). Erstmals wurde hier in einem Vergleich zehn professioneller DolmetscherInnen mit drei anderen Gruppen, zu denen zehn im zweiten Studienjahr befindliche DolmetschstudentInnen (noch ohne Erfahrungswerte im Simultandolmetschen), zehn im dritten Studienjahr befindliche DolmetschstudentInnen (mit Grundkenntnissen im Simultandolmetschen) und eine Kontrollgruppe aus zehn Nicht-DolmetscherInnen gehörten, der RST verwendet, um AG-Kapazitätsunterschiede abhängig von Dolmetscherfahrung aufzudecken. Die Ergebnisse stimmten mit den Erwartungen überein: Die Differenzen waren signifikant ($F[3,38] = 4.05, p < .02$), wobei die Gruppe der BerufsdolmetscherInnen tatsächlich besser abschnitt als alle anderen. Nach Ansicht von Padilla *et al.* (1995) wird dadurch die Schlussfolgerung nahegelegt, dass das AG eine zentrale Rolle beim Simultandolmetschen spielt und sich durch ständige Beanspruchung auch entsprechend weiterentwickelt.

Im Rahmen einer empirischen Studie (Bajo *et al.* 2000), die von einer personell fast wie bei Padilla *et al.* (1995) zusammengesetzten ForscherInnengruppe betreut wurde und eine ganze Reihe kognitionspsychologischer Tests umfasste, um erfolgreiches Dolmetschen mit kognitiven Determinanten in Verbindung zu bringen, war unter anderem auch ein Lesespannentest vorgesehen. Die TeilnehmerInnen entfielen auf drei Gruppen: zehn BerufsdolmetscherInnen, zehn DolmetschstudentInnen ohne Simultandolmetscherfahrung (zweites Studienjahr) und zehn Nicht-DolmetscherInnen. Die Datenanalyse ergab ein Bild, dass schon aus Padilla *et al.* (1995) bekannt ist: Die Unterschiede zwischen den Versuchsgruppen waren signifikant ($F[2,27] = 5.32, p < .01$). BerufsdolmetscherInnen erwiesen sich als überlegen.

Dass die ersten Hinweise einer AG-Überlegenheit von BerufsdolmetscherInnen nicht unangefochten bleiben konnten, wurde durch den Forschungsbeitrag von Nordet und Voegtlin (1998, zit. n. Köpke & Signorelli 2011: 9 und Köpke & Nespoulous 2006: 3) klar ersichtlich. Als experimentelle Grundlage für ihre an der *École de traduction et d'interprétation* in Genf verfasste Diplomarbeit diente ein LST. Der Einsatz des akusti-

schen Zwillings-tests des RST stellte ein Novum in der Dolmetschwissenschaft dar. Die VersuchleiterInnen erhofften sich dadurch eine erhöhte externe Validität, da beim Simultandolmetschen eben akustische und nicht visuelle Reize verarbeitet werden. Es absolvierten sechs BerufsdolmetscherInnen, sieben DolmetschstudentInnen und 22 PsychologiestudentInnen den Versuch. Im Unterschied zu den bisher berichteten Forschungsarbeiten waren die Ergebnisse in diesem Fall viel weniger eindeutig. Es wurden keine signifikanten Gruppenunterschiede festgestellt. Allerdings konnte eine Tendenz dahingehend abgelesen werden, dass DolmetschstudentInnen in der Regel besser abschnitten als professionelle DolmetscherInnen und die zwei mit dem Dolmetschen vertrauten Gruppen die Kontrollgruppe üblicherweise übertrafen.

In das Lager jener DolmetschwissenschaftlerInnen, welche die Hypothese eines domänenunabhängigen AG-Vorteils für DolmetscherInnen verwerfen und stattdessen eine Interpretation des Simultandolmetschens als domänenspezifische Expertenleistung bevorzugen, gehört Liu (2001; s. auch Liu *et al.* 2004). Auch sie führte einen LST durch. An ihrer Studie nahmen elf BerufsdolmetscherInnen und 22 Dolmetschstudierende mit unterschiedlich weit fortgeschrittenem Studium (elf im ersten Studienjahr, elf im zweiten Studienjahr) teil. Beim Vergleich der drei Gruppen traten keine signifikanten Unterschiede hinsichtlich der Hörspanne auf. Gleichzeitig stellte die Versuchsleiterin fest, dass die Gruppe der BerufsdolmetscherInnen die ebenfalls zum Experiment gehörige Simultandolmetschaufgabe signifikant besser absolvierte als die fortgeschrittenen StudentInnen und diese wiederum signifikant besser abschnitten als die StudienanfängerInnen. Damit erweist sich nach Ansicht von Liu *et al.* (2004) die Behauptung, dass bessere Simultandolmetschleistungen durch ein effizienteres oder größeres AG bedingt sind, als unhaltbar. Vielmehr weisen die im Rahmen ihrer Studie gelieferten und nach genau festgelegten Parametern ausgewerteten Dolmetschleistungen darauf hin, dass sich BerufsdolmetscherInnen durch besser entwickelte domänenspezifische Expertenfertigkeiten auszeichnen. Dazu zählt die zielorientierte Aufmerksamkeitsallokation oder die effiziente Auswahl der zu dolmetschenden Ausgangstextsegmente.

Ein weiterer LST wird im Rahmen einer breit angelegten Studie (Köpke & Nespoulous 2006) mit der Zielsetzung einer gründlichen Erforschung der Entwicklung kognitiver Ressourcen mit wachsender Berufserfahrung beschrieben. Eine besondere Eigenschaft dieser Forschungsunternehmung im Vergleich zu anderen dolmetschwis-

senschaftlichen Beiträgen ist in der außergewöhnlich großen Anzahl an TeilnehmerInnen zu sehen, die den Ergebnissen selbstverständlich auch eine gesteigerte statistische Aussagekraft verschafft. Für den Test konnten 21 BerufsdolmetscherInnen, 18 im zweiten Ausbildungsjahr befindliche Dolmetschstudierende, 20 mehrsprachige Personen, die als erste Kontrollgruppe dienten, und 20 Studierende aus dolmetschfernen Studienrichtungen, die als zweite Kontrollgruppe dienten, gewonnen werden. Auch durch diese Studie ließ sich die Hypothese eines AG-Kapazitätsvorteils erfahrener DolmetscherInnen nicht bestätigen. Zwar lieferte der Hörspannentest signifikante Gruppenunterschiede ($F = 5.511$, $p = .0018$). Allerdings war dafür das signifikant bessere Abschneiden der DolmetschstudentInnen gegenüber den Kontrollgruppen ursächlich ($p < .05$). Die VertreterInnen der höchsten Erfahrungsstufe, die BerufsdolmetscherInnen, schafften nur die zweitbeste AG-Leistung und unterschieden sich damit von keiner Gruppe signifikant. Auch Köpke und Nespoulous (2006) äußern die Vermutung, dass dieses Ergebnismuster auf die Rolle von Expertenwissen hinweisen könnte. Demnach wären unerfahrene DolmetscherInnen noch stark von den kognitiven Ressourcen des AG abhängig, um die für sie neue Tätigkeit zu bewältigen. Da in dieser Einstiegsphase so intensiv auf das AG zurückgegriffen werde, sei auch eine Kapazitätssteigerung zu erwarten, die sich in den Ergebnissen ihres LST äußere (Köpke & Nespoulous 2006: 15-16). Der Erwerb von Expertise durch langjährige Erfahrung führe zu einer grundlegenden Veränderung der Prozesse. Dank der Entwicklung hochspezialisierter Expertenstrategien seien BerufsdolmetscherInnen weniger auf die kognitiven Ressourcen des AG angewiesen, was auch durch die Daten ihres Experiments nahegelegt werde (Köpke & Nespoulous 2006: 18).

In jüngster Zeit stand wieder vermehrt der RST im Mittelpunkt dolmetschwissenschaftlicher Forschungsanstrengungen. Christoffels *et al.* (2006) befassten sich mit der Rolle kognitiver Grundfertigkeiten beim Simultandolmetschen und setzten dazu eine Reihe verschiedener Wortabruf- und Gedächtnisaufgaben, darunter den Lesespannentest, ein. In einer ersten Phase wurden 39 BachelorstudentInnen mit Niederländisch als L1 und Englisch als L2 und 13 BerufsdolmetscherInnen mit Niederländisch als A-Sprache und Englisch als B-Sprache getestet. Anschließend fügte man eine weitere Versuchsgruppe hinzu, die 15 Englischlehrende im sekundären Bildungsbereich umfasste und die gleichen Experimente zu absolvieren hatte. Der RST wurde sowohl auf Nieder-

ländisch als auch auf Englisch abgewickelt und ergab ähnlich wie in früheren Studien (Bajo *et al.* 2000; Padilla *et al.* 1995) eine signifikante AG-Überlegenheit der BerufsdolmetscherInnen gegenüber den Studierenden ($F [1,50] = 4.76, p = .03$), ebenso wie gegenüber den EnglischlehrerInnen ($F [1,26] = 10.92, p < .01$). In Summe werden die Ergebnisse als Indiz dafür betrachtet, dass eine enge Beziehung zwischen Erfahrung im Dolmetschen und AG-Kapazität besteht. Christoffels *et al.* (2006) sind dank ihres Versuchsdesigns in der Lage, kognitive Leistungsfähigkeit und Fremdsprachenkompetenz als mögliche Determinanten erfolgreichen Simultandolmetschens zu dissoziieren. Denn die Gruppe der Lehrenden, die den professionellen DolmetscherInnen betreffend Englischniveau glich, schnitt bei Sprachverarbeitungsaufgaben mit lexikalischen Abrufkomponenten gleich gut ab. Beim RST lag sie signifikant hinter den BerufsdolmetscherInnen, was für besonders gut ausgebildete kognitive Fähigkeiten der letzteren Gruppe spricht. In der Untersuchung von Christoffels *et al.* (2006: 326) bleibt allerdings die Frage ungeklärt, ob ein gutes AG eine Voraussetzung für die Ausübung des DolmetscherInnenberufs ist oder erst allmählich in Abhängigkeit von Erfahrung erworben wird.

Der von Christoffels *et al.* (2006) entworfene RST in englischer Sprache wurde als Teil einer weiteren Experimentreihe bei Signorelli *et al.* (2011) repliziert. Zusätzlich zur Frage, ob sich BerufsdolmetscherInnen im Allgemeinen durch AG-Kapazitätsvorteile auszeichnen, wurde erstmals auch dem Faktor „Alterung“ Rechnung getragen. Dementsprechend wurden die 38 ProbandInnen in vier Gruppen unterteilt: zwölf junge BerufsdolmetscherInnen, acht junge Nicht-DolmetscherInnen, sieben ältere BerufsdolmetscherInnen und elf ältere Nicht-DolmetscherInnen. Aus der statistischen Analyse der Daten ging hervor, dass DolmetscherInnen ein signifikant besseres AG aufweisen als Nicht-DolmetscherInnen ($F [1,32] = 4.98, p = .033, MSE = 44.25$). Der in diesem RST inkorporierte Altersfaktor scheint allerdings keinen Einfluss auf die Leistungen zu haben. Eine mögliche Begründung für das Fehlen eines Alterseffekts ist nach Ansicht der AutorInnen (Signorelli *et al.* 2011: 13) einerseits die Tatsache, dass die TeilnehmerInnen die Tests in einer Nicht-L1-Sprache absolvierten. Andererseits weisen sie darauf hin, dass alle Versuchspersonen mehrsprachig waren. Laut vorherrschender Forschungsmeinung (s. z.B. Kavé *et al.* 2008) sind bei multilingualen Personen mit fortschreitendem Alter andere kognitive Entwicklungen zu beobachten.

Die Versuchsanordnung von Liu (2001) kann als Vorbild für eine weitere erst jüngst publizierte Untersuchung gelten (Tzou *et al.* 2011). Dabei wurde wie bei Liu (2001) ein Gedächtnisspannentest mit einer Simultandolmetschaufgabe gekoppelt. Abgesehen davon, dass es sich hier anders als bei Liu (2001) um den visuellen RST handelte, wurde die Studie in einigen wichtigen Punkten modifiziert und um weitere Aspekte bereichert. Tzou *et al.* (2011) legten besonderes Augenmerk auf die L2-Sprachkompetenz der TeilnehmerInnen. Ziel war es, die Beziehungen zwischen Dolmetschtrainingsniveau, L2-Sprachkompetenz, Dolmetschkompetenz und AG-Kapazität zu erkunden. Dabei dienten Trainingsniveau und Sprachkompetenz als unabhängige Variablen. In einem ersten Schritt wurden ProbandInnen daher nach Trainingsniveau rekrutiert: elf im ersten Studienjahr befindliche DolmetschstudentInnen, neun im zweiten Studienjahr befindliche DolmetschstudentInnen und 16 zweisprachige StudentInnen aus anderen Disziplinen als Kontrollgruppe. Die L1 aller TeilnehmerInnen war Mandarin, ihre L2 Englisch. Mit Hilfe verschiedener Messinstrumente und Selbsteinschätzungsskalen wurde anschließend eine Zweiteilung der Versuchspersonen nach L2-Sprachniveau getroffen. Der Hauptversuch bestand aus einem Simultandolmetschtest und einem RST, die beide auf Mandarin und auf Englisch abgewickelt wurden, wobei die Anforderung, dass auch eine Nicht-DolmetscherInnengruppe Simultandolmetschaufgaben absolvieren musste, bereits bei Christoffels *et al.* (2003) als experimentelles Werkzeug zu finden ist. Insgesamt stellten Tzou *et al.* (2011) beim RST einen signifikanten Gruppeneffekt auf der Variablen „Trainingsniveau“ fest ($F [2,33] = 6.74, p < .004$), der durch eine signifikante AG-Überlegenheit der beiden DolmetscherInnengruppen gegenüber der Kontrollgruppe erklärt wurde. Der Unterschied zwischen den beiden mit dem Simultandolmetschen vertrauten Gruppen erreichte jedoch nicht Signifikanz. Die Analyse der Daten nach L2-Kompetenz ergab ein Bild, das in starkem Kontrast zu den Schlussfolgerungen von Christoffels *et al.* (2006) steht. Es schien Christoffels *et al.* (2006: 339) eine Dissoziation zwischen L2-Sprachniveau und AG-Kapazität gelungen zu sein. Obwohl dieser Aspekt nicht wirklich von Tzou *et al.* (2011) thematisiert wird, ist es dennoch wichtig festzuhalten, dass die von Christoffels *et al.* (2006) vertretene Auffassung einer klaren Trennung der zwei Variablen nun aber in Frage gestellt ist. Schließlich stellen Tzou *et al.* (2011: 13) die genau entgegengesetzte Vermutung auf, dass die L2-Kompetenz der ausschlaggebende Faktor hinsichtlich AG-Kapazität und

Simultandolmetschkompetenz sei. Zwar korrelieren die Ergebnisse des RST und des Simultandolmetschtests signifikant und scheinen damit einen Trainings- oder Erfahrungseffekt anzudeuten. Allerdings werde dieser Effekt tatsächlich durch eine darunterliegende Drittvariable, die L2-Kompetenz, hervorgerufen (Tzou *et al.* 2011: 13).

Insgesamt lässt sich sagen, dass bereits ein beträchtlicher Grundstock an experimentellen Daten vorhanden ist, der aus RST/LST-Untersuchungen mit DolmetscherInnen gewonnen wurde. Nichtsdestoweniger erlauben die bisherigen Ergebnisse keine klaren Aussagen hinsichtlich der Beziehungen zwischen AG-Kapazität und Dolmetscherfahrung bzw. -ausbildungsniveau (Köpke & Signorelli 2011: 9; Signorelli *et al.* 2011: 2-3; Tzou *et al.* 2011: 3-4). Bei genauerer Betrachtung des Datenmaterials drängt sich allerdings ein Gesichtspunkt auf, der schon von Köpke und Signorelli (2011: 10) erwähnt wurde und für eine sinnvolle Sortierung der Resultate sorgen würde. Verwendet man nämlich die Modalität der Durchführung (visuell oder akustisch) als Unterscheidungskriterium, so zeichnet sich ein bemerkenswertes Muster ab. Die fünf Untersuchungen, die auf den RST zurückgreifen, scheinen eindeutig zu belegen, dass größere Erfahrung im Simultandolmetschen mit einer größeren AG-Kapazität in Verbindung gebracht werden kann (Bajo *et al.* 2000; Christoffels *et al.* 2006; Padilla *et al.* 1995; Signorelli *et al.* 2011; Tzou *et al.* 2011). Den drei LST-Studien (Köpke & Nespoulous 2006; Liu 2001, vgl. auch Liu *et al.* 2004; Nordet & Voegtlin 1998) ist hingegen gemeinsam, dass sie ein wesentlich facettenreicheres Bild vermitteln, das in der Disziplin auch weiterhin für Kopfzerbrechen sorgt. Bei Liu (2001) wie bei Nordet und Voegtlin (1998) waren keine signifikanten Differenzen bezüglich AG-Kapazität feststellbar. Die Untersuchung von Köpke und Nespoulous (2006) ergab zwar einen signifikanten Gruppeneffekt, jedoch beruhte dieser entgegen den Erwartungen auf der Überlegenheit der noch in der Ausbildung befindlichen DolmetscherInnen. In der Forschung (Köpke & Signorelli 2011; Signorelli 2008) ist in jüngster Zeit darauf hingewiesen worden, dass methodische Einzelaspekte für die Diskrepanzen in RST/LST-Studien verantwortlich gemacht werden können. Methodische Facetten, die dabei in Frage kommen könnten, sind deshalb Thema des nächsten Unterkapitels.

3.3.3. Methodische Faktoren

Bei der Analyse methodischer Faktoren im Zusammenhang mit dolmetschwissenschaftlichen RST/LST-Untersuchungen sind zwei allgemeine Vorbemerkungen angebracht. Erstens ist der noch sehr zaghafte Austausch zwischen Dolmetschwissenschaft und Kognitionspsychologie dafür verantwortlich zu machen, dass die Implementierung der neuesten Standards aus der RST/LST-Forschung nur schleppend vorangeht und oft noch veraltete, auf Daneman und Carpenter (1980) zurückgehende Praktiken eingesetzt werden. Zweitens muss vorausgeschickt werden, dass nicht immer alle methodischen Aspekte optimal untersucht werden können, da sie oft überhaupt nicht oder nur unzureichend beschrieben werden. Nicht selten ist die Ursache dieser Einschränkung die Tatsache, dass die Versuche im Rahmen von Dissertationen (Liu 2001; Signorelli 2008) oder Diplomarbeiten (Nordet & Voegtlin 1998) durchgeführt wurden, die nicht in vollem Umfang zugänglich, sondern meist nur in Form von zusammenfassenden Zeitschriftenartikeln (Liu *et al.* 2004; Signorelli *et al.* 2011; Tzou *et al.* 2011) greifbar sind.

Einige Unterschiede zwischen den Studien sind bereits hinsichtlich des Stimulusmaterials erkennbar. Beispielsweise variieren die Gesamtanzahl der verwendeten Stimuli und ihre Gliederung in verschiedene Satzblöcke. Liu (2001) arbeitete mit 70 Sätzen, die in fünf Serien aus 2- bis 5-Satz-Blöcken aufgeteilt waren. Bei Köpke und Nespoulous (2006) waren es 60 Stimulusstrings, die allerdings in 3 Serien aus 2- bis 6-Satz-Blöcken gruppiert waren. Christoffels *et al.* (2006), Signorelli *et al.* (2011) und Tzou *et al.* (2011) sind in diesem Punkt identisch. In allen drei Untersuchungen wurde pro Sprache ein Grundstock von 42 Sätzen eingesetzt, der in 3 Serien aus 2- bis 5-Satz-Blöcken gegliedert war. Es ist also eine Variation sowohl bezüglich Serienanzahl als auch Anzahl der Satzniveaus zu beobachten.

Ein weiterer uneinheitlich ausgeprägter Parameter in der Kategorie „Material“ bezieht sich auf die Länge der Satzstimuli. Liu (2001) operierte mit 13- bis 16-Wort-Sätzen. In den methodisch eng verwandten Studien von Christoffels *et al.* (2006) und Signorelli *et al.* (2011) waren es 11- bis 13-Wort-Sätze. Eine Länge von 13 bis 17 Worteinheiten hatten die Satzstrings bei Tzou *et al.* (2011). Den besten Einblick betreffend Materialerzeugung gewähren Köpke und Nespoulous (2006), die bei ihrem LST

auf Desmette *et al.* (1995) zurückgreifen. Hier wurde die Länge der Sätze sowohl nach Worteinheiten (12 bis 17 Wörter) als auch nach Silbenanzahl (22 bis 26) normiert.

Was die Merkmale der Endwörter betrifft, so erhält man nicht immer ausführlich Auskunft. Zu den unterschiedlich berücksichtigten Aspekten zählt beispielsweise, ob darauf geachtet wurde, dass sich die Endwörter innerhalb eines Blocks nicht reimen (Christoffels *et al.* 2006; Signorelli *et al.* 2011), dass sich die Wortfrequenz innerhalb einer gewissen Spanne bewegt (Christoffels *et al.* 2006, s. Beschreibung bei Desmette *et al.* 1995; Köpke & Nespoulous 2006; Signorelli *et al.* 2011), oder dass die Konkretheit der Wortitems festgelegten Vorgaben entspricht (Desmette *et al.* 1995). Eine Berücksichtigung der Wortlänge in Silben und eine Kontrolle semantischer Gesichtspunkte innerhalb von Satzblöcken werden nur von Desmette *et al.* (1995) berichtet, die die Grundlage für den LST von Köpke und Nespoulous (2006) bilden. Welcher Wortart die Endwörter angehören, wird in keiner der hier verglichenen Arbeiten erwähnt. Meist wird in der Tradition von Daneman und Carpenter (1980: 458) einfach betont (Liu 2001: 46; Tzou *et al.* 2011: 5), dass die Satzitems Texten des täglichen Lebens entnommen wurden, eine relativ einfache Struktur haben und kein besonderes Fachwissen voraussetzen.

Eine bei der Erklärung der angesprochenen Widersprüche durchaus relevante Facette steht mit dem Versuchsablauf in Zusammenhang. Es wäre möglicherweise aufschlussreich, in einer eigenen Untersuchung der Frage nachzugehen, ob unterschiedliche Mechanismen bei der Präsentation der Items einen Einfluss auf die Ergebnisse haben. Denn bei allen hier berichteten dolmetschwissenschaftlichen RST-Studien währt die Darbietung eines Satzes so lange, wie die Testperson braucht, um ihn mit „normaler“ Geschwindigkeit zu lesen, oder beim LST von Köpke und Nespoulous (2006) so lange, wie das Nachsprechen des akustisch präsentierten Stimulus in Anspruch nimmt. Es kann also nicht ausgeschlossen werden, dass die TeilnehmerInnen in der Lage sind, durch eine geschickte Dosierung des Lese- oder Sprechtempos gewisse Strategien ins Spiel zu bringen, um ihre Behaltensleistung zu steigern. Beim LST ist die Präsentationsdauer hingegen von der Audioaufnahme vorgegeben, sodass ein ähnlicher Zeitdruck herrscht wie bei einem Simultandolmetscheinsatz. Wie in der vorliegenden Studie demonstriert wird, ist es auch in der visuellen Modalität möglich, einheitliche Zeitbeschränkungen bei der Präsentation zu schaffen.

Nicht zu vernachlässigen ist bei der Aufklärung der Diskrepanzen auch der Wiedergabemodus. Der Ansicht von Köpke und Signorelli (2011) zufolge ist in diesem Punkt der entscheidende Faktor zur Entschlüsselung der widersprüchlichen Daten zu sehen. Zwei Arten der Wiedergabe sind im Rahmen von RST/LST-Versuchen möglich. Beim freien Modus können die Endwörter in einer beliebigen Reihenfolge reproduziert werden – oft mit der Einschränkung, dass das letzte Endwort nicht an erster Stelle stehen darf (Christoffels *et al.* 2006; Signorelli *et al.* 2011). Der serielle Modus ist mit dem Erfordernis verbunden, dass die Endwörter streng nach der ursprünglichen Reihenfolge wiedergegeben werden. Ein Mittelweg besteht darin, die serielle Wiedergabe in der Instruktion vorzuschreiben, bei der Auswertung aber auch frei reproduzierte Wörter einzurechnen (Padilla *et al.* 1995). Wie bereits oben besprochen, vermuten Köpke und Signorelli (2011), dass die freie Wiedergabe den Einsatz elaborativer Rehearsalmechanismen wesentlich erleichtert, wodurch SimultandolmetscherInnen bevorzugt wären.

Als zentraler methodischer Gesichtspunkt können in weiterer Folge auch die Entscheidungen der VersuchsleiterInnen bezüglich der ProbandInnenwahl und Gruppenzusammensetzung gelten. Allein schon was die Gesamtzahl der teilnehmenden Personen anbelangt, sind große Unterschiede anzutreffen. Köpke und Nespoulous (2006) und Christoffels *et al.* (2006) befinden sich dabei mit 79 bzw. 67 TeilnehmerInnen am oberen Ende des Spektrums. Alle anderen ForscherInnen arbeiten mit wesentlich kleineren Stichproben, deren Umfang sich zwischen 30 und 40 Personen bewegt: Bajo *et al.* (2000) mit 30 Personen; Liu (2001) mit 33; Nordet und Voegtlin (1998) mit 35; Padilla *et al.* (1995) mit 40; Signorelli *et al.* (2011) mit 38 und Tzou *et al.* (2011) mit 36. Dass kleine Versuchsgruppen mit Einbußen hinsichtlich der statistischen Zuverlässigkeit verbunden sind, bedarf selbstverständlich keiner eingehenden Erläuterung. Folglich ist es durchaus zulässig, eine Schwäche darin zu sehen, dass die letztgenannte Gruppe von Untersuchungen eher bescheidene TeilnehmerInnenzahlen aufzuweisen hat. Erschwerend kommt in manchen Fällen hinzu, dass gerade jene Gruppe, deren besondere kognitiven Eigenschaften im Mittelpunkt stehen, nämlich die der BerufsdolmetscherInnen, kleiner bemessen ist als die anderen. Bei Nordet und Voegtlin (Nordet & Voegtlin 1998, zit. n. Köpke & Signorelli 2011) findet man beispielsweise nur sechs BerufsdolmetscherInnen, aber sieben Dolmetschstudierende und 22 Psychologiestudierende. Bei

Christoffels *et al.* (2006) sind es nur 13 professionelle DolmetscherInnen, aber 15 Englischlehrende und 39 zweisprachige Nicht-DolmetscherInnen.

Als zusätzlich verwirrend kann die Tatsache betrachtet werden, dass die in dolmetschwissenschaftlichen Untersuchungen so zentrale Gruppe der BerufsdolmetscherInnen oft nicht einheitlich definiert ist. Das Anforderungsprofil variiert erheblich hinsichtlich der erforderlichen Jahre an Erfahrung, wie Köpke und Signorelli (2011: 11-12) in ihrer Analyse zu bedenken geben. Man könnte die Studien diesbezüglich in zwei Gruppen teilen. Bei Padilla *et al.* (1995, frische Absolventen oder ca. fünf Jahre Erfahrung), Nordet und Voegtlin (1998, zit. n. Köpke & Nespoulous 2006)¹, Bajo *et al.* (2000, frische Absolventen oder ca. 5 Jahre Erfahrung) und Liu (2001, mind. einjährige Ausbildung abgeschlossen und mind. zwei Jahre Berufserfahrung mit 40 Arbeitstagen pro Jahr) sind unter der Bezeichnung „BerufsdolmetscherInnen“ relativ junge VertreterInnen der Zunft zu verstehen. Christoffels *et al.* (2006, $M = 16$ Jahre Erfahrung), Köpke und Nespoulous (2006, $M = 16,9$ Jahre Erfahrung) und Signorelli *et al.* (2005, jüngere und ältere Gruppe zusammen: $M = 13,11$ Jahre) verwenden diese Benennung für Personen, die auf einem deutlich höheren Erfahrungsniveau anzusiedeln sind. Die Bemerkung von Köpke und Nespoulous (2006: 15), dass ihre als *novices* bezeichneten DolmetschstudentInnen mit den professionellen DolmetscherInnen von Padilla *et al.* (1995) gleichzusetzen seien, da sie hinsichtlich Erfahrung ähnlich seien und schließlich auch vergleichbare Ergebnisse erzielen, ist in diesem Zusammenhang bezeichnend.

Ein letzter Punkt, der in dieser methodischen Frage die Vergleichbarkeit erschwert, betrifft die Wahl der Versuchsgruppen. Abgesehen von Tzou *et al.* (2011) umfassen alle Forschungsarbeiten eine BerufsdolmetscherInnengruppe. Bei allen Untersuchungen außer Christoffels *et al.* (2006) kommt ferner mindestens eine DolmetschstudentInnengruppe hinzu, die es erlauben soll, die AG-Kapazität mit verschiedenen Trainings- bzw. Erfahrungsniveaus in Verbindung zu bringen. Im Rahmen von Christoffels *et al.* (2006) sind folglich auch keine Aussagen über longitudinale Effekte des Erfahrungszuwachses auf kognitive Fähigkeiten möglich. Und schließlich lässt sich festhalten, dass bei allen Studien außer Liu (2001) mindestens eine Nicht-DolmetscherInnengruppe verwendet wird. Bei Liu (2001) fehlt eine derartige Kontrollgruppe,

¹ Das Durchschnittsalter (35,8 Jahre) der sieben BerufsdolmetscherInnen in dieser Studie lässt vermuten, dass es sich um Personen mit vergleichsweise eher wenig professioneller Erfahrung handelt.

weshalb nicht ausgeschlossen werden kann, dass bei Verwendung einer Nicht-DolmetscherInnengruppe vielleicht doch ein Gruppeneffekt aufgetreten wäre.

Erst in den neueren Publikationen wird dem Aspekt der Sprache mehr Beachtung geschenkt. Durchaus von Belang wäre die Frage, in welcher Sprache die Lesespannen- und Hörspannentests bisher durchgeführt wurden. Möglicherweise spielt es eine Rolle, ob das Experiment in der Muttersprache oder einer Zweitsprache abgewickelt wurde. Zwar scheinen die Versuche von Service *et al.* (2002) zu belegen, dass bei hoher L2-Kompetenz, wie sie bei BerufsdolmetscherInnen anzunehmen wäre, keine signifikanten Unterschiede zwischen L1-AG-Kapazität und L2-AG-Kapazität festzustellen sind. Dass diese Auffassung wahrscheinlich nicht haltbar ist, scheinen allerdings vor Kurzem Tzou *et al.* (2011) gezeigt zu haben. Bei ihnen wurde ein RST sowohl in der L1 als auch in der L2 durchgeführt. Die Analyse der Daten ergab, dass die Gruppen beider L2-Sprachniveaus in der L1-Variante deutlich besser abschnitten. Nach Service *et al.* (2002) wäre zu erwarten gewesen, dass die Teilnehmer mit höherer L2-Kompetenz in beiden Varianten gleich gut sind. Von den hier berücksichtigten Studien zählen Liu (2001) und Signorelli *et al.* (2011) zu jenen, die den LST bzw. RST in der L2 (Englisch) durchführten (L1: Mandarin bzw. Spanisch). Einen besonderen Stellenwert hatte der Aspekt „Sprache“ bisher nur im Rahmen der Untersuchungen von Christoffels *et al.* (2006) und Tzou *et al.* (2011), die folglich auch die einzigen sind, die eine L1-Version (Niederländisch bzw. Mandarin) mit einer L2-Version (in beiden Fällen Englisch) des RST innerhalb einer Versuchsreihe koppelten.

Scoringmethoden wurden bereits an anderer Stelle besprochen (s. Kap. 3.2.4). Zweifellos wäre eine eigene Untersuchung zur Vorgehensweise bei der Punktevergabe sehr aufschlussreich im Hinblick auf die Enträtselung der bisher sehr widersprüchlichen Ergebnisse. Ein Teil der Forschung (Bajo *et al.* 2000; Köpke & Nespoulous 2006; Liu 2001; Padilla *et al.* 1995) stützt sich auf die von Daneman und Carpenter (1980) entwickelte Scoringmethode. Vor allem in neueren Publikationen (Christoffels *et al.* 2006; Nordet & Voegtlin 1998; Signorelli *et al.* 2011; Tzou *et al.* 2011) hat sich ein methodischer Ansatz durchgesetzt, bei dem pro richtig wiedergegebenem Endwort ein Punkt vergeben wird und als AG-Score die Gesamtzahl erinnelter Endwörter gilt. Die Forschung scheint darin übereinzustimmen, dass die Präzision und statistische Aussagekraft des Instrumentes dadurch erhöht wird (Conway *et al.* 2005: 774-776; Köpke &

Signorelli 2011: 11). Trotzdem schließen es Köpke und Signorelli (2011: 11) aus, dass sich das allgemeine Ergebnismuster in der dolmetschwissenschaftlichen RST/LST-Literatur bei einer durchgehenden Verwendung dieser Scoringmethode grundlegend verändern würde.

Ein auffälliges Merkmal der hier in Augenschein genommenen Studien tritt bei der Analyse der als Teil des Doppelanforderungsparadigmas eingesetzten Bearbeitungsaufgabe zu Tage. Wie bereits dargestellt (s. Kap. 3.2.3) hatten Daneman und Carpenter (1980) in ihrer ursprünglichen Versuchsreihe nur ein lautes Vorlesen der Satzitems vorgesehen, um die Bearbeitungskapazitäten des AG anzuzapfen. Bereits in ihrer zweiten Experimentreihe wurde jedoch eine zusätzliche Satzverifikationskomponente („richtig oder falsch“) eingeführt, die gewährleisten sollte, dass die Satzstrings vollständig verarbeitet wurden und nicht nur auf das Memorieren der Endwörter geachtet wurde. Da diese zusätzliche Maßnahme in der Kognitionspsychologie auf breite Zustimmung stößt und in maßgeblichen Arbeiten – freilich in verschiedentlich abgeänderter Form – eingesetzt wurde (Conway *et al.* 2005), erscheint es als bemerkenswert, dass in der dolmetschwissenschaftlich orientierten RST/LST-Forschung Liu (2001) die einzige ist, die auf diese anderweitig so übliche Methode zurückgreift. In allen anderen Publikationen, wo die Bearbeitungsaufgabe überhaupt dokumentiert ist (Christoffels *et al.* 2006; Köpke & Nespoulous 2006; Signorelli *et al.* 2011; Tzou *et al.* 2011), wird die Artikulation der Stimuli mit lauter Stimme offensichtlich für ausreichend gehalten. Köpke und Nespoulous (2006) verlangten bei ihrem LST von den Testpersonen, dass sie die Stimuli laut nachsprechen.

Wie die Analyse der Ergebnisse bisheriger Untersuchungen im vorigen Kapitel naheulegen scheint, könnte zuletzt ein ganz anderer methodischer Faktor herangezogen werden, um Klarheit zu schaffen. Es wurde vorgeschlagen, eine Aufgliederung der Daten nach Modalität vorzunehmen. Dabei kam ein bezeichnendes Muster zum Vorschein. RST-Studien bestätigen die Hypothese eines AG-Vorteils der BerufsdolmetscherInnen, Arbeiten mit LST hingegen nicht. Bei den Hörspannentests sind die Unterschiede entweder nicht signifikant oder scheinen eine Überlegenheit der Gruppe der Dolmetschstudierenden (Köpke & Nespoulous 2006) zu belegen. In der vorliegenden Studie wird aufbauend auf diese Beobachtung versucht, einen Beitrag zur Aufklärung des unübersichtlichen dolmetschwissenschaftlichen Ergebnisstandes zu leisten.

4. Experiment

4.1. Allgemeine Zielsetzung

Aus dem bisher Gesagten lässt sich ableiten, dass die Durchführung einer weiteren Untersuchung der AG-Kapazität bei DolmetscherInnen durchaus gerechtfertigt erscheint. Ob eine kausale Beziehung zwischen Dolmetscherfahrung und dem Niveau der kognitiven Leistungsfähigkeit – ausgedrückt durch die mit Hilfe des RST/LST operationalisierte AG-Kapazität – besteht, konnte nämlich bisher noch nicht endgültig geklärt werden, wovon die Widersprüchlichkeit der bisherigen Ergebnisse zeugt.

Im Vordergrund soll im Rahmen der vorliegenden empirischen Untersuchung daher eine methodisch möglichst saubere und an die neuesten methodischen Standards angepasste Ermittlung der AG-Kapazität dreier Versuchsgruppen mit unterschiedlich ausgeprägter Dolmetscherfahrung stehen.

Der Bedeutung methodischer Faktoren als Ursache widersprüchlicher Resultate wurde in letzter Zeit vermehrt Beachtung geschenkt. Eine eingehende Beschäftigung mit diesem Gesichtspunkt ließ es als nützlich erscheinen, die Erforschung der möglichen Auswirkungen des Faktors „Modalität“ auf das Abschneiden der VersuchsteilnehmerInnen als zusätzlichen Schwerpunkt zu setzen. Die Analyse der bisherigen Forschungsarbeiten gibt nämlich Anlass zur Vermutung, dass die Modalität des Gedächtnisspannentests für unterschiedliche Ergebnisse verantwortlich gemacht werden kann. In der Grundlagenforschung zum Thema RST/LST (Daneman & Carpenter 1980) wurde zwar demonstriert, dass diese zwei Maße stark korrelieren. Dennoch ist nicht auszuschließen, dass sich BerufsdolmetscherInnen gerade durch bessere Leistungen in einer der beiden Modalitäten auszeichnen, während sie sich in der anderen Modalität nicht signifikant unterscheiden.

Köpke und Signorelli (2011: 10) halten es für unwahrscheinlich, dass die Modalität bei Gedächtnisspannentests mit DolmetscherInnen ein entscheidender methodischer Aspekt ist. Um dies zu untermauern, wiederholten sie sogar den LST von Köpke und Nespoulous (2006) in Form eines RST mit einer Gruppe von StudentInnen. Dieser Test ergab keine signifikanten Unterschiede, tendenziell aber einen leichten Vorteil in der akustischen Modalität. Einschränkend ist anzumerken, dass es sich hierbei nicht wirk-

lich um eine vollwertige Versuchsanordnung handelte, da nur die Gruppe der StudentInnen berücksichtigt wurde.

Erstmals soll hier also eine dolmetschwissenschaftliche Untersuchung durchgeführt werden, bei der alle Versuchsgruppen sowohl einen RST als auch einen LST absolvieren. Es ist bemerkenswert, dass diese Versuchsanordnung nicht schon früher eingesetzt wurde. Modalität kann nämlich als ein methodischer Faktor betrachtet werden, der relativ unkompliziert zwischen Studien zu vergleichen ist, da er bei jeder experimentellen Arbeit schon an der Bezeichnung des eingesetzten Versuchs ersichtlich ist (Reading Span Task oder Listening Span Task). Ein Vergleich anderer methodischer Punkte ist oft auf Grund der knappen Dokumentation im Rahmen zusammenfassender Zeitschriftenartikel wesentlich schwieriger.

Es erscheint sinnvoll, drei Versuchsgruppen mit unterschiedlichem Erfahrungsniveau im Simultandolmetschen heranzuziehen. Nach dem Vorbild anderer Studien soll daher eine Gruppe ohne jegliche Erfahrung (Nicht-Dolmetschstudierende), eine in der Ausbildung befindliche (Dolmetschstudierende) und eine mit mehrjähriger Berufspraxis (BerufsdolmetscherInnen) verwendet werden. Insbesondere soll auch darauf Wert gelegt werden, dass der allgemein in diesem Forschungsbereich laut gewordenen Forderung nach größeren Stichproben nachgekommen wird.

4.2. Forschungsfragen

In der vorliegenden Arbeit gilt es, auf der Grundlage eines kombinierten deutschsprachigen RST/LST-Versuchs mit drei Gruppen, die hinsichtlich der Erfahrung im Simultandolmetschen eindeutig variieren, folgende Fragenkomplexe zu erörtern:

- *Haupteffekt von Erfahrung:* Gibt es abhängig von der Erfahrung im Simultandolmetschen signifikante Unterschiede in der AG-Kapazität, wenn RST- und LST-Ergebnisse gemeinsam betrachtet werden? Ist also ein Haupteffekt von Erfahrung feststellbar? Ist allgemein ein größeres AG mit mehr Erfahrung in Verbindung zu bringen, wie von einem Teil der Literatur angedeutet wird? Schneiden also BerufsdolmetscherInnen signifikant besser ab als Dolmetschstudierende und diese ihrerseits besser als Nicht-DolmetscherInnen?

- *Einfache Haupteffekte von Erfahrung:* Lassen sich bei der getrennten Analyse der modalitätsspezifischen Teilscores signifikante Unterschiede in der AG-Leistung der Gruppen abhängig von Erfahrung im Simultandolmetschen feststellen? Ist ein überzufälliger Vorteil der BerufsdolmetscherInnen sowohl beim Les- als auch beim Hörspantentest vorhanden? Oder zeichnen sich professionelle DolmetscherInnen durch einen AG-Vorteil aus, der auf eine bestimmte Modalität beschränkt ist? Kann also die sich bei der Analyse bisheriger Ergebnisse aufdrängende Vermutung bestätigt werden, dass BerufsdolmetscherInnen nur in der visuellen Variante des Tests signifikant vorne liegen, während der LST keine signifikanten Unterschiede aufzeigt oder vielleicht sogar einen Vorteil der „Neulinge“ wie bei Köpke und Nespoulous (2006) nahelegt?
- *Haupteffekt von Modalität:* Liegt ein Haupteffekt von Modalität vor? Ist unabhängig von der Erfahrungsstufe ein allgemein besseres Abschneiden aller Gruppen in einer bestimmten Modalität zu beobachten?
- *Einfache Haupteffekte von Modalität:* Ergeben sich signifikante Differenzen, wenn die modalitätsspezifischen Teilscores isoliert nach Versuchsgruppen (= Erfahrungsstufen) verglichen werden? Ist der AG-Mittelwert einer Gruppe in einer Modalität deutlich besser als in der anderen?
- *Wechselwirkung zwischen Erfahrung und Modalität:* Sind signifikante Interaktionseffekte zwischen Erfahrung im Simultandolmetschen und Modalität zu beobachten?

4.3. Methode

Es kam ein ursprünglich von Daneman und Carpenter (1980) vorgestelltes Versuchsparadigma zum Einsatz. Diese zwei wegbereitenden Forscherinnen prägten den Namen „Reading Span Task“ (RST) für die visuelle und „Listening Span Task“ (LST) für die akustische Version dieses AG-Tests. Seither hat sich das Experiment erfolgreich als

maßgebliches Instrument zur Ermittlung der AG-Kapazität durchgesetzt. Die große Beliebtheit dieses Versuchs trug auch zur ihrer methodischen Entfaltung bei. Im Laufe der Jahre wurden regelmäßig Verbesserungen vorgeschlagen, um eine optimale Kalibrierung des Instruments zu erreichen. Bei der Durchführung dieser Forschungsarbeit wurde deshalb darauf geachtet, dass eine möglichst moderne Methodenpalette zum Einsatz kam. Aufschlussreiche Zusammenstellungen methodischer Aspekte aus der Perspektive der neueren kognitionspsychologischen RST/LST-Forschung sind bei Conway *et al.* (2005) und Van den Noort *et al.* (2008) zu finden. Sie bilden auch gleichzeitig in vielerlei Hinsicht die Grundlage der vorliegenden Untersuchung, wie in den folgenden Unterkapiteln näher erläutert wird.

4.3.1. TeilnehmerInnen

Bei der TeilnehmerInnenauswahl wurde die Methode der anfallenden Stichproben verwendet, d.h. die Versuchspersonen wurden nicht zufällig aus der Grundgesamtheit der MerkmalsträgerInnen rekrutiert. Trotzdem wurde durch die Festlegung möglichst durchdachter Anforderungsprofile versucht, repräsentative und homogene Stichproben zu bilden. Der Stichprobenziehung lag die organismische Variable „Simultandolmetscherfahrung“ zu Grunde. Dieses Merkmal lag in drei verschiedenen Ausprägungen vor: viel, wenig und keine Erfahrung.

Für die Gruppe mit „viel Erfahrung“ wurde vorausgesetzt, dass es sich um BerufsdolmetscherInnen mit mindestens fünf Jahren Berufspraxis und vorzugsweise mit abgeschlossener universitärer Dolmetscherausbildung handelt. Die Versuchsgruppe mit „wenig Erfahrung“ sollte aus Studierenden des Masterstudiengangs „Konferenzdolmetschen“ des Zentrums für Translationswissenschaft (Universität Wien) zusammengesetzt sein. Um zu gewährleisten, dass alle einigermaßen mit dem Simultandolmetschen vertraut waren, wurde als Aufnahmekriterium zusätzlich festgelegt, dass die Angehörigen dieser Gruppe entweder vor kurzer Zeit die praktische Abschlussprüfung abgelegt oder mindestens schon vier Semester aktiv „Konferenzdolmetschen“ studiert hatten².

² Im Curriculum des Masterstudiengangs „Dolmetschen“ (Schwerpunkt: Konferenzdolmetschen) am Zentrum für Translationswissenschaft (Universität Wien) ist eine Mindeststudiendauer von zwei Jahren festgeschrieben. In diesem Zeitraum soll durch den Besuch von praktisch orientierten „Übungen“ ein berufsrelevantes Niveau im simultanen und konsekutiven Dolmetschen erworben werden. Der Erwerb der beiden Modi erfolgt gleichzeitig und nicht, wie an manch anderen Ausbildungseinrichtungen für Dolmet-

Schließlich sollten für eine Kontrollgruppe „ohne jegliche Erfahrung“ im Dolmetschen Studierende aus dolmetschfernen Studienrichtungen rekrutiert werden. Da im Rahmen dieser Untersuchung keine Spracheffekte studiert werden sollten, wurde als weitere Teilnahmeanforderung festgelegt, dass alle ProbandInnen Deutsch als Muttersprache haben sollten.

Es wurden alle Kanäle der Kontaktaufnahme genutzt. Die Rekrutierung von Dolmetschstudierenden und Studierenden aus anderen Fächern wurde wesentlich durch die Tatsache erleichtert, dass der Versuchsleiter selbst noch dieser Personengruppe angehört und in diesem Kreis deshalb schneller Teilnahmezusagen erhielt. Es gestaltete sich erwartungsgemäß erheblich schwieriger, eine zufriedenstellend große Anzahl an BerufsdolmetscherInnen zu rekrutieren.

Nach einer ersten Kontaktaufnahme, bei der versichert wurde, dass die Anonymität im Rahmen der Forschungsarbeit gewahrt bleiben würde, erhielten die TeilnehmerInnen eine Nachricht, in der sich ein Link zur Terminkoordinationsplattform „congregar.com“ befand. Hier wurde auf einer eigens eingerichteten Internet-Seite das Forschungsvorhaben nochmals überblicksmäßig beschrieben. Sie bot den Interessenten überdies die Möglichkeit, sich für eine Sitzung ihrer Wahl einzutragen.

Es wurde eine Gesamtzahl an TeilnehmerInnen angestrebt, die nicht unter dem Mittelwert der bisherigen dolmetschwissenschaftlichen Studien (ca. 48) liegen sollte. Ursprünglich waren daher Versuchsgruppen zu je 16 Personen ($n = 16$) vorgesehen. Schlussendlich gelang es, die Gruppen noch ein wenig zu vergrößern: 19 BerufsdolmetscherInnen (= BD), 19 Dolmetschstudierende (= DS) und 18 Nicht-Dolmetschstudierende (= ND).

Nach der Datenerhebung mussten zwei Fälle ausgeschlossen werden, da festgestellt wurde, dass sie die Anforderungskriterien nicht erfüllten. Versuchsperson 12 (P12) aus der Gruppe der BD hatte zwar langjährige Erfahrung im Gesprächsdolmetschen, das Simultandolmetschen gehörte allerdings seit vielen Jahren nicht mehr zu ihrem Berufsalltag, sodass es ratsam schien, diesen Fall nicht weiter zu berücksichtigen. Auch die Ergebnisse von P17 aus der Gruppe der DS mussten ausgeklammert werden, da sie be-

scherInnen, in einer festgelegten Reihenfolge. Es ist ferner vorgesehen, dass Studierende nach dem erfolgreichen Abschluss des praktischen Lehrveranstaltungskontingents zu einer „Modulprüfung“ antreten, bei der im Rahmen verschiedener Teilprüfungen die Dolmetschkompetenz getestet wird. Unabhängig davon ist eine dolmetschwissenschaftliche Forschungsleistung in Form einer Masterarbeit zu erbringen.

reits ein Jahr beruflich dolmetschte und ihre Ausbildung nicht am Zentrum für Translationswissenschaft absolviert hatte.

Die Gruppen umfassten daher schließlich jeweils 18 Personen; insgesamt flossen also 54 Datensätze in die Analyse ein:

- **18 BerufsdolmetscherInnen (= BD)**, davon 16 weiblich, wurden bewertet. Sie waren zwischen 33 und 60 Jahre alt und hatten ein Durchschnittsalter von 47 Jahren ($SD = 7.26$). Die Erfahrung im Simultandolmetschen betrug minimal 6 Jahre, maximal 35 Jahre, wobei der Mittelwert bei 21 Jahren ($SD = 7.90$) lag.
- **18 Dolmetschstudierende (= DS)**, alle weiblich, wurden bei der Auswertung berücksichtigt. Sie hatten ein Alter zwischen 23 und 28 Jahren. Das Durchschnittsalter dieser Gruppe lag bei 25 Jahren ($SD = 1.42$). Minimal belief sich ihre Erfahrung auf 2 Jahre (6 Fälle), maximal auf 3,5 Jahre (2 Fälle). Die durchschnittliche Erfahrung betrug 2,5 Jahre ($SD = 0.50$).
- **18 Nicht-Dolmetschstudierende (= ND)**, davon 13 weiblich, nahmen als Kontrollgruppe am Versuch teil. Ihr Alter lag zwischen 21 und 30 Jahren, wobei das Durchschnittsalter 24 Jahre ($SD = 2.07$) betrug.

In Tabelle 1 sind die wichtigsten demographischen Daten der drei Versuchsgruppen nochmals dargestellt. Weitere deskriptivstatistische Details sind im Anhang enthalten.

Tabelle 1

Demographische Daten der Versuchsgruppen

Versuchsgruppe ($N = 54$)	Alter		Erfahrung ^{a)}		Geschlecht ^{b)}	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	w	m
BD ($n = 18$)	47.06	7.26	20.89	7.90	16	2
DS ($n = 18$)	25.44	1.42	2.53	0.50	18	-
ND ($n = 18$)	24.17	2.07	-	-	13	5

Anmerkungen. BD = BerufsdolmetscherInnen. DS = Dolmetschstudierende. ND = Nicht-Dolmetschstudierende.

^{a)} Die Erfahrung ist zwecks Vergleichbarkeit bei den DS in Jahren und nicht in Semestern angegeben.

^{b)} w = weiblich, m = männlich.

4.3.2. *Material*

Der Versuch wurde auf Deutsch, der L1-Sprache der ProbandInnen, abgewickelt. Da in der dolmetschwissenschaftlichen Fachliteratur keine deutschsprachigen RST/LST-Untersuchungen greifbar sind, setzte sich der Versuchsleiter mit dem federführenden Autor der Publikation Van den Noort *et al.* (2008) in Verbindung. Dort wird unter besonderer Berücksichtigung neuester methodischer Erkenntnisse eine mehrsprachige standardisierte Version des RST vorgestellt. Diese überarbeitete, computergestützte Variante des Tests wurde neben Englisch, Niederländisch und Norwegisch auch auf Deutsch realisiert. Nach Rücksprache mit Maurits van den Noort konnte schließlich das deutschsprachige Satzmaterial, das im Rahmen der Studie erzeugt worden war, beschafft werden.

Van den Noort *et al.* (2008) beabsichtigten mit dieser neuen Version des RST zahlreiche Schwächen des ursprünglichen Tests von Daneman und Carpenter (1980), die im Laufe der Jahre offensichtlich geworden waren, zu beseitigen. Sie heben einige verbesserungswürdige Aspekte hervor und schlagen Lösungen vor. Ein Teil der Kritikpunkte betrifft die psycholinguistische Zusammensetzung der Sätze. An dieser Stelle erscheint eine kurze Besprechung dieser linguistischen Aspekte zusammen mit den von Van den Noort *et al.* (2008: 36-37) befürworteten Lösungen sinnvoll:

- **Satzlänge:** Ein Großteil der Studien berücksichtigt einzig die Anzahl der Wörter pro Stimulussatz. Wichtig wäre außerdem eine Standardisierung der Silbenanzahl, wie bei Desmette *et al.* (1995), sowie der Buchstabenanzahl. Die von Van den Noort *et al.* (2008: 36) erzeugten Strings umfassen daher in allen Sprachen zwischen 12 und 17 Wörter, 20 bis 22 Silben und 55 bis 73 Buchstaben.
- **Endwortlänge:** Ein weiterer, ursprünglich nicht bedachter Faktor ist die Länge der zu memorierenden Endwörter. Diesen Aspekt zu kontrollieren, wäre wichtig, da gemäß dem Wortlängeneffekt (Gathercole 2008: 35) generell mehr kurze als lange Wörter im AG behalten werden können. Die Silben- und Buchstabenanzahl wurde also von Van den Noort *et al.* (2008) über alle Serien hinweg ausbalanciert.

- **Häufigkeitswerte der Endwörter:** Ferner geben Van den Noort *et al.* (2008) zu bedenken, dass die Häufigkeitswerte der Endwörter in der jeweiligen Sprache eine große Rolle bei der Reproduktion spielen. Ein höherer Frequenzwert erleichtert den Abruf aus dem AG. Es wurde also bei der Wahl der Endwörter die CELEX-Datenbank verwendet, um eine ausgeglichene Frequenz zu gewährleisten.
- **Konkretheitsgrad der Endwörter:** Nach Ansicht von Van den Noort *et al.* (2008: 36) wurde der Konkretheit der Endwörter bislang noch zu wenig Bedeutung beigemessen. Sie betonen, dass Items mit größerer Bildhaftigkeit durch die Verwendung visueller Speicherstrategien besser behalten werden. Folglich ließ man in einem Kontrollverfahren mit 120 TeilnehmerInnen (Van den Noort *et al.* 2008: 38) die Konkretheit der Endwörter auf einer sechsstufigen Skala bewerten und verteilte sie gleichmäßig nach diesem Parameter auf die fünf Serien. Der Konkretheitsgrad war dabei in den verschiedenen Sprachen vergleichbar.
- **Semantische Beziehungen:** Um auszuschließen, dass semantische Beziehungen zwischen benachbarten Endwörtern oder ganzen Sätzen innerhalb von Satzblöcken oder Serien bei der Behaltensleistung im Spiel sind, wurde auch dieser Aspekt von Van den Noort *et al.* (2008) unter Zuhilfenahme dreier MuttersprachlerInnen kontrolliert.
- **Mehrsprachige RST:** Van den Noort *et al.* (2008: 36) merken an, dass in vielen Sprachen einfach Übersetzungen des englischen Stimulusmaterials von Daneman und Carpenter (1980) zum Einsatz kommen, was den Nachteil habe, dass viele der linguistischen Kriterien in der Zielsprache nicht mehr erfüllt werden. Wie Hacker *et al.* (1994: 298) folgerichtig feststellen, müsste das durch eine Übersetzung produzierte Satzmaterial nochmals in einem eigenen Verfahren an die Kriterien angepasst werden. Van den Noort *et al.* (2008) entwickelten daher von Grund auf neues Material.
- **Wortart der Endwörter:** Die Endwörter waren allesamt Nomen.

Insgesamt ist das Ergebnis ein nach strengeren Kriterien als zuvor üblich erstellter Satzkatalog, der die Reliabilität des Verfahrens erhöht und den Vergleich von RST-Ergebnissen zwischen mehreren Sprachen erleichtert (Van den Noort *et al.* 2008: 36).

Es standen insgesamt 105 deutsche Stimulussätze zur Verfügung. Da ein kombinierter RST/LST-Test durchgeführt werden sollte und nicht nur ein Lesespannentest wie bei Van den Noort *et al.* (2008), musste das Material in zwei Hälften geteilt werden. Es wurden schlussendlich 47 Sätze für den RST verwendet und ebenso viele für den LST, wobei darauf geachtet wurde, dass die besprochenen Kriterien eingehalten wurden. Von den jeweils 47 Strings wurden fünf für einen Probelauf verwendet. Dieser bestand aus einem 2- und einem 3-Satz-Block. Die restlichen 42 Stimuli wurden in drei Serien gespalten, die jeweils einen 2-Satz-Block, einen 3-Satz-Block, einen 4-Satz-Block und einen 5-Satz-Block umfassten. Den Hauptteilen des Experiments (RST bzw. LST) war jeweils eine Probeserie vorgeschaltet. Den TeilnehmerInnen wurde damit die Möglichkeit gegeben, sich bereits vor Beginn der Hauptserien zu orientieren. Dann folgten die drei Serien.

Für den LST war es erforderlich, die Stimuli als Tondateien aufzunehmen. Zu diesem Zweck dienten zwei verschiedene weibliche Stimmen, die auf Grund ihrer Klarheit und neutralen Aussprache für besonders geeignet gehalten wurden. Als Vorsichtsmaßnahme erschien es sinnvoll, nicht nur die 47 für den LST notwendigen Sätze, sondern wesentlich mehr Stimuli zu vertonen, um für den Fall fehlerhafter, verzerrter oder abgehackter Aufnahmen gewappnet zu sein. Insgesamt wurden daher mit jeder Stimme 58 Sätze in normalem Sprechtempo aufgenommen. Die Dateien wurden mit Hilfe der Audiosoftware Audacity 1.3.13-beta nachbearbeitet und alle einheitlich auf die von Van den Noort *et al.* (2008: 37) empfohlene Länge von 6,5 Sekunden gebracht. Unbrauchbare Aufnahmen wurden durch Reservesätze ersetzt. Die Aufnahmen mit Stimme 1 (= V1) wurden abends (18:00-19:30 Uhr) in ruhiger Umgebung auf einem Apple MacBook Pro mit der Software Garageband durchgeführt. Als Dateiformat wurde MP3 bei einer Datenrate von 320 kbit/s gewählt. Mit Stimme 2 (= V2) erfolgte die Aufnahme ebenfalls abends (18:00-19:00 Uhr) in ruhiger Umgebung. Allerdings wurde ein hochwertiges Aufnahmegerät von Sony, der Zoom H4n MP3-Recorder, verwendet. Dateiformat und Datenrate blieben unverändert.

Eine Auflistung einiger der verwendeten Satzstimuli ist im Anhang zu finden.

4.3.3. Versuchsaufbau und -ablauf

Der Versuch wurde mit Hilfe der Experimentsteuerungssoftware Paradigm der Firma Perception Research Systems Inc. programmiert.

Dem Grundkonzept des Forschungsansatzes entsprechend wurde zuerst eine Zweiteilung des Experiments nach Modalität festgelegt. Um möglichen Sequenzeffekten (vgl. Hussy *et al.* 2010) vorzubeugen, wurde darauf geachtet, dass die Reihenfolge, in der RST und LST absolviert wurden, vollständig interindividuell ausbalanciert war. Auf eine Versuchsperson, die zuerst den RST und dann den LST machte, folgte also immer eine Versuchsperson, die in der umgekehrten Reihenfolge getestet wurde.

Die Doppelaufgabenanforderung ist als wesentliches Charakteristikum der Methode zu nennen. Dabei müssen die Versuchspersonen einerseits einzelne unzusammenhängende Sätze lesen bzw. hören und beurteilen. Andererseits sind die präsentierten Endwörter zu behalten und blockweise wiederzugeben. Derartige Doppelverfahren, die aus konkurrierenden Bearbeitungs- und Behaltenskomponenten bestehen, werden in der Forschung als reliable und valide Operationalisierung der AG-Kapazität betrachtet (Conway *et al.* 2005: 783).

Jeder der zwei modalitätsspezifischen Teile begann mit Instruktionen, in denen der grundsätzliche Ablauf beschrieben und konkrete Anweisungen gegeben wurden (s. Anhang A: Instruktionen für TeilnehmerInnen). Es wurde dabei besonders betont, dass die Satzbeurteilungsaufgabe per Mausklick für die Auswertung wichtig sei und ferner dass die Reihenfolge, in der die Endwörter wiedergegeben wurden, keine Rolle spiele. Für das Lesen dieser Anleitungen konnten sich die ProbandInnen beliebig lange Zeit lassen. Bei beiden Teilen des Experiments folgten danach zwei Probeblöcke, ein 2- und ein 3-Satz-Block. An dieser Stelle hatten die Versuchspersonen bei Unklarheiten noch die Möglichkeit, Fragen persönlich an den Versuchsleiter zu richten. Es begannen anschließend die drei Hauptserien, wobei vor jeder Serie für 1,5 Sekunden eine Einstellung mit der Nummer der Serie erschien. Die drei Serien waren bei beiden Modalitäten gleich gestaltet. Die Stimuli wurden einzeln entweder visuell oder akustisch präsentiert und dauerten stets 6,5 Sekunden. Die visuellen Items erschienen zentriert in schwarzer Schrift auf weißem Bildschirmhintergrund (s. Anhang E: Screenshots). Es war hier also im Gegensatz zu anderen in der dolmetschwissenschaftlichen Literatur beschriebenen

RST-Studien ein Zeitlimit in der visuellen Modalität vorhanden, wodurch verhindert werden sollte, dass die TeilnehmerInnen absichtlich ein langsames Lesetempo wählen, um die Endwörter besser speichern zu können. Van den Noort *et al.* (2008: 36) geben zu bedenken, dass es sich dann streng genommen nicht mehr um einen AG-Test handle. Es bestand folglich sowohl beim RST als auch beim LST ein Zeitdruck wie ihn Conway *et al.* (2005: 773) fordern.

RST und LST setzten sich jeweils aus drei Hauptserien zusammen. Eine Serie beinhaltete einen 2-, einen 3-, einen 4- und einen 5-Satz-Block. Das niedrigste Satzniveau umfasste also zwei, das höchste fünf Items. In Summe absolvierte jede Testperson 105 Stimuli. In der Forschungsliteratur werden Blockgrößen von 2 bis 6 beschrieben. Im Rahmen der hier behandelten dolmetschwissenschaftlichen Arbeiten werden überwiegend 2- bis 5-Satz-Gruppen verwendet (Christoffels *et al.* 2006; Liu 2001; Signorelli *et al.* 2011; Tzou *et al.* 2011) und auch Conway *et al.* (2005: 773) halten diese Größe bei Untersuchungen mit Studierenden für angemessen.

Nach jedem Block wurde ein Textfeld für die Wiedergabe der Endwörter angezeigt. Es sollten so viele Endwörter wie möglich ungeachtet der Reihenfolge abgerufen werden. In den Reproduktionsphasen stand den TeilnehmerInnen beliebig viel Zeit zur Verfügung. Es oblag ihrem Ermessen, wann sie per Tastendruck die Fortsetzung einleiteten. Drei Serien wurden durchlaufen, dann erschienen bereits die Instruktionen für Teil 2 des Tests und es folgte wieder eine Probeserie und drei Hauptserien, diesmal in der anderen Modalität. Wie bereits oben erläutert (s. Kap. 593.2.6), ist die vorliegende Studie durch die von Van den Noort *et al.* (2008) angeregte Besonderheit charakterisiert, dass die Satzebenen nicht in schrittweise zunehmender Größe dargeboten werden, sondern in randomisierter Reihenfolge. Die ProbandInnen können die nächste Satzebene nicht antizipieren und daher auch nicht an die jeweilige Blockgröße angepasste Strategien entwickeln.

Von grundlegender Bedeutung war auch die Wahl einer geeigneten Bearbeitungsaufgabe, die entsprechend dem hier angewandten Paradigma für eine repräsentative und nachvollziehbare kognitive Doppelbelastung des AG sorgen sollte. Es wurden bereits Zweifel geäußert (s. Kap. 3.3.3), ob das bloße Artikulieren der Stimuli durch lautes Vorlesen oder Nachsprechen eine ausreichende inhaltliche und kognitiv sorgfältige Verarbeitung der Satzstrings garantiert. Es ist nämlich zu vermuten, dass die Auf-

merksamkeit der TeilnehmerInnen in dieser Anordnung wohl vorwiegend auf das Einspeichern der Endwörter gerichtet ist. Um einer derartigen Verhaltensweise vorzubeugen, erschien es zweckmäßig, ein ähnliches Vorgehen wie bei Liu (2001) zu wählen. Bei ihr wurde allerdings eine Beurteilung der Sätze nach dem Kriterium der Sinnhaftigkeit verlangt. In der vorliegenden Studie hatten die TeilnehmerInnen demgegenüber zu beurteilen, ob der jeweilige Satz ihrem subjektiven Sprachgefühl zufolge syntaktisch und stilistisch „gut“ oder „schlecht“ sei. Diese Wahl der Bearbeitungsaufgabe wurde vor dem Hintergrund der Tatsache getroffen, dass ein Kriterium bei der Materialerzeugung durch Van den Noort *et al.* (2008) die ausschließliche Verwendung von Nomen als Endwörter war. Diese Vorgabe hatte zur Folge, dass oft Stimuli produziert wurden, bei denen die Satzstellung für Deutsch-MuttersprachlerInnen ungewöhnlich, stilistisch „schlecht“ anmutet. Von dieser Eigenschaft wurde bei der Bearbeitungsaufgabe Gebrauch gemacht. Die ProbandInnen hatten bei jedem Satzitem per Mausklick ein stilistisches Urteil nach „gut“ oder „schlecht“ abzugeben. Zusammen mit den Sätzen wurden am Bildschirm immer ein rotes und ein grünes Feld angezeigt. Sie gaben an, welche Maustaste für welches Urteil zu verwenden war („gut“ = grün, linke Maustaste; „schlecht“ = rot, rechte Maustaste; s. auch Anhang E: Screenshots).

Das Klickverhalten der Versuchspersonen konnte auf Grund technischer Beschränkungen bei der Programmierung nur beim RST-Teil des Experiments registriert werden. Wurden bei der Überprüfung dieser Daten auffällige Klickmuster festgestellt, so konnte vermutet werden, dass die betreffende Versuchsperson strategisch vorgeht und sich vielleicht vorwiegend auf das effiziente Behalten der Endwörter konzentrierte, was einen Ausschluss des Datensatzes rechtfertigen würde. Anzumerken ist darüber hinaus, dass es bei der Satzbeurteilung kein vom Versuchsleiter festgelegtes Lösungsschema gab. Die Ergebnisse der Bearbeitungsaufgabe wurden bei der Punktevergabe daher auch nicht berücksichtigt.

4.3.4. Durchführung

Nach der Fertigstellung einer vorläufigen Testversion des Experiments mit Paradigm wurde eine zweistufige Vorversuchsphase durchlaufen. Eine erste Vorversuchssitzung fand am 16. November 2011 mit drei Personen in Medienlabor 1 des Zentrums für

Translationswissenschaft (= ZTW) der Universität Wien statt. Da es sich um den Raum handelte, der auch für das eigentliche Experiment vorgesehen war, konnte bei dieser Gelegenheit eine erste Bestandsaufnahme der örtlichen Gegebenheiten gemacht werden. Dies war wichtig, um Informationen hinsichtlich technischer Ausstattung und Verfügbarkeit des Saales einzuholen. Am 21. November 2011 wurde eine weitere Sitzung mit drei TeilnehmerInnen durchgeführt, diesmal in privaten Räumlichkeiten. Insgesamt half diese Probephase dabei, die optische Gestaltung des Experiments sowie die sprachliche Formulierung der Instruktionen zu verbessern. Ferner konnten letzte technische Pannen und Programmierfehler behoben werden. In Zusammenarbeit mit dem Medienteam des ZTW wurde sichergestellt, dass die notwendige Software auf den PCs in Medienlabor 1 installiert wurde und einwandfrei funktionierte. Diese Sitzungen wurden ebenfalls dazu genutzt, eine Wahl hinsichtlich der für das Experiment zu verwendenden Hauptstimme zu treffen. Die Mehrheit der Vorversuchspersonen sprach sich für V2 aus. Beim Hauptversuch kamen daher fast ausschließlich die V2-Aufnahmen zum Einsatz. Nur in einem Fall wurde auf V1 zurückgegriffen, da zu befürchten war, dass die Versuchsperson V2 erkennen könnte und die Leistung dadurch beeinträchtigt werden könnte.

Der Hauptversuch wurde im Zeitraum vom 23. November bis 16. Dezember 2011 in 32 jeweils ungefähr 20-minütigen Sitzungen durchgeführt. Davon fanden 28 am ZTW statt (27 in Medienlabor 1, eine in Seminarraum 2). Die Zahl der externen Durchführungen belief sich auf vier. Das Medienlabor 1 am ZTW ist mit PCs ausgestattet, die das Betriebssystem Windows 7 Enterprise 64-Bit verwenden. Den Versuchspersonen standen in diesem Raum geräuschreduzierende Kopfhörer zur Verfügung, deren Lautstärke nach Belieben eingestellt werden konnte. Bei den externen Sitzungen sowie im Seminarraum 2 wurde ein mit Windows 7 Professional 64-Bit und Standardkopfhörern ausgerüstetes Dell Latitude E6410 Notebook benutzt.

Die Sitzungen wurden mit möglichst kleinen TeilnehmerInnengruppen abgewickelt. Damit wurde gewährleistet, dass der Versuchsleiter das Verhalten aller ProbandInnen gleichzeitig im Auge behalten konnte. Die Anzahl der Personen pro Sitzung schwankte zwischen 1 (16 Fälle) und 4 (1 Fall) bei einem Mittelwert von 1,7. Außer den Versuchspersonen und dem Versuchsleiter war niemand im Saal anwesend. Um möglichst gleiche Bedingungen für alle TeilnehmerInnen zu schaffen, wurde die Mehrheit der Sitzungen am späten Nachmittag zwischen 16:00 und 18:00 Uhr durchgeführt, wo-

bei der früheste Termin um 13:30 Uhr (1 Fall) und der späteste um 19:00 Uhr (1 Fall) begann.

Die Versuchspersonen wurden zu Beginn der Sitzung vom Versuchsleiter empfangen, der ihnen einen Platz zuwies und die technische Ausstattung erklärte. Die Sitzordnung wurde so gewählt, dass die TeilnehmerInnen möglichst weit voneinander entfernt waren und keinen Blickkontakt hatten. In Medienlabor 1 wurden folgende PCs verwendet: LAB1STUD03, LAB1STUD07, LAB1STUD16 und LAB1STUD19. Das Experiment selbst betreffend gab der Versuchsleiter vorerst keine persönlichen Anleitungen, da ohnehin gleich zu Beginn des computergestützten Versuchs ausführliche Instruktionen vorgesehen waren. Gab es danach noch offene Fragen, so beantwortete der Versuchsleiter diese persönlich.

Nach Abschluss des Experiments hatten die TeilnehmerInnen ein Protokoll auszufüllen, das in gewissen Punkten gezielt an die jeweilige Versuchsgruppe angepasst war (s. Anhang D: Protokolle).

4.3.5. *Scoring*

Im Kapitel über Scoringmethoden (Kap. 3.2.4) wurde ein Überblick über das breite Spektrum der Punktevergabeverfahren und die diesbezüglichen Veränderungen seit Daneman und Carpenter (1980) gegeben. Die in dieser Studie gewählte Vorgehensweise beruht auf den empirischen Erkenntnissen von Conway *et al.* (2005), die die Verwendung des sog. Partial-Credit Unit Scoring (PCU) nahelegen.

Wesentlich ist bei der hier verwendeten Methode einerseits, dass alle reproduzierten Endwörter zum Gesamtergebnis beitrugen (*partial credit*), d.h. auch sehr lückenhaft bewältigte Blöcke berücksichtigt und nicht wie bei Daneman und Carpenter (1980) einfach eliminiert wurden, wenn der Anteil der richtig wiedergegebenen Elemente unter eine gewisse Schwelle fiel (2/3 beim RST oder 3/5 beim LST; *all-or-nothing*). Andererseits zeichnet sich dieses Scoringverfahren dadurch aus, dass alle Satzebenen gleich gewichtet werden (*unit scoring*). Dazu wird für jeden Block ein relativer Score zwischen 0 und 1 kalkuliert. Werden beispielsweise in einem 4-Satz-Block drei Endwörter richtig wiedergegeben, so beträgt der Wert für diesen Block 0,75 oder 75 %. Die Ergebnisse der einzelnen Blöcke werden anschließend addiert und an der

Gesamtzahl der Blöcke relativiert. Insgesamt wird also der mittlere Anteil richtig reproduzierter Endwörter pro Satzblock ermittelt. Auf diese Weise wurden getrennte Gesamtscores für den RST und den LST berechnet. Bemerkenswert ist in den Ausführungen von Conway *et al.* (2005: 775) die Tatsache, dass das Ergebnis bei Verwendung der PCU-Methode höher ausfällt als bei den anderen Verfahren, sodass leicht der Eindruck eines Deckeneffekts entstehen könnte. Diese Schwäche wird allerdings durch das höhere Maß an Empfindlichkeit, das durch das PCU-Scoring erreicht wird, aufgewogen.

Die Ergebnisse der Satzbeurteilungsaufgabe hatten, wie bereits beschrieben, keinen direkten Einfluss auf die Berechnung der Scores. Es war aber festgelegt worden, dass die RST-Klickprofile der TeilnehmerInnen im Rahmen der Auswertung auf Auffälligkeiten geprüft werden sollten. Augenscheinlich strategisches Verhalten würde die Disqualifikation eines Datensatzes rechtfertigen. In der Untersuchung wurden jedoch keine auffälligen Klickmuster entdeckt und folglich auch keine Datensätze auf Grundlage dieses Kriteriums ausgeschlossen.

4.3.6. *Versuchsplan und statistisches Analysekonzept*

Die Versuchsprotokolle dienten als Vorlage für die Entwicklung eines Codeplans. Auf dieser Grundlage erfolgte die Eingabe der Rohdaten in SPSS.

Die vorliegende empirische Untersuchung basiert auf einem quasiexperimentellen Versuchsplan. Diese Forschungsmethode ist dadurch charakterisiert, dass das Vorgehen in bestimmter Hinsicht weniger randomisiert ist als bei anderen Experimenten (vgl. z.B. Hussy *et al.* 2010). Im vorliegenden Fall bedeutete die Einführung der organisatorischen Variablen „Erfahrung“ eine systematische Zuteilung der ProbandInnen zu einer Versuchs- oder Kontrollgruppe und somit eine Einschränkung der Randomisierung. Ein weiterer quasiexperimenteller Aspekt ist in der intraindividuellen Bedingungsvariation an der unabhängigen Variablen (= UV) „Modalität“ zu sehen, d.h. es wurde eine Messwiederholung (RST und LST) mit gleichen Gruppen durchgeführt. Man kann daher von abhängigen Stichproben sprechen. Um Sequenzeffekten entgegenzuwirken, erfolgte – wie bereits an anderer Stelle ausgeführt – eine vollständige interindividuelle Ausbalancierung (wechselnde Anordnung von RST und LST).

Das Versuchsdesign ist darüber hinaus als zweifaktoriell zu beschreiben. Die zwei beteiligten UV waren „Erfahrung“ als dreistufiger Zwischensubjektsfaktor und „Modalität“ als zweistufiger Innersubjektsfaktor. Als abhängige Variable diente die durch den RST/LST ermittelte AG-Kapazität.

4.4. Ergebnisse

Insgesamt wurde kein signifikanter Haupteffekt von Erfahrung beobachtet. Die Gruppe der BD hob sich also hinsichtlich der über RST und LST ermittelten AG-Kapazität nicht signifikant von den anderen TeilnehmerInnen ab. Dem steht ein signifikanter Haupteffekt von Modalität gegenüber, der auf ein allgemein besseres Abschneiden in der akustischen Variante des Tests zurückgeführt werden kann. Zunächst werden hier deskriptivstatistische Aspekte dargestellt. Anschließend sollen die Erkenntnisse aus den Signifikanztests ($\alpha = 0.05$) anhand der oben aufgelisteten Forschungsfragen erläutert werden.

4.4.1. Deskriptivstatistische Auswertung

Wie aus Tabelle 2 (s. auch Anhang F: Statistische Datenauswertung) und Abbildung 1 hervorgeht, schnitten die BD am mittleren Gesamtergebnis gemessen ($M = 0.91$, $SD = 0.08$) besser ab als die DS ($M = 0.90$, $SD = 0.05$, $\Delta = 0.01$) und ND ($M = 0.86$, $SD = 0.07$, $\Delta = 0.05$). Vergleicht man die Gesamtmittelwerte des RST ($M = 0.81$) und LST ($M = 0.90$) über alle drei Gruppen hinweg, so fällt auf, dass beim Hörspannentest durchschnittlich höhere Ergebnisse erzielt wurden ($\Delta = 0.04$).

Betrachtet man die Ergebnisse im Detail, indem man die verschiedenen Versuchsbedingungen kreuzt, so entsteht ein differenzierteres Bild. Eine Aufschlüsselung nach Modalität ergab beim Lesespannentest einen Vorteil der BD ($M = 0.91$) gegenüber den DS ($M = 0.87$, $\Delta = 0.04$) und insbesondere gegenüber den ND ($M = 0.81$, $\Delta = 0.10$). Am stärksten streuten beim RST die Ergebnisse bei den ND ($SD = 0.10$), gefolgt von den BD ($SD = 0.08$) und den DS ($SD = 0.06$). Darüber hinaus ist anzumerken, dass drei Personen der BD-Gruppe im RST fehlerfreie Leistungen erbrachten (BD insgesamt: $min = 0.74$, $max = 1.00$, $R = 0.26$). In den anderen zwei Gruppen gelang dies niemandem (DS: $min = 0.73$, $max = 0.95$, $R = 0.21$; ND: $min. = 0.62$, $max. = 0.98$, $R = 0.37$).

Tabelle 2

RST- und LST-Ergebnisse nach Gruppen

Versuchsgruppe (<i>N</i> = 54)	RST		LST		Gesamt	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
BD (<i>n</i> = 18)	0.91	0.08	0.90	0.10	0.91	0.08
DS (<i>n</i> = 18)	0.87	0.06	0.92	0.05	0.90	0.05
ND (<i>n</i> = 18)	0.81	0.10	0.90	0.06	0.86	0.07
Gesamt	0.87	0.09	0.91	0.07	0.89	0.07

Anmerkungen. BD = BerufsdolmetscherInnen. DS = Dolmetschstudierende. ND = Nicht-Dolmetschstudierende. RST = Reading Span Task. LST = Listening Span Task.

Der Hörspantentest hingegen lieferte weitaus weniger klare Unterschiede. Hier lagen die DS vorne ($M = 0.92$), während die BD ($M = 0.90$) und ND ($M = 0.90$) ungefähr auf gleicher Höhe abschnitten. Tendenziell waren sogar leichte Vorteile der ND festzustellen ($\Delta = 0.01$). Wie beim RST war die größte Streuung bei den BD ($SD = 0.10$) und den ND ($SD = 0.06$) zu beobachten. Weniger hoch fiel dieser Wert bei den DS ($SD = 0.05$) aus. Dass die TeilnehmerInnen beim LST insgesamt besser abschnitten, ist daran zu erkennen, dass vier VertreterInnen der BD-Gruppe (BD insgesamt: $min. = 0.73$, $max. = 1.00$, $R = 0.27$) und eine VertreterIn der ND-Gruppe (ND insgesamt: $min. = 0.76$, $max. = 1.00$, $R = 0.24$) die höchstmögliche Punkteanzahl erreichten. Die DS wiesen dafür die niedrigste Variationsbreite auf ($min. = 0.79$, $max. = 0.98$, $R = 0.19$).

Auffallend ist ferner, dass die Differenz der Gruppenmittelwerte zwischen RST und LST bei den BD am niedrigsten ausfiel ($\Delta = 0.01$). Im Gegensatz zu den anderen Gruppen ist bei ihnen ein Vorteil in der visuellen Modalität zu beobachten. Die DS und ND schnitten beim Hörspantentest besser ab als beim Lesespantentest (DS: $\Delta = 0.05$; ND: $\Delta = 0.09$).

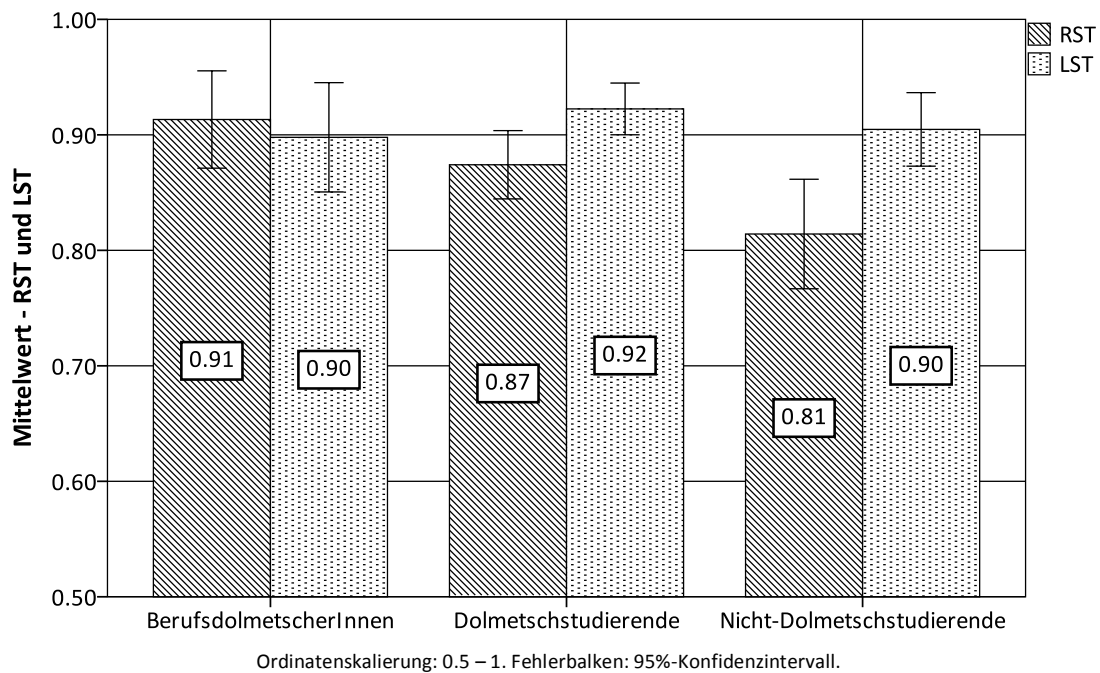


Abbildung 3

Mittlere RST- und LST-Scores nach Gruppen

4.4.2. Inferenzstatistische Auswertung

Die Ergebnisse der Signifikanztests sollen an dieser Stelle beschrieben werden. Eine tabellarische Übersicht ist im Anhang zu finden (s. Anhang F: Statistische Datenauswertung).

a. Haupteffekt von Erfahrung

Die Analyse der Daten ergab keinen signifikanten Haupteffekt der unabhängigen Variablen „Erfahrung“ ($F [2,51] = 2.59, p = .09$). Der kombinierten Betrachtung der RST- und LST-Mittelwerte nach Gruppe waren also keine Hinweise zu entnehmen, die auf eine größere AG-Kapazität mit zunehmender Erfahrung im Simultandolmetschen rück schließen lassen würden. Lediglich eine Tendenz war zu erkennen, die leichte Vorteile der BD anzudeuten scheint, jedoch nicht überzufällig war.

b. *Einfache Haupteffekte von Erfahrung*

Wie schon die Mittelwertsdifferenzen bei der deskriptivstatistischen Analyse vermuten ließen, kamen bei der getrennten Betrachtung der beiden Faktorstufen der Variablen „Modalität“ wesentlich aufschlussreichere Ergebnismuster zum Vorschein.

Beim LST konnten zwar keine signifikanten Gruppenunterschiede festgestellt werden und der p -Wert blieb sogar weit hinter jenem des Haupteffekts von Erfahrung zurück ($F [2,51] = 0.57, p = .57$). Bei den RST-Ergebnissen traten dafür hochsignifikante Unterschiede auf ($F [2,51] = 6.80, p < 0.01$). Ein Post-Hoc-Test nach Bonferroni zeigte, dass BD signifikant besser abschnitt als ND ($p < 0.01$). Alle anderen Differenzen waren nicht signifikant. Die Überlegenheit der BD gegenüber den ND verursachte also das hochsignifikante Ergebnis bei der Varianzanalyse.

c. *Haupteffekt von Modalität*

Die zweifaktorielle Varianzanalyse ergab einen hochsignifikanten Haupteffekt von Modalität ($F [1,51] = 14.69, p < .001$). Der Gesamtmittelwert des LST lag somit hochsignifikant über jenem des RST.

d. *Einfache Haupteffekte von Modalität*

Der signifikante Haupteffekt von Modalität konnte durch einzelne t -Tests, bei denen die modalitätsabhängigen Mittelwertsunterschiede isoliert nach Gruppe überprüft wurden, näher untersucht werden. Dabei zeigte sich, dass die DS ($t [17] = 4.06, p \leq .001$) und die ND ($t [17] = 4.10, p \leq .001$) beim Hörspannentest einen signifikant besseren Mittelwert hatten als beim Lesespannentest. Der Unterschied zwischen RST- und LST-Mittelwert war bei den BD nicht signifikant ($t [17] = 4.10, p \leq .001$).

e. *Wechselwirkung zwischen Erfahrung und Modalität*

Die Interaktion zwischen Erfahrung und Modalität war signifikant ($F [2,51] = 8.23, p \leq .001$). Diese Wechselwirkung ist in Abbildung 4 illustriert. Bei einer Nullinteraktion

würden die Kurven parallel verlaufen. Zurückzuführen ist die Interaktion auf das signifikant bessere Abschneiden der BD gegenüber den ND beim RST sowie auf die signifikanten einfachen Haupteffekte von Modalität.

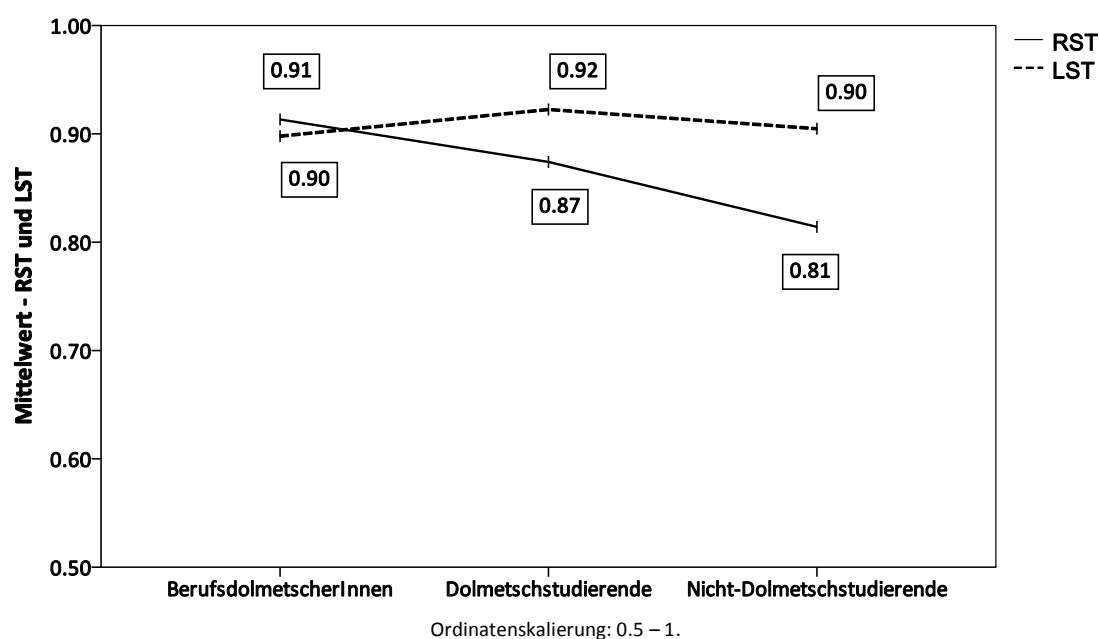


Abbildung 4

Interaktion zwischen Erfahrung und Modalität

4.5. Diskussion

In Summe spiegelt sich in den Ergebnissen dieses Experiments die bei der Analyse der dolmetschwissenschaftlichen RST/LST-Literatur vorgefundene modalitätsspezifische Dichotomie.

Wie schon bei Liu (2001) und Nordet und Voegtlin (1998, zit. n. Köpke & Signorelli 2011: 9) waren beim Hörspannentest keine signifikanten Unterschiede feststellbar. Wollte man den geringfügigen Differenzen Bedeutung beimessen, so ließen sie sich am ehesten im Sinne von Köpke und Nespoulous (2006) interpretieren. Denn es sind leichte Vorteile der DS gegenüber den anderen zwei Gruppen zu erkennen, was wohl damit im Zusammenhang stehen könnte, dass sich DS als Neulinge beim Simultandolmetschen ständig in einem kognitiven Ausnahmezustand befinden und ihre AG-Reserven regelmäßig auf das Äußerste beanspruchen, um den komplexen Anforderun-

gen dieser Tätigkeit gerecht zu werden. Dadurch könnte ein Trainingseffekt entstehen, der den DS bei komplexen Gedächtnisspannungsaufgaben einen Vorteil verschaffen würde. Laut Köpke und Nespoulous (2006) ist diese enge Beziehung zwischen Simultandolmetschen und gesteigerter AG-Kapazität allerdings nur auf die Einstiegsphase beschränkt. Bei BD ist das Simultandolmetschen bereits eine Expertenleistung. Dies würde bedeuten, dass sie Automatismen und Strategien entwickelt haben, die das AG entlasten. Da die kognitiven Ressourcen also nicht mehr ständig bis an ihre Grenzen gefordert sind, würden sie wieder auf ein normales Niveau zurückkehren. Die Ergebnisse dieser Untersuchung scheinen diese Hypothese zu bestätigen: BD und ND haben tatsächlich einen fast identischen Mittelwert beim LST.

Trotzdem muss betont werden, dass die AG-Leistung der DS beim LST keineswegs statistisch signifikant besser ist als die der anderen Gruppen. Plausibler ist daher möglicherweise eine Interpretation, wie sie von Liu (2001) vertreten wird. Ihrer Ansicht nach besteht keine Beziehung zwischen AG-Kapazität und Erfahrung im Simultandolmetschen. Mehr Erfahrung sei nur mit mehr domänenspezifischer Expertise in Verbindung zu bringen. DolmetscherInnen seien also keine TrägerInnen eines bestimmten AG-relevanten Merkmals. Bei Maßen der allgemeinen AG-Kapazität ist Liu (2001) zufolge keine Überlegenheit von DolmetscherInnen zu erwarten. Die LST-Resultate der vorliegenden Studie sind mit dieser Auffassung durchaus vereinbar.

Mit den Erkenntnissen zahlreicher KognitionspsychologInnen (vgl. z.B. Goolkasian & Foos 2002; Murdock 1968) übereinstimmend ist das Auftreten eines allgemeinen Modalitätseffekts zugunsten akustisch präsentierten Materials. Unter BeobachterInnen dieses Phänomens herrscht die Ansicht vor, dass akustische Reize im sensorischen Speicher generell langsamer zerfallen als visuelle Stimuli (Goolkasian & Foos 2002: 1096). Die Ergebnisse des LST sind daher erwartungsgemäß signifikant höher als in der visuellen Modalität.

Richtet man das Augenmerk nun auf die Resultate des RST, so ist festzustellen, dass hier wiederum starke Parallelen zu jenen dolmetschwissenschaftlichen Untersuchungen vorzufinden sind, die in der gleichen Modalität durchgeführt wurden. Der generelle Eindruck eines signifikanten AG-Vorteils in Abhängigkeit von Simultandolmetscherfahrung wird bestätigt. Die hier ermittelten RST-Scores scheinen die klassische Annahme eines AG-Kapazitätsvorteils durch Training und Erfahrung oder auf Grund

natürlicher Veranlagung zu belegen. Größere Erfahrung im Simultandolmetschen geht in der vorliegenden RST-Studie mit einer größeren AG-Kapazität einher. Der Unterschied ist allerdings nur zwischen BD und ND signifikant.

Insgesamt erlaubt es also der Faktor Modalität, ein Muster in den bisherigen RST/LST-Studien mit DolmetscherInnen aufzuspüren, das im Rahmen dieser Arbeit zusätzlich empirisch untermauert wurde. Selbstverständlich wirft diese Aufschlüsselung nach Modalität zahlreiche andere Fragen auf. Schließlich ist bei Daneman und Carpenter (1980) sowie Conway *et al.* (2005) von einer stark positiven Korrelation zwischen allen komplexen Gedächtnisspannungsaufgaben unabhängig von Modalität die Rede. Es wäre also nicht ein so deutlicher Kontrast zwischen den RST- und den LST-Ergebnissen zu erwarten gewesen.

Es erscheint die Vermutung legitim, dass BD beim RST auf eine besondere Fähigkeit zurückgreifen können, die ihnen beim LST nicht zur Verfügung steht. Möglicherweise sind sie beim Lesespannentest in der Lage, einen bestimmten Aspekt ihres Expertenwissens gewinnbringend einzusetzen. Generelle Vorteile der BD hinsichtlich Sprachkompetenz sind wohl kaum als ausschlaggebend zu betrachten, da alle TeilnehmerInnen in ihrer L1 arbeiteten. Möchte man dem expertise-spezifischen Erklärungsansatz treu bleiben, so müssen die Ergebnisse der vorliegenden Untersuchung mit einer besser ausgeprägten visuellen Textverarbeitung bei DolmetscherInnen in Verbindung gebracht werden. In diesem Fall wäre der RST kein reines Maß des kognitiven Leistungsvermögens, da die BD und in geringerem Ausmaß die DS bestimmte durch Erfahrung erworbene Strategien einsetzen, um die Aufgabe effizienter zu bewältigen. Diese besondere Kompetenz könnte damit im Zusammenhang stehen, dass es DolmetscherInnen von ihrem Alltag gewohnt sind, vor Einsätzen große Mengen an schriftsprachlichen Informationen zu verarbeiten. Vermutlich gelingt es ihnen, diese Expertise auf den RST zu übertragen. Wie auch aus Shlesingers (2000b: 7) Überlegungen zur Erforschung der kognitiven Grundprozesse des Dolmetschens hervorgeht, wäre dieses Verfahren also wenig geeignet, die kognitiven Grundmerkmale von DolmetscherInnen unabhängig von domänenspezifischen Automatismen und Strategien zu erforschen.

In Anbetracht der inhärent akustischen Natur ihres Berufs wäre zu erwarten, dass DolmetscherInnen gesprochenes Textmaterial ebenfalls effizienter verarbeiten. Doch die Ergebnisse zeigen, dass sich ihr kognitives Leistungsniveau beim LST nicht

von den anderen Gruppen unterscheidet. Gemäß der soeben dargelegten Interpretation darf daher angenommen werden, dass beim Hörspantentest keine Expertenstrategien aus dem Berufsalltag der DolmetscherInnen angewendet werden. Ein möglicher Grund ist darin zu sehen, dass das Abspielen der Stimuli beim LST für alle ProbandInnen die gleiche Zeit in Anspruch nimmt, während beim Lesespantentest das ganze Satzitem auf einmal präsentiert wird und daher verschieden schnell gelesen bzw. verarbeitet werden kann.

Von Bedeutung ist allgemein die Tatsache, dass beim RST wie auch beim LST eine Textverarbeitungsleistung zu erbringen ist. Da nun DolmetscherInnen zweifellos als ExpertInnen der Textverarbeitung zu betrachten sind, so ist grundsätzlich davon auszugehen, dass sich ihr Expertenwissen auf die Ergebnisse auswirken könnte. Sinnvoller für die saubere Erkundung AG-relevanter Aspekte im Zusammenhang mit DolmetscherInnen wären daher vielleicht komplexe AG-Aufgaben, bei denen diese Kompetenz nicht zum Tragen kommen kann, wie beispielsweise beim *counting span task* oder beim *operation span task*. Hier müssen Objekte in verschiedenen visuellen Anordnungen gezählt werden bzw. arithmetische Operationen durchgeführt werden.

Darüber hinaus sollte die Möglichkeit eines Alterseffekts im Rahmen dieser Studie nicht verschwiegen werden, insbesondere angesichts der Tatsache, dass das mittlere Alter der BD fast doppelt so hoch wie das der anderen zwei Gruppen war. Die Hypothese einer AG-Leistungsabnahme mit zunehmendem Alter besitzt in der kognitionswissenschaftlichen Forschung eine breite empirische Grundlage (s. z.B. Just & Carpenter 1992: 129f.; MacDonald & Christiansen 2002: 47f.; Meguro *et al.* 2000). Dieses Phänomen wird verschiedentlich durch einen AG-Kapazitätsschwund oder eine allgemeine Verlangsamung der Verarbeitungsprozesse erklärt. Sicherlich wäre die Verwendung einer zusätzlichen, altersangepassten Kontrollgruppe im Rahmen dieser Arbeit hilfreich gewesen, um diesen Faktor zu kontrollieren. Allerdings konnten Signorelli *et al.* (2011) in einer jüngst publizierten RST-Studie keine altersrelevanten Unterschiede zwischen DolmetscherInnen und Nicht-DolmetscherInnen feststellen. Es kann daher als unwahrscheinlich bezeichnet werden, dass die vorliegenden Ergebnisse mit einem Alterseffekt im Zusammenhang stehen.

Außerdem scheint bei näherer Betrachtung der Ergebnisse das Vorliegen eines Deckeneffekts (*ceiling effect*) nicht gänzlich auszuschließen zu sein. Die TeilnehmerIn-

nen schneiden insgesamt sehr hoch ab. Der prozentuelle Anteil der richtig reproduzierten Endwörter bewegt sich zwischen 81 und 92 %. Hinzu kommt, dass die BD in drei Fällen beim RST die volle Punktezahl erreichten und beim LST sogar vier Mal fehlerlose Wiedergabeleistungen schafften. Bei der hier gewählten Experimentanordnung wurde das AG dieser TeilnehmerInnen offensichtlich nicht bis an die Grenzen strapaziert. Trotz der Einschätzung von Conway *et al.* (2005: 773), dass Satzblockgrößen von zwei bis fünf für Versuche mit Studierenden ausreichen würden, wäre in der vorliegenden Studie also wohl die Verwendung einer weiteren Satzebene mit sechs Stimuli angebracht gewesen. Gleichwohl muss hinzugefügt werden, dass eine Scoringmethode (PCU-Methode) gewählt wurde, die das Zustandekommen von hohen Scores begünstigt. Es deutet zuletzt nichts darauf hin, dass sich das Ergebnismuster mit der Einführung eines zusätzlichen Satzniveaus wesentlich in eine andere Richtung als bisher angenommen entwickeln würde. Vielmehr ist davon auszugehen, dass die Überlegenheit der BD beim RST dann noch stärker ausgeprägt sein würde, während beim Hörspannentest der Vorteil der DS möglicherweise Signifikanz erreichen würde.

Dass manchen TeilnehmerInnen die fehlerlose Wiedergabe aller Endwörter gelang, könnte auch auf ein weiteres Problem dieses experimentellen Verfahrens hinweisen, das bereits von Daneman und Carpenter (1980: 457) thematisiert wurde und auch in der späteren Forschungsliteratur aufgegriffen wird (s. z.B. Hacker *et al.* 1994: 314f.). Es ist nämlich vorstellbar, dass es manche ProbandInnen schaffen, eine bewusst gewählte und an diese experimentelle Methode angepasste Vorgehensweise einzusetzen, sodass sie deutlich besser abschneiden, aber sich nicht stark genug abheben, um sie als Ausreißer zu bewerten. Die individuelle Wahl einer bestimmten Strategie oder Strategienkombination könnte also zu gesteigerten Leistungen führen. Nicht zu verwechseln ist diese Art der Strategienanwendung allerdings mit dem Einsatz des oben besprochenen Expertenwissens. Bei ersterer ist eine ad hoc zurechtgelegte Vorgangsweise gemeint, bei letzterem handelt es sich um eingeübte Expertenmechanismen, die über viele Jahre hinweg erworben wurden und vermutlich bei der Bewältigung des RST behilflich sind.

Um ad hoc-Strategien aufzudecken, wurden nach Abschluss jeder Sitzung noch informelle Gespräche mit ausgewählten ProbandInnen geführt. Aus den Aussagen der Testpersonen waren zwei konkrete Prozeduren abzuleiten, die auch schon von Daneman und Carpenter (1980) und Hacker *et al.* (1994) beschrieben wurden. Einerseits versuch-

ten die TeilnehmerInnen bewusst, bildhafte Assoziationen zwischen den Endwörtern herzustellen, was einem Chunking-Prozess gleichzusetzen ist. Andererseits bedienten sie sich kontrollierter Rehearsalprozesse, um die Speicherung der Elemente durch Auffrischung der phonologischen Spuren zu erleichtern. Wie sich die Entwicklung einer individuellen Strategie auf RST/LST-Ergebnisse auswirkt, bleibt allerdings weiterhin unklar. Das hier verwendete Verfahren wurde jedenfalls so konzipiert, dass der Strategieinsatz sehr erschwert wird. Vollständig eliminieren lässt er sich vermutlich nicht.

4.6. Schlussfolgerungen

In der dolmetschwissenschaftlichen Forschung herrscht auf Grund der widersprüchlichen empirischen Ergebnisse Uneinigkeit hinsichtlich der Frage der Beziehungen zwischen dem AG und der Erfahrung im Simultandolmetschen. Ein Teil der Studien liefert Indizien, die die Ansicht bekräftigen, dass sich BerufsdolmetscherInnen durch ein außergewöhnliches Leistungsvermögen des AG auszeichnen. Die VertreterInnen dieser Auffassung halten also eine große AG-Kapazität für unerlässlich, um die kognitiv sehr anspruchsvolle Tätigkeit des Simultandolmetschens zu meistern. Bei anderen Untersuchungen wiederum war empirisch kein AG-Vorteil der DolmetscherInnen festzustellen, was die Vermutung nahelegt, dass sich diese Berufsgruppe vielleicht eher durch spezielles Expertenwissen und nicht durch besondere kognitive Merkmale hervorhebt.

Eine Analyse der bisherigen empirischen Arbeiten zu diesem Thema erlaubte es, einen Faktor ausfindig zu machen, der möglicherweise für die Heterogenität der Erklärungsversuche verantwortlich gemacht werden kann. Die Beobachtung, dass eine saubere Trennung der Ergebnisse nach der Modalität des eingesetzten Verfahrens möglich ist, erscheint nämlich nicht unwesentlich. RST-Studien belegen dabei einen AG-Vorteil der BerufsdolmetscherInnen. In LST-Untersuchungen ist hingegen zwischen professionellen DolmetscherInnen und den anderen Versuchsgruppen kein Unterschied zu erkennen. Der bisher vernachlässigte Faktor „Modalität“ versprach daher ein aufschlussreicher Forschungsparameter zu sein. Vor diesem Hintergrund wurde eine Experimentalanordnung gewählt, bei der die AG-Kapazität professioneller DolmetscherInnen sowohl in der visuellen Modalität als auch in der akustischen Modalität ermittelt wurde und wie in

Vorgängerstudien mit einer Gruppe von Dolmetschstudierenden und einer Kontrollgruppe verglichen wurde.

Die im Rahmen dieses Experiments gelieferten Hinweise reichen zusammengekommen letztendlich nicht aus, um eine kausale Beziehung zwischen AG-Kapazität und Erfahrung im Simultandolmetschen als empirisch verifiziert zu betrachten. Zwar schneiden die BerufsdolmetscherInnen beim RST tatsächlich signifikant besser ab als die Nicht-Dolmetschstudierenden, sodass eine Übereinstimmung mit den bisher berichteten Lesespannentests gegeben ist. Allerdings wäre eine solche Interpretation mit den Ergebnissen des LST-Teils unvereinbar. Hier konnten nämlich keine signifikanten Unterschiede festgestellt werden. Es wäre daher plausibler, nicht von einem kognitiven Ressourcenvorteil der BerufsdolmetscherInnen auszugehen, sondern anzunehmen, dass eine mit dem Dolmetschen in Verbindung stehende Expertenfähigkeit beim Lesespannentest zum Einsatz kam, während beim Hörspannentest keine vergleichbare Übertragung von Expertise stattfinden konnte. Möglicherweise entwickelt sich bei DolmetscherInnen eine besondere visuelle Textverarbeitungskompetenz, die ihnen beim Lesespannentest einen Vorteil verschafft.

Um weitere Fortschritte in der Erforschung der kognitiven Grundlagen des Simultandolmetschens zu erzielen, wären vor dem Hintergrund der hier gewonnenen Erkenntnisse experimentelle Verfahren aussichtsreich, bei denen der Störfaktor Expertise besser kontrolliert ist. Da DolmetscherInnen im Allgemeinen als TextverarbeitungsexpertInnen zu betrachten sind, wären also AG-Experimente empfehlenswert, die ohne Textverarbeitungskomponente operieren – beispielsweise der *operation span task* oder der *counting span task*. Erst dann ließe sich ein AG-Vorteil stichhaltig argumentieren.

5. Zusammenfassung

Die vorliegende Masterarbeit stellt einen Beitrag zur AG-Forschung aus dolmetschwissenschaftlicher Perspektive dar. Dieser interdisziplinäre Forschungsbereich ist von kognitionspsychologischen Überlegungen, Modellen und Verfahren stark geprägt. Besondere Bedeutung kommt an dieser Schnittstelle wissenschaftlicher Bestrebungen dem von Daneman und Carpenter (1980) vorgestellten Versuchsparadigma, dem *reading span task* bzw. *listening span task*, zu. Dieses Verfahren ist eng mit der Entstehung des AG-Begriffs verbunden und kann der vorherrschenden Ansicht in der Kognitionspsychologie zufolge als Maß der AG-Kapazität verstanden werden.

VertreterInnen des kognitionspsychologischen Zweigs der Dolmetschwissenschaft schrieben das Gelingen der kognitiv äußerst anspruchsvollen Tätigkeit des Simultandolmetschens wiederholt einem großen oder besonders effizienten AG zu. Um diese Hypothese zu überprüfen, bot sich das experimentelle Verfahren von Daneman und Carpenter (1980) an. Ein einhelliges Urteil ist den RST/LST-Studien mit DolmetscherInnen in dieser Frage nicht zu entnehmen. Es gelang in der vorliegenden Arbeit allerdings, einen Faktor ausfindig zu machen, der angesichts der widersprüchlichen Ergebnisse möglicherweise Klarheit schaffen könnte. Bei der Betrachtung bisheriger Untersuchungen fällt nämlich auf, dass alle Tests, die einen AG-Vorteil der erfahrenen DolmetscherInnen vorfinden, in der visuellen Modalität (RST) durchgeführt wurden. Im Rahmen von Studien, die in der akustischen Modalität (LST) abgewickelt wurden, konnte diese AG-Überlegenheit aber nie bestätigt werden. Dieses modalitätsspezifische Muster galt es in der vorliegenden Arbeit mit Hilfe eines gekoppelten RST/LST-Versuchs, der als Zwischensubjektsfaktor „Erfahrung im Simultandolmetschen“ und als Innersubjektsfaktor „Modalität“ hatte, zu testen. Die in der Literatur festgestellte Dichotomie in diesem Verfahren konnte tatsächlich reproduziert werden. Beim RST lagen die BerufsdolmetscherInnen signifikant vor Nicht-DolmetscherInnen, während beim LST keine signifikanten Unterschiede zu Tage traten.

Auf dieser empirischen Grundlage scheint es daher zulässig, einen Zusammenhang zwischen dem visuellen RST und Erfahrung im Dolmetschen zu postulieren. Angesichts der Ergebnisse des LST muss es jedoch als unwahrscheinlich gelten, dass sich professionelle DolmetscherInnen grundsätzlich durch eine bessere kognitive Verarbei-

tungskapazität auszeichnen. Eher nachvollziehbar wäre die Erklärung, dass die Ursache für das bessere Abschneiden beim RST im Expertenwissen dieser Personengruppe in Form eines Vorteils in der visuellen Textverarbeitung liegt. Beim LST kommt den BerufsdolmetscherInnen ihre Expertise dieser Interpretation zufolge weniger zugute.

Aus dieser Abhandlung lässt sich somit eine umfassende Schlussfolgerung und eine damit verbundene Forderung an zukünftige AG-Experimente mit DolmetscherInnen ableiten. Das RST-Verfahren scheint kein sauberes Messinstrument zur Untersuchung AG-relevanter Phänomene bei BerufsdolmetscherInnen zu sein, da Indizien vorliegen, dass die bei dieser Gruppe besonders ausgeprägte visuelle Textverarbeitungsexpertise zu irreführenden Ergebnissen führt, die einen möglicherweise gar nicht vorhandenen AG-Kapazitätsvorteil suggerieren. Allgemein wäre es daher wünschenswert, in Zukunft bei der Untersuchung des AG im Zusammenhang mit dem Dolmetschen von experimentellen Methoden abzukommen, die mit Textverarbeitungsaufgaben operieren.

Bibliographie

- Atkinson, R. C. & R. M. Shiffrin (1968) Human memory: a proposed system and its control processes. In K. W. Spence & J. T. Spence (Hgg.) *The Psychology of Learning and Motivation*. Bd. 2. New York: Academic Press, 89-195.
- Baddeley, A. D. (2000a) The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences* 4 (11), 417-423.
- Baddeley, A. D. (2000b) Working memory and language processing. Interdisciplinary perspectives. In B. Englund Dimitrova & K. Hyltenstam (Hgg.) *Language Processing and Simultaneous Interpreting* (Benjamins Translation Library 40). Amsterdam/Philadelphia: John Benjamins, 1-16.
- Baddeley, A. D. ([2001]/2002) Is working memory still working? *European Psychologist* 7 (2), 85-97.
- Baddeley, A. D. (2003) Working memory: looking back and looking forward. *Nature Reviews Neuroscience* 4, 829-839.
- Baddeley, A. D. & G. J. Hitch (1974) Working memory. In G. H. Bower (Hg.) *The Psychology of Learning and Motivation*. Bd. 8. New York: Academic Press, 47-89.
- Baddeley, A. D. & R. H. Logie (1999) Working memory: the multiple-component model. In A. Miyake & P. Shah (Hgg.) *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Cambridge: Cambridge University Press, 28-61.
- Bajo, M. T., F. Padilla & P. Padilla (2000) Comprehension processes in simultaneous interpreting. In A. Chesterman, N. Gallardo San Salvador & Y. Gambier (Hgg.) *Translation in Context*. Amsterdam: John Benjamins, 127-142.
- Bajo, M. T., P. Padilla, M. Ricardo, F. Padilla, C. Gómez, M. C. Puerta, P. Gonzalvo & P. Macizo (2001) Comprehension and memory processes in translation and interpreting. *Quaderns. Revista de traducció* 6, 27-31.
- Chincotta, D. & G. Underwood (1998) Simultaneous interpreters and the effect of concurrent articulation on immediate memory. *Interpreting* 3 (1), 1-20.
- Christoffels, I. K., A. M. B. De Groot & L. J. Waldorp (2003) Basic skills in a complex task: a graphical model relating memory and lexical retrieval to simultaneous interpreting. *Bilingualism: Language and Cognition* 6, 201-211.
- Christoffels, I. K., A. M. B. De Groot & J. F. Kroll (2006) Memory and language skills in simultaneous interpreters: the role of expertise and language proficiency. *Journal of Memory and Language* 54, 324-345.

- Conway, A. R. A., M. J. Kane, M. F. Bunting, D. Z. Hambrick, O. Wilhelm & R. W. Engle (2005) Working memory span tasks: a methodological review and user's guide. *Psychonomic Bulletin & Review* 12, 769-786.
- Cowan, N. (1988) Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin* 104 (2), 163-191.
- Cowan, N. (1995) *Attention and Memory: An Integrated Framework*. New York: Oxford University Press.
- Cowan, N. (1999) An embedded-processes model of working memory. In A. Miyake & P. Shah (Hgg.) *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control*. Cambridge: Cambridge University Press, 62-101.
- Cowan, N. (2000) Processing limits of selective attention and working memory, potential implications for interpreting. *Interpreting* 5 (2), 117-146.
- Cowan, N., E. M. Elliott, J. S. Sauls, C. C. Morey, S. Mattox, A. Hismjatullina & A. R. A. Conway (2005) On the capacity of attention: its estimation and its role in working memory and cognitive aptitudes. *Cognitive Psychology* 51, 42-100.
- Craik, F. I. M. & R. S. Lockhart (1972) Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, 671-684.
- Daneman, M. & P. A. Carpenter (1980) Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior* 19, 450-466.
- Danks, J. H., G. M. Shreve, S. B. Fountain & M. K. McBeath (Hgg.) (1997) *Cognitive Processes in Translation and Interpreting* (Applied Psychology 3). Thousand Oaks: Sage Publications.
- Darò, V. (1995) Attentional, auditory, and memory indexes as prerequisites for simultaneous interpreting. In J. Tömmola (Hg.) *Topics in Interpreting Research*. Turku: University of Turku, Centre for Translation and Interpreting, 3-10.
- Darò, V. (1997) Experimental studies on memory in conference interpretation. *Meta* 42 (4), 622-628.
- Darò, V. & F. Fabbro (1994) Verbal memory during simultaneous interpretation: effects of phonological interference. *Applied Linguistics* 15, 365-381.
- De Groot, A. M. B. (1997) The cognitive study of translation and interpretation. In J. H. Danks, G. M. Shreve, S. B. Fountain & M. K. McBeath (Hgg.) *Cognitive Processes in Translation and Interpreting*. Thousand Oaks: Sage Publications, 25-56.

- Desmette, D., M. Hupet, M.-A. Schelstraete & V. d. L. Martial (1995) Adaptation en langue française du "Reading Span Test" de Daneman et Carpenter (1980). *L'Année psychologique* 95, 459-482.
- Ericsson, K. A. (2000) Expertise in interpreting: an expert-performance perspective. *Interpreting* 5 (2), 189-222.
- Gathercole, S. E. (2008) Working memory. In H. L. Roediger III. (Hg.) *Cognitive Psychology of Memory* (Learning and Memory: A Comprehensive Reference 2). Oxford/San Diego: Elsevier, 33-51.
- Gerver, D. (1976) Empirical studies of simultaneous interpretation: a review and a model. In R. W. Brislin (Hg.) *Translation: Applications and Research*. New York: Gardner Press, 165-207.
- Gile, D. (1997) Conference interpreting as a cognitive management problem. In J. H. Danks, G. M. Shreve, S. B. Fountain & M. K. McBeath (Hgg.) *Cognitive Processes in Translation and Interpreting*. Thousand Oaks: Sage Publications, 196-214.
- Goolkasian, P. & P. W. Foos (2002) Presentation format and its effect on working memory. *Memory & Cognition* 30 (7), 1096-1105.
- Gruber, T. (2011) *Gedächtnis* (Basiswissen Psychologie). Wiesbaden: VS Verlag.
- Hacker, W., T. Veres & E. Wollenberger (1994) Verarbeitungskapazität für Text: Ergebnisse der Entwicklung eines deutschsprachigen Prüfverfahrens des Arbeitsgedächtnisses. *Zeitschrift für Psychologie* 202, 295-320.
- Hussy, W., M. Schreier & G. Echterhoff (2010) *Forschungsmethoden in Psychologie und Sozialwissenschaften für Bachelor*. Berlin/Heidelberg: Springer Verlag.
- Isham, W. P. (2000) Phonological interference in interpreters of spoken-languages: an issue of storage or process? In B. Englund Dimitrova & K. Hytenstam (Hgg.) *Language Processing and Simultaneous Interpreting* (Benjamins Translation Library 40). Amsterdam/Philadelphia: John Benjamins, 133-149.
- James, W. (1890) *Principles of Psychology*. Bd. 1. New York: Henry Holt.
- Just, M. A. & P. A. Carpenter (1992) A capacity theory of comprehension: individual differences in working memory. *Psychological Review* 99 (1), 122-149.
- Kavé, G., N. Eyal, A. Shorek & J. Cohen-Mansfield (2008) Multilingualism and cognitive state in the oldest old. *Psychology and Aging* 23 (1), 70-78.
- Köpke, B. & J.-L. Nespoulous (2006) Working memory performance in expert and novice interpreters. *Interpreting* 8 (1), 1-23.

- Köpke, B. & T. M. Signorelli (2011) Methodological aspects of working memory assessment in simultaneous interpreters. *International Journal of Bilingualism*, 15. Juni 2011, 1-15. DOI: 10.1177/1367006911402981
- Liu, M. (2001) *Expertise in Simultaneous Interpreting: A Working Memory Analysis*. Dissertation, University of Texas at Austin.
- Liu, M. (2008) How do experts interpret? Implications from research in interpreting studies and cognitive science. In G. Hansen, A. Chesterman & H. Gerzymisch-Arbogast (Hgg.) *Efforts and Models in Interpreting and Translation Research*. Amsterdam/Philadelphia: John Benjamins, 159-177.
- Liu, M., D. L. Schallert & P. J. Carroll (2004) Working memory and expertise in simultaneous interpreting. *Interpreting* 6 (1), 19-42.
- MacDonald, M. C. & M. H. Christiansen (2002) Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review* 109, 35-54.
- MacWhinney, B. (1997) Simultaneous interpretation and the competition model. In J. H. Danks, G. M. Shreve, S. B. Fountain & M. K. McBeath (Hgg.) *Cognitive Processes in Translation and Interpretation*. Thousand Oaks: Sage Publications, 215-232.
- Meguro, Y., T. Fujii, A. Yamadori, T. Tsukiura, K. Suzuki, J. Okuda & M. Osaka (2000) The nature of age-related decline on the Reading Span Task. *Journal of Clinical and Experimental Neuropsychology* 22 (3), 391-398.
- Miller, G. A. ([1956]/1994) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* 101 (2), 343-352.
- Mizuno, A. (2005) Process model for simultaneous interpreting and working memory. *Meta* 50 (2), 739-752.
- Moser-Mercer, B., U. H. Frauenfelder, B. Casado & A. Künzli (2000) Searching to define expertise in interpreting. In B. Englund Dimitrova & K. Hyltenstam (Hgg.) *Language Processing and Simultaneous Interpretation: Interdisciplinary Perspectives*. Amsterdam: John Benjamins, 107-131.
- Murdock, B. B. J. (1968) Modality effects in short-term memory: storage or retrieval? *Journal of Experimental Psychology* 77 (1), 79-86.
- Nordet, L. & L. Voegtlin (1998) *Les tests d'aptitude pour l'interprétation de conférence : la mémoire*. Diplomarbeit, Université de Genève (Ecole de traduction et d'interprétation).

- Padilla, F., M. T. Bajo & P. Macizo (2005) Articulatory suppression in language interpretation: working memory capacity, dual tasking and word knowledge. *Bilingualism: Language and Cognition* 8 (3), 207-219.
- Padilla, P., M. T. Bajo, J. J. Cañas & F. Padilla (1995) Cognitive processes of memory in simultaneous interpretation. In J. Tömmola (Hg.) *Topics in Interpreting Research*. Turku: University of Turku, Centre for Translation and Interpreting, 61-71.
- Service, E., M. Simola, O. Metsänheimo & S. Maury (2002) Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology* 14 (3), 383-408.
- Shlesinger, M. (2000a) *Strategic Allocation of Working Memory and Other Attentional Resources*. Dissertation, Bar-Ilan University.
- Shlesinger, M. (2000b) Interpreting as a cognitive process: how can we know what really happens? In S. Tirkkonen-Condit & R. Jääskeläinen (Hgg.) *Tapping and Mapping the Processes of Translation and Interpreting: Outlooks on Empirical Research*. Amsterdam: John Benjamins, 3-15.
- Shlesinger, M. (2003) Effects of presentation rate on working memory in simultaneous interpreting. *The Interpreters' Newsletter* 12, 37-49.
- Signorelli, T. M. (2008) *Working Memory in Simultaneous Interpreters*. Dissertation, City University of New York.
- Signorelli, T. M. & H. J. Haarmann (2005) Are there working memory differences between simultaneous interpreters and non-interpreter multilinguals? *2005 Proceedings of the Monterey Journal of Translation and Interpreting Studies*.
- Signorelli, T. M., H. J. Haarmann & L. K. Obler (2011) Working memory in simultaneous interpreters: effects of task and age. *International Journal of Bilingualism*, 7. Juni 2011, 1-15. DOI: 10.1177/1367006911403200
- Timarová, Š. (2008) Working memory and simultaneous interpreting. In P. Boulogne (Hg.) *Translation and its Others: Selected papers of the CETRA Research Seminar in Translation Studies 2007*. Online: <http://www2.arts.kuleuven.be/info/bestanden-div/Timarova.pdf> (11.03.2012)
- Tzou, Y.-Z., Z. R. Eslami, H.-C. Chen & J. Vaid (2011) Effect of language proficiency and degree of formal training in simultaneous interpreting on working memory and interpreting performance: evidence from Mandarin-English speakers. *International Journal of Bilingualism*, 14. Juni 2011, 1-15. DOI: 10.1177/1367006911403197

- Van den Noort, M., P. Bosch, M. Haverkort & K. Hugdahl (2008) A standard computerized version of the reading span test in different languages. *European Journal of Psychological Assessment* 24, 35-42.
- Viezzi, M. (1990) Sight translation, simultaneous interpretation and information retention. In L. Gran & C. Taylor (Hgg.) *Aspects of Applied and Experimental Research of Conference Interpretation*. Udine: Campanotto, 54-60.

Anhang

Anhang A: Instruktionen für TeilnehmerInnen

Anhang B: Stimulussätze des Reading Span Task

Anhang C: Stimulussätze des Listening Span Task

Anhang D: Protokolle

Anhang E: Screenshots

Anhang F: Statistische Datenauswertung

Anhang A: Instruktionen für TeilnehmerInnen

Instruktionen für den RST

(Instruktion 1)

In diesem Teil werden einzelne, unzusammenhängende Sätze kurz am Bildschirm präsentiert.

Ihre Aufgaben:

- Lesen: Lesen Sie den Satz aufmerksam durch.
- Beurteilen: Während der Satz präsentiert wird, mittels Mausklick (links/rechts) entscheiden, ob der Satz stilistisch (Satzbau, Ausdruck) gut oder schlecht ist (für die Auswertung wichtig).
- Speichern: Merken Sie sich das letzte Wort (= Endwort) dieses Satzes.

Nach jedem Block (2, 3, 4 oder 5 Sätze) erscheint ein Eingabefeld für die Wiedergabe der Endwörter des Blocks. Die Reihenfolge der Wörter ist unwesentlich.

Drücken Sie bitte die LEERTASTE (Fortsetzung der Instruktion)...

(Instruktion 2)

Es wird geraten, die Sätze zügig zu lesen, aber dennoch so, dass Sie den Inhalt verstehen und den Satz stilistisch beurteilen können. Parallel dazu sollen die Endwörter gespeichert werden.

Vergessen Sie nicht: Die Satzbeurteilung mittels Mausklick muss rasch erfolgen, da die Sätze knapp aufeinander folgen.

Es folgen nun zunächst zwei Probeblöcke.

Drücken Sie bitte die LEERTASTE...

Instruktionen für den LST

(Instruktion 1)

In diesem Teil werden Ihnen einzelne, unzusammenhängende Sätze von einer Stimme vorgelesen.

Ihre Aufgaben bleiben gleich:

- Zuhören: Hören Sie aufmerksam zu.
- Beurteilen: Während der Satz vorgelesen wird, mittels Mausklick (links/rechts) entscheiden, ob der Satz stilistisch (Satzbau, Ausdruck) gut oder schlecht ist (für die Auswertung wichtig).
- Speichern: Merken Sie sich das letzte Wort (Endwort) dieses Satzes.

Wie im vorangegangenen Teil erscheint nach jedem Block ein Eingabefeld für die Wiedergabe der Endwörter des Blocks. Die Reihenfolge der Wörter ist unwesentlich.

Drücken Sie bitte die LEERTASTE (Fortsetzung der Instruktion)...

(Instruktion 2)

Es wird geraten, sich auf den Inhalt des Satzes zu konzentrieren um ihn stilistisch besser beurteilen können. Parallel dazu sollen die Endwörter gespeichert werden.

Vergessen Sie nicht, die Sätze mittels Mausklick stilistisch zu beurteilen!

Es folgen nun zunächst zwei Probeblöcke.

Bitte setzen Sie die Kopfhörer auf und drücken Sie danach die LEERTASTE...

Anhang B: Stimulussätze des Reading Span Task (Auszug)³

Probeserie

2-Satz-Block	
1	Der schnelle Niedergang der Kurse war für den Aktienberater ein großes Problem.
2	Viele Leute liefen mit in der Demonstration um zu protestieren gegen den Krieg.
3-Satz-Block	
1	Der Notar hat mit seinem Kunden einen Termin gemacht für nächste Woche.
2	Es gibt immer viel Feuerwerk zu Neujahr und das bringt Freude.
3	Auf dem Gipfel wurde über vieles gesprochen und es war auch ein Erfolg.

Serie 1

2-Satz-Block	
1	Durch den Sturm war das Schiff nicht mehr zu lenken und es prallte gegen eine Wand.
2	Der Geburtstag war sehr fröhlich und es gab wie immer sehr viel herrlichen Kuchen.
3-Satz-Block	
1	Die Leute fanden die Renovierung störend, aber jetzt gibt es einen schönen Turm.
2	Er ging sofort zu seinem Arzt um etwas zu unternehmen gegen die Schmerzen.
3	Das junge Mädchen fuhr stolz mit ihrem neuen Wagen zu ihrer neuen Liebe.
4-Satz-Block	
1	Das letzte Spiel des Jahres musste abgesagt werden wegen eines Gewitters.
2	Um drei Uhr stürmte die Polizei unter Einsatz von viel Gewalt das Gebäude.
3	Eine Bemerkung von unserem Chef wurde der Anfang einer großen Debatte.
4	Das kleine Kind hatte eine Ohrinfektion und lag deswegen drei Wochen im Bett.
5-Satz-Block	
1	Als der Zug in den Bahnhof fuhr, rannte mein guter Freund aus der gemütlichen Bar.
2	Die Hochzeit war sehr einzigartig, und der Höhepunkt war das Fest auf dem Schloss.
3	Im Sommer haben wir herrlich im Garten gegessen, trotz der vielen Fliegen.
4	Es schmeckt mir nicht, aber das Wichtigste ist das Abendessen nach der Konferenz.
5	Sein schneller Durchbruch in der Wissenschaft wurde zum Teil verursacht durch sein Buch.

³ Aus Urheberrechtsgründen werden die Stimulussätze hier nur auszugsweise veröffentlicht. Das Material wurde dankenswerterweise von Prof. Dr. Maurits van den Noort zur Verfügung gestellt, an den man sich für Nutzungsbewilligungen und weitere Informationen wende:

Prof. Dr. Maurits van den Noort

Adresse: Kyung Hee University, Seoul, Republic of Korea & Free University of Brussels, Pleinlaan 2,
1050 Brussels, Belgium

E-Mail: info@mauritsvandennoort.com

Anhang C: Stimulussätze des Listening Span Task (Auszug)

Probeserie

2-Satz-Block	
1	Nach dem Fall der Mauer gab es in der Stadt Berlin eine einzigartige Feier.
2	Als ich gestern am Strand war, sah ich einen schönen Sonnenuntergang im Westen.

3-Satz-Block	
1	Die Schüler langweilten sich und fingen an Witze zu machen mit dem Lehrer.
2	Heute Nachmittag kann man nicht umsonst parken, denn hier ist ein sehr großes Fest.
3	Gestern war es ein schöner Tag, weil ich eine Reise gemacht habe mit meinem Sohn.

Serie 1

2-Satz-Block	
1	Der Termin war zu Ende, und der Minister dankte seinem großen Publikum.
2	Für die Kinder war der Klassenausflug zum Flughafen ein großes Abenteuer.

3-Satz-Block	
1	Als die junge Frau das Strafverfahren gewonnen hatte, dankte sie dem Richter.
2	Der Politiker drohte nach dem großen Streit in der Partei mit seinem Rückzug.
3	Der Mitarbeiter war wütend, als er informiert wurde über den schweren Unfall.

4-Satz-Block	
1	Das Mädchen stand zufrieden auf und dachte, was ist mein Freund manchmal für ein Schatz.
2	Diese Pferde hatten damals oft gewonnen, deswegen gab es viele Käufer.
3	Viele arme Leute aus dem Stadtviertel lebten von dem Handel mit Tabak.
4	Weil es heute Nachmittag so unglaublich warm war, waren die meisten Leute am Strand.

5-Satz-Block	
1	Mein Freund war sehr müde, denn er hatte hart gearbeitet in seinem neuen Garten.
2	Sie bekam Geld von ihrer Mutter, weil sie so gut geholfen hatte mit dem Umzug.
3	Die junge Geschäftsfrau hatte ihr Vermögen verdient mit dem Handel von Kunst.
4	Der Junge fuhr mit dem Fahrrad in die Stadt und war noch rechtzeitig bei seiner Freundin.
5	Wegen des starken Westwinds war Schluss mit dem Segeln auf dem See.

Anhang D: Protokolle

Protokoll für BerufsdolmetscherInnen

B1-Protokoll

Vom Versuchsleiter auszufüllen:	
Sq/V:	Sq1/V2
Vp-Code:	P1
Vp-Gruppe:	B1
Datum:	TT.MM.JJ
Uhrzeit:	
Dauer:	
Ort:	ML1, ZTW
Betriebssyst.:	Windows 7 Enterprise, 64-Bit
Computername:	LAB1STUD[]

Bitte ausfüllen!

Alter:

Geschlecht (m/w):

Erfahrung im Dolmetschen (in Jahren):

Sprachkombination:

Universitätsabschluss (ja/nein + Studienrichtung):

Feedback:

Protokoll für Dolmetschstudierende

B2-Protokoll

Vom Versuchsleiter auszufüllen:

Sq/V:	Sq2/V2
Vp-Code:	P17
Vp-Gruppe:	B2
Datum:	TT.MM.JJ
Uhrzeit:	
Dauer:	
Ort:	ML1, ZTW
Betriebssyst.:	Windows 7 Enterprise, 64-Bit
Computername:	LAB1STUD[]

Bitte ausfüllen!

Alter:

Geschlecht (m/w):

**Erfahrung im Dolmetschen (in Semestern,
ggf. Berufserfahrung in Jahren):**

Sprachkombination:

Feedback:

Protokoll für Nicht-Dolmetschstudierende

B3-Protokoll

Vom Versuchsleiter auszufüllen:

Sq/V:	Sq1/V2
Vp-Code:	P33
Vp-Gruppe:	B3
Datum:	TT.MM.JJ
Uhrzeit:	
Dauer:	
Ort:	ML1, ZTW
Betriebssyst.:	Windows 7 Enterprise, 64-Bit
Computername:	LAB1STUD[]

Bitte ausfüllen!

Alter:

Geschlecht (m/w):

Studienrichtung:

Semester:

Muttersprache:

Fremdsprache(n):

Feedback:

Anhang E: Screenshots

Stimulusdarbietung

RST

Der Notar hat mit seinem Kunden einen Termin gemacht für nächste Woche.

Gut (linke Maustaste) Schlecht (rechte Maustaste)

LST

Gut (linke Maustaste) Schlecht (rechte Maustaste)

Wiedergabephase

Wiedergabe

Geben Sie nun die ENDWÖRTER ein.
(Großbuchstaben sind gesperrt.)
Drücken Sie anschließend die linke Maustaste.

A horizontal orange bar intended for the user to enter the end words.

Wiedergabe

Geben Sie nun die ENDWÖRTER ein.
(Großbuchstaben sind gesperrt.)
Drücken Sie anschließend die linke Maustaste.

woche, freude, erfolg

Anhang F: Statistische Datenauswertung

Tabelle F-1

Demographische Daten der Versuchsgruppen

Versuchsgruppe (<i>N</i> = 54)	Alter					Erfahrung ^{a)}					Geschlecht ^{b)}	
	<i>M</i>	<i>SD</i>	<i>min.</i>	<i>max.</i>	<i>R</i>	<i>M</i>	<i>SD</i>	<i>min.</i>	<i>max.</i>	<i>R</i>	w	m
BD (<i>n</i> = 18)	47.06	7.26	33	60	27	20.89	7.90	6	35	29	16	2
DS (<i>n</i> = 18)	25.44	1.42	23	28	5	2.53	0.50	2	3.5	1.5	18	-
ND (<i>n</i> = 18)	24.17	2.07	21	30	9	-	-	-	-	-	13	5

Anmerkungen. BD = BerufsdolmetscherInnen. DS = Dolmetschstudierende. ND = Nicht-Dolmetschstudierende.

^{a)} Die Erfahrung ist zwecks Vergleichbarkeit bei den DS in Jahren und nicht in Semestern angegeben.

^{b)} w = weiblich, m = männlich.

Tabelle F-2*RST- und LST-Ergebnisse nach Gruppen*

Versuchsgruppe (N = 54)	RST						LST						Gesamt		
	<i>M</i>	<i>SEM</i>	<i>SD</i>	<i>min.</i>	<i>max.</i>	<i>R</i>	<i>M</i>	<i>SEM</i>	<i>SD</i>	<i>min.</i>	<i>max.</i>	<i>R</i>	<i>M</i>	<i>SEM</i>	<i>SD</i>
BD (<i>n</i> = 18)	0.91	0.02	0.08	0.74	1.00	0.26	0.90	0.02	0.10	0.73	1.00	0.27	0.91	0.02	0.08
DS (<i>n</i> = 18)	0.87	0.01	0.06	0.73	0.95	0.21	0.92	0.01	0.05	0.79	0.98	0.19	0.90	0.01	0.05
ND (<i>n</i> = 18)	0.81	0.02	0.10	0.62	0.98	0.37	0.90	0.02	0.06	0.76	1.00	0.24	0.86	0.02	0.07
Gesamt	0.87	0.01	0.09	0.62	1.00	0.39	0.91	0.01	0.07	0.73	1.00	0.27	0.89	0.01	0.07

Anmerkungen. BD = BerufsdolmetscherInnen. DS = Dolmetschstudierende. ND = Nicht-Dolmetschstudierende. RST = Reading Span Task. LST = Listening Span Task.

Tabelle F-3

Ergebnisse der Varianzanalysen: Effekte von Erfahrung und Modalität, Interaktion Erfahrung x Modalität

	Freiheitsgrade	F-Wert	p-Wert
Haupteffekt Erfahrung	2, 51	2.59	.09
Einfacher Haupteffekt Erfahrung in RST	2, 51	6.80*	.002
Einfacher Haupteffekt Erfahrung in LST	2, 51	0.57	.57
Haupteffekt Modalität	1, 51	14.69**	< .001
Interaktion Erfahrung x Modalität	2, 51	8.23**	.001

Anmerkungen. RST = Reading Span Task. LST = Listening Span Task.

* $p \leq .01$.

** $p \leq .001$.

Tabelle F-4

Ergebnisse der t-Tests: Einfache Haupteffekte von Modalität

	Freiheitsgrade	t-Wert	p-Wert
Einfacher Haupteffekt Modalität in BD	17	-0.76	.455
Einfacher Haupteffekt Modalität in DS	17	4.06**	.001
Einfacher Haupteffekt Modalität in ND	17	4.10**	.001

Anmerkungen. BD = BerufsdolmetscherInnen. DS = Dolmetschstudierende. ND = Nicht-Dolmetschstudierende.

* $p \leq .01$.

** $p \leq .001$.

Abstract (Deutsch)

Die vorliegende Masterarbeit ist der Rolle des Arbeitsgedächtnisses (AG) beim Simultandolmetschen gewidmet. In zahlreichen wissenschaftlichen Beiträgen wird es für wahrscheinlich gehalten, dass das Arbeitsgedächtnis professioneller DolmetscherInnen durch Ausbildung und Berufspraxis an Effizienz oder Kapazität gewinnt. Der von Daneman und Carpenter (1980) konzipierte Reading Span Task (RST) bzw. Listening Span Task (LST) ist in diesem Zusammenhang vielfach eingesetzt worden, um den postulierten Arbeitsgedächtnisvorteil dieser Berufsgruppe zu beweisen. Bislang ist es allerdings nicht gelungen, sich mit Hilfe dieses Verfahrens ein klar interpretierbares Bild der Beziehungen zwischen Erfahrung im Simultandolmetschen und kognitiver Leistungsfähigkeit – operationalisiert durch die Arbeitsgedächtniskapazität – zu verschaffen. Wie durch eine Analyse bisheriger Untersuchungen demonstriert wird, ist ein sinnvolles Muster erst dann erkennbar, wenn die Ergebnisse nach Modalität aufgeschlüsselt werden. Dann fällt nämlich auf, dass stets nur beim visuellen RST eine Überlegenheit der BerufsdolmetscherInnen in Erscheinung tritt, während der akustische LST keine signifikanten AG-Vorteile dieser Gruppe liefert.

Es wurde vor dem Hintergrund dieser Überlegungen ein gekoppelter RST/LST-Versuch mit 18 BerufsdolmetscherInnen, 18 Dolmetschstudierenden und 18 Nicht-Dolmetschstudierenden durchgeführt. Mit dieser Versuchsanordnung erhoffte man sich, diesen scheinbaren Modalitätseffekt, der bei der Durchsicht der Literatur so augenfällig war, reproduzieren und anschließend begründen zu können. Die Datenauswertung ergab tatsächlich das erwartete Muster. Beim RST wurde ein signifikanter Vorteil der BerufsdolmetscherInnen gegenüber Nicht-Dolmetschstudierenden festgestellt. Beim LST unterschieden sich die Gruppen nicht signifikant voneinander.

Eine mögliche Erklärung dieser Ergebnisse könnte mit dem Expertenwissen der DolmetscherInnen in Verbindung stehen. Demnach wäre die visuelle Textverarbeitungsexpertise, die eine für diese Personengruppe unerlässliche Fertigkeit bei der Vorbereitung von Dolmetscheinsätzen ist, als Störvariable für den AG-Vorteil im RST verantwortlich zu machen. In Zukunft wären also AG-Experimente empfehlenswert, die ohne Textverarbeitungsaufgaben operieren. Erst dann ließe sich ein AG-Vorteil stichhaltig argumentieren.

Abstract (Englisch)

This thesis explores the role of working memory (WM) in the context of simultaneous interpretation. Scholars in the field of interpretation studies have repeatedly advanced the notion that professional interpreters are able to increase the capacity or efficiency of WM through training and experience. Daneman and Carpenter's (1980) dual task paradigm has been used widely to substantiate the hypothesis of a WM superiority among interpreters. Thus far, studies operating with Daneman and Carpenter's (1980) reading span task (RST) or listening span task (LST) have failed to provide a clear pattern of evidence to elucidate the exact nature of the relationship between simultaneous interpretation experience and cognitive ability as instantiated by WM capacity. An analysis of the research based on this method seems to indicate that modality might be a helpful avenue of examination to disentangle the inconsistent results this work has yielded. Indeed, on closer inspection, it is striking that only experiments using the visual RST have attested a WM advantage of professionals, whereas the results from the auditory LST do not confirm this hypothesis. Accordingly, it might be considered insightful to assume that modality underlies the contradictory empirical findings.

To test the hypothesis that the choice of modality might account for the discordant body of evidence, an experiment was conducted that consisted of both an LST and an RST. 18 professional interpreters, 18 student interpreters, and 18 students from other fields (control group) were solicited for this investigation. It was expected that an effect of modality similar to the one discovered by dissecting the empirical findings in other studies would emerge. As predicted, the RST yielded a significant advantage of the professional group compared to the control group, while no significant differences could be found in the LST.

In considering these results, it might be hypothesized that the expertise of professional interpreters is a determining factor. It is important to emphasize in this context that interpreters are experts at quickly processing printed text, an indispensable skill when preparing for assignments. It is therefore not unlikely that this sort of expertise underlies the WM advantage of interpreters in the RST. In light of this, it might be expedient for future WM research with professional interpreters to avoid experimental designs that operate with text processing components.

Lebenslauf

Adrian Carl HIGH, Bakk.phil.

Adresse Dorfstraße 16, 2465 Höflein
Telefon +43 699 103 183 67
E-Mail adrian.high@gmail.com

Persönliche Daten

Geburtsdatum 11. Januar 1986
Geburtsort Wien
Staatsbürgerschaft Vereinigte Staaten von Amerika

Ausbildung

seit 04/2009 **Zentrum für Translationswissenschaft, Universität Wien**
Masterstudium Dolmetschen
Schwerpunkt: Konferenzdolmetschen
Sprachenkombination: Deutsch, Englisch, Französisch

09/2007 – 06/2008 **Université Nancy 2, Frankreich**
Auslandsstudienjahr

10/2004 – 04/2009 **Zentrum für Translationswissenschaft, Universität Wien**
Bakkalaureatsstudium Übersetzen und Dolmetschen
Sprachenkombination: Deutsch, Englisch, Französisch

seit 10/2004 **Universität Wien**
Diplomstudium Alte Geschichte und Altertumskunde

09/1996 – 06/2004 **Akademisches Gymnasium Wien**
Sprachlicher Schwerpunkt: Französisch
Matura mit ausgezeichnetem Erfolg bestanden

Maturaschwerpunktprüfungen u.a. Französisch, Latein und Altgriechisch

Facheinschlägige Berufserfahrung

- | | |
|----------------|---|
| 01/2012 | Simultandolmetscher beim „Campaigning Summit 2012“ im Studio 44, Wien
Sprachen: Deutsch – Englisch |
| 05/2011 | Simultandolmetscher bei der internationalen Tagung „Daker und Dakismus“ im kleinen Festsaal der Universität Wien
Sprachen: Deutsch – Englisch, Englisch – Deutsch |
| 05/2011 | Simultandolmetscher bei der 6. internationalen CEE-Konferenz des Instituts für Interne Revision Österreich im Hotel Arcotel Wimberger
Sprachen: Deutsch – Englisch, Englisch – Deutsch |

Kenntnisse und Fähigkeiten

- | | |
|-----------------|--|
| Sprachen | Deutsch: Mutter- bzw. Bildungssprache
Englisch: Muttersprache
Französisch: fließend in Wort und Schrift |
| EDV | Adobe Dreamweaver CS5
CorelDraw Graphics Suite X5
IBM SPSS Statistics 20
Microsoft Office 2007
SDL Trados Studio 2009
SDL Multiterm 2009
WordPerfect Office X3 |