

Synthese von Kurz- und Langzeitgedächtnisspanne:
Experimentelle Überprüfung eines mathematischen Modells

Inaugural-Dissertation
zur
Erlangung der Doktorwürde
der
Philosophischen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu Bonn

vorgelegt von

Silke Hamm

aus

Würzburg

Bonn 2001

INHALTSVERZEICHNIS

1 Einleitung	7
2 EIN DESKRIPTIV-MODULARES GEDÄCHTNISMODELL: Das Arbeitsgedächtnis	9
<hr/>	
2.1 Aufbau des Arbeitsgedächtnisses nach Baddeley	10
2.2 Die Zentrale Exekutive	15
2.2.1 Adaptive und/oder voluntäre Strategiewechsel	18
2.2.2 Regulation des Informationsflusses	19
2.2.3 Verbindung zum Langzeitgedächtnis	21
2.3 Der Visuell-Räumliche Notizblock	25
2.4 Die Phonologische Schleife	29
2.4.1 Die Technik Artikulatorischer Unterdrückung	31
2.4.2 Der Effekt Irrelevanter Sprache	32
2.4.3 Der Phonologische Ähnlichkeitseffekt	36
2.4.4 Der Wortlängeneffekt	37
3 EIN QUANTITATIV- MATHEMATISCHES GEDÄCHTNISMODELL: die modifizierte Total-Time Hypothese (mTTH)	50
<hr/>	
3.1 Ableitung und Darstellung der mTTH	50
3.2 Verknüpfung der mTTH mit dem Arbeitsgedächtnismodell	56
3.3 Experimentelle Befunde	58
3.3.1 Experimentelle Evidenz für die Annahme einer hyperbolischen Lernfunktion	59
3.3.2 Experimentelle Evidenz für die Äquivalenzhypothese	60
3.3.3 Experimentelle Befunde zur Verarbeitungszeit im Langzeitgedächtnispeicher	61
4 Ableitung und Formulierung der Hypothesen	65
<hr/>	
4.1 Empirische und statistische Hypothesen	66
4.1.1 Arbeitshypothesen	66
4.1.2 Zentrale Hypothesen	75

4.1.3 Explorative Hypothesen	79
4.2 Aufbau der Studie	82
5 Versuchsplanung	85
<hr/>	
5.1 Aufbau der Experimentreihen	85
5.1.1 Experiment zur Messung der Artikulationszeit	87
5.1.2 Das Subspannen-Experiment	89
5.1.3 Das Supraspannen-Experiment	91
5.1.4 Das Subspannen-Experiment mit Artikulatorischer Unterdrückung	94
5.1.5 Das Supraspannen-Experiment mit Artikulatorischer Unterdrückung	95
5.2 Materialkonstruktion	95
5.2.1 Zahlenmaterial	95
5.2.2 Wortmaterial	97
5.3 Probanden	103
5.4 Operationalisierung der abhängigen Variablen	104
6 Intraexperimentelle Hypothesenprüfung	107
<hr/>	
6.1 Kontrolle der Experimentgüte und deskriptive Ergebnisse	107
6.1.1 Reliabilität der Zeitmessung	108
6.1.2 Durchführungsobjektivität	109
6.1.3 Validität und konvergente Validierung einzelner Parameter	113
6.1.4 Diskussion und Ergebnisse	116
6.2 Überprüfung der modifizierten Total-Time Hypothese (mTTH)	118
6.2.1 Experiment 1 (Stille, Deutsch, Zahlen, Symbol)	119
6.2.2 Experiment 2 (Art. Unterdrückung, Deutsch, Zahlen, Symbol)	120
6.2.3 Experiment 3 (Stille, Spanisch, Zahlen, Symbol)	121
6.2.4 Experiment 4 (Stille, Chinesisch, Zahlen, Symbol)	121
6.2.5 Experiment 5 (Art. Unterdrückung, Chinesisch, Zahlen, Symbol)	121
6.2.6 Experiment 6 (Stille, Deutsch, Phonolog. ähnl. Worte, Schrift)	121
6.2.7 Experiment 7 (Stille, Deutsch, Phonolog. unähnl. Worte, Schrift)	122
6.2.8 Experiment 8 (Stille, Deutsch, Zahlen, Schrift)	122
6.2.9 Überblick über die Ergebnisse	122
6.3 Prüfung der Äquivalenzhypothese der mTTH	126
6.3.1 Experiment 1 (Stille, Deutsch, Zahlen, Symbol)	128
6.3.2 Experiment 2 (Art. Unterdrückung, Deutsch, Zahlen, Symbol)	128

6.3.3 Experiment 3 (Stille, Spanisch, Zahlen, Symbol)	129
6.3.4 Experiment 4 (Stille, Chinesisch, Zahlen, Symbol)	129
6.3.5 Experiment 5 (Art. Unterdrückung, Chinesisch, Zahlen, Symbol)	129
6.3.6 Experiment 6 (Stille, Deutsch, Phonolog. ähnl. Worte, Schrift)	130
6.3.7 Experiment 7 (Stille, Deutsch, Phonolog. unähnl. Worte, Schrift)	130
6.3.8 Experiment 8 (Stille, Deutsch, Zahlen, Schrift)	130
6.3.9 Überblick über die Ergebnisse	131
6.4 Ergebnisse der intraexperimentellen Auswertung	131
7 Interexperimentelle Hypothesenprüfung	133

7.1 Voraussetzungen interexperimenteller Vergleiche	133
7.2 Überprüfung der Arbeitshypothesen	136
7.2.1 Der Spracheffekt	137
7.2.2 Der Effekt Artikulatorischer Unterdrückung	139
7.2.3 Der Geschlechtseffekt	141
7.2.4 Der Materialeffekt	142
7.2.5 Der Effekt des Präsentationsmodus	144
7.2.6 Überblick über die Ergebnisse	145
7.2.7 Einzelfalluntersuchungen	148
7.3 Überprüfung der Invarianzhypothese von Baddeley	149
7.4 Überprüfung der Invarianzhypothese von Doshier und Ma	153
7.5 Experimentelle Manipulationen und die Asymptotenhöhe der mTTH	157
7.6 Experimentelle Manipulation und der Quotient a/b sowie die Verarbeitungszeit im Langzeitgedächtnisspeicher	159
7.6.1 Der Spracheffekt	160
7.6.2 Der Effekt Artikulatorischer Unterdrückung	161
7.6.3 Der Materialeffekt	162
7.6.4 Überblick über die Ergebnisse	163
7.7 Überblick über die Ergebnisse zu den Invarianzhypothesen	165
8 Diskussion	168

8.1 Diskussion zum Arbeitsgedächtnismodell	168
8.2 Diskussion zur Invarianzhypothese von Baddeley	169
8.3 Diskussion zur Invarianzhypothese von Doshier und Ma	170

8.4 Diskussion zur mTTH	172
8.4.1 Diskussion der Gültigkeit der mTTH	173
8.4.2 Diskussion der Äquivalenzhypothese	175
8.4.3 Diskussion zur Verarbeitungszeit im Langzeitgedächtnisspeicher	177
8.4.4 Zusammenfassende Diskussion der mTTH	180
8.5 Zusammenfassung und Ausblick	181

Literaturverzeichnis	184
-----------------------------	------------

Anhang I Instruktionen zu den Hauptexperimenten

Anhang II Ergebnisse zu den Vorexperimenten

1 Einleitung

Das Phänomen „Gedächtnis“ übte schon immer eine besondere Faszination aus. Wie aufgrund der langen Beschäftigung mit diesem Thema zu erwarten ist, bestehen sehr verschiedene Interessenschwerpunkte nebeneinander. Gegenstand dieser Arbeit ist die Kapazitätsbegrenzung des Gedächtnisses, wie sie durch Kurz- und Langzeitgedächtnisspanne beschrieben wird. Unter dem Terminus „Kurzzeitgedächtnisspanne“ wird die Anzahl von Items verstanden, die unmittelbar nach einem Lernversuch wiedergegeben werden können. Der Begriff „Langzeitgedächtnisspanne“ bezeichnet die aufgenommene Information in Lernversuchen, in denen die Kapazität des Kurzzeitgedächtnisses überschritten wird (Kapitel 3.1). Schon Platon entwickelte nicht nur das Wachstafel-Gleichnis als Gedächtnismodell, sondern beschrieb explizit Gedächtnisfehler, die - in der heutigen Terminologie - als Konsequenzen von Interferenz interpretiert werden können:

Ich kenne dich und Theodorus und habe in dem Wachsblock von euch beiden die Abdrücke wie von Siegelringen; nun sehe ich aber euch beide nur von weitem und nicht genau und bemühe mich daher, den jeweils passenden Abdruck der entsprechenden Wahrnehmung zuzuordnen, in die zugehörige Spur zu bringen und ihr somit anzupassen, damit ein Wiedererkennen stattfindet; dabei aber greife ich daneben und vertausche beides wie jemand, der seine Schuhe verkehrt herum anzieht, in dem ich die Gesichtswahrnehmung von beiden mit dem falschen Abdruck zusammenbringe (Platon, Theaitet, 193 b f.).

In vorliegender Arbeit wird anhand eines mathematischen Modells experimentell die Wirkung verschiedener Untersuchungsbedingungen – eben auch der Einfluß von Interferenz - auf Kurz- und Langzeitgedächtnisleistungen sowie auf die Verbindung zwischen Kurz- und Langzeitgedächtnisspanne überprüft.

Im folgenden wird ein Überblick über die Gliederung dieser Arbeit gegeben. Nach dem Klassifikationsschema von Grainger und Jacobs (1998) können drei Arten von Modellen in der psychologischen Forschung unterschieden werden: verbale, mathematische und

algorithmische. Entsprechend dieser Einteilung und unter Ausschluß algorithmischer Modelle sind Kapitel 2 und 3 konzipiert. In *Kapitel 2* wird ausführlich das Arbeitsgedächtnismodell von Alan D. Baddeley vorgestellt, das den Ausgangspunkt für die Wahl der experimentellen Variation in den eigenen Experimenten bildet. Die meisten Studien der gegenwärtigen Gedächtnisforschung berücksichtigen dieses psychologische Modell mit deskriptivem Schwerpunkt.

In *Kapitel 3* wird dem verbalen Modell ein psychologisches Gedächtnismodell mit mathematischem Schwerpunkt, die Invarianzhypothesen und deren Verknüpfung, gegenübergestellt. Invarianzhypothesen beschreiben die limitierte Gedächtniskapazität dadurch, daß aus dem Produkt oder der Summe zweier Parameter ein Wert resultiert, von dem angenommen wird, daß er unter differierenden Umständen konstant bleibt. Durch die Verknüpfung verschiedener Invarianzhypothesen wird die Verbindung zwischen Kurz- und Langzeitgedächtnisspanne quantifiziert.

Der methodische Teil wird durch die beiden ersten Schritte empirischer Untersuchungen nach Erdfelder (1994, S. 48) definiert: „Festlegung einer empirischen Domäne“ und „Festlegung eines Untersuchungsdesigns“. In *Kapitel 4* wird die Fragestellung dieser Arbeit formuliert. Es folgt eine Ableitung der Hypothesen aus den dargestellten Theorien und die Bestimmung des Aufbaus dieser Studie.

In *Kapitel 5* werden die Versuchsanordnung der einzelnen Experimente sowie Präliminarien und Vorexperimente zur Konstruktion des Stimulusmaterials dargestellt. Am Ende dieses Kapitels wird auf die Charakteristika der Stichprobenselektion und die Operationalisierung der abhängigen Variablen eingegangen.

In den folgenden beiden Kapitel 6 und 7 werden die Hypothesen anhand der beobachteten Daten überprüft. Um Übersichtlichkeit zu gewährleisten, werden zwei Arten von Auswertungen unterschieden. In *Kapitel 6* „Intraexperimentelle Hypothesenprüfung“ erfolgt die Hypothesenprüfung separat für jede Experimentreihe. In *Kapitel 7* „Interexperimentelle Hypothesenprüfung“ werden die Parameter verschiedener Experimentreihen miteinander verglichen.

Den Abschluß dieser Arbeit bildet die Diskussion in *Kapitel 8*. Die Hypothesen und deren Vereinbarkeit mit den empirischen Daten werden für die wichtigsten Annahmen einzeln diskutiert. In Kapitel 8.5 folgt eine kurze Zusammenfassung der Diskussion, die auch einen Ausblick auf weitere Studien enthält.

2 EIN DESKRIPTIV-MODULARES GEDÄCHTNISMODELL: DAS ARBEITSGEDÄCHTNIS

Die Entwicklung deskriptiver Gedächtnismodelle wurde weitgehend von zwei Paradigmen dominiert: neben Ansätzen mit dem Schwerpunkt eines modularen Aufbaus der Gedächtnisspeicher (beispielsweise das Modal-Modell von Atkinson & Shiffrin, 1968), bestanden Ansätze, die prozessurale Komponenten fokussierten (zum Beispiel die Levels-of-Processing Theorie von Craik & Lockhart, 1972). Die Leistung des im folgenden näher erläuterten Arbeitsgedächtnis-Konzepts besteht unter anderem in der extensiven Integration der Ansätze. In der Konzeption des Arbeitsgedächtnisses finden im Vergleich zum Modal-Modell verstärkt aktive Kontroll- und Transformationsprozesse Berücksichtigung. Gleichzeitig werden separate Module zum einen von Lang- und Kurzzeitgedächtnisspeichern sowie zum anderen - erstmals - auch innerhalb des Kurzzeitgedächtnisspeichers postuliert. Nach der Modularitätshypothese von Fodor (1985) sind sogenannte „modulare“ Komponenten unabhängig voneinander, während die behaviorale Leistung durch das Zusammenwirken aller Systeme zustande kommt¹. Der Kurzzeitgedächtnisspeicher wird - unabhängig vom zugrundeliegenden Modell - durch vier Punkte charakterisiert: durch die nur kurzzeitige Speicherung, durch die Begrenztheit der Ressourcen und der Informationsaufnahme sowie als letztes Kriterium durch den schnellen und wenig Aufmerksamkeit beanspruchenden Abruf.

Die Definition des Begriffs „Arbeitsgedächtnis“ ist umstritten. Im folgenden wird ein sehr umfassender und relativ vager Bestimmungsversuch von Baddeley und Logie (1999, S. 28f.) zitiert:

Our own definition of working memory is that it comprises those functional components of cognition that allow humans to comprehend and mentally represent their immediate environment, to retain information about their immediate past experience, to support the acquisition of new knowledge, to solve problems, and to formulate, relate, and act on current goals.

¹ „Modularität“ im strengen Sinne muß für die Konzeption eines Arbeitsgedächtnisses nicht erfüllt sein (Zimmer, 1993).

Baddeley (1992) weist darauf hin, daß sein Gedächtnismodell nicht dem Kriterium der Falsifizierbarkeit nach Popper (1969) genügt. „I tend to develop a theory on the basis of experimental results, rather than set up an abstract theory and then find ways of testing it“ (Baddeley, 1992, S. 3). Seiner Meinung nach ist das nicht erforderlich, da er nur zwei Anforderungen an sein Modell stellt: erstens sollen Befunde einfach und plausibel erklärt werden. Zweitens soll weitere Forschung angeregt werden. Beides scheint ihm wohl gelungen, fraglich allein bleibt, ob dieser Anspruch für eine wissenschaftliche Theorie ausreichend ist, denn „ein Kausalgesetz ist aber auf jeden Fall mehr als eine empirische Regelmäßigkeit“ (Westermann, 2000, S. 151). Induktiv-statistische Erklärungen müssen folgende Adäquatheitsbedingungen erfüllen, um wissenschaftliche Erklärungen zu sein (Hempel, 1977)

1. Explanans und Explanandum sind empirisch
2. Das Explanans enthält nur wahre Sätze
3. Das Explanandum ist wahr
4. Das Explanans enthält mindestens eine gesetzesartige Aussage, das Explanandum ist nicht gesetzlich
5. Das Explanans macht das Explanandum sehr wahrscheinlich
6. Das Explanandum folgt nicht allein aus den nicht-gesetzlichen Sätzen des Explanans

Für das Arbeitsgedächtnis-Modells sind diese Kriterien wissenschaftlicher Erklärung in spezifischen Teilbereichen erfüllt. Beispielsweise kann die Erklärung zum Wortlängeneffekt in der Formulierung als Invarianzhypothese (Kapitel 3) als gesetzliche Aussage interpretiert und somit im Begründungszusammenhang untersucht werden. In diesem wissenschaftstheoretischen Kontext können auch die in dieser Arbeit vorgestellten Modifikationen zu dem ursprünglichen Erklärungsansatz zum Wortlängeneffekt nach Baddeley als Realisationen der Forderung nach maximaler Spezifizierung der Bezugsklasse interpretiert werden (Kapitel 2.4.4).

2.1 AUFBAU DES ARBEITSGEDÄCHTNISSES NACH BADDELEY

Baddeley und Hitch (1974) postulieren eine Aufteilung des Kurzzeitgedächtnisspeichers (Abbildung 1) in drei interagierende Module: die Zentrale Exekutive, die Phonologische Schleife und der Visuell-räumliche Notizblock. Der Kurzzeitgedächtnisspeicher wird dabei nicht als reines Speichermedium verstanden, sondern als „Working Memory“, das mit dem gespeicherten Material zusätzlich Operationen durchführt.

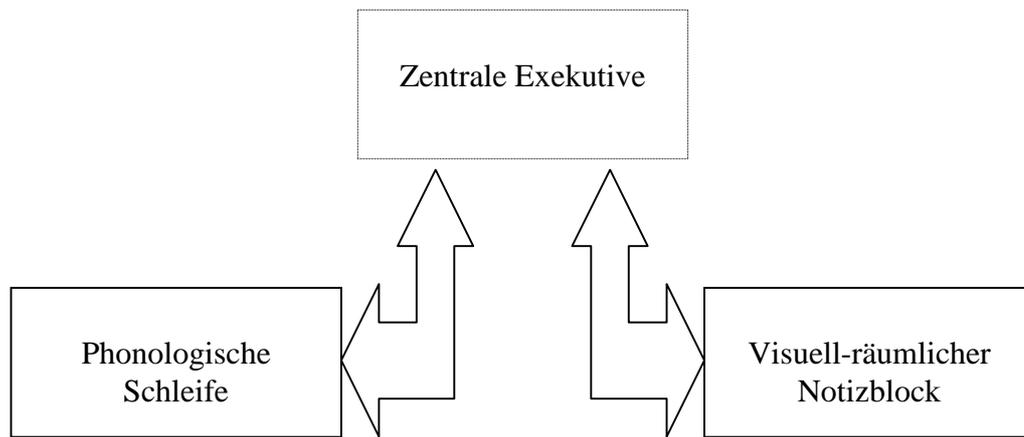


Abbildung 1 Schematische Darstellung des Arbeitsgedächtnisses nach Baddeley (1999)

Die Erforschung des Arbeitsgedächtnisses läßt sich charakterisieren durch das Paradigma der Dual-Task-Situation, durch die Logik doppelter Dissoziation und durch die Bandbreite der Untersuchungsmethoden.

Unter Dual-Task-Paradigma wird die experimentelle Methode der simultanen Bewältigung von zwei verschiedenen Aufgaben verstanden. Wenn beide Aufgaben dasselbe Modul beanspruchen, ist ein Leistungszusammenbruch wegen lokaler Interferenz zu erwarten. Ein typisches Beispiel innerhalb des Konzepts der Phonologischen Schleife ist die signifikante Reduktion der Gedächtnisspanne bei zusätzlicher Artikulatorischer Unterdrückung (Chincotta & Underwood, 1997b). Belegen Primär- und Sekundäraufgabe unterschiedliche Systeme, entsteht keine wesentliche, lokale Interferenz und kein Leistungszusammenbruch. Beide Aufgaben können simultan durchgeführt werden. Eine eventuell stattfindende Leistungsminderung ist durch Ressourcenallokation zu erklären. Da die Gesamtkapazität begrenzt ist, sinkt bei hoher Ressourcenabsorbtion einer Aufgabe die Ausführungsqualität der anderen. Dies gilt besonders für die Arbeitsweise des Notizblocks, die weniger automatisiert zu sein scheint als die der Phonologischen Schleife (Morris, 1987).

Die Untersuchungen zum Aufbau des Arbeitsgedächtnisses folgen der Logik funktionaler Dissoziation: wenn durch eine Manipulation der unabhängigen Variablen (oder durch eine bestimmte Sekundäraufgabe oder durch eine spezifische, zerebrale Schädigung eines Patienten) die Bewältigung einer Aufgabe X gestört wird, nicht jedoch das Lösen von Aufgabe Y, so muß, bevor zwei unabhängige Einheiten zur Lösung der beiden Aufgaben postuliert werden, eine weitere experimentelle Manipulation (oder eine Zweitaufgabe oder ein anderer Patient mit komplementärem, neuropsychologischen Ausfall)

gesucht werden, bei der (dem) nur die Ausführung von Aufgabe Y, nicht aber auch von Aufgabe X beeinträchtigt ist. Allerdings ist auch bei Nachweis des Vorliegens einer doppelten Dissoziation die Möglichkeit ähnlicher, durch die Aufgaben evozierter Anforderungen nicht auszuschließen.

Ein drittes Charakteristikum bei der Erforschung des Arbeitsgedächtnisses ist die Bandbreite der verwendeten Methoden und die Kombination von beispielsweise experimentellen, neuropsychologischen, physiologischen oder entwicklungspsychologischen Daten. Im folgenden wird anhand von drei ausgewählten Methoden Evidenz für eine Zurückweisung der Annahme eines unitären Kurzzeitspeichers und das Postulat der Unabhängigkeit der Subsysteme nach Baddeley an je einem Beispiel skizziert.

Experimentelle Belege zur Aufspaltung in funktionale Einheiten beruhen auf einer Serie von Untersuchungen mit dem Dual-Task-Paradigma. Eine Belastung der Zentralen Exekutive oder eine visuell-räumliche Zusatzaufgabe evozierten keinen gravierenden Leistungszusammenbruch bei einer verbalen Gedächtnisaufgabe, für deren Bewältigung allein die Phonologische Schleife angenommen wird (Baddeley, 1997). Ebenso konnte gezeigt werden, daß visuell-räumliche Aufgaben trotz Belastung der Zentralen Exekutive oder Belegung der Phonologischen Schleife erledigt werden können. Die Leistung verminderte sich dagegen signifikant bei Induktion von lokaler Interferenz durch räumliche oder visuelle Zusatzaufgaben (Tresch, Sinnamon & Seamon, 1993).

Um die Notwendigkeit der Trennung zwischen den beiden Subkomponenten nachzuweisen, wurden im Sinne einer doppelten Dissoziation neuropsychologische Einzelfalluntersuchungen an Patienten mit Schädel-Hirn-Verletzung durchgeführt. Bei einer Patientin wurde in Baddeleys Termini eine Beeinträchtigung des Visuell-räumlichen Notizblocks bei gleichzeitig intakter Phonologischer Schleife diagnostiziert (Hanley, Young & Pearson, 1991). Bei einer anderen Patientin zeigte sich das gegenläufige Bild: verbale Leistungen waren beeinträchtigt, die Arbeiten, die dem Notizblock zugeschrieben werden, waren nicht defizitär (Belleville, Peretz & Arguin, 1992).

Studien zu ereigniskorrelierten Potentialen sprechen ebenfalls für das Bestehen von zwei relativ unabhängigen Subsystemen: Beispielsweise zeigten sich systematische Unterschiede in der Lateralität und Lokalisation der Aktivierung bei Beanspruchung verschiedener Subsysteme. Während bei verbalen Gedächtnisaufgaben der linke Frontallappen negativierte, trat bei räumlichen Aufgaben eine langsame Negativierung des rechten Parietallappen auf (Ruchkin, Johnson, Grafman, Canoune & Ritter, 1992).

Der empirische Nachweis von drei Modulen des Kurzzeitspeichers ist problematisch. Baddeley (1986) selbst merkt an, daß die Logik doppelter Dissoziationen bei der hier realisierten Dreiteilung mit dem Modul der Zentralen Exekutive praktisch nicht realisierbar ist. Zudem kann jede der drei vorgestellten Methoden im Hinblick auf das Untersuchungsziel des Nachweises der Unabhängigkeit der beiden Subsysteme kritisiert werden.

Bei experimentellen Studien mit dem Dual-Task-Paradigma muß das Axiom gültig sein, daß Zusatzaufgaben spezifische Wirkung auf ein einziges Subsystem ausüben. Dies ist eine sehr starke Vorannahme, die zwar bezüglich der Trennung der beiden Subsysteme gerechtfertigt zu sein scheint, nicht jedoch im Hinblick auf die Zentrale Exekutive. Bei Nachweis von Konfundierungen ist die Aussagekraft der Interpretation eingeschränkt. Eine zweite Vorannahme des Dual-Task-Paradigmas ist ebenso wichtig: Es muß unbedingt sichergestellt werden, daß die Zweitaufgabe von allen Probanden zu jeder Zeit korrekt durchgeführt wird. Allein, die Überprüfung der Bearbeitungsgüte für die Zweitaufgabe und die Auswertungsobjektivität können jedoch sehr problematisch sein (Kapitel 2.2.1).

Klatte (1996) merkt an, daß bei neuropsychologischen Studien unmittelbar auf verschiedene Gedächtnissysteme geschlossen werden kann, weil durch das Auftreten spezifischer Ausfälle die Erklärungsmöglichkeiten von allgemeiner Interferenz oder Ähnlichkeit der Prozesse ausgeschlossen werden können. Da es sich bei neuropsychologischen Studien meist um Einzelfallstudien handelt, die weder den Ansprüchen einer Längsschnittstudie gerecht werden, noch denen einer Querschnittstudie einer Population existieren zwei Probleme. Erstens ist die Höhe der Kurzzeitgedächtnisspanne materialabhängig: Beispielsweise wurden bei einem Rechenkünstler extrem hohe Gedächtniswerte für Zahlen als Stimuli nachgewiesen. Seine Leistungen im Bereich der Buchstaben waren im interindividuellen Vergleich durchschnittlich, respektive es zeigte sich eine signifikante Differenz im intraindividuellen Vergleich (Bredenkamp, Klein, van Hayn & Vaterrodt, 1988). Zweitens kann die Gedächtnisspanne nicht als projizierbares Merkmal im Sinne Goodmans angesehen werden. Dies führt bei induktivem Vorgehen nach dem Kriterium von Nicod zu Paradoxien (Sainsbury, 1994). „Projizierbar“ bedeutet, daß es sich um eine substantielle und nicht-kontingente Eigenschaft des Untersuchungsobjekts handelt, die keinen Veränderungen mit der Zeit unterzogen ist. Die Gedächtnisspanne ist kein Dispositionsbegriff einer Person (Westermann, 2000), im Lauf der Zeit kann sich die Gedächtnisspanne auch ohne zerebrale Schädigung verändern – somit handelt es sich um kein projizierbares Merkmal. Bei neuropsychologischen Studien kommt hinzu, daß das Niveau der Gedächtnisspanne vor Schädigung des Gehirns nicht bekannt ist. Wenn also bei

einem Patienten nach einer Schädel-Hirn-Verletzung unterschiedliche Gedächtnisspannen zwischen verschiedenen Materialien festgestellt werden, ist das meines Erachtens kein zwingender Beleg für die Unabhängigkeit verschiedener Subsysteme.

Auch Ergebnisse physiologischer Methoden sind, zumindest in Bezug auf den Nachweis der Zentralen Exekutiv (Parkin, 1998), nicht eindeutig: „Some of the executive functions have been linked to the frontal lobes, although many tasks thought to involve the central executive may involve other areas of the brain as well” (Baddeley & Logie, 1999, S. 31).

Neben diesen Schwierigkeiten bei der Operationalisierung besteht für das beschriebene Vorgehen neuropsychologischer und physiologischer Studien das wissenschaftstheoretische Problem, daß induktive Schlüsse wegen der epistemischen Mehrdeutigkeit nicht wahrheitsfunktional sind (Hempel, 1977).

Der Erfolg des Baddeley-Modells kann durch drei Merkmale charakterisiert werden: Erstens: das Modell hat einen klaren und einfachen Aufbau aus dem spezifische Hypothesen abgeleitet werden können, die für jede Aufgabe ein adaptives Optimum postulieren. Zweitens ist der Anwendungsbereich groß: Dieses Gedächtnismodell kann auf einfache und komplexe Aufgaben wie Sprachproduktion (Dilger, 2000), Spracherwerb (Baddeley, Gathercole & Papagno, 1998), Lesen (Gathercole & Baddeley, 1993) oder Rechnen (Seitz, 1998) angewendet werden. Ein dritter Vorzug besteht in der Möglichkeit unabhängiger Operationalisierung durch experimentelle, neuropsychologische und physiologische Methoden. Allerdings zeigt sich hier – neben den aufgeführten Problemen - das bereits erwähnte Phänomen, daß das globale Modell auf konzeptueller Ebene nicht prüfbar ist: so ist etwa die Annahme modaler Repräsentationen im Kurzzeitspeicher als latente Variable schwer zu belegen (Kintsch, Healy, Hegarty, Pennington & Salthouse, 1999), auch wenn das Postulat eines amodalen, unitären Kurzzeitspeichers (Jones, 1993) widerlegt zu sein scheint (Meiser, 1997).

Für das Konzept des Arbeitsgedächtnisses sammeln sich auf vier Ebenen kritische Befunde und alternative Erklärungsmodelle an:

- Auf granularer Ebene modifizierten weiterführende Forschungsarbeiten den von Baddeley vorgeschlagenen Erklärungsansatz. Als Beispiel kann die Erweiterung des Effekts Irrelevanter Sprache zum Effekt Irrelevanter Töne (Kapitel 2.4.2) angeführt werden oder die Modifikationen zur Erklärung des Wortlängeneffekts (Kapitel 2.4.4).

- Die Diskussion auf prozessualer Ebene bezieht sich einerseits auf das Postulat eines Kontrollprozesses und andererseits auf die potentielle Existenz alternativer Prozesse. Brown und Hulme (1995) zeigten in Simulationsstudien, daß verschiedene Phänomene auch ohne Annahme eines Kontrollprozesses erklärt werden können. Ericsson und Delaney (1999) postulieren, daß der Kontrollprozeß in ungewohnten Laborexperimenten eine Rolle spielen mag, nicht aber bei komplexen, alltagsrelevanten Aufgaben. Zur Bewältigung dieser Anforderungen werden aufgaben- und bereichsspezifische „memory skills“ angenommen (Kapitel 2.2.3).
- Auf konzeptueller Ebene wird die Dreiteilung des Gedächtnisses diskutiert: von allgemein multimodalen Alternativen (Neath & Nairne, 1995) zu amodalen Modellen (Jones, 1993) und sich adaptiv entwickelnden, zusätzlichen Subsystemen, wie der tonalen Schleife bei Musikern (Pechmann & Mohr, 1992). Des weiteren ist eine Forschungsrichtung der Abkehr von statischen Einheiten hin zu einer Interpretation der Subsysteme als prozessurale Gedächtnisstrategien zu verzeichnen (Reisberg, Rappaport & O’Shaugnessy, 1984).
- Die Schwierigkeiten auf der Ebene der Operationalisierung wurden bereits angesprochen: das Paradigma der Zweitaufgaben kann nur funktionieren, wenn die Zusatzaufgabe wirklich prozeßrein systemspezifisch wirkt. Dies scheint mitunter fraglich: so konnten Seitz, Zoelch und Schumann-Hengsteler (2000) zeigen, daß Artikulatorische Unterdrückung mit einem einzigen Wort andere Effekte evoziert als wenn mit stereotypen Abfolgen wie „der- die –das“ unterdrückt wird (Kapitel 2.4.1).

Summa summarum ist der Arbeitsgedächtnis-Ansatz von Baddeley zur Zeit das einflußreichste Gedächtnismodell, obwohl es mit spezifischen Problemen behaftet ist, die in dieser Arbeit nur angedeutet werden können. Im folgenden wird detailliert auf die einzelnen Komponenten des Arbeitsgedächtnis-Modells eingegangen.

2.2 DIE ZENTRALE EXEKUTIVE

Das Konzept der Zentralen Exekutive ist weit weniger elaboriert als das der Phonologischen Schleife. Durch die rege Forschungsarbeit der letzten Jahre wurde jedoch das Manko behoben, daß die Zentrale Exekutive als Sammelstelle aller nicht-kategorisierbarer Phänomene verstanden wurde (Baddeley, 1986). Im folgenden wird die Konzeption der Zentralen Exekutive nach Baddeley vorgestellt und mit Ergänzungen durch die Gedächtnismodelle von Engle, Kane und Tuholsky (1999) sowie Ericsson und Delaney (1999) versehen.

Baddeley (1986) stellt einen Bezug zur Aufmerksamkeitsforschung her, indem er das Supervisory Attentional System (SAS) des Modells zur Handlungskontrolle von Norman und Shallice (1986) mit der Zentralen Exekutive vergleicht. Das Modell nimmt drei Kontrollebenen eigener Tätigkeit an. Die erste Ebene ist definiert als automatische Handlungskontrolle bei absoluter Kompatibilität zwischen externem Umweltreiz und internem Schema. Auf der zweiten angenommenen Ebene wird das Konkurrieren mehrerer Schemata durch Inhibition, respektive verstärkte Aktivierung eines Schemas gelöst (contention scheduling). Die dritte postulierte Stufe bezieht sich auf Situationen, in denen ein übergeordnetes und ressourcenlimitiertes Aufmerksamkeitssystem, das SAS, benötigt wird. Solche Situationen sind beispielsweise folgende Konstellationen:

- Es ist kein passendes Handlungskonzept vorhanden (Baddeley, 1999).
- Ein Fehler muß korrigiert werden (Baddeley, 1999).
- Eine adaptive oder voluntäre Entscheidung wird gefällt (Baddeley, 1999).
- Die Aufgabenlösung wird durch Interferenz oder Distraktoren gestört (Engle et al., 1999).
- Irrelevante Informationen müssen unterdrückt werden (Engle et al., 1999).
- Der eigene Output wird aktiv auf Fehler hin untersucht (Engle et al., 1999).
- Planvolle und aufmerksamkeitsfordernde mentale Operationen werden benötigt (Engle et al., 1999).

Evidenz für das SAS - Modell ergibt sich einerseits aus Alltagsfehlern, andererseits aus neuropsychologischen Untersuchungen (Baddeley, 1997). Unter Alltagsfehlern versteht man erstens die Aktivierung inadäquater Schemata, zum Beispiel das Bezeichnen von Prüflingen als Versuchspersonen. Zweitens fallen Perseverationen unter diesen Begriff, etwa wenn man Einkäufe auspackt und eine Banane in der Hand hält, unter Einfluß eines interessanten Gesprächs aber die Banane nicht in die Schale legt, sondern anfängt sie zu essen. Bei Patienten mit geschädigtem Frontalhirn findet sich nicht nur eine erhöhte Auftretenswahrscheinlichkeit der Alltagsfehler, sondern auch das Phänomen fehlender Hemmung bereits ausgeführter Handlungen: so spielen Musiker mit Schädel-Hirn-Verletzung stundenlang ein Stück, weil sie jedes Mal, wenn das Wiederholungszeichen kommt, vergessen haben, daß sie bereits die Wiederholung gespielt haben.

Der Zentralen Exekutive werden nach Baddeley (1996) vier Funktionen zugeordnet, die im weiteren besprochen werden:

- Adaptive oder voluntäre Strategiewechsel
- Regulation des Informationsflusses und Koordination der beiden Subsysteme

- Selektion der Aufmerksamkeit
- Verbindung zum Langzeitgedächtnis

Weder der Inhalt dieser Aufgaben noch die Definition als „Funktion“ sind erschöpfend theoretisch fundiert. Sie dienen lediglich der Systematisierung der Befunde und als Ansatz zur Diskussion. So schließen etwa Baddeley und Logie (1999) nicht aus, daß zum einen der Zentralen Exekutive weitere Funktionen zugeordnet werden können. Oder zum anderen, daß diese „Funktionen“ nur verschiedene Operationen weniger latenter, fundamentaler Prozesse sind. Diese Basisprozesse können entweder hierarchisch organisiert sein oder als relativ autonome Prozesse interagieren. Baddeley und Logie (1999) legen sich nicht eindeutig fest, ob die Zentrale Exekutive als unitäres System angesehen werden sollte oder ob nicht eine Fraktionierung angemessen wäre. Nur in einem Punkt herrscht innerhalb der Diskussion um den Zentralen Prozessor Kongruenz: für die Zentrale Exekutive wird keine Speicherfunktion postuliert.

Die „Funktionen“ der Zentralen Exekutive können meiner Ansicht nach anhand einer dichotomisierten Systematik beschrieben werden. Einerseits wird eine Meta-Ebene postuliert, die durch Kontrollmechanismen und bewußte Entscheidungen gekennzeichnet ist. Andererseits wird eine Produktionsebene eingeführt. Auf der Produktionsebene werden nicht-bewußte Prozesse angesetzt, die nach Engle (1996) durch Aktivierung und Inhibition beschrieben werden können. Diese Aufteilung hat den Vorteil, daß neben der Begrenzung der Ressourcen auch periphere, datenlimitierte Prozesse hinzugefügt werden können, auf deren Relevanz bereits in Simulationsstudien hingewiesen wurde (Meyer & Kieras, 1997 a & b).

Die Aufgabe „adaptiver oder voluntärer Strategiewechsel“ (Kapitel 2.2.1) kann als Interaktion von Meta- und Produktionsebene beschrieben werden: inadäquate Schemata werden inhibiert, während angemessene Schemata simultan, etwa durch bewußte Entscheidung, aktiviert werden. Die Funktion „Regulation des Informationsflusses“ (Kapitel 2.2.2) kann interpretiert werden als Aktivierungsprozeß eines Subsystems und Inhibition von Reizen, die nicht für die Aufgabenlösung relevant sind. Als drittes wird detailliert auf die Aufgabe „Verbindung zum Langzeitgedächtnis“ eingegangen: nach Engle (1996) ist die Aktivierung von Langzeitgedächtnis-Inhalten ein automatischer Prozeß, dem die Funktion der Inhibition irrelevanter Information gegenübersteht. Wegen der Strukturähnlichkeit zwischen den von Baddeley (1996) genannten Funktionen „Selektive Aufmerksamkeit“ und „Verbindung zum Langzeitgedächtnis“ innerhalb dieser Beschreibungssystematik wird auf diese beiden Aufgaben in einem gemeinsamen Kapitel eingegangen (Kapitel 2.2.3).

2.2.1 ADAPTIVE UND/ODER VOLUNTÄRE STRATEGIEWECHSEL

Aus dem SAS-Modell läßt sich das Generieren von Zufallsfolgen als eine Anforderung ableiten, die spezifisch die Zentrale Exekutive belastet. Das Zufallskonzept ist definiert als die Gleichwahrscheinlichkeit jedes innerhalb des Suchraums gültigen Ereignisses in jeder Produktion, unabhängig vom Antecedens. Die instruktionsgemäße Ausführung kann durch die beiden postulierten Ebenen dargestellt werden. Die Produktionsebene wird einerseits durch exzitatorische Prozesse der Schema-Aktivierung und andererseits durch Inhibition von Automatismen beschrieben. Auf der Meta-Ebene wird die Kontrollfunktion angesetzt: Durch kontinuierliche Überwachung wird Redundanz diagnostiziert und bei Bedarf auf der Produktionsebene ein Strategiewechsel initiiert.

Die Qualität der Aufgabenlösung bei Zufallsgenerierung variiert in den beiden abhängigen Variablen der Redundanz und der Latenzzeit. Diese Variationen werden als Folgen limitierter Kapazität interpretiert: „je höher die Produktionsrate ist, desto größer ist der Anteil redundanter Aktionen“ und „je größer der Suchraum ist, desto langsamer ist die Produktionsrate“ (Baddeley, 1996).

Realisationen dieser den Zentralen Prozessor belastenden Aufgabe sind beispielsweise das Generieren zufälliger Tastenfolgen, Zahlen- oder Buchstabensequenzen. Problematisch an diesen Operationalisierungen sind zwei Punkte. Erstens kann erst post-hoc ausgewertet werden, ob die Versuchspersonen wirklich instruktionsgemäß ihre Sekundäraufgabe durchgeführt haben. Die Redundanzanalyse bei Zahlenkombinationen weist zudem das Problem der geringen Auswertungsobjektivität auf. Das Generieren einer dem Probanden bekannten Telefonnummer entspricht zwar intern einer stereotypen Abfolge, die Wahrscheinlichkeit der Entdeckung durch den Versuchsleiter konvergiert aber gegen null (Baddeley, 1996). Der zweite Kritikpunkt stellt die Prozeßreinheit der Zufallsgenerierung in Frage (Vandierendonck, de Vooght & van der Goten, 1998 a und b). Das Drücken von Tasten in zufälligen Folgen beeinflusst Zentrale Exekutive und Notizblock, respektive das verbale Generieren von Zahlen oder Buchstaben den Prozessor und die Phonologische Schleife.

Eine mögliche Lösung zur Gewährung der Prozeßreinheit bietet die *Random-Interval-Repetition-Task* (RIR, Vandierendonck, de Vooght & van der Goten, 1998b). Diese Aufgabe ist so konzipiert, daß in variierenden Zeitabständen ein Computerton mit einem Tastendruck beantwortet werden soll. Durch die variablen Intervalle zwischen den Tönen soll stereotypes Drücken unterbunden werden. Der Einfluß des Notizblocks wird minimiert,

da nur eine einzige Taste zu drücken ist. Eine zweite Aufgabe ist die *Random-Time-Interval-Generation-Task* (RIG, Vandierendonck, de Vooght & van der Goten, 1998a). Während bei RIR auf einen Ton hin der Tastendruck erfolgt, kann bei RIG der Zeitpunkt des Tastendrucks selbst bestimmt werden. Die Intervalle zwischen den Tastendrücker sollten in ihrer Länge variieren. Diese Aufgabe entspricht eher der Baddeleyschen Intention der Zufallsgenerierung, da sowohl ein rhythmisches Tippen unterdrückt als auch eine zufällige Intervall-Länge aktiviert werden muß. Bei RIR dagegen ist nur die Aktivierungskomponente relevant, da nicht agiert, sondern nur reagiert wird.

Analog der Zufallsgenerierung wird die variierende Leistungshöhe als Effekt der Höhe der Beanspruchung limitierter Ressourcen interpretiert: Bei hoher Belastung der Zentralen Exekutive erhöht sich einerseits die Fehlerrate (bei RIG: rhythmisches Drücken, bei RIR: Auslassen von Tönen bzw. Drücken ohne diskriminativen Hinweisreiz). Andererseits sinkt bei RIG die Produktionsrate, respektive bei RIR steigt die Reaktionszeit.

Die Autoren betonen jedoch, daß auch bei RIR und RIG nicht ausschließlich die Zentrale Exekutive belastet wird, was als ein unrealisierbares Ideal angesehen wird, aber daß es dennoch gelungen ist „keeping the load on other processing to a minimum“ (Vandierendonck, de Vooght & van der Goten, 1998a, S. 417). Die Wirksamkeit von RIG und RIR wurde indirekt nachgewiesen durch Dissoziation der Interferenzmuster im Vergleich mit Aufgaben, die spezifisch die Subkomponenten belasten.

2.2.2 REGULATION DES INFORMATIONSFLUSSES

Bevor die exekutive Funktion „Regulation des Informationsflusses und Koordination der beiden Subsysteme“ untersucht werden kann, müssen nach Baddeley (1996) zwei spezifische Fragen geklärt werden. Erstens, ist die Zentrale Exekutive ein von den Subsystemen unabhängig arbeitendes Modell? Zweitens, ist die Zentrale Exekutive ein einheitliches System?

Zur Überprüfung der Hypothese, daß der zentrale Prozessor unabhängig von den beiden Subsystemen operiert, wurden zwei Aufgaben vorgegeben, wobei eine Aufgabe spezifisch die Zentrale Exekutive belastete und die andere ein Subsystem. So störte Artikulatorische Unterdrückung die Leistungen einer Zufallsgenerierung von Tastendrücker nicht (Baddeley, 1996). Da kein Leistungszusammenbruch wegen lokaler Interferenz auftrat, wird nach der Logik des Dual-Task-Paradigmas von verschiedenen Modulen der Zentralen Exekutive und den Subsystemen ausgegangen.

Eine zweite Untersuchung sollte durch Induktion lokaler Interferenz klären, ob die Zentrale Exekutive ein einheitliches System ist. Das Leistungsniveau in einem Test zur Wortflüssigkeit, operationalisiert durch das Generieren von Mitgliedern einer bestimmten Kategorie, verschlechterte sich bei simultanem Drücken verschiedener Tasten nach dem Zufallskonzept. Es wird angenommen, daß beide Aufgaben Anforderungen an den zentralen Prozessor stellen (Baddeley, 1996). Demnach könnte die Zentrale Exekutive als einheitliches System angesehen werden. Zu einem gegenteiligen Fazit kommt Lehto (1996). Er untersuchte den Zusammenhang zwischen Zentraler Exekutive und neuropsychologischen Tests, die mit exekutiven Funktionen in Zusammenhang gebracht werden. Von drei Tests korrelierte nur einer signifikant mit der Zentralen Exekutive, die Interkorrelationen zwischen den einzelnen Tests waren ebenfalls nicht signifikant. Dies spricht gegen die Existenz eines einheitlichen, kapazitätsbegrenzten Prozessors, wobei aber der Autor selbst konstatiert, daß allein wegen dieser Studie das Konzept einer einheitlichen Zentralen Exekutive nicht verworfen werden kann: „The present study cannot totally rule out the possibility that there might still be a general factor“ (Lehto, 1996, S. 48).

Folgt man der Annahme eines eigenen, einheitlichen Moduls der Zentralen Exekutive, entsteht die Frage, ob das Dual-Task-Paradigma auch auf diese Komponente übertragbar ist. Die Untersuchung der Funktion „Regulation des Informationsflusses und Koordination der Subsysteme“ bezieht sich nicht auf den Nachweis von Haupteffekten, sondern beschäftigt sich mit statistischen Interaktionen. Als Sekundäraufgabe wurde Zufallsgenerierung zur Belastung des zentralen Prozessors eingesetzt, als Primäraufgabe ein verbaler Kurzzeitgedächtnistest, der in der Schwierigkeit graduell von einem bis acht Items variierte. Es wird angenommen, daß die Kurzzeitgedächtnisspanne durch die Phonologische Schleife determiniert wird (zur Diskussion dieser Annahme vgl. Kapitel 2.4.4). Bei Erhöhung des Schwierigkeitsgrades werden mehr Ressourcen von der Zentralen Exekutive benötigt. Es konnte die Erwartung empirisch gestützt werden, daß sich bei steigender Sequenzlänge die Leistungen in der den zentralen Prozessor belastenden Sekundäraufgabe signifikant verschlechterten (Baddeley, 1996). Dieses Ergebnis spricht dafür, daß das Dual-Task-Paradigma auch für die Erforschung der Zentralen Exekutive geeignet ist.

Da neuropsychologische Störungsbilder und physiologische Lokalisationen im Gehirn weder ein notwendiges noch ein hinreichendes Kriterium für den Nachweis der Zentralen Exekutive sind, führte Baddeley (1996) Studien zur Herstellung korrelativer Zusammenhänge durch. Neuropsychologische Evidenz für die Wirkung der Zentralen Exekutive ergab sich bei Alzheimer-Patienten, die bei Doppeltätigkeit wesentlich schlechter abschnitten als eine gesunde Vergleichsgruppe. Dies war aufgrund der Parallelisierung nach individuellen Möglichkeiten der Einzeltätigkeit nicht erwartet worden (Baddeley, Bressi,

Della Sala, Logie & Spinnler, 1991). In physiologischen Studien mit PET zeigte sich ein Zusammenhang zwischen exekutiven Funktionen und dem präfrontalem Kortex (D'Esposito, Detre, Alsop, Shin, Atlas & Grossmann, 1995). Der präfrontale Kortex wird nur bei Koordination von zwei Tätigkeiten aktiviert, nicht während der separaten Durchführung der beiden Aktionen.

Diese Studien geben Hinweise darauf, daß das System der Zentralen Exekutive für die Regulation des Informationsflusses bei Kurzzeitgedächtnis-Aufgaben zuständig ist, auch wenn die zugrundeliegende Annahme einer unitären Komponente der Zentralen Exekutive kritisch zu sein scheint. Im folgenden Kapitel wird der Funktion „Verbindung zum Langzeitgedächtnis“ nachgegangen, die ebenfalls der Zentralen Exekutive zugeschrieben wird.

2.2.3 VERBINDUNG ZUM LANGZEITGEDÄCHTNIS

Die Zentrale Exekutive wird innerhalb des Arbeitsgedächtnis-Modells als das System angenommen, das die Verbindung zwischen Arbeitsgedächtnis und Langzeitgedächtnisspeicher herstellt. Diese Verknüpfung ermöglicht das Lösen komplexer Aufgaben wie Satzverstehen oder Schachspielen. Nur mit Hilfe des Langzeitspeichers kann die begrenzte Kapazität des Kurzzeitgedächtnisses überwunden werden. So können beispielsweise Gedächtnisphänomene wie das Expertentum des Rechenkünstlers (Bredenkamp, Klein, van Hayn & Vaterrodt, 1988) erklärt werden.

Es stellt sich die Frage, ob das Kurzzeitgedächtnis als Aktivierungszustand des Langzeitgedächtnisses angesehen werden kann (Cowan, 1999; Engle et al., 1999), wie auch physiologische Studien nahelegen (Fuster, 1998). Baddeley (1996) verneint diese Frage mit den Hinweisen darauf, daß beispielsweise das Arbeitsgedächtnis nur limitierte Ressourcen zur Verfügung hat. Zudem bestehen dissoziative Effekte zwischen sofortiger und verzögerter Wiedergabe: bei unmittelbar erfolgter Wiedergabe spielt die semantische Ähnlichkeit zwischen den Items einer Lernliste kaum eine Rolle (Baddeley, 1966) im Gegensatz zu stark verzögerten Gedächtnistests (Baddeley, 1999). Auch die Abrufmechanismen scheinen unterschiedliche Ansprüche an die Aufmerksamkeit zu stellen, je nachdem ob die Information im Kurzzeitspeicher aktiviert ist oder ob ein Rückgriff auf Langzeitgedächtnis-Repräsentationen erfolgt.

Die individuelle Kapazität des Arbeitsgedächtnisses kann durch verschiedene Methoden ermittelt werden: etwa durch Messung der Arbeitsgedächtnisspanne (Daneman & Carpenter,

1980), durch das FAN-Verfahren (Anderson, 1983) oder durch die Operation-Aufgabe (Turner & Engle, 1989). Die Aufgaben beziehen sich jeweils auf eine Merkaufgabe und eine komplexe, aufmerksamkeitsbeanspruchende Aufgabe (Daneman & Carpenter: das Hörverständnis von Sätzen; Anderson: Verifikation von Sätzen mit variierender Satzgröße; Turner & Engle: mathematische Operationen). Die Höhe der Kapazität des zentralen Prozessors wird durch die Anzahl korrekt reproduzierter Items bestimmt. Es besteht eine hohe Korrelation dieses Indikators mit der Bewältigungsgüte komplexer Aufgaben (für einen Überblick: vgl. Engle, 1996).

Interindividuelle Unterschiede in der Arbeitsgedächtniskapazität basieren nach Ericsson und Delaney (1999) auf drei potentiellen Faktoren. Erstens: Reichhaltigkeit und Grad der Vernetztheit des semantischen und episodischen Wissens im Langzeitspeicher variieren zwischen den Individuen. Zweitens unterscheiden sich erworbene Kompetenzen und angewendete Strategien im Umgang mit mentalen Operationen, basierend auf verschiedenen individuellen Erfahrungen. Drittens werden Differenzen bezüglich „sustaining, maintaining, and shifting attention“ (Engle et al., 1999, S. 126) angenommen. Diese drei Punkte werden im folgenden besprochen. Der Fokus wird dabei auf Differenzen in der Aufmerksamkeit und den Forschungsansatz von Engle (1996) gerichtet, da hier ein Bezug zu Wiedergabeprozessen hergestellt wird, der für die Diskussion der empirischen Untersuchung (Kapitel 8) relevant ist.

Die Bedeutsamkeit von Existenz und Grad der Verknüpfung der Repräsentationen im Langzeitspeicher wird durch systematische Vergleiche belegt. Intraindividuelle Vergleiche zeigen, daß nur in den spezifischen Wissensgebieten die Assoziationen zwischen den Items oder zwischen den Items und den Repräsentationen leichter fallen (Ericsson & Delaney, 1999). Interindividuelle Vergleiche zwischen Experten und Novizen belegen Vorteile für Experten in der Verfügbarkeit semantischer Hinweisreize und in der Leichtigkeit der Manipulation bereits bestehender Repräsentationen (Ericsson & Delaney, 1999).

Innerhalb des Ansatzes von Ericsson und Delaney (1999) wurde der zweite Punkt bezüglich der sogenannten „memory skills“ elaboriert. Diese Fertigkeiten entwickeln sich materialspezifisch aufgrund einer iterierten Beschäftigung mit einer Aufgabe. Durch Antizipation der Abrufsituation und dem Wissen um aufgabenrelevante Informationen wird sowohl die Effizienz der Enkodierung als auch die Effektivität des Abrufs maximiert. Zudem werden aktiv Strategien zur Vermeidung von proaktiver Interferenz entwickelt, beispielsweise durch Implementation des temporalen Kontextes oder durch zusätzliche elaborative Assoziation zwischen den Chunks. Innerhalb des Theoriemodells von Ericsson

und Delaney (1999) wird eine Integration der neuen Informationen in die bereits vorhandenen komplexen Strukturen im Langzeitspeicher postuliert. Während die Aktivierung im Kurzzeitspeicher durch Zusatzaufgaben reduziert wird, bleiben die Repräsentationen im Langzeitgedächtnisspeicher resistent gegenüber einer Wirkung von Simultanaufgaben. Die neu integrierten Items können bei Bedarf redintegriert werden. Diese Gedächtnisfertigkeiten sind im Alltag von besonderer Bedeutung: die Limitierung der Gesamtkapazität kann transzendierte werden und die simultane Bewältigung mehrerer Aufgaben wird ermöglicht.

Die dritte potentielle Ursache interindividueller Unterschiede bezieht sich auf die Aufmerksamkeitskapazität. Baddeley (1996) berichtet von einer entwicklungspsychologischen Studie, in der konstatiert wird, daß trotz Ausparialisierung der Faktoren „fluide Intelligenz“ und „generelle Informationsverarbeitungszeit“ Unterschiede in der Leistung eines Reaktionszeitexperiments von jungen Erwachsenen und Senioren bestehen. Nur wenn Störreize die gleiche Modalität wie die Testreize hatten, konnte dieser Effekt nachgewiesen werden, nicht jedoch bei Dimensionswechseln. Baddeley (1996, S. 21) interpretiert diesen Befund, der in Kongruenz zu der im folgenden dargestellten Theorie von Engle et alii (1999) steht, durch die im Alter nachlassenden Inhibitionsmechanismen „less highly peaked attentional focus“.

Die Theorie von Engle et alii (1999) über Aktivations- und Inhibitionsmechanismen wird im folgenden am Beispiel der Wiedergabe erklärt: die Wiedergabeleistung in Gedächtnistests wird als Produkt von drei komplementären und teilweise optionalen Prozessen (Engle et al., 1999) interpretiert:

1. Automatische Aktivationsausbreitung
2. Optional: intentionale Kategorien-Suche
3. Bei freier Wiedergabe: Blockierung der Aktivierung bereits genannter Items

Engle (1996) schloß zur Interpretation interindividueller Unterschiede zwischen Probanden mit hoher, respektive niedriger Arbeitsgedächtniskapazität unterschiedliche Aktivierungsressourcen aus. Die Aktivierungszeit wurde ungefähr mit 600 Millisekunden angegeben und variierte in seiner Studie nicht systematisch zwischen den Versuchspersonen.

In puncto intentionale Kategorien-Suche unterschieden sich Probanden mit hoher und niedriger Arbeitsgedächtnisspanne: durch aufmerksamkeitsbeanspruchende Zusatz Tätigkeiten wurde die Leistung von Probanden mit niedriger Spanne nicht tangiert, die Leistung von Versuchspersonen mit hoher Spanne dagegen wurde reduziert (Engle et al.,

1999). Eine Untersuchung nach der Prozeß-Dissoziations-Prozedur von Jacoby (1991) zeigte Unterschiede zwischen den beiden Gruppen bezüglich der bewußten Parameter, die in der Theorie von Engle den Inhibitionsmechanismen entsprechen. Kein Gruppeneffekt trat jedoch bei den Parameterschätzungen für unbewußte Leistungen auf, die nach Engle (1996) als Aktivationsprozesse interpretiert werden. Angenommen wird, daß bei Personen mit schwächeren, exekutiven Leistungen die Inhibition irrelevanter Informationen schlechter gelingt.

Dies belegt auch das Ergebnis einer Untersuchung zur Wortflüssigkeit (Engle et al., 1999). Die Probanden sollten so viele Mitglieder der Kategorie „Tiere“ aufschreiben, wie ihnen in einem fixen Zeitintervall einfiel. Wiederholungen sollten vermieden werden. Es zeigte sich, daß Probanden mit hoher Arbeitsgedächtnisspanne weniger redundante Items wiedergaben als Personen mit niedriger Arbeitsgedächtnisspanne. Ein weiterer Unterschied der Probanden mit hoher, respektive geringer Kapazität zeigte sich in einer Untersuchung zum Sternberg-Paradigma. Die Überprüfungszeit bei Personen mit hoher Arbeitsgedächtnisspanne war schneller als bei denen mit niedriger Spanne (Conway & Engle, 1994). Allerdings trat dieser Effekt nur auf, wenn einzelne Zielitems Mitglieder verschiedener Sets waren. Dieser Befund wurde so erklärt, daß unter Interferenzbedingungen die Probanden mit hoher Kapazität die irrelevante Alternative leichter unterdrücken konnten.

Aus dieser Interpretation ergibt sich meines Erachtens ein nur scheinbarer Widerspruch zu Studien, die belegen, daß Probanden mit hoher Arbeitsgedächtnisspanne mehrere Alternativen gleichzeitig parat halten können, beispielsweise bei Aufgaben mit Sackgassen-Sätzen (Mac Donald, Just & Carpenter, 1992). Wenn beim Textverstehen irrelevante Interpretationen unterdrückt würden, dann müßten die Personen mit hoher Arbeitsgedächtnisspanne weniger Interpretationsalternativen aktiviert haben. Dementsprechend müßte eine Uminterpretation eines Satzes länger dauern als bei Probanden mit niedrigeren Arbeitsgedächtnisspannen, die ja laut Engle et al. (1999) weniger effektiv irrelevante Informationen unterdrücken können. Die empirischen Befunde zeigen eindeutig in die gegenteilige Richtung: Probanden mit hoher Arbeitsgedächtnisspanne gelingt eine Neuinterpretation des Satzes wesentlich schneller. Meines Erachtens sind die Befunde für die Theorie von Engle (1996) nicht notwendigerweise kritisch. Der Widerspruch könnte durch methodische Unterschiede zwischen den Aufgaben entstanden sein, da die optimale Aufgabenlösung bei verschiedenen Untersuchungszielen nicht invariant sein muß. Bei Gedächtnisaufgaben mit freier Wiedergabe ist die Inhibierung bereits genannter Items optimal, um Redundanzen zu vermeiden. Beim Textverstehen ist

wegen der Gefahr von Sackgassen-Sätzen das simultane Bereithalten mehrerer Interpretationsalternativen von Vorteil.

Insgesamt kann resümiert werden, daß die Konzeption der Zentralen Exekutive nicht mehr lediglich als ein „homunculus“ bewertet werden kann. Die Erforschung einzelner Funktionen ermöglichen meines Erachtens das Generieren prüfbarer Hypothesen und ein solides Verständnis der Zentralen Exekutive in diesen Teilbereichen. Forschungsbedarf scheint zu bestehen in Hinblick auf die Gesamtkonzeption der Zentralen Exekutive bezüglich der Begriffsintension und für den Funktionsbereich bezüglich der Begriffsextension.

Im den folgenden beiden Kapiteln wird auf die zwei Subsysteme eingegangen: auf den Visuell-räumlichen Notizblock (Kapitel 2.3) und die Phonologische Schleife (Kapitel 2.4).

2.3 DER VISUELL-RÄUMLICHE NOTIZBLOCK

Der Visuell-räumliche Notizblock, dem im Arbeitsgedächtnismodell die Funktion der Speicherung und Bearbeitung visuell-räumlicher Informationen zugeschrieben wird, ist nicht Part der in dieser Arbeit vorgestellten experimentellen Überprüfung. Aus diesem Grund erfolgt lediglich ein komprimierter Überblick. Vorab muß erwähnt werden, daß der Notizblock (noch) weniger elaboriert ist als die Phonologische Schleife, und daß die Ergebnisse verschiedener Studien widersprüchlich sind.

Die uneinheitliche Ergebnislage könnte auf einer stärkeren Interaktion zwischen Notizblock und Zentraler Exekutive basieren, da die Arbeit des Notizblocks nicht so hoch automatisiert zu sein scheint, wie das für die Phonologische Schleife angenommen werden kann (Quinn, 1988). Eine vergleichende Studie zwischen einer visuell-räumlichen und einer artikulatorischen Aufgabe zeigt, daß bei Arbeit des Notizblocks mehr Ressourcen benötigt werden (Klauer & Stegmaier, 1997). Analog resümieren Smyth und Scholey (1994), daß bei den für den Notizblock typischen Sekundäraufgaben, wie Tapping oder Tracking, räumliche Orientierung und visuelle Rückmeldung für Bewegungskontrolle benötigt werden. Diese Orientierung beansprucht Aufmerksamkeit, die im Arbeitsgedächtnismodell dem Funktionsbereich des Zentralen Prozessors zugeordnet wird. Die Operationalisierung einer prozeßreinen Zusatzaufgabe für den Notizblock ist bisher noch nicht gelöst (Quinn & McConnell, 1996a). Primäraufgaben wie das Generieren von Vorstellungen aus dem Langzeitgedächtnisspeicher oder mentale Rotation sind ebenfalls Operationen, die nicht nur durch den Notizblock belastende Zusatzaufgaben gestört werden, sondern auch in großem

Ausmaß durch die Zentrale Exekutive belastende Aktionen (Logie, 1995). Somit ist das gesamte Forschungsgebiet der visuellen Vorstellung und die Diskussion um analoge, respektive propositionale Repräsentationen nicht eindeutig dem Notizblock zuzuordnen (Baddeley, 1999).

Für die enge Verbindung zwischen Zentralem Prozessor und Subsystem spricht neben diesen Befunden auch die wesentliche Bedeutung des Langzeitgedächtnisses: so können bildhafte Worte besser gemerkt werden als abstrakte (Paivio, 1986). Dieser Effekt kann durch räumliche Zweitaufgaben nicht eliminiert werden (Baddeley, 1999). Zweitens sind schon seit der Antike eidetische Mnemotechniken bekannt, deren Wirksamkeit auch bei zusätzlicher, räumlicher Sekundäraufgabe nicht verschwindet (Baddeley, 1999). Die Verbindung zwischen den Kurzzeitmodulen und dem Langzeitgedächtnisspeicher wird innerhalb des Arbeitsgedächtnis-Modells dem System der Zentralen Exekutive zugeordnet. Die Besondere Relevanz der Zentralen Exekutive für den Notizblock wird in Abbildung 2 nach Logie (1995) dargestellt.

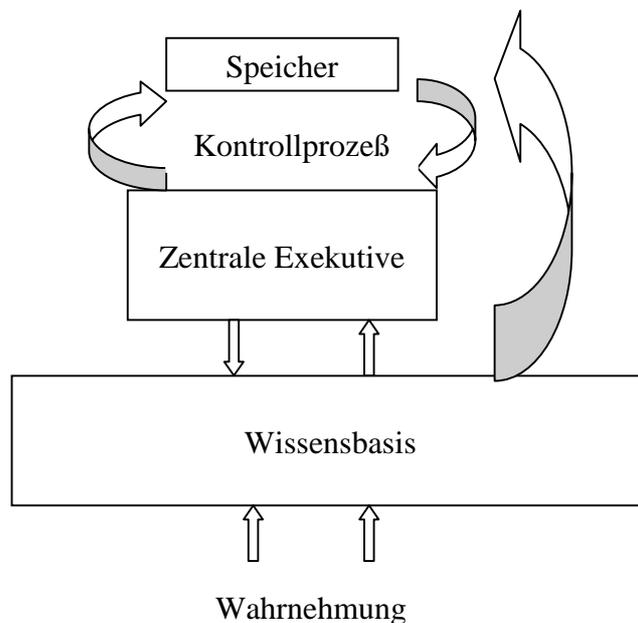


Abbildung 2 Visuell-räumliches Arbeitsgedächtnis nach Logie (1995)

Ein weiterer Problempunkt des Konzepts eines Notizblocks ist die Aufgabenkomplexität, die im Vergleich zur Phonologischen Schleife wesentlich höher ist: kontinuierlich müssen räumliche und visuelle Informationen verarbeitet und integriert werden (Baddeley, 1992). Die komplexen Anforderungen manifestieren sich in Ermüdungserscheinungen und einer langsamer verlaufenden Automatisierung als zum Beispiel bei Artikulatorischer Unterdrückung nachgewiesen wurde.

Ein Kernpunkt der Erforschung des Notizblocks ist die Parallelisierung zum phonologischen Subsystem. Ob Analogieschlüsse korrekt sind, ist umstritten, wie an einem Beispiel verdeutlicht werden soll. Serielle Reihenfolgen werden bei verbalem Material zeitlich kodiert (Jones, 1993). Als logische Folge ist die serielle Wiedergabe der umgekehrten Reihenfolge bei diesen Stimuli signifikant reduziert im Vergleich zur Reproduktion der dargebotenen Reihenfolge (Cowan et al., 1992). Bei räumlichen Gedächtnisaufgaben dagegen widersprechen sich die Studien: eine Untersuchung verzeichnete keinen Effekt der Variation in der Richtung der seriellen Wiedergabe (Isaacs & Vargha-Kadem, 1989), während in einer weiteren Studie doch eine Analogie zur Phonologischen Schleife gefunden wurde (Farrand & Jones, 1996).

Logie (1986) stellte ein Zwei-Komponenten-Modell des Visuell-räumlichen Notizblocks vor. Es wird ein visueller und ein räumlicher Speicher angenommen und zudem Kontrollprozesse, die unabhängig von der realen Ausführung von Augen- oder Körperbewegungen sind (Smyth, 1996). Während visuelle Informationen direkt in den Speicher gelangen, müssen räumliche oder verbale Informationen (Gathercole & Baddeley, 1993) durch einen aktiven Kontrollprozeß transformiert werden. Existenz, Wirkung und Notwendigkeit der postulierten Kontrollprozesse wird in der Forschung kritisch diskutiert.

Es wird angenommen, daß die räumliche und die visuelle Komponente spezifisch auf das jeweilige Stimulusmaterial reagieren. Die beiden Speichertypen interagieren stark miteinander, sind in ihrer Funktionsweise aber relativ unabhängig (Logie, 1995). Dennoch wird die Eigenständigkeit von visuellem und räumlichem Bereich von manchen Forschern als existentiell diskutiert, so daß erwogen wird, ob statt eines Notizblocks nicht je ein räumliches und ein visuelles Subsystem postuliert werden sollte (Smith & Jonides, 1995). Antithetisch dazu stehen Forscher, die das visuell-räumliche Arbeitsgedächtnis als so abhängig von der Zentralen Exekutive sehen, daß sie sogar die Existenzberechtigung eines Subsystems in Frage stellen (Phillips & Christie, 1977).

Für die Annahme von getrennten visuell und räumlichen Systemen sprechen experimentelle, neuropsychologische und physiologische Studien.

In Dual-Task-Experimenten ergeben sich funktionale Dissoziationen zwischen räumlichen und visuellen Anforderungen (Tresch et al., 1993): Die visuelle Sekundäraufgabe der Farbdiskrimination interferiert nur mit der visuellen Primäraufgabe der Rekognition geometrischer Figuren, kein Einfluß kann auf die räumliche Erstaufgabe „Lokalisation“ verzeichnet werden. Das Lokalisieren von Objekten wurde signifikant durch das Erkennen

von Targets in einem sich bewegenden Balken als räumliche Zusatzaufgabe erschwert, im Gegensatz zur visuellen Primäraufgabe. Auch in neuropsychologischen Studien wurde eine doppelte Dissoziation nachgewiesen: es gibt Patienten mit defekter visueller und intakter räumlicher Verarbeitung (Farah, Hammond, Levine & Calvanio, 1988). Bei einer anderen Patientin wurde von Problemen mit räumlichen Leistungen berichtet, während visuelle Aufgaben gut bewältigt werden konnten (Hanley et al., 1991). Zudem spricht diese Studie für die Trennung von Kurz- und Langzeitgedächtnis: denn alles, was vor der erworbenen Hirnschädigung gelernt worden ist, war auch nach der Krankheit zugänglich. Physiologische Studien belegen die Bedeutung des temporal, parietal-posterioren Bereichs für die Verarbeitung räumlicher Informationen sowie die Aktivierung parieto-okzipitaler Regionen bei visuellen Anforderungen (De Renzi, Faglioni & Previdi, 1977).

Die räumliche Komponente des Notizblocks kann durch zwei Effekte charakterisiert werden. Es besteht ein Dimensionalitätseffekt, der besagt, daß die Erinnerungsleistung für Bewegungssequenzen mit steigender Anzahl involvierter Dimensionen sinkt (Baddeley & Logie, 1999). Zudem kann der postulierte Kontrollprozeß durch zusätzliche Bewegungen blockiert werden (Baddeley & Lieberman, 1980), analog der Technik Artikulatorischer Unterdrückung (Kapitel 2.4.1). Der räumliche Part des Notizblocks ist eng verknüpft mit dem visuellen Vorstellungssystem des Langzeitgedächtnisses (Quinn & McConnell, 1996a). Wiederum vertreten Baddeley und Logie (1999) mit dem Verweis auf eine Studie von Morton und Moris die Meinung, daß das Arbeitsgedächtnis nicht ein Aktivierungszustand des Langzeitgedächtnisses ist. Diese berichten von einer Patientin, deren räumlich-visuelle Arbeitsgedächtnisleistungen vergleichbar zu gesunden Probanden war, eine Leistungsreduktion dagegen zeigte sich bei Vorstellungsaufgaben, die einen Abruf von Repräsentationen aus dem Langzeitgedächtnis erfordern. Das erhaltene Ergebnismuster sollte sich nicht ergeben, wenn das Arbeitsgedächtnis nur ein bestimmter Zustand des Langzeitgedächtnisses wäre.

Der visuelle Part des Notizblocks kann durch vier Effekte beschrieben werden. Erstens scheint ein Visueller Ähnlichkeitseffekt zu existieren (Logie, 1995). Dieser wird gekennzeichnet durch eine Leistungsreduktion bei hoher visueller Ähnlichkeit und durch eine erhöhte Anzahl von Vertauschungsfehlern. Zweitens zeigt sich ein Effekt der Merkmalskomplexität visueller Stimuli (Broadbent & Broadbent, 1981). Je höher die visuelle Komplexität der Stimuli ist, desto mehr ist die Höhe der Gedächtnisspanne reduziert. Eng mit dem Komplexitätseffekt hängt drittens der Effekt Irrelevanter Bilder zusammen (Quinn & McConnell, 1996b). Im Gegensatz zur Wirksamkeit Irrelevanter Sprache erzeugen Irrelevante Bilder nicht eine ebenso stabile Interferenzwirkung. Der vierte

Effekt bezieht sich auf die Möglichkeit der Blockierung des visuellen Kontrollprozesses (Baddeley & Logie, 1999).

Der Nachweis dieser Effekte darf nicht darüber hinwegtäuschen, daß die Befundlage sehr uneinheitlich ist und die Wirkung dieser Effekte bei weitem nicht die Stabilität und das Ausmaß der Störwirkung analoger Effekte bei phonologischem Material erreichen, das Thema des nächsten Kapitels ist.

2.4 DIE PHONOLOGISCHE SCHLEIFE

Der Phonologischen Schleife wird innerhalb des Arbeitsgedächtnisansatzes die Funktion und Verarbeitung sprachbasierter Informationen zugeschrieben. Zu diesem Bereich des Arbeitsgedächtnismodell liegen die meisten Forschungsbefunde vor. So wurde beispielsweise die Rolle der Phonologischen Schleife beim Spracherwerb (Baddeley et al., 1998) oder beim Lesen und Textverstehen (Gathercole & Baddeley, 1993) ausführlich untersucht.

Die Phonologische Schleife wird von Baddeley (1986) in zwei Komponenten zerlegt: in aktive Prozesse und in einen passiven Speicher, in dem die Itemrepräsentationen für circa zwei Sekunden bestehen, sofern kein Auffrischen der verblässenden Gedächtnisspur durch den Kontrollprozeß erfolgt. Für die Notwendigkeit dieses Aufbaus besteht experimentelle, neuropsychologische und physiologische Evidenz, die im folgenden vorgestellt wird.

Die experimentelle Bekräftigung des Postulats zweier Komponenten innerhalb des Konzepts der Phonologischen Schleife wurde unter anderem aus dem Forschungsgebiet zum Effekt Irrelevanter Sprache abgeleitet (Kapitel 2.4.2). Salamé und Baddeley (1982) wiesen nach, daß ein zu ignorierender Störschall die Lernleistung beeinträchtigt, obwohl die Versuchspersonen das Gehörte nicht selber aussprechen. Das bisherige Konzept einer unitären Artikulatorischen Schleife, in die nur artikuliertes Material gelangt, wurde modifiziert durch die bereits erwähnte Fraktionierung: in einen Phonologischen (nicht mehr rein artikulatorischen) Speicher gelangen Informationen durch aktive Wiederholung des Materials oder direkt durch auditive Wahrnehmung. Follow-up Studien bestätigen diese Annahme durch Nachweis dissoziativer Effekte von aktivem Artikulationsprozeß und passivem Speicher: der Phonologische Ähnlichkeitseffekt (Kapitel 2.4.3) bleibt trotz Artikulatorischer Unterdrückung (Kapitel 2.4.1) bei auditiver Präsentation erhalten. Bei visueller Präsentation dagegen wird dieser Effekt bei zusätzlicher artikulatorischer Sekundäraufgabe eliminiert (Baddeley, Lewis & Vallar,

1984). Innerhalb der Arbeitsgedächtnis-Theorie wird die Dissoziation folgendermaßen erklärt: Durch Belegung des Kontrollprozesses wird die Graphem-Phonem-Transformation inhibiert, so daß visuell präsentierte Stimuli nicht mehr in den Phonologischen Speicher gelangen können. Wenn die Items phonologisch nicht repräsentiert werden, verliert die phonologische Ähnlichkeit ihre Relevanz. Dies manifestiert sich in der erwähnten Elimination des Effekts Phonologischer Ähnlichkeit bei visueller Präsentation und Artikulatorischer Unterdrückung. Bei auditiver Darbietung dagegen werden direkt phonologische Repräsentationen gebildet, die anfällig gegenüber phonologischer Interferenz sind.

Auch neuropsychologische Einzelfallstudien sprechen für die Trennung der beiden Bestandteile der Phonologischen Schleife. Vallar und Baddeley (1984) berichten von einer Patientin, bei der (bleibt man bei Baddeleys Terminologie) der Phonologische Speicher geschädigt ist, die aktive Komponente dagegen intakt blieb. Im Sinne einer doppelten Dissoziation komplettiert eine Studie von Belleville et alii (1992) die neuropsychologische Evidenz: nach zahlreichen Experimenten an einem Patienten diagnostizierten sie einen intakten Speicher und defizitäre Kontrollprozesse.

Einen weiteren Hinweis auf die Gültigkeit der modifizierten Form der Phonologischen Schleife geben physiologische Studien: so konnte in PET-Studien gezeigt werden, daß aktive Kontrollprozesse andere Regionen aktivieren als Speicherprozesse. Beide Teile zeigen Veränderungen der Aktivierungsmuster in der linken Hemisphäre, jedoch werden bei Speicherfunktionen hintere Regionen aktiviert und bei den aktiven Prozessen vordere Strukturen (Jonides et al., 1996).

Es werden drei Zugangsmöglichkeiten zum Phonologischen Speicher angenommen: durch die auditive Wahrnehmung, durch den Kontrollprozeß sowie durch das Langzeitgedächtnis. Der erste Weg ist durch Studien mit Irrelevanter Sprache belegt, der zweite durch den Effekt Artikulatorischer Unterdrückung. Der dritte Zugang wurde generiert aus einem Experiment mit visueller Darbietung der Items und gleichzeitiger Artikulatorischer Unterdrückung. Selbst unter dieser Bedingung konnten zutreffende Urteile über den Reim der Wortpaare abgegeben werden (Baddeley, 1997). Auch neuere Forschungen und Ergebnisse zum Wortlängeneffekt legen eine direkte Verbindung zwischen Langzeitgedächtnisspeicher und Phonologischer Schleife nahe. Abbildung 3 veranschaulicht den postulierten Aufbau der Phonologischen Schleife und die Zugänge zum Speicher.

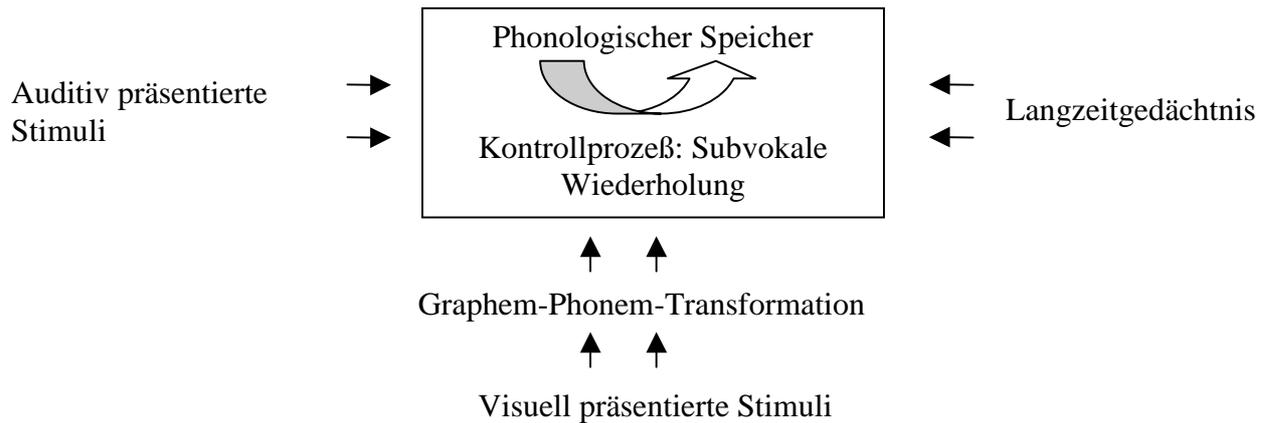


Abbildung 3 Zugangsmöglichkeiten zum Phonologischen Speicher

Die Phonologische Schleife wird im folgenden näher beschrieben anhand der Effekte der Artikulatorischen Unterdrückung, der Irrelevanten Sprache, der Phonologischen Ähnlichkeit und der Wortlänge.

2.4.1 DIE TECHNIK ARTIKULATORISCHER UNTERDRÜCKUNG

Unter Artikulatorischer Unterdrückung versteht man die iterierte Produktion eines irrelevanten Wortes oder einer Wortsequenz, die simultan zur Primäraufgabe ausgeführt wird. Die Phonetik dieses Wortes spielt keine Rolle (Baddeley, 1991), jedoch zeigen sich differentielle Effekte, je nachdem ob ein Wort oder eine Sequenz wiederholt werden soll (Seitz et al., 2000). Eine typische Wirkung der artikulatorischen Sekundäraufgabe ist die signifikante Reduktion der Kurzzeitgedächtnisspanne (Bredenkamp & Klein, 1998). Wie die Dissoziation der Ergebnismuster mit anderen Zweitaufgaben belegt, kann die Wirkung Artikulatorischer Unterdrückung nicht als Aufmerksamkeitseffekt interpretiert werden (Baddeley, 1991).

Für die Wirksamkeit der Suppression werden vier komplementäre Erklärungsmöglichkeiten angegeben. So wird eine Blockierung beider aktiven Kontrollprozesse, nämlich der subvokalen Wiederholung und Graphem-Phonem-Transformation, angenommen. Eine dritte Interpretation betont die Induktion von Interferenz im Phonologischen Speicher, äquivalent zu dem Effekt Irrelevanter Sprache. Ein vierter Effekt Artikulatorischer Unterdrückung zeigt sich in der für die Aktivierung von Repräsentationen im Langzeitgedächtnisspeicher benötigten Zeit.

Bei Artikulatorischer Unterdrückung während Präsentations- und Wiedergabephase wird der Wortlängeneffekt eliminiert (Baddeley, 1991). Dies spricht für eine Blockierung des

Kontrollprozesses: die Wiederauffrischung der Items wird verhindert, der Unterschied in der Schnelligkeit der Artikulation zur Bewahrung der Items im Phonologischen Speicher kann nicht mehr zum Tragen kommen (Kapitel 2.3.4). Zudem existiert neuropsychologische Evidenz für diese Annahme: bei Patienten, die in der Primäraufgabe die Strategie der Wiederholung nicht anwenden, zeigen sich keine Leistungsreduktionen, wenn simultan artikulatorisch unterdrückt wird (Vallar & Baddeley, 1984). Die zweite potentielle Wirkungsweise der Suppression bezieht sich auf die Blockierung der Graphem-Phonem-Übersetzung: bei visuell präsentem Material und zusätzlicher Unterdrückung kann kein Effekt Irrelevanter Sprache nachgewiesen werden (Salamé & Baddeley, 1982). Die Interferenz im Phonologischen Speicher zwischen Stör- und Lernsequenz kann nicht entstehen, weil das Lernmaterial bei visueller Darbietung und Suppression nicht in einen phonologischen Code transformiert werden kann. Bei auditiver Präsentation der Stimuli, simultaner Suppression und Irrelevanter Sprache dagegen, wird trotz Artikulatorischer Unterdrückung der Effekt Irrelevanter Sprache nachgewiesen (Salamé & Baddeley, 1982). Die Wirkung des selbst-produzierten Sprachschalls als Irrelevante Sprache wurde nur theoretisch als mögliche zusätzliche Störvariable generiert, nicht aber empirisch überprüft (Baddeley, 1986). Über einen vierten Effekt Artikulatorischer Unterdrückung wird innerhalb des Modells der Invarianzhypothesen (Kapitel 3) berichtet: Suppression verlangsamt signifikant die Verarbeitungszeit im Langzeitgedächtnisspeicher (Bredenkamp & Klein, 1998), deren Varianz zum Großteil durch den Kontrollprozeß aufgeklärt werden kann (Bredenkamp & Hamm, 2001).

2.4.2 DER EFFEKT IRRELEVANTER SPRACHE

Der Effekt Irrelevanter Sprache (im folgenden ISE) ist durch zwei Merkmale definiert: Während der Durchführung einer Primäraufgabe, beispielsweise einer Lernaufgabe, werden simultan auditive Stimuli präsentiert. Zusätzlich werden die Versuchspersonen per instructionem dazu angehalten, den Sprachklang zu ignorieren. Trotz dieser Instruktion evoziert Irrelevante Sprache signifikante Leistungsverschlechterungen in Gedächtnistests mit serieller Reproduktion (Colle & Welsh, 1976). Analog der Forschungsdiskussion wird im folgenden der Schwerpunkt auf die Erklärungsansätze von Baddeley und Jones gelegt.

Um einen ISE zu provozieren, muß die Primäraufgabe notwendig drei Charakteristika aufweisen. Erstens müssen Gedächtnisprozesse beansprucht werden. Irrelevante Sprache stört nur komplexe kognitive Aufgaben, wie beispielsweise Rechnen (Banbury & Berry, 1998), nicht jedoch einen simplen Vergleich von Buchstabenpaaren (Salamé & Baddeley, 1986). Zweitens muß die Phonologische Schleife in die Aufgabenlösung involviert sein.

Irrelevante Sprache stört ausschließlich die Arbeit der Phonologischen Schleife, nicht jedoch Lernen von räumlich-visuellem Material, dessen Bewältigung durch das visuell-räumliche Arbeitsgedächtnis angenommen wird (Klatte & Hellbrück, 1993). Drittens muß in der Primäraufgabe die Reihenfolge der Stimuli eine Rolle spielen: nur serielle Reproduktion wird durch Irrelevante Sprache signifikant gestört, nicht jedoch freie Wiedergabe (Beaman & Jones, 1998).

Der ISE kann durch verschiedene Materialien des Störschalls evoziert werden:

- Unverständliche Fremdsprache (Colle & Welsh, 1976)
- Lange und kurze Worte (Salamé & Baddeley, 1982)
- Nonworte und Worte (Buchner, Irmen & Erdfelder, 1996)
- Vokalmusik (Salamé & Baddeley, 1989).

Aus der umfangreichen Forschung kristallisieren sich drei Eigenschaften des ISE heraus. Erstens ist der ISE unabhängig von physikalischen Momenten. Ein Sprecherwechsel innerhalb des Störschalls erzeugt keine zusätzliche Störwirkung (Jones, 1993). Dasselbe gilt für Lokalisation und Lautstärke des Schalls (Colle, 1980). Ein zweiter Punkt ist die Unabhängigkeit des ISE von der Semantik der Irrelevanten Störreize. Drittens, eine Habituation an Irrelevante Sprache ist weder durch die Dauer der Exposition (Ellermeier & Zimmer, 1997) noch durch eine massierte Anzahl der Durchgänge (Tremblay & Jones, 1998) möglich. Allerdings zeigen sich große interindividuelle Unterschiede in der Resistenz gegenüber dem Störschall (Hamm, 1998), die zeitlich stabil sind (Ellermeier & Zimmer, 1997).

Der ISE ist kein Effekt visueller Enkodierung (Broadbent, 1983), wie Colle und Welsh (1976), respektive Miles, Jones und Madden (1991) durch systematische experimentelle Variation belegen. Erklärungsversuche des ISE als Phänomen eines Orientierungsreflexes oder als Effekt der Zentralen Exekutive konnten nicht durch empirische Daten gestützt werden (Salamé & Baddeley, 1982). Beispielsweise verschwand der ISE bei visueller Darbietung der Items und Artikulatorischer Unterdrückung (Salamé & Baddeley, 1982). Da es nicht einsichtig ist, warum unter Suppression der Störschall leichter zu ignorieren ist, kann die Annahme des ISE als Aufmerksamkeitseffekt vernachlässigt werden.

Innerhalb des Rahmenmodells des Arbeitsgedächtnisses wird der ISE folgendermaßen gedeutet: Der störende Sprachschall gelangt über einen auf Periodik sensitiven Filter (Salamé & Baddeley, 1989) direkt in den Phonologischen Speicher, in dem sich auch das Lernmaterial befindet. Die zu ignorierenden und die relevanten Stimuli interferieren. Dies manifestiert sich in reduzierten Leistungen im Gedächtnistest. Nach Baddeleys Theorie ist

der entscheidende Faktor für die Höhe der Störung der Grad phonologischer Ähnlichkeit zwischen Irrelevanter Sprache und Lernstimuli: je höher die Ähnlichkeit ist, desto höher ist die Störwirkung (Salamé & Baddeley, 1982).

Sowohl die Annahme eines Filters als auch die Ähnlichkeit zwischen Stör- und Lernsequenz wurden in Frage gestellt:

1. Besteht der Störschall lediglich aus einer einzigen Silbe, tritt der ISE nicht auf (Jones & Macken, 1995b). Silben sind per se sprachliches Material und haben Zugang über den Filter in den Phonologischen Speicher. Der Nachweis gelang dann und nur dann, wenn mehr als eine Silbe dargeboten wurde (Tremblay & Jones, 1998). Die Dissoziation, warum mehrere Silben den Filter passieren und Interferenz erzeugen, eine Silbe allein dagegen nicht, kann durch die Theorie von Baddeley und Salamé nicht erklärt werden. Sprache scheint also keine hinreichende Bedingung für die Entstehung des ISE zu sein. Kritisch muß angemerkt werden, daß der beschriebene Befund von Jones und Macken (1995b) von LeCompte (1995) nicht repliziert werden konnte.
2. Eine Fülle von Studien belegt, daß auch Töne den ISE erzeugen (Spitzbarth, 2000): Sprache ist für diesen Effekt also weder eine hinreichende noch eine notwendige Bedingung. Nach Salamés und Baddeleys (1982) Erklärung dürften Töne keine Interferenz erzeugen, da sie keine Ähnlichkeit zum Material der Primäraufgabe aufweisen. Aus der Störwirkung der Töne leiten Jones und Macken (1995a) die Äquipotenzhypothese sprachlicher und nicht-sprachlicher Schalle ab: Unterschiedliche Frequenzen bei Tönen sollen eine zu verschiedenen Buchstaben im Störschall äquivalente Wirkung verursachen (Jones & Macken, 1995a). Dies konnte von Spitzbarth (2000) nicht repliziert werden. In verschiedenen Experimenten konnte sie Störeinflüsse von Tönen nachweisen, die Höhe der Leistungsreduktion durch irrelevante Töne erreichte jedoch niemals das Niveau des ISE. Sprache störte signifikant stärker.
3. Das Experiment von Salamé und Baddeley (1982, Experiment fünf) zum Nachweis des Einflusses der Ähnlichkeit zwischen Stör- und Lernmaterial konnte von LeCompte und Shaibe (1997) nicht repliziert werden. Bridges und Jones (1996) stellen eine Konfundierung der Ähnlichkeit in der Studie von 1982 mit der Variable der Wortdichte (Word-Dose) im Störschall fest. Das Postulat dieses Word-Dose-Effects besagt, daß analog zur Höhe der Änderungsrate in der Störsequenz der Grad der Störung ansteigt (Bridges & Jones, 1996). Diesem Effekt steht der Befund gegenüber, daß die Störsequenz mit dem Muster ABCABC keinen anderen Effekt erzielt als eine alternative Reihenfolge AABCC (Tremblay & Jones, 1998). Auch wenn in Experiment fünf von

Salamé und Baddeley (1982) eine Konfundierung mit der Variable Wortdichte vorliegt, so ist deren Bedeutung nicht so eindeutig, wie von Bridges und Jones deklariert.

4. Gegen das Argument der Ähnlichkeit der Materialien spricht der Nachweis, daß kein Unterschied bezüglich der Störwirkung entsteht, wenn die Fremdsprache im Störschall vorwärts oder rückwärts abgespielt wird (Jones, 1993). Dieser Effekt konnte in zwei unabhängigen Studien nicht repliziert werden (LeCompte, Neely & Wilson, 1997; Hamm, 1998). Ebenso zeigten sich differentielle Effekte, je nachdem ob die störende Fremdsprache dem Lernmaterial ähnlich oder unähnlich war (Bredenkamp & Hamm, 2001). Die Ansicht von Jones (1993), daß die Ähnlichkeit zwischen Lern- und Störstimuli keinerlei Relevanz für die Entstehung des ISE haben, scheint angesichts der hier referierten Befunde nicht haltbar zu sein.

Als Alternative zu dem Erklärungsansatz von Baddeley und Salamé schlägt Jones (1993) das Object-Oriented Episodic Record Modell (O-OER) vor. Postuliert wird ein unitäres, amodales Gedächtnissystem („Blackbord“), auf dem alle Items als stabile Objekte eingetragen werden, die kaum Vergessensprozessen ausgesetzt sind. Zwischen den Objekten werden durch das zeitliche Sucedere des Sprachstroms, respektive durch gestalttheoretische Prinzipien bei visueller Darbietung, Verbindungen zwischen den Objekten angelegt, die sowohl dem Spurverfall als auch Interferenz ausgesetzt sind. Der Kontrollprozeß fungiert zur Herstellung der Verknüpfung und zur Wiederauffrischung der verfallenden Verbindung. Den ISE erklärt Jones durch die Changing-State-Hypothese. Als Changing-State wurde experimentell die Variation physikalischer Merkmale definiert, wie beispielsweise Frequenzänderung, Buchstaben- oder Tonhöhenwechsel (Tremblay & Jones, 1998). Die Lernsequenz und die gehörte Störsequenz werden beide als Ereignisketten zwischen den einzelnen Objekten auf dem Blackbord präsentiert. Bei Überkreuzen der Objektverbindungen entsteht Interferenz zwischen den Objektverbindungen. Abbildung 4 stellt graphisch die Kernaussage des Changing-State-Ansatzes dar. Die Lernsequenz ist durch die Buchstaben LS gekennzeichnet, die Irrelevante Sprache durch IS. Die Verbindungen innerhalb des Sprachstroms werden automatisch durch perzeptuelle Informationsverarbeitung gebildet. Je nach Wahrnehmung wird das Gedächtnis organisiert (Jones, Alford, Bridges, Tremblay & Macken, 1999). Dabei zeigt sich, daß eine entscheidende Variable der Genese des ISE die Unähnlichkeit innerhalb des Sprachstroms ist. Je unähnlicher IS_1 bis IS_2 sind, desto höher ist der Störeinfluß. Je weniger Unterschiede wahrgenommen werden, desto undeutlicher sind die Objektverbindungen und desto geringer ist die Störwirkung (Jones & Macken, 1995c). Wird beispielsweise die Irrelevante Sprache als ein Objekt wahrgenommen, so erfolgt eine Gedächtnisrepräsentation als Steady-State,

die keine Interferenz zu den Lernstimuli erzeugen kann. Ein Steady-State Ereignis ist mit der Kennzeichnung SS in Abbildung 4 eingezeichnet.

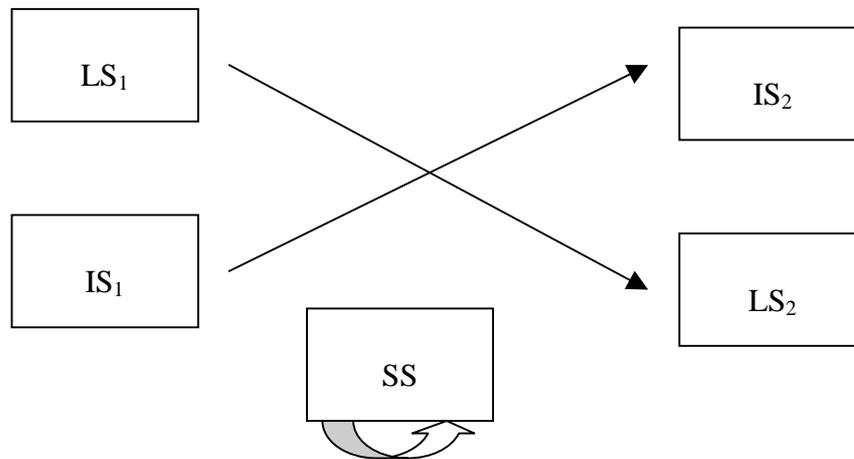


Abbildung 4 Darstellung von Changing-State und Steady-State

Als empirische Evidenz für sein Modell führt Jones die Studien zum Steady-State an, zur Äquipotenzhypothese, zum Word-Dose-Effekt und weitere Untersuchungen, die belegen, daß die Leistungen bei freier Wiedergabe nicht durch Irrelevante Sprache beeinflusst wird (Beaman & Jones, 1998). Wie bereits ausgeführt, werden alle diese Studien in der Forschungsdiskussion kritisiert, so daß die Changing-State-Hypothese zur alleinigen Erklärung des ISE ebensowenig geeignet erscheint wie die Hypothese der Ähnlichkeit der Materialien von Baddeley und Salamé. Das O-OER Modell wird zudem wegen seiner argumentativen Zirkularität angegriffen und bezüglich der Annahme der Amodalität falsifiziert (Meiser, 1997). Auch das Feature-Modell von Neath und Nairne (1995) wird zur Erklärung des ISE diskutiert. Allerdings zeigen sich hier ebenfalls Probleme (Spitzbarth, 2000). Bis heute scheint kein Gedächtnismodell den ISE adäquat erklären zu können.

2.4.3 DER PHONOLOGISCHE ÄHNLICHKEITSEFFEKT

Charakteristikum des Phonologischen Ähnlichkeitseffekts ist die signifikante Leistungsreduktion in einem unmittelbar folgenden Gedächtnistest beim Erlernen von phonologisch ähnlichem Material im Gegensatz zu distinkten Items (Conrad, 1964; Conrad & Hull, 1964). Für visuelle, respektive semantische Ähnlichkeit konnte kein äquivalenter Effekt nachgewiesen werden (Baddeley, Lewis & Vallar, 1984). Die Wirkung ähnlicher Phonologie kann nicht als akustisches Problem interpretiert werden. Der Effekt tritt auch bei visueller Darbietung der Items auf, die sich graphemisch unähnlich sind und bei tauben

Kindern, die von ihren Lehrern als gute Sprecher eingestuft wurden (Baddeley, 1986). Innerhalb des Arbeitsgedächtnis-Modells wird der Phonologische Ähnlichkeitseffekt durch Interferenz in der Phonologischen Schleife erklärt. Die Repräsentationen ähnlich klingender Items im phonologischen Speicher sind wegen der Ähnlichkeit der akustischen Codes schwieriger diskriminierbar. Als Folge der Interferenz sinkt die Gedächtnisleistung. Diese These stützt sich auf experimentelle Evidenz: wie bereits berichtet, wird dieser Effekt bei zusätzlicher Artikulatorischer Suppression nur unter der Bedingung des visuellen Darbietungsmodus eliminiert (Baddeley, Lewis & Vallar, 1984). Zudem ist bei phonologisch ähnlichem Material kein Effekt Irrelevanter Sprache nachweisbar (Colle & Welsh, 1976): Da Irrelevante Sprache als Induktor von Interferenz im Speicher interpretiert wird, steht das Ausbleiben der Wirkung eines Störschalls bei ähnlich klingenden – interferierenden - Items in Kongruenz mit der Auslegung des Phonologischen Ähnlichkeitseffekts als Provokation von Interferenz im Speicher. Ein ungelöstes Problem ist derzeit die absolute Berechnung der Itemähnlichkeit, die meist nur intuitiv als graduelle Abstufung relativ zum Vergleichsmaterial bestimmt wird.

Schweickert, Guentert und Hersberger (1990) verbinden die Forschungszweige der Effekte Phonologischer Ähnlichkeit und Wortlänge (Kapitel 2.4.4): die Persistenz der Gedächtnisspur für phonologisch ähnliche Stimuli ist kürzer als bei distinktem Material. Itemrepräsentationen im Phonologischen Speicher verfallen um so schneller, je mehr sie sich ähneln. Diese Studie greift fundamental das Postulat einer invarianten Dauer der Gedächtnisspur an, ist aber kongruent mit dem Baddeleyschen Ansatz zur Erklärung des Effekts Phonologischer Ähnlichkeit: die Verursachung dieses Effekts wird in dem Phonologischen Speicher, nicht jedoch im Kontrollprozeß vermutet.

2.4.4 DER WORTLÄNGENEFFEKT

Der Wortlängeneffekt (im folgenden WLE) bezeichnet das Phänomen, daß die Höhe der Gedächtnisspanne bei serieller Reproduktion von Items mit schneller Artikulationszeit signifikant höher ist als bei langsamer Aussprache. Die offene Artikulation wird als Indikator der für die subvokale Wiederholung benötigten Zeit interpretiert (Baddeley, Thomson & Buchanan, 1975). Der WLE tritt nur bei serieller, nicht jedoch bei freiem Wiedergabemodus auf, wie experimentelle (Avons, Wright & Pammer, 1994), neuropsychologische (Romani, 1992) und entwicklungspsychologische (Henry, 1991) Evidenz belegt. Da der WLE direkt zu den Zielen der eigenen Untersuchung (Kapitel 4) relatiert ist, werden im folgenden der Baddeleysche Erklärungsansatz und zwei Modifikationen ausführlich dargestellt.

Baddeley et alii (1975; Baddeley, 1986) legen der Erklärung des WLE vier Annahmen zugrunde:

1. Die Items sind im Phonologischen Speicher einem Zerfallsprozeß ausgesetzt.
2. Die Dauer der Gedächtnisspur im Phonologischen Speicher ist konstant.
3. Die Items können vor dem Zerfall durch Wiederauffrischung per Repetition bewahrt werden.
4. Die Artikulationszeit pro Item ist eine adäquate Operationalisierung der inneren Kontrollprozesse. Diese Wiederholung ist schnell, benötigt kaum Aufmerksamkeit und ist unabhängig von der Listenlänge (Baddeley, Lewis & Vallar, 1984). Im folgenden wird durchgehend der Parameter der Artikulationszeit pro Item beschrieben, allein für die Abbildungen erfolgt eine Umrechnung in die Artikulationsrate als reziproker Artikulationszeit.

Nach Baddeley ist der WLE ein Epiphänomen des Kontrollprozesses der Phonologischen Schleife, der durch die Artikulationsdauer der Items determiniert wird: Ist die Artikulationszeit kurz, können viele Items erinnert werden, respektive bei langsamer Aussprache werden mehr Items vergessen. Der WLE ist nicht modalitätsspezifisch und kann sowohl für auditive als auch für visuelle Itempräsentation nachgewiesen werden (Baddeley et al., 1975). Abbildung 5 veranschaulicht den linearen Zusammenhang zwischen Artikulationsrate pro Item und Höhe der Gedächtnisspanne.

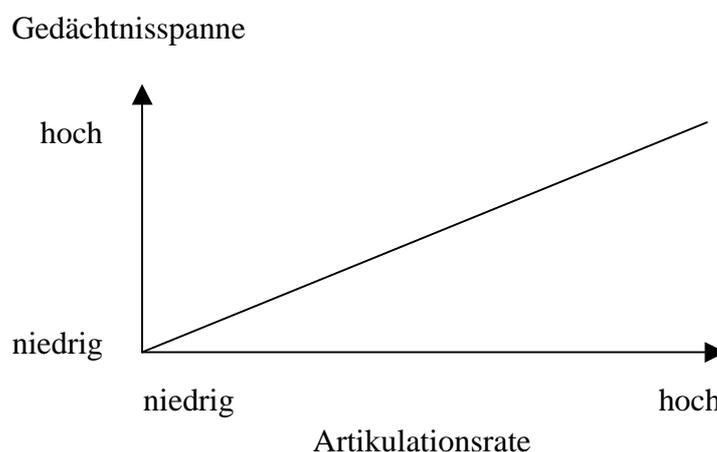


Abbildung 5 Darstellung des Wortlängeneffekts

Sechs Untersuchungsparadigmen, die im folgenden näher beschrieben werden, wurden in einer Vielzahl von Studien realisiert: Variation des Stimulusmaterials, der Stichprobenkohorte bezüglich des Alters, neuropsychologischer Diagnosen oder der

Muttersprache und zwei Methoden zur Elimination des Effekts durch schwer verbalisierbares Material, respektive durch zusätzliche Artikulatorische Unterdrückung.

Wortlängeneffekt und Stimulusmaterial Verschiedene Materialarten bieten sich zur Untersuchung des WLE an. Da es mehr Zeit beansprucht mehrsilbige Worte zu artikulieren als einsilbige, können bei konstanter Dauer der Gedächtnisspur weniger mehrsilbige Items vor dem Verfall bewahrt werden: die Höhe der Gedächtnisspanne sinkt. Eine Reihe von Experimenten prüfte systematisch diese Hypothese (Baddeley et al., 1975). Dabei zeigte sich, daß die relevante Determinante des WLE nicht per se die Anzahl der Silben ist, sondern die Dauer der Artikulationszeit. Die Gedächtnisspanne für zweisilbige Worte mit kurzer Artikulationszeit war höher als bei zweisilbigen Substantiven mit längerer Aussprechdauer. Die Artikulationsdauer erwies sich intraindividuell als stabil und interindividuell als variabel. Die individuelle Lernleistung kann durch die Anzahl der Targets vorhergesagt werden, die innerhalb von 1.6 Sekunden gelesen werden können, respektive innerhalb von 1.3 Sekunden artikulierbar sind (Baddeley et al., 1975). Ebenso wurde eine interexperimentelle Stabilität nachgewiesen: Gedächtnisspanne und Leserate korrelieren über die Probanden hinweg. Personen, die schnell lesen, haben auch eine höhere Gedächtnisspanne (Baddeley et al., 1975). Hoosain (1982) gibt für die Korrelation zwischen Artikulationszeit und Gedächtnisspanne bei chinesischen Probanden einen hoch signifikanten Wert von $-.66$ an, analog bei einer englischsprachigen Stichprobe von $-.70$. Der Vergleich verschiedener Materialien ist in der letzten Zeit stark kritisiert worden (Service, 1998), da bei Änderung des Stimulusmaterials auch Konfundierungen mit anderen Faktoren als der Artikulationszeit vorliegen, beispielsweise mit der Wortkomplexität.

Wortlängeneffekt und Alterskohorten Die Ergebnisse der Studie von Hasselhorn (1986) bezüglich des Zusammenhangs zwischen Artikulationszeit und Gedächtnisspanne im Verlauf der Lebensspanne sprechen insgesamt für die Gültigkeit der Hypothese von Baddeley: über alle Altersstufen hinweg konnte eine signifikante, negative Korrelation zwischen Gedächtnisspanne und Artikulationszeit nachgewiesen werden. Erwartungskonform stiegen die Gedächtnisleistungen und Artikulationsgeschwindigkeiten bis in das junge Erwachsenenalter an und fielen danach deutlich wieder ab. Die Dauer der Gedächtnisspur blieb statistisch altersinvariant trotz des deskriptiven Trends, daß sie kontinuierlich mit dem Alter abnimmt (Hasselhorn, 1986, Tabelle 1, S. 330). Diese Studie muß um eine weitere Arbeit ergänzt werden: während Hasselhorn Kinder ab sechs Jahren untersuchte, wählte Henry (1991) vierjährige Probanden als Versuchspersonen. Auch sie konnte den WLE nachweisen, obwohl angenommen wird, daß Kinder dieses Alters noch nicht innerlich wiederholen. Dies spricht gegen den Ansatz von Baddeley und verweist auf

die Relevanz von Faktoren außerhalb der Zeit subvokaler Wiederholung, beispielsweise auf Prozesse während der Wiedergabe.

Wortlängeneffekt und neuropsychologische Einzelfalluntersuchungen Die Forschungsdiskussion um den aktiven Kontrollprozeß umfaßt auch neuropsychologische Studien. Nachweislich tritt bei Patienten mit Dysarthrie der WLE auf (Vallar & Cappa, 1987). Der artikulatorische Kontrollprozeß ist also unabhängig von muskulärem Feedback bei der Artikulation. Patienten mit Sprachapraxie dagegen zeigen keinen WLE (Waters, Rochon & Caplan, 1992). Als Determinante wird nicht die Artikulation per se angesehen, sondern die Wortplanung. Diese These wird durch experimentelle Evidenz an gesunden Probanden untermauert (Caplan, Rochon & Waters, 1992): Wenn sich die Wortmaterialien nicht in Silben- oder Phonemanzahl unterscheiden, sondern nur in der Artikulationszeit der Vokale, dann wird der WLE eliminiert. Dieses Ergebnis steht in Dissonanz zu zwei fundamentalen Annahmen Baddeleys: Innere und offene Artikulation scheinen, erstens, nicht äquivalent zu sein. Möglicherweise ist die Artikulationszeit ein Indikator für einen generellen Prozeß: Smyth und Scholey (1996) wiesen nach, daß durch die Schnelligkeit der Artikulation auch visuell-räumliche Leistungen gut vorausgesagt werden können. Eine alternative Möglichkeit wäre die Fraktionierung der Phonologischen Schleife in einen Input-Speicher, welcher sensitiv gegenüber der Sprachrezeption ist, und einen Output-Speicher der Sprachproduktion (Burani, Vallar & Bottoni, 1991). Der WLE wäre dann ein Effekt des Wiedergabespeichers: lange Worte benötigen mehr Ressourcen für die Sprachproduktion (Monsell, 1987). Zweitens ist die Artikulationszeit anscheinend nicht die (einzig) relevante Variable bei der Provokation des WLE, wie im Anschluß an die Darstellung der sechs Untersuchungsparadigmen dargelegt wird.

Wortlängeneffekt und linguistische Gruppe Im strengen Sinne ist die Untersuchung an Zahlenmaterial in verschiedenen Sprachen eine Spezialform der Materialmanipulation, weil die Benennung der Zahlen in den Sprachen variiert. Da aber das präsentierte Stimulusmaterial identisch bleibt und deshalb weniger Konfundierungen mit weiteren Faktoren als bei der originalen Stimulusvariation vorliegen, wird der Vergleich linguistischer Gruppen als eigenständiger Ansatz behandelt. Um den Störfaktor interindividueller Varianz zu eliminieren, wurden in einer ersten Periode der Erforschung sprachabhängiger Gedächtnisunterschiede bilinguale Probanden untersucht. Es zeigte sich durchweg, daß die Merkleistung in der Sprache besser war, in der auch die Artikulationsgeschwindigkeit höher war. Da die sprachabhängigen Differenzen sehr stabil waren und auch bei Sprechern auftraten, bei denen eine Sprache dominant ist (Chincotta & Hoosain, 1995), wurden in einer zweiten Untersuchungsperiode die Sprachen

interindividuell variiert. Dies birgt den Vorteil, daß statt paarweisen Vergleichen Relationen über mehrere Sprachen aufgestellt werden können. Aus der Flut der Studien zu sprachbasierten Unterschieden in Artikulationszeit und Gedächtnisspanne (beispielsweise Naveh-Benjamin & Ayres, 1986; Ellis & Hennelly, 1980) werden zwei Ergebnisse, die für die vorliegende Arbeit relevant sind, genannt. Chincotta & Underwood (1997a) zeigten, daß Gedächtnisspanne und Artikulationsgeschwindigkeit bei chinesischen Probanden höher sind als bei spanischen Probanden. Lüer et al. (1998) belegten, daß chinesische Versuchspersonen bezüglich dieser beiden Parameter auch deutsche Probanden übertrafen.

In Tabelle 1 werden aus verschiedenen Studien die empirischen Werte von Artikulationszeit (in Millisekunden) und Gedächtnisspanne für Zahlenmaterial verschiedener Sprachgruppen (Chinesisch, Deutsch und Spanisch) vergleichend gegenübergestellt, die auch in der hier vorliegenden Arbeit untersucht werden (Kapitel 5.3). Die absolute Vergleichbarkeit ist allerdings wegen verschiedener Erhebungs- und Auswertungsmethoden stark eingeschränkt. Dies gilt besonders für Definition und Erfassung der Artikulationszeit (Chincotta & Underwood, 1997b), weniger für die Ermittlung der Gedächtnisspanne (Puckett & Kausler, 1984).

Tabelle 1 Empirische Werte der Artikulationszeit und Gedächtnisspanne an exemplarisch ausgewählten Sprachgruppen.

	Chinesisch ²	Deutsch	Spanisch ³
Gedächtnisspanne	7.33	6.22	6.50
Artikulationszeit	236	363	439

Ein noch ungeklärtes Phänomen wurde von Chincotta und Underwood (1997b) beschrieben: Es ergaben sich unterschiedliche Gedächtnisleistungen und eine Interaktion zwischen der Sekundäraufgabe Artikulatorischer Unterdrückung und der untersuchten Sprache, wenn der Präsentationsmodus (Symbole versus Worte) der Zahlen variiert.

Wortlängeneffekt und schwer verbalisierbares Material Die beiden letzten vorgestellten Untersuchungsparadigmen zielen auf die Elimination des Wortlängeneffekts ab: Eine Arbeitsgruppe aus Göttingen (Lüer et al., 1998) zeigte, daß die Überlegenheit der chinesischen Probanden gegenüber deutschen Versuchspersonen bezüglich Artikulationszeit und Gedächtnisspanne nicht universell, sondern materialspezifisch ist: bei Verwendung von

² Werte entnommen für das Chinesische und das Deutsche aus Lüer et alii (1998).

³ Werte entnommen aus Chincotta und Hoosain (1995).

schwer oder nicht verbalisierbarem Material werden zwischen den Sprachen keine signifikanten Unterschiede mehr in der Gedächtnisspanne nachgewiesen.

Wortlängeneffekt und Artikulatorische Unterdrückung Die Intention dieser Untersuchungsmethode gleicht der vorigen. Es wird erwartet, daß bei Reduktion des Einflusses der subvokalen Wiederholung keine Unterschiede zwischen verschiedenen Materialarten, Kohorten oder Sprachgruppen auftreten. Baddeley et alii (1975) belegten experimentell diese Erwartung für verschiedene Materialarten. Allerdings wird der WLE bei auditiver Präsentation nur dann eliminiert, wenn während Enkodierungs- und Wiedergabephase artikulatorisch unterdrückt wird (Baddeley, Lewis & Vallar, 1984). Dies wird darauf zurückgeführt, daß der subvokale Artikulationsprozeß ein so automatisierter und wohlgelernter Prozeß ist, daß er sogar noch während der Wiedergabe eingesetzt werden kann. Chincotta und Underwood (1997b) konnten die sprachbasierten Differenzen durch zusätzliche Artikulatorische Unterdrückung eliminieren. Allerdings zeigte sich dabei die bereits erwähnte Abhängigkeit von der Präsentationsart.

Wie die Erläuterungen zu den Untersuchungsmethoden zeigen, wird der Erklärungsansatz von Baddeley, der als einzige Variable zur Determination der Gedächtnisspanne die Artikulationszeit berücksichtigt, kontrovers diskutiert. In den letzten Jahren zeigte sich ein Impetus an Studien, der der Annahme einer alleinigen Provokation des WLE durch die Artikulationszeit widerspricht. Im Sinne einer maximalen Spezifizierung der Bezugsklasse (Hempel, 1977) müssen die folgenden als relevant identifizierten Einflußfaktoren ausgeschlossen werden, damit der von Baddeley postulierte Erklärungsansatz Gültigkeit behält. Die Literaturangabe bezieht sich beispielhaft auf je eine Studie pro Variable.

- Phonologische Ähnlichkeit (Schweickert et al., 1990).
- Graphemische Komplexität (Zhang & Feng, 1990).
- Lexikalität (Hulme, Maughan & Brown, 1991).
- Sprechplanung (Caplan et al., 1992).
- Frequenz (Hulme et al., 1997).
- Präsentationsmodus (Chincotta & Underwood, 1997b).
- Wortkomplexität (Service, 1998).
- Konkretheit (Walker & Hulme, 1999).
- Informationsgehalt, definiert durch die Größe des Itempools (Kapitel 3, Bredenkamp & Hamm, 2001).
- Assoziationsgehalt (Lovatt, Avons & Masterson, 2000).

Alle Studien eint ein Faktum: in jedem der zitierten Artikel wird nachgewiesen, daß bei ausschließlicher Berücksichtigung der Artikulationszeit die Gedächtnisspur nicht als konstant angesehen werden kann. Im folgenden werden zwei Modifikationen des Interpretationsansatzes von Baddeley vorgestellt, die den hier aufgelisteten kritischen Befunden eher gerecht werden.

ERSTE MODIFIKATION: VERGESSEN ALS SPURVERFALL UND INTERFERENZ

Der Ansatz von Doshier und Ma (1998) bezieht sich auf die Verbindung zwischen schneller Artikulation und hoher Gedächtnisspanne. In Abbildung 6 entspricht dies dem letzten Drittel auf der x-Achse. Meines Erachtens sprechen die oben aufgelisteten Befunde gegen Baddeleys Annahme der Linearität: bei vielen, schnell artikulierbaren Items entsteht Interferenz, so daß der Steigungsparameter mit sinkender Artikulationszeit immer geringer wird und nicht konstant bleibt.

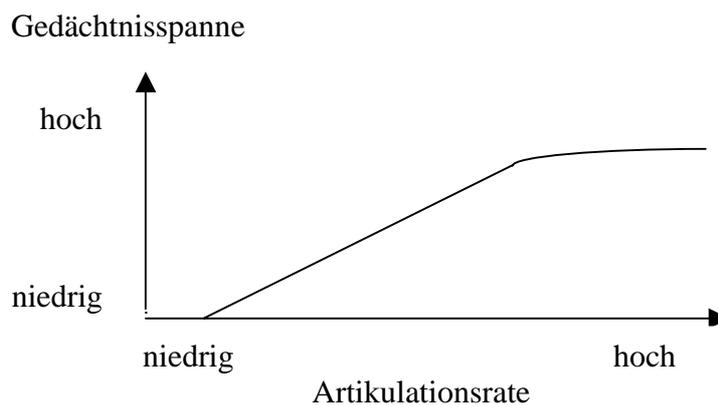


Abbildung 6 Modifikation des Erklärungsansatzes von Baddeley durch Berücksichtigung von Interferenzprozessen.

Zum Nachweis der Bedeutsamkeit von Interferenzprozessen während der Reproduktionsphase wird exemplarisch eine Studie von Cowan et alii (1992) erwähnt. Das experimentelle Design (Experiment 3) sieht im Gegensatz zu den bisher erwähnten Studien keine homogenen Lernsequenzen vor, sondern gesplittete Listen: die Reihe besteht in der ersten Hälfte aus kurzen Items und in der zweiten Hälfte aus langen Stimuli oder vice versa. In Verbindung mit den Versuchsbedingungen serieller Wiedergabe der Gedächtnisspanne, respektive Wiedergabe der Sequenz in umgekehrter Reihenfolge, werden direkt Interferenztheorie und Spurverfalltheorie gegeneinander getestet. Da innere Wiederholung immer vorwärts geht, weil die Probanden erst nach der Präsentationsphase über den Wiedergabemodus informiert werden, wird aus der Spurverfalltheorie die Prognose abgeleitet, daß immer die Länge der präsentierten ersten Hälfte die relevante Determinante

ist. Vorhergesagt wird, daß bei Wiedergabe der umgekehrten Reihenfolge die Itemkonstruktion kurz/lang, ausgesprochen lang/kurz, höhere Gedächtnisspannenwerte evoziert als die konträre Bedingung. Die Interferenztheorie dagegen sieht die erste Hälfte der auszusprechenden Items als relevante Variable an, da Interferenz bei prononcierten Stimuli niemals retrograd wirken kann. Analog zu Cowan et al. (1997) sowie zu Brown und Hulme (1995) wird postuliert, daß kurze Items störanfälliger sind, ihrerseits aber weniger Interferenz erzeugen. Die aus der Interferenztheorie deduzierte Prognose besagt, daß bessere Leistungen bei dem Itemmaterial lang/kurz, ausgesprochen kurz/lang, zu erwarten sind. Die empirischen Daten sprechen für die Interferenztheorie und belegen somit eindrucksvoll die Bedeutsamkeit von Prozessen während der Wiedergabephase. Ein Ansatz von Gathercole und Hitch (1993) kann meines Erachtens auch für die Interferenztheorie diskutiert werden: Interferenz entsteht, wenn ein Item aus dem Speicher aktiviert und in den Kontrollprozeß involviert wird. Wenn Interferenz nicht auf das ganze Wort an sich wirkt, sondern auf Wortfragmente, dann ist die Wahrscheinlichkeit eines Übertragungsfehlers bei mehrsilbigen Worten höher als bei einsilbigen, die nur aus einem Segment bestehen.

Cowan et alii (1997) zeigen, daß zweisilbige Worte eine höhere Gedächtnisspanne evozieren als einsilbige Worte gleicher Artikulationsdauer. Erklärt wird der sogenannte Wortkomplexitätseffekt durch steigende Resistenz gegen Interferenz mit Zunahme der Komplexität: die Effizienz der Enkodierung steigt mit morphologischer Komplexität. Eine Re-Analyse der Daten von Baddeley et al. (1975) zeigte, daß Wortkomplexität (im Sinne der Anzahl von Betonungen) und Artikulationszeit in der Studie zum Materialvergleich unter Variation der Artikulationszeit konfundiert waren. Leider ist das Problem einer Definition von Wortkomplexität noch nicht gelöst: „We make no attempt here to distinguish between the number of syllables and the number of stresses in defining word complexity; that is left for future work“ (Cowan et al., 1997, S.292).

Dosher und Ma (1998) operationalisieren die Vergessensprozesse durch Messung der Wiedergabezeit der gesamten Reproduktion anstatt der Artikulationszeit. Die Reproduktionszeit umschreibt in einem Gedächtnisexperiment das – auf die Sequenzlänge relativierte - Zeitintervall von Präsentation des Wiedergabesignals an bis zur Beendigung der Reproduktion durch den Probanden. Diese On-Line Messung ist ein besserer Prädiktor für die Höhe der Gedächtnisspanne als die Off-Line Messung, die von Baddeley vorgeschlagen wird, da die Reproduktionszeit mehr Fehlervarianz aufklärt, als die Artikulationszeit. Die x-Achse von Abbildung 6 müßte also umbenannt werden in „Reproduktionszeit“. Es können nach Dosher und Ma (1998) soviel Items behalten werden wie in vier bis sechs Sekunden wiedergegeben werden können. Dieser Wert ergibt sich aus

empirischer Generalisierung und basiert nicht auf theoretischer Erkenntnis. Im Gegensatz dazu steht die Annahme von Cowan (1994), daß die Reproduktionszeit trotz variierender Artikulationszeiten bei differierendem Material invariant ist. Cowan (1994) verweist darauf, daß beispielsweise ältere Kinder schneller die Items artikulieren können, was in Off-Line Messungen der Artikulationszeit sichtbar wird, daß sich aber bei einer On-Line Erfassung der Reproduktionszeit pro Item die altersbedingten Unterschiede nivellieren. Differenzen machen sich nur in der Länge der gesamten Wiedergabezeit bemerkbar – ganz im Gegensatz zu der Annahme von Doshier und Ma. Auch Schweickert et alii (1990) berichten, daß je nach Grad der phonologischen Ähnlichkeit des Materials die Artikulationszeit variiert, daß jedoch in der Variablen „Reproduktionszeit“ keine Unterschiede aufzuweisen sind. Die offene Frage, ob nun die Wiedergabezeit eine Konstante ist oder eher die Reproduktionszeit, wurde bisher nicht beantwortet. Sie wird in dieser Arbeit aufgegriffen.

Einige der aufgelisteten Befunde, die der Baddeley-Theorie widersprechen, können durch die Berücksichtigung von Wiedergabeprozessen erklärt werden, beispielsweise die Relevanz von Phonologischer Ähnlichkeit, von Wortplanung und Wortkomplexität. Für die kritischen Effekte von Lexikalität, Assoziationsgehalt, graphemischer Komplexität, Frequenz, Konkretheit, Informationsgehalt oder Präsentationsmodus bieten sich keine Erklärungsmöglichkeiten. Die Reduktion der Artikulationszeit auf die Reproduktionszeit, so sie denn weiteren empirischen Überprüfungen standhalten kann, ist mit dem Working-Memory Modell von Baddeley auf konzeptueller Ebene verträglich. Allein die Erklärung Baddeleys für den WLE bezieht sich spezifisch auf die Spurverfalltheorie. Die Interferenztheorie findet beispielsweise bei Erklärung des Phonologischen Ähnlichkeitseffekts Berücksichtigung (Kapitel 2.4.2). Die Relevanz der Prozesse während der Wiedergabe läßt sich meines Erachtens auch direkt aus Experimenten von Baddeley, Lewis & Vallar (1984) ableiten: Der WLE wird nur dann eliminiert, wenn Artikulatorische Unterdrückung während Präsentations- und Wiedergabe erfolgt. Eine Reduktion der Baddeley-Hypothese durch den Ansatz von Doshier und Ma (1998) kann nicht erfolgen, da Output-Prozesse allein nicht den WLE provozieren können. Der WLE tritt auch bei Rekognitionsaufgaben (Baddeley & Wilson, 1985) und Probe-Wiedergabe auf (Henry, 1991).

ZWEITE MODIFIKATION: ERINNERN ALS ABWESENHEIT VON VERGESSEN UND ALS AKTIVE REKONSTRUKTION

Meiner Ansicht nach betrifft der zweite Kritikpunkt an Baddeleys Modell besonders den Bereich extrem langer Artikulationszeiten. Nach der Baddeleyschen Theorie müßte die

fiktive Gedächtnisspanne für Worte, deren Artikulation länger als zwei Sekunden dauert, gleich null sein. Schweickert (1993) würde dagegen halten, daß wegen des Einflusses des Langzeitgedächtnisses die Gedächtnisspanne höher sein könnte. Selbstverständlich liegt dieses Exempel im spekulativen Bereich, aber es bringt die Unterschiede in den Prognosen von Baddeley und Schweickert auf den Punkt, wie Abbildung 7 veranschaulicht.

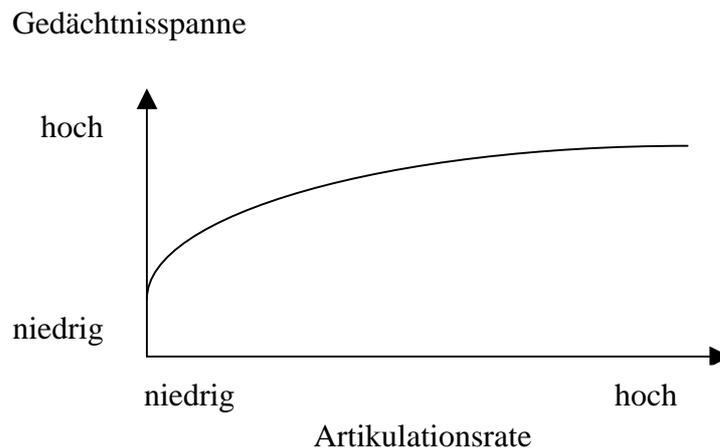


Abbildung 7 Zweite Modifikation des Baddeleyschen Erklärungsansatzes durch Berücksichtigung von Interferenz- und Langzeitgedächtnisprozessen

Im folgenden soll der Erklärungsansatz der Forschergruppe um C. Hulme am Lexikalitätseffekt erläutert werden. Hulme, Newton, Cowan, Stuart und Brown (1999) kontrastieren die Wirkungen von Lexikalität und Länge der Worte auf die Höhe der Gedächtnisspanne. Es zeigt sich erstens, daß Nonworte (unabhängig von der Artikulationszeit) schlechter erinnert werden als sinnvolle Worte und – analog zu Baddeley – kurze Stimuli bessere Gedächtnisleistungen evozieren als lange.

Wie in der ersten Modifikation des Ansatzes von Baddeley gezeigt wurde, werden zwei mögliche Ursachen für das Vergessen postuliert. Hulme et alii (1999) gehen zudem von zwei Alternativen zur korrekten Wiedergabe aus: Ein Item wird erinnert, wenn die Spur noch nicht zerfallen ist oder wenn die verfallende Gedächtnisspur korrekt rekonstruiert werden kann. Die korrekte Wiedergabe besteht aus zwei Phasen: zuerst wird die Gedächtnisspur selektiert, dann bei Bedarf wiederhergestellt. Diese sogenannte „Redintegration“ ist bei Nonworten, die auf keine Repräsentationen im Langzeitgedächtnisspeicher verweisen, weniger effizient als bei Worten. Zudem ist die Vertrautheit mit Nonworten geringer, was zusätzlich den Suchprozeß erschwert (Hulme, Roodenrys, Brown & Mercer, 1995).

Meßbar wird die zur Redintegration benötigte Zeit durch die Interstimulusintervalle während der Wiedergabe. Die charakteristischen Befunde für die Pausenzeiten werden im weiteren aufgelistet.

- Pausen sind bei Worten länger als bei Nonworten (Hulme et al., 1999).
- Zwischen Anzahl der Silben und Lexikalität wurde eine Interaktion gefunden: Bei nicht-lexikalem Material beeinflußt die Silbenzahl nicht die Pausenlänge (Bosshardt & Laug, 1995). Bei langen Worten dagegen sind die Interstimulusintervalle kürzer (Hulme et al., 1999). Erklärt wird dies durch den höheren Informationsgehalt von langen Worten, der trotz partiellen Informationsverlusts die Wahrscheinlichkeit einer korrekten Rekonstruktion erhöht.
- Bei steigender Listenlänge verlängert sich das Interstimulusintervall (Cowan, 1992).
- Die Länge der Interstimulusintervalle ist nicht sensitiv gegenüber der seriellen Position, die aktualisiert wird (Cowan, 1994).
- Je höher die Gedächtnisspanne ist, desto kürzer sind die Pausen (Cowan et al., 1994).
- Die Pausenlänge variiert in Abhängigkeit von der realisierten Sprache (Lüer et al., 1998).

Nach Hulme et alii (1999) kann die Wiedergabezeit in drei Vorgänge zerlegt werden: Während des Präparationsintervalls vor Initiierung der Sprachproduktion finden Prozesse zur Suche der Gedächtnisspuren statt. Analog zu den Sternberguntersuchungen und dem Cavanagh-Gesetz sinkt die Suchgeschwindigkeit mit Erhöhung des Suchraums. Dieser Zusammenhang mit der Suchzeit nach Sternberg (1966) auf die Gedächtnisspanne konnte von Laß, Lüer, Becker, Fang und Chen (2000) allerdings nicht gefunden werden. Dem Präparationsintervall folgt der Sprechvorgang, der wiederum unterteilt wird: einerseits in die Artikulationszeit pro Item und andererseits in die bereits erwähnten Pausenintervalle.

Alle aufgelisteten Determinanten des Wortlängeneffekts, nämlich Assoziationsgehalt, Frequenz, Konkretheit, Phonologische Ähnlichkeit, Wortkomplexität und Informationsgehalt, können meiner Ansicht nach durch das Redintegrations-Modell erklärt werden: Die Reichhaltigkeit (bestimmt durch Assoziationsgehalt und Frequenz) und Trennschärfe (evoziert durch Phonologische Ähnlichkeit, Konkretheit, Wortkomplexität und Listenlänge) der Assoziationen der semantischen Codes im Langzeitgedächtnisspeicher erhöht den Informationsgehalt der Items und steigert somit die Wahrscheinlichkeit einer korrekten Wiederherstellung, unabhängig von der Artikulationszeit. Bei Gültigkeit dieses Ansatzes würde die von Schweickert et alii (1990) postulierte kürzere Persistenz der

Gedächtnisspur bei klangähnlichen Items ein Phänomen der schlechten Redintegrationsbedingung wegen der hohen Interferenz sein.

Faktoren wie Phonologische Ähnlichkeit und Informationsgehalt pro Item wirken nicht nur auf die Wahrscheinlichkeit einer erfolgreichen Redintegration, sondern auch auf die Suchzeit, die im Modell von Hulme mit dem Präparationsintervall in Verbindung gebracht wird. Wie in zahlreichen Sternberg-Experimenten nachgewiesen wurde, steigt die Suchzeit mit Zunahme potentieller Alternativen (Klein, 1995). Der Effekt graphemischer Komplexität auf die Kurzzeitgedächtnisspanne (Zhang & Feng, 1990) kann durch die Size-Hypothese erklärt werden: Je mehr Merkmale zur Identifikation durchmustert werden müssen, desto länger dauert die Suchzeit (Cavanagh, 1972). Der Einfluß von Sprachplanungsprozessen scheint weder relevant für Redintegrationsmöglichkeiten noch für die Suchzeit zu sein. Diese Prozesse lassen sich meines Erachtens nur als kritisch für die Pausenzeiten interpretieren.

Als relevante Variable zur Prognose der Gedächtnisspanne unter Berücksichtigung von Redintegrationsprozessen könnte also - statt der Artikulations- oder Reproduktionszeit - der Informationsgehalt angesehen werden, der mit der Artikulationszeit korreliert (Bredenkamp & Hamm, 2001). Die Artikulationszeit verliert scheinbar bei Ausparialisierung des Informationsgehalts ihre Güte als Prädiktor der Gedächtnisspanne. Es kann vermutet werden, daß der Informationsgehalt erstens die Dauer der Durchmusterung des Suchraums in den Intervallen der Präparation und Pausen beeinflusst. Zweitens könnte er hoch mit der Artikulationszeit pro Item korrelieren und drittens könnte er die Wahrscheinlichkeit einer Rekonstruktion determinieren.

Die Forschergruppe um Hulme dagegen zieht ein alternatives Gedächtnismodell zur Erklärung von Redintegrationsprozessen heran, das nicht mehr eine rein modalitätsspezifische Speicherung annimmt: das Feature-Modell von Neath & Nairne (1995). Demnach werden sowohl modalitätsspezifische als auch amodale Merkmale als Repräsentationen integriert. Sowohl die Qualität der Repräsentationen, beispielsweise operationalisierbar als Grad der Interferenz durch Ähnlichkeit, als auch die Quantität der gespeicherten Merkmale beeinflussen die Wahrscheinlichkeit einer Rekonstruktion. Die Einführung des Feature-Modells wirft weitere Fragen auf, die im Rahmen dieser Arbeit nicht behandelt werden können, etwa das Problem mehrerer Speicherebenen, wie es die linguistischen Annahmen der Sprachproduktionsforschung vorschlagen.

Das Postulat des Einflusses vom Langzeitgedächtnisspeicher auf die Kurzzeitgedächtnisspanne hat für die Konzeption des Arbeitsgedächtnis-Modells weitreichende Folgen. Die Kurzzeitgedächtnisspanne wird nicht prozeßrein durch das jeweilige Subsystem determiniert, sondern sie wird von Langzeitgedächtnisspeicher und Zentraler Exekutive beeinflusst. Unter Annahme der Determination der Gedächtnisspanne durch den Zentralen Prozessor können auch „rätselhafte“ Befunde erklärt werden: Tattersall und Broadbent (1991) zeigen, daß der verbale Reproduktionsmodus dem schriftlichen nur dann überlegen ist, wenn die Versuchsperson vor der Lernphase nicht über die Art der Wiedergabe informiert wurde. Bei Information über den Reproduktionsmodus vor der Präsentationsphase verschwindet dieser Vorteil. Dies kann kein Effekt der Phonologischen Schleife sein, da sie keine Elemente voluntärer Einflußnahme berücksichtigt – im Gegensatz zur Konzeption der Zentralen Exekutive. Auch der Bias bezüglich des Präsentationsmodus von Chincotta und Underwood (1997b) könnte als Effekt der Zentralen Exekutive interpretiert werden: mit Zahlenmaterial hat jede Person eine große Lernerfahrung, allerdings meist nur mit Zahlensymbolen und nicht mit Zahlworten. Eine mögliche Erklärung für die signifikant reduzierte Gedächtnisleistung bei Zahlworten ist die reduzierte Vertrautheit und vor allem die erschwerte Chunkbildung.

Summa summarum läßt sich sagen, daß die Erklärung des WLE von Baddeley nicht ausreichend zu sein scheint und nur für eine mittlere Geschwindigkeit der Artikulation oder bei nicht-lexikalem Material gilt. Die Modifikation, die Vergessen sowohl als Spurverfall als auch wegen Interferenz berücksichtigt, scheint eine mit dem Arbeitsgedächtnismodell insgesamt kongruente Verbesserung zu sein. Allerdings findet eine Vielzahl von nachgewiesenen Einflußfaktoren keine Erklärungsmöglichkeit innerhalb dieses Modells. Eine weitere Reduktion dieses Ansatzes durch Schweickert (1993) dagegen scheint den aufgedeckten Befunden gerecht zu werden: Unter Implementation von Langzeitgedächtnisprozessen und dem Einfluß der Zentralen Exekutive können die Prognosen für die Gedächtnisspanne verbessert werden. Während das Feature-Modell die modalitätsspezifische Speicherung an sich in Frage gestellt, steht die Annahme einer verstärkten Wirkung der Zentralen Exekutive in Kongruenz mit der Konzeption des Arbeitsgedächtnisses nach Baddeley.

Im nächsten Kapitel wird das Baddeley-Modell als Invarianzhypothese innerhalb eines quantitativ-mathematischen Gedächtnismodells vorgestellt.

3 EIN QUANTITATIV-MATHEMATISCHES GEDÄCHTNISMODELL: DIE MODIFIZIERTE TOTAL-TIME HYPOTHESE (MTTH)

The computational-modeling approach is perforce more explicit, formal and complete and is associated with well-known advantages: it enforces systematicity and consistency, draws attention to gaps in the research evidence, and permits one to derive the formal implications of the model, which may contain some surprises and new discoveries (Kintsch, Healy, Hegarty, Pennington & Salthouse, 1999, S. 437).

Für die modifizierte Total-Time Hypothese (mTTH) als ein mathematisches Modell, das Lang- und Kurzzeitgedächtnisprozesse beim Lernen quantifiziert, gelten die oben genannten Vorteile der Formalisierung. Die besondere Attraktivität der mTTH besteht darin, daß aus hypothesengenerierenden Theorien, wie beispielsweise dem Arbeitsgedächtnismodell, präzise und prüfbare Prognosen abgeleitet werden können, die dem Kriterium der Falsifizierbarkeit genügen. Bredenkamp (1993) weist daraufhin, daß mit der Verknüpfung mehrerer Invarianzhypothesen nicht nur der empirische Gehalt steigt (Gadenne, 1994), sondern auch strengere Prüfungskriterien angelegt werden können. Dies steht im Gegensatz zu deskriptiven Theorien, die oft bei Berücksichtigung eines größeren Phänomenbereichs an Transparenz und Falsifizierbarkeit verlieren.

Im folgenden wird zuerst auf eine Ableitung und Darstellung der Invarianzhypothesen eingegangen, dann wird die Integrationsmöglichkeit der mTTH in das Arbeitsgedächtnismodell von Baddeley (Kapitel 2) beschrieben. Den Abschluß bildet eine Zusammenfassung der bisherigen experimentellen Evidenz für die Gültigkeit der mTTH.

3.1 ABLEITUNG UND DARSTELLUNG DER MTTH

Die mTTH ergibt sich aus der Verknüpfung der Total-Time Hypothese von Cooper und Pantle (1967) und der Hypothese einer invarianten Langzeitgedächtnisspanne von Nevelsky (1970).

Die Total-Time Hypothese (TTH) postuliert die Gesamtlernzeit zum Bewältigen eines bestimmten Lernmaterials als invariant ohne Interdependenz zu der temporalen

Segmentierung der Präsentationsphase. Benötigt man beispielsweise eine Minute zum Erlernen einer Stimulussequenz, dann besteht kein Unterschied, ob dreimal für zwanzig Sekunden gelernt wird oder einmal sechzig Sekunden am Stück. Baddeley (1997) spezifiziert die Verbindung in dem Sinne, daß zusätzlich Zeit benötigt wird, um die in den einzelnen Lernphasen aufgenommene Information in eine kohärente Gesamtrepräsentation zu integrieren. Diese Annahme wurde jedoch weder theoretisch fundiert, noch experimentell überprüft. Formel 1 beschreibt den mathematischen Zusammenhang zwischen Gesamtlernzeit, Anzahl der Lernversuche und der Präsentationszeit nach Cooper und Pantle (1967).

Formel 1 $T_v = Y_{vj} * t_j$

T_v entspricht der für Material v spezifischen Gesamtlernzeit, Y_{vj} steht für die benötigte Anzahl der Lerndurchgänge, spezifisch für Material v und Darbietungszeit j , und mit t_j ist die jeweilige Darbietungszeit gemeint. Sämtliche Parameter und Indices werden durchgängig in der gleichen Bedeutung verwendet. Aus diesem Grund erfolgt nur eine einmalige Erläuterung.

Implizite Voraussetzung der Gültigkeit von Formel 1 ist die aktive Nutzung der Lernzeit. Diese Prämisse ist nicht trivial, wie Studien an Kindern und an Patienten, die an Demenz leiden, zeigen (Baddeley 1997).

Die Befundlage zu der TTH ist heterogen: Cooper und Pantle (1967) berichten über experimentelle Belege der TTH bei freier und serieller Wiedergabe und bei paarassoziertem Lernen. Dieser von der TTH postulierte Zusammenhang verschwindet bei sehr kurzer Darbietungszeit (Bugelski, Kidd, Segmen, 1968). Ebenso wird die TTH durch alle Studien widerlegt, die für eine Superiorität verteilten Lernens über massierte Exposition sprechen. Zwischen den Studien, in denen die TTH bestätigt wird, und denen, in denen sie verworfen wird, besteht ein substantieller Unterschied: Die TTH kann akzeptiert werden für mittlere Präsentationszeiten, für sehr lange oder für sehr kurze Materialpräsentation ist sie nicht adäquat. Bei langen Darbietungszeiten wird die Menge verarbeiteter Information überschätzt (Bredenkamp & Hager, 1979). Bevor die Modifikation der TTH, die mTTH, vorgestellt wird, wird die Invarianzhypothese von Nevelsky (1970) eingeführt.

Die Hypothese von Nevelsky bezieht sich auf die Langzeitgedächtnisspanne. Diese Spanne ist definiert als die Informationsmenge in bit, die man in einem Lernversuch verarbeiten

kann, wenn der Gesamtinformationsgehalt des Lernmaterials die begrenzte Kapazität des Kurzzeitgedächtnisses überschreitet.

Nevelsky (1970) definiert eine informationstheoretisch limitierte Langzeitgedächtnisspanne: So soll in jedem Lernversuch, in Abhängigkeit von der Präsentationszeit, die aufgenommene Information invariant sein, unabhängig vom insgesamt zu verarbeitendem Informationsgehalt. Eine Invarianz der Langzeitgedächtnisspanne wird bei Messung in bit erreicht, nicht jedoch bei Berücksichtigung der Anzahl von Items. Formel 2 beschreibt den Zusammenhang zwischen verarbeiteter Information, Gesamtinformationsgehalt und Anzahl der Lernversuche.

Formel 2 $r_j = I_v * Y_{jv}^{-1}$

r_j steht für die in der Zeit j verarbeitete Informationsmenge und I_v für den Gesamtinformationsgehalt des Materials v . Kintsch (1970) berichtet über experimentelle Evidenz für diese Hypothese.

Aus der Verknüpfung der TTH (Formel 1) mit der Nevelsky-Hypothese (Formel 2) resultiert das Postulat eines linearen Zuwachses der verarbeiteten Information mit der Darbietungszeit. In Formel 3 wird die Verbindung zwischen verarbeiteter Information, Gesamtinformationsgehalt, Gesamtlernzeit und Darbietungszeit pro Lernversuch nach der TTH und Nevelsky beschrieben.

Formel 3 $r_j = I_v * T_v^{-1} * t_j$

Dieser Zusammenhang wird in der gesamten Literatur zur limitierten Kapazität des Gedächtnisses widerlegt. Dies drückte Ebbinghaus schon im Jahre 1885 folgendermaßen aus.

Es ist hinreichend bekannt, daß die Einprägung von Vorstellungsreihen, die zu einer bestimmten späteren Zeit reproduziert werden sollen, um so schwieriger ist, je länger die Reihen sind. Das heißt, diese Einprägung erfordert nicht nur absolut genommen mehr Zeit bei größerer Länge der Reihen, weil eben jede Wiederholung länger dauert, sondern sie beansprucht auch verhältnismäßig

mehr Zeit, weil eine wachsende Anzahl von Wiederholungen nötig wird. Sechs Verse eines Gedichts kosten, um gelernt zu werden, nicht nur dreimal soviel Zeit wie zwei, sondern erheblich mehr (Ebbinghaus, 1966, S. 62).

Bredenkamp (1975) zeigte in Re-Analysen, daß Formel 4 an die Daten angepaßt werden kann. Die These einer konstanten Gesamtlernzeit unabhängig von der zeitlichen Segmentierung der TTH wird aufgegeben und durch den in Formel 4 dargestellten Zusammenhang zwischen der Anzahl der Lernversuche, den beiden Parametern a und b sowie der Darbietungszeit pro Lernversuch modifiziert (mTTH).

Formel 4 $Y_{jv} = a_v + b_v * t_j^{-1}$

a und b sind für das Material v spezifische Konstanten. Da nach der Nevelsky-Hypothese die verarbeitete Informationsmenge als unabhängig vom Gesamtinformationsgehalt angenommen wird, kann in Experimenten mit variierendem Gesamtinformationsgehalt einer Materialart der Index v eliminiert werden. Allein die Mittelwerte von a_v und b_v gehen in die weitere Berechnung ein. Ermittelt werden a und b über eine Regressionsgleichung. Dabei gilt, daß a im positiven Wertebereich über Null liegen muß. Voraussetzung ist, daß Informationsgehalt und Darbietungszeit unabhängig voneinander verändert werden können.

Für eine Definition der Langzeitgedächtnisspanne ergibt sich folgende Formalisierung nach der TTH (Formel 5), die Formel 3 ersetzt. Formel 5 drückt die Verbindung zwischen verarbeiteter Information, Gesamtinformationsgehalt, Darbietungszeit und den beiden Konstanten nach der mTTH aus.

Formel 5 $r_j = \frac{I}{a + b * t_j^{-1}}$

Diese Gleichung beschreibt - als logisches Theorem aus Gleichung 2 und 4 - eine hyperbolische Funktion, an die Thurstone schon 1919 die Lernkurve anpaßte. Dividiert man Nenner und Zähler von Formel 5 durch a, sieht man, daß diese Funktion (wenn $t \rightarrow \infty$) eine Asymptote hat, die in Formel 6 beschrieben wird durch den Quotienten aus dem Gesamtinformationsgehalt (I) und Parameter a.

Formel 6 $r_{\max} = I * a^{-1}$

r_{\max} entspricht der maximal pro Versuch verarbeitbaren Informationsmenge. Als fundamentale Annahme der Theorie der mTTH wird diese Asymptote mit der Kurzzeitgedächtnisspanne in bit gleichgesetzt (Kapitel 3.2 und 3.3). Abbildung 8 stellt graphisch die Differenz der Prognosen von TTH und mTTH in Bezug auf die Langzeitgedächtnisspanne dar.

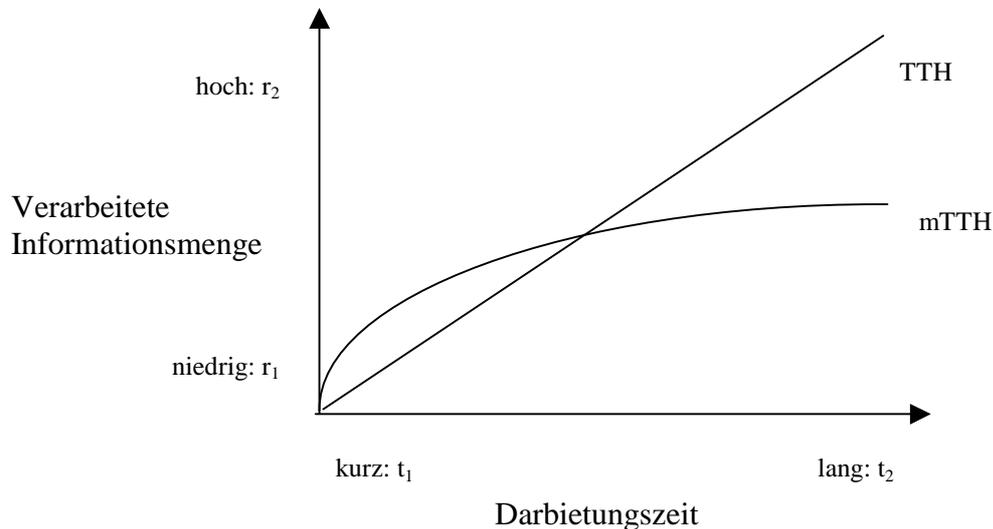


Abbildung 8 Beziehung zwischen Menge der verarbeiteten Information und der Darbietungszeit nach TTH und mTTH

Für die beiden Konstanten der mTTH, a und b , ergeben sich innerhalb des Arbeitsgedächtnis-Modells folgende Interpretationsansätze (Bredenkamp & Klein, 1998): a entspricht, deduziert aus Formel 6, der Anzahl der Lernversuche, die mindestens durchgeführt werden müssen, um das Lernkriterium im Kurzzeitgedächtnis zu erfüllen. Wie aus Formel 4 ersichtlich ist, ist a abhängig vom Informationsgehalt und unabhängig von der Darbietungszeit.

bt_j^{-1} wird interpretiert als die Anzahl von Lernversuchen, die zusätzlich zur minimal erforderlichen Anzahl a benötigt wird. b entspricht der Verarbeitungszeit im Langzeitgedächtnisspeicher, die mit dem Gesamtinformationsgehalt variiert und unabhängig von der Darbietungszeit ist.

Die beiden Konstanten können statt durch ein modulares Modell auch durch Integration des Levels-of-Processing Ansatzes (Craik & Lockhart, 1972) interpretiert werden (Bredenkamp, 1975): Parameter a entspricht der Anzahl der Lernversuche für Typ-1 der erhaltenden Verarbeitung, b entspricht der Zeit, die zur Typ-2 der vertiefenden Verarbeitung benötigt wird.

Bevor im nächsten Kapitel auf die Verknüpfung der beiden Parameter der mTTH mit dem Arbeitsgedächtnismodell von Baddeley eingegangen wird, werden in einem Exkurs verschiedene Möglichkeiten der Bestimmung der Kurzzeitgedächtnisspanne vorgestellt.

EXKURS DEFINITION DER KURZZEITGEDÄCHTNISSPANNE

Die Kurzzeitgedächtnisspanne ist eine Gedächtnisleistung, die dem Umfang aufgenommener Information in einem Lernversuch entspricht. Es gibt verschiedene Einheiten zur Messung der Kurzzeitgedächtnisspanne. Wie in diesem Kapitel gezeigt wird, sind jedoch denotative Bedeutung und Begriffsextension äquivalent.

1. Miller (1956) beschrieb die Begrenzung der Kurzzeitgedächtnisleistung durch eine konstante Anzahl von Items, die maximal gemerkt werden können „the magical number seven, plus or minus two“ (Seite 81). Diese Anzahl ist fix und kann nicht verändert werden. Die enge Kapazitätsbegrenzung des Kurzzeitgedächtnisses kann ausgeweitet werden durch Organisation und Strukturierung der Items. So bleibt die Anzahl der gemerkten Chunks invariant, aber der Informationsgehalt pro Item ändert sich. Miller berichtet über experimentelle Evidenz.
2. Baddeley et alii (1975) prüften die Hypothese von Miller: Die gemessene Höhe der Gedächtnisspanne für verschiedene Materialien war nicht durch eine invariante Anzahl von Items zu beschreiben. Ein weiteres Experiment (Baddeley et al., 1975, Experiment 4) zeigte, daß die Kurzzeitgedächtnisleistung auch nicht durch die Anzahl konstituierender Phoneme determiniert wird. Als Alternative zur verworfenen Hypothese von Miller entwickelten Baddeley et alii den Ansatz, daß die Gedächtnisspanne einer zeitlichen Determination durch die Artikulationszeit der Items unterliegt. „The probability of correctly recalling a list equals the probability that the time to recite the list is less than the variable duration of the trace“ (Schweickert & Boruff, 1986, S. 419). Für diese – zeitlich – begrenzte Kapazität des Gedächtnisses existieren zahlreiche empirische Bestätigungen, in jüngster Zeit häuft sich Evidenz gegen diese Behauptung (Kapitel 2.4.4).
3. Die dritte Art, wie die Kurzzeitgedächtnisspanne gemessen werden kann, wird beispielsweise durch Invarianzhypothesen realisiert. Postuliert wird weder eine Beschränkung auf eine fixe Anzahl von Items noch eine zeitliche Begrenzung, sondern ein informationstheoretisches Limit. Der Begriff „informationstheoretisch“ wird in

dieser Arbeit immer nur in Bezug auf den Informationsgehalt der Items gebraucht und impliziert keine weiteren Annahmen der Informationstheorie. Die Messung der Kurzzeitgedächtnisspanne in bit birgt zwei Vorteile: Erstens wird dieselbe Einheit wie auch für die Langzeitgedächtnisspanne verwendet (Nevelsky, 1970) und zweitens kann man über die Variable des Informationsgehalts materialabhängige oder interindividuelle Unterschiede erklären. Relevant scheint dabei nicht nur der objektiv bestimmbare Informationsgehalt der Items zu sein, sondern auch der subjektive Informationsgehalt. Dieser kann interindividuell variieren, da für seine Bestimmung auch das Verfügen über Strategien zur Reduktion der Information relevant ist, wie auch Miller annimmt. Dies stellt ein Anknüpfungspunkt an die Arbeit von Ericsson und Delaney (1999, Kapitel 2.2.3) dar.

3.2 VERKNÜPFUNG DER MTTH MIT DEM ARBEITSGEDÄCHTNISMODELL

Die fundamentale Annahme der mTTH besagt, daß die Asymptote (Formel 6) mit der Kurzzeitgedächtnisspanne in bit gleichgesetzt werden kann. Die Kurzzeitgedächtnisspanne kann aus mehreren unabhängigen Theorien geschätzt werden: beispielsweise durch die Definition der Spanne nach Cavanagh (1972), durch den Ansatz von Baddeley zum Wortlängeneffekt, der von Bredenkamp (1993) als Invarianzhypothese formuliert wurde, durch die Modifikationen zum Wortlängeneffekt durch Doshier und Ma (1998) oder Hulme et alii (1999). Auf die Integration von mTTH und Cavanagh oder Hulme wird im folgenden nicht eingegangen. Diese Arbeit bezieht sich auf die Verknüpfung von mTTH und der Invarianzhypothese von Baddeley sowie der Hypothese von Doshier und Ma.

Wie bereits in Kapitel 2.4.4 beschrieben, entspricht die Gedächtnisspanne (gemessen in der Einheit „Anzahl der Items“) nach Baddeleys Erklärungsansatz zum Wortlängeneffekt dem Quotienten zwischen der konstanten Dauer der Gedächtnisspur und der Artikulationszeit pro Item (Formel 7).

Formel 7 $g_v = B * u_v^{-1}$

g_v steht für die materialspezifische Höhe der Gedächtnisspanne, B für die von Baddeley auf ungefähr zwei Sekunden geschätzte Persistenz der Gedächtnisspur und mit u_v ist die durchschnittliche Artikulationszeit des Materials v gemeint (Kapitel 2.4.4). Im folgenden wird durch die Verwendung der Termini „Artikulationszeit“, „Artikulationsgeschwindigkeit“ oder „Schnelligkeit der Artikulation“ immer auf Parameter u verwiesen.

Zunächst ist die Einheit der Gedächtnisspanne eine zeitlich begrenzte Anzahl erinnerter Items, doch durch Multiplikation mit dem Informationsgehalt pro Item (m_v) wird die gleiche Einheit zwischen der Baddeleyschen Invarianzhypothese und der Asymptote erreicht. Somit wird als Maßeinheit für Kurz- und Langzeitgedächtnisspanne das informationstheoretische bit definiert.

Formel 8 $s_v = B * u_v^{-1} * m_v$

Mit s_v wird auf die Gedächtnisspanne in bit verwiesen. Durch Konjunktion der mTTH mit der Invarianzhypothese von Baddeley wird eine Interpretation der Parameter a und b innerhalb des Arbeitsgedächtnis-Modells ermöglicht. So wird vermutet, daß Parameter a von der Phonologischen Schleife determiniert wird, respektive daß Parameter b als Verbindung zum Langzeitgedächtnisspeicher von der Zentralen Exekutive beeinflusst wird. Über experimentelle Evidenz zu diesen Annahmen wird im nächsten Kapitel berichtet.

Die Hypothese von Doshier und Ma (1998) stellt eine Modifikation des Erklärungsansatzes von Baddeley dar. Statt einer konstanten Dauer der Gedächtnisspur wird eine Invarianz und zeitliche Limitierung der Wiedergabezeit postuliert. Variable des Quotients ist nicht die Artikulationszeit, sondern die On-Line erhobene Reproduktionszeit. Die Theorie von Doshier und Ma (1998) kann durch Formel 9, bestehend aus den Variablen Kurzzeitgedächtnisspanne in bit, Gesamtwiedergabe-Zeit und Reproduktionszeit, als Invarianzhypothese ausgedrückt werden. Um eine gleiche Maßeinheit zu der Asymptote der Lernfunktion herzustellen, wird die Gedächtnisspanne, die nach Doshier und Ma als Anzahl korrekt reproduzierter Items definiert ist, durch Multiplikation mit dem Informationsgehalt pro Item in die Einheit bit transformiert.

Formel 9 $s_v = D * d_v^{-1} * m_v$

D entspricht der konstanten Wiedergabezeit und d_v der Reproduktionszeit pro Item. Im folgenden wird mit den beiden Termini „Gesamtwiedergabe-Zeit“ oder „Wiedergabezeit“ immer Parameter D bezeichnet. Mit „Reproduktionszeit“ ist durchgängig das bei der Wiedergabe benötigte Intervall pro Item gemeint. Im Gegensatz zur Definition der Persistenz der Gedächtnisspur ist es fraglich, durch welches theoretische oder physiologische Konstrukt die Begrenzung der Wiedergabezeit fundiert werden kann.

3.3 EXPERIMENTELLE BEFUNDE

Die Forschung zur modifizierten Invarianzhypothese umfaßt bis jetzt – neben den Re-Analysen - zehn Experimente. Der erste Forschungsschwerpunkt lag, neben dem Beleg der Gültigkeit der mTTH, auf der explorativen Erkundung des Zusammenhanges zum Arbeitsgedächtnismodell. In sechs Experimenten mit Silben wurden systematisch Dual-Task Aufgaben (Kapitel 2.4) durchgeführt. Konsonant - Vokal - Verbindungen bieten sich als Untersuchungsmaterial an, da der Informationsgehalt (2, 3 und 4 bit pro Item) unabhängig von der Änderung des Umfangs und der Artikulationsdauer variiert werden kann. Des weiteren ist bei dieser Materialart der von Hulme et alii (1995) beschriebene Einfluß des Langzeitgedächtnisses minimiert. Die beschriebenen Experimentreihen bestehen jeweils aus Messungen der Kurz- und Langzeitgedächtnisspanne. Der Versuchsaufbau der beiden Gedächtnisexperimente entspricht weitgehend dem in Kapitel 5.1.2 und 5.1.3 beschriebenen Vorgehen.

In Experiment 1 war keine Sekundäraufgabe simultan zum Lernexperiment auszuführen (Bredenkamp & Hamm, 2001), Experiment 2 verlangte zusätzlich beim Lernen, daß artikulatorisch unterdrückt wurde (Bredenkamp & Klein, 1998), Experiment 3, 4 und 5 explorierten den Einfluß verschiedener Irrelevanter Sprachen (Polnisch vorwärts und rückwärts für das Supraspannen-Experiment: Hamm, 1998; Rumänisch Sub- und Supraspannenexperiment sowie Polnisch vorwärts/Subspannen-Experiment: Bredenkamp & Hamm, 2001) und Experiment 6 (Bredenkamp & Hamm, 2001) bezog sich auf die Untersuchung des Einflusses der Zentralen Exekutive via der Random-Interval-Repetition-Task von Vandierendonck et alii (1998b, Kapitel 2.2.1). Wie in Kapitel 2.4 beschrieben, sollten diese Treatments die Kurzzeitgedächtnisspanne und damit die Asymptotenhöhe reduzieren. Bezüglich der Veränderung für die Verarbeitungszeit im Langzeitspeicher lagen keine Hypothesen außer der vor, daß sich bei Belastung der Zentralen Exekutive (Experiment 6) diese Zeit signifikant verlängern sollte (Bredenkamp & Hamm, 2001).

Weitere Experimente wurden mit lexikalem Material durchgeführt: Experiment 7 und 8 bezogen sich auf bildhafte, verbalisierbare Stimuli mit und ohne artikulatorische Unterdrückung (Schiffer, 2000). Variiert wurde der Informationsgehalt mit 2.81 bit, respektive 3.81 bit pro Item. Die Ausprägung der Kurzzeitgedächtnisspanne wurde unabhängig von diesen Experimenten mit identischem Material erhoben. Die letzten beiden Studien überprüften den Einfluß von Chunking auf die hyperbolische Funktion (Fankam, 1999) bei Zahlenmaterial. Dabei unterschieden sich die beiden Experimente darin, ob die

Lernreihe in Zweier- oder in Dreier-Chunks zusammengefaßt werden sollten. Die Kodierung wurde vorgegeben und in einem Vorversuch geübt. Es wurden keine zusätzlichen Kurzzeitgedächtnis - Experimente durchgeführt.

Im folgenden werden die Resultate der zehn Experimente in einen gemeinsamen Kontext integriert. Die experimentelle Evidenz wird einzeln für jede der drei Hauptannahmen besprochen: Zuerst wird auf Ergebnisse der Studien zur Annahme einer hyperbolischen Funktion eingegangen, dann einzeln auf Befunde zu den Parametern a und b.

3.3.1 EXPERIMENTELLE EVIDENZ FÜR DIE ANNAHME EINER HYPERBOLISCHEN LERNFUNKTION

Dieser Punkt läßt sich kurz abhandeln: In allen zehn Experimenten konnte die Lernkurve einer hyperbolischen Funktion unabhängig vom Informationsgehalt erwartungsgemäß angepaßt werden. Dies war unabhängig von der Materialart (sinnloses versus sinnvolles Material) und vom Treatment (mit / ohne artikulatorische Unterdrückung, Belastung der Zentralen Exekutive, Irrelevante Sprachen).

Abbildung 9 veranschaulicht graphisch den Verlauf der aufgenommenen Information unabhängig vom Gesamtinformationsgehalt bei Variation der Darbietungszeit exemplarisch für Experiment 1 (Stille), Experiment 2 (Art. Unterdrückung), Experiment 5 (Rumänisch) und Experiment 6 (RIR).

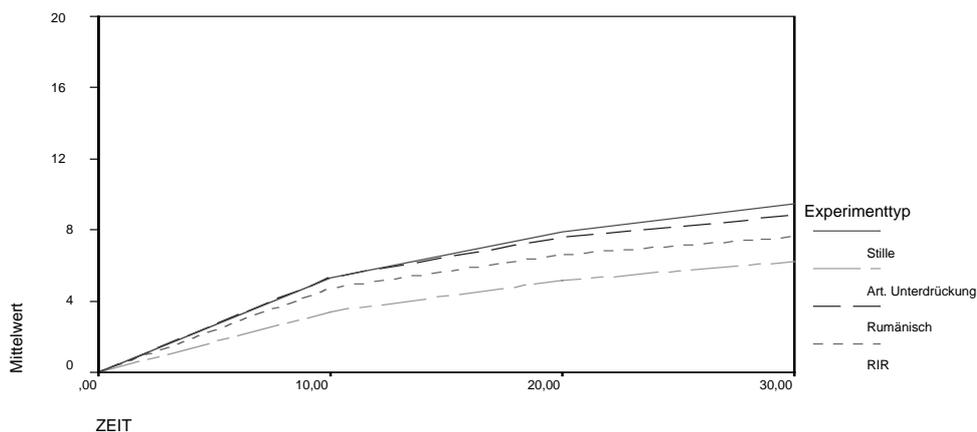


Abbildung 9 Aufgenommene Information der Silben - Experimente 1 (Stille), 2 (Art. Unterdrückung), 5 (Irrelevante Sprache: Rumänisch) und 6 (RIR, zur Belastung der Zentralen Exekutive) in Abhängigkeit von der Präsentationszeit

3.3.2 EXPERIMENTELLE EVIDENZ FÜR DIE ÄQUIVALENZHYPOTHESE

Der Parameter a ist definiert als Quotient aus Gesamtinformationsgehalt und Kurzzeitgedächtnisspanne in bit (Kapitel 3.1, Formel 6). In den bisher durchgeführten Studien fand allein die Baddeleysche Invarianzhypothese Eingang in die Verknüpfung. Zur Integration von mTTH und Invarianzhypothese von Doshier und Ma (1998) liegen derzeit noch keine Daten vor.

Die Interpretation der Asymptote als Kurzzeitgedächtnisspanne in bit wird dadurch geprüft, daß – wie in Formel 10 dargestellt wird – Parameter a durch den Quotienten $I * s_v^{-1}$ substituiert wird, wobei s_v der Kurzzeitgedächtnisspanne (r_{max}) in bit entspricht. Durch einen F-Test wird geprüft, ob die so berechneten Werte von Y''_j systematisch von den empirischen Werten Y_j abweichen. In dieser Arbeit bedeutet „ ‘ “ hinter einem Parameter durchgängig, daß der Parameter geschätzt und nicht empirisch beobachtet wurde.

Formel 10 $Y''_j = I * r_{max}^{-1} + b * t_j^{-1}$

Ein Vergleich der nach Formel 10 prognostizierten Daten mit den empirischen führt zu folgenden Ergebnissen:

- In allen Experimenten mit Silbenmaterial konnte die Äquivalenzhypothese beibehalten werden.
- Weder in Experiment 7 (bildhaftes Material, Stille) noch in Experiment 8 (bildhaftes Material, Artikulatorische Unterdrückung) konnte die Asymptote mit der Kurzzeitgedächtnisspanne in bit gleichgesetzt werden (Schiffer, 2000): Die Asymptote überschätzte die Kurzzeitgedächtnisleistung signifikant. Deskriptiv zeigte sich bei Artikulatorischer Unterdrückung sowohl eine signifikant reduzierte Asymptote als auch eine signifikant niedrigere Kurzzeitgedächtnisspanne. Die relative Beziehung zwischen den beiden Parametern blieb erhalten, allein die absolute Gleichstellung gelang nicht. Schiffer (2000) vermutete, daß es sich hierbei um den Einfluß der Lexikalität und/oder der Repräsentation im Langzeitgedächtnisspeicher handelt.
- Die beiden Studien mit Zweier-, respektive Dreier-Chunking wurde nur als Supraspannen-Experiment durchgeführt. Die Äquivalenzhypothese konnte nicht überprüft werden.

Ein zweiter Ansatzpunkt der Überprüfung dieser Gleichstellung ergab sich aus der systematischen Manipulation der Versuchsbedingungen, die auf die Leistungen im

Kurzzeitgedächtnis wirken sollten (Kapitel 2.4). Signifikante Reduktionen im interexperimentellen Vergleich der Subspannen-Experimente sollten sich auch in der Asymptotenhöhe widerspiegeln. Wiederum entsprachen die Ergebnisse für Silbenmaterial den Erwartungen (Bredenkamp & Klein, 1998; Bredenkamp & Hamm, 2001). Im Vergleich der beiden Chunking-Experimente zeigte sich erwartungsgemäß, daß bei Dreier-Chunks die Asymptote signifikant erhöht war.

Durch die Verknüpfung der mTTH mit der Invarianzhypothese von Baddeley eröffnet sich die Möglichkeit, die Konstanten a und b mit den im Arbeitsgedächtnis-Modells postulierten Modulen zu verbinden. Es zeigte sich, daß sich Parameter a bei Manipulation der Phonologischen Schleife im Dual-Task Paradigma sowohl durch Artikulatorische Unterdrückung als auch durch Irrelevante Sprache erhöhte. Ebenso sank die Asymptotenhöhe bei Belastung der Zentralen Exekutive. Die Ergebnisse sprechen in toto dafür, daß sowohl die Phonologische Schleife als auch die Zentrale Exekutive als determinierende Systeme für die Asymptotenhöhe verantwortlich sind. Dies scheint auf den ersten Blick verwunderlich, da Baddeley (1997) postulierte, daß die Kurzzeitgedächtnisspanne vom jeweiligen Subsystem allein abhängt, nicht jedoch vom zentralen Prozessor. Es besteht jedoch Kongruenz zu der Studie von Hulme et alii (1995), in der ein Einfluß des Langzeitgedächtnisspeichers auch auf die Kurzzeitgedächtnisspanne nachgewiesen wurde (Kapitel 2.4.4). Die Verbindung zwischen dem Kurz- und dem Langzeitgedächtnismodul wird innerhalb des Arbeitsgedächtnis-Modells dem Funktionsbereich der Zentralen Exekutive zugeordnet.

3.3.3 EXPERIMENTELLE BEFUNDE ZUR VERARBEITUNGSZEIT IM LANGZEITGEDÄCHTNISPEICHER

Parameter b war von Bredenkamp (1975) als die zur Verarbeitung im Langzeitgedächtnisspeicher benötigte Zeit interpretiert worden. Die Befundlage der Silben - Experimente erbrachte eine erstaunliche Konstanz des b -Parameters unabhängig von Sekundäraufgaben, die bei 1.2 Sekunden Verarbeitungszeit pro bit im Langzeitgedächtnisspeicher liegt (Tabelle 2). Diese Zeitangabe errechnet sich aus dem Quotienten zwischen Parameter b und dem Gesamtinformationsgehalt I . Eine Ausnahme stellt das Experiment zur Artikulatorischen Unterdrückung dar: Die Verarbeitungszeit stieg auf 2 Sekunden an (Bredenkamp & Klein, 1998). Das Ausführen dieser Sekundäraufgabe veränderte signifikant das Verhältnis zwischen durchschnittlich verarbeiteter Information und maximaler Informationsmenge pro Zeitpunkt (Bredenkamp & Hamm, 2001).

Es stellte sich heraus, daß die Verarbeitungszeit - entgegen der aus dem Arbeitsgedächtnismodell abgeleiteten Erwartung - nicht durch die Zentrale Exekutive beeinflußt wird, deren Belastung keine Änderung der Größe bewirkte. Die Phonologische Schleife, respektive der Prozeß subvokaler Wiederholung scheint diesen Parameter zu determinieren. Dies ist, meiner Meinung nach, ein Hinweis darauf, daß das Baddeley-Modell erweitert werden muß: Es gibt anscheinend vom Subsystem aus einen direkten Zugang zum Langzeitgedächtnisspeicher. Die direkte Verbindung vom Langzeitspeicher zur Phonologischen Schleife wird auch von Baddeley angenommen (Kapitel 2), der umgekehrte Weg jedoch sollte nach Baddeley mit Hilfe der Zentralen Exekutive erfolgen.

Ausgangspunkt der beiden Experimente ohne, respektive mit Artikulatorischer Unterdrückung bei bildhaften Stimuli, war der Gedanke, daß das Sketchpad den Ausfall der Phonologischen Schleife kompensiert, der durch die Sekundäraufgabe induziert wurde. Demnach sollte eine Zusatzfähigkeit nicht die Verarbeitungszeit im Langzeitspeicher verlängern. Die Ergebnisse zeigten für Parameter b deskriptiv eine Verlängerung. Für die Auslegung der Veränderung des Parameters b bieten sich aus meiner Sicht drei Alternativen an. Die erste Möglichkeit wäre, daß die Phonologische Schleife komplett ausgeschaltet wurde und daß die deskriptiv verlängerte Verarbeitungszeit im Langzeitspeicher ein Phänomen des Sketchpads ist, das langsamer als die Phonologische Schleife arbeitet. Eine weitere Möglichkeit wäre, daß - wie in Baddeleys Modell postuliert - das Sketchpad keinen direkten Zugang zum Langzeitgedächtnisspeicher hat, sondern nur mittelbar über die Zentrale Exekutive. Die beiden Erklärungen wurden spekulativ generiert, können aber durch ein entsprechendes Experiment überprüft werden. Eine dritte, und die meines Erachtens wahrscheinlichste Möglichkeit ist die, daß Artikulatorische Unterdrückung die Artikulation nicht immer vollständig unterdrückt, sondern nur erschwert - zu demselben Schluß basierend auf einem anderen Untersuchungsparadigma kommt Dilger (2000). Dies würde erklären, warum sich deskriptiv eine Verlängerung der Verarbeitungszeit im Langzeitspeicher bei Belegung des Kontrollprozesses auch bei visuellem Material zeigt. Ein möglicher Nutzen der - wenn auch durch die Sekundäraufgabe erschwerten - Artikulation des Materials liegt eventuell im Versuchsaufbau begründet: Nach der Lernphase sollten die Versuchspersonen laut aussprechen, woran sie sich erinnern. Spätestens bis zur Wiedergabe mußten also phonologische Repräsentationen gebildet werden, so daß die Nutzung der Phonologischen Schleife adaptiv erscheint.

Eine Gegenüberstellung der Silben- mit den Bildexperimenten zeigt deskriptiv das Auftreten eines Lexikalitätseffekts. Die Verarbeitungszeit pro bit im Langzeitgedächtnisspeicher ist

bei Silben im Langzeitspeicher länger als bei bereits bekannten Stimuli, wie in Tabelle 2 deskriptiv dargestellt wird.

**Tabelle 2 Geschätzte Verarbeitungszeit im Langzeitgedächtnis-
speicher pro bit**

Experiment	Verarbeitungszeit
1 (Stille, Silben)	1.25
2 (art.Unt., Silben)	1.99
3 (Pol.rü., Silben)	1.11
4 (Pol.vo., Silben)	1.22
5 (Rumä., Silben)	1.10
6 (RIR, Silben)	1.21
7 (Stille, Figuren)	0.84
8 (art.Unt., Figuren)	1.14
9 (2er-Chunk, Zahlen)	1.66
10 (3er-Chunk, Zahlen)	1.56

Der Unterschied bezüglich der Langzeitgedächtnisprozesse zwischen den beiden Chunking-Experimenten ist signifikant: Bei dem Bilden von Dreier-Chunks wird durchschnittlich mehr Information im Bezug zur maximalen Informationsmenge pro Zeitpunkt verarbeitet als beim Bilden von Zweier-Chunks (Fankam, 1999). Die beiden Chunking-Experimente geben Aufschluß über den Einfluß der Zentralen Exekutive, zu deren Funktionen auch der Einsatz von Mnemotechniken gerechnet werden kann. Wie Tabelle 2 belegt, liegen die Werte der Verarbeitungszeit pro bit bei Chunking deskriptiv eher in dem Intervall der Silbenexperimente. Das ist in Zusammenhang mit dem deklarierten Lexikalitätseffekt erstaunlich. Für sinnvolles Material wäre eher eine Ausprägung der Verarbeitungszeit im Bereich von Experiment 7 und 8 erwartet worden, da in den Silbenexperimenten keine Beeinflussung von Parameter b durch die Zentrale Exekutive nachgewiesen werden konnte (Experiment 6). Meiner Meinung nach bieten sich zwei Erklärungsmöglichkeiten an. Erstens könnte dieses Ergebnis zustande kommen, weil die Chunks nicht internalisiert waren. Sie wurden vom Versuchsleiter oktroyiert und manifestierten sich in Rückübersetzungsfehler. Eine weitere Möglichkeit wäre, daß die Zentrale Exekutive nicht ein einheitliches Konstrukt ist, sondern ein Konglomerat interagierender Prozesse (Kapitel 2.2.2). Eine weitere Ausdifferenzierung oder Entscheidung zwischen den Hypothesen ist bei dieser Datenlage nicht möglich.

Summa summarum zeigt das Ergebnismuster der Untersuchungen zur mTTH ein relativ homogenes Bild für die Äquivalenzhypothese: Parameter I/a entspricht bei Silben der Kurzzeitgedächtnisspanne in bit. Alle Variationen der Kurzzeitgedächtnis-Experimente konnten (wenn auch in einem Fall nur tendenziell) in der Asymptotenhöhe nachgewiesen werden. Parameter a ist entgegen dem Baddeley - Modell nicht nur durch die spezifischen Subsysteme determiniert, sondern auch durch die Zentrale Exekutive. Bei verbalisierbarem, bildhaftem Material gelang die Gleichsetzung von Asymptote und Kurzzeitgedächtnisspanne nicht. In vorliegender Arbeit wird die Konfundierung der variierten Faktoren Lexikalität und Präsentationsmodus aufgelöst.

Parameter b , dessen Untersuchung im Gegensatz zu Parameter a hypothesenexplorierenden Charakter hatte, weist die Befunde eines Lexikalitätseffekts und der Möglichkeit eines phonologischen Zugangs zum Kurzzeitgedächtnis auf. Fragen, wie beispielsweise nach der spezifischen Wirkungsweise Artikulatorischer Unterdrückung bei figuralem Material oder nach der Relevanz der Zentralen Exekutive, können nur aufgedeckt, aber noch nicht beantwortet werden. In den folgenden beiden Kapiteln werden die aus den dargestellten Theorien abgeleiteten Hypothesen sowie das methodische Vorgehen beschrieben.

4 Ableitung und Formulierung der Hypothesen

Diese Arbeit verfolgt drei globale Ziele:

1. Nachdem die INVARIANZHYPOTHESE VON BADDELEY in letzter Zeit stark kritisiert worden ist (Kapitel 2), soll untersucht werden, ob diese kritischen Befunde repliziert werden können. Außerdem wird geprüft, ob der von HULME ET AL. (1995) postulierte Einfluß des Langzeitgedächtnisses auf die Kurzzeitgedächtnisspanne nachweisbar ist.
2. Als Alternative zur Invarianzhypothese von Baddeley wird eine INVARIANZHYPOTHESE NACH DEM ANSATZ VON DOSHER UND MA (1998) formuliert. Diese Formalisierung wird erstmals empirisch überprüft. Zudem wird untersucht, ob eine Generalisierung der Hypothese auf einen – im Vergleich zu der originalen Studie von Doshier und Ma (1998) - erweiterten Anwendungsbereich möglich ist. Eine systematische Gegenüberstellung mit den Ergebnissen zur Baddeley-Hypothese soll den Stellenwert aufzeigen. Ein weiteres Ziel der Analysen ist die Untersuchung intraexperimenteller Varianz der Gesamtwiedergabe-Zeit nach dem Ansatz von COWAN (1994, Kapitel 2.4.4).
3. Der Schwerpunkt dieser Arbeit liegt auf der Überprüfung der INVARIANZHYPOTHESE VON BREDENKAMP. Dabei werden vier Ziele verfolgt. Erstens wird die Möglichkeit einer Ausdehnung des intendierten Anwendungsbereichs geprüft. Die Gültigkeit der mTTH ist nachgewiesen für Stimulusmaterialien wie Silben und geometrische Figuren. Angestrebt wird eine Ausweitung des Gültigkeitsbereichs auf die in der Forschung am häufigsten realisierten Stimulusmaterialien: Zahlen und Worte. Das zweite Ziel bezieht sich auf den Nachweis universeller Gültigkeit der mTTH in Bezug auf die untersuchte Stichprobe. Bisher wurden alle Experimente zur mTTH an deutschen Studenten durchgeführt, in dieser Arbeit werden auch Probanden mit spanischer respektive chinesischer Muttersprache untersucht. Drittens wird die Möglichkeit der Verknüpfung der Invarianzhypothesen des Kurz- und Langzeitgedächtnisses durch die Äquivalenzhypothese geprüft. Als vierter Punkt wird explorativ dem Einfluß experimenteller Manipulation auf den Parameter der Verarbeitungszeit im Langzeitgedächtnisspeicher nachgegangen.

Im weiteren werden zuerst die empirischen und statistischen Hypothesen formuliert, dann folgt ein Überblick über den Aufbau der Studie. Die deskriptiven Charakteristika der Untersuchung sowie die Operationalisierung der abhängigen Variablen werden in Kapitel 5 dargestellt.

4.1 EMPIRISCHE UND STATISTISCHE HYPOTHESEN

In diesem Kapitel werden drei Arten von Hypothesen unterschieden. Durch ARBEITSHYPOTHESEN soll geprüft werden, ob die experimentelle Manipulation, die aus den Annahmen des Working-Memory-Modells abgeleitet wird, wirksam ist. Die Gruppe der ZENTRALEN HYPOTHESEN bezieht sich auf die aus der Theorie abgeleiteten Erwartungen zu den drei Invarianzhypothesen. Die EXPLORATIVEN HYPOTHESEN schließlich besitzen heuristischen Wert, sind aber beim derzeitigen Stand der Forschung noch nicht theoretisch abgesichert.

4.1.1 Arbeitshypothesen

Im folgenden werden Hypothesen zu sieben Effekten formuliert. Der Aufbau der Unterkapitel zur Begründung der Hypothesen ist für alle sieben Punkte äquivalent: Zuerst wird auf die empirische Basis der Effekte und deren theoretische Begründung hingewiesen. Dann folgt die Formulierung der empirischen Hypothese und deren Formalisierung als statistische Hypothese. In Klammern ist das Kapitel vorliegender Arbeit genannt, in dem sich die relevanten Auswertungen der empirischen Untersuchung zu der betreffenden Hypothese befinden.

Der Zusammenhang zwischen der Reproduktionszeit und der Höhe der Gedächtnisspanne beruht auf einer schmalen empirischen Evidenz mit wenig theoretischen Verknüpfungen (Kapitel 2.4.4). Aus diesem Grund können a priori keine Hypothesen über das Wirken experimenteller Variation auf den Parameter der Reproduktionszeit aufgestellt werden, mit Ausnahme der Wirkung phonologischer Ähnlichkeit. Dennoch wird versucht, durch Integration von Beobachtungen verschiedener Studien Prognosen zu erstellen; dies ist jedoch nicht für alle folgenden Effekte möglich.

1 ES GIBT EINEN EFFEKT DER LINGUISTISCHEN GRUPPE AUF DIE HÖHE DER KURZZEITGEDÄCHTNISSPANNE SOWIE AUF DIE SCHNELLIGKEIT DER ARTIKULATION UND REPRODUKTION

Der Effekt der linguistischen Gruppe wird in dieser Arbeit anhand von drei verschiedenen Sprachen überprüft: an der Studie nehmen Probanden mit deutscher, spanischer und chinesischer Muttersprache teil (Kapitel 4.2: Experiment 1, 3, 4).

Systematische Unterschiede in der Höhe der Kurzzeitgedächtnisspanne und der Schnelligkeit der Artikulation in Abhängigkeit von der Muttersprache sind in zahlreichen

Studien nachgewiesen worden. Relevant für die vorliegende Studie sind die Befunde von Lüer et al. (1998; Laß et al., 1999; Laß et al., 2000) sowie von Chincotta und Underwood (1997a).

Meines Wissens existiert keine vergleichende Studie zwischen deutschen und spanischen Probanden. Vergleicht man die Studie von Chincotta und Underwood (1997a) mit der Arbeit von Lüer et alii (1998), so kann wegen der verschiedenen Meßmethoden keine Prognose über diesen Sprachvergleich abgeleitet werden: Einerseits ist die Gedächtnisspanne im Spanischen höher als im Deutschen, andererseits aber ist die Artikulationszeit langsamer. Lüer et alii (1998) berichten beim Vergleich deutscher und chinesischer Probanden von Unterschieden in Bezug auf die Höhe der Gedächtnisspanne sowie von signifikant schnelleren Artikulationszeiten im Chinesischen für das auch in dieser Studie verwendete Stimulusmaterial von Zahlen. Bezüglich der Reproduktionszeit nach Doshier und Ma liegen keine Befunde vor. Aus bisherigen Ergebnissen kann aber eine Hypothese abgeleitet werden, wenn man die Gesamtwiedergabe-Zeit in ihre drei Komponenten zerlegt (Präparationsintervall, Pausenintervall und Artikulationszeit per se, Kapitel 2.4.4), da für jeden Teil bereits von sprachbedingten Unterschieden berichtet wurde:

- Laß (1997) wies signifikant schnellere Scanning-Zeiten für chinesische Probanden im Vergleich zu deutschen Probanden nach. Diese Zeit wird mit dem Präparationsintervall in Verbindung gebracht.
- Lüer et alii (1998) referieren, daß die Pausenintervalle im Deutschen deskriptiv fast dreimal länger sind als im Chinesischen.
- Höhere Artikulationsgeschwindigkeiten im Chinesischen wurden bereits erwähnt.

Alle drei Befunde lassen darauf schließen, daß ein Effekt der linguistischen Gruppe für die Reproduktionszeit zu erwarten ist in die Richtung, daß die Reproduktionsgeschwindigkeit pro Item bei Chinesen schneller ist. Für den Vergleich zwischen chinesischen und spanischen Probanden belegen Chincotta und Underwood (1997a), daß chinesische Probanden signifikant höhere Gedächtnisspannenwerte und erheblich schnellere Artikulationszeiten aufweisen. Für die Reproduktionszeit wird vermutet, daß derselbe Zusammenhang wie zwischen den deutschen und chinesischen Versuchspersonen besteht.

Theoretisch begründet wird der Effekt der linguistischen Gruppe durch die Invarianzhypothese von Baddeley (Kapitel 3). Ob eine Erklärung auch durch die Invarianzhypothese von Doshier und Ma erfolgen kann, wird geprüft.

Als EMPIRISCHE UND STATISTISCHE HYPOTHESEN lassen sich folgende Prognosen formulieren, die auch eine Aussage über die globale Fragestellung zur Invarianzhypothese von Baddeley erlauben:

- Zwischen den verschiedenen Sprachgruppen bestehen signifikante Unterschiede in der Artikulationszeit. Es wird erwartet, daß die Artikulationszeit im Chinesischen kürzer ist als die der beiden anderen Sprachgruppen. Als statistische Hypothese wird folgender Unterschied der Mittelwerte (u) erwartet (Kapitel 7.2.1), wobei die Parameterbezeichnung jeweils den in Kapitel 3 eingeführten Abkürzungen entspricht. Der Index verweist auf die Experimentalgruppe: c steht für die linguistische Gruppe des Chinesischen, respektive d für das Deutsche.

$$H_0: \quad u_c \geq u_d \text{ und } u_c \geq u_s$$

$$H_1: \quad u_c < u_d \text{ und } u_c < u_s$$

Falls sich zwischen dem Chinesischen und den beiden anderen Sprachen Unterschiede entgegen der Erwartung ergeben, wird die empirische Hypothese aufgegeben. Als bedingt bewährt wird die empirische Hypothese angesehen, wenn der Unterschied zwischen den chinesischen Probanden und den beiden anderen Gruppen deskriptiv in die erwartete Richtung zeigt, nicht aber die Signifikanzgrenze überschreitet. Die empirische Hypothese wird akzeptiert, falls die Unterschiede zwischen der chinesischen Gruppe und den beiden anderen Gruppen signifikant sind.

- Ein Effekt der linguistischen Gruppe auf die Höhe der Gedächtnisspanne wird erwartet. Es wird prognostiziert, daß bei chinesischen Probanden die Kurzzeitgedächtnisspanne signifikant höher ist als bei den beiden anderen Sprachgruppen. Die Umsetzung in eine statistische Hypothese sieht vor, daß die empirische Hypothese abgelehnt wird, wenn zwischen den beiden Sprachgruppen und dem Chinesischen der Erwartung gegenläufige Unterschiede aufgedeckt werden. Bedingt bewährt sich diese Hypothese, wenn die Abweichung zwischen dem Chinesischen und den beiden anderen Gruppen nicht signifikant ist, aber deskriptiv in die erwartete Richtung zeigt. Angenommen wird die empirische Hypothese, wenn Unterschiede zwischen den beiden Gruppen bestehen. Realisiert wird als statistische Hypothese folgende Überprüfung:

$$H_0: \quad g_c \leq g_d \text{ und } g_c \leq g_s$$

$$H_1: \quad g_c > g_d \text{ und } g_c > g_s$$

- Bezüglich der Reproduktionszeit wird folgende Prognose abgeleitet: Wenn ein Effekt der linguistischen Gruppe auf die Schnelligkeit der Artikulation und auf die

Kurzzeitgedächtnisspanne aufgedeckt wird, wird ein Effekt im Parameter der Reproduktionszeit pro Item erwartet. Dies ist darin begründet, daß bei unterschiedlichen Gedächtnisspannen auch die Reproduktionszeiten abweichen müssen, damit der Ansatz von Doshier und Ma (1998) Bestand hat. Zudem ist die Artikulationszeit ein integraler Bestandteil der Reproduktionszeit. Auf statistischer Ebene werden Mittelwertunterschiede erwartet:

$$H_0: d_c \geq d_d \text{ und } d_c > d_s$$

$$H_1: d_c < d_d \text{ und } d_c < d_s$$

Da keine Erwartungen über die Größe des Effekts vorliegen, wird die empirische Hypothese auch dann angenommen, wenn lediglich deskriptiv das erwartete Verhältnis der drei Sprachgruppen eintritt. Abgelehnt wird die Hypothese, falls sich die Rangfolge im Gegensatz zu den Auswertungen zur Gedächtnisspanne verändert.

2 BEI SIMULTANER ARTIKULATORISCHER UNTERDRÜCKUNG REDUZIEREN SICH DIE KURZZEITGEDÄCHTNISSPANNEN

Die Reduktion der Kurzzeitgedächtnisspanne durch Artikulatorische Unterdrückung ist in zahlreichen Experimenten belegt worden (Chincotta & Underwood, 1997a). Zu einer Wirkung der Sekundäraufgabe auf die Reproduktionszeit liegen keine Erkenntnisse vor.

Die *theoretische Begründung* findet sich in der Erklärung innerhalb des Arbeitsgedächtnis-Modells: Wenn subvokale und laute Artikulation dieselben Prozesse sind, dann blockiert das Aussprechen einer irrelevanten Silbe das subvokale Wiederholen der Items (Baddeley et al., 1975). Die Gedächtnisspuren sind im Phonologischen Speicher kürzer präsent (Kapitel 2.4.1).

- Als EMPIRISCHE HYPOTHESE wird die Prognose formuliert, daß die Höhe der Kurzzeitgedächtnisspanne absinkt, wenn simultan die Artikulation des Zielitems unterdrückt wird. Auf Ebene der statistischen Hypothese ergibt sich folgender Zusammenhang (Kapitel 7.2.2).

$$H_0: g_{\text{STILLE}} \leq g_{\text{ARTU}}$$

$$H_1: g_{\text{STILLE}} > g_{\text{ARTU}}$$

Der Index „Stille“ steht für die Experiment ohne Zusatzaufgabe, die Abkürzung „Artu“ bezeichnet die Experimentalbedingung „Artikulatorische Unterdrückung“. Die Hypothese wird akzeptiert, wenn sich eine signifikante Reduktion der Kurzzeitgedächtnisspanne bei zusätzlicher Unterdrückung ergibt. Als bedingt bewährt wird die Hypothese interpretiert, wenn sich nur deskriptiv Abweichungen erkennen

lassen, die aber nicht die Signifikanzgrenze erreichen. Abgelehnt wird die empirische Hypothese, wenn sich ein erwartungskonträrer Zusammenhang ergibt.

3 EINE SIMULTAN AUSZUFÜHRENDE ARTIKULATORISCHE SEKUNDÄRAUFGABE ELIMINIERT DEN EFFEKT DER LINGUISTISCHEN GRUPPE

Empirische Evidenz belegt die Existenz eines Interaktionseffektes zwischen Sekundäraufgabe und Muttersprache (Chincotta & Underwood, 1997a). Ohne Zusatzaufgabe konnten bedeutsame Unterschiede zwischen den Sprachgruppen in der Höhe der Gedächtnisspanne nachgewiesen werden. Dieser Effekt verschwand bei simultaner Artikulatorischer Unterdrückung.

Eine *theoretische Erklärung* für dieses Phänomen liefert das Modell von Baddeley zur Erklärung des Wortlängeneffekts: durch die Blockade des Prozesses subvokaler Wiederholung können Unterschiede in der Artikulationszeit, die zwischen den Sprachen bestehen, nicht mehr zum Tragen kommen. An dieser Stelle wird der Bedeutungsüberschuß des Arbeitsgedächtnis-Modells gegenüber der Invarianzhypothese deutlich: Durch das mathematische Modell wird dieser Interaktionseffekt nicht impliziert.

- Als EMPIRISCHE HYPOTHESE wird prognostiziert, daß der Spracheffekt in Bezug auf die Kurzzeitgedächtnisspanne – so er in den Daten nachweisbar ist – bei artikulatorischer Zusatzfähigkeit verschwindet (Kapitel 7.2.2). Untersucht wird folgende statistische Hypothese. Der Index „c,artU“ steht für Experiment 5 und „d,artU“ für Experiment 2.

$$H_0: g_{c,artU} = g_{d,artU}$$

$$H_1: g_{c,artU} < g_{d,artU} \text{ oder } g_{c,artU} > g_{d,artU}$$

Die empirische Hypothese wird beibehalten, wenn zwischen den Sprachgruppen keine systematischen Abweichungen nachweisbar sind. Die Hypothese wird verworfen, wenn Unterschiede bestehen, unabhängig von der deskriptiven Richtung.

4 ES BESTEHEN UNTERSCHIEDE JE NACH GESCHLECHT DER VERSUCHSPERSON

Garai und Scheinfeld (1968) berichten von empirisch aufgedeckten Unterschieden in den Parametern „Artikulationszeit“ und „Höhe der Gedächtnisspanne“ in Abhängigkeit vom Geschlecht der Probanden. Es zeigten sich sowohl eine schnellere Artikulationsgeschwindigkeit als auch höhere Werte für die Kurzzeitgedächtnisspanne bei weiblichen Versuchspersonen. Die *theoretische Begründung* findet sich in der

Invarianzhypothese von Baddeley. Die Prüfung des Geschlechtseffekts dient der Steigerung der Validität, wenn nachgewiesen werden kann, daß die globale Hypothese auch in relevanten Untergruppen Bestand hat (Westermann, 2000).

- Folgende Prognose wird als EMPIRISCHE HYPOTHESE aufgestellt: Wenn nachgewiesen werden kann, daß Frauen im Mittel schneller artikulieren als Männer, dann wird auch eine höhere Kurzzeitgedächtnis-Leistung bei den Teilnehmerinnen erwartet. Eine statistische Hypothese wird nur formuliert, wenn Unterschiede in der Artikulationszeit aufgedeckt werden (Kapitel 7.2.3).

$$\text{falls } u_{\text{Frauen}} > u_{\text{Männer}}: \quad H_0: \quad g_{\text{Frauen}} \leq g_{\text{Männer}}$$

$$\quad \quad \quad H_1: \quad g_{\text{Frauen}} > g_{\text{Männer}}$$

Die Hypothese wird auf empirischer Ebene zurückgewiesen, wenn sich ein erwartungskonträrer, signifikanter Zusammenhang ergibt, daß bei einer Untergruppe sowohl die schnelleren Artikulationszeiten als auch die niedrigeren Werte in dem Parameter „Kurzzeitgedächtnisspanne“ auftreten.

- Falls sich Artikulationszeitunterschiede zwischen männlichen und weiblichen Teilnehmern nachweisen lassen, wird auch vermutet, daß die Reproduktionszeiten variieren. Falls jedoch keine Unterschiede in der Off-Line Messung bestehen, existiert keine Hypothese für die On-Line Messung.

$$\text{falls } u_{\text{Frauen}} > u_{\text{Männer}}: \quad H_0: \quad d_{\text{Frauen}} \geq d_{\text{Männer}}$$

$$\quad \quad \quad H_1: \quad d_{\text{Frauen}} < d_{\text{Männer}}$$

Die empirische Hypothese wird zurückgewiesen, falls sich in einer Gruppe höhere Artikulationszeiten und niedrigere Reproduktionszeiten ergeben.

5 DER GESCHLECHTSEFFEKT KANN DURCH ARTIKULATORISCHE UNTERDRÜCKUNG ELIMINIERT WERDEN

Der Effekt zur Eliminierung der geschlechtsspezifischen Unterschiede durch eine artikulatorische Zusatzaufgabe folgt der gleichen Logik wie der Interaktionseffekt zwischen Sekundäraufgabe und Muttersprache. Aus diesem Grund wird zur theoretischen Begründung auf Punkt 3 verwiesen. Ein Effekt der Zusatzaufgabe kann nur dann geprüft werden, wenn geschlechtsspezifische Unterschiede nachgewiesen werden können. Als EMPIRISCHE PROGNOSE wird formuliert, daß der Geschlechtseffekt in Bezug auf die Kurzzeitgedächtnisspanne verschwindet, wenn simultan unterdrückt wird.

$$H_0: \quad g_{\text{Frauen}} = g_{\text{Männer}}$$

$$H_1: \quad g_{\text{Frauen}} < g_{\text{Männer}} \text{ oder } g_{\text{Frauen}} > g_{\text{Männer}}$$

Wenn sich auf statistischer Ebene keine signifikanten Unterschiede zeigen, werden die statistische und die empirische Hypothese angenommen. Bestehen dagegen weiterhin systematische Abweichungen unabhängig vom Trend, werden die Hypothesen verworfen.

6 ZWISCHEN VERSCHIEDENEN MATERIALIEN BESTEHEN UNTERSCHIEDE IN ARTIKULATIONSZEIT, HÖHE DER GEDÄCHTNISSPANNE UND REPRODUKTIONSZEIT

Bei der Hypothesenformulierung wird der Effekt verschiedener Materialien als ein Effekt klassifiziert. In der Auswertung werden dagegen zwei verschiedene Materialeffekte unterschieden. Der erste Effekt bezieht sich auf Unterschiede innerhalb der Materialklasse „Worte“ und der zweite auf Abweichungen zwischen den Materialklassen „Worte“ und „Zahlen“.

Für beide Materialeffekte existiert empirische Evidenz: Innerhalb der Wortklasse wird der Effekt Phonologischer Ähnlichkeit nachgewiesen (Kapitel 2.4.3). Dieser Effekt wird durch das Arbeitsgedächtnis-Modell theoretisch so begründet, daß bei steigendem Grad phonologischer Verwandtschaft die Interferenz zwischen den Items im Phonologischen Speicher ansteigt und deshalb die Kurzzeitgedächtnisspanne sinkt. Der Effekt bezüglich des Einflusses der Ähnlichkeit zwischen den Items auf die Artikulationszeit ist fraglich: In der Literatur wird sowohl über abweichende als auch über äquivalente Sprechgeschwindigkeiten berichtet (Schweickert et al., 1990). Nach der Studie von Doshier und Ma (1998) sind die Befunde für die Reproduktionszeit bei variierender Itemähnlichkeit eindeutig: Bei phonologisch ähnlichen Items sinkt die Reproduktionszeit pro Item (Kapitel 3).

Für Materialunterschiede ohne Variation der Klangähnlichkeit der Items wird der Erklärungsansatz von Baddeley bezüglich des Wortlängeneffekts herangezogen (Kapitel 2.4.4). Bei differierender Artikulationszeit zwischen den Materialien werden Unterschiede bezüglich der Kurzzeitgedächtnisspannen dahingehend prognostiziert, daß bei schnellerer Artikulationszeit die Höhe der Gedächtnisspanne steigt.

Folgende EMPIRISCHE UND STATISTISCHE HYPOTHESEN werden für den Materialeffekt innerhalb der Wortklasse formuliert (Kapitel 7.2.4):

- Innerhalb der Materialklasse „Worte“ bestehen Unterschiede in den Kurzzeitgedächtnisspannen dahingehend, daß die Leistung mit zunehmendem Grad an Itemähnlichkeit sinkt. Als statistische Formulierung ergeben sich folgende

Erwartungen über Mittelwertsunterschiede, wobei mit den Indices „unähnlich“ und „ähnlich“ jeweils die Distinktheit der Stimuli gemeint ist.

$$H_0: \quad g_{\text{unähnlich}} \leq g_{\text{ähnlich}}$$

$$H_1: \quad g_{\text{unähnlich}} > g_{\text{ähnlich}}$$

Die empirische Hypothese wird bei Ergebnissen entgegen der erwarteten Richtung verworfen, bei nicht signifikanten Unterschieden mit erwartetem Trend bewährt sie sich bedingt und bei Annahme von H_1 wird auch die empirische Hypothese akzeptiert.

- Für die Reproduktionszeit wird erwartet, daß die Geschwindigkeit bei geringer Ähnlichkeit höher ist als bei ähnlichen Items. Statistisch wird dieser Zusammenhang durch folgende Hypothesen ausgedrückt:

$$H_0: \quad d_{\text{unähnlich}} \geq d_{\text{ähnlich}}$$

$$H_1: \quad d_{\text{unähnlich}} < d_{\text{ähnlich}}$$

Analog den Hypothesen zur Gedächtnisspanne gilt auch die empirische Hypothese als falsifiziert, wenn sich ein der Erwartung gegenläufiger Trend nachweisen läßt. Bei Nachweis eines Trends der Unterschiede in die erwartete Richtung wird die Hypothese beibehalten.

Bezüglich der phonologisch unähnlichen Materialien „Worte“ und „Zahlen“ besteht nur eine Hypothese (Kapitel 7.2.4): Wenn sich zwischen den beiden Materialarten unterschiedliche Artikulationszeiten ergeben ($u_{\text{material1}} < u_{\text{material2}}$), dann werden auch Abweichungen in der Höhe der Gedächtnisspanne erwartet. Eine Hypothese dazu, ob Unterschiede zwischen den hier realisierten Materialien in der Schnelligkeit der Aussprache bestehen, existiert nicht, da mit den Wortstimuli von Experiment 7 (Kapitel 5.2.2) noch keine Studie durchgeführt wurde. „Material 1“ und „Material 2“ werden erst durch die empirischen Ergebnisse des Experiments zur Artikulationszeitmessung definiert.

Wenn es sich ergibt, daß die Artikulationszeit bei Zahlen schneller ist, so werden diese als „Material 1“ und die Worte als „Material 2“ bezeichnet, vice versa.

$$\text{falls } u_{\text{material1}} > u_{\text{material2}}: \quad H_0: \quad g_{\text{material1}} \leq g_{\text{material2}}$$

$$H_1: \quad g_{\text{material1}} > g_{\text{material2}}$$

Die empirische Hypothese wird zurückgewiesen, wenn ein der Erwartung gegenläufiger Trend beobachtet wird.

Falls keine Unterschiede in der Artikulationszeit nachweisbar sind ($u_{\text{material1}} = u_{\text{material2}}$), werden auch keine Differenzen in den Kurzzeitgedächtnisspannen erwartet. Diese Erwartung entspricht auf statistischer Ebene der Nullhypothese.

falls $u_{\text{material1}} = u_{\text{material2}}$: $H_0: g_{\text{material1}} = g_{\text{material2}}$

$H_1: g_{\text{material1}} \neq g_{\text{material2}}$

Die empirische Hypothese wird zurückgewiesen, wenn sich ergibt, daß bei einer Materialart sowohl eine längere Artikulationszeit als auch eine höhere Gedächtnisspanne nachweisbar ist. Die Hypothese wird als bewährt angesehen, wenn sich im interexperimentellen Vergleich Abweichungen in Artikulationszeit und Kurzzeitgedächtnisspanne ergeben, die den Erwartungen entsprechen, auch, wenn die Unterschiede nicht in beiden Parametern die Signifikanzgrenze überschreiten, sondern nur in einem von beiden Tests.

- Bezüglich der Reproduktionszeit wird vermutet, daß bei verschiedenen Artikulationszeiten auch die Reproduktionszeiten variieren. Falls die Artikulationsgeschwindigkeiten für die verschiedenen Materialien nicht voneinander abweichen, liegt keine Hypothese vor.

7 ES EXISTIERT KEIN EFFEKT DES PRÄSENTATIONSMODUS

Wie in der Studie von Lürer et alii (1998) berichtet wird, treten keine Veränderungen in den Parametern „Gedächtnisspanne“ und „Artikulationszeit“ für Zahlen auf, wenn die Stimuli als Symbole oder als ausgeschriebene Worte präsentiert werden. Andererseits zeigt die Studie von Chincotta und Underwood (1997b), daß zwischen den Präsentationsarten Unterschiede bestehen bei Einführung einer Sekundäraufgabe.

Im Rahmenmodell des Arbeitsgedächtnisses wird *theoretisch begründet*, warum es keinen Effekt des Präsentationsmodus geben sollte: Unabhängig von der Darbietungsart ändert sich die phonologische Kodierung der Stimuli nicht. Wenn sich die Artikulationszeit nicht ändert, werden äquivalente Kurzzeitgedächtnisspannen erwartet. Der divergente Effekt Artikulatorischer Unterdrückung läßt sich durch das Arbeitsgedächtnismodell nicht erklären. Die empirischen und statistischen Hypothesen für den Präsentationseffekt lauten folgendermaßen (Kapitel 7.2.5):

- Bezüglich der Artikulationszeit besteht kein Effekt des Präsentationsmodus. Statistisch werden keine Mittelwertunterschiede je nach Darbietung der Ziffern als Wort oder als Symbol erwartet. Dies entspricht der Nullhypothese:

$H_0: u_{\text{Wort}} = u_{\text{Symbol}}$

$H_1: u_{\text{Wort}} \neq u_{\text{Symbol}}$

- Analog werden nach Baddeley keine Abweichungen in der Kurzzeitgedächtnisspanne je nach Präsentationsmodus erwartet. Sowohl für die Artikulationszeit als auch für die Höhe der Kurzzeitgedächtnisspanne wird die empirische Hypothese abgelehnt, sollten sich signifikante Unterschiede (unabhängig von der Richtung) ergeben. Im Zusammenhang mit dem globalen Ziel der Überprüfung des Arbeitsgedächtnis-Modells würde bei Ablehnung der empirischen Hypothesen ein sehr kritisches Datum für das Modell an sich entstehen.

$$H_0: g_{\text{Wort}} = g_{\text{Symbol}}$$

$$H_1: g_{\text{Wort}} \neq g_{\text{Symbol}}$$

- Für den Parameter der Reproduktionszeit liegen keine Hypothesen vor. Eventuell erzeugen die beiden Präsentationsmodi unterschiedliche Effekte bezüglich nicht näher spezifizierter Output-Prozesse, die in die Messung der Reproduktionszeit einfließen.

8 ALLE GEFUNDENEN EFFEKTE SIND AUCH IN EINZELFALLSTUDIEN NACHWEISBAR

Diese Hypothese bezieht sich auf die Äquivalenz zwischen Aggregat- und Einzelfalluntersuchungen. Die Analysen resultieren aus der Kritik Westmeyers (1987), daß aufgrund interindividueller Unterschiede Aussagen über Prozesse innerhalb von Individuen abgeleitet werden. Ein methodisches Problem entsteht aus dieser Praxis nicht, wenn die Hypothesen – wie in dieser Arbeit versucht – so formuliert werden, daß man bei deren Prüfung sich zu einem Zurückweisen der psychologischen Theorie entscheiden kann (Erdfelder & Bredenkamp, 1994). Im folgenden wird für vier exemplarisch ausgewählte Effekte (Effekt der Sprache, der Artikulatorischen Unterdrückung, der Materialien und des Präsentationsmodus) die Äquivalenz zwischen intra- und interindividuellen Bedingungsvariation geprüft.

4.1.2 ZENTRALE HYPOTHESEN

Die Zentralen Hypothesen werden zur Untersuchung der Gültigkeit der drei Invarianzhypothesen formuliert (Kapitel 3). Die Auswertung wird in zwei Kapiteln dargestellt: in Kapitel 6 finden sich Analysen, die innerhalb einer Experimentreihe durchgeführt werden, während in Kapitel 7 zwischen verschiedenen Experimentreihen verglichen wird. Die Trennung zwischen den beiden Kapiteln 6 und 7 liegt allein in der Übersichtlichkeit begründet und basiert nicht auf inhaltlichen Divergenzen.

Die Invarianzhypothesen des Kurzzeitgedächtnisses von Baddeley sowie von Doshier und Ma beziehen sich auf einen Vergleich verschiedener Experimente, die in dem Kapitel zur interexperimentellen Hypothesenprüfung (Kapitel 7) ausgewertet werden. Für die Invarianzhypothese zur Verbindung von Kurz- und Langzeitgedächtnisprozessen, der mTTH, liegt der Schwerpunkt der Überprüfung auf dem Nachweis der Gültigkeit innerhalb jeder einzelnen Experimentreihe (Kapitel 6: Intraexperimentelle Hypothesenprüfung). Wenn die mTTH bestätigt werden kann, dann können durch interexperimentelle Vergleiche weitere Hypothesen, die aus der mTTH abgeleitet werden, einer Überprüfung unterzogen werden (Kapitel 7: Interexperimentelle Hypothesenprüfung).

1 DIE INVARIANZHYPOTHESE VON BADDELEY IST UNIVERSELL GÜLTIG

Diese Hypothese basiert auf reicher empirischer Evidenz sowohl für verschiedene Materialien als auch für verschiedene Populationen (Kapitel 2.4.4). Das Postulat der invarianten Persistenz der Gedächtnisspur ist nicht theoretisch abgeleitet, sondern resultiert aus induktiver Beobachtung des Zusammenhangs zwischen Artikulationszeit und Gedächtnisspanne. Als plausible Theorie, wie diese Relation zustande kommt, postuliert Baddeley das in Kapitel 2.4.4 beschriebene Erklärungsmodell zum Wortlängeneffekt. Durch die Formalisierung als Invarianzhypothese werden die Adäquatheitsbedingungen wissenschaftlicher Erklärung nach Hempel (1977) erfüllt: Die Invarianzhypothese ist ein empirisch bestätigtes, quantitatives Gesetz, das eine allgemeine Aussage über eine unerschöpfliche Anzahl von Individuen erlaubt (Kapitel 2). Neben einer „starken“ Auslegung der Invarianzhypothese, die auch eine interindividuelle Konstanz vermutet, kann auch eine „schwache“ Interpretation formuliert werden, in denen Unterschiede zwischen Individuen zugelassen werden.

Unter der Annahme, daß die Dauer der Gedächtnisspur nicht interindividuell variiert, wird die EMPIRISCHE HYPOTHESE formuliert, daß zwischen den Parametern „Gedächtnisspanne“ und „Artikulationszeit“ ein negativer Zusammenhang besteht. Statistisch sollte sich diese Hypothese als negativer Korrelationskoeffizient sowohl innerhalb der einzelnen Experimentreihen als auch im Verbund der Experimente manifestieren. Die postulierte Invarianz der Persistenz der Gedächtnisspur entspricht statistisch folgender Nullhypothese:

$$H_0: \quad B_1 = B_3 = B_4 = B_7 = B_8$$

$$H_1: \quad \neg H_0$$

Die Indices beziehen sich auf die durchgängige Numerierung der Experimente (Kapitel 4.2). Falls sich auf statistischer Ebene bedeutsame Unterschiede ergeben, wird die empirische Hypothese als falsifiziert betrachtet, unabhängig vom spezifischen Ergebnismuster. Von der

Auswertung ausgeschlossen werden die Experimente, für die nach Baddeley der beschriebene Zusammenhang zwischen Artikulationszeit und Kurzzeitgedächtnisspanne nicht besteht: für Experimente mit Sekundäraufgabe, respektive mit phonologisch ähnlichen Items.

2 DIE INVARIANZHYPOTHESE VON DOSHER UND MA IST UNIVERSELL GÜLTIG

In der Studie von Doshier und Ma (1998) wird belegt, daß die Gesamtwiedergabe-Zeit von phonologisch ähnlichem, respektive unähnlichem Material eine Konstante ist. Wissenschaftstheoretisch sind die Invarianzhypothesen von Baddeley sowie Doshier und Ma nicht äquivalent. Während die Hypothese von Baddeley als gesetzesartiger Satz in das Modell wissenschaftlicher Erklärung integriert wird, gilt es für die Hypothese von Doshier und Ma, ihre Gesetzesartigkeit empirisch zu belegen.

Begründet wird der Phonologische Ähnlichkeitseffekt von den Autoren damit, daß bei Interferenz die Reproduktionszeit pro Item steigt. Wenn die Wiedergabezeit eine Konstante ist, muß bei langsamerer Reproduktionszeit die Kurzzeitgedächtnisspanne sinken (Kapitel 2.4.4). Im Gegensatz zur Theorie von Baddeley wird nicht theoretisch begründet, warum die Wiedergabezeit eine Konstante sein soll oder warum sie in dem Wertebereich zwischen vier und sechs Sekunden liegen soll. Auch ist nicht spezifiziert, welche „Output-Prozesse“ (Cowan et al., 1992) außer Interferenz und Artikulationszeit die Höhe der Reproduktionszeit determinieren. Im Prinzip bedeutet das, daß aufgrund von Beobachtung zwar ein Gesetz induktiv formuliert, aber noch nicht mit Theorien verknüpft wurde. Bei der Diskussion und Interpretation der Ergebnisse wird dies noch einmal aufgegriffen. Für die Überprüfung des Gesetzes dagegen ergibt sich kein Problem: Aus der Invarianzhypothese und einem empirisch beobachteten Sachverhalt lassen sich eindeutige Prognosen ableiten.

Nachdem in den Arbeitshypothesen die Wirkung experimenteller Variation auf die Reproduktionszeit formuliert wurde, wird nun prognostiziert, daß folgender Zusammenhang besteht: Wenn für ein bestimmtes Experiment im Vergleich zu einem anderen im Mittel kürzere Reproduktionszeiten gemessen werden, wird eine höhere Gedächtnisspanne für diesen Materialtyp erwartet. Analog der Interpretation zur Invarianzhypothese von Baddeley sind zwei Auslegungen möglich. Falls zwischen den Probanden keine bedeutsamen Unterschiede bestehen, sollte sich ein negativer Zusammenhang von Gedächtnisspanne und Reproduktionszeit ergeben. Zur Überprüfung der Hypothese, daß die Wiedergabezeit aller Experimente konstant ist, werden auf statistischer Ebene keine Mittelwertsunterschiede erwartet (Kapitel 7.4). Dies entspricht der Nullhypothese.

$$H_0: D_1 = D_2 = D_3 = D_4 = D_5 = D_6 = D_7 = D_8$$

$$H_1: \neg H_0$$

Die Verbindung zwischen statistischer und empirischer Hypothese ist im Fall der Hypothese von Doshier und Ma wegen der eingeschränkten theoretischen Verknüpfung nicht eindeutig. In einer abschließenden Analyse zur Invarianzhypothese von Doshier und Ma soll deren Stellenwert gegenüber dem Modell von Baddeley bestimmt werden.

3 DIE mTTH IST UNIVERSELL GÜLTIG

Die Prüfung der universellen Gültigkeit der mTTH ist die wichtigste Hypothese dieser Arbeit. Es besteht empirische Evidenz für die Gültigkeit der mTTH bei Silbenmaterial und geometrischen Figuren (Kapitel 3). Geprüft wird, ob eine Übertragung auf die lexikalen Materialien (Worte und Zahlen) sowie auf verschiedene Populationen möglich ist. Auf die empirische Hypothese der Anpassung einer hyperbolischen Funktion pro Material unabhängig vom Informationsgehalt wird verzichtet, da sich diese Hypothese durchweg gut bestätigt hat.

Als empirische Hypothese wird formuliert, daß die Formel der mTTH auch für lexikales Material angepaßt werden kann. Auf statistischer Ebene werden keine Abweichungen zwischen den empirisch ermittelten Daten und den aus der mTTH prognostizierten Werten erwartet (Kapitel 6.2). Werden keine signifikanten Unterschiede aufgedeckt, wird eine hyperbolische Funktion der verarbeiteten Information mit der Darbietungszeit angenommen. Aus der mTTH können noch weitere Hypothesen abgeleitet werden: die Äquivalenzhypothese und – explorativ - Aussagen über die Verarbeitungszeit im Langzeitgedächtnisspeicher.

4 DIE ÄQUIVALENZHYPOTHESE DER mTTH IST GÜLTIG

Die Äquivalenzhypothese besagt, daß die Asymptote der Lernfunktion mit der Kurzzeitgedächtnisspanne in bit gleichgesetzt werden kann. Für diese Hypothese besteht für Silbenmaterial empirische Evidenz, für figurationales Material dagegen mußte diese Annahme zurückgewiesen werden.

Es werden zwei empirische und statistische Hypothesen entwickelt:

- Für jede Experimentreihe wird erwartet, daß die Kurzzeitgedächtnisspanne in bit und die Asymptote der Lernkurve äquivalent sind (Kapitel 6.3). Die empirische Hypothese entspricht in diesem Fall der statistischen Nullhypothese.

H_0 : Gedächtnisspanne = Asymptote

H_1 : Gedächtnisspanne \neq Asymptote

Der Zusammenhang zwischen statistischer und empirischer Hypothese ist eindeutig: Falls sich signifikante Abweichungen ergeben, muß die empirische Hypothese abgelehnt werden.

- Zweitens wird überprüft, ob die experimentell induzierten Variationen in der Kurzzeitgedächtnisspanne (Kapitel 7.2) auch für die Asymptotenhöhe nachweisbar sind (Kapitel 7.5). A priori können keine spezifischen Hypothesen gebildet werden, da erst in den Arbeitshypothesen geprüft werden muß, wie die experimentelle Variation die Kurzzeitgedächtnisspanne affiziert. Insgesamt sind zwei Zusammenhänge zwischen diesem Parameter und der Asymptote erwartungskonform:

Falls nach experimenteller Manipulation der unabhängigen Variablen keine Veränderung in der Kurzzeitgedächtnisspanne beobachtet wird, so werden auch keine Unterschiede in der Asymptotenhöhe erwartet. Diese Hypothese wird abgelehnt, wenn sich signifikante Unterschiede in der einen Variablen zeigen, nicht jedoch in der anderen.

Falls ein signifikanter Einfluß der Bedingung für die Kurzzeitgedächtnisspanne nachgewiesen wird, wird prognostiziert, daß äquivalente Unterschiede in der Höhe der Asymptote bestehen. Akzeptiert wird diese Annahme immer dann, wenn äquivalente Unterschiede auftreten, unabhängig vom Erreichen der Signifikanzgrenze.

4.1.3 EXPLORATIVE HYPOTHESEN

Als explorative Hypothesen werden Annahmen formuliert, die zwar nicht aus einer Theorie generiert, aber aufgrund der bisherigen empirischen Evidenz zu erwarten sind.

1 FÜR DIE KURZZEITGEDÄCHTNISSPANNE IST EIN EINFLUSS DES LANGZEITGEDÄCHTNISSES NACHWEISBAR

Hulme et alii (1995) belegen in ihrer Studie, daß es einen Einfluß des Langzeitgedächtnisses auf die Kurzzeitgedächtnisspanne gibt. Da in der vorliegenden Arbeit lexikales Material untersucht wird, kann vermutet werden, daß auch hier ein Effekt des Langzeitgedächtnisses nachweisbar ist. Statistisch wird der Effekt des Langzeitgedächtnisses geprüft durch den Ordinatenabschnitt (a_{Hulme}), der sich bei der Regressionsanalyse mit dem Kriterium der Artikulationsrate und dem Prädiktor der Gedächtnisspanne ergibt. In der Studie von Bosshardt und Laug (1995) war bei lexikalem

Material a_{Hulme} signifikant von Null verschieden. Dieses Ergebnis wird als Hypothese für die Auswertung vorliegender Experimentreihen formuliert. Verworfen wird die Hypothese, wenn a_{Hulme} statistisch gleich 0 ist (Kapitel 6.2.6). Der Index „Hulme“ bei den Parametern dieser Regression dient der Abgrenzung gegenüber den Regressionparametern der mTTH (Kapitel 4.1.2).

$$H_0: a_{\text{Hulme}} = 0$$

$$H_1: a_{\text{Hulme}} > 0$$

Explorativ wird geprüft, ob ein variierender Einfluß des Langzeitgedächtnisses je nach Materialart und Population entsteht (Kapitel 7.3).

2 ES BESTEHEN BEZÜGLICH DER REPRODUKTIONSZEIT KEINE INTERINDIVIDUELLEN UNTERSCHIEDE (COWAN, 1994)

„Instead, subjects with different memory spans recalled their span-length lists at comparable rates of about 1 s for each correctly recalled item, resulting in longer lasting recall for more capable subjects.” (Cowan, 1994, S. 188).

Interpretiert wird dieser Befund so, daß während der Wiedergabe an sich andere Prozesse (beispielsweise die Rekonstruktion) eine größere Rolle spielen als die Fähigkeit zur schnellen Artikulation. Interessant für diese Studie ist die Beobachtung von Cowan deshalb, weil sie erstens der Invarianzhypothese von Doshier und Ma widerspricht und zweitens, weil durch sie Binnenvarianz innerhalb der Experimentreihen teilweise aufgeklärt werden könnte.

Als empirische Hypothese wird untersucht, ob die Reproduktionszeit interindividuell invariant und unabhängig von den Kurzzeitgedächtnisleistungen ist (Kapitel 7.4). Auf statistischer Ebene werden keine Mittelwertunterschiede im Parameter der Reproduktionszeit pro Item zwischen den Gruppen mit einer hohen Wiedergabezeit (Gruppe A) und der Gruppe mit niedriger Gesamtwiedergabe-Zeit (Gruppe B) erwartet. Die Gruppeneinteilung wird errechnet durch das erste und letzte Quartil für die Variable „Gesamtwiedergabe-Zeit“. Die prognostizierte Hypothese entspricht der Nullhypothese.

$$H_0: \gamma_{\text{gruppeA}} = \gamma_{\text{gruppeB}}$$

$$H_1: \gamma_{\text{gruppeA}} \neq \gamma_{\text{gruppeB}}$$

Das Auftreten statistischer Unterschiede besagt, daß die Beobachtung von Cowan hier nicht zutrifft.

3 DIE VERARBEITUNGSZEIT IM LANGZEITGEDÄCHNISSPEICHER IST INVARIANT, SOFERN DIE MÖGLICHKEIT ZUR ARTIKULATION BESTEHT

In den Experimenten zur Prüfung der Gültigkeit der mTTH mit Silbenmaterial ergab sich, daß die zur Verarbeitung im Langzeitgedächtnisspeicher benötigte Zeit eine Konstante ist, die weder durch Belastung der Zentralen Exekutive noch durch Irrelevante Sprache beeinflußt wird. Es zeigt sich lediglich eine signifikante Verlängerung der Verarbeitungszeit und ein – im Gegensatz zu Parameter a überproportionalen – Zuwachs für Parameter b, wenn keine subvokale Wiederholung möglich ist (Kapitel 3.3). Diese Beobachtung wird als empirische Hypothese auf diese Arbeit übertragen, ohne Berücksichtigung der Variation linguistischer Gruppen und Stimulusmaterialien. Zwischen den realisierten Materialien findet sich kein Unterschied in der Schnelligkeit der Informationsverarbeitung im Langzeitgedächtnisspeicher, der nicht durch Unterschiede der Artikulationszeit erklärt werden kann. Es werden drei Hypothesen geprüft (Kapitel 7.6):

Erstens wird erwartet, daß zwischen den Experimenten, in denen die Möglichkeit zur Artikulation gegeben ist (Gruppe 1), keine Mittelwertsunterschiede bestehen. Wie schon in Kapitel 4.1.2 bei Überprüfung der Gültigkeit der mTTH ist hierbei die Alternativhypothese der Testausgang, bei dem die Hypothese verworfen wird.

$$H_0: \quad b_1' / I = b_3' / I = b_4' / I = b_6' / I = b_7' / I = b_8' / I$$

$$H_1: \quad \neg H_0$$

Die Verarbeitungszeit im Langzeitgedächtnisspeicher pro bit ist abgekürzt durch b' / I und errechnet sich durch den Quotienten des nach der mTTH errechneten Steigungsparameters b und dem insgesamt zu verarbeitendem Informationsgehalt I (Kapitel 3.3), die Indices folgen der Numerierung der Experimente (Kapitel 4.2).

Zweitens werden zwischen den beiden Experimenten mit Artikulatorischer Sekundäraufgabe keine Unterschiede in der Verarbeitungszeit im Langzeitgedächtnisspeicher erwartet (Gruppe 2).

$$H_0: \quad b_2' / I = b_5' / I$$

$$H_1: \quad b_2' / I \neq b_5' / I$$

In der dritten Hypothese wird angenommen, daß zwischen diesen beiden Gruppen (subvokale Wiederholung möglich/nicht) signifikante Abweichungen dahingehend bestehen, daß Artikulatorische Unterdrückung die Verarbeitungszeit im Langzeitgedächtnisspeicher signifikant verlängert.

$$H_0: \quad b_{\text{gruppe1}} / I \leq b_{\text{gruppe2}} / I$$

$$H_1: \quad b_{\text{gruppe1}} / I > b_{\text{gruppe2}} / I$$

Falls die statistischen und empirischen Hypothesen zurückgewiesen werden, muß die Fragestellung modifiziert werden.

4.2 AUFBAU DER STUDIE

Zur Untersuchung dieser Fragestellungen werden insgesamt acht Experimentreihen konzipiert, die jeweils aus drei Untersuchungen pro Proband bestehen, die ausführlich in Kapitel 5.1 geschildert werden. Um die Universalität der drei Invarianzhypothesen möglichst umfassend und streng zu prüfen, werden die folgenden vier Faktoren variiert.

- Anhand des Zahlenmaterials werden drei verschiedene LINGUISTISCHE GRUPPEN untersucht (deutsche, chinesische und spanische Probanden).
- Zusätzlich wird mit deutschen und chinesischen Probanden je ein Experiment mit ARTIKULATORISCHER UNTERDRÜCKUNG durchgeführt.
- Insgesamt werden drei STIMULUSMATERIALIEN realisiert (Zahlensymbole, phonologisch ähnliche und unähnliche Worte).
- Als vierte Variable wird der PRÄSENTATIONSMODUS variiert (Darbietung der Ziffern als Symbol oder als Wort).

Im folgenden werden die Experimentreihen anhand dieser vier Charakteristika definiert. Für jedes Experiment erfolgt eine Zuordnung der untersuchten Hypothesen.

EXPERIMENTREIHE 1 Das Basisexperiment - 1 (Stille, d, ZS) - ist charakterisiert durch Zahlensymbole als Materialart, deutsche Muttersprache der Probanden und durch das Fehlen einer Sekundäraufgabe. Anhand dieser Experimentreihe werden, neben den ZENTRALEN HYPOTHESEN und den EXPLORATIVEN HYPOTHESEN, der *Effekt der linguistischen Gruppe, der Artikulatorischen Unterdrückung, des Geschlechts* sowie des *Präsentationsmodus* untersucht.

EXPERIMENTREIHE 2 Die Konzeption der Experimentreihe 2 (art.U, d, ZS) ist analog aufgebaut zu Experimentreihe 1 mit dem Unterschied, daß während der Gedächtnisexperimente als Sekundäraufgabe Artikulatorische Unterdrückung verlangt wird. Weder die Muttersprache der Probanden (deutsch) noch die Materialart (Zahlen) oder der Präsentationsmodus (Symbol) werden verändert. Untersuchungsziele sind die Überprüfung

der *mTTH*, der *Äquivalenzhypothese* sowie der Explorativen Hypothese zur *verlängerten Verarbeitungszeit im Langzeitgedächtnisspeicher*. Die drei Arbeitshypothesen *Effekt Artikulatorischer Unterdrückung* sowie *Eliminierung des Effekts der linguistischen Gruppe*, respektive des *Geschlechtseffekts durch diese Sekundäraufgabe* beziehen sich ebenfalls auf Experimentreihe 2.

EXPERIMENTREIHE 3 (Stille, s, ZS) unterscheidet sich vom Basisexperiment darin, daß Probanden mit spanischer Muttersprache akquiriert werden. Das Stimulusmaterial besteht aus Zahlensymbolen, eine Sekundäraufgabe ist nicht Bestandteil der Untersuchung. Geprüft werden anhand dieses Experiments die ZENTRALEN und die EXPLORATIVEN HYPOTHESEN sowie die Arbeitshypothese zum *Effekt der linguistischen Gruppe*.

EXPERIMENTREIHE 4 Die Intention der Durchführung von Experimentreihe 4 (Stille, c, ZS)- mit chinesischen Probanden, Zahlensymbolen und dem Verzicht auf eine Sekundäraufgabe entspricht der des Basisexperiments: Es werden sowohl die ZENTRALEN als auch die EXPLORATIVEN HYPOTHESEN geprüft. Ebenso werden die Arbeitshypothesen *Effekt der linguistischen Gruppe*, *der Artikulatorischen Unterdrückung*, *des Präsentationsmodus* und *des Geschlechts* auf diese Experimentreihe bezogen.

EXPERIMENTREIHE 5 Die Bedingung 5 (art.U, c, ZS) ist eine Wiederholung von Experimentreihe 2 mit dem Unterschied, daß chinesische statt deutsche Probanden an der Studie teilnehmen. Die Zahlensymbole werden unter der Bedingung Artikulatorischer Unterdrückung gelernt. Neben der Untersuchung der *mTTH* und der *Äquivalenzhypothese* wird die Explorative Hypothese zur *verlängerten Verarbeitungszeit im Langzeitgedächtnisspeicher* geprüft. Ferner gehen die Daten von Bedingung 5 in die Analysen zu den drei Arbeitshypothesen *Effekt Artikulatorischer Unterdrückung* sowie *Eliminierung des Effekts der linguistischen Gruppe*, respektive des *Geschlechtseffekts durch diese Sekundäraufgabe* ein.

EXPERIMENTREIHE 6 UND 7 Für Bedingung 6 (Stille, d, Pae) wird die Materialart im Vergleich zum Basisexperiment geändert: Statt Zahlen fungieren phonologisch ähnliche Worte als Stimulusmaterial. Experimentreihe 7 (Stille, d, Pu) entspricht der vorigen Reihe, allein der Grad Phonologischer Ähnlichkeit zwischen den Items wird minimiert. Mit der Materialart ist der Präsentationsmodus konfundiert: Die Darbietung der Worte erfolgt in alphabetischer Schrift. An beiden Experimenten nehmen nur deutsche Probanden teil. Anhand von Experimentreihe 6 und 7 werden alle ZENTRALEN sowie EXPLORATIVEN HYPOTHESEN und die Arbeitshypothese *des Materialeffekts* geprüft.

EXPERIMENTREIHE 8 Die letzte Reihe - **8** (Stille, d, ZW) – ist bis auf eine Ausnahme mit dem Basisexperiment identisch: Die Ziffern werden als Zahlworte präsentiert. Auf eine Zusatzaufgabe wird verzichtet. Wiederum werden die ZENTRALEN und die EXPLORATIVEN HYPOTHESEN geprüft. Zusätzlich beziehen sich noch zwei Arbeitshypothesen auf Experimentreihe 8: der *Materialeffekt* sowie der *Effekt des Präsentationsmodus*.

Tabelle 3 gibt einen Überblick über die acht realisierten Bedingungen dieser Studie.

Tabelle 3 Aufbau der Studie

Experiment	art.Unterdr.	linguist.Gruppe	Material	Präsent.art
1 (Stille, d, ZS)	-	d	Z	S
2 (art.U, d, ZS)	+	d	Z	S
3 (Stille, s, ZS)	-	s	Z	S
4 (Stille, c, ZS)	-	c	Z	S
5 (art.U, c, ZS)	+	c	Z	S
6 (Stille, d, Pae)	-	d	Pae	W
7 (Stille, d, Pu)	-	d	Pu	W
8 (Stille, d, ZW)	-	d	Z	W

Die Abkürzung „art.U“, respektive „art. Unterdr.“ steht für „Artikulatorische Unterdrückung“ als Sekundäraufgabe. Mit „linguist. Gruppe“ d-s-c ist die Muttersprache der Probanden gemeint (deutsch, spanisch, chinesisch). Die drei Materialien sind Zahlen (Z), phonologisch ähnliche Worte (Pae) und phonologisch unähnliche Worte (Pu). Mit „Präsent.art“ ist der Darbietungsmodus der Materialien als Symbol (S) oder Wort (W) beschrieben.

Der Aufbau der Experimentreihen sowie Informationen zur Materialkonstruktion und zum Akquirieren der Probanden finden sich im nächsten Kapitel.

5 VERSUCHSPLANUNG

In diesem Kapitel werden die deskriptiven Charakteristika der Versuchsanordnung behandelt. Die für jede der acht beschriebenen Experimentreihen definierte Bedingungskonstellation wird innerhalb der Reihe nicht geändert. Alle acht Experimentreihen bestehen aus folgenden drei Teilen:

- Experiment zur Messung der Artikulationszeit
- Subspannen-Experiment zur Messung der Kurzzeitgedächtnisspanne
- Supraspannen-Experiment zur Messung der Langzeitgedächtnisspanne

Da der Aufbau der Experimentreihen völlig identisch ist, wird in Kapitel 5.1 die Darstellung des experimentellen Designs für alle acht Reihen zusammengefaßt. Die Materialkonstruktion für jede realisierte Stimulusart inklusive der nötigen Vorexperimente wird in Kapitel 5.2 behandelt. In Kapitel 5.3 werden die deskriptiven Charakteristika der Stichproben beschrieben.

5.1 AUFBAU DER EXPERIMENTREIHEN

In diesem Kapitel werden die drei Untersuchungen innerhalb der Experimentreihen besprochen. Wie bereits erwähnt, besteht jede Reihe aus einem Artikulationszeit-, einem Sub- und einem Supraspannen-Experiment.

Begonnen wird bei allen Probanden mit dem Experiment zur Messung der Artikulationszeit. Die beiden folgenden Gedächtnisexperimente werden in ihrer Darbietungsreihenfolge variiert. Fünfzig Prozent der Probanden führen erst das Langzeit- und dann das Kurzzeit-Gedächtnisexperiment durch, bei den anderen kehrt sich der Ablauf der Gedächtnisexperimente um. Die beiden Reihenfolgen werden a priori systematisch ausbalanciert, um Ermüdungs-, respektive Übungseffekte kontrollieren zu können. Die Zuordnung der Versuchspersonen zu der Gruppe erfolgt nach dem Zufallsprinzip. In Kapitel 6.1.2 wird eine statistische Überprüfung der Durchführungsobjektivität in bezug auf die Darbietungsreihenfolge der Gedächtnistests vorgenommen.

Der Versuchsaufbau ist durch folgende Merkmale charakterisiert:

- Die Dauer einer Untersuchungssitzung pro Experimentreihe beträgt insgesamt 35 bis 40 Minuten.

- Die Durchführung erfolgt in Einzelsitzungen.
- Sowohl Instruktion, Itempräsentation als auch die Auswertungen der Leistungen und Latenzzeiten erfolgt computergestützt.
- Die Beurteilung der Korrektheit der Antworten (in den Gedächtnistests) und das Signal zum Stoppen der Zeituhr (im Artikulationszeit- und Subspannen-Experiment) werden durch den Versuchsleiter eingegeben.
- Vor Präsentation der Items wird ein ACHTUNG-Signal auf dem ansonsten schwarzen Bildschirm in weißer Schrift in der Mitte des Bildschirms dargeboten.
- Die Präsentation der Items erfolgt zentriert in schwarzer Schrift in einem weißen Rahmen (in der Mitte des Bildschirms, Höhe: 2.4 cm), der sich über den ganzen Bildschirm zieht. Ober- und unterhalb dieses Rahmens ist der Bildschirm schwarz.
- Die Items werden mit einer Höhe von 0.60 Zentimetern dargeboten und sind gut lesbar. Zwischen den Items wird ein Abstand von 0.50 Zentimetern gehalten.
- Bei den Gedächtnistests wird die Aufforderung zur Reproduktion innerhalb des weißen Rahmens in schwarzer Schrift mit rechtsseitiger Ausrichtung gezeigt.
- Zwischen den einzelnen Durchgängen zum Vorlesen, respektive der Lernversuche in den Gedächtnistests, ist der Bildschirm schwarz.
- Sowohl das Wortmaterial als auch die Signale ACHTUNG und WIEDERGABE werden in Großbuchstaben präsentiert.

Als Versuchsausrüstung dienen ein Compaq PC 486 (25 MHz, 4 MB RAM, MS-DOS 6.22) mit einem schwarz-weiß LCD-Display ohne Nachflimmern und ein Olivetti-Laptop (Pentium 100, 8 MB RAM Windows 3.1). Zur Erhebung der Probanden in Spanien wird ein Compaq-Laptop (Pentium 100, 4 MB RAM, Windows 1998) verwendet.

Die Äquivalenz zwischen den Experimenten mit deutschen, spanischen und chinesischen Probanden soll durch drei Punkte gewährleistet werden: Erstens wird die Instruktion in der Muttersprache des Probanden präsentiert. Zweitens wird darauf geachtet, daß der Proband und der Versuchsleiter dieselbe Muttersprache sprechen und drittens findet die Erhebung im Herkunftsland der Probanden statt (mit Ausnahme von 14 spanisch sprechenden Teilnehmern, die in Deutschland leben). Diese Maßnahmen sollen gewährleisten, daß jeder Proband die Instruktionen gut versteht und eine Möglichkeit zur Rückfrage hat. Zudem werden Verständnisschwierigkeiten des Versuchsleiters in der Beurteilung der Korrektheit der Leistung des Probanden minimiert. Die Erhebung der Daten im Herkunftsland der Probanden erleichtert nicht nur die praktische Durchführung, sondern wirkt auch einem möglichen Selektionseffekt entgegen. Da Auslandsaufenthalte meist durch Stipendien

vergeben werden, könnte eventuell die Vergleichbarkeit der Stichproben eingeschränkt sein, wenn im Spanischen und Chinesischen nur besonders ausgewählte Probanden in die Stichprobe eingehen.

Nachteil der Parallelisierung der Muttersprache von Versuchsleiter und Proband innerhalb der Experimentreihen ist, daß mehr als ein Versuchsleiter zur Erhebung der Daten benötigt wird. Um Varianz durch verschiedene Personen zu reduzieren, nimmt erstens jeder Versuchsleiter als Proband einmal an den Experimenten teil. Zweitens erfolgt eine schriftliche Information über den Ablauf der Experimente sowie die Aufgaben des Versuchsleiters und drittens absolviert jeder Versuchsleiter einen ersten Durchlauf an einer Versuchsperson, deren Daten nicht ausgewertet werden. In den Experimenten, in denen mehrere Personen als Versuchsleiter fungieren, wird auf ein über die Darbietungszeiten im Supraspannen-Experiment balanciertes Erheben geachtet, um systematische Verzerrungen zu vermeiden.

In den Kapiteln 5.1.1 bis 5.1.3 werden die einzelnen Teile einer Experimentreihe vorgestellt, wie sie in den Experimenten ohne Sekundäraufgabe realisiert werden (Experiment 1, 3, 4, 6, 7 & 8). In den Kapitel 5.1.4 und 5.1.5 findet sich eine Darstellung der Gedächtnisexperimente unter der Bedingung Artikulatorischer Unterdrückung (Experiment 2 & 5).

5.1.1 EXPERIMENT ZUR MESSUNG DER ARTIKULATIONSZEIT

In dem Experiment zur Messung der Artikulationszeit werden den Probanden Stimulussequenzen variabler Länge vorgegeben. Die Aufgabe besteht in lautem Vorlesen dieser Reihen.

Da dieses Experiment bei allen Probanden die Erhebung eröffnet, wird der Instruktion eine allgemeine Einführung vorangestellt (Anhang I). Diese Einführung informiert den Probanden über den Ablauf der Studie, über das verwendete Stimulusmaterial und über die vertrauliche Behandlung der Daten. Zudem appelliert sie an die Bereitschaft zur konzentrierten Teilnahme und zum sorgfältigen Lesen der Instruktionen.

Nach dieser Einführung wird die Instruktion für das Artikulationszeit-Experiment präsentiert (Anhang I). Dem Probanden werden ausführlich die Aufgabe und der Experimentablauf beschrieben. Es folgt ein Beispiel und ein mündliches Feedback durch

den Versuchsleiter, ob die Aufgabe korrekt durchgeführt wurde. Dann wird schriftlich auf folgende vier Punkte hingewiesen, die beim Lesen berücksichtigt werden sollen:

1. Auskommen mit einer Atemlänge
2. Vermeiden von Versprecher
3. Deutliche Aussprache
4. Lesen so schnell als möglich.

Der erste und der letzte Punkt werden eingeführt, um die Versuchsperson unter Zeitdruck zu setzen. Da mit erhöhtem Zeitdruck ein häufigeres Auftreten von Versprechern und ein undeutlicheres Sprechen konfundiert ist, wird durch die Punkte zwei und drei versucht, diese unerwünschten Nebeneffekte zu minimieren. Falls dennoch ein Versprecher auftritt, geht die Reihe nicht in die Auswertung ein, sondern wird mit anderen Stimuli gleicher Sequenzlänge wiederholt. Wenn der Proband die Instruktion gelesen hat und keine Fragen offen sind, wird das Experiment gestartet.

Zuerst ist der Bildschirm schwarz, dann erscheint in der Mitte mit weißer Schrift das Signal „ACHTUNG“. Während das ACHTUNG-Signal als visueller Reiz auf dem Bildschirm verbleibt, wird nach einer Pause von 800 Millisekunden ein Ton 200 Millisekunden lang präsentiert (400 Hz). Diese Kombination zwischen Pause und Ton wird dreimal wiederholt. Nach dem dritten Ton tritt eine Pause von einer Sekunde ein und das ACHTUNG-Signal verschwindet. Die Präsentation der Items auf dem Bildschirm folgt. Wie bereits beschrieben, werden die Items zentriert in einem weißen Balken in schwarzer Schrift gezeigt (Kapitel 5.1). Die Items werden so lange dargeboten, bis die Versuchsperson die gesamte Sequenz vorgelesen hat. Nach Beendigung des Lesens drückt der Versuchsleiter eine beliebige Taste und stoppt damit die computer-interne Uhr der Zeitmessung. Mit diesem Tastendruck verschwindet die Reihe. Auf dem schwarzen Bildschirm erscheint die Frage „Wiedergabe korrekt (j/n)?“. Diese Frage wird von dem Versuchsleiter mit Tastendruck beantwortet. Danach wird der Bildschirm durchgehend schwarz und nach einer Pause von fünf Sekunden erscheint in weißer Schrift das ACHTUNG-Signal. Der nächste Durchgang beginnt. Nach dem Durchlauf aller 18 Reihen ist der Versuch beendet.

Abbildung 10 zeigt den zeitlichen Verlauf des Experiments zur Messung der Artikulationszeit.

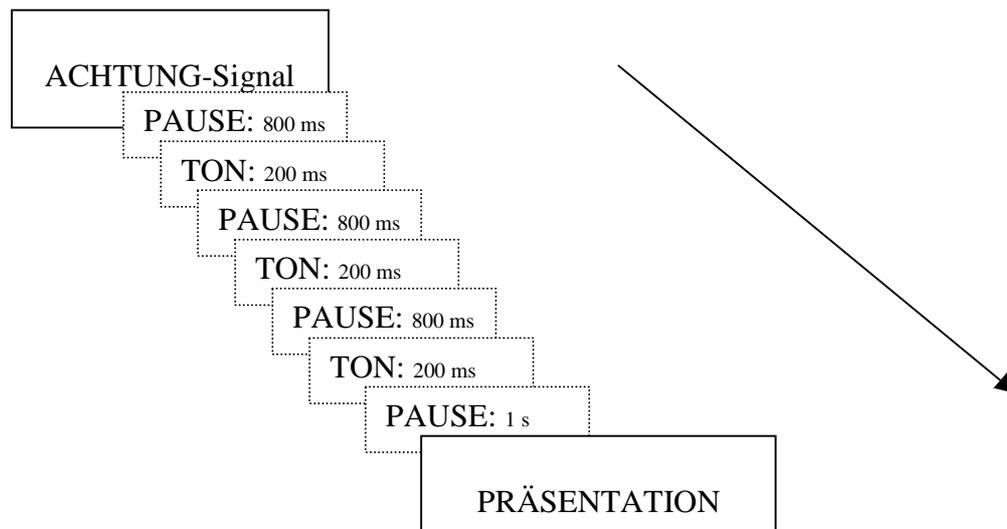


Abbildung 10 Zeitlicher Verlauf des Experiments zur Messung der Artikulationszeit

Insgesamt werden 18 Reihen präsentiert. Die ersten drei Sequenzreihen sind Übungsreihen (dies wird dem Probanden nicht mitgeteilt), die nicht protokolliert werden. Dadurch wird dem möglichen Phänomen Rechnung getragen, daß bei den ersten Items wegen der Neuheit der Experimentalsituation die Latenzzeit bis zum Beginn des Vorlesens verlängert sein kann. Allein die zur Artikulation benötigten Zeiten der verbleibenden 15 Sequenzen gehen in die Auswertung ein.

Präsentiert werden nach den Übungsreihen variabler Länge je fünf Reihen mit sechs, acht und zehn Items in zufälliger Reihenfolge. Dieses Vorgehen hat sich bereits bewährt (Bredenkamp & Klein, 1998). Die Konstruktion der Sequenzen erfolgt durch randomisiertes Ziehen von Items aus dem Itempool mit Zurücklegen. Als Restriktion wird eine Vermeidung von dreifachen Itemwiederholungen eingeführt. Bei Zahlenmaterial wird zusätzlich die Entstehung stereotyper Dreier-Reihen (beispielsweise 4, 5, 6) unterbunden.

Nach 15 fehlerfrei vorgelesenen Sequenzen wird das Experiment zur Artikulationszeitmessung beendet. In fünfzig Prozent der Fälle wird im Anschluß das Subspannen-Experiment (Kapitel 5.1.2) dargeboten, in den anderen fünfzig Prozent das Supraspannen-Experiment (Kapitel 5.1.3).

5.1.2 DAS SUBSPANNEN-EXPERIMENT

Aufgabe in dem Subspannen-Experiment ist es, sich verschiedene Sequenzen variabler Länge einzuprägen und nach einer Zeitverzögerung von zwei Sekunden wiederzugeben.

In der Instruktion zu diesem Experiment (Anhang I) werden die Versuchspersonen mit ihrer Aufgabe und dem Verlauf des Experiments vertraut gemacht. Zur Veranschaulichung wird an einem Beispiel die Durchführung geübt, wobei der Versuchsleiter ein mündliches Feedback gibt. Falls der Proband die Items während der visuellen Präsentation laut vorliest, wird er gebeten, dies zu unterlassen. Begründet wird diese Restriktion mit dem Befund von Frankish (1985), daß bei lautem Vorlesen während der Präsentationsphase die Gedächtnisleistung gegenüber einer stillen Bedingung signifikant erhöht ist. Wenn der Proband nach Durchführung des Übungsbeispiels keine Fragen hat, wird das Experiment gestartet.

Auf dem schwarzen Bildschirm erscheint in weißer Schrift zentriert das ACHTUNG-Signal. Dieses Signal bleibt 2.5 Sekunden auf dem Bildschirm sichtbar. Nach Verschwinden des Signals ist der Bildschirm eine Sekunde lang schwarz, dann erscheint der erste Durchlauf des Experiments. Die Items werden, wie in Kapitel 5.1 bereits beschrieben, zentriert in schwarzer Schrift innerhalb eines weißen Balkens präsentiert. Die Darbietung erfolgt sukzessiv. Jedes Item ist für jeweils 0.80 Sekunden zu sehen, das Interstimulusintervall beträgt 0.20 Sekunden. Der weiße Balken bleibt auch während der Interstimulusintervalle sichtbar. Lokal werden die Items jeweils um 0.5 Zentimeter versetzt gezeigt. Nach Präsentation des letzten Items folgt eine Pause von zwei Sekunden bis zum Erscheinen des WIEDERGABE-Signals am rechten Rand des weißen Balkens. Der Modus einer verzögerten Wiedergabe hat sich in anderen empirischen Untersuchungen bewährt (Baddeley et al., 1984). Sobald das Signal zur Wiedergabe erscheint, soll der Proband die erinnerten Items aussprechen. Der weiße Balken und das WIEDERGABE-Signal sind bis zur Beendigung der Wiedergabe durch die Versuchsperson auf dem Bildschirm zu sehen. Nach Beenden der Reproduktion drückt der Versuchsleiter eine beliebige Taste, wenn die Antwort falsch ist, respektive die Taste „j“ wenn die Antwort korrekt ist. Durch den Tastendruck wird einerseits die computer-interne Zeitmessung der Reproduktionszeit gestoppt und andererseits der neue Durchgang gestartet. Nach dem Tastendruck erscheint wiederum der schwarze Bildschirm und nach fünf Sekunden das weiße ACHTUNG-Signal.

Als Abbruchkriterium fungiert der zehnte Wechsel zwischen korrekter und fehlerhafter Reproduktion. Auf ein zweites Kriterium (Bredenkamp & Klein, 1998) zur Einschränkung des Konfidenzintervalls des Fehlers wird verzichtet, da andere Experimente zeigen, daß die Genauigkeit der Messung durch Verzicht auf dieses Kriterium nicht wesentlich beeinflußt wird (Bredenkamp & Hamm, 1999). In Abbildung 11 wird der zeitliche Verlauf graphisch dargestellt.

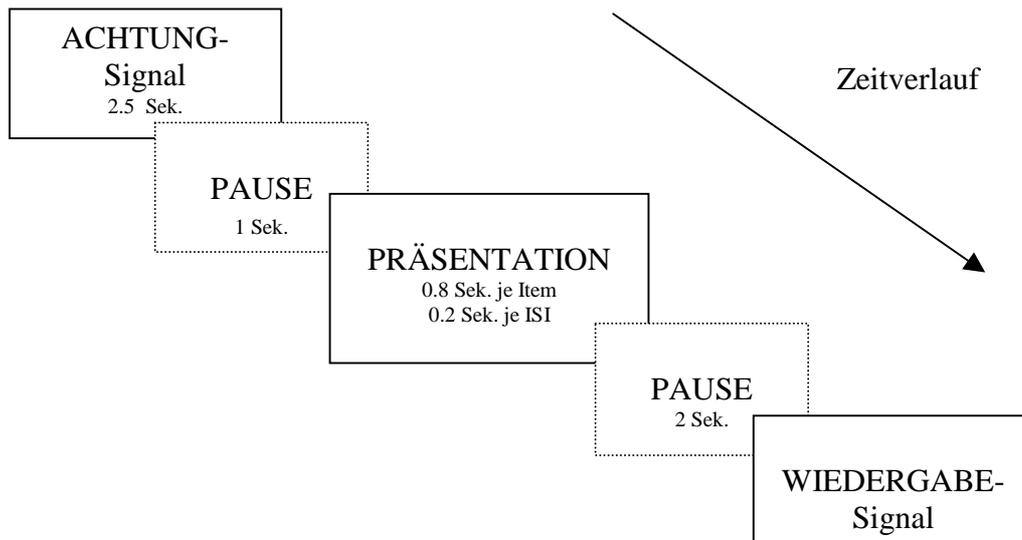


Abbildung 11 Zeitlicher Verlauf des Subspannen-Experiments

Nach der Logik des up-and-down-Verfahrens von Ericsson, Chase und Faloon (1980) wird dieser adaptive Gedächtnistest in Analogie zu dem Verfahren von Bredenkamp und Klein (1998) durchgeführt. Im ersten Versuch werden drei Items vorgegeben. Bei korrekter Wiedergabe der Itemsequenz in der richtigen Reihenfolge wird im nächsten Durchgang die Reihe um ein Item vergrößert, bei fehlerhafter Reproduktion wird ein Item weniger vorgegeben. Das bedeutet, daß sich die Länge der Sequenz nicht in zwei aufeinanderfolgenden Durchgängen wiederholen kann. Die Richtigkeit der Antwort wird vom Versuchsleiter beurteilt.

Die Konstruktion der Sequenzen ist analog zu dem in Kapitel 5.1.1 beschriebenen Vorgehen konzipiert: Aus dem Itempool werden in randomisierter Reihenfolge Items „mit Zurücklegen“ gezogen. Dreifache Wiederholungen eines Items sind ebensowenig zugelassen wie auf- oder absteigende Tripel bei Zahlenmaterial. Zusätzlich zur Ermittlung der Gedächtnisspanne, die definiert ist als der Mittelwert der lokalen Maxima und Minima, wird die Reproduktionszeit erfaßt. Diese wird gemessen vom Erscheinen des WIEDERGABE – Signals bis zum Tastendruck des Versuchsleiters nach Beendigung der Reproduktion durch den Probanden. Für fünfzig Prozent der Probanden schließt an dieses Subspannen-Experiment das Supraspannen-Experiment an, für die restlichen Versuchspersonen bildet diese Untersuchung den Abschluß der Sitzung.

5.1.3 DAS SURPASPANNEN-EXPERIMENT

Aufgabe im Supraspannen-Experiment ist es, sich eine Reihe von Items innerhalb von 10, 20 oder 30 Sekunden einzuprägen und diese unmittelbar nach der Präsentation

wiederzugeben. Die Reihe wird unverändert so oft hintereinander präsentiert bis sie der Proband zweimal hintereinander korrekt reproduziert hat.

Zum Subspannen-Experiment bestehen fünf fundamentale Unterschiede:

1. Die Sequenz ändert sich interindividuell, nicht aber intraindividuell: Es gibt pro Person nur eine einzige Reihe, die weder bezüglich der Stimuli noch bezüglich der Länge variiert.
2. Die Darbietung der Items erfolgt nicht sukzessiv, sondern simultan.
3. Gemessen wird nicht die Anzahl gemerkter Items, sondern die Anzahl der Lernversuche, bis die fixe Anzahl von Items korrekt wiedergegeben wird.
4. Es erfolgt eine unmittelbare Reproduktion ohne Zeitverzögerung.
5. Abbruchkriterium ist die zweimalige korrekte Wiedergabe der Itemreihe.

Die Darbietungszeit variiert interindividuell zwischen 10, 20 und 30 Sekunden, entsprechend dem Versuchsaufbau von Bredenkamp und Klein (1998). Während dieser Zeit wird das Lernmaterial en bloc dargeboten. Jede Versuchsperson wird zufällig einer Darbietungszeit-Bedingung zugeordnet.

Vor Beginn des Versuchs wird der Proband in der Instruktion (Anhang I) mit der Aufgabe und dem Ablauf vertraut gemacht. Nach einem Übungsbeispiel besteht für den Probanden die Möglichkeit zur Rückfrage. Dann wird der Versuch gestartet.

Auf dem schwarzen Bildschirm wird zentriert in weißer Schrift das Signal „ACHTUNG“ gezeigt. Nach 1.5 Sekunden verschwindet das Signal, der Bildschirm bleibt für eine halbe Sekunden schwarz. Dann erscheint en bloc die Itemreihe, je nach Bedingung für 10, 20 oder 30 Sekunden. Die Items werden, wie in Kapitel 5.1 geschildert, in schwarzer Schrift zentriert in einem weißen Balken präsentiert. Nach Ablauf der Darbietungszeit verschwinden alle Items und gleichzeitig erscheint das WIEDERGABE-Signal auf der rechten Seite des weißen Balkens. Dies bleibt solange sichtbar, bis der Proband die Reproduktion beendet und der Versuchsleiter per Tastendruck beurteilt hat, ob die Wiedergabe korrekt ist oder nicht. Dann wird der Bildschirm schwarz und nach fünf Sekunden erscheint das weiße ACHTUNG-Signal. Der Versuch ist beendet, wenn der Proband zweimal hintereinander die Reihe korrekt wiedergegeben hat. Eine zeitliche Darstellung des Experiments findet sich in Abbildung 12.

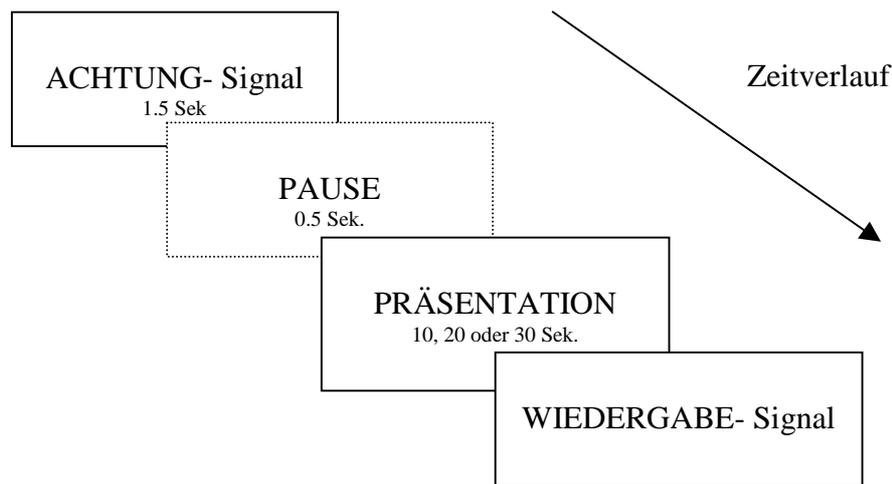


Abbildung 12 Zeitlicher Verlauf des Supraspannen-Experiments

Im Supraspannen-Experiment finden sich Unterschiede zwischen den Experimentreihen: In allen Experimenten mit Zahlensymbolen werden 15 Items dargeboten (Experiment 1, 2, 3, 4 & 5). In den anderen drei Experimenten dagegen werden nur 10 Items präsentiert (Experiment 6, 7 & 8). Dies entspricht dem optimalen Umfang, der weder Decken- noch Bodeneffekte evoziert, wie in vier Vorexperimenten an je zehn Versuchspersonen nachgewiesen wurde.

In *Vorexperiment 1* zeigt sich, daß bei zehn Zahlensymbolen und zehn Sekunden Darbietungszeit vier von zehn (deutschsprachigen) Probanden auf Anhieb die Itemreihe richtig wiedergaben und drei weitere Personen nur einen Fehlversuch benötigten. Es handelt sich um einen Deckeneffekt, da mit einer Verlängerung der Darbietungszeit keine Leistungsverbesserung eintreten kann. Aus diesem Grund wird der Umfang auf 15 Items festgesetzt, der sich als adäquat erweist (Kapitel 6.4). In den drei anderen Vorexperimenten mit Wortmaterial und einem Umfang von 15 Items bei zehn Sekunden Präsentationsdauer zeigt sich ein gegenteiliger Effekt: Von zehn Probanden unter der Bedingung klangähnlicher Substantive brachen drei Personen das Experiment ab (*Vorexperiment 2*). Bei phonologisch unähnlichen Worten konnten alle zehn Probanden die Aufgabe bewältigen, benötigten im Schnitt aber 17 Durchgänge bis zur korrekten Wiedergabe (*Vorexperiment 3*). In *Vorexperiment 4* mit Zahlworten zeigt sich der Befund, daß die Probanden zum Erlernen von 15 Zahlworten im Durchschnitt 15 Versuche zum Erreichen des Kriteriums benötigen. Zwei Versuchspersonen brachen das Experiment vorzeitig ab. Dieses Ergebnis ist in Zusammenhang mit *Vorexperiment 1* überraschend. Um die Schwierigkeit der Aufgabe zu reduzieren, wird in allen Experimenten mit Wortmaterialien die Sequenzlänge auf zehn Items fixiert.

Die individuellen Sequenzen sind so aufgebaut, daß an jeder Position jedes Item dieselbe Auftretenswahrscheinlichkeit hat mit der Beschränkung, daß ein Item nicht mehr als drei Mal wiederholt werden darf. Bei Zahlenmaterial werden zusätzlich auf-, respektive absteigende Dreier-Sequenzen ausgeschlossen.

5.1.4 DAS SUBSPANNEN-EXPERIMENT MIT ARTIKULATORISCHER UNTERDRÜCKUNG

Sowohl bei chinesischen als auch bei deutschen Probanden wird je eine Experimentreihe mit Zahlensymbolen als Stimulusmaterial und simultaner Artikulatorischer Unterdrückung entworfen. Das Experiment zur Messung der Artikulationszeit ist nicht von einer Änderung im Vergleich zu den Stille-Bedingungen betroffen.

Die Einführung in die Untersuchung ändert sich im Vergleich zu den Experimenten ohne Zusatzfähigkeit wenig: Den Probanden wird derselbe Einführungstext und dieselbe Instruktion zum Artikulationszeitexperiment vorgegeben. Es wird lediglich zwischen diese beiden Teile zusätzlich eine Instruktion zur Artikulatorischen Unterdrückung eingefügt (Anhang I). Dem Probanden wird erklärt, daß während der Gedächtnisexperimente im Gleichklang mit einem – durch den Computer präsentierten - Ton die Silbe TAU ausgesprochen werden soll. Der Ton wird im Takt von einer halben Sekunden präsentiert. Anhand eines Beispiels wird das Sprechen im Gleichklang mit dem Ton geübt. Die Silbe TAU hat sich bereits als zu artikulierendes Material bewährt (Bredenkamp & Hamm, 1999). TAU bietet sich auch für diese Untersuchungen an, da sie für Deutsche und Chinesen leicht zu artikulieren ist. Ferner ist TAU unabhängig von der Sprache der Probanden phonologisch unähnlich und ohne semantischen Bezug zum Lernmaterial.

Die ursprüngliche Instruktion zum Subspannen-Experiment wird um die Aufforderung zur Artikulation von TAU erweitert (Anhang I). An Aufgabe und Aufbau wird keine Veränderung im Vergleich zum Vorgehen in Stille vorgenommen. Um Wiederholungen zu vermeiden, sei auf Kapitel 5.1.2 hingewiesen. Die einzige Abwandlung betrifft das Einfügen eines Computertones (400 Hz, 100 ms) alle halbe Sekunde von Beginn des ACHTUNG-Signals über die Präsentationsphase der zu lernenden Sequenz bis zum Erscheinen des WIEDERGABE-Signals. Die Ausführung der Artikulation wird vom Versuchsleiter überwacht. Da relativ leise Artikulatorische Unterdrückung eventuell andere Effekte evoziert als die laute Aussprache der irrelevanten Silbe (Dilger, 2000), wird der Proband durch den Versuchsleiter dazu angehalten, laut und deutlich zu sprechen.

5.1.5 DAS SUPRASPANNEN-EXPERIMENT MIT ARTIKULATORISCHER UNTERDRÜCKUNG

Das Supraspannen-Experiment mit Sekundäraufgabe unterscheidet sich von dem unter Stillebedingung durch die Ergänzung der Instruktion um den Hinweis der Artikulatorischen Unterdrückung und durch das Ertönen eines Computertons im Takt einer halben Sekunde.

Die Instruktion enthält die Umschreibung der Aufgabe, Hinweise zum Aufbau der Untersuchung und die Bitte um Artikulation von TAU im Takt mit dem Computerton sowie ein Übungsbeispiel (Anhang I). Da sich im Aufbau des Supraspannen-Experiments nichts verändert, sei auf Kapitel 5.1.3 zur Beschreibung des Ablaufs verwiesen. Zusätzlich wird alle halbe Sekunde ein Computerton (400 Hz, 100 ms) eingeführt, der vor der Präsentationsphase mit Erscheinen des ACHTUNG-Signals startet und mit dem WIEDERGABE-Signal stoppt. Wie bereits in Kapitel 5.1.4 beschrieben, überwacht der Versuchsleiter die instruktionsgemäße Ausführung der Artikulatorischen Unterdrückung und die laute Aussprache.

5.2 MATERIALKONSTRUKTION

Das Stimulusmaterial dieser Studie besteht ausschließlich aus lexikalischen Items, nämlich aus Zahlen und Worten. Die Struktur dieses Kapitels folgt nicht einer Abhandlung der acht Experimentreihen, sondern der Materialart. So wird für alle sechs Experimente mit Zahlenmaterial das Vorgehen der Stimulusauswahl in Kapitel 5.2.1 dargestellt. Wie bereits in Kapitel 4 erwähnt, werden zwei Wortmaterialien verwendet, die sich im Grad der phonologischen Ähnlichkeit der Items unterscheiden. Da diese beiden Wortreihen parallelisiert werden, bietet sich eine zusammenfassende Darstellung innerhalb eines Kapitels an (Kapitel 5.2.2). Im folgenden werden für jede Materialart Präliminarien und Vorexperimente geschildert.

5.2.1 ZAHLENMATERIAL

Das Zahlenmaterial wird in sechs Experimenten (Experiment 1-5 & 8) verwendet, die sich hinsichtlich *Präsentationsmodus* (Zahlensymbol versus Zahlenworte), *Population* (deutsche, spanische und chinesische Probanden) und *Sekundäraufgabe* (ohne, respektive mit Artikulatorischer Unterdrückung) unterscheiden (Kapitel 4.2). Bei Zahlenmaterial ist die Wahrscheinlichkeit eines „cultural bias“ minimiert:

- Man kann davon ausgehen, daß die individuellen Erfahrungen im Umgang mit Zahlen unabhängig von der Kultur annähernd gleich sind.
- Zusätzlich bietet Zahlenmaterial den Vorteil, daß die Präsentation von Zahlensymbolen interkulturell nicht variiert, so daß keine Konfundierung mit Schrift oder graphischer Komplexität vorliegt.
- Zahlen sind ein Material, für das der Informationsgehalt unabhängig von der Sprache bestimmt werden kann.
- Der vierte Vorteil liegt darin, daß Ziffern ein gängiges Versuchsmaterial sind, das in vielen Studien verwirklicht und validiert wurde (Kapitel 2).

Der Umfang des Zahlenmaterials wird mit acht Items über alle Experimentreihen hinweg konstant gehalten. Dies entspricht einem Informationsgehalt von 3 bit pro Item. Die Versuchspersonen werden über die Extension vor Beginn des Experiments in der Instruktion informiert.

Um die Vergleichbarkeit zwischen den Sprachgruppen zu gewährleisten, wird die phonologische Distinktheit zwischen den Items maximiert. Aus diesem Grund wird bei allen Experimenten mit deutschen Versuchspersonen auf die Zahl DREI wegen ihrer phonologischen Ähnlichkeit zur ZWEI verzichtet. Um dennoch den Umfang bei acht Items zu halten, wird die Ziffer DREI durch die Ziffer NULL ersetzt. Diese Substitution wird für die spanischen und chinesischen Probanden nicht durchgeführt, da dort alle Benennungen der Zahlen eins bis neun phonologisch unähnlich sind.

Die Zahlenworte zwischen null und neun sind im Deutschen einsilbig, mit Ausnahme der Zahl SIEBEN. Im Chinesischen sind alle Benennungen einsilbig, im Spanischen gibt es mehr zwei- als einsilbige Benennungen. Mit Verweis auf die Arbeit von Chincotta und Underwood (1997a) wird zur Maximierung des Spracheffekts das Material homogenisiert: Die Zahl SIEBEN wird als einzige zweisilbige Zahlenbenennung im Deutschen aus allen Experimenten ausgeschlossen. Grund dieses Vorgehens ist der sogenannte Silbeneffekt (Service, 1998): Je mehr Silben in einem Wort vorkommen, desto weniger Worte werden – unabhängig von der Artikulationszeit – gemerkt. Dies wird unter anderem mit Planungsprozessen beim Sprechen in Zusammenhang gebracht (Caplan et al. 1992). Diese Konfundierung von Sprache und Silbenzahl soll für den Vergleich zwischen den chinesischen und den deutschen Probanden durch den Verzicht auf die Zahl SIEBEN eliminiert werden.

In Tabelle 4 wird das verwendete Stimulusmaterial für jede Sprache als Symbol und als Wort aufgeführt. Für die chinesischen Zahlen wird die offizielle Umschrift Hànyu pinyin

verwendet. Hier deutet sich ein Unterschied zwischen den Sprachen an, der nicht kontrolliert werden kann, aber in der Interpretation der Ergebnisse (Kapitel 8) Beachtung finden muß: Die Wortkomplexität (Kapitel 2) ist bei chinesischen Zahlenbenennungen geringer als im Deutschen oder Spanischen. Im Chinesischen bestehen die Worte aus Konsonant – Vokal - Konsonant, respektive Konsonant - Vokal – Verbindungen. In den beiden anderen Sprachen kommen auch Konsonantencluster vor, deren Aussprache schwieriger ist (Service, 1998), beispielsweise die Artikulation der „Zwei“ im Deutschen.

Tabelle 4: Zahlenmaterial pro Sprachgruppe

Deutsche Probanden	Chinesische Probanden	Spanische Probanden
0 : null	1 : yi	1 : uno
1 : eins	2 : èr	2 : dos
2 : zwei	3 : san	3 : tres
4 : vier	4 : sì	4 : cuatro
5 : fünf	5 : wu	5 : cinco
6 : sechs	6 : liù	6 : seis
8 : acht	8 : ba	8 : ocho
9 : neun	9 : jiu	9 : nueve

5.2.2 WORTMATERIAL

Der Umfang der beiden verwendeten Wortmaterialien, die sich bezüglich des Grads phonologischer Ähnlichkeit zwischen den Items unterscheiden, umfaßt je zehn Stimuli. Dies entspricht einem Informationsgehalt von 3.24 bit pro Item. Entgegen der chronologischen Reihenfolge werden zuerst die ausgewählten Stimuli in Tabelle 5 präsentiert, dann folgt eine ausführliche Schilderung der Materialkonstruktion.

Tabelle 5: Stimuli der beiden Wortlisten für Experiment 6 und 7

Phonologisch ähnliche Worte	Phonologisch unähnliche Worte
GOLD	BANK
GELD	ZIEL
HELD	LUFT
ZELT	BAUM
FELD	CHEF
WELT	PLAN
WALD	WERT
WILD	KIND
BILD	FORM
KILT	BUCH

Um eine Vergleichbarkeit der Stimuli zu gewährleisten, werden die beiden Wortlisten in sieben Merkmalen parallelisiert, die sich in der Literatur als relevant für Untersuchungen zum Kurz-, respektive Langzeitgedächtnis herausgestellt haben: *Ort der Betonung*, *Aufbau der Worte*, *graphische Komplexität*, *Wortfrequenz*, *Bedeutungsgehalt*, *Konkretheit* und *Bildhaftigkeit*. Zusätzlich wird für jedes Wortpaar der Wortliste die *phonologische* und die *semantische Ähnlichkeit* ermittelt. Die Parallelisierung der Wortmaterialien wird im folgenden geschildert.

Da bisher das Konstrukt "*Wortkomplexität*" noch nicht präzise definiert ist, aber Studien zum Einfluß dieser Variablen auf Leistungen im Kurzzeitgedächtnis vorliegen (Service, 1998), werden aus der Literatur zwei Kriterien der Wortkomplexität übernommen: erstens Ort der Betonung und zweitens Aufbau der Worte. Der Faktor *Ort der Betonung* hängt eng mit der Silbenzahl der Worte zusammen (Lovatt et al., 2000), während unter dem Faktor *Aufbau der Worte* beispielsweise der Beginn oder das Ende der Worte mit Konsonant oder Vokal verstanden wird (Service, 1998). In diese Studie werden nur einsilbige Worte aufgenommen, die mit einem Konsonanten anfangen und enden. Bei einsilbigen Worten mit einem Konsonant am Anfang und einem Konsonanten am Ende kann sich der Ort der Betonung nicht unterscheiden. Zudem gibt der identische Aufbau der Worte (Konsonant – Vokal – Konsonant mit einer Verdopplung pro Wort) Grund zu der Vermutung, daß sich die Schwierigkeit der Aussprache zwischen den Worten nicht gravierend unterscheidet. Einschränkend muß angemerkt werden, daß die Kontrolle der *Wortkomplexität* nur annähernd gewährleistet ist. Nicht beachtet wird, daß die Schwierigkeit der Aussprache eines Wortes auch von der Artikulation des vorangegangenen Items abhängig sein kann.

Die Relevanz zur Kontrolle *graphischer Komplexität* leitet sich aus einer Untersuchung von Zhang und Feng (1990) ab, die diese Variable als kritische Determinante der Gedächtnisspanne identifizieren. Umgesetzt wird die Kontrolle dieses physikalischen Merkmals durch die Beschränkung auf Worte, die aus vier Buchstaben bestehen. Dadurch wird vermieden, daß einsilbige Worte mit drei, respektive fünf Buchstaben optisch des homogene Bild stören und so durch Aufmerksamkeitslenkung Störeinflüsse erzeugen.

Die vierte kontrollierte Variable ist die *Wortfrequenz*, deren Berücksichtigung seit der Publikation von Hulme et alii (1997) unverzichtbar ist. Für jedes Wort wird aus der CELEX-Datenbank (Baayen, Piepenbrock & Gulikers, 1995) die Wortfrequenz ermittelt. In Kongruenz mit dem konventionellen Vorgehen zum Erstellen von parallelisierten Itemlisten werden die Worte pro Sequenz so gewählt, daß die Abweichung der Mittelwerte möglichst gering gehalten wird. Deskriptiv ergibt sich für phonologisch ähnliche Worte ein Mittelwert

von 176.90, bei phonologisch unähnlichen Worten beträgt er 177.00 Nennungen pro eine Million Worte. Die Frequenzen der einzelnen Worte sind im Anhang II dargestellt.

Die Parallelisierung der beiden Wortlisten in den drei Faktoren *Bedeutungsgehalt*, *Konkretheit* und *Bildhaftigkeit* wird durch Vorexperimente sichergestellt.

Vorexperiment 5: Bedeutungsgehalt von Substantiven

Als Bedeutungsgehalt ist definiert, in welchem Ausmaß die Substantive Assoziationen mit anderen Substantiven hervorrufen. Kontrolliert wird diese Variable, weil in der Studie von Lovatt et alii (2000) auf die Relevanz semantischer Assoziationen hingewiesen wird.

Zur Erfassung des *Bedeutungsgehalts* von Substantiven wird ein Fragebogen von Baschek, Bredenkamp, Ohrle und Wippich (1977) eingesetzt. Eine Wiederholung dieser Fragebogenuntersuchung ist nötig, weil in den von Baschek et alii untersuchten Worten zu wenig Items enthalten sind, die den oben genannten Kriterien bezüglich Anzahl der Silben, Phonemabfolge sowie Anzahl der Buchstaben entsprechen. Die Instruktion wird aus der zitierten Studie übernommen. Als Untersuchungsmaterial werden 57 Worte eingesetzt, die aus vier Buchstaben bestehen, einsilbig sind und mit einem Konsonanten beginnen und enden. Die Beurteilung auf einer siebenstufigen Skala wird von zwanzig Studenten vorgenommen.

Der Mittelwert des Alters der Probanden beträgt 26.65 Jahre, der Range liegt bei 17 Jahren. Die empirischen Werte für alle untersuchten Worte finden sich im Anhang.

In den Vorexperimenten sind Kontrollitems eingebaut, die aus der Liste von Baschek et alii (1977) übernommen worden sind: BILD, GOLD, GRAB, GRAS, PELZ, WALD und WELT. Es zeigt sich eine hohe Beurteilungsstabilität zwischen den beiden Experimenten: Der Mittelwert der Studie von 1977 beträgt für den Bedeutungsgehalt 4.91, bei der jetzigen Studie liegt der Wert bei 4.78. Das mittlere Abweichungsquadrat ist 0.10. Für jedes Kontrollitem sind die aktuellen Werte und die Daten von 1977 in Anhang II dargestellt. Wegen der hohen Beurteiler-Übereinstimmung der beiden Studien werden Items aus der alten Liste, die im Aufbau den oben genannten Anforderungen entsprechen, in das Experiment aufgenommen.

Für die zehn Worte je Liste, die für diese Studie ausgewählt werden, ergeben sich folgende durchschnittliche Beurteilungen des Bedeutungsgehalt auf einer siebenstufigen Skala: phonologisch ähnliche Worte 4.45, respektive unähnliche Worte: 4.42.

Vorexperiment 6: Konkretheit von Substantiven

Als „konkret“ wird ein Wort definiert, das sich auf „sinnlich erfahrbare Merkmale von Dingen oder Personen bezieht“ (Baschek et al., 1977, 358). Eine Parallelisierung der beiden Wortlisten bezüglich Konkretheit leitet sich aus einer Untersuchung von Walker und Hulme (1999) ab. Die Autoren stellen fest, daß nicht nur Faktoren die Gedächtnisleistung verbessern, die eine effizientere, phonologische Kodierung ermöglichen, sondern auch Faktoren, die als semantische Determinanten gelten.

Das Vorgehen zur Bestimmung der Konkretheit von Substantiven ist analog zu Vorexperiment 5 konzipiert. Wiederum wird eine Fragebogenuntersuchung durchgeführt. Der Fragebogen zur *Konkretheit* von Baschek et alii (1977) wird mit den 57 Items von Vorexperiment 5 zwanzig Studenten vorgelegt, die nicht an Vorexperiment 5 teilgenommen haben.

Mittelwert (26.45) und Range (18 Jahre) des Alters zeigen eine annähernde Äquivalenz der beiden Stichproben von Vorexperiment 5 und 6. Zwischen den eigenen Ergebnissen und denen von Baschek et alii bezüglich der Kontrollitems zeigt sich erneut eine hohe Stabilität: 5.88 (1977) versus 5.43 (2001). Das mittlere Abweichungsquadrat liegt bei 0.41.

Als Kriterium für die Erstellung der Wortlisten der eigenen Experimente wird die Forderung einer minimalen Abweichung der Mittelwerte in ihrer Beurteilung der Konkretheit gestellt. Dies wird mit den Werten von 5.10 für klangähnliche Worte und 5.19 für unähnliche Worte realisiert.

Vorexperiment 7: Bildhaftigkeit von Substantiven

Bildhaftigkeit wird definiert als der Grad, wie leicht oder schwer die Substantive bildhafte Vorstellungen initiieren. Der Einfluß von Bildhaftigkeit auf die Höhe der Gedächtnisleistung wurde oft repliziert (Paivio, 1986). Dies gilt nicht nur für Merkleistungen, die dem Langzeitgedächtnis zugeschrieben werden, sondern auch für das Kurzzeitgedächtnis (Bourassa & Besner, 1994).

Wie schon in den bereits vorgestellten Vorexperimenten, muß auch zur Erfassung der Bildhaftigkeit der 57 Worte eine Fragebogenerhebung durchgeführt werden. Entsprechend dem Vorgehen in Vorexperiment 5 werden die eigenen Items mit dem Fragebogen zur *Bildhaftigkeit* von Baschek et alii (1977) kombiniert. In Anhang II werden die

durchschnittlichen Ratings der zwanzig Probanden auf der siebenstufigen Skala pro Wort dargestellt.

Die zwanzig befragten Studenten haben ein mittleres Alter von 24.65 Jahren (Range: 18 Jahre). Keiner der Probanden hat an Vorexperiment 5 oder 6 teilgenommen. Wie schon in den ersten beiden Vorexperimenten zeigt sich in einer einfaktoriellen Varianzanalyse mit Meßwiederholung kein bedeutsamer Unterschied in der Beurteilung von 1977 zu den heutigen Einschätzungen: 6.10 (1977), respektive 5.98 (2001).

Zur Erstellung der Wortlisten in der vorliegenden Arbeit wird eine möglichst hohe Parallelisierung des Mittelwerts der Bildhaftigkeit der Wortlisten angestrebt. Der Vergleich der Mittelwerte bezüglich Bildhaftigkeit zeigt deskriptiv kaum Unterschiede zwischen phonologisch ähnlichen und unähnlichen Substantiven: 5.46 und 5.45.

Die Listen scheinen parallelisiert zu sein, wie Tabelle 6 im Überblick über die Vorexperimente zeigt.

Tabelle 6: Zusammenfassung der Ergebnisse der Voruntersuchungen

Variable	Grad der phonologischen Ähnlichkeit	
	hoch (Experiment 6)	niedrig (Experiment 7)
Frequenz	177	177
Bedeutungsgehalt	4.45	4.42
Konkretheit	5.10	5.19
Bildhaftigkeit	5.46	5.45

Die sieben bisher untersuchten Kriterien zur Parallelisierung der Wortlisten beziehen sich auf einzelne Worte. Da in den Experimenten nie singuläre Worte, sondern immer Wortsequenzen präsentiert werden, werden zwei weitere Fragebogenstudien mit Substantivpaaren durchgeführt. In Vorexperiment 8 wird die semantische Ähnlichkeit aller möglichen Wortpaare innerhalb einer Liste untersucht, während Vorexperiment 9 als Kontrolluntersuchung dient, ob die beiden Listen zurecht das Label „Phonologisch ähnliche Items“, respektive „Phonologisch unähnliche Items“ tragen.

Vorexperiment 8: Semantische Ähnlichkeit von Wortpaaren

Ziel von Vorexperiment 8 ist herauszufinden, ob bei den ausgewählten Worten (Tabelle 5) eine Konfundierung zwischen phonologischer und semantischer Nähe vorliegt. Eine

Konfundierung würde die Interpretation der Ergebnisse beeinträchtigen, da in einer Vielzahl von Studien belegt ist, daß die Kurzzeitgedächtnisleistung nicht nur sensitiv gegenüber phonologischen Veränderungen ist, sondern auch gegenüber semantischen Variationen (Kapitel 2).

Zur Untersuchung der semantischen Ähnlichkeit der Wortpaare einer Liste wird ein Fragebogen konstruiert, bei dem jedes Wortpaar pro Liste bezüglich seiner Bedeutungsüberschneidung auf einer dreistufigen Skala beurteilt wird. Bei je zehn Worten und einem vollständigen Paarvergleich der Worte innerhalb jeder Liste ergeben sich insgesamt für beide Listen zusammen 90 Wortpaare zur Beurteilung. Der Fragebogen wird 18 Studenten vorgegeben. Das mittlere Alter liegt bei 25.3 Jahren, der Range beträgt 14 Jahre.

Die durchschnittliche Beurteilung der einzelnen Wortpaare auf ihre Bedeutungsüberschneidung hin findet sich im Anhang II. Für die Wortliste mit ähnlichen Items ergibt sich auf der dreistufigen Skala ein Wert für die Bedeutungsüberschneidung von 1.36. Für die andere Liste liegt der Wert bei 1.35. Deskriptiv zeigt sich also kaum ein Unterschied zwischen den beiden Wortlisten. Bei beiden Listen liegen die Werte deutlich im Bereich minimaler Ähnlichkeit. Eine *Varianzanalyse* mit der unabhängigen Variablen „Wortliste“ und der abhängigen Variablen „Beurteilung auf der dreistufigen Skala“ deckt keine signifikanten Unterschiede auf ($F_{(1,88)} = 0.02$, $s_e^2 = 0.142$, n.s.). Die beiden Wortlisten werden bezüglich der semantischen Relation der Wortpaare als äquivalent angesehen.

Vorexperiment 9: Phonologische Ähnlichkeit von Wortpaaren

Die ausgewählten Wortlisten (Tabelle 5) werden von zwei unabhängigen Personen im Vergleich als eher phonologisch ähnlich (linke Spalte), respektive als eher unähnlich bewertet (rechte Spalte). Problematisch bei der Beurteilung der Phonologischen Ähnlichkeit ist, daß erstens Ähnlichkeit keine kategoriale, sondern eine kontinuierliche Variable ist, und zweitens, daß keine definitiven Kriterien zur Beurteilung bestehen. Um die subjektive Einschätzung der beiden Gutachter durch eine unabhängige Quelle zu belegen und zu quantifizieren, wird ein Fragebogen konstruiert, in dem jedes Wortpaar pro Liste bezüglich seiner Klangähnlichkeit auf einer dreistufigen Skala beurteilt werden soll. Der Aufbau des Fragebogens ist äquivalent zu Vorexperiment 8. Die durchschnittliche Beurteilung pro Wortpaar finden sich im Anhang II.

Ziel ist nicht die absolute Feststellung von Ähnlichkeit und Unähnlichkeit in der Phonologie, sondern lediglich der Nachweis einer relativen Differenz zwischen den Wortpaaren. Der Fragebogen wird den 18 Studenten vorgelegt, die bereits an Vorexperiment 8 teilgenommen haben.

Deskriptiv und statistisch zeigt sich der erwartete Unterschied in der Beurteilung der Klangähnlichkeit (einfaktorielle Varianzanalyse: $F_{(1,88)} = 177$, $s_e^2 = 0.153$, $p < .05$). Der Mittelwert bei als klangunähnlich angenommen Worten von 1.06 liegt in der dreistufigen Rating-Skala nahe am Minimum von 1.00. Dagegen werden die Wortpaare der als „ähnlich“ bezeichneten Liste als mittlere Klangähnlichkeit eingestuft mit einem Mittelwert von 2.16 gegenüber einem potentiellen Maximum von 3.00. Für die Studie hat dieses deskriptive Ergebnis keine Relevanz, weil eine relative Differenz als hinreichendes Kriterium definiert wurde.

Zusammenfassend kann man festhalten, daß die Wortlisten mit variierendem Grad der Klangähnlichkeit in sieben Merkmalen, die sich in der Literatur als relevant herauskristallisiert haben, parallelisiert sind. Es liegt auch keine Konfundierung mit semantischer Nähe vor. Zudem wird durch eine unabhängige Quelle belegt, daß die Materialien sich bezüglich phonologischer Ähnlichkeit unterscheiden.

5.3 PROBANDEN

Die Stichprobe mit insgesamt 624 Probanden wird aus drei Populationen gezogen: Es werden deutsche, spanische und chinesische Studenten untersucht. Die Restriktion auf Studenten gründet sich in Studien, die einen Zusammenhang zwischen Gedächtnisleistung und Intelligenz belegen (Klein, 1995). Die Korrelation zwischen dem Ergebnis des Untertests „Höhe der Gedächtnisspanne“ im Wechsler-Intelligenztest und dem Gesamtwert der Intelligenzmessung in diesem Test beträgt .55 (Wechsler, 1956). Ferner wird ein Zusammenhang zwischen kontrollierter Aufmerksamkeit als Leistung der Zentralen Exekutive und fluider Intelligenz nachgewiesen (Engle et al., 1999). Da sowohl für das Sub- als auch für das Supraspannen-Experiment die Relevanz der Zentralen Exekutive für die Aufgabenlösung gezeigt wurde (Bredenkamp & Hamm, 2001) und die ohnehin sehr anstrengende Untersuchungssitzung nicht durch einen zusätzlichen Intelligenztest verlängert werden sollte, wird die Stichprobe durch die alleinige Auswahl von Studenten homogenisiert. Dies hat zudem den Effekt, daß die Altersvarianzen reduziert werden. Neben dem Alter des Probanden wird auch die Anzahl der an der Universität verbrachten Jahre

erfaßt. Dieser Parameter wird als Quelle kultureller Unterschiede betrachtet, der als Störvariable auftreten könnte (Stigler, Lee & Stevenson, 1986).

Insgesamt wird für jede Experimentreihe eine Erhebung von 57 Probanden geplant. Dieser Stichprobenumfang erweist sich als optimal für das Supraspannen-Experiment bei Annahme einer mittleren Effektstärke und Fixierung der beiden Fehlerarten auf .10 (Erdfelder, Faul & Buchner, 1996). Bei Experimentreihe 1 wurde der Stichprobenumfang um weitere 30 Probanden erweitert. Die a priori Fehlerwahrscheinlichkeiten reduzierten sich auf .05 bei angenommener mittlerer Effektstärke. Die Probanden nahmen jeweils an allen drei Untersuchungen einer Experimentreihe teil, mit Ausnahme von Experimentreihe 8. In dieser Reihe führten 57 Probanden das Supraspannen-Experiment durch und weitere 20 Versuchspersonen das Artikulationszeit- und das Subspannen-Experiment.

Das Akquirieren der deutschen Studenten erfolgte an den Universitäten Bonn und Würzburg. Für die Teilnahme an den Hauptexperimenten wurden entweder acht DM gezahlt oder Versuchspersonenstunden bescheinigt, die zur Anmeldung zum Vordiplom nötig sind. Die Teilnehmer der neun Vorexperimente erhielten lediglich Süßigkeiten. Insgesamt nahmen 453 deutsche Studenten verschiedener Studienrichtungen an der Erhebung teil. 40 Probanden wirkten bei den Vorexperimenten eins bis vier mit, 78 Personen nahmen insgesamt an den Fragebogenuntersuchungen teil, die sie nicht von einer Teilnahme der Experimente ausschloß.

Die 114 chinesischen Probanden wurden in Peking an der Academy of Science untersucht. Die Zuordnung zu den beiden Experimenten erfolgte zufällig. Das Versuchspersonenhonorar liegt bei umgerechnet 4.50 DM. Dies entspricht den dortigen Konventionen.

An der Universität in Madrid und an der Fachhochschule für Übersetzen und Dolmetschen in Köln wurden die 57 spanischen Probanden akquiriert, deren Teilnahme umgerechnet mit zehn DM vergütet wurde.

In Kapitel 7.1 findet sich neben einer Darstellung der deskriptiven Charakteristika auch eine Überprüfung der Stichprobenäquivalenz.

5.4 OPERATIONALISIERUNG DER ABHÄNGIGEN VARIABLEN

Die Operationalisierung der abhängigen Variablen ist so gewählt, daß die Variablenvalidität möglichst hoch ist. Zudem bewährte sich diese Operationalisierung – mit Ausnahme der Reproduktionszeit – bereits in anderen Studien (Bredenkamp & Klein, 1998). Zur

Überprüfung der Adäquatheit der Operationalisierung wird in Kapitel 6.1.3 eine konvergente Validierung der Parameter durchgeführt.

Zuerst wird die *Artikulationszeit* in jeder Experimentreihe in einem eigenen Versuch erhoben. Wie bereits in Kapitel 5.1.1 berichtet, werden jedem Probanden je fünf Reihen mit sechs, acht und zehn Items vorgegeben. Über diese 15 Datenpunkte wird intraindividuell eine *lineare Regression* mit der unabhängigen Variablen „Sequenzlänge“ und dem abhängigen Faktor „Dauer der Artikulation“ berechnet. Der Steigungsparameter wird interpretiert als Artikulationszeit pro Item.

Die Höhe der *Gedächtnisspanne* wird im Subspannen-Experiment ermittelt und ergibt sich aus dem Mittelwert der zehn lokalen Maxima und Minima. Diese Punkte sind definiert als Umkehrpunkte zwischen richtigen und fehlerhaften Reproduktionen (Kapitel 5.4.2).

Die *Reproduktionszeit* errechnet sich aus der Reproduktionsdauer der korrekt wiedergegebenen Sequenzen im Subspannen-Experiment. Als Reproduktionsintervall wird das gesamte Intervall zwischen Beginn des WIEDERGABE-Signals und dem Tastendruck durch den Versuchsleiter definiert. Nach drei Kriterien wird ein Ausschluß von Daten aus der Analyse vorgenommen:

- Alle fehlerhaften Reproduktionen werden ausgeschlossen.
- Die Zeiten aller richtigen Antworten, deren Sequenzlänge die ermittelte Gedächtnisspanne des Probanden übersteigt, werden in der Analyse nicht berücksichtigt. Diese Entscheidung liegt darin begründet, daß die Varianz der Reproduktionszeit extrem steigt im Vergleich zu den Antworten, deren Sequenzlänge niedriger als die der Gedächtnisspanne ist. Zudem ist der Anstieg der Reproduktionszeiten von Reihen, die über der mittleren Gedächtnisspanne des Probanden liegen, nicht mehr linear.
- Nach dem Ausschlußverfahren von Tukey (1977), das intraindividuell angewendet wird (Formel 11), werden Ausreißer eliminiert. Die Ausreißeranalyse wird deshalb eingeführt, um Verzerrungen der Zeitmessung durch Itemwiederholung, Verbesserungen der Reproduktion, Nachfragen oder Räuspern zu eliminieren.

Formel 11: $Q_3 \pm 1.5 * (Q_3 - Q_1)$

Q_1 und Q_3 stehen für das erste, respektive dritte Quartil der Verteilung der Zeitmessung. Alle Reproduktionszeiten außerhalb des nach Formel 11 definierten Intervalls wurden

ausgeschlossen. Insgesamt ergab sich nur bei wenigen Probanden eine Datenreduktion durch diese Analyse.

Aus den bereinigten Rohdaten wird die Reproduktionszeit durch den Quotienten zwischen benötigter Reproduktionsdauer und Sequenzlänge separat für jeden Probanden ermittelt. Eine Regressionsanalyse in Analogie zur Artikulationszeit ist nicht möglich, da in Experimenten mit Phonologischer Ähnlichkeit und Artikulatorischer Unterdrückung bei manchen Probanden die Gedächtnisspanne so niedrig ist, daß nur die Länge einer einzigen Sequenz richtig reproduziert wird. Bei konstanter Reihenzahl kann keine Regression durchgeführt werden.

Der vierte relevante Parameter „*Anzahl der Durchgänge*“ bezieht sich auf das Superspannen-Experiment: Er bezeichnet die Anzahl der Lernversuche, die zum Erreichen des Kriteriums einer zweimaligen korrekten Wiedergabe benötigt werden.

In die Auswertungen der intra-, repektive interexperimentellen Hypothesenprüfung gehen vier Parameter pro Versuchsperson ein: die Artikulationszeit, die Höhe der Gedächtnisspanne, die Reproduktionszeit sowie die bis zum Kriterium benötigte Anzahl von Lerndurchgängen im Supraspannen-Experiment für eine fixierte Darbietungszeit.

In den nächsten beiden Kapiteln werden die Hypothesen in bezug auf diese vier Parameter intra- und interexperimentell ausgewertet und in Kapitel 8 diskutiert.

6 INTRAEXPERIMENTELLE HYPOTHESENPRÜFUNG

In diesem Kapitel werden drei Arten von Hypothesen separat für jede Experimentreihe untersucht: Kontrollhypothesen, Zentrale und Explorative Hypothesen. Die in Kapitel 4 formulierten Arbeitshypothesen sowie die Invarianzhypothesen von Baddeley und Doshier und Ma beziehen sich auf den interexperimentellen Vergleich und werden in Kapitel 7 geprüft. Im folgenden werden zuerst die *Kontrollhypothesen* untersucht. Diese Hypothesen wurden nicht in Kapitel 4 dargestellt, da es sich um Anforderungen handelt, die generell an Experimente gestellt werden, und deshalb keiner speziellen Begründung bedürfen (Kapitel 6.1). Als zweites werden die *Zentralen Hypothesen* zur mTTH (Kapitel 6.2) und zur Äquivalenzhypothese (Kapitel 6.3) auf ihre intraexperimentelle Gültigkeit hin überprüft. Als drittes wird die *Explorative Hypothese* zur Verarbeitungszeit im Langzeitgedächtnisspeicher untersucht (Kapitel 6.4). Die Ergebnisse werden in Kapitel 6.5 kurz zusammengefaßt.

Zur Terminologie müssen folgende Anmerkungen gemacht werden:

- Bei F- oder t-Tests wird jeweils der empirisch ermittelte Testwert angegeben, wobei jeweils die Anzahl der Freiheitsgrade in Klammern angefügt werden: $t_{(109)}$, $F_{(2,34)}$.
- Mit „ s_e^2 “ wird der Quotient zwischen der Quadratsumme innerhalb der Gruppen und der Anzahl der Freiheitsgrade bezeichnet.
- Wenn ein Wert geschätzt wird, sei es die tatsächlich beobachtete Effektgröße oder die Anzahl der Lernversuche im Supraspannen-Experiment, wird der Parameter durch dieses Zeichen „ $\hat{}$ “ markiert.
- Unter „ s_{emp}^2 “ wird die Varianz der Abweichung zwischen den erhaltenen und den erwarteten Werten verstanden.
- Die Abkürzung „n.s.“ steht für „nicht signifikant“, respektive $p < .05$ für „signifikant bei einem alpha-Fehler von .05“.
- Die Angabe der Ausprägung des Fehlers zweiter Art bezieht sich immer auf eine a priori Analyse nach Gpower (Erdfelder et al., 1996).
- Auf die Schätzung einer Effektgröße wird verzichtet, wenn sich ein Testresultat ergibt, das die Signifikanzgrenze nicht überschreitet.
- Die Benennungen der Parameter weichen nicht von den in Kapitel 3 eingeführten ab.

6.1 KONTROLLE DER EXPERIMENTGÜTE UND DESKRIPTIVE ERGEBNISSE

Sowohl die Kontrollanalysen als auch die Darstellung der Ergebnisse beziehen sich auf vier abhängige Variablen: die der Artikulationszeit, Höhe der Gedächtnisspanne,

Reproduktionszeit und Anzahl der Lerndurchgänge im Supraspannen-Experiment (Kapitel 5.4). Alle ermittelten Werte jedes Experiments sind in den Tabellen 11 und 12 am Ende des Kapitels 6.1.4 aufgeführt. Im folgenden werden stichprobenartig Hinweise auf das Gelingen der Experimente untersucht. Berücksichtigt wird die Reliabilität der Zeitmessung (Kapitel 6.1.1), die Durchführungsobjektivität bezüglich der Untersuchungsapparatur und der dargebotenen Reihenfolge der Gedächtnistests (Kapitel 6.1.2) sowie die Validität und Validierung (Kapitel 6.1.3).

6.1.1 RELIABILITÄT DER ZEITMESSUNG

Die Untersuchung der Reliabilität der Zeitmessung dient der Bewertung der Aussagekraft der Parameter „Artikulationszeit“ und „Reproduktionszeit“. Die konkreten Ausprägungen dieser beiden Parameter ergeben sich nach den in Kapitel 5.4 vorgestellten Verfahren. Für jede Versuchsperson existiert nur ein Wert der Artikulationszeit und nur ein Wert der Reproduktionszeit. Da sich die Zeitmessungen dieser beiden Parameter im Bereich von Millisekunden befinden und somit sensitiv gegenüber geringfügigen Störungen sind, scheint eine Reliabilitätsanalyse sinnvoll.

Als Maß wird das standardisierte Cronbach-Alpha (Cronbach, 1951) gewählt und intraindividuell ausgewertet (Formel 12).

$$\text{Formel 12} \quad \alpha = \frac{k * \overline{Corr}}{1 + (k - 1) * \overline{Corr}} \quad \text{mit } \overline{Corr} = \frac{2 * \sum_{i < j} r_{ij}}{[k * (k - 1)]}$$

k entspricht der Anzahl der Gruppen, hier wurden die Item-Interkorrelation ausgewertet. Als Gruppe wird der Umfang der Sequenzlänge definiert. Die Anzahl der Gruppen variiert für die Auswertung der Reproduktionszeit zwischen den Probanden in Abhängigkeit der Höhe der Gedächtnisspanne.

Aus allen folgenden Analysen werden die Daten von Versuchspersonen der Artikulations-, respektive Reproduktionszeit ausgeschlossen, deren Reliabilitätswerte niedriger als .75 sind. Der Wert von .75 ist eine willkürlich gesetzte Grenze, die sich an Konventionen orientiert (für eine Diskussion über Konventionen in der Psychologie: Westmeyer, 1998). Tabelle 7 gibt einen Überblick über die Höhe des Cronbach-Alpha für die beiden Parameter nach Ausschluß der Werte mit geringen Reliabilitätswerten, über die Anzahl der eingegangenen Fälle und über die empirisch ermittelte Höhe von Artikulations- und Reproduktionszeit sowie die Standardabweichungen. Bezüglich der Auswertung der Daten spanischer Probanden (Experiment 3) muß auf die Analyse der Durchführungsobjektivität hinsichtlich der

Testbedingung vorgegriffen werden (Kapitel 6.1.3). In Tabelle 7 werden nur die in Deutschland erhobenen Daten berücksichtigt.

Tabelle 7 Empirische Mittelwerte (Wert) und Standardabweichungen (in eckigen Klammern) in Millisekunden, Anzahl gültiger Fälle (n) und Mittleres Cronbach-Alpha (α) für die Parameter Artikulationszeit und Reproduktionszeit der acht Experimente.

Experiment	Artikulationszeit			Reproduktionszeit		
	Wert	<i>n</i>	α	Wert	<i>n</i>	α
1 (Stille, d, ZS)	303 [62]	84	.96	527 [97]	64	.88
2 (art.U, d, ZS)	290 [54]	56	.96	709 [150]	28	.88
3 (Stille, s, ZS)	356 [42]	11	.95	727 [163]	6	.88
4 (Stille, c, ZS)	244 [63]	50	.89	328 [54]	36	.88
5 (art.U, c, ZS)	269 [65]	48	.88	430 [119]	34	.87
6 (Stille, d, Pae)	412 [73]	57	.96	727 [148]	51	.88
7 (Stille, d, Pu)	353 [62]	57	.97	635 [99]	50	.87
8 (Stille, d, ZW)	328 [37]	20	.97	498 [70]	18	.86

In allen folgenden Tabellen wird dieses Kodierungssystem zur Beschreibung der Experimente verwendet, deswegen erfolgt lediglich an dieser Stelle eine genaue Definition der Abkürzungen.

- „Stille“ bedeutet ohne Zweitaufgabe
- „art.U“ steht für Artikulatorische Unterdrückung
- d – s – c meint deutsche, spanische, respektive chinesische Muttersprache
- „ZS“ steht für Zahlensymbole, „ZW“ für Zahlworte
- Mit „Pae“ wird phonologisch ähnliches Material gekennzeichnet, mit „Pu“ phonologisch unähnliche Stimuli.

6.1.2 DURCHFÜHRUNGSOBJEKTIVITÄT

Von wissenschaftlicher Beobachtung soll gesprochen werden, wenn abgrenzbare empirische Sachverhalte in *objektiver* und *prinzipiell wiederholbarer* Weise unter Zugrundelegung eines vorgegebenen Vokabulars in Beobachtungsprotokolle überführt werden (Erdfelder, 1994, S. 51).

Die Überführung in Beobachtungsprotokolle erfolgt computergesteuert und bezieht sich auf klar definierte Zeitmessungen, respektive den Mittelwert von präzise bestimmten Extrema im Subspannen-Experiment oder die Anzahl der Lernversuche im Supraspannen-Experiment. Im Subspannen-Experiment wird lediglich die Korrektheit der Wiedergabe gemessen, ohne Auswertungen bezüglich der seriellen Positionskurve.

Um der zitierten Beschreibung gerecht zu werden, werden im folgenden drei Objektivitätskriterien analysiert: Untersuchungen zur Testbedingung, zum Reihenfolge-Effekt und zur Replizierbarkeit.

DURCHFÜHRUNGSOBJEKTIVITÄT BEZÜGLICH DER TESTBEDINGUNG

Unter „Testbedingung“ wird die Variation von Versuchsleiter, Computerapparaturen und Testumgebungen verstanden. Die Überprüfung der Durchführungsobjektivität soll Störeffekte durch variierende Testbedingungen aufdecken. Da kein Meßwiederholungsdesign vorliegt, sind die Faktoren „Testbedingung“ und „Versuchsperson“ konfundiert. Bei der Erhebung der Daten zu den Experimentreihen mit chinesischen Probanden (Experiment 4 & 5) sowie mit deutschen Probanden und den drei Wortmaterialien (Experiment 6, 7 & 8) gibt es nur eine Durchführungsmodalität. Eine Überprüfung der Testbedingung entfällt. In den Experimentreihen mit variierenden Untersuchungsbedingungen werden die Parameter „Artikulationszeit“, „Gedächtnisspanne“ und „Anzahl der Lerndurchgänge im Supraspannen-Experiment“ hinsichtlich der Objektivität verglichen. Alle Probanden einer Experimentreihe führten das Artikulationszeit-Experiment und das Subspannen-Experiment unter identischen Bedingungen durch. Anders dagegen ist das Vorgehen beim Supraspannen-Experiment: Die Stichprobe einer Experimentreihe wurde nach dem Zufallsprinzip in drei gleich große Untergruppen geteilt, denen jeweils eine andere Darbietungszeit (10, 20 oder 30 Sekunden) zugeordnet wurde. Für die Prüfung der Durchführungsobjektivität erfolgt für das Artikulationszeit- und das Subspannen-Experiment eine einfaktorielle Varianzanalyse, für das Supraspannen-Experiment entsprechend eine zweifaktorielle Auswertung.

In den Experimenten 1 und 2 decken *Varianzanalysen* mit der dreifach (Experiment 1), respektive zweifach (Experiment 2 & 3) gestuften, unabhängigen Variablen „Testbedingung“ und der abhängigen Variablen „Artikulationszeit“ keine signifikanten Effekte auf. Die empirischen F-Werte liegen durchweg unter dem kritischen F-Wert bei einem Alpha-Niveau von .10 (Experiment 1: $F_{(2;81)} = 0.161$, Fehlervarianz: $s_e^2 = 3928$, $\beta = .05$, n.s.; Experiment 2: $F_{(1;54)} = 1.89$, $s_e^2 = 2883$, $\beta = .11$, n.s.). Dasselbe Verfahren zeigt bei den Artikulationszeiten der spanischen Probanden eine Verletzung der Durchführungsobjektivität bei Messung dieser

Zeit unter zwei verschiedenen Bedingungen (Experiment 3: $F_{(1,55)} = 10.33$, $s_e^2 = 2907$, $p < .05$). Die geschätzte Effektstärke liegt bei $f' = 0.16$, die Power des Tests bei dieser Effektstärke ist .80.

Sowohl für das *Sub- als auch für das Supraspannen-Experiment* werden keine Verletzungen der Durchführungsobjektivität bezüglich der Testbedingung festgestellt, wie die Ergebnisse der Varianzanalysen in Tabelle 8 zeigen.

Tabelle 8 Überprüfung der Durchführungsobjektivität bezüglich der Testbedingung in den beiden Gedächtnistests: Ergebnis der Varianzanalysen und Angabe der Fehlervarianz

Testbedingung Experiment	Subspannen-Exp.	Supraspannen-Exp.	
	Gedächtnisspanne	Anz. Durchgänge	Interaktion
1 (Stille, d, ZS)	$F_{(2,84)} = 0.137$ [0.585]	$F_{(2,78)} = 0.754$ [4.705]	$F_{(4,78)} = 1.211$ [4.705]
2 (art.U, d, ZS)	$F_{(1,55)} = 0.243$ [0.517]	$F_{(1,51)} = 1.989$ [10.10]	$F_{(2,51)} = 0.785$ [10.10]
3 (Stille, s, ZS)	$F_{(1,55)} = 2.098$ [0.497]	$F_{(1,51)} = 1.037$ [6.512]	$F_{(2,51)} = 0.464$ [6.512]

In den Zeilen sind die drei Experimentreihen eingetragen, in denen die Testbedingungen variiert werden. In der Spalte „Gedächtnisspanne“ sind die Ergebnisse der *Varianzanalysen* zum Subspannen-Experiment mit der unabhängigen Variablen „Testbedingung“ und der abhängigen Variablen „Gedächtnisspanne“ aufgeführt. Die Spalte „Anz. Durchgänge“ gibt den Haupteffekt des unabhängigen Faktors „Testbedingung“ auf die Anzahl der Lernversuche bis zum Erreichen des Kriteriums im Supraspannen-Experiments wieder (*zweifaktorielle Varianzanalyse*: unabhängige Variablen „Darbietungszeit“ und „Testbedingung“; abhängige Variable „Lernversuche“). Die Spalte „Interaktion“ nimmt Bezug auf den Test einer möglichen Wechselwirkung zwischen den Faktoren „Testbedingung“ und „Darbietungszeit“. In den eckigen Klammern ist jeweils die Fehlervarianz eingetragen.

In keiner der neun Auswertungen wird eine Verletzung der Durchführungsobjektivität aufgedeckt. Bei einer Festlegung des Fehlers erster Art auf .10 und Annahme einer mittleren Effektstärke ($f = 0.25$) ergibt sich für die Auswertungen mit der abhängigen Variablen „Gedächtnisspanne“ ein Fehler zweiter Art in der Höhe von .05 in Experiment 1 und in den beiden anderen Experimentreihen von .11. Die Analysen zum Supraspannen-Experiment ($\alpha = .10$, $f^2 = .15$) zeigen für Experiment 2 und 3 eine Power von .89 bei Prüfung des Haupteffekts (Experiment 1: .94), respektive ergibt sich bei Überprüfung der Interaktion ein beta-Fehler in der Höhe von .18 (Experiment 1: .12).

DURCHFÜHRUNGSOBJEKTIVITÄT BEZÜGLICH DER EXPERIMENTREIHENFOLGE

Die zweite Überprüfung der Durchführungsobjektivität bezieht sich auf Auswirkungen variierender *Darbietungsreihenfolgen*: Fünfzig Prozent der Probanden führten erst das Supraspannen-Experiment aus, die andere Hälfte begann mit dem Subspannen-Experiment. Dieses Design wird in allen Experimentreihen realisiert. Lediglich in der letzten Experimentreihe (Experiment 8) bearbeiteten die Versuchspersonen entweder das Supraspannen-Experiment oder das Subspannen-Experiment.

Die folgenden Analysen werden analog zu denen der Durchführungsobjektivität bezüglich der Testbedingung durchgeführt. Für jede der verbleibenden sieben Experimentreihen wird separat eine *Varianzanalyse* mit der unabhängigen Variablen „Reihenfolge“ und dem abhängigen Faktor „Höhe der Gedächtnisspanne“ durchgeführt. Die Ergebnisse sind in Tabelle 9 dargestellt, deren Aufbau äquivalent zu Tabelle 8 ist.

Die *Varianzanalysen* beziehen als unabhängige Variable die „Reihenfolge der Gedächtnistests“ ein. Keine der hier dargestellten 21 Auswertungen deckt einen systematischen Einfluß der Bearbeitungsreihenfolge der Gedächtnistests auf. Dieses Resultat spricht dafür, daß das Auftreten von Sequenzeffekten in dieser Studie unwahrscheinlich ist. Die Irrtumswahrscheinlichkeit des Fehlers erster Art beträgt hier .10, die Prüfung der Varianzhomogenität deckt keine signifikanten Unterschiede auf.

- Für die Auswertungen mit der abhängigen Variablen „Gedächtnisspanne“ liegt bei Annahme eines großen Effekts ($f = .40$) die Power des Test bei Experiment 2-8 für die Haupteffekte bei .91, in Experiment 1 bei .98.
- Bei Prüfung des Haupteffekts der Reihenfolge auf die abhängige Variable „Anz. Durchgänge“ ergibt sich eine Power bei den Experimenten 2-8 von .70, bei Experiment 1 von .88 ($f^2 = .15$).
- Für die Analyse der Interaktionseffekte zwischen den beiden unabhängigen Variablen „Reihenfolge der Gedächtnistests“ und „Darbietungszeit“ ergibt sich in Experiment 1 ein Fehler zweiter Art von .05 und in den verbleibenden Experimenten ein Wert von .18.

Tabelle 9 Überprüfung der Durchführungsobjektivität bezüglich der Reihenfolge der beiden Gedächtnistests: Ergebnis der Varianzanalysen und Angabe der Fehlervarianz (in eckigen Klammern).

Reihenfolge Experiment	Subspannen-Exp.	Supraspannen-Exp.	
	Gedächtnisspanne	Anz. Durchgänge	Interaktion
1 (Stille, d, ZS)	$F_{(1,85)} = 0.310$ [0.578]	$F_{(1,81)} = 0.035$ [4.76]	$F_{(2,81)} = 1.61$ [4.76]
2 (art.U, d, ZS)	$F_{(1,55)} = 2.09$ [0.500]	$F_{(1,51)} = 1.39$ [10.05]	$F_{(2,51)} = 1.20$ [10.05]
3 (Stille, s, ZS)	$F_{(1,55)} = 1.64$ [0.501]	$F_{(1,51)} = 0.022$ [6.03]	$F_{(2,51)} = 3.14$ [6.03]
4 (Stille, c, ZS)	$F_{(1,55)} = 0.283$ [1.14]	$F_{(1,51)} = 0.817$ [2.73]	$F_{(2,51)} = 1.53$ [2.73]
5 (art.U, c, ZS)	$F_{(1,55)} = 0.269$ [1.54]	$F_{(1,51)} = 0.856$ [3.12]	$F_{(2,51)} = 1.14$ [3.12]
6 (Stille, d, Pae)	$F_{(1,55)} = 0.379$ [0.272]	$F_{(1,51)} = 2.71$ [5.89]	$F_{(2,51)} = 1.83$ [5.89]
7 (Stille, d, Pu)	$F_{(1,55)} = 0.890$ [0.478]	$F_{(1,51)} = 2.79$ [6.25]	$F_{(2,51)} = 1.63$ [6.25]

REPLIZIERBARKEIT

Das Kriterium der *Replizierbarkeit* wird anhand von Experimentreihe 1 überprüft. Sieben Monate nach Erhebung der ersten 57 Probanden wurden weitere 30 Probanden getestet (für eine Begründung dieser Nacherhebung: Kapitel 6.2.1). *Varianzanalysen* mit dem unabhängigen Faktor „Zeitpunkt der Erhebung“ und den abhängigen Variablen „Artikulationszeit“ ($F_{(1,82)} = 0.008$, $s_e^2 = 3895$, n.s.), „Gedächtnisspanne“ ($F_{(1,85)} = 0.023$, $s_e^2 = 0.580$, n.s.) und „Reproduktionszeit“ ($F_{(1,62)} = 1.78$, $s_e^2 = 9215$, n.s.) decken keine signifikanten Unterschiede zwischen den beiden Erhebungszeiträumen auf. Auch die mehrfaktorielle *Varianzanalyse* („Darbietungszeit“ und „Erhebungszeitpunkt“, abhängige Variable „Anzahl der Lernversuche“) zur Überprüfung der prinzipiellen Replizierbarkeit des Supraspannen-Experiments zeigt weder einen Haupteffekt des Faktors „Erhebungszeitraum“ noch eine signifikante Interaktion mit der Variablen „Darbietungszeit“: $F_{(1,81)} = 0.374$, n.s. und $F_{(2,81)} = 0.758$, $s_e^2 = 4.787$, n.s..

6.1.3 VALIDITÄT UND KONVERGENTE VALIDIERUNG EINZELNER PARAMETER

Aus der umfangreichen Literatur zum Thema „Validität“ werden nach Erdfelder (1994) interne und externe Validität ausgewählt und exemplarisch behandelt. Aus dem Monotherapy-Multimethod Ansatz (Westmeyer, Hannemann, Nell, Völkel & Winkelmann, 1987) wird das Konzept konvergenter Validierung entlehnt. Dieses Verfahren dient einer Abschätzung über das Gelingen der Experimente im Vergleich zu bereits publizierten Studien.

Für die Auswertungen zu den explorativen Hypothesen wird die *interne Validität* durch Stichprobenhomogenität gesichert. So können an dem Experiment nur Studierende teilnehmen. Dies bedeutet keineswegs eine Einschränkung der Repräsentativität dieser Untersuchung und der *externen Validität*, denn bei nomologischen Aussagen ist jedes Individuum gleich repräsentativ, unabhängig davon ob eine Zufallsstichprobe gezogen wird (Erdfelder & Bredenkamp, 1994). Als potentieller Störfaktoren werden die an der Universität verbrachten Jahre und das Lebensalter der Probanden erfaßt und miteinander verglichen: In Kapitel 7.1 findet sich eine ausführliche Überprüfung der Stichprobenäquivalenz. Zudem werden die Untersuchungsbedingungen möglichst konstant gehalten (Kapitel 6.1.2). Für die Prüfung der Arbeitshypothesen und der Zentralen Hypothesen wird jeweils a priori die Teststärke und Testfairneß berechnet und durch den Stichprobenumfang möglichst maximiert. Die konkreten Ausprägungen dieser zwei Indikatoren der Validität werden in jedem Kapitel angegeben.

Die *konvergente Validierung* der erhaltenen Daten erfolgt durch einen Vergleich mit bereits veröffentlichten Werten. Da die Hypothese von Doshier und Ma sowie die mTTH in den hier realisierten Kontexten erstmals untersucht werden, existieren für die Parameter „Reproduktionszeit“ und „Anzahl der Lerndurchgänge im Supraspannen-Experiment“ keine Vergleichsdaten, so daß die konvergente Validierung nur exemplarisch an den Parametern der Gedächtnisspanne und Artikulationszeit durchgeführt werden kann. In Tabelle 10 werden die erhaltenen und aus der Literatur entlehnten Werte deutscher, spanischer und chinesischer Probanden und die jeweiligen Standardabweichungen (in Klammern) deskriptiv gegenübergestellt. In die Darstellung der empirischen Werte der Artikulationszeit gehen nur Daten dieser Arbeit ein, deren Reliabilität .74 übersteigt (Kapitel 6.1.1).

Tabelle 10 Konvergente Validierung von Artikulationszeit in Millisekunden und Gedächtnisspanne bei deutschen, spanischen und chinesischen Probanden. Die Standardabweichung findet sich in den eckigen Klammern.

Experiment	Artikulationszeit	Gedächtnisspanne
1 (Stille, d, ZS): empirisch	303 ms [62]	6.15 Items [0.76]
Lüer et al. (1998)	363 ms [79]	6.22 Items [1.23]
8 (Stille, d, ZW): empirisch	328 ms [37]	5.82 Items [0.84]
Lüer et al. (1998)	377 ms [73]	5.88 Items [0.98]
3 (Stille, s, ZS): empirisch	356 ms [42]	5.76 Items [0.71]
Chincotta & Hoosain (1995)	439 ms [78]	6.50 Items [0.91]
4 (Stille, c, ZS): empirisch	244 ms [63]	8.37 Items [1.04]
Lüer et al. (1998)	236 ms [34]	7.33 Items [1.10]
7 (Stille, d, Pu): empirisch	353 ms [62]	5.29 Items [0.69]
Bosshardt & Laug (1995)	340 ms [67]	5.34 Items [0.78]

Hinsichtlich der Artikulationszeit scheint eine Vergleichbarkeit der Daten aus den verschiedenen Quellen gegeben zu sein:

- Für die Experimente mit Zahlensymbolen und deutschen, respektive chinesischen Probanden und bei der Bedingung phonologisch unähnlicher Worte (Experiment 1, 4, & 7) liegen die der Literatur entlehnten Werte innerhalb des Intervalls des hier ermittelten Mittelwerts plus/minus der Standardabweichung. Der Wert der Artikulationszeit, der aus der Studie von Bosshardt und Laug (1995) übernommen wurde, wurde als Artikulationsrate angegeben. Für die Vergleichbarkeit innerhalb dieser Tabelle wird dieser Wert in die Artikulationszeit umgerechnet.
- Der Vergleich der hier gemessenen Werte mit denen von Lüer et alii (1998) ist trotz verschiedener Meßmethoden zulässig: Die Daten der Artikulationszeit nach Lüer ergeben sich aus einer On-Line Erhebung während des Gedächtnisexperimentes. Jedoch gehen in diese Messung nur die Zeiten der Aussprache an sich in die Auswertung ein, nicht jedoch die Pausenintervalle.
- Für die Experimente mit spanischen Probanden und Zahlensymbolen (Experiment 3) sowie deutschen Probanden und Zahlworten (Experiment 8) liegen die aus der Literatur entnommenen Werte außerhalb des Intervalls einer Standardabweichung. Dieser Effekt kann aber wegen der kleinen realisierten Stichprobe und der resultierenden geringen Standardabweichung zustande kommen. Dafür spricht, daß die empirischen Werte dieser Arbeit in den aus den zitierten Studien entlehnten Intervalle enthalten sind.

Obwohl Puckett und Kausler (1984) darauf verweisen, daß unterschiedliche Erhebungsmethoden in Bezug auf die Artikulationszeit großen Einfluß auf die Höhe der erhaltenen Parameter ausüben, zeigt sich hier eine gute Übereinstimmung zwischen den verschiedenen Studien.

Für die Messungen der Gedächtnisspanne bestehen für alle drei Experimente mit deutschen Probanden kaum Unterschiede zu den Werten vergleichbarer Untersuchungen. Dies gilt nicht für die Messungen mit den beiden anderen Muttersprachen. Für chinesische Probanden zeigen sich deskriptiv große Unterschiede zwischen den Studien. Diese beiden Werte aber weichen nicht mehr als eine Standardabweichung vom Mittelwert des jeweils anderen Experiments ab. Für spanische Probanden dagegen reicht das Intervall der hier erhaltenen Ergebnisse von 5.05 bis 6.47 Items. Der aus der Literatur übernommene Wert mit 6.50 (Tabelle 10) liegt außerhalb dieses Wertebereichs. Da bei Chincotta und Hoosain (1995) kein adaptiver Test verwirklicht wurde und nur in der Einheit von halben Items gemessen wird, scheint der Unterschied zwischen 6.47 und 6.50 vernachlässigbar zu sein.

6.1.4 DISKUSSION UND ERGEBNISSE

Zusammenfassend läßt sich feststellen, daß die *Reliabilität* der Artikulationszeitmessung zufriedenstellend ist, nicht jedoch die der Reproduktionszeit. Durch den Ausschluß aller Zeiten, deren Cronbach-Alpha unter .75 liegen, wird die Reliabilität wirksam erhöht. Der gesunkene Stichprobenumfang und eine ungleich große Anzahl gültiger Fälle muß dafür in Kauf genommen werden. Nachteil dieser Ausschlußmethode ist zudem die nachlassende Übersichtlichkeit: je nach abhängiger Variable ändert sich die Anzahl gültiger Fälle. Beispielsweise stehen in Experiment 1 drei verschiedene Mittelwerte der Anzahl erinnerter Items nebeneinander, je nachdem welche Hypothese untersucht wird. Bei Prüfung der Baddeley-Hypothese werden 84 Daten akzeptiert. Die Gedächtnisspanne liegt bei 6.16 Items. Untersucht man dagegen den Ansatz von Doshier und Ma, gehen nur die Daten von 64 Probanden in die Auswertung ein, deren mittlere Gedächtnisspanne 6.26 Items beträgt. Bei Prüfung der Äquivalenzhypothese muß auf keine Zeitmessung zurückgegriffen werden, so daß alle erhobenen 87 Daten berücksichtigt werden. Der Mittelwert der Gedächtnisspanne liegt dann bei 6.15 Items. In Anbetracht der reduzierten Aussagekraft von Zeitmessungen mit niedriger Reliabilität muß diese erschwerte Übersichtlichkeit vorgezogen werden. Zudem scheint der Ausschluß von Probanden aufgrund eines niedrigen Reliabilitätskriteriums in einer der Zeitmessungen keine systematische Verzerrung der Mittelwerte der Gedächtnisspanne zu bewirken, wie der Vergleich der oben aufgeführten Daten zeigt (6.16, 6.26 und 6.15 Items).

Die Auswertung der *Durchführungsobjektivität bezüglich der Testbedingung* offenbart Probleme bei der Erhebung der spanischen Probanden. Es werden signifikante Abweichungen je nach Testbedingung aufgedeckt. Diese Verletzung der Durchführungsobjektivität ist als gravierend anzusehen und kann nicht auf unterschiedliche Artikulationsgeschwindigkeiten der Probanden hin attribuiert werden. Um zwischen einem Versuchsleitereffekt und dem Effekt der Computerapparatur zu unterscheiden, wurde nach einem 2 x 2 Design mit den beiden Faktoren „Versuchsleiter“ und „Rechner“ ein weiterer Proband untersucht. Beide Versuchsleiter führten das Artikulationszeitexperiment an beiden Experimentalrechnern durch. Das Resultat zeigt, daß die Zeitmessungen zwischen den Computern deutlich voneinander abweichen. Die Messungen der beiden Versuchsleiter an ein und demselben Computer unterscheiden sich nur minimal, unabhängig davon an welchem PC das Experiment durchgeführt wird.

Ein weiterer Vergleich zwischen der Computerzeitmessung und einer Stoppuhr zeigt, daß die Zeitmessungen des Laptops, auf dem die Probanden in Spanien getestet wurden, deutlich von den mit der Stoppuhr gemessenen Zeiten abweichen. Eine Systematik in der Abweichung wird nicht erkannt. Dieser computerbedingten Meßungenauigkeit wird durch den Ausschluß aller Zeitdaten, die am Laptop erhoben wurde, Rechnung getragen. Dies gilt auch für die Werte der Reproduktionszeit.

Die Forderungen nach *Durchführungsobjektivität der Reihenfolge* und *Replizierbarkeit* der Ergebnisse scheinen erfüllt zu sein. Auswertungs- und Interpretationsobjektivität sind ohnehin gegeben, so daß man annehmen kann, daß die realisierten Experimente wohl dem Kriterium der Objektivität gerecht werden.

Die exemplarisch durchgeführte *konvergente Validierung* belegt, daß die erhaltenen Werte relativ kongruent zu den in der Literatur referierten Ergebnisse sind. Nach Ausschluß der ungenauen oder fehlerbehafteten Zeitmessungen werden die Parameter adäquat erfaßt. Die Experimente scheinen der stichprobenhaften Überprüfung der Experimentgüte zu genügen, so daß die Aussagekraft der Daten nicht eingeschränkt werden muß.

In Tabelle 11 werden die erhaltenen gültigen Mittelwerte und Streuung von Artikulationszeit, Höhe der Gedächtnisspanne und Reproduktionszeit im Überblick dargestellt. Die Ergebnisse des Supraspannen-Experimentes in Abhängigkeit von der Darbietungszeit sind in Tabelle 12 aufgelistet.

Tabelle 11 Deskriptive Ergebnisse I: Mittelwert und Standardabweichung von Artikulationszeit, Höhe der Gedächtnisspanne und Reproduktionszeit der acht Experimentreihen

Experiment	Artikulationszeit	Gedächtnisspanne	Reproduktionszeit
1 (Stille, d, ZS)	303 [62.04]	6.15 [0.76]	527 [96.59]
2 (art.U, d, ZS)	290 [54.12]	5.16 [0.71]	709 [149.52]
3 (Stille, s, ZS)	356 [42.44]	5.76 [0.71]	727 [163.08]
4 (Stille, c, ZS)	244 [63.17]	8.37 [1.04]	328 [53.68]
5 (art.U, c, ZS)	269 [65.46]	7.22 [1.23]	430 [119.01]
6 (Stille, d, Pae)	412 [72.90]	4.42 [0.52]	727 [148.09]
7 (Stille, d, Pu)	353 [62.04]	5.29 [0.69]	635 [98.79]
8 (Stille, d, ZW)	328 [36.55]	5.82 [0.84]	498 [70.14]

Tabelle 12 Deskriptive Ergebnisse II: Mittelwerte der Supraspannen-Experimente der acht Experimentreihen

Experiment	Darbietungszeit		
	10 Sekunden	20 Sekunden	30 Sekunden
1 (Stille, d, ZS)	4.45	2.41	2.34
2 (art.U, d, ZS)	7.32	4.00	3.47
3 (Stille, s, ZS)	6.42	3.68	2.53
4 (Stille, c, ZS)	3.37	1.79	0.63
5 (art.U, c, ZS)	3.42	1.95	1.21
6 (Stille, d, Pæ)	5.32	3.42	3.21
7 (Stille, d, Pu)	3.95	3.37	1.95
8 (Stille, d, ZW)	3.63	1.89	1.21

6.2 ÜBERPRÜFUNG DER MODIFIZIERTEN TOTAL-TIME-HYPOTHESE (MTTH)

Folgende Analysen beruhen auf den Daten des Supraspannen-Experiments. Die Überprüfung der modifizierten Total-Time-Hypothese bezieht sich auf den Nachweis der hyperbolischen Lernfunktion. Angenommen wird, daß die verarbeitete Information mit der Darbietungszeit hyperbolisch statt linear (wie in der TTH vermutet) steigt. Nach der mTTH wird zudem postuliert, daß unabhängig vom Informationsgehalt pro Materialart eine einzige Funktion bestimmt werden kann. Da sich diese Annahme in den bisher durchgeführten experimentellen Studien (Kapitel 3.3) bewährte, wird in dieser Arbeit auf die Variation des Informationsgehaltes verzichtet, so daß die Hypothese einer materialspezifischen Funktion nicht Gegenstand dieser Arbeit ist.

Gepprüft wird, ob die empirischen Werte für die Anzahl der Lernversuche pro Darbietungszeit signifikant von den nach der mTTH prognostizierten Werten abweichen. Zur Ermittlung der materialspezifischen Parameter a und b (Formel 4) wird eine lineare *Regression* durchgeführt. Als unabhängige Variable fungieren die Inversen der Darbietungszeit (t_j^{-1}) und als abhängiger Faktor werden die empirischen – über alle Probanden pro Darbietungszeit gemittelten - Werte der Anzahl der Lernversuche pro Darbietungszeit definiert. Aus den Parametern a , b und der Darbietungszeit wird durch Formel 13 die nach der mTTH erwarteten Werte (Y'_j) ermittelt.

Formel 13 $Y'_j = a + b * t_j^{-1}$

Realisiert wird der Vergleich von erhaltenen und erwarteten Werten durch einen *F-Test*. Dabei gehen im Zähler zwei Freiheitsgrade durch die Schätzung der beiden Parameter „additive Konstante“ und „Steigungsparameter“ verloren. Im Nenner geht in jeder der drei Gruppen ein Freiheitsgrad verloren.

Experimentreihe 1 umfaßt 29 Fälle pro Darbietungszeit, bei den anderen Experimenten 19 Datenpunkte. Bei einem alpha-Niveau von .05 und einer angenommenen mittleren Effektstärke ($f^2 = 0.15$) ergibt sich in Experiment 1 eine Power von .95. Bei den anderen sieben Experimenten dagegen steigen sowohl Fehler erster Art als auch Fehler zweiter Art auf je .10, basierend auf einer mittleren Effektstärke (Erdfelder et al., 1996). Diese Berechnungen gelten für alle folgenden Unterkapitel und werden an den entsprechenden Stellen nicht noch einmal wiederholt.

Wird keine systematische Differenz zwischen den empirischen und den erwarteten Werten aufgedeckt, so kann eine hyperbolische Funktion angenommen werden. Die Nullhypothese entspricht der erwarteten Hypothese. Jede Analyse wird pro Experiment durchgeführt, ein Überblick wird am Ende des Kapitel (Kapitel 6.2.9) präsentiert. Der Aufbau der Unterkapitel variiert nicht: Zuerst wird über die Ergebnisse der Parameterschätzung von a und b berichtet, dann folgt die Prognose von Y_j durch die Parameter und ein Vergleich dieser Prognosen mit den tatsächlichen Werten per *F-Test*.

6.2.1 EXPERIMENT 1 (STILLE, DEUTSCH, ZAHLEN, SYMBOL)

Für Experiment 1 war ursprünglich analog zu den anderen Experimentreihen die Teilnahme von insgesamt 57 Versuchspersonen geplant. Nach Erhebung einer Stichprobe dieser Größe wurde über die empirischen Mittelwerte der Anzahl der Lernversuche und die Inversen der Darbietungszeit eine *lineare Regression* durchgeführt. Für die beiden Parameter ergeben sich folgende Schätzungen: $a = 0.552$; $b = 39.508$. Deskriptiv deutet sich an dieser Stelle ein Problem an: a ist kleiner als eins. Wenn a gleich eins ist, bedeutet dies, daß der insgesamt zu verarbeitende Informationsgehalt dem Umfang der Kurzzeitgedächtnisspanne entspricht. Alle weiteren Interpretationen werden fraglich, wenn man davon ausgehen muß, daß in diesem Experiment keine Langzeitgedächtnisleistungen erforderlich sind. Bei einem Stichprobenumfang von 57 Probanden, also 19 Daten pro Darbietungszeit, liegen alpha und beta-Fehler bei mittlerer Effektstärke ($f^2 = 0.15$) bei je .10. Um als Erklärungsmöglichkeit für den niedrigen Parameter a die beiden Fehlerarten auszuschließen, werden weitere 30 Probanden getestet. Die erhaltenen Mittelwerte pro Darbietungszeit sind für die Gesamtstichprobe in Tabelle 12 aufgelistet.

Die *lineare Regression* über die Inversen der Darbietungszeit und die gemittelte Anzahl der Lernversuche pro Darbietungszeit der erweiterten Stichprobe resultieren in den Parameterschätzungen $a = 1.003$ und $b = 33.76$, $R^2 = .96$. Parameter a erreicht knapp das theoretisch fundierte Kriterium, daß a größer als 1 sein muß. Eine Diskussion dieses Befunds folgt in Kapitel 8.4.

Zur Überprüfung der hyperbolischen Funktion werden nach Formel 13 für den Stichprobenumfang von 87 Teilnehmern die Prognosen der Mittelwerte pro Darbietungszeit errechnet. Diese Werte aller Experimentreihen sind in Tabelle 13 dargestellt.

Tabelle 13 Anzahl der erwarteten Lernversuche in allen Experimenten

Experiment	Darbietungszeit		
	10 Sekunden	20 Sekunden	30 Sekunden
1 (Stille, d, ZS)	4.38	2.69	2.13
2 (art.U, d, ZS)	7.25	4.27	3.27
3 (Stille, s, ZS)	6.45	3.57	2.61
4 (Stille, c, ZS)	3.44	1.50	0.85
5 (art.U, c, ZS)	3.45	1.83	1.30
6 (Stille, d, Pæ)	5.27	3.62	3.06
7 (Stille, d, Pu)	4.10	2.81	2.38
8 (Stille, d, ZW)	3.64	1.84	1.24

Ein *F-Test* vergleicht die empirischen Daten von Experiment 1 (Tabelle 12) mit den prognostizierten Werten. Die empirische Abweichungsquadratvarianz zwischen erhaltenen und erwarteten Werten beträgt $s_{\text{emp}}^2 = 3.77$. Die Fehlervarianz innerhalb der Darbietungszeiten beträgt $s_e^2 = 4.72$. Der empirische F-Wert als Quotient dieser beiden Varianzen liegt mit einem Wert von $F_{(1,84)} = 0.798$, n.s., unter dem kritischen F-Wert.

6.2.2 EXPERIMENT 2 (ART. UNTERDRÜCKUNG, DEUTSCH, ZAHLEN, SYMBOL)

Analog dem Vorgehen unter Kapitel 6.2.1 werden für die beiden Parameter per *Regression* folgende Realisationen geschätzt: $a = 1.28$, $b = 59.75$. Das R-Quadrat der Anpassungsgüte liegt bei .99. Für die Fehlervarianz wird $s_e^2 = 10.20$ ermittelt, die Abweichungsvarianz liegt bei $s_{\text{emp}}^2 = 2.28$. Der *F-Wert* zeigt, daß die empirischen Daten (Tabelle 12) nicht signifikant von den prognostizierten Werten (Formel 13 sowie Tabelle 13) abweichen: $F_{(1,53)} = 0.224$, n.s..

6.2.3 EXPERIMENT 3 (STILLE, SPANISCH, ZAHLEN, SYMBOL)

Durch *lineare Regression* werden für die beiden materialspezifischen Konstanten folgende Werte geschätzt: $a = 0.694$, $b = 57.53$, $R^2 = .998$. Für die Fehlervarianz wird $s_e^2 = 6.40$ ermittelt und die Abweichungsvarianz errechnet sich mit $s_{\text{emp}}^2 = 0.38$. Der empirische *F-Wert* $F_{(1;53)}$ ist kleiner als 1. Die empirischen Daten (Tabelle 12) unterscheiden sich nicht signifikant von den nach Formel 13 prognostizierten Werten (Tabelle 13).

6.2.4 EXPERIMENT 4 (STILLE, CHINESISCH, ZAHLEN, SYMBOL)

Für das Experiment mit chinesischen Probanden, bei dem während des Lernvorgangs keine Zusatzfähigkeit gefordert ist, ergeben sich durch *lineare Regression* folgende Parameterschätzungen: $a = -0.448$ und $b = 38.91$ ($R^2 = .96$). Die mit diesen Koeffizienten nach Formel 13 prognostizierten Werte (Tabelle 13) weichen nicht systematisch von den empirischen Daten (Tabelle 12) ab, wie der *F-Test* zeigt ($s_e^2 = 2.78$, $s_{\text{emp}}^2 = 2.66$, $F_{(1;53)} = 0.987$, n.s.).

6.2.5 EXPERIMENT 5 (ART. UNTERDRÜCKUNG, CHINESISCH, ZAHLEN, SYMBOL)

In dem Supraspannen-Experiment 5 nahmen 57 Probanden mit chinesischer Muttersprache teil. Im Unterschied zu Experimentreihe 4 sollte jedoch während der Präsentationsphase der Lernitems artikulatorisch unterdrückt werden. Per *linearer Regression* werden die beiden Werte $a = 0.220$ und $b = 32.29$ geschätzt. Die Anpassungsgüte ist sehr gut ($R^2 = .99$). Der anschließende *F-Test* deckt keine signifikanten Unterschiede zwischen empirischen Daten und nach Formel 13 vorhergesagten Werten auf: $F_{(1;53)} = 0.121$, n.s., mit der Fehlervarianz: $s_e^2 = 3.13$ und der Abweichungsvarianz $s_{\text{emp}}^2 = 0.038$.

6.2.6 EXPERIMENT 6 (STILLE, DEUTSCH, PHONOLOG. ÄHNL. WORTE, SCHRIFT)

In Experiment 6 werden phonologisch ähnliche Worte präsentiert. Die per *Regression* ermittelten Werte für die Konstanten liegen bei $a = 1.96$, $b = 33.12$, $R^2 = .98$. Die Fehlervarianz entspricht $s_e^2 = 6.26$. Der Quotient aus Abweichungsvarianz ($s_{\text{emp}}^2 = 1.14$) und Fehlervarianz ergibt einen Wert von $F_{(1;53)} = 0.182$ im *F-Test*. Der errechnete F-Wert ist niedriger als der kritische F-Wert. Die Nullhypothese wird beibehalten.

6.2.7 EXPERIMENT 7 (STILLE, DEUTSCH, PHONOLOG. UNÄHNL. WORTE, SCHRIFT)

In Experiment 7 werden als Stimulusmaterial phonologisch unähnliche Worte verwendet. Die nach der *linearen Regression* errechneten Schätzungen für a und b entsprechen $a = 1.52$, respektive $b = 25.75$ bei einem quadrierten R von 0.75. Die Berechnung *des F-Wertes* ($s_e^2 = 6.60$, $s_{emp}^2 = 9.88$) resultiert in der Gültigkeit der Nullhypothese. Der empirische F-Wert $F_{(1;53)} = 1.50$, n.s., ist kleiner als der theoretische Wert. Die empirischen Daten (Tabelle 12) und die Prognosen (Tabelle 13) unterscheiden sich nur zufällig.

6.2.8 EXPERIMENT 8 (STILLE, DEUTSCH, ZAHLEN, SCHRIFT)

Experiment 8 verknüpft die beiden vorangegangenen Stimulusarten: Die Ziffern werden nicht als Symbole präsentiert, sondern als ausgeschriebene Worte. Die *lineare Regression* ergibt für Parameter a einen geschätzten Wert von $a = 0.0462$ und für Parameter b einen Wert von $b = 35.95$. Die Güte der Anpassung ist mit $R^2 = 0.999$ ausgezeichnet. Im *F-Test* mit einer Fehlervarianz von 3.21 und einer Abweichungsvarianz von 0.0665 ist der empirische F-Wert ($F_{(1;53)} = 0.0207$, n.s.) deutlich kleiner als die kritische Grenze. Die Abweichungen der erhaltenen von den prognostizierten Werten (Formel 13, Tabelle 13) ist nicht systematisch.

6.2.9 ÜBERBLICK ÜBER DIE ERGEBNISSE

Die Auswertung bezüglich der Gültigkeit der mTTH sind eindeutig: In allen Experimentreihen kann die mTTH beibehalten werden. Allerdings deuten sich hier spezifische Probleme in der Ausprägung des Parameters a bei Zahlenmaterial an, die in Kapitel 8.4.1 diskutiert werden. Tabelle 14 gibt im Überblick die Parameter und Ergebnisse der F-Tests an.

Tabelle 14 Zusammenfassung der Auswertung zur mTTH

Experiment	Parameter a	Parameter b	F-Wert
1 (Stille, d, ZS)	1.00	33.76	$F_{(1,84)} = 0.798$
2 (art.U, d, ZS)	1.28	59.75	$F_{(1,54)} = 0.224$
3 (Stille, s, ZS)	0.694	57.53	$F_{(1,54)} = 0.0594$
4 (Stille, c, ZS)	-0.448	38.91	$F_{(1,54)} = 0.958$
5 (art.U, c, ZS)	0.220	32.29	$F_{(1,54)} = 0.121$
6 (Stille, d, Pae)	1.96	33.12	$F_{(1,54)} = 0.182$
7 (Stille, d, Pu)	1.52	25.75	$F_{(1,54)} = 1.50$
8 (Stille, d, ZW)	0.0462	35.95	$F_{(1,54)} = 0.0207$

Innerhalb der mTTH wird postuliert, daß die im Langzeitgedächtnis verarbeitete Informationsmenge mit der Darbietungszeit nicht linear ansteigt, sondern nach der hyperbolischen Funktion. Nachdem im vorangegangenen Kapitel bereits die entsprechenden Analysen zur Gültigkeit der mTTH durchgeführt wurden, folgt nun erstens eine Berechnung der Höhe verarbeiteter Information in Abhängigkeit der Darbietungszeit (Tabelle 15) und zweitens graphische Darstellungen dieser Ergebnisse, die für die Diskussion der Daten (Kapitel 8.4) relevant sind.

Die Werte der Höhe verarbeiteter Information pro Darbietungszeit ergeben sich aus Formel 14. Diese Gleichung ist äquivalent zu Formel 5 mit dem Unterschied, daß hier Nenner und Zähler jeweils durch a dividiert wurden.

Formel 14 $r_j = \frac{I * a^{-1}}{1 + b * a^{-1} * t_j^{-1}}$

Tabelle 15 Deskriptive Werte der Langzeitgedächtnisspanne

Experiment	Darbietungszeit		
	10 Sekunden	20 Sekunden	30 Sekunden
1 (Stille, d, ZS)	10.28	16.72	21.14
2 (art.U, d, ZS)	6.20	10.55	13.76
3 (Stille, s, ZS)	6.98	12.60	17.23
4 (Stille, c, ZS)	13.07	30.05	53.01
5 (art.U, c, ZS)	13.05	24.53	34.72
6 (Stille, d, Pæ)	6.30	9.19	10.84
7 (Stille, d, Pu)	8.11	11.83	13.97
8 (Stille, d, ZW)	8.24	16.27	24.10

Abbildung 13 gibt graphisch die verarbeitete Informationsmenge pro Präsentationszeit separat für die Experimente wieder, die als Stimulusmaterial Zahlensymbole verwenden. Die beiden Experimente mit chinesischen Probanden (Experimente 4 und 5) heben sich deutlich von den anderen Untersuchungen ab. Die graphische Darstellung zeigt, daß das Anwachsen verarbeiteter Information nahezu linear verläuft, bei Experiment 4 ergibt sich eine positive Beschleunigung der Kurve. Bei den Experimenten mit deutschen (Experiment 1 & 2) und spanischen (Experiment 3) Versuchsteilnehmern flacht der Anstieg bei steigender Darbietungszeit ab, dennoch zeigt sich auch hier ein linearer Trend. Die Annäherung an eine Asymptote ist bei den realisierten Darbietungszeiten kaum zu erkennen.

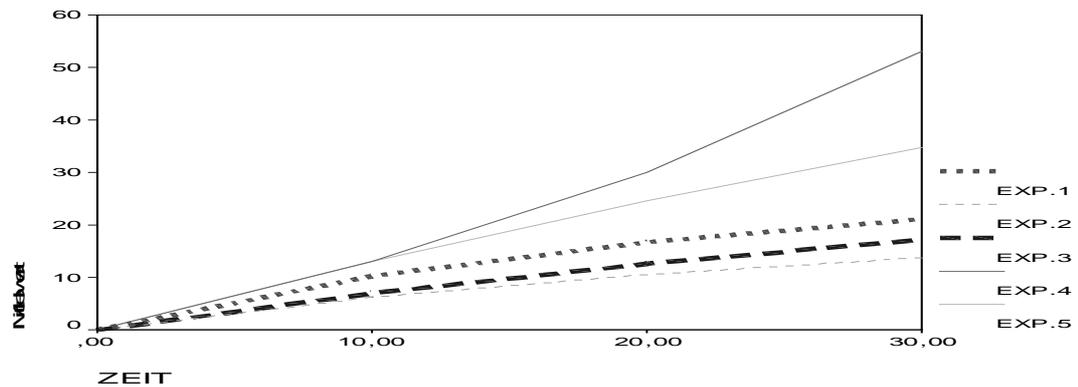


Abbildung 13 Graphischer Verlauf verarbeiteter Information in bit (y-Achse: Mittelwert) in Abhängigkeit von der Darbietungszeit (x-Achse: Zeit) für die Experimente mit Zahlensymbolen

Experiment 1 steht für deutsche Probanden, die unter Stillebedingungen das Supraspannen-Experiment ausführten. Experiment 2 bezeichnet dieselbe Kohorte unter der Testbedingung zusätzlicher Artikulatorischer Unterdrückung. Experiment 3 und 4 markieren eine Experimentdurchführung ohne Zusatzaufgabe für Probanden mit spanischer (Experiment 3), respektive chinesischer Muttersprache (Experiment 4). Als Experiment 5 ist die Untersuchung von chinesischen Versuchspersonen mit Artikulatorischer Unterdrückung definiert.

Abbildung 14 stellt die verarbeitete Information pro Darbietungszeit in den drei Experimenten mit verschiedenen Wortmaterialien dar. Für Experiment 8, in dem Zahlworte als Stimuli verwendet werden, steigt die mit der Zeit verarbeitete Information nahezu linear an. Bei den beiden anderen Experimenten mit phonologisch ähnlichen (Experiment 6), respektive unähnlichen Worten (Experiment 7) dagegen zeigt sich die erwartete hyperbolische Funktion und die Annäherung an eine Asymptote.

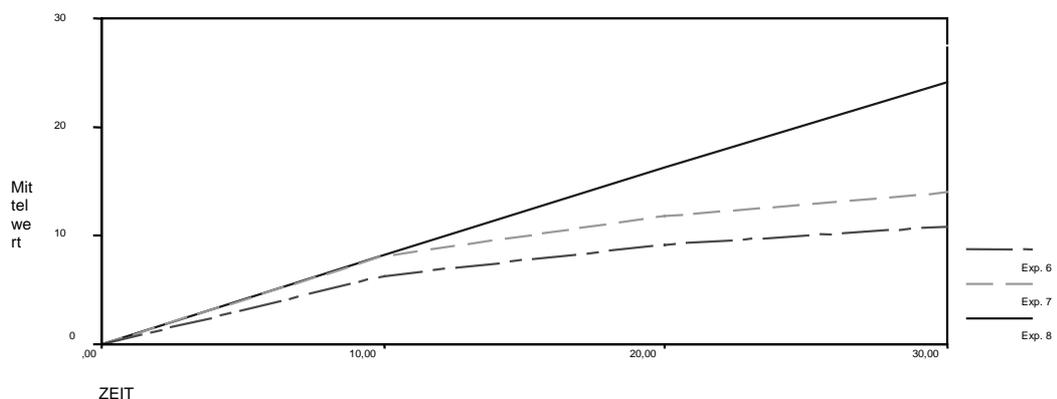


Abbildung 14 Graphischer Verlauf verarbeiteter Information in bit (y-Achse: Mittelwert) in Abhängigkeit von der Darbietungszeit (x-Achse: Zeit) für die Experimente mit Wortmaterialien

Nach Betrachtung der Schaubilder könnte man annehmen, daß in allen Experimenten mit Zahlenmaterial nicht die mTTH gilt, sondern die TTH. Um diese Annahme zu prüfen, wird für jedes Experiment einzeln aus dem Produkt von Darbietungszeit und Anzahl der Lernversuche eine als invariant postulierte, mittlere Lernzeit errechnet. Nach Umformung von Formel 1 wird die erwartete Anzahl der Lernversuche pro Darbietungszeit berechnet, die in Tabelle 16 eingetragen sind. Zum deskriptiven Vergleich ist in Klammern die quadrierte Abweichung der empirischen von den erwarteten Werten angefügt.

Tabelle 16: Prognose der Anzahl der Lernversuche im Supraspannen-Experiment nach der TTH, in Klammern ist die Abweichungsvarianz angegeben.

Experiment	Darbietungszeit		
	10 Sekunden	20 Sekunden	30 Sekunden
1 (Stille, d, ZS)	5.44 [0.97]	2.72 [0.10]	1.81 [0.28]
2 (art.U, d, ZS)	8.58 [1.59]	4.29 [0.08]	2.86 [0.37]
3 (Stille, s, ZS)	7.12 [0.49]	3.56 [0.01]	2.37 [0.02]
4 (Stille, c, ZS)	2.95 [0.18]	1.47 [0.10]	0.98 [0.12]
5 (art.U, c, ZS)	3.65 [0.05]	1.82 [0.02]	1.22 [0.00]
6 (Stille, d, Pæ)	7.26 [3.78]	3.63 [0.04]	2.42 [0.62]
7 (Stille, d, Pu)	5.51 [2.43]	2.75 [0.38]	1.84 [0.01]
8 (Stille, d, ZW)	3.68 [0.00]	1.84 [0.00]	1.23 [0.00]

Wendet man analog zu den Auswertungen der mTTH einen *F-Test* für jedes dieser Experimente an, so zeigen sich folgenden Ergebnisse. Bei den Experimenten 2 bis 8 liegen Fehler erster und zweiter Art bei je .10, für Experiment 1 liegen die beiden Fehlerarten bei .05 bei mittlerer Effektstärke (Erdfelder et al., 1996).

- In Experiment 1 (Stille, Deutsch, Zahlen, Symbol) weichen die nach der TTH prognostizierten Daten signifikant von den tatsächlich erhaltenen ab: $F_{(1,84)} = 8.29$.
- Ebenso kann die TTH nicht an die Daten von Experiment 2 (Art. Unterdrückung, Deutsch, Zahlen, Symbol) angepaßt werden: $F_{(1,54)} = 3.80$.
- Der kritische F-Wert von 2.80 bei Prognose der Daten durch die TTH wird in Experiment 3 (spanische Probanden) nicht überschritten: $F_{(1,54)} = 1.57$.
- Die Prognosen nach der TTH schätzen die empirischen Daten von Experiment 4 (Stille, Chinesisch) treffend: $F_{(1,54)} = 2.73$.
- Die TTH gilt auch in Experiment 5 (Art. Unterdrückung, Chinesisch): $F_{(1,54)} = 0.42$.

- In Experiment 6 (Stille, Deutsch, Phonolog. ähnliche Worte, Schrift) muß die TTH abgelehnt werden: $F_{(1,54)} = 13.48$.
- Die TTH muß für Experiment 7 (Stille, Deutsch, Phonolog. unähnliche Worte, Schrift) ebenfalls zurückgewiesen werden: $F_{(1,54)} = 8.12$.
- Angenommen wird die TTH in Experiment 8: $F_{(1,54)} = 0.059$.

In den vier Experimenten (Experiment 1, 2 6 und 7), in denen die TTH zurückgewiesen werden muß, zeigt sich, daß bei kurzer Leistung die Darbietungszeit unterschätzt wird. Dies steht in Kongruenz zu den Analysen von Bredenkamp (1975).

6.3 PRÜFUNG DER ÄQUIVALENZHYPOTHESE DER MTTT

Eine fundamentale Annahme zur Interpretation der mTTT ist die Gleichsetzung der Asymptote der Lernfunktion mit der Kurzzeitgedächtnisspanne in bit (Kapitel 3, Formel 6). Die Überprüfung ist nur sinnvoll für die Experimente 1, 2, 6 und 7. In den Experimenten, in denen die Gültigkeit der TTH angenommen wurde, erübrigt sich eine Auswertung zur Äquivalenzhypothese. In diesen Reihen (Experiment 3, 4, 5 und 8) wird eine alternative Analyse vorgenommen zur Prüfung, ob wegen des Versuchsdesigns nur ein Ausschnitt der hyperbolischen Funktion erfaßt wurde (Kapitel 8.4).

Das Procedere für die Experimentreihen, in denen die TTH verworfen wurde, ist folgendermaßen: Als Asymptote wird der analoge Wert der Höhe der Kurzzeitgedächtnisspanne eingesetzt. Im Falle, daß sich die Äquivalenzhypothese und die Hypothese von Baddeley, respektive Doshier und Ma bewähren, ist die Verknüpfung der Invarianzhypothesen gelungen. Da die Messung der Kurzzeitgedächtnisspannen in dem Subspannen-Experiment in der Einheit „Anzahl der Items“ erfolgt, muß vor der Gleichstellung mit der Asymptote dieser Wert in die informationstheoretische Einheit bit umgerechnet werden. Dazu wird die mittlere Anzahl erinnerter Items pro Experimentreihe mit dem Informationsgehalt pro Item multipliziert. Dieser Wert der Kurzzeitgedächtnisspanne in der informationstheoretischen Einheit (r_{\max}) wird in Formel 15 eingesetzt. Parameter b wird aus den Berechnungen der *linearen Regression* (Kapitel 6.2) übernommen. Formel 15 (entspricht Formel 10) gibt die Operation wieder, wie die im Supraspannen-Experiment erhaltenen Werte (Tabelle 12) durch die Implementation der Kurzzeitgedächtnisspannen aus dem Subspannen-Experiment vorhergesagt werden können. Die geschätzten Werte (Y'_{ij}) werden dann mit den empirischen Werten (Y_{ij}) verglichen.

$$\text{Formel 15} \quad Y'_{j} = I * r_{\max}^{-1} + b * t_j^{-1}$$

Analog dem Vorgehen in Kapitel 6.4 werden die nach Formel 15 errechneten Werte für die erwartete Anzahl der Lernversuche mit den tatsächlich beobachteten Daten durch einen *F-Test* verglichen. Da durch Integration der Daten aus dem Subspannen-Experiment Parameter *b* aus den Werten des Supraspannen-Experimentes geschätzt wird, geht im Gegensatz zu den Auswertungen von Kapitel 6.2 nur ein Freiheitsgrad verloren. Dementsprechend ändert sich die Power. Bei $\alpha = .10$ und $f^2 = 0.15$ ist die Teststärke bei den Experimentreihen 2, 6 und 7 gleich .82. In der ersten Reihe (Stille, Deutsch, Zahlen, Symbol) liegt bei dem Alpha-Niveau von .05 und einer postulierten mittleren Effektstärke ($f^2 = 0.15$) der Fehler zweiter Art bei .10 (Erdfelder et al., 1996). Um Wiederholungen zu vermeiden werden diese Berechnungen nicht ein zweites Mal in den Unterkapiteln aufgeführt.

Eine Anwendung von Formel 15 ist bei den Experimenten, in denen auch die TTH gilt, nicht sinnvoll. Für diese Experimente werden sowohl Parameter *a* als auch Parameter *b* geschätzt. Parameter *a'* ergibt sich dabei aus dem Quotienten des Gesamtinformationsgehalts und der Kurzzeitgedächtnisspanne in bit aus dem Subspannen-Experiment. Parameter *b'* wird nach Formel 16, relativiert auf die Darbietungszeiten von zehn (t_1) und dreißig Sekunden (t_3), geschätzt. Y_1 und Y_3 stehen für die mittlere Anzahl von Lernversuchen je nach Darbietungszeit.

$$\text{Formel 16} \quad b' = \frac{t_1 * t_3 (Y_3 - Y_1)}{t_1 - t_3}$$

Mit den Werten von *a'* und *b'* werden die Werte für die Experimente 3, 4, 5 und 8 einzeln prognostiziert und per *F-Test* mit den tatsächlich erhaltenen Werten verglichen. Die Power der Tests liegt jeweils bei .90, der alpha-Fehler bei .10 unter Annahme einer mittleren Effektstärke (Erdfelder et al., 1996).

Der Aufbau der Unterkapitel wird durch folgenden Schritte gekennzeichnet:

- Angabe der Parameter für r_{\max} und *b*, respektive für *a'* und *b'*
- Berechnung der Schätzungen von Y'_{j}
- *F-Test* zum Vergleich zwischen den erwarteten und den empirischen Werten.

In Kapitel 6.3.9 werden die Ergebnisse zusammengefaßt.

6.3.1 EXPERIMENT 1 (STILLE, DEUTSCH, ZAHLEN, SYMBOL)

In Experiment 1 läßt sich durch *lineare Regression* über die Inversen der Darbietungszeit als unabhängige Variable und die Anzahl der benötigten Lernversuche im Supraspannen-Experiment als abhängige Variable ein Steigungsparameter von $b = 33.76$ schätzen. In dem Subspannen-Experiment ergibt sich eine über alle Teilnehmer gemittelte Höhe der Gedächtnisspanne $r_{\max} = 18.45$ bit. Setzt man diese beiden Werte in Formel 15 ein, so erhält man die aus dem Subspannen-Experiment erwarteten Werte für das Supraspannen-Experiment. Die Prognosen für alle acht Experimente sind in Tabelle 17 dargestellt.

Tabelle 17 Anzahl der durch die Kurzzeitgedächtnisspanne prognostizierten Lernversuche (Experiment 1, 2, 6 und 7), respektive durch die Kurzzeitgedächtnisspanne und b' erwartete Anzahl der Lernversuche (Experiment 3, 4, 5 und 8)

Experiment	Darbietungszeit		
	10 Sekunden	20 Sekunden	30 Sekunden
1 (Stille, d, ZS)	5.82	4.13	3.56
2 (art.U, d, ZS)	8.88	5.89	4.90
3 (Stille, s, ZS)	8.44	5.52	4.55
4 (Stille, c, ZS)	5.90	3.85	3.16
5 (art.U, c, ZS)	5.40	3.74	3.19
6 (Stille, d, Pæ)	5.57	3.92	3.37
7 (Stille, d, Pu)	4.47	3.18	2.75
8 (Stille, d, ZW)	5.35	3.54	2.93

Wiederum wird durch einen *F-Test* geprüft, ob Prognosen und empirische Werte signifikant voneinander abweichen. Die Abweichungsvarianz liegt bei $s_{\text{emp}}^2 = 91.50$, die Fehlervarianz bei $s_e^2 = 4.72$. Der empirische F-Wert mit zwei Freiheitsgraden im Zähler und 84 Freiheitsgraden im Nenner $F_{(2;84)} = 19.37$, $p < .05$, überschreitet den kritischen F-Wert. Die Unterschiede zwischen den erwarteten und erhaltenen Daten sind systematischer Natur. Dies widerspricht den in Kapitel 3 geschilderten Erwartungen, war aber nach Abbildung 13 zu erwarten.

6.3.2 EXPERIMENT 2 (ART. UNTERDRÜCKUNG, DEUTSCH, ZAHLEN, SYMBOL)

Aus der Auswertung von Kapitel 6.4 hat sich eine Höhe von $b = 59.75$ für Parameter b ergeben. Der Mittelwert für die Gedächtnisspanne aus dem Subspannen-Experiment liegt bei $r_{\max} = 15.48$ bit. Aus diesen beiden Koeffizienten wird die Anzahl der Versuche pro

Darbietungszeit nach Formel 15 geschätzt (Tabelle 17). Im *F-Test* überschreitet bei einer Abweichungsvarianz von 76.67 und einer Fehlervarianz von 10.20 der empirische F-Wert ($F_{(2;54)} = 7.52$, $p < .10$) den theoretischen F-Wert. Das bedeutet, daß die Anzahl der Lernversuche pro Darbietungszeit (Tabelle 12) nicht zutreffend vorhergesagt werden können. Wiederum kann die Kurzzeitgedächtnisspanne in bit nicht mit der Asymptote der Lernfunktion gleichgesetzt werden.

6.3.3 EXPERIMENT 3 (STILLE, SPANISCH, ZAHLEN, SYMBOL)

In dem Kurzzeitgedächtnis-Experiment mit spanischen Probanden wird ein Mittelwert der Gedächtnisspanne von $r_{\max} = 17.28$ bit ermittelt. Da die TTH für dieses Experiment zu zutreffenden Prognosen führt, werden sowohl a' als auch b' (Formel 16) geschätzt: $a' = 2.60$, $b' = 58.35$. Die Integration dieser Werte in Formel 15 prognostiziert die Daten, die in Tabelle 17 aufgeführt sind. Der *F-Test* ($s_e^2 = 6.40$ und $s_{abw}^2 = 219$) zeigt signifikante Unterschiede zwischen den erhaltenen und den prognostizierten Werten: $F_{(1;54)} = 34.14$, $p < .10$. Die Anzahl der Lernversuche pro Darbietungszeit (Tabelle 12) kann nicht aus der Kurzzeitgedächtnisleistung und dem geschätzten b-Parameter vorhergesagt werden.

6.3.4 EXPERIMENT 4 (STILLE, CHINESISCH, ZAHLEN, SYMBOL)

Durch die Ausführung einer *linearen Regression* ergeben sich für Parameter a und b die geschätzten Werte von $a' = 1.79$ und $b' = 41.13$ (Formel 16). Die Prognosen nach Formel 15 für Experiment 4 sind wiederum in Tabelle 17 eingetragen. Der *F-Test* deckt signifikante Unterschiede zu den empirischen Daten des Supraspannen-Experimentes auf: $F_{(1;54)} = 116.5$, $p < .10$, $s_e^2 = 2.78$, $s_{abw}^2 = 324$. Die Gleichsetzung der Asymptote mit der Kurzzeitgedächtnisspanne kann analog zu den deutschen und spanischen Probanden für das Stimulusmaterial „Zahlen“ auch bei chinesischen Probanden nicht beibehalten werden.

6.3.5 EXPERIMENT 5 (ART. UNTERDRÜCKUNG, CHINESISCH, ZAHLEN, SYMBOL)

Für das Experiment mit chinesischen Probanden und Sekundäraufgabe während der Präsentationsphase werden folgende Parameter ermittelt: $a' = 2.08$, $b' = 33.15$ (Formel 16). Die Höhe der Gedächtnisspanne im Subspannen-Experiment beträgt $r_{\max} = 21.66$ bit. Die nach Formel 15 prognostizierten Daten sind in Tabelle 17 enthalten. Der *F-Test* zum Vergleich dieser Prognosen mit den tatsächlich erhaltenen Daten (Tabelle 12) deckt systematische Abweichungen auf: $F_{(1;54)} = 66.77$, $p < .10$, mit einer Fehlervarianz von 3.13 und der Abweichungsvarianz von 210.

6.3.6 EXPERIMENT 6 (STILLE, DEUTSCH, PHONOLOG. ÄHNL. WORTE, SCHRIFT)

Nachdem in den Experimentreihen mit Zahlen die Äquivalenzhypothese nicht nachgewiesen werden kann, ist es von Interesse, die Konfundierung zwischen dem Präsentationsmodus als Symbol und der Lexikalität aufzuheben. Das in Experiment 6 und 7 realisierte Wortmaterial wird genau wie Silbenmaterial, für das die Gleichsetzung Bestand hatte (Bredenkamp & Klein, 1998), in Schriftform dargeboten und verweist dennoch auf lexikale Repräsentationen im Langzeitgedächtnisspeicher.

Aus der Auswertung von Kapitel 6.2 ergibt sich für Parameter b in Experiment 6 ein Wert von $b = 33.12$. Die Höhe der Gedächtnisspanne liegt bei $r_{\max} = 14.68$ bit. Die aus diesen beiden Werten nach Formel 15 errechnete Anzahl der Lernversuche bis zum Kriterium ist in Tabelle 17 dargestellt. Mit der Abweichungsvarianz von 3.23, respektive der Fehlervarianz von 6.26 ist in einem *F-Test* der empirische F -Wert ($F_{(2;54)} = 0.52$, n.s.) kleiner als die kritische Grenze. Die Anzahl der Lernversuche pro Darbietungszeit (Tabelle 12) kann aus der Kurzzeitgedächtnisleistung (Tabelle 17) vorhergesagt werden. Die Interpretation der Asymptote der mTTH als Kurzzeitgedächtnisspanne in bit kann in Experiment 6 beibehalten werden.

6.3.7 EXPERIMENT 7 (STILLE, DEUTSCH, PHONOLOG. UNÄHNL. WORTE, SCHRIFT)

In Experiment 7 werden ebenfalls lexikale Worte als Stimulusmaterial eingesetzt, die im Gegensatz zu Experiment 6 phonologisch unähnlich sind. Die Kurzzeitgedächtnisspanne in bit liegt bei $r_{\max} = 15.57$ bit, für Parameter b ergibt sich $b = 25.75$. Die Implementation der Kurzzeitgedächtnisspanne in bit als Asymptote (Tabelle 17) erbringt keine signifikante Abweichungen der Prognosen (Formel 15) von den empirischen Daten im *F-Test*. Der erhaltenen F -Wert $F_{(2;54)} = 1.35$, n.s., überschreitet nicht den kritischen Wert ($s_e^2 = 6.60$, $s_{\text{emp}}^2 = 8.93$). Die Interpretation der Asymptote der Lernfunktion als Kurzzeitgedächtnisspanne in bit kann beibehalten werden.

6.3.8 EXPERIMENT 8 (STILLE, DEUTSCH, ZAHLEN, SCHRIFT)

Die letzte Analyse zur Gleichsetzungshypothese prognostiziert die Mittelwerte pro Darbietungszeit für Experiment 8 (Tabelle 17), indem in Formel 15 folgende Werte eingesetzt werden: $a' = 1.72$, $b' = 36.32$ (Formel 16). Der *F-Test* deckt signifikante Unterschiede zwischen den empirischen und den erwarteten Werten auf: $F_{(1;54)} = 51.14$, $p < .10$, $s_{\text{emp}}^2 = 164$,

$s_e^2 = 3.21$. Wie schon bei den anderen Experimentreihen mit Zahlenmaterial gelingt die Anpassung nicht.

6.3.9 ÜBERBLICK ÜBER DIE ERGEBNISSE

Die Ergebnisse sind eindeutig: Wenn Ziffern als Stimulusmaterial dienen, ist die Äquivalenzhypothese nicht haltbar, unabhängig vom Präsentationsmodus oder dem Ausführen einer Sekundäraufgabe. Bei den Experimentreihen, in denen auch die TTH gültig ist, kann ebenfalls keine Asymptote angepaßt werden. Allein bei Vorlage von Wortmaterial können Gedächtnisspanne und Asymptote gleichgesetzt werden. Die in Kapitel 3 beschriebene Konfundierung zwischen symbolischer Darbietung und Lexikalität ist aufgelöst. Keiner der beiden Faktoren ist verantwortlich für das Scheitern der Äquivalenzhypothese. Lexikalität ist nicht die relevante Variable, weil auch in Experiment 6 und 7 lexikale Stimuli verwendet werden. Das Scheitern der Interpretation entsteht auch nicht durch die zweite vermutete Alternative: Der Präsentationsmodus verändert nicht das Ergebnismuster, daß bei Zahlenmaterial die Interpretation der Asymptote nicht beibehalten werden kann. In Kapitel 8.4 wird verschiedenen Erklärungsmöglichkeiten dieses Phänomens nachgegangen. Tabelle 18 werden die Ergebnisse kurz zusammen, wobei das Überschreiten der Signifikanzgrenze durch „*“ gekennzeichnet wird.

Tabelle 18 Zusammenfassung der intraexperimentellen Auswertung der mTTH für die Äquivalenzhypothese

Experiment	Gedächtnisspanne	Parameter b	F-Wert
1 (Stille, d, ZS)	18.45 bit	33.76	$F_{(2,84)} = 19.38 *$
2 (art.U, d, ZS)	15.48 bit	59.75	$F_{(2,54)} = 7.52 *$
3 (Stille, s, ZS)	17.28 bit	58.35	$F_{(1,54)} = 34.14*$
4 (Stille, c, ZS)	25.11 bit	41.13	$F_{(1,54)} = 116.5*$
5 (art.U, c, ZS)	21.66 bit	33.15	$F_{(1,54)} = 66.77*$
6 (Stille, d, Pæ)	14.68 bit	33.12	$F_{(2,54)} = 0.52$
7 (Stille, d, Pu)	15.57 bit	25.75	$F_{(2,54)} = 1.35$
8 (Stille, d, ZW)	17.46 bit	36.32	$F_{(1,54)} = 51.14*$

6.4 ERGEBNISSE DER INTRAEXPERIMENTELLEN AUSWERTUNG

Die intraexperimentelle Auswertung ist angelegt zur Prüfung der Gütekriterien, der Gültigkeit der mTTH und der Äquivalenzhypothese. Im folgenden sind kurz die wichtigsten Resultate und die sich daraus ergebenden Konsequenzen für die Auswertungen in Kapitel 7 skizziert.

- Aus der *Reliabilitätsanalyse* folgt, daß zur Erhöhung der Reliabilität Daten der Artikulations-, respektive Reproduktionszeit ausgeschlossen werden, deren Cronbach-Alpha unter .75 liegt.
- Für Experimentreihe 3 mit spanischen Probanden können für Analysen, in die Zeitmessungen eingehen, nur die in Deutschland erhobenen Daten berücksichtigt werden, da die *Durchführungsobjektivität* verletzt ist.
- Weitere stichprobenartig ausgewählte Analysen (*Reihenfolge-Effekt, prinzipielle Replizierbarkeit, Validität und Validierung*) sprechen dafür, daß die Experimentgüte gesichert zu sein scheint.
- Die *mTTH* ist in allen Experimenten gültig, so daß interexperimentelle Vergleiche ohne Einschränkung möglich sind. In vier der acht Experimente ist der Wert für Parameter a kleiner als das theoretisch geforderte Minimum von 1.
- Der *graphische Verlauf* der Zunahme an verarbeiteter Information mit der Darbietungszeit zeigt bei Wortmaterial die erwartete Annäherung an eine Asymptote, nicht jedoch für Zahlenmaterial. Die *TTH* ist in vier der acht Experimente ebenfalls gültig (Experimente 3, 4, 5 und 8).
- Die Prüfung der *Äquivalenzhypothese* resultiert in einem deutlichen Materialbias: für Wortmaterial können Kurzzeitgedächtnisspanne in bit und Asymptote der Lernfunktion gleichgesetzt werden. Bei allen Experimenten mit Zahlenmaterial dagegen ist die Äquivalenz nicht gegeben, so daß ein wichtiges Ziel dieser Arbeit nicht überprüft werden kann, nämlich die Frage, ob experimentell induzierte Veränderungen in der Kurzzeitgedächtnisspanne sich auch in der Höhe der Asymptote niederschlagen (Kapitel 7).

7 INTEREXPERIMENTELLE HYPOTHESENPRÜFUNG

Die interexperimentelle Auswertung der empirischen Ergebnisse fokussiert drei Themengebiete: als erstes werden die Vergleichbarkeit der Experimente (Kapitel 7.1) und die Wirkung der experimentellen Manipulation (Kapitel 7.2) untersucht. Als nächstes werden die Invarianzhypothesen des Kurzzeitgedächtnisses von Baddeley (Kapitel 7.3) sowie von Doshier und Ma (Kapitel 7.4) hinsichtlich ihres Anspruches auf Universalität geprüft. Den Schluß bildet die Überprüfung der mTTH. Analysiert werden differentielle Wirkungsweisen der manipulierten, unabhängigen Variablen auf die Parameter Asymptotenhöhe (Kapitel 7.5) und Verarbeitungszeit im Langzeitgedächtnisspeicher (Kapitel 7.6). Am Abschluß dieses Kapitels werden die Auswertungen zu den Invarianzhypothesen zusammengefaßt.

7.1 VORAUSSETZUNGEN INTEREXPERIMENTELLER VERGLEICHE

Für die interexperimentelle Vergleichbarkeit müssen vier Voraussetzungen notwendig erfüllt sein: Unabhängigkeit und Äquivalenz der Stichproben, ebenso Varianzhomogenität und Normalverteilung des Fehlers. Die Überprüfung dieser Voraussetzungen scheint in Anbetracht der ungleich großen Stichproben notwendig.

Unabhängigkeit der Stichproben. Dies wird durch die Versuchsplanung ohne Meßwiederholung zwischen den Experimenten gewährleistet.

Äquivalenz der Stichproben. Durch Restriktion der Stichprobe auf Studenten soll der Einfluß von Intelligenzunterschieden nivelliert werden (zum Einfluß von Intelligenz auf die Gedächtnisspanne: Klein, 1995). In Tabelle 19 werden – separat für jede Experimentreihe - die Stichprobenkennwerte durchschnittliches Alter, mittleres Studienjahr und Verhältnis der Geschlechter inklusive der Standardabweichung (in Klammern) dargestellt.

Tabelle 19 Stichprobenkennwerte: Mittelwert und Standardabweichung

Experiment	Alter	Studienjahr	Frauenanteil
1 (Stille, d, ZS)	24.5 [4.67]	3 [2.06]	57 %
2 (art.U, d, ZS)	24.5 [5.63]	2 [1.81]	74 %
3 (Stille, s, ZS)	25.4 [4.94]	4 [1.54]	75 %
4 (Stille, c, ZS)	19.4 [3.44]	3 [3.22]	63 %
5 (art.U, c, ZS)	19.4 [0.99]	2 [0.44]	65 %
6 (Stille, d, Pæ)	25.6 [6.58]	3 [2.88]	72 %
7 (Stille, d, Pu)	26.4 [6.78]	3 [2.77]	75 %
8 (Stille, d, ZW)	25.5 [3.52]	4 [2.44]	67 %

Während die Kennwerte bezüglich Studienjahr und Frauenanteil relativ homogen wirken, zeigt sich, daß ein Altersunterschied im Durchschnitt von sechs Jahren zwischen chinesischen Probanden und den anderen Teilnehmern der Studie besteht. Dieser Effekt ist statistisch signifikant: $F_{(7,478)} = 18.17$, $p < .05$ und mit einer geschätzten Effektstärke von $\eta^2 = 0.42$ als ein kleiner bis mittlerer Effekt zu interpretieren. Ein post-hoc Test zeigt, daß zwei homogene Untergruppen bestehen: einerseits die chinesischen Probanden und andererseits die deutschen und spanischen Probanden. Dieser Altersunterschied könnte die interne Validität gefährden und wird deswegen in den folgenden Analysen als Kovariate berücksichtigt.

Um die von Stigler et alii (1986) erwähnten differierenden kulturellen Unterschiede und Lernerfahrungen zu minimieren, wurde der Parameter der Lernjahre eingeführt. Deskriptiv sind zwar chinesische Probanden jünger als die deutsche Gruppe, die Unterschiede der Lernjahre sind aufgrund verschiedener Bildungssysteme dagegen vergleichbar.

Da in der Literatur auf Geschlechtsunterschiede in Artikulationszeit, respektive Gedächtnisleistungen hingewiesen wird (Garai & Scheinfeld, 1968), wird die Geschlechterverteilung zwischen den Experimentreihen überprüft: Auch hier scheinen sich kaum Unterschiede zu ergeben. In allen Experimentreihen ist eine relativ deutlich ausgeprägte Mehrheit von Teilnehmerinnen zu konstatieren.

Varianzhomogenität. Zur Prüfung dieser Forderung wird für jede abhängige Variable einzeln ein *Levene-F-Test* durchgeführt. Mit sechs Zählerfreiheitsgraden und 355 Nennerfreiheitsgraden wird nach der Levene-Statistik ($F_{(6,355)} = .997$) auf einem Alpha-Niveau von .05 keine Signifikanz für die abhängige Variable „Artikulationszeit“ erreicht. Die Erfüllung der Varianzhomogenität aller Experimente für diesen Parameter kann angenommen werden.

Derselbe Test erbringt für den abhängigen Faktor „Gedächtnisspanne“ bedeutsame Unterschiede. Dieses Ergebnis deutet sich deskriptiv bereits in Tabelle 11 an: Die Varianz der Gedächtnisspannendaten chinesischer Probanden ist wesentlich höher als die der anderen Experimentreihen. Schließt man Experiment 4 und 5 aus der Analyse aus, so erhält man Varianzhomogenität für die Variable „Gedächtnisspanne“: $F_{(5,329)} = 1.365$, n.s.

Die dritte abhängige Variable zum Test auf Varianzhomogenität ist die Reproduktionszeit. Ein Vergleich über alle Experimentreihen deckt signifikante Unterschiede in den Varianzen auf. Deskriptiv fallen dabei besonders vier Experimente mit großer Standardabweichung auf (Tabelle 11): Die beiden Experimente mit Artikulatorischer

Unterdrückung, das Experiment mit phonologisch ähnlichen Items sowie das Experiment mit spanischen Probanden haben im Gegensatz zu den vier anderen Experimenten eine sehr viel höhere Varianz.

Normalverteilung des Fehlers. Der *Kolmogorov-Smirnov-Test* zeigt für kein Experiment in keiner der abhängigen Variablen Artikulationszeit, Gedächtnisspanne oder Reproduktionszeit eine signifikante Abweichung von der Normalverteilungsannahme. Tabelle 20 zeigt den *Kolmogorov-Smirnov-Z-Wert* für alle drei Variablen in den acht Experimenten.

Tabelle 20 Ergebnisse der *Kolmogorov-Smirnov-Tests*

Experiment	<i>Kolmogorov-Smirnov-Z-Wert</i>		
	Artikulationszeit	Gedächtnisspanne	Reproduktionszeit
1 (Stille, d, ZS)	1.19	1.31	0.732
2 (art.U, d, ZS)	0.696	0.685	0.697
3 (Stille, s, ZS)	0.622	0.431	0.382
4 (Stille, c, ZS)	0.880	0.621	0.512
5 (art.U, c, ZS)	0.434	1.11	0.887
6 (Stille, d, Pæ)	0.510	0.964	0.973
7 (Stille, d, Pu)	0.820	0.510	0.814
8 (Stille, d, ZW)	0.729	0.550	0.982

Die Prüfung der Voraussetzungen interexperimenteller Vergleiche zeigt, daß einige Vergleiche nur unter Berücksichtigung von Kovariaten durchgeführt werden dürfen, respektive daß nicht in allen Fällen Varianzanalysen angewendet werden können.

- Bezüglich des Spracheffekts muß der *Altersunterschied* zwischen den Probanden als Kovariate in die weiteren Analysen einbezogen werden, weil die Stichprobenäquivalenz nicht gesichert ist.
- Da die *Varianzhomogenität* des Parameters „Gedächtnisspanne“ für Vergleiche mit den Experimenten 4 und 5 nicht gegeben ist, darf nur ein *t-Test* mit Berücksichtigung ungleicher Varianzen durchgeführt werden.
- Ebenfalls wegen *mangelnder Homogenität* der Varianzen für die Variable „Reproduktionszeit“ muß bei jedem spezifischen Vergleich nochmals ein *Levene-F-Test* durchgeführt werden. Falls Unterschiede bestehen, wird der *Welch-t-Test* angewendet.

Alle anderen Überprüfungen genügen den Forderungen und gefährden die Aussagekraft interexperimenteller Vergleiche nicht.

7.2 ÜBERPRÜFUNG DER ARBEITSHYPOTHESEN

In folgendem Kapitel soll das Vorliegen der nach dem Arbeitsgedächtnismodell prognostizierten Effekte überprüft werden zur Kontrolle der Wirksamkeit experimenteller Manipulation. Insgesamt werden die bereits in Kapitel 4.1 beschriebenen Phänomene untersucht:

- Mit Spracheffekt wird der Einfluß der Muttersprache auf die abhängigen Variablen beschrieben. Realisiert werden in dieser Studie drei verschiedene Sprachen: deutsch (Experiment 1), spanisch (Experiment 3) und chinesisch (Experiment 4).
- Unter Effekt Artikulatorischer Unterdrückung wird die Wirksamkeit dieser Zusatzfähigkeit überprüft, getrennt für deutsche und chinesische Probanden (Experiment 1 versus 2, Experiment 4 versus 5).
- Eliminierung des Spracheffekts durch Artikulatorische Unterdrückung beschreibt das Phänomen, daß die spezifischen durch die Muttersprache (deutsch versus chinesisch) evozierten Effekte (so sie in den vorangegangenen Analysen nachgewiesen werden können) durch die Zusatzaufgabe eliminiert werden (Experiment 2 versus 5).
- Zwischen Männern und Frauen sollen Unterschiede in der Merkleistung und in der Sprechgeschwindigkeit bestehen. Dies wird unter dem Begriff „Geschlechtseffekt“ untersucht (Experiment 1 & 4).
- Analog dem dritten Effekt wird eine Eliminierung des Geschlechtseffekt durch Artikulatorische Unterdrückung erwartet (Experiment 2 & 5).
- Der Materialeffekt wird in zwei verschiedene Vergleiche unterteilt. Zuerst wird untersucht, ob systematische Unterschiede in den abhängigen Variablen bestehen, je nachdem ob phonologisch unähnliches, respektive ähnliches Stimulusmaterial verwendet wird (Experiment 6 versus 7). Der zweite Paarvergleich bezieht sich auf die Materialklassen „Worte“ und „Zahlen“. Da bei der Materialkonstruktion auf maximale phonologische Distinktheit der Zahlenstimuli geachtet wurde, wird aus der Klasse „Worte“ nur die Realisation mit phonologisch unähnlichen Worten berücksichtigt. Aus der Klasse der „Zahlen“ wird das Experiment mit Zahlworten ausgewählt, damit der

Präsentationsmodus zwischen den Materialklassen nicht verändert wird (Experiment 7 versus 8).

- Zum Schluß wird auf den Effekt des Präsentationsmodus eingegangen. Damit ist die unterschiedliche Darbietung von Zahlenmaterial entweder als Ziffer oder als Wort gemeint (Experiment 1 versus 8).

Das Procedere der Überprüfung der Effekte ist in allen folgenden Kapiteln gleich. Der Einfluß der unabhängigen Variablen, die durch obenstehende Erläuterungen definiert sind, wird einzeln auf jede der drei abhängigen Variablen „Artikulationszeit“, „Höhe der Gedächtnisspanne“ und „Reproduktionszeit“ geprüft. Als Restriktion wird die Erfüllung der Voraussetzungen der *Varianzanalyse* (Kapitel 7.1) und für die Zeitmessung das Überschreiten der Reliabilitätsgrenze von .75 eingeführt. Falls sich signifikante Unterschiede ergeben, wird ein *Scheffé-Test* zum post-hoc Paarvergleich angewendet, wenn keine spezifischen Hypothesen vorliegen. Falls aus den bisherigen Forschungsergebnissen gerichtete Hypothesen abgeleitet werden können, wird ein einseitiger *t-Test* durchgeführt. Sollte die Varianzhomogenität nicht erfüllt sein, so wird der *Welch-t-Test* angewendet.

In den Kapitel 7.2.1 bis 7.2.6 wird untersucht, ob die oben aufgeführten Phänomene in diesen Experimenten nachgewiesen werden können. In Kapitel 7.2.7 folgt ein zusammenfassender Überblick über die Auswertungen.

7.2.1 DER SPRACHEFFEKT

Der Spracheffekt bezieht sich auf die Überprüfung von Unterschieden zwischen den Experimentreihen 1 (deutsche Probanden), 3 (spanische Probanden) und 4 (chinesische Probanden). In bezug auf die Faktoren „Artikulationszeit“ und „Gedächtnisspanne“ können aus der Literatur gerichtete Hypothesen abgeleitet werden (Kapitel 4.1.2). Dagegen ist das Vorgehen hinsichtlich der Variable „Reproduktionszeit“ explorativ. Die signifikanten Unterschiede bezüglich des Alters zwischen den Gruppen werden in die Analysen einbezogen. In Auswertungen, in die Zeitmessungen eingehen, wird das Reliabilitätskriterium berücksichtigt.

Zur Prüfung von Unterschieden in der Artikulationszeit wird eine *Varianzanalyse* mit der dreifach gestuften unabhängigen Variablen „Sprache“ und der Kovariate „Alter“ durchgeführt. Die Kovariate zeigt keinen signifikanten Haupteffekt ($F_{(1,139)} = 0.786$, n.s.) und

keine Interaktion mit der Muttersprache, die als Haupteffekt signifikant wird: $F_{(2,139)} = 18.12$, $s_e^2 = 3748$, $p < .05$, $\eta^2 = 0.21$.

Der post-hoc *Scheffé-Test* belegt signifikant unterschiedliche Artikulationsgeschwindigkeiten für das Deutsche und das Spanische. Ebenso bedeutsame Abweichungen zeigt der gerichtete *t-Test* für die Vergleiche Chinesisch versus Deutsch ($t_{(131)} = 5.07$, $p < .05$) und Chinesisch versus Spanisch ($t_{(57)} = 5.32$, $p < .05$). Der Trend des Unterschieds entspricht den aus der Literatur abgeleiteten Hypothesen: mit 244 Millisekunden Artikulationszeit pro Item wird im Chinesischen die schnellste Geschwindigkeit verzeichnet. Für die deutschen Probanden wird ein Wert von 303 Millisekunden ermittelt, bei der spanischen Gruppe ergibt sich eine Geschwindigkeit von 356 Millisekunden.

Analog wird für die Prüfung sprachlich bedingter Unterschiede in der Höhe der Gedächtnisspanne vorgegangen. Da keine Varianzhomogenität von Experiment 1 und 3 zu der chinesischen Kohorte besteht, werden Einzelvergleiche durchgeführt. Der Trend zeigt erwartungsgemäß für das Chinesische die höchsten Werte (8.37 Items versus 6.15 Items bei deutschen Probanden und 5.76 Items bei spanischen Versuchspersonen). Der *Welch-t-Test* zum Vergleich von Experiment 1 und Experiment 4 deckt auf, daß der Unterschied in der Merkleistung zwischen deutschen und chinesischen Probanden von 2.22 Items statistisch signifikant ist ($t_{(94)} = -13.81$, $p < .05$). Die geschätzte Effektgröße von $\eta^2 = 0.60$ läßt nach den Konventionen einen großen Effekt vermuten. Ebenso ist die Differenz von 2.61 Items zwischen chinesischen und spanischen Probanden in einem *Welch-t-Test* mit Annahme ungleicher Varianzen signifikant ($t_{(98.81)} = -15.55$, $p < .05$). Die geschätzte Effektstärke ist mit $\eta^2 = 0.68$ erwartungsgemäß höher als beim Vergleich deutscher und chinesischer Probanden

Die *Varianzanalyse* für die abhängige Variable „Gedächtnisspanne“ im Vergleich deutscher und spanischer Probanden deckt signifikante Unterschiede auf $F_{(1,142)} = 9.49$, $s_e^2 = 0.547$, $p < .05$, wenn auch die geschätzte Effektstärke mit $\eta^2 = 0.06$ gering ist.

Die Varianzen der Reproduktionszeiten in den drei Experimentreihen verteilen sich nicht gleich. Aus diesem Grund werden drei *Welch-t-Tests* durchgeführt. Bei 97.99 Freiheitsgraden ergibt sich für den Vergleich deutscher und chinesischer Probanden ein t-Wert von $t_{(97.99)} = 13.24$, der auf dem alpha-Niveau von .05 signifikant ist. Die geschätzte Effektgröße liegt bei $\eta^2 = 0.57$.

Im zweiten Vergleich zwischen spanischen und chinesischen Probanden deckt der *t-Test* wiederum bedeutsame Unterschiede auf $t_{(5.18)} = 5.95$. Die Effektgröße wird mit $\eta^2 = 0.78$ geschätzt,

Der dritte Paarvergleich bezieht sich auf deutsche und spanische Probanden im *t-Test*. Der t-Wert $t_{(5,33)} = 2.96$ überschreitet den kritischen Wert. Für die Effektstärke errechnet sich ein Wert von $\eta^{2'} = 0.23$. Insgesamt ergibt sich folgender Trend: Die langsamste Reproduktionszeit wird bei der spanischen Gruppe mit 727 Millisekunden pro Item gemessen, am schnellsten ist die Reproduktion bei der chinesischen Gruppe mit 328 Millisekunden. Der Wert der deutschen Versuchspersonen liegt mit 527 Millisekunden in der Mitte zwischen den beiden anderen Sprachgruppen.

Zusammenfassend bleibt festzuhalten, daß die Muttersprache erwartungsgemäß sowohl für die Geschwindigkeit der Artikulation als auch auf die Höhe der Gedächtnisspanne eine relevante Determinante ist. Ferner läßt sich festhalten, daß sich auch die Reproduktionszeiten signifikant unterscheiden und daß der Trend der Veränderungen der Richtung der Artikulationsunterschiede entspricht. Die Rangfolge verändert sich nicht zwischen den drei abhängigen Variablen: die kürzesten Zeitintervalle und die besten Gedächtnisleistungen finden sich in den chinesischen Daten. Die längsten Zeitmessungen und die niedrigste, mittlere Gedächtnisleistung treten in der spanischen Gruppe auf. Die Daten deutscher Versuchspersonen liegen in allen drei Parametern dazwischen.

7.2.2 DER EFFEKT ARTIKULATORISCHER UNTERDRÜCKUNG

Der Effekt Artikulatorischer Unterdrückung wird in drei Auswertungen geprüft: Zuerst werden für die deutschen Probandengruppen Unterschiede in Höhe der Gedächtnisspanne und Reproduktionszeit untersucht. In einem zweiten Schritt werden dieselben Tests für die chinesischen Gruppen durchgeführt. Als letztes wird die Eliminierung des in Kapitel 7.2.1 nachgewiesenen Spracheffekts durch Artikulatorische Unterdrückung getestet. Die Auswertungen bezüglich der abhängigen Variablen „Artikulationszeit“ entfallen, da bei diesen Messungen keine Zusatztätigkeit eingeführt werden kann.

DER EFFEKT ARTIKULATORISCHER UNTERDRÜCKUNG BEI DEUTSCHEN PROBANDEN

Eine einfaktorielle *Varianzanalyse* (unabhängige Variable: mit / ohne Zusatztätigkeit; abhängige Variable: Gedächtnisspanne) belegt, daß die Gedächtnisleistung bei deutschen Probanden durch die artikulatorische Zusatztätigkeit signifikant beeinträchtigt wird (6.15 Items versus 5.16): $F_{(1,142)} = 61.22$, $s_e^2 = 0.548$, $p < .05$, $\eta^{2'} = .30$.

Deskriptiv zeigt sich eine erhebliche Verlangsamung der Reproduktionszeit bei artikulatorischer Zusatztätigkeit in der Enkodierphase: 709 versus 525 Millisekunden. Da

keine interexperimentelle Varianzhomogenität gegeben ist, wird ein *Welch-t-Test* durchgeführt ($t_{(37.23)} = 5,92, p < .05$). Die Effektgröße wird auf $\eta^{2'} = .35$ geschätzt.

Als Resultat dieser Auswertungen kann festgehalten werden, daß bei deutschen Probanden die Zusatztätigkeit auf die Höhe der Gedächtnisspanne in erwarteter Weise wirkt. Zudem zeigen sich verlängerte Reproduktionszeiten in der Wiedergabephase, wenn während der Präsentationsphase artikulatorisch unterdrückt wird.

DER EFFEKT ARTIKULATORISCHER UNTERDRÜCKUNG BEI CHINESISCHEN PROBANDEN

Die Auswertungen der chinesischen Probanden sind analog zu den vorherigen der deutschen Probanden.

- Die Varianzhomogenität zwischen den Werten der Gedächtnisspanne von Experiment 4 und 5 ist bei $\alpha = .05$ signifikant. Wiederum wurde ein *t-Test* mit Berücksichtigung ungleicher Varianzen berechnet. Die Anzahl der Freiheitsgrade reduziert sich von 112 auf 109. Der resultierende t-Wert überschreitet den kritischen Wert ($t_{(109)} = 5.38, p < .05$). Es kann festgestellt werden, daß Zusatztätigkeit die Gedächtnisleistung reduziert (8.37 versus 7.22 Items). Die Größe des Effekts wird auf $\eta^{2'} = 0.21$ geschätzt.
- Zur Prüfung von Unterschieden in der Reproduktionszeit wird ein *t-Test* für unabhängige Stichproben durchgeführt mit Berücksichtigung des Ergebnisses der fehlenden Varianzhomogenität. Mit 45.23 Freiheitsgraden ergibt sich ein signifikanter t-Wert von $t_{(45.23)} = -4.61, p < .05$. Die Wiedergabezeit ist bei Sekundärtätigkeit verlängert: 430 versus 328 Millisekunden, die Größe des Effekts wird mit $\eta^{2'} = 0.24$ beschrieben.

Zusammenfassend ergeben sich sowohl bei deutschen als auch bei chinesischen Probanden unter der Bedingung Artikulatorischer Unterdrückung äquivalente Resultate: In beiden Kohorten wird die Gedächtnisleistung durch die Sekundäraufgabe signifikant reduziert. Für die Reproduktionszeit ergibt sich in beiden Paarvergleichen eine signifikante Verlängerung dieses Parameters.

ELIMINIERUNG DES SPRACHEFFEKTS DURCH ARTIKULATORISCHE SUPPRESSION

Folgende Auswertung vergleicht die beiden Experimente mit Artikulatorischer Unterdrückung von deutschen, respektive chinesischen Probanden (Experiment 2 & 5). Erwartet wird eine Eliminierung des Spracheffekts (Kapitel 7.2.1) bezüglich der Gedächtnisspanne durch diese Sekundäraufgabe. Da die Kovariate „Alter“ keinen

signifikanten Beitrag zur Varianzaufklärung leistet, wie die Auswertungen zum Spracheffekt nahelegen, wird auf eine Berücksichtigung dieser Variablen im weiteren verzichtet.

Rein deskriptiv ist diese Hypothese der Eliminierung bezüglich der Gedächtnisspanne schon widerlegt. Chinesische Probanden merken sich trotz Suppression im Durchschnitt wesentlich mehr Items als deutsche Probanden: 7.22 Items versus 5.16. Ein *t-Test* für unabhängige Stichproben mit ungleichen Varianzen belegt, daß die Werte auch statistisch verschieden sind ($t_{(89,80)} = -10.88$, $p < .05$). Dabei liegt die Größe des Effekts bei $\eta^2 = 0.51$.

Eine *Varianzanalyse* mit der abhängigen Variablen „Reproduktionszeit“ resultiert in signifikanten Abweichungen zwischen den Sprachgruppen. Die Varianzhomogenität ist gegeben ($F_{(1,60)} = 1.616$, n.s.). Die Reproduktionsgeschwindigkeit der deutschen Probanden ist gegenüber den chinesischen deutlich retardiert, $F_{(1,60)} = 66.59$, $s_e^2 = 17849$, $p < .05$, bei einer Effektgröße von $\eta^2 = 0.53$. Die Werte für chinesische Probanden liegen mit 430 Millisekunden im Mittel niedriger als bei den deutschen Teilnehmern mit 709 Millisekunden als Mittelwert.

Summa summarum muß konstatiert werden, daß der in Studien berichtete Effekt (Kapitel 2) der Eliminierung von Spracheffekten durch die artikulatorische Sekundäraufgabe nicht repliziert werden kann. Trotz Artikulatorischer Unterdrückung bestehen weiterhin systematische Unterschiede zwischen den Sprachgruppen in der Höhe der Gedächtnisspanne. Die Stärke des Effekts kann als groß interpretiert werden. Für die Werte der Reproduktionszeit ergibt sich ebenfalls ein bedeutsamer Unterschied zwischen den Sprachgruppen, der in der Tendenz analog zu den Untersuchungen ohne Sekundäraufgabe ist, und schnellere Wiedergabezeiten für chinesische Probanden zeigt.

7.2.3 DER GESCHLECHTSEFFEKT

Im folgenden wird geprüft, ob Unterschiede zwischen den Geschlechtern bestehen. Erwartet wird für Teilnehmerinnen im Durchschnitt sowohl höhere Gedächtnisspannen als auch schnellere Artikulationszeiten. Für die Variable „Reproduktionszeit“ kann keine Hypothese aus der Literatur abgeleitet werden.

Der Aufbau dieses Kapitels gleicht dem des letzten: Getestet wird wiederum erst getrennt für die beiden Experimente, die ohne Zusatztätigkeit durchgeführt werden, ob ein Geschlechtsunterschied besteht. In einem dritten Schritt soll untersucht werden, ob die geschlechtsspezifischen Unterschiede unter der Sekundäraufgabe Artikulatorischer Unterdrückung verschwinden.

Experiment 1: In einer *Varianzanalyse* innerhalb des ersten Experiments mit deutschen Probanden wird getestet, ob die unabhängige Variable „Geschlecht der Versuchsperson“ einen Effekt auf den abhängigen Faktor „Artikulationszeit“ ausübt. Entgegen der Erwartung zeigt sich kein Geschlechtseffekt: $F_{(1,82)} = 0.219$, $s_e^2 = 3885$, n.s.. Eine zweite *Varianzanalyse* zeigt, daß zwischen den beiden Geschlechtern keine bedeutsame Abweichung in bezug auf den Faktor „Gedächtnisspanne“ besteht: $F_{(1,85)} = 1.26$, $s_e^2 = 0.572$, n.s.. Eine dritte *Varianzanalyse* mit dem abhängigen Faktor „Reproduktionszeit“ deckt ebenfalls keine geschlechtsabhängigen Differenzen auf: $F_{(1,62)} = 2.73$, $s_e^2 = 9080$, n.s..

Experiment 4: Für chinesische Probanden ergibt sich ein äquivalentes Ergebnismuster wie bei der deutschen Gruppe. Auch in Experiment 4 ist kein Geschlechtseffekt nachweisbar.

- Die erste *Varianzanalyse* zeigt keinen Unterschied zwischen Chinesen und Chinesinnen in der Artikulationszeit $F_{(1,48)} = 0.624$, $s_e^2 = 4022$, n.s.
- Die Höhe der Gedächtnisspanne variiert in der *Varianzanalyse* $F_{(1,55)} = 0.024$, $s_e^2 = 1.11$, n.s. ebenfalls nicht zwischen den beiden Geschlechtern.
- Für die abhängige Variable „Reproduktionszeit“ zeigt sich in der *Varianzanalyse* mit dem unabhängigen Faktor „Geschlecht des Probanden“, daß der empirische F-Wert $F_{(1,34)} = 0.298$, $s_e^2 = 2940$, n.s. nicht die kritische Grenze überschreitet.

Eine Analyse zur Überprüfung der Baddeley-Hypothese, daß die geschlechtsspezifischen Unterschiede bei zusätzlicher artikulatorischer Tätigkeit verschwinden, erübrigt sich, weil in keinem der beiden Experimente in keinem der drei Parameter der Geschlechtseffekt nachgewiesen werden kann.

7.2.4 DER MATERIALEFFEKT

Innerhalb der vorliegenden Experimentreihen werden zwei Arten von Materialeffekten unterschieden: erstens, differentielle Wirkungsmechanismen innerhalb der Materialklasse "Worte" in Abhängigkeit von der Phonologischen Ähnlichkeit der Items und zweitens, durch verschiedene Stimuli-Arten (Zahlen versus Worte) evozierte Diskrepanzen.

DER MATERIALEFFEKT INNERHALB DER KLASSE „WORTE“

Die Auswertungen bezüglich des Materialeffekts innerhalb der Wortklasse beziehen sich auf alle drei abhängigen Variablen „Artikulationszeit“, „Gedächtnisspanne“ und „Reproduktionszeit“. Bezüglich der Artikulationszeit existieren keine Hypothesen, weil die Befunde widersprüchlich sind - wie in Kapitel 2 berichtet. Unterschiede in der Reproduktionszeit werden von Doshier und Ma (1998) für englische Versuchspersonen

berichtet. Oft repliziert ist die reduzierende Wirkung der Ähnlichkeit auf die Merkleistung (Kapitel 2).

Die Variation der Artikulationszeiten zwischen den Materialien spricht bei den Werten von 353 versus 412 Millisekunden deskriptiv für einen Einfluß der Ähnlichkeit auf diesen Parameter. Eine einfaktorielle *Varianzanalyse* bestätigt diesen Befund statistisch: $F_{(1,112)} = 21.89$, $s_e^2 = 4582$, $\eta^2 = 0.16$, $p < .05$. Die Artikulationszeit für klangunähnliche Items ist kürzer als für ähnliche Items.

Da die Varianzhomogenität zwischen Experiment 6 und 7 bei Messung der Gedächtnisspanne nicht gegeben ist, wird der *Welch-t-Test* durchgeführt. Der t-Wert von $t_{(103.9)} = -7.59$ ist signifikant. Die Effektstärke wird mit $\eta^2 = 0.34$ geschätzt. Die Richtung des Unterschieds entspricht der Erwartung: Bei phonologischer Ähnlichkeit ist die Merkleistung deutlich reduziert (4.42 versus 5.29 Items).

Varianzhomogenität ist auch für die abhängige Variable der Reproduktionszeit nicht gegeben. Ein *Welch-t-Test* unabhängiger Stichproben resultiert bei reduzierter Anzahl der Freiheitsgrade in einem Überschreiten der kritischen Grenze durch den empirischen t-Wert ($t_{(87.71)} = 3.68$, $p < .05$, $\eta^2 = 0.12$). Der deskriptive Trend entspricht den Daten von Doshier und Ma (1998): die Reproduktionszeiten bei ähnlichem Material sind länger (727 Millisekunden) als die bei distinktem Material (635 Millisekunden).

Zusammenfassend replizieren die hier vorliegenden Daten bereits bestehende Untersuchungen zur Reduktion der Gedächtnisleistung bei steigender Itemähnlichkeit. Die Annahmen von Doshier und Ma können auf deutsche Probanden übertragen werden. Zudem werden zwischen ähnlichem und unähnlichem Material signifikante Abweichungen in der Artikulationszeit beobachtet. Der Trend zeigt immer eine Verlangsamung der Zeitmessungen und eine Beeinträchtigung der Merkleistung bei ähnlichem Items.

DER MATERIALEFFEKT ZWISCHEN DEN KLASSEN „WORTE“ UND „ZAHLEN“

Der Materialvergleich zwischen Worten und Zahlen berücksichtigt zwei Experimentreihen: Experiment 7 und Experiment 8.

Der deskriptive Vergleich der Mittelwerte der Artikulationszeitmessungen der beiden Experimente mit Zahlworten, respektive phonologisch unähnlichen Worten, zeigt eine schnellere Geschwindigkeit für Zahlenmaterial: Im Durchschnitt werden 328 Millisekunden

zur Artikulation eines Zahlwortes benötigt, im Kontrast zu 353 Millisekunden bei unähnlichen Worten. Dieser Effekt ist statistisch nicht signifikant ($F_{(1,75)} = 2.75$, n.s., $.05$, $s_e^2 = 3212$) wie eine *Varianzanalyse* mit dem unabhängigen Faktor „Materialklasse“ und dem abhängigen Faktor „Artikulationszeit“ belegt.

Die *Varianzanalyse* zur Überprüfung des Materialklasseneffekts auf den Faktor „Höhe der Gedächtnisspanne“ deckt signifikante Unterschiede auf: $F_{(1,75)} = 7.67$, $p < .05$, $s_e^2 = 0.534$. Die geschätzte Effektgröße ist mit $\eta^2 = 0.09$ gering. Der Trend zeigt eine höhere Gedächtnisspanne für Zahlenmaterial auf (5.82 versus 5.29 Items).

Als letzter Vergleich des Materialeffekts werden die Reproduktionszeiten der beiden Experimente miteinander verglichen. In Experiment 7 werden zur Reproduktion eines Items 635 Millisekunden im Durchschnitt benötigt, im Experiment mit Zahlworten dagegen nur 498 Millisekunden pro Item. Da nach dem *Levene-F-Test* keine Varianzhomogenität gegeben ist ($F_{(1,66)} = 4.97$, $p < .05$) wird ein *t-Test* mit Berücksichtigung ungleicher Varianzen durchgeführt. Dieser Test weist auf signifikante Unterschiede hin ($t_{(42.44)} = 6.31$, $p < .05$). Die Effektstärke wird auf $\eta^2 = .31$ geschätzt.

Zusammenfassend ergeben sich bedeutsame Unterschiede in der Schnelligkeit der Reproduktion und in der Höhe der Gedächtnisspanne von Zahlworten und Worten, wenn auch nur mit einer geringen geschätzten Effektgröße, während die zur Artikulation pro Item benötigten Zeiten statistisch äquivalent sind.

7.2.5 DER EFFEKT DES PRÄSENTATIONSMODUS

Aus der Studie von Chincotta und Underwood (1998b) geht hervor, daß es einen Effekt des Präsentationsmodus gibt. Dieser Effekt zeigt sich nicht in einem Haupteffekt, sondern in einem Interaktionseffekt in einer Dual-Task Situation. Auch in den von Lüer et alii (1998) publizierten Daten, konnte kein Haupteffekt der Präsentationsart von Ziffern als Zahlen oder als Symbole auf die Lernleistung und auf die Artikulationszeit nachgewiesen werden. In den folgenden Auswertungen sollen diese Ergebnisse durch den Vergleich von Experimentreihe 1 (Stille, Deutsch, Zahlen, Symbol) und Reihe 8 (Stille, Deutsch, Zahlen, Wort) repliziert und auf den Parameter der „Reproduktionszeit“ ausgeweitet werden.

Die erste *Varianzanalyse* bezieht sich auf die abhängige Variable der Artikulationszeit. Der deskriptive Unterschied zwischen den beiden Experimentreihen beträgt 25 Millisekunden, die die Probanden im Durchschnitt langsamer sind, wenn Zahlworte dargeboten werden (328 versus 303 Millisekunden). Bei einem F-Wert von $F_{(1,102)} = 3.12$, n.s., $s_e^2 = 3380$, wird der kritische Wert nicht überschritten.

Die zweite Auswertung vergleicht die „Höhe der Gedächtnisspanne“ der beiden Experimentreihen per *Varianzanalyse* miteinander. Im Durchschnitt merken sich die Probanden 6.15 Items, wenn Symbole präsentiert werden, und 5.82 Items bei Zahlworten. Da Varianzhomogenität zwischen den beiden Untersuchungen gegeben ist, wird eine Varianzanalyse ausgeführt: Es zeigt sich kein signifikanter Unterschied in der Anzahl gemerkter Items in Abhängigkeit vom Präsentationsmodus $F_{(1,105)} = 3.04$, n.s., $s_e^2 = 0.597$.

Als dritte abhängige Variable wird die Reproduktionszeit auf Abweichungen zwischen Experimentreihe 1 und 8 untersucht: Der Unterschied beträgt 29 Millisekunden (498 versus 527 Millisekunden). Dieser Unterschied ist statistisch gesehen zufällig, wie eine *Varianzanalyse* belegt. Der empirische F-Wert $F_{(1,80)} = 1.33$, $s_e^2 = 8392$, n.s., überschreitet nicht den kritischen Wert.

Zusammenfassend ist anzumerken, daß sich in allen drei Parametern (Artikulationszeit, Höhe der Gedächtnisspanne und Reproduktionszeit) nur deskriptive Unterschiede in Abhängigkeit der Darbietungsart zeigen.

7.2.6 ÜBERBLICK ÜBER DIE ERGEBNISSE

Im folgenden wird ein Überblick über die Resultate gegeben. Entgegen den Auswertungen wird nicht nach Effekten sortiert, sondern nach abhängigen Variablen.

ÜBERBLICK ÜBER DIE ERGEBNISSE ZUR ARTIKULATIONSZEIT

- *Der Effekt der linguistischen Gruppe* In der Variablen „Artikulationszeit“ zeigt sich ein deutlicher Effekt der Sprache. Am schnellsten ist die Artikulationsgeschwindigkeit im Chinesischen. Die geschätzte Effektgröße im Vergleich deutscher und chinesischer

- Probanden liegt bei $\eta^{2'} = 0.17$. Für den Unterschied zwischen deutscher und spanischer Gruppe wird die Effektgröße mit $\eta^{2'} = 0.07$ geschätzt.
- Die Untersuchungen zum *Effekt Artikulatorischer Unterdrückung* beziehen sich nicht auf die Variable „Artikulationszeit“.
- In den Experimenten mit deutschen, respektive chinesischen Probanden kann kein *Geschlechtseffekt* nachgewiesen werden.
- *Der Materialeffekt* Die Überprüfung zeigt einen signifikanten Unterschied in der Artikulationszeit innerhalb der Wortklassen zwischen ähnlichen und unähnlichen Items ($\eta^{2'} = 0.16$). Zwischen den Materialklassen „Zahlen“ und „Worte“ dagegen wird kein Effekt auf die Artikulationszeit aufgedeckt.
- Bezüglich der Artikulationszeit bei Zahlenmaterial zeigt sich kein *Effekt des Präsentationsmodus*.

ÜBERBLICK ÜBER DIE ERGEBNISSE ZU DEN KURZZEITGEDÄCHTNISLEISTUNGEN

- In dem Parameter der „Gedächtnisspanne“ zeigen sich weder Unterschiede in Abhängigkeit vom *Geschlecht* der Versuchsperson noch ein *Effekt des Präsentationsmodus* bei Zahlenmaterial.
- *Der Effekt der linguistischen Gruppe*

In der Höhe der Gedächtnisspanne bestehen zwischen allen drei Muttersprachen signifikante Unterschiede: Am besten schneiden die chinesischen Probanden ab ($\eta^{2'} = .60$ im Vergleich zu der deutschen Gruppe). Der Unterschied zwischen der deutschen und spanischen Stichprobe ist ebenfalls signifikant. Die Effektgröße wird mit $\eta^{2'} = .06$ geschätzt.

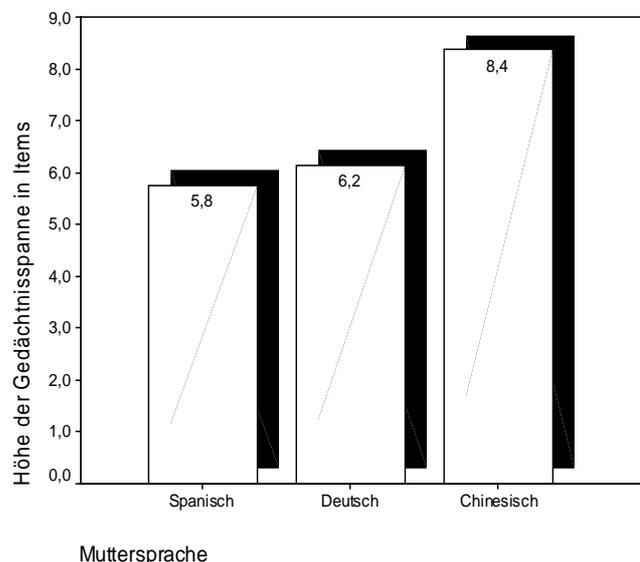


Abbildung 16 Spracheffekt & Gedächtnisleistung

- *Der Effekt Artikulatorischer Unterdrückung*

Im Vergleich der beiden Experimente mit/ ohne Sekundäraufgabe bei deutschen und bei chinesischen Probanden zeigt sich eine deutliche Reduktion der Höhe der Gedächtnisspanne bei Zusatzaktivität ($\eta^{2'} = .35$, respektive $.21$). Der Effekt der linguistischen Gruppe dagegen wird nicht durch Artikulatorische Unterdrückung eliminiert.

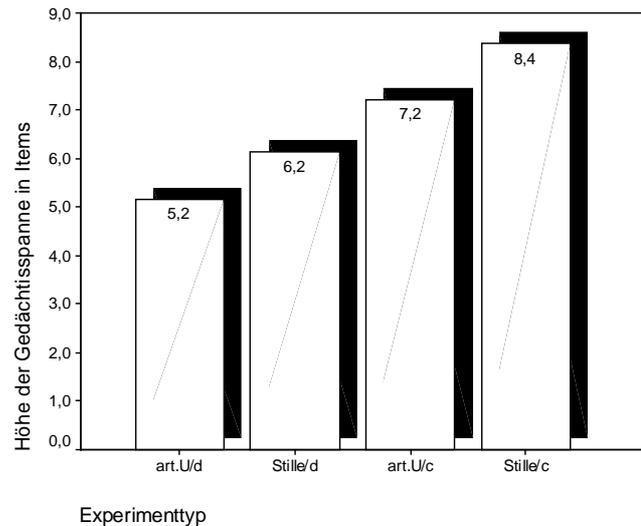


Abbildung 17 Sekundäraufgabe & Gedächtnisleistung

- *Der Materialeffekt*

Für den Materialeffekt zeigen sich sowohl innerhalb der Wortklasse als auch zwischen Wort- und Zahlenklasse signifikante Unterschiede in der Gedächtnisleistung. Der Effekt zwischen den beiden Wortexperimenten ist relativ deutlich ($\eta^{2'} = 0.34$), während der Effekt zwischen den Materialklassen zwar signifikant, aber sehr niedrig ist ($\eta^{2'} = .09$)

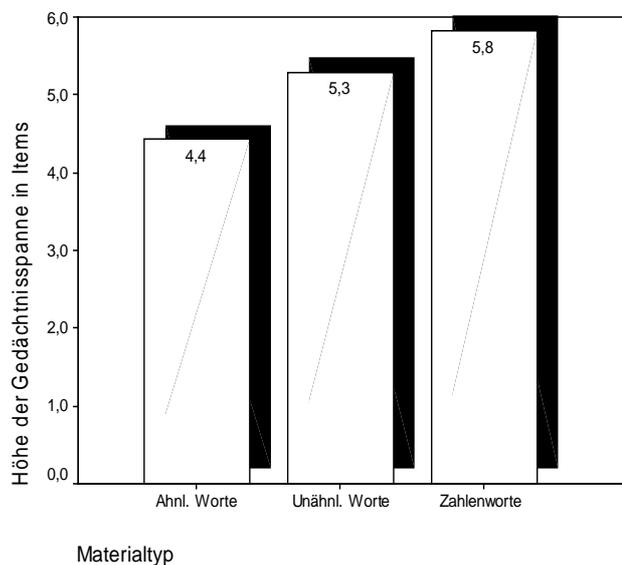


Abbildung 18 Materialeffekt & Gedächtnisleistung

ÜBERBLICK ÜBER DIE ERGEBNISSE ZUR REPRODUKTIONSZEIT

- Auch für den Faktor „Reproduktionszeit“ ist weder ein *Geschlechtseffekt* noch ein *Effekt des Präsentationsmodus* nachzuweisen.
- Zwischen den verschiedenen *Sprachen* bestehen große Effekte im Parameter der „Reproduktionszeit“. Das Wiedergabeintervall pro Item ist bei chinesischen Probanden wesentlich kürzer als bei deutschen Versuchspersonen ($\eta^{2'} = .57$). Diese benötigen im Mittel wiederum weniger Zeit als die spanische Gruppe ($\eta^{2'} = .23$).
- Die *artikulatorische Sekundäraufgabe* resultiert in einer Veränderung der Reproduktionszeit im Vergleich zu den Stillebedingungen: Sowohl bei deutschen als auch bei chinesischen Probanden verlängert sich die zur Reproduktion benötigte Zeit pro Item ($\eta^{2'} = .35$, respektive $\eta^{2'} = .24$).
- Eine *Eliminierung des Spracheffekts durch Artikulatorische Unterdrückung* zeigt sich weder im Parameter der Gedächtnisspanne noch im Parameter der Reproduktionsleistung. Wie schon unter der Experimentalbedingung „ohne Zusatzaufgabe“ finden sich in den Daten der chinesischen Stichprobe wesentlich kürzere Zeitintervalle.
- Der *Materialeffekt* dagegen affiziert die Reproduktionsdauer sowohl innerhalb der Wortklasse als auch zwischen Wort- und Zahlenmaterial: Die Reproduktionszeit bei Zahlworten ist signifikant kürzer als bei phonologisch unähnlichen Worten ($\eta^{2'} = .31$), die wiederum wesentlich kürzer ist als bei phonologisch ähnlichen Stimuli ($\eta^{2'} = .12$).

7.2.7 EINZELFALLUNTERSUCHUNGEN

Um die Äquivalenz zwischen Aggregat- und Individualdesign zu prüfen, werden für die Effekte der linguistischen Gruppe, der Artikulatorischen Unterdrückung, des Materials und des Präsentationsmodus Einzelfalluntersuchungen durchgeführt. Zur Minimierung von Interferenz-, Ermüdungs- oder Gewöhnungseffekten liegt zwischen den Untersuchungszeitpunkten ein Intervall von sechs Wochen bis drei Monate.

- Zur Überprüfung des *Effekts der linguistischen Gruppe* wurde ein Proband mit zwei Muttersprachen untersucht. Spanisch ist die dominante Sprache, aber wegen seines Aufenthalts in Deutschland wird Deutsch zum gegenwärtigen Zeitpunkt häufiger gesprochen. Für die Artikulationszeit werden für das Deutsche ein Wert von 299 Millisekunden und für das Spanische ein Wert von 354 Millisekunden ermittelt. Die Höhe der Gedächtnisspanne unterscheidet sich ebenfalls zwischen den untersuchten Sprachen: 5.45 Items im Deutschen versus 4.91 Items im Spanischen. Die Reproduktionszeiten

dagegen unterscheiden sich kaum: 692 Millisekunden bei Experiment 1 und 677 Millisekunden im Spanischen.

- Der *Effekt Artikulatorischer Unterdrückung* weist bei Zahlensymbolen auch in der Einzelfalluntersuchung in die gleiche Richtung wie bei den Aggregatuntersuchungen: Bei einem (deutschsprachigen) Probanden wird in der Stillebedingung (Experiment 1) eine Gedächtnisspanne von 5.73 Items gemessen, bei Zusatztätigkeit sinkt die Höhe auf 4.82 Items. Die Reproduktionszeit liegt bei 472 Millisekunden pro Item, respektive bei 958 Millisekunden.
- Zur Untersuchung des *Materialeffekts* führte ein Proband alle drei Experimente durch: mit phonologisch ähnlichen, respektive unähnlichen Worten sowie mit Zahlworten. Der Übersichtlichkeit wegen sind die erhaltenen Werte in Tabelle 21 gegenübergestellt. Die Zeitangaben sind in Millisekunden gemessen, die Höhe der Gedächtnisspanne bezieht sich auf die Anzahl erinnerter Items.

Tabelle 21 Einzelfalluntersuchung zum Materialeffekt

Experiment	Artikulationszeit	Gedächtnisspanne	Reproduktionszeit
6 (Stille, d, Pæ)	398	4.36	629
7 (Stille, d, Pu)	333	5.36	552
8 (Stille, d, ZW)	301	5.45	419

- Als Einzelfalluntersuchung zum *Effekt des Präsentationsmodus* nahm eine weibliche Versuchsperson an Experiment 1 (Zahlensymbol) und an Experiment 8 (Zahlworte) teil. Die Messung in Experiment 1 liegen bei 211 Millisekunden für die Artikulation, 442 Millisekunden zur Reproduktion pro Item und einer Gedächtnisspanne von acht Items. Bei Experiment 8 zeigt sich im Vergleich dazu eine Verlangsamung der Geschwindigkeit, respektive Reduktion der Gedächtnisspanne: 310 Millisekunden Artikulationszeit, 589 Millisekunden Reproduktionszeit. Die Gedächtnisleistung liegt bei 7.45 Items.

Summa summarum zeigen die Einzelfalluntersuchungen den gleichen Trend wie die Aggregatuntersuchungen in allen realisierten Effekten.

7.3 ÜBERPRÜFUNG DER INVARIANZHYPOTHESE VON BADDELEY

Der Erklärungsansatz von Baddeley zum WLE (Kapitel 2.4.4) muß zurückgewiesen werden, da der Effekt der linguistischen Gruppe nicht durch Artikulatorische Unterdrückung eliminiert werden konnte. Obwohl – bleibt man bei Baddeleys

Terminologie – der Prozeß subvokaler Wiederholung durch die Zusatzfähigkeit blockiert ist, haben die Unterschiede zwischen den Sprachen dennoch Bestand. Im folgenden wird geprüft, ob dennoch die aus dem Erklärungsansatz zum WLE abgeleitete Invarianzhypothese gültig ist. Diese Hypothese (Kapitel 3) beruht auf der Annahme einer konstanten Dauer der Gedächtnisspur, die sich aus dem Produkt von Artikulationszeit und Gedächtnisspanne ergibt. In einer strengen Auslegung wird eine interindividuelle und materialunabhängige Konstanz angenommen, in einer schwächeren Form dagegen werden interindividuelle Unterschiede zugelassen. Es wird allein eine materialunabhängige Konstanz postuliert. Die Überprüfung der Invarianzhypothese und der beiden Interpretationsansätze gliedert sich in vier Auswertungsschritte: Zuerst wird intraexperimentell und interexperimentell die *Korrelation* zwischen Artikulationszeit und Gedächtnisspanne ermittelt. Der Test auf Signifikanz wird wegen der gerichteten Hypothesen einseitig durchgeführt: Erwartet wird ein negativer Zusammenhang. Zweitens werden durch *Intervallschätzung* die empirischen Werte der Haltbarkeit der Gedächtnisspur interexperimentell verglichen und in einem weiteren Auswertungsschritt in einer *Varianzanalyse* auf eine Invarianz hin geprüft. Abschließend werden zwei Auswertungen nach dem Ansatz von Hulme (Kapitel 2.4.4) vorgenommen.

Die Experimentreihen mit Artikulatorischer Unterdrückung, respektive Phonologischer Ähnlichkeit der Items gehen nicht in die folgenden Auswertungen ein, da für diese Untersuchungsbedingungen keine Gültigkeit der Invarianzhypothese erwartet wird (Kapitel 2.4). In Tabelle 22 sind jeweils die empirisch ermittelten Werte der Persistenz der Gedächtnisspur aufgeführt. Deskriptiv zeigt sich eine relative Konstanz.

Innerhalb der einzelnen Experimente ist die *Produkt-Moment-Korrelation* der Parameter „Artikulationszeit“ und „Gedächtnisspanne“ sehr gering ausgeprägt: Die Werte sind ebenfalls in Tabelle 22 dargestellt. Allein in Experiment 1 (Stille, deutsch, Zahlen, Symbol) ist der Zusammenhang der beiden Parameter auf einem Alpha-Niveau von .05 bedeutsam. Über alle Versuchspersonen unabhängig vom Experimenttyp ist die Korrelation, mit einem Wert von $-.504$ ($n = 221$) auf einem Alpha-Niveau von .01 signifikant. Die niedrige Korrelationen im intraexperimentellen Vergleich können einerseits bedeuten, daß die „starke“ Auslegung der Invarianzhypothese nicht zutrifft und interindividuelle Unterschiede zugelassen werden müssen und/oder andererseits, daß die Kurzzeitgedächtnisspanne durch weitere Faktoren determiniert wird, wie beispielsweise durch Interferenzprozesse bei der Wiedergabe (Kapitel 7.4) oder durch einen Langzeitgedächtniseinfluß. Diesem Punkt wird am Ende des Kapitels nachgegangen.

Tabelle 22 Deskriptive Werte der Dauer der Gedächtnisspur (B), Höhe der Korrelation zwischen Artikulationszeit und Gedächtnisspanne und 95 %-Konfidenzintervall für B (in Klammern die Standardabweichung)

	B in Sekunden	Korr.koeffizient	Konfidenzintervall
1 (Stille, d, ZS)	1.85 [.37]	-.302 * ($n = 84$)	1.77 – 1.93
3 (Stille, s, ZS)	1.98 [.40]	.041 ($n = 10$)	1.73 – 2.23
4 (Stille, c, ZS)	2.04 [.65]	.093 ($n = 50$)	1.86 – 2.22
7 (Stille, d, Pu)	1.86 [.39]	-.163 ($n = 57$)	1.76 – 1.96
8 (Stille, d, ZW)	1.90 [.29]	-.312 ($n = 20$)	1.77 – 2.03
1, 3, 4, 7 und 8	1.91 [.44]	-.504 ** ($n = 221$)	1.85 – 1.97

Als zweite Auswertung wird für jede Experimentreihe einzeln das Konfidenzintervall bestimmt (Tabelle 22). Der interexperimentelle Vergleich zeigt, daß Unterschiede zwischen den Sprachgruppen in den Ausprägungen der Haltbarkeit der Gedächtnisspur bestehen: die Dauer der Gedächtnisspur ermittelt aus den Daten spanischer und chinesischer Probanden ist bei einem Konfidenzintervall von .05 signifikant länger als bei deutschen Probanden.

In einem dritten Auswertungsschritt wird die Dauer der Gedächtnisspur auf ihre Invarianz hin untersucht. Falls diese zwischen den Probanden variiert – und dafür sprechen die niedrigen Korrelationskoeffizienten-, kann die Beziehung zwischen Artikulationszeit und Gedächtnisspanne nicht für arithmetische Mittel gültig sein (Bredenkamp, 1993). Deshalb werden die individuellen Werte der beiden Parameter über den logarithmus naturalis transformiert. Aus den resultierenden, logarithmierten Werten der individuellen Dauer der Gedächtnisspur wird der Antilogarithmus gezogen: Diese Ausprägungen von Parameter B werden im folgenden auf ihre interexperimentelle Invarianz hin untersucht. Eine einfaktorielle *Varianzanalyse* mit der fünffach gestuften, unabhängigen Variablen „Experiment 1, 3, 4, 7 und 8“ und der abhängigen Variablen „Logarithmen der Dauer der Gedächtnisspur“ deckt keine signifikanten Unterschiede auf: $F_{(4,216)} = 0.311$, $s_e^2 = 0.048$, $\eta^2 = 0.02$, n.s.. Um die Power zu maximieren, wurde der Fehler erster Art auf .10 gesetzt. Der Fehler zweiter Art beträgt .09, wenn man von einem Effekt mittlere Stärke ausgeht ($f = 0.25$).

Abschließend werden explorative Analysen zu möglichen weiteren Determinanten der Kurzzeitgedächtnisspanne durchgeführt: Berechnet man die *Produkt-Moment-Korrelation* zwischen Artikulationszeit und Gedächtnisspanne unter Auspartialisierung der Variablen „Reproduktionszeit“, „Experimenttyp“ und „Alter“, so sinkt der Korrelationskoeffizient auf einen Wert von $-.26$ ($n = 139$). Trotzdem bleibt dieser Wert bei $\alpha = .01$ bedeutsam. Die

Stichprobe für diese Korrelation enthält 139 Datenpunkte, weil nur Daten der Probanden eingehen, deren Zeitmessung sowohl der Artikulation als auch der Reproduktion den Reliabilitätsanforderungen genügen. Die meiste aufgeklärte Varianz geht durch Auspartialisierung der Reproduktionszeit verloren: der Wert sinkt auf $-.28$. Eine weitere Reduktion um $.02$ geht zu Lasten der beiden zusätzlichen Kontrollvariablen „Alter“ und „Experimenttyp“.

Ein zweiter Faktor, der die Höhe der Kurzzeitgedächtnisspanne beeinflusst, könnte der von Baddeley vernachlässigte Einfluß des Langzeitgedächtnisses (Hulme et al., 1991) sein. Die Analyse nach Hulme wird nur für Experiment 1 ausgeführt, da dies das einzige Experiment ist, in dem ein substantieller Zusammenhang zwischen Artikulationszeit und Gedächtnisspanne besteht. Zur Prüfung des Einflusses des Langzeitgedächtnisses nach Hulme werden die Artikulationszeiten in Artikulationsraten umgerechnet (gesprochene Items pro Sekunde), dann wird eine *Regression* mit dem Prädiktor der Artikulationsrate und dem Kriterium „Höhe der Gedächtnisspanne“ berechnet. Die additive Konstante dieser Regressionsanalyse wird interpretiert als Rekonstruktionsprozeß oder als Einfluß des Langzeitgedächtnisses und sollte – entgegen den Annahmen von Baddeley – signifikant von Null verschieden sein. Hulme et alii (1991) berichten über unterschiedliche a_{Hulme} -Parameter im Vergleich zwischen Worten und Nonworten und eine Invarianz des b_{Hulme} -Parameters. Für den Steigungsparameter b_{Hulme} wird angenommen, daß er „den Zusammenhang zwischen Behalten und Silbenzahl beschreibt“ (Bosshardt & Laug, 1995, S. 586). Für Experiment 1 ergibt sich der Wert von 5.04 als additive Konstante der Regressionsanalyse und ein Steigungsparameter von 0.325 . Der Ordinatenabschnitt ist signifikant von Null verschieden. Nach dem Ansatz von Hulme ist in Experiment 1 (Deutsch, Stille, Zahlen, Symbol) also ein bedeutsamer Beitrag des Langzeitgedächtnisses zu verzeichnen.

Zusammenfassend zeigen sich folgende Ergebnisse:

- Die Korrelation zwischen Kurzzeitgedächtnisspanne und Artikulationszeit ist niedrig. In zwei Experimentreihen besteht entgegen der Erwartung ein positiver Zusammenhang. Lediglich in Experiment 1 kann eine substantielle, negative Korrelation nachgewiesen werden.
- Die Intervallschätzung der Haltbarkeit der Gedächtnisspur zeigt, daß zwischen den Experimentreihen mit variierenden Muttersprachen Unterschiede bestehen.
- Die Varianzanalyse deckt keine bedeutsamen Unterschiede zwischen den Experimentreihen bezüglich der Persistenz auf.
- Bei Auspartialisierung der Reproduktionszeit aus der Artikulationszeit sinkt der Korrelationskoeffizient zwischen Artikulationszeit und Gedächtnisspanne deutlich ab.

- In Experiment 1 (Stille, Deutsch, Zahlen, Symbol) ist der Ordinatenabschnitt der linearen Regression (Prädiktor: Artikulationsrate, Kriterium: Gedächtnisspanne) entgegen der Annahme von Baddeley statistisch von Null verschieden. Nach dem Ansatz von Hulme et alii (1991) wird dies als Einfluß des Langzeitgedächtnisses interpretiert.

7.4 ÜBERPRÜFUNG DER INVARIANZHYPOTHESE VON DOSHER UND MA

Im folgenden wird die Invarianzhypothese von Doshier und Ma überprüft. Es wird angenommen, daß die entscheidende Variable zur Prognose der Gedächtnisspanne die gegenüber Interferenzprozessen sensitive Reproduktionszeit pro Item (d) ist (Kapitel 2.4.4).

Das Vorgehen der Überprüfung entspricht dabei weitgehend Kapitel 7.3: Zuerst wird die *Produkt-Moment-Korrelation* zwischen Reproduktionszeit und Höhe der Gedächtnisspanne bestimmt. Da die Richtung des Zusammenhangs durch die Theorie spezifiziert wird – es soll sich um eine negative Beziehung handeln – wird der Test auf Signifikanz einseitig durchgeführt. Als nächster Auswertungsschritt folgt eine *Intervallschätzung* der Gesamtwiedergabe-Zeit und ein interexperimenteller Vergleich. Zudem wird durch eine *Varianzanalyse* geprüft, ob eine invariante Output-Zeit über alle Probanden und Experimente vorliegt. Dazu werden, wie auch schon im vorigen Kapitel, die logarithmierten Werte von Gedächtnisspanne und Reproduktionszeit verwendet (Bredenkamp, 1993). Abschließend wird erstens eine Analyse zum Ansatz einer invarianten Reproduktionszeit nach Cowan durchgeführt und zweitens werden die Prädiktoren der Gedächtnisspanne - Artikulationszeit und Reproduktionszeit- gegenübergestellt.

Die Korrelationskoeffizienten der *Produkt-Moment-Korrelation* zwischen „Reproduktionszeit“ und „Höhe der Gedächtnisspanne“ sind in Tabelle 23 eingetragen. In der mit „D“ gekennzeichneten Spalte werden die empirischen Ausprägungen der Gesamtwiedergabe-Zeit für jedes Experiment zusammen mit der Standardabweichung (in eckigen Klammern) aufgeführt, dann folgt eine Auflistung der Korrelationskoeffizienten nach Pearson. Die Richtung des Zusammenhangs zwischen den beiden Parametern „Gedächtnisspanne“ und „Reproduktionszeit“ ist in sechs der acht Auswertungen erwartungskonform negativ. In Experimentreihe 2 und 3 ist die Korrelation positiv. Allerdings muß angemerkt werden, daß durch das Selektionskriterium der Gütekriterien die Anzahl gültiger Fälle in Experimentreihe 3 erheblich reduziert wird. Interexperimentell liegt die Höhe des Korrelationskoeffizienten bei $-.693$. Dieser Wert ist auf einem Alpha-Niveau von $.01$ signifikant.

Tabelle 23 Deskriptive Werte der Dauer der Gesamtwiedergabe-Zeit (D), Höhe der Korrelation zwischen Reproduktionszeit und Gedächtnisspanne und 95 %-Konfidenzintervall für D

Experiment	D in Sekunden	Korr.koeffizient	Konfidenzintervall
1 (Stille, d, ZS)	3.28 _[.63]	-.208 * (n = 64)	3.12 – 3.44
2 (art.U, d, ZS)	3.82 _[.96]	.002 (n = 29)	3.46 – 4.18
3 (Stille, s, ZS)	4.37 _[1.28]	.658 (n = 6)	3.14 – 5.60
4 (Stille, c, ZS)	2.76 _[.54]	-.148 (n = 36)	2.58 – 2.94
5 (art.U, c, ZS)	3.28 _[.90]	-.293 * (n = 34)	2.97 – 3.59
6 (Stille, d, Pæ)	3.18 _[.55]	-.498 * (n = 51)	3.03 – 3.33
7 (Stille, d, Pu)	3.37 _[.58]	-.322 * (n = 50)	3.21 – 3.53
8 (Stille, d, ZW)	2.88 _[.57]	-.054 (n = 18)	2.60 – 3.16
1 - 8	3.26 _[.75]	-.693** (n = 287)	3.17 – 3.35

In einem zweiten Auswertungsschritt werden intraexperimentell Konfidenzintervalle für die Gesamtwiedergabe-Zeit berechnet (Tabelle 23). Es zeigt sich erstens, daß die Gesamtwiedergabe-Zeit in beiden Sprachgruppen in einem Konfidenzintervall von .05 bei zusätzlicher Artikulatorischer Unterdrückung signifikant verlängert ist. Zweitens liegen die beiden empirischen Werte der Gesamtwiedergabe-Zeit spanischer und chinesischer Probanden nicht in dem Intervall, das für die deutsche Kohorte ermittelt wurde. Drittens bestehen Unterschiede in Abhängigkeit vom Grad phonologischer Verwandtschaft zwischen den Items (Experiment 6 & 7) und vom Präsentationsmodus (Experiment 1 & 8).

Der Test, ob die Wiedergabezeit in allen Experimentreihen eine Konstante ist, wird durch eine einfaktorielle *Varianzanalyse* mit der unabhängigen, achtfach gestuften, Variablen „Experimenttyp“ und dem abhängigen Faktor „Logarithmen der Wiedergabezeit“ realisiert. Die Nullhypothese entspricht der Prognose von Doshier und Ma. Die Forderung nach Varianzhomogenität ist erfüllt. Das Ergebnis ist ein F-Wert mit $F_{(7,279)} = 8.33$, $\eta^2 = 0.17$, der größer als der theoretische F-Wert ist. Die Power des Tests liegt bei Annahme eines mittleren Effekts $f = 0.25$ und eines Fehlers erster Art von .10 bei $\beta = .06$. Der post-hoc *Scheffé Test* deckt folgende Befundlage auf: Die Wiedergabezeiten der deutschen Probanden bilden eine homogene Untergruppe unabhängig davon ob Zahlensymbole, phonologisch ähnliche oder unähnliche Stimuli, Zahlworte oder eine Sekundäraufgabe vorgegeben wurden (3.28, 3.18, 3.37, 2.88 versus 3.82 Sekunden). Von diesem Wert weichen signifikant die Wiedergabezeiten der anderen Sprachgruppen ab: Die Wiedergabezeit bei den chinesischen Probanden ist signifikant kürzer (Experiment 4: 2.76 Sekunden) als bei Experiment 1 (3.28

Sekunden), respektive die der spanischen Kohorte signifikant länger (Experiment 3: 4.37 Sekunden). Das letzte Ergebnis ist nur mit äußerster Vorsicht zu interpretieren, denn erstens gehen in diese Analyse nur sechs gültige Daten ein und zweitens ist die Standardabweichung der Gesamtwiedergabe-Zeit bei der spanischen Kohorte mit 1.28 Sekunden mehr als doppelt so groß im Vergleich zu den anderen Experimentreihen (Tabelle 11).

Im folgenden wird die Annahme einer konstanten Gesamtwiedergabe-Zeit der Beobachtung Cowans (1994) gegenübergestellt. Diese besagt, daß die Reproduktionszeit pro Item eine interindividuelle Konstante ist und die Wiedergabezeit zwischen den Probanden variiert: Probanden mit höherer Behaltensspanne benötigen ein längeres Reproduktionsintervall. Zur Überprüfung der These von Cowan werden die Wiedergabezeiten pro Experiment in vier Quartile zerlegt. Für die Gruppe mit den niedrigsten Zeitintervallen sowie für die Gruppe mit dem längsten Reproduktionsintervall werden jeweils die Mittelwerte für Gedächtnisspanne, Wiedergabezeit und Reproduktionszeit berechnet. In Tabelle 24 findet sich eine Gegenüberstellung der drei abhängigen Variablen „Reproduktionszeit“ (Rep.zeit in Millisekunden), „Gesamtwiedergabe-Zeit“ (GWZ in Sekunden) und „Gedächtnisspanne in Items“ (Ged.s.) separat für jede Experimentreihe. In den Klammern ist jeweils die Standardabweichung angegeben. Experiment 3 wird wegen des geringen Stichprobenumfangs aus der Auswertung ausgeschlossen.

Tabelle 24 Mittelwerte und Standardabweichungen der Probanden des ersten und letzten Quartils der Gesamtwiedergabe-Zeit für die Messungen der Reproduktionszeit (Rep.zeit), Gesamtwiedergabe-Zeit (GWZ) und Gedächtnisspanne (Ged.s.)

	Erstes Quartil			Viertes Quartil		
	Rep.zeit	GWZ	Ged.s.	Rep.zeit	GWZ	Ged.s.
1 (Stille, d, ZS)	427 _[66]	2.41 _[.28]	5.69 _[.51]	653 _[74]	4.23 _[.47]	6.52 _[.68]
2 (art.U, d, ZS)	604 _[64]	2.77 _[.16]	4.64 _[.58]	839 _[96]	4.81 _[.47]	5.78 _[.72]
4 (Stille, c, ZS)	273 _[37]	2.12 _[.18]	7.83 _[.78]	407 _[48]	3.56 _[.45]	8.79 _[.92]
5 (art.U, c, ZS)	339 _[53]	2.37 _[.14]	7.14 _[1.18]	582 _[124]	4.39 _[.63]	7.72 _[1.20]
6 (Stille, d, Pae)	599 _[64]	2.53 _[.22]	4.25 _[.38]	893 _[126]	3.90 _[.33]	4.44 _[.70]
7 (Stille, d, Pu)	564 _[64]	2.64 _[.12]	4.73 _[.57]	716 _[72]	4.08 _[.35]	5.71 _[.34]
8 (Stille, d, ZW)	439 _[60]	2.27 _[.20]	5.20 _[.30]	534 _[66]	3.64 _[.47]	6.84 _[.51]

Der deskriptive Vergleich allein zeigt, daß zwischen den Reproduktionszeiten – mit Ausnahme von Experiment 8 - der beiden Extremgruppen erhebliche Abweichungen

bestehen. Schon durch diese Betrachtung ist klar, daß die Beobachtung einer invarianten Reproduktionszeit nicht in den hier realisierten Experimenten haltbar ist, so daß auf eine Darstellung der detaillierten statistischen Analysen verzichtet wird. Allein in Experiment 8 sind die Reproduktionsintervalle zwischen den Extremgruppen statistisch äquivalent, aber der Trend entspricht dem der anderen Experimente. Ein weiteres, in Tabelle 24 dargestelltes Resultat ist für die Hypothese von Doshier und Ma (1998) kritisch: Der Extremgruppenvergleich zeigt in allen Experimenten, daß die Probanden des vierten Quartils, das durch eine lange Wiedergabezeit definiert ist, im Durchschnitt sowohl höhere Behaltensspannen haben als auch längere Reproduktionszeiten. Diese Beobachtung widerspricht der Hypothese von Doshier und Ma (1998) und wird in Kapitel 8.3 nochmals aufgenommen.

Die letzten beiden Auswertungen beziehen sich auf den Vergleich der Invarianzhypothesen des Kurzzeitgedächtnisses. In Kapitel 7.2 kann gezeigt werden, daß die durch die Artikulationszeit aufgeklärte Varianz deutlich reduziert ist, wenn Einflüsse des Experimenttyps und der Reproduktionszeit auspartialisiert werden. Die Varianzaufklärung durch die Reproduktionszeit wird durch Auspartialisierung der Artikulationszeit dagegen kaum reduziert: Der Korrelationskoeffizient liegt nur mehr bei -.61 statt bei -.69. Über die Artikulationszeit hinaus wird also durch die Reproduktionszeit ein signifikanter Beitrag zur Prognose der Kurzzeitgedächtnisspanne geleistet. Nach dem Modell von Hulme kann dies den Prozessen während des Präparationsintervall und während der Pausenzeiten zugeschrieben werden (Kapitel 8.3).

Ein *Korrelations-Differenzen-Test* (Formel 17, McNemar, 1962) soll zeigen, ob durch die Variable "Reproduktionszeit" signifikant mehr Varianz der Gedächtnisleistung aufgeklärt werden kann als durch den Faktor „Artikulationszeit“.

Formel 17
$$t = \frac{(r_{12} - r_{13}) \sqrt{(n-3)(1+r_{23})}}{\sqrt{2(1-r_{12}^2 - r_{13}^2 - r_{23}^2 + 2r_{12}r_{13}r_{23})}}$$

Da für diese Formel auf die gleiche Anzahl von Daten zurückgegriffen werden muß, ändern sich die in den bisherigen Auswertungen angegebenen Koeffizientenwerte (Tabelle 22 & 24). Erstens wird der Vergleich der zwei Invarianzhypothesen über Experimentreihe 1, 3, 4 und 7 berechnet und zweitens werden nur die Daten berücksichtigt, deren Reliabilitätswerte beider Messungen zugleich dem Kriterium ($\geq .75$) entsprechen. Aus dieser reduzierten Anzahl von Werten ($n = 207$) ergeben sich folgende drei Korrelationskoeffizienten nach Pearson:

- Für den Zusammenhang zwischen Gedächtnisspanne und Artikulationszeit: $r_{12} = -.638$
- Für den Zusammenhang zwischen Gedächtnisspanne und Reproduktionszeit $r_{13} = -.731$
- Für den Zusammenhang zwischen Artikulationszeit und Reproduktionszeit $r_{23} = .572$

Nach Formel 17 erhält man einen empirischen t-Wert von 2.29. Da hier ein großer Stichprobenumfang ($n = 207$) vorliegt, kann zum Vergleich mit dem theoretischen Wert die Normalverteilung herangezogen werden. Der Wert von 2.29 überschreitet die kritische Grenze von 1.65 bei $\alpha = .05$. Somit wird durch die Einführung der Reproduktionszeit signifikant mehr Varianz aufgeklärt als durch alleinige Berücksichtigung der Artikulationszeit. Durch den Ansatz von Doshier und Ma wird also erstens in den Experimentreihen, für die auch die Baddeley-Hypothese gilt, mehr Varianz aufgeklärt. Zweitens ändert sich der Anteil durch die Reproduktionszeit aufgeklärte Varianz auch nach Auspartialisierung der Artikulationszeit kaum und drittens bewährt sich die Hypothese von Doshier und Ma ebenfalls in Experimenten mit Interferenzinduktion durch Ähnlichkeit der Items.

Insgesamt zeigen sich folgende Ergebnisse:

- Mit drei Ausnahmen besteht intraexperimentell und interexperimentell ein negativer Zusammenhang zwischen Reproduktionszeit und Kurzzeitgedächtnisspanne.
- In der Intervallschätzung zeigen sich fast durchgehend systematische Unterschiede in der Gesamtwiedergabe-Zeit zwischen den Experimentreihen.
- Die Varianzanalyse deckt bedeutsame Unterschiede zwischen den Experimentreihen in der Gesamtwiedergabe-Zeit auf. Die Invarianzhypothese ist nur innerhalb der Sprachgruppen gültig.
- Die Beobachtung Cowans, daß die Reproduktionszeit eine Konstante ist, konnte in keinem Experiment bestätigt werden.
- Trotz Auspartialisierung der Artikulationszeit aus der Reproduktionszeit nach dem Modell von Hulme, leistet die Reproduktionszeit einen signifikanten Beitrag zur Varianzaufklärung in der Kurzzeitgedächtnisspanne.
- Der Korrelations-Differenzen-Test zeigt, daß durch die Reproduktionszeit mehr Varianz an der Kurzzeitgedächtnisspanne aufgeklärt wird als durch die Artikulationszeit.

7.5 EXPERIMENTELLE MANIPULATIONEN UND DIE ASYMPTOTENHÖHE DER MTT

Ziel dieses Kapitel ist es, die Wirkung induzierter Effekte, deren Wirksamkeit bereits für die Gedächtnisspanne nachgewiesen wurde (Kapitel 7.2), auf die Asymptotenhöhe zu überprüfen.

Hintergrund dieser Analysen ist die Äquivalenzhypothese zwischen Asymptote und Gedächtnisspanne in bit. Da diese Hypothese in der intraexperimentellen Auswertung (Kapitel 6.4) für die Untersuchungen mit Zahlenmaterial nicht akzeptiert werden kann, entfallen sämtliche Überprüfungen für diese Experimentreihen. Schiffer (2000) konnte eine prozentuale Äquivalenz der Veränderung in Asymptote und Gedächtnisspanne nachweisen, obwohl Asymptote und Kurzzeitgedächtnisspanne in bit nicht gleichgesetzt werden konnten. Doch auch dieser Effekt wird in diesen Daten mit Zahlenmaterial nicht gefunden.

Im folgenden kann nur auf den Materialeffekt innerhalb der Wortklasse eingegangen werden. In den Auswertungen von Kapitel 7.2 ergab sich, daß mit zunehmendem Grad phonologischer Ähnlichkeit die Höhe der Gedächtnisspanne signifikant reduziert wird. Diese Analyse soll zeigen, ob die signifikante Reduktion auch für die Asymptote nachgewiesen werden kann. Die Überprüfung geht in zwei Schritten vor sich: Zuerst werden die Daten von Experiment 6 (ähnliche Worte) nach Formel 15 prognostiziert. Dabei geht in die Berechnung der empirische Parameter $b = 33.12$ von Experiment 6 (ähnliche Worte) ein und der empirische Parameter $a = 1.52$ von Experiment 7 (unähnliche Worte). Die so prognostizierten Daten (Tabelle 25) werden in einem weiteren Schritt per *F-Test* mit den tatsächlichen Werten (Tabelle 12) verglichen. Mit einem Zählerfreiheitsgrad und 54 Freiheitsgraden im Nenner liegt bei einem Alpha-Niveau von .10 und postulierter mittlerer Effektstärke die Power .89 (Erdfelder et al., 1996).

Tabelle 25: Vergleich der Asymptoten von Experiment 6 und 7

Experiment	Darbietungszeit		
	10 Sekunden	20 Sekunden	30 Sekunden
6 mit a_7 und b_6	4.83	3.17	2.62
7 mit a_6 und b_7	4.54	3.25	2.82

Index 6 steht für den Parameter des Experiments mit ähnlichen Stimuli, mit Index 7 sind die Parameter des Experiments mit unähnlichen Worten bezeichnet.

Der *F-Test* zum Vergleich der tatsächlichen Daten von Experiment 6 und den aus der Asymptote von Experiment 7 errechneten Werten deckt nur zufällige Unterschiede auf. Bei einer Abweichungsvarianz von 15.2 und einer Fehlervarianz von 6.26 liegt der F-Wert unter dem kritischen Wert, $F_{(1, 54)} = 2.43$, n.s.. Entsprechend wird zur Prognose der Daten von Experiment 7 ($b = 25.75$) aus der Asymptote von Experiment 6 ($a = 1.96$) vorgegangen. Die nach Formel 15 errechneten Werte stehen in Tabelle 25. Der *F-Test* deckt signifikante

Unterschiede zwischen den Prognosen und den empirischen Daten auf ($s_e^2 = 6.60$, $s_{emp}^2 = 24.13$, $F_{(1,45)} = 3.66$, $p < .05$).

Das Ergebnislage ist für die beiden Vorhersagerichtungen nicht homogen: Die signifikanten Unterschiede, die in Abhängigkeit vom Ähnlichkeitsgrad der Items für die Gedächtnisspannenwerte gefunden werden (Kapitel 7.2), finden sich statistisch in den Asymptoten nur für eine Prognoserichtung wieder. Allerdings zeigt sich ein deutlicher Trend in die erwartete Richtung: 16.95 bit bei ähnlichen Worten unterschreitet den Wert von 21.85 bit bei distinktem Material.

7.6 EXPERIMENTELLE MANIPULATION UND DER QUOTIENT A/B SOWIE DIE VERARBEITUNGSZEIT IM LANGZEITGEDÄCHNISSPEICHER

In diesem Kapitel werden zwei Phänomene untersucht: Erstens wird der Einfluß experimenteller Manipulation auf den Quotienten von Parameter a und b geprüft, um Unterschiede der Lernleistung relativ zum möglichen Maximum aufzudecken (Bredenkamp & Klein, 1998). Geprüft wird, ob - durch experimentelle Variation induziert - sich die Ausprägungen der Parameter der mTTH a und b proportional verändern, oder ob nur einer der Parameter durch das experimentelle Design beeinflusst wird. Der Terminus „Quotient“ bezeichnet im folgenden durchgehend das Verhältnis a/b der beiden materialspezifischen Konstanten der mTTH. Ein zweites Ziel dieses Kapitels ist der deskriptive, interexperimentelle Vergleich der Verarbeitungszeit im Langzeitgedächtnisspeicher, analog den Ausführungen in Kapitel 3.

Ein Vergleich des Quotient der beiden materialspezifischen Konstanten kann wegen der spezifischen Probleme mit Parameter a bei dem Stimulusmaterial von Zahlen nur für die Experimente vorgenommen werden, deren Wert für a größer eins ist, respektive deren Parameter a sich auf eins setzen läßt.

Dem ersten Kriterium genügen die Experimente 1, 2, 6 und 7. In Experiment 3 kann die mTTH angepaßt werden, wenn Parameter a auf 1 fixiert wird. Bei Berechnung der Prognosen der Daten für Experiment 3 (Stille, spanisch, Zahlen, Symbol) nach Formel 15 mit den Werten: $a = 1.00$ und $b = 57.53$ zeigt sich im F-Test keine signifikante Abweichung zu den beobachteten Daten: $F_{(1,54)} = 0.89$, n.s., $s_e^2 = 6.40$, $s_{emp}^2 = 5.70$. In den Experimenten 4, 5 und 8 mußte bei Festsetzung von a auf das theoretische Minimum die mTTH verworfen werden. In diesen drei Experimentreihen entfällt die Bestimmung des Quotienten der beiden materialspezifischen Konstanten.

Für die Experimentreihen 1, 2, 3, 6 und 7 wird die Prognose der Daten des einen Experiments durch Implementation des Quotienten des anderen Experiments errechnet. Das im folgenden beschriebene Vorgehen ist für alle Paarvergleiche identisch. Der Quotient des einen Experiments wird in das andere Experiment integriert ohne daß dabei Parameter a verändert wird. Die auf diesem Weg berechneten Werte werden dann per *F-Test* mit den empirischen Daten verglichen. Bei zwei Freiheitsgraden im Zähler und 54 Freiheitsgraden im Nenner liegt der alpha-Fehler bei .10 und der beta-Fehler bei .18, wenn man eine mittlere Effektstärke annimmt ($f^2 = .15$). Für Experiment 1 mit 87 Probanden ergeben sich folgende Fehlerwerte bei mittlerer Effektgröße: $\alpha = .05$, $\beta = .10$ (Erdfelder et al., 1996). Die Werte werden nur an dieser Stelle erwähnt, um Wiederholungen zu vermeiden.

Da die Auswertungen zum Quotient der Parameter der mTTH auf die Experimente 1, 2, 3, 6 und 7 beschränkt ist, können nicht alle Effekte, die in den Arbeitshypothesen definiert wurden, geprüft werden: die Untersuchung zur Eliminierung des Effekts der linguistischen Gruppe durch Artikulatorische Unterdrückung muß ebenso entfallen wie eine Auswertung zum Effekt des Präsentationsmodus. Die Untersuchung des Effekts der linguistischen Gruppe kann nur für den Vergleich deutscher und spanischer Probanden ausgeführt werden (Kapitel 7.6.1). Der Effekt einer artikulatorischen Zusatzaufgabe auf den Quotienten wird anhand der deutschen Population untersucht (Kapitel 7.6.2) und der Materialeffekt kann allein innerhalb der Wortklasse überprüft werden (Kapitel 7.6.3). Zu Abschluß des Kapitel wird ein Überblick gegeben über die Auswertungen. Zudem wird auf die deskriptiven Unterschiede in der Verarbeitungszeit im Langzeitgedächtnisspeicher eingegangen.

7.6.1 DER SPRACHEFFEKT

Zur Untersuchung des Spracheffekts kann nur der Paarvergleich zwischen deutschen und spanischen Probanden durchgeführt werden. Da die Prognose der Daten in beide Richtungen berechnet werden kann, ergeben sich zwei Auswertungen.

Zuerst werden die Daten von Experiment 1 prognostiziert. Unter der Vorannahme, daß der empirisch erhaltene Wert von Parameter a in Experiment 3 mit dem theoretisch geforderten Minimum gleichgesetzt wird, kann Parameter b von Experiment 3 direkt in Experiment 1 überführt werden. In Formel 15 wird der empirische Parameter $a = 1.00$ von Experiment 1 und der Parameter $b = 57.53$ von Experiment 3 übernommen. Die daraus resultierenden Werte sind in Tabelle 26 dargestellt, a/b kürzt die Bezeichnung für den Quotienten der beiden Parameter der mTTH ab, die Indices richten sich nach der durchgängigen Numerierung der Experimente.

Tabelle 26 Prognose der Daten mit Implementation des Quotienten des jeweils anderen Experiments

	Darbietungszeit		
	10 Sekunden	20 Sekunden	30 Sekunden
1 mit a_1 und a_3/b_3	6.76	3.88	2.92
3 mit a_3 und a_1/b_1	4.07	2.38	1.82

Die Abweichungsvarianz dieser erwarteten Werte von den empirischen Werten (Tabelle 12) liegt bei einem Wert von 113. Die Fehlervarianz bei Experiment 1 ist 4.72. Der aus diesen beiden Werten resultierende F-Wert des *F-Tests* überschreitet deutlich die kritische Grenze: $F_{(2,84)} = 23.99$, $p < .05$.

Die umgekehrte Prognoserichtung dient der Vorhersage der Mittelwerte von Experiment 3 ($a = 0.694$) aus der Verarbeitungszeit im Langzeitgedächtnisspeicher von Experiment 1 ($b = 33.76$). Für diese Überprüfung wird der empirisch errechnete Wert von a übernommen und nicht das theoretische Minimum. Die nach Formel 15 geschätzten Daten finden sich in Tabelle 26. In den *F-Test* geht eine Abweichungsvarianz von 73.25 und eine Fehlervarianz von 6.40 ein. Der empirische F-Wert ($F_{(2,54)} = 11.44$) ist größer als die kritische Grenze, die sich bei beiden Fehlerarten von .10 ergibt.

Die Befundlage ist homogen: Es bestehen systematische Unterschiede im Quotienten von deutschen und spanischen Probanden. Die Verarbeitungszeit im Langzeitgedächtnisspeicher ist bei der spanischen Gruppe mit 1.28 Sekunden je bit wesentlich länger als bei deutschen Probanden mit 0.75 Sekunden.

7.6.2 DER EFFEKT ARTIKULATORISCHER UNTERDRÜCKUNG

Ein Befund der bisherigen mTTH-Forschung ist, daß Parameter b der mTTH bei Artikulatorischer Unterdrückung ansteigt. Der Zuwachs ist nicht proportional zum Steigen von Parameter a (Bredenkamp & Klein, 1998). Im folgenden wird geprüft, ob sich dieser Effekt replizieren läßt. Dazu kann nur der Paarvergleich innerhalb der Gruppe deutscher Probanden ohne und mit Artikulatorische Unterdrückung vorgenommen werden.

Zuerst werden die Mittelwerte von Experiment 1 (Stillebedingung) nach Formel 15 aus $a = 1.00$ (Experiment 1) und $b = 46.68$ (Quotient von Experiment 2 mit Asymptote von Experiment 1) berechnet. Diese Prognosen sind in Tabelle 27 aufgeführt.

Tabelle 27 Prognose der Daten mit Implementation des Quotienten des jeweils anderen Experiments

Experiment	Darbietungszeit		
	10 Sekunden	20 Sekunden	30 Sekunden
1 mit a_1 und a_2/b_2	5.67	3.33	2.56
2 mit a_2 und a_1/b_1	4.66	2.97	2.41

Es ergibt sich eine Abweichungsvarianz von 34.51 und eine Fehlervarianz von 4.72. Der *F-Test* zeigt, daß die in Zeile 1 dargestellten errechneten Werte signifikant von den realen Daten abweichen (Tabelle 12) ab: $F_{(2,84)} = 7.31, p < .05$.

Zur Vorhersage der Werte von Experiment 2 werden in Formel 15 die Asymptote von Experiment 2 ($a = 1.28$) und der im Langzeitgedächtnis benötigten Zeit von Experiment ($b = 33.76$) eingesetzt. In Zeile zwei von Tabelle 27 sind die detaillierten Mittelwerte eingetragen. Diese Werte weichen nach dem *F-Test* signifikant von den empirischen Werten ab ($s_e^2 = 10.20, s_{emp}^2 = 88.35, F_{(2,54)} = 8.66, p < .05$).

Die Ergebnisse der beiden Auswertungen zum Paarvergleich von deutschen Probanden mit, respektive ohne Artikulatorische Unterdrückung zeigen, daß die erwartete Veränderung des Quotienten a/b bei Zusatztätigkeit nachzuweisen ist.

7.6.3 DER MATERIALEFFEKT

Für die beiden Wortmaterialien ist die Äquivalenzhypothese gültig, der Quotientenvergleich wird nach dem oben beschriebenen Verfahren durchgeführt. Ein Überprüfung des Materialeffekts zwischen der Klasse der „Worte“ und der Klasse der „Zahlen“ ist nicht möglich, da in Experiment 8 Parameter a nicht auf das theoretische Minimum eins erhöht werden kann, ohne daß dabei die mTTH ihre Gültigkeit verliert.

Legt man den Quotienten von a und b aus Experiment 7 zugrunde und verknüpft diese mit der Asymptote von Experiment 6, so errechnet sich für b ein Wert von $b = 33.20$. Aus diesem neuen b -Parameter und der Asymptote von Experiment 6 ($a = 1.96$) werden nach Formel 15 die Mittelwerte geschätzt (Tabelle 28).

Tabelle 28 Prognose der Daten mit Implementation des Quotienten des jeweils anderen Experiments

Experiment	Darbietungszeit		
	10 Sekunden	20 Sekunden	30 Sekunden
6 mit a_6 und a_7/b_7	5.28	3.62	3.07
7 mit a_7 und a_6/b_6	4.09	2.80	2.38

Ein Vergleich der Werte aus der ersten Spalte von Tabelle 28 mit den tatsächlich erhaltenen Daten von Experiment 6 (Tabelle 12) deckt keine systematischen Unterschiede auf. Die Abweichungsvarianz liegt bei 0.57 und die Fehlervarianz bei 6.26. Der *F-Wert* liegt mit $F_{(2,54)} = 0.0911$, n.s., unter dem kritischen Wert.

Für die zweite Auswertung dieses Paarvergleichs ergeben sich zum Einsetzen in Formel 15 folgende Werte: $a = 1.52$ (Experiment 7) und $b = 25.68$ (Quotient von a und b aus Experiment 6 bei konstanter Asymptote von Experiment 7). Die Auswertung mit dem *F-Test* resultiert wieder in zufälligen Unterschieden. Bei der Abweichungsvarianz von 4.94 und der Fehlervarianz von 6.60, n.s., ist der *F-Wert* niedriger als eins. Die Ergebnisse beider Auswertungen des Paarvergleichs belegen, daß zwischen den beiden Wortexperimenten keine signifikanten Quotientenunterschiede bestehen.

7.6.4 ÜBERBLICK ÜBER DIE ERGEBNISSE

Die Befunde zum *Effekt der linguistischen Gruppe* zeigen, daß signifikante Unterschiede in dem Quotienten von a und b zwischen dem Deutschen und dem Spanischen bestehen. Die spanischen Probanden verarbeiteten weniger Information relativiert an der maximal zu verarbeitenden Information im Vergleich zu den deutschen Versuchspersonen. Ebenso trat bei artikulatorischer Zusatzfähigkeit eine Veränderung des Quotienten a/b gegenüber der Stillebedingung auf: die Verarbeitungszeit im Langzeitgedächtnisspeicher sank. *Dieser Effekt Artikulatorischer Unterdrückung*, der schon für Silbenmaterial nachgewiesen worden ist (Bredenkamp & Klein, 1998), kann - bei deutschen Probanden - auch auf lexikales Material übertragen werden. Für den Quotienten der beiden Parameter der mTTH konnte kein *Materialeffekt* nachgewiesen werden. In Tabelle 29 werden kurz die Ergebnisse der Auswertungen zusammengefaßt.

Tabelle 29 Überblick über die Auswertungen zum Quotientenvergleich

Experiment	a	b	F-Wert
1 mit Quotient ₃	1.00	57.53	$F_{(2,84)} = 23.99 *$
3 mit Quotient ₁	1.00	33.76	$F_{(2,54)} = 11.44 *$
1 mit Quotient ₂	1.00	59.75	$F_{(2,84)} = 7.31 *$
2 mit Quotient ₁	1.28	33.76	$F_{(2,54)} = 8.66 *$
6 mit Quotient ₇	1.96	33.20	$F_{(2,54)} = 0.0911$
7 mit Quotient ₆	1.52	25.68	$F_{(2,54)} = 0.748$

Für die Experimente 4, 5 und 8 konnten keine Untersuchungen bezüglich der Quotienten der beiden Parameter der mTTH durchgeführt werden, dennoch soll deskriptiv die zur Verarbeitung im Langzeitgedächtnis benötigte Zeit zwischen den Experimenten verglichen werden. Tabelle 30 gibt einen Überblick über die Verarbeitungszeit im Langzeitgedächtnisspeicher pro bit.

Tabelle 30 Verarbeitungszeit im Langzeitgedächtnisspeicher pro bit in Sekunden

Experiment	Verarbeitungszeit pro bit
1 (Stille, d, ZS)	0.75
2 (art.U, d, ZS)	1.33
3 (Stille, s, ZS)	1.28
4 (Stille, c, ZS)	0.89
5 (art.U, c, ZS)	0.72
6 (Stille, d, Pae)	1.00
7 (Stille, d, Pu)	0.78
8 (Stille, d, ZW)	1.20

Deskriptiv deutet der Vergleich zwischen den drei linguistischen Gruppen an, daß zwischen den chinesischen und den deutschen Probanden erwartungsgemäß kaum ein Unterschied in der Verarbeitungszeit im Langzeitgedächtnisspeicher besteht. Wie auch statistisch nachgewiesen wurde, weicht der Wert der spanischen Probanden signifikant von dem der deutschen Gruppe ab.

Der Effekt Artikulatorischer Unterdrückung zeigt für den Vergleich innerhalb der deutschen Stichprobe den erwartungsgemäßen Trend einer Verlangsamung der Verarbeitungszeit. Der deskriptive Vergleich der beiden chinesischen Gruppe dagegen belegt, daß diese

Verlangsamung bei chinesischen Probanden nicht eintritt. Der Wert für die Verarbeitungszeit pro bit bei Artikulatorischer Unterdrückung in dieser Gruppe liegt sogar etwas unter dem des Lernens in Stille. Ein Vergleich der beiden Experimente mit Artikulatorischer Unterdrückung weist einen Effekt der linguistischen Gruppe auf: die Verarbeitungszeit bei deutschen Probanden ist bei Artikulatorischer Unterdrückung mit 0.61 Sekunden deutlich höher als bei den Chinesen.

Für den Materialeffekt innerhalb der Wortklassen wurden keine signifikanten Unterschiede im Verhältnis der Parameter a und b nachgewiesen: a und b stiegen proportional an bei hohem Grad phonologischer Ähnlichkeit. Der Effekt zwischen den Materialklassen „Zahlen“ und „Worte“ wird überlagert durch den Effekt des Präsentationsmodus. Je nach Darbietung der Ziffern als Symbole, respektive als Worte ergibt sich ein Unterschied in der Verarbeitungszeit pro bit von .45 Sekunden. Kontrastiert man die Transferzeit in das Langzeitgedächtnis von phonologisch unähnlichen Worten mit der benötigten Zeit für Zahlensymbole, so ergibt sich deskriptiv kein Unterschied. Die Gegenüberstellung der Experimente mit unähnlichen Worten und Zahlworten dagegen spricht für eine schnellere Verarbeitungszeit für das Wortmaterial.

7.7 ÜBERBLICK ÜBER DIE ERGEBNISSE ZU DEN INVARIANZHYPOTHESEN

Im folgenden werden kurz die Ergebnisse zu den Auswertungen der Invarianzhypothesen zusammengefaßt.

Invarianzhypothese von Baddeley Folgende Ergebnisse sprechen für die Invarianzhypothese:

- Die Auswertung zum Effekt des Geschlechts der Versuchsperson belegt für die beiden Untergruppen, daß eine schnellere Artikulationszeit deskriptiv auch mit höheren Gedächtnisleistungen zusammenhängt.
- Mit einer schnelleren Artikulationszeit geht im interexperimentellen Vergleich auch eine höhere Gedächtnisspanne einher.
- Die empirisch ermittelte Persistenz der Gedächtnisspur mit Werten von 1.85 bis 2.04 liegt in dem erwarteten Wertebereich.
- Zwischen Artikulationszeit und Gedächtnisspanne besteht der erwartete negative Zusammenhang, sowohl im intra- als auch im interexperimentellen Vergleich – mit zwei Ausnahmen.
- Interexperimentell kann über alle Experimente eine konstante Persistenz der Gedächtnisspur angenommen werden.

Gegen die Hypothese und den Erklärungsansatz von Baddeley sprechen folgende Resultate:

- Die Korrelation zwischen Artikulationszeit und Gedächtnisspanne ist intraexperimentell verschwindend gering. Auch interexperimentell ist der Anteil aufgeklärter Varianz niedrig.
- In der Intervallschätzung zeigen sich zwischen den Muttersprachen systematische Unterschiede zwischen den Experimenten.
- Die Höhe des Korrelationskoeffizienten sinkt deutlich ab, wenn die Variable „Reproduktionszeit“ auspartialisiert wird.
- Der Ordinatenabschnitt der linearen Regression mit dem Prädiktor „Artikulationsrate“ und dem Kriterium „Gedächtnisspanne“ ist in Experiment 1 signifikant von 0 verschieden.

Summa summarum wird nur die „schwache“ Interpretation der Invarianzhypothese von Baddeley bestätigt, die auch interindividuelle Unterschiede zwischen den Probanden zulässt. Zudem legen die Daten die Annahme der Existenz von weiteren Faktoren zur Determination der Kurzzeitgedächtnisspanne nahe (Kapitel 8.2).

Invarianzhypothese von Doshier und Ma Folgende Ergebnisse sprechen für die Invarianzhypothese nach Doshier und Ma:

- Deskriptiv ist die Struktur zu erkennen, daß eine schnellere Reproduktionszeit mit einer höheren Kurzzeitgedächtnisspanne einhergeht. Dieses Verhältnis ist nicht durchgängig im Vergleich aller Experimente - Ausnahmen bilden Experiment 3 und 8. Weil diese beiden Untersuchungsreihen nur wenig gültige Fälle enthalten (Experiment 3: $n = 6$, Experiment 8: $n = 18$), wird hier ein Stichprobeneffekt vermutet.
- Die Intervalle der Wiedergabezeit befinden sich in dem Intervall zwischen 2.76 und 4.37 Sekunden.
- Die Korrelation zwischen Gedächtnisspanne und Reproduktionszeit ist mit Ausnahme von Experiment 3 und den beiden Untersuchungen mit artikulatorischer Sekundäraufgabe negativ.
- Parameter D ist unabhängig von der Materialart innerhalb der Sprachen konstant.
- Die Varianzaufklärung ist höher als durch die Invarianzhypothese von Baddeley.
- Die Reproduktionszeit leistet trotz Auspartialisierung der Artikulationszeit einen bedeutsamer Beitrag zur Prognose der Kurzzeitgedächtnisspanne. Zudem ändert sich die Höhe des Korrelationskoeffizienten kaum bei der Auspartialisierung.

Die folgenden Befunde dagegen stellen die Gültigkeit in Frage:

- Die Korrelation zwischen Gedächtnisspanne und Reproduktionszeit ist intraexperimentell sehr gering.

- In der Intervallschätzung finden sich systematische Unterschiede zwischen einigen Experimenten.
- Die Output-Zeit ist nicht universell invariant.
- Eine interindividuelle Detailanalyse zeigt, daß der postulierte negative Zusammenhang zwischen Reproduktionszeit und Gedächtnisspanne in sechs der acht Experimente nicht auftritt.

Zusammenfassend kann festgehalten werden, daß die Annahme einer universellen Invarianz der Wiedergabezeit zurückgewiesen werden muß wegen eines Effekts der linguistischen Gruppe. Innerhalb der Sprachgruppen dagegen stimmt die Hypothese von Doshier und Ma gut mit den empirischen Daten überein und scheint insgesamt eine empirisch haltvollere Hypothese zu sein als der Ansatz von Baddeley.

Die modifizierte Total-Time-Hypothese Die mTTH stimmt gut mit den empirischen Daten überein:

- Die Daten aller Supraspannen-Experimente können durch die mTTH zutreffend vorhergesagt werden.
- In den Experimenten mit Wortmaterial ist die Äquivalenzhypothese gültig.
- Auch der interexperimentelle Vergleich der Untersuchung mit Wortmaterialien spricht für die Gültigkeit der Äquivalenzhypothese, wenn auch nicht durchweg statistische Signifikanz erreicht wird.
- Der Effekt artikulatorischer Zusatzfähigkeit auf den Quotienten a/b und die Verarbeitungszeit im Langzeitgedächtnisspeicher kann für deutsche Probanden repliziert werden.

Unerwartet dagegen sind folgende Ergebnisse:

- In vier der acht Experimente (Experiment 3, 4, 5 und 8) ist die TTH gültig.
- Der Parameter a ist in den Experimenten mit Ziffern nicht interpretierbar. Die Äquivalenzhypothese gilt nicht für Experimente mit Zahlenmaterial.
- Artikulatorische Unterdrückung verlängert bei chinesischen Probanden nicht die Verarbeitungszeit im Langzeitgedächtnisspeicher, sondern verkürzt sie.

Zusammenfassend muß festgehalten werden, daß die mTTH insgesamt gültig ist, daß aber bei Zahlenmaterial spezifische Probleme bestehen.

8 DISKUSSION

Im folgenden werden die Ergebnisse, die in den Kapitel 6 und 7 dargestellt wurden, diskutiert. Der Aufbau dieses Kapitels orientiert sich an den formulierten Hypothesen, die nach inhaltlicher Kohärenz zusammengefaßt werden, und integriert die Ergebnisse der intra- und interexperimentellen Auswertung: Zuerst wird die Wirkung der aus dem Arbeitsgedächtnis abgeleiteten Effekte diskutiert. Dann wird auf die drei Invarianzhypothesen eingegangen. In einer abschließenden Diskussion werden die Einzelergebnisse einander gegenübergestellt

8.1 DISKUSSION ZUM ARBEITSGEDÄCHTNISMODELL

Aus dem Arbeitsgedächtnismodell waren verschiedene Arbeitshypothesen abgeleitet worden (Kapitel 4.1). Dabei zeigte sich eine gute Übereinstimmung der empirischen Daten mit den aus dem Arbeitsgedächtnismodell abgeleiteten Prognosen für die Effekte der linguistischen Gruppe, der Artikulatorischen Unterdrückung, der Materialien und des Präsentationsmodus.

Kritisch für die Erklärung Baddeleys zum Wortlängeneffekt (Kapitel 2.4.4) – nicht aber für das Arbeitsgedächtnismodell an sich – ist das Ausbleiben eines Interaktionseffekts zwischen linguistischer Gruppe und Artikulatorischer Unterdrückung. Der Effekt Artikulatorischer Unterdrückung wurde sowohl für deutsche als auch für chinesische Probanden nachgewiesen, so daß der fehlende Effekt nicht auf mangelnde Wirksamkeit der Sekundäraufgabe zurückgeführt werden kann: Artikulatorische Unterdrückung reduzierte bei beiden Kohorten signifikant die Höhe der Gedächtnisspanne. Diese Reduktion war aber proportional, so daß also die Gruppenunterschiede zwischen den Sprachen trotz artikulatorischer Zusatzaufgabe bestehen blieben. Studien zeigen, daß die Gedächtnisspanne von Chinesen gegenüber deutschen Probanden dann erhöht ist, wenn die Items verbalisiert werden können. Bei ausschließlich visuell-räumlich kodierbaren Materialien besteht kein Effekt der linguistischen Gruppe (Lüer et al., 1998). Das erhaltene Ergebnis eines Fortbestehens von Sprachunterschieden auch bei Blockade des Kontrollprozesses deutet also darauf hin, daß in den durchgeführten Experimenten die Funktionstüchtigkeit der Phonologischen Schleife nicht ausgeschaltet wurde. Dies wird unter folgenden zwei Annahmen diskutiert: erstens, Artikulatorische Unterdrückung blockiert nicht völlig den Kontrollprozeß oder zweitens, die Items wurden durch Abruf aus dem Langzeitspeicher phonologisch kodiert.

Die Annahme, daß Artikulatorische Unterdrückung die artikulatorischen Kontrollprozesse nicht völlig ausschaltet, wurde schon bei Untersuchungen zur Sprachproduktion formuliert (Dilger, 2000). Es gibt mehrere Möglichkeiten diese Vermutung zu spezifizieren: es könnte angezweifelt werden, ob die offene Artikulation wirklich eine adäquate Operationalisierung der latenten, inneren Artikulation ist oder es könnte angenommen werden, daß mehrere Kontrollprozesse nebeneinander bestehen (Kapitel 2). Diese Punkte würden Aufbau und Funktionsweise der Phonologischen Schleife nach Baddeley angreifen. Das Bestehen von Sprachunterschieden trotz Artikulatorischer Unterdrückung kann auch innerhalb des Arbeitsgedächtnismodells interpretiert werden: nach Baddeley können die Items durch Abruf aus dem Langzeitspeicher in die Phonologische Schleife gelangen. Durch das Verwenden von Zahlen als Stimulusmaterial sollte ein "cultural bias" minimiert werden: zwischen den Sprachen sollten also keine Unterschiede bezüglich der Langzeitgedächtnisrepräsentationen oder der "memory skills" bestehen (Ericsson & Delaney, 1999). Lass (1997) weist in ihren Studien jedoch daraufhin, daß chinesische Probanden ihre Kapazitäten effektiver nutzen. Dies könnte – besonders im Zusammenhang mit der vermuteten flexibleren Kodierung bei Chinesen aufgrund ihrer Schrift (Kapitel 8.4.3) – zu einem Fortbestehen der Sprachunterschiede trotz Blockade des Kontrollprozesses führen.

Zusammenfassend kann festgehalten werden, daß sich alle aus dem Arbeitsgedächtnismodell abgeleiteten Prognosen gut bewährt haben. Allein die Erklärung zum Wortlängeneffekt von Baddeley, Thomson und Buchanan (1975) muß zurückgewiesen werden. Dies ist aber kein kritischer Punkt für die Überprüfung der Invarianzhypothese nach Baddeley, da aus dem formalen Modell nicht die Prognose einer Eliminierung des Spracheffekts bei Artikulatorischer Unterdrückung abzuleiten ist. Die Ergebnisse zur Invarianzhypothese nach Baddeley werden im nächsten Kapitel diskutiert.

8.2 DISKUSSION ZUR INVARIANZHYPOTHESE VON BADDELEY

Die Annahme einer invarianten Haltbarkeit der Gedächtnisspur stimmt mit den empirischen Ergebnissen gut überein. Ferner zeigt die deskriptive Betrachtung der Ausprägung von Artikulationszeit und Höhe der Kurzzeitgedächtnisspanne eine gute Übereinstimmung mit den Prognosen der Invarianzhypothese: bessere Gedächtnisleistungen gehen mit höheren Artikulationsgeschwindigkeiten einher. Dennoch ist die Höhe der Korrelation dieser beiden Parameter sehr gering und auch die Intervallschätzungen decken bedeutsame Unterschiede in der Persistenz der Gedächtnisspur zwischen den Experimentreihen auf. Drei Möglichkeiten können zur Erklärung dieses Phänomens herangezogen werden:

Erstens, die niedrige Korrelation könnte ein Effekt der Operationalisierung sein: Die Listenlänge im Experiment zur Messung der Artikulationszeit überstieg durchweg den Umfang einer Gedächtnisspanne. Lass et alii (2000a) berichten in ihrer Studie, in der die Sequenzlängen zur Artikulationszeitmessung der Höhe der Gedächtnisspanne entsprachen, von einer hohen Korrelation zwischen Artikulationszeit und Gedächtnisspanne. Eventuell könnte also hier die niedrige Ausprägung des Korrelationskoeffizienten dadurch verursacht sein, daß die Annahme von Baddeley et alii (1984) nicht zutrifft, daß die Artikulationszeit unabhängig von der Listenlänge ist (Kapitel 2.4.4). Da die intraexperimentell ausgewerteten Prognosen der Gedächtnisspanne durch die Artikulationszeit meist zu einer Überschätzung der empirischen Werte führte, könnte von einer Beschleunigung der Artikulation beim Vorlesen gegen Ende der Reihe ausgegangen werden, die noch nicht beim Umfang einer Gedächtnisspanne auftritt. Zweitens, die niedrigen Korrelationskoeffizienten sprechen gegen eine materialunabhängige, interindividuelle Persistenz der Gedächtnisspur. Das Zulassen von interindividuellen Unterschieden und eine "schwache" Auslegung der Konstanten B scheint angemessen auch vor Hintergrund der Tatsache, daß Baddeley selbst immer nur Intervallschätzungen angibt. Drittens, die Daten legen die Vermutung nahe, daß außerhalb der Artikulationszeit noch weitere Determinanten der Kurzzeitgedächtnisspanne existieren. Beispielsweise zeigte die Analyse nach Hulme, daß ein signifikanter Beitrag des Langzeitgedächtnisses in Experiment 1 (Stille, Deutsch, Zahlen, Symbol) besteht. Eine weitere mögliche Determinante der Kurzzeitgedächtnisspanne könnte die Reproduktionszeit sein (Kapitel 8.3).

Zusammenfassend kann die Auslegung der Invarianzhypothese, die auch interindividuelle Unterschiede zuläßt, als bestätigt angesehen werden. Doch zeigte sich auch die Bedeutsamkeit weiterer Faktoren, wie der Einfluß des Langzeitgedächtnisses oder die Reproduktionszeit, die Thema des nächsten Kapitels ist.

8.3 DISKUSSION ZUR INVARIANZHYPOTHESE VON DOSHER UND MA

Deskriptiv konnte der - durch die Invarianzhypothese von Doshier und Ma postulierte - Zusammenhang zwischen Reproduktionszeit und Höhe der Gedächtnisspanne mit zwei Ausnahmen nachgewiesen werden. Da in diesen beiden Experimenten der Stichprobenumfang sehr gering war, ist ein Stichprobeneffekt sehr wahrscheinlich.

Ferner wurde eine Überlegenheit der Reproduktionszeit zur Prognose der Kurzzeitgedächtnisspanne im Vergleich zur Artikulationszeit nachgewiesen. Reproduktionszeit und Artikulationszeit stellen nach Lass et alii (2000a) äquivalente Anforderungen an die Motorik oder – meines Erachtens – an Sprachproduktionsprozesse im

Allgemeinen, so daß ähnliche Wirkungen der beiden Parameter zu erwarten sind. Darüber hinaus aber wachsen die Effekte im Mittel prozentual bei den On-Line Messungen an. Beispielsweise ist für den Materialeffekt zwischen phonologisch unähnlichen Worten und Zahlworten die Artikulationsgeschwindigkeit der Klasse „Worte“ um acht Prozent erhöht, bei der Reproduktionsgeschwindigkeit dagegen um 20 Prozent. Deutlicher ist dieses Ergebnis beim Sprachvergleich: Die Artikulationszeit im Deutschen ist im Mittel um 21 Prozent langsamer als im Chinesischen, die Reproduktionszeit dagegen um 61 Prozent. Die Varianzaufklärung für die Prognose der Kurzzeitgedächtnisspanne durch die Reproduktionszeit ist weit besser als durch die Artikulationszeit. Ferner ändert sich kaum die Höhe des Zusammenhangs zwischen Reproduktionszeit und Gedächtnisspanne, wenn nach dem Modell von Hulme die Artikulationszeit aus der Reproduktionszeit auspartialisiert wird.

Die intra- und interexperimentell nachgewiesene, signifikante Korrelation steht in Kontrast zur Studie von Lass et alii (2000a), die keinen systematischen Zusammenhang nachweisen konnten. Es bestehen - im Vergleich mit vorliegender Arbeit – Unterschiede in der Operationalisierung der Variablen: Bei Lass et alii (2000a) wurde die Reproduktionszeit pro Item ohne Berücksichtigung des Präparationsintervalls und unter Ausschluß langer Interstimulusintervalle ermittelt. Es handelt sich also nicht um eine Operationalisierung im Sinne von Doshier und Ma (1998). In diese Studie ging nach der Definition von Doshier und Ma (1998) das insgesamt benötigte Zeitintervall von Darbietung des Wiedergabesignals bis zur Beendigung der Reproduktion ein. Die unterschiedlichen Operationalisierungen könnten also verantwortlich für die widersprüchlichen Ergebnisse sein.

Unabhängig von der Prognosekraft der Reproduktionszeit, mußte das Postulat einer universell invarianten Gesamtwiedergabe-Zeit zurückgewiesen werden. Der statistische Test deckte bedeutsame Unterschiede zwischen den linguistischen Gruppen auf. Die Daten legen die Modifikation nahe, daß nur innerhalb einer linguistischen Gruppe über alle experimentellen Manipulationen eine konstante Wiedergabezeit zu erwarten ist. Dies könnte auch erklären, warum die hier ermittelten, empirischen Werte der Wiedergabezeit nicht mit den von Doshier und Ma an englischen Probanden beobachteten Werten übereinstimmen. Auf die Frage, warum sich die Gesamtwiedergabe-Zeiten zwischen den verschiedenen linguistischen Gruppen unterscheiden, kann mit der Invarianzhypothese selbst geantwortet werden: Die zeitliche Begrenzung der Wiedergabezeit sollte dadurch entstehen, daß sich während des Wiedergabeintervalls Interferenz anhäuft. Da beispielsweise im Chinesischen schneller artikuliert wird und die Pausenzeiten kürzer sind (Lüer et al., 1998), könnte der Grad an Interferenz, der den Leistungsabbruch bewirkt, schneller erreicht werden als bei

deutschen Probanden. Unklar bleibt, worin sich die linguistischen Gruppen über die Schnelligkeit der Reproduktionsgeschwindigkeit hinaus weiter unterscheiden (Kapitel 8.5).

Die Beobachtung von Cowan (1994), daß die Reproduktionszeit interindividuell invariant ist, konnte nicht bestätigt werden. Bei Gegenüberstellung von Probanden mit hoher, respektive niedriger Wiedergabezeit ergab sich der für die Hypothese von Doshier und Ma kritische Befund, daß in fünf der sieben Experimente höhere Gedächtnisleistungen mit langsameren Reproduktionszeiten verbunden waren. An dieser Stelle könnte eine Parallele zu dem Ansatz zur Zentralen Exekutive (Kapitel 2.2.3) von Engle et alii (1999) gezogen werden, in dem den Probanden mit hoher Arbeitsgedächtniskapazität während der Wiedergabephase neben Prozessen automatischer Aktivationsausbreitung auch intentionale Suchprozesse zugeschrieben werden. Falls das auf das Paradigma der Kurzzeitgedächtnisspanne übertragen werden kann, dann würde eine verlängerte Reproduktionszeit durch intentionale Suchprozesse als Funktion der Zentralen Exekutive entstehen, die die Gedächtnisleistung erhöhen.

Zusammenfassend stellt die Invarianzhypothese von Doshier und Ma einen Fortschritt gegenüber der Hypothese von Baddeley dar:

- Erstens wird mehr Varianz an dem Zustandekommen der Gedächtnisleistung aufgeklärt.
- Zweitens ist die modifizierte Invarianzhypothese von Doshier und Ma mit einer größeren Anzahl von Befunden vereinbar, da sie auch zur Erklärung der Effekte Phonologischer Ähnlichkeit und Artikulatorischer Unterdrückung herangezogen werden kann.
- Drittens leistet die Reproduktionszeit auch nach Auspartialisierung der Artikulationszeit einen bedeutsamen Beitrag zur Varianzaufklärung der Kurzzeitgedächtnisspanne.

Dennoch wurde in dieser Arbeit auch über die - für die universelle Gültigkeit der Hypothese - kritischen Befunde der Unterschiede in der Gesamtwiedergabe-Zeit zwischen linguistischen Gruppen und dem interindividuellen Vergleich mit längeren Reproduktionszeiten trotz höherer Gedächtnisspanne berichtet.

8.4 DISKUSSION ZUR MTTH

Schwerpunkt der vorliegenden Arbeit war die Untersuchung der mTTH unter verschiedenen Bedingungen. Aufgrund der vielfältigen Ergebnisse erfolgt eine Unterteilung der

Diskussion: Zuerst wird die Annahme der mTTH und die Abgrenzung zur TTH dargestellt, es folgen Interpretationen der Ergebnisse zur Äquivalenzhypothese sowie zur Verarbeitungszeit im Langzeitgedächtnisspeicher. Den Abschluß bildet eine zusammenfassende Diskussion.

8.4.1 DISKUSSION DER GÜLTIGKEIT DER MTTH

Die intraexperimentellen Auswertungen bestätigen die Gültigkeit der mTTH. Die Anzahl der Lernversuche im Supraspannen-Experiment pro Untersuchungsreihe können durch die aus der mTTH abgeleiteten Prognosen zutreffend geschätzt werden.

Die mTTH reduziert die TTH. Das bedeutet, daß das Reducens einen höheren empirischen Gehalt hat und sowohl neue als auch genauere Prognosen erlaubt (Popper, 1965). Der logische Spielraum ist für das Reducens erhöht, das heißt, daß mehr Ereignisse definiert werden können, bei deren Existenz die Hypothese verworfen wird. Für die Vergleiche zwischen mTTH und TTH bedeutet das einerseits, daß die mTTH mit Befunden vereinbar ist, die auch kompatibel mit der TTH sind und andererseits, daß andere Befunde nur durch die mTTH nicht aber auch durch die TTH erklärt werden können. Unerwartet zeigen sich in den vorliegenden Ergebnissen beide Facetten des Reduzierens: vier Experimente (1, 2, 6 und 7) sind nur mit der mTTH vereinbar. Bei den anderen vier Experimenten (3, 4, 5 und 8) liefern sowohl die mTTH als auch die TTH zutreffende Prognosen. Die Gültigkeit der TTH könnte durch verschiedene Merkmale des Versuchsaufbaus und/oder die Stichprobengröße verursacht sein:

Sequenzlänge Für die Unterschätzung von Parameter a könnte der Versuchsaufbau hinsichtlich der Sequenzlänge verantwortlich sein: in Vorexperiment 1 wurde gezeigt, daß zehn Zahlensymbole keine adäquate Sequenzlänge im Supraspannen-Experiment darstellen. Eine Erhöhung auf 15 Items pro Reihe schien optimal zu sein, da nur drei deutsche Teilnehmer der Studie, respektive ein spanischer Proband direkt beim ersten Lernversuch die Reihe vollständig und in der richtigen Reihenfolge wiedergeben konnten. Anders verhält es sich für die Untersuchung von chinesischen Probanden: Bei Experiment 4 (Stille, Chinesisch, Zahlen, Symbol) konnten bereits bei einer Darbietungszeit von zehn Sekunden zwei Probanden die 15 Items korrekt wiedergeben, bei der Präsentationszeit von dreißig Sekunden waren es acht Versuchspersonen. Vier Probanden bei zwanzig Sekunden Darbietungszeit und fünf Probanden bei dreißigsekündigem Präsentationsintervall merkten sich im ersten Lernversuch die Reihe in Experiment 5 (Artikulatorische Unterdrückung, Chinesisch, Zahlen, Symbol). In Experiment 8 verhält es sich entsprechend: In

Vorexperiment 4 wurde nachgewiesen, daß ein Umfang von 15 Zahlworten pro Sequenz zu schwer ist. Das Erlernen der auf zehn reduzierten Anzahl von Items scheint für zehn Probanden zu einfach gewesen zu sein. Gegen die Erklärungsmöglichkeit einer zu leichten Aufgabe sprechen zwei Befunde. Erstens, die spanische Gruppe zeigte im Vergleich zu den deutschen Probanden schlechtere Kurzzeitgedächtnisleistungen - für die deutsche Gruppe aber mußte trotz identischen Gesamtinformationsgehalts im Supraspannen-Experiment die TTH zurückgewiesen werden. Zweitens, ein Vergleich mit den beobachteten Kurzzeitgedächtnisspannenwerten zeigt, daß in keinem Experiment bei keinem Probanden der Umfang der Sequenzreihe des Supraspannen-Experiments der Höhe der Kurzzeitgedächtnisspanne entsprach. Auch wenn insgesamt 7.6 Prozent der Probanden direkt beim ersten Versuch die Sequenz richtig wiedergaben, so überstieg doch für die restlichen 92.4 Prozent der Umfang der Sequenz die Kurzzeitgedächtnisleistung.

Vorhersehbarkeit der Sequenzlänge und Darbietungsform Zwei weitere mögliche Ursachen des Ergebnisses, daß für Parameter a in vier Experimenten ein Wert kleiner Eins ermittelt wurde, könnten die Berechenbarkeit der Reihenzlänge und die simultane Präsentation der Stimuli sein. Im Subspannen-Experiment wurden die Items sukzessive dargeboten. Wie Nachbefragungen ergaben, durchschauten die Probanden das adaptive Vorgehen nicht. Demnach war der Umfang der Reihe nicht vorhersehbar. Im Supraspannen-Experiment dagegen wurde die Sequenzlänge in der Instruktion angekündigt, die Items wurden simultan präsentiert. Da die Blickbewegungen nicht registriert wurden, kann nicht geprüft werden, ob die Probanden seriell jedes Items bei jeder Darbietung enkodierten. Dies scheint aber unwahrscheinlich. Es wurde beobachtet, daß Probanden beim ersten Versuch den Anfang der Sequenzreihe korrekt wiedergaben, sich beim nächsten Versuch aber nur an die letzten Items erinnerten. Die en bloc Darbietung erleichtert zudem die Chunkbildung. Für diese Erklärungsmöglichkeit spricht der Präsentationseffekt im Supraspannen-Experiment, wie er im Vergleich der Vorexperimente 1 und 4 aufgedeckt wurde (Kapitel 8.4.4).

Darbietungszeit Eine weitere Erklärung für das Bestehen der TTH in vier der acht Experimente könnte in der Wahl der Darbietungszeiten liegen. Gerade für Zahlenmaterial (Kapitel 8.4.5), mit dessen Umgang die Probanden wohlvertraut sind, waren wohl die Darbietungszeiten zu lang: Innerhalb einer hyperbolischen Lernfunktion sind eventuell nur Zeitpunkte realisiert worden, die Abschnitte des insgesamt hyperbolischen Verlaufs abbilden, in denen ein linearer Trend nachzuweisen ist. Die alternativen Auswertungen zur Äquivalenzhypothese für die Experimente, in denen die TTH gilt, konnten keine Anpassung an die Asymptote nachweisen, doch könnte dies ein Effekt der Parameterschätzung sein.

Auf eine zweite potentielle Problematik bei der Wahl der Darbietungszeiten wird in Kapitel 8.4.4 eingegangen.

Stichprobengröße Für diesen Punkt zur Erklärung der niedrigen Ausprägung von Parameter a spricht, daß bei Experiment 1 mit dem Stichprobenumfang von 57 Probanden a ebenfalls kleiner als 1 war. Erst nach einer Erweiterung der Stichprobe auf insgesamt 87 Versuchspersonen wurde das theoretisch abgeleitete Minimum überschritten.

8.4.2 DISKUSSION DER ÄQUIVALENZHYPOTHESE

Die Asymptote der Lernfunktion, in deren Definition Parameter a enthalten ist, wird innerhalb der Theorie der mTTH mit der Kurzzeitgedächtnisspanne in bit gleichgesetzt. Diese Annahme kann nur für die Experimente mit ähnlichen und unähnlichen Worten beibehalten werden (Experiment 6 & 7). Bei der Überprüfung, ob sich die signifikanten Unterschiede in der Kurzzeitgedächtnisspanne bei phonologisch un-, respektive ähnlichem Wortmaterial auch in der Asymptote widerspiegeln, ergab sich Evidenz.

Zurückgewiesen wird die Äquivalenzhypothese in allen Experimenten mit Zahlenmaterial: Die Asymptote überschätzt systematisch die Höhe der Kurzzeitgedächtnisspanne. Dies gilt auch für die Experimente mit Zahlenmaterial, in denen die TTH zurückgewiesen werden mußte. Im folgenden werden fünf Faktoren diskutiert, die mit diesem Materialeffekt in Verbindung gebracht werden können: Einfluß der Lexikalität, des Präsentationsmodus, Möglichkeit zur doppelten Kodierung, Einfluß der semantischen Nähe und Möglichkeit zur Gruppierung. Die Erklärungsansätze schließen sich nicht gegenseitig aus, sondern ergänzen sich.

Lexikalität Das Scheitern der Äquivalenzhypothese kann nicht durch den Faktor „Lexikalität“ der Items verursacht sein, denn sowohl Zahlen als auch Worte sind im Langzeitgedächtnisspeicher repräsentiert.

Präsentationsmodus Der Präsentationsmodus als Ursache für das Scheitern der Äquivalenzhypothese bei Zahlenmaterial kann ebenfalls nicht zur Erklärung des Materialeffekts herangezogen werden. Zahlensymbole können phonologisch oder visuell als Bild kodiert werden. Falls die Überschätzung der Kurzzeitgedächtnisspanne durch die Asymptote aufgrund des Zusammenwirkens der beiden Subsysteme - Phonologische Schleife und Visuell-räumlicher Notizblock - zustande gekommen wäre, so müßte bei schriftlicher Darbietung der Zahlen als Worte die Überschätzung verschwinden, da eine

räumlich-visuelle Verarbeitung für vier-, respektive fünfstellige Buchstabenkombinationen der Zahlworte im Gegensatz zu einem Symbol pro Item erschwert sein sollte. Die empirische Evidenz verweist jedoch darauf, daß unabhängig von der Darbietungsart bei Zahlenmaterial die Kurzzeitgedächtnisspanne durch die Asymptote systematisch überschätzt wird.

Möglichkeit zur doppelten Kodierung Nachdem der Präsentationsmodus an sich als Erklärungsmöglichkeit des Materialunterschieds verworfen wurde, bleibt die Alternative, daß die Möglichkeit zur doppelten Kodierung entscheidend ist. Nach der Dual-Coding-Theory von Paivio (1986) besteht eine Superiorität im Gedächtnistest für die Stimulusart, die sowohl visuell als auch phonologisch kodiert werden kann. Für Zahlenmaterial scheint per se die Möglichkeit zur doppelten Kodierung zu bestehen. Bei Wortmaterial dagegen ist die Bildhaftigkeit, die in Vorexperiment 7 für phonologisch ähnliche und unähnliche Items auf einer siebenstufigen Rating-Skala erhoben wurde, nur auf einem mittleren Skalenniveau angesetzt. Für die Relevanz der Dual-Coding-Theory spricht das Resultat, daß auch bei Verwendung geometrischer Figuren die Äquivalenzhypothese zurückgewiesen werden mußte (Schiffer, 2000). Kritisch für diese Erklärung ist der Befund, daß bei Artikulatorischer Unterdrückung nach Baddeley keine phonologischen Repräsentationen bei visueller Darbietung gebildet werden können. Somit müßte bei dieser Sekundäraufgabe die Äquivalenzhypothese bestätigt werden. Dies zeigte sich jedoch weder bei Schiffer (2000) noch in Experiment 2 und 5. Zwei Erklärungsmöglichkeiten sind vorstellbar: Erstens kann die Wirkung Artikulatorischer Unterdrückung als absolute Blockade der Graphem-Phonem-Transformation in Frage gestellt werden und zweitens kann der fehlende Effekt Artikulatorischer Unterdrückung durch Abruf von Repräsentationen aus dem Langzeitgedächtnisspeicher entstanden sein (Kapitel 8.1).

Semantische Nähe Eine weitere Erklärungsmöglichkeit zum Materialeffekt ist, daß Worte als Stimulusmaterial semantisch unrelatiert sind, während Zahlen eine eigene Kategorie bilden. Dieser Effekt „diffuse versus eindeutige“ Kategorie sollte durch die Instruktion und das Artikulationszeit-Experiment minimiert werden: Vor Beginn der Sitzung wurden die Probanden in der Einführung (Anhang I) mit dem Stimulusmaterial vertraut gemacht. Dabei könnten qualitative Unterschiede zwischen den Materialien entstanden sein. Man kann sich meines Erachtens relativ sicher sein, daß bereits nach dem ersten Durchlesen des Zahlenmaterials die Ziffern als Stimulusmaterial memoriert wurden. Dagegen scheint es unwahrscheinlich, daß sich die Probanden direkt beim Lesen die zehn phonologisch un-, respektive ähnlichen Worte einprägten. Obwohl in der Untersuchungssitzung auf diese Einführung hin das Artikulationszeit-Experiment folgte, könnte es möglich sein, daß für die

Probanden der beiden Experimente mit Wortmaterial der Suchraum nicht in gleichem Ausmaß wie beim Zahlenmaterial definiert war. Diese Erklärungsmöglichkeit kann jedoch nicht die einzige sein, denn in den Experimenten mit figuralem Material (Schiffer, 2000) mußten die Probanden vor Beginn des Supraspannen-Experiments die Figuren und deren Bezeichnung lernen. Man kann also relativ sicher sein, daß der Itempool klar definiert war. Dennoch konnte auch für geometrische Figuren die Äquivalenzhypothese nicht beibehalten werden.

Möglichkeit zur Gruppierung Durch Zugehörigkeit zu einer Kategorie könnte durch Assoziationen zwischen den einzelnen Items das Herstellen semantischer Gruppierungen erleichtert werden. Bei Zahlenmaterial bestehen mehr Gruppierungsmöglichkeiten, da Zahlensequenzen im Alltag sehr präsent sind, während Wortaufzählungen ohne syntaktische Verknüpfungen eine Ausnahme darstellen. Auch bei geometrischen Figuren läßt sich Alltagsnähe vermuten. Ericsson und Delaney (1999) postulieren zwei Folgen einer dauernden Exposition mit einer bestimmten Stimulusart: Erstens werden die Wissensseinheiten im Langzeitspeicher informativer und das Verknüpfungssystem dichter. Zweitens ist anzunehmen, daß durch forcierte Exposition spezifische Gedächtnisstrategien erworben werden, die nicht nur effektiv und effizient, sondern auch unter Zeitdruck schnell anwendbar sind. Diese beiden Faktoren ermöglichen ein Expertentum im Umgang mit dem Material, das die typischen Kapazitätsbegrenzungen des Einprägens und Erinnerns transzendiert. Durch das Postulat eines Expertentums der Probanden, spezifisch für Ziffern, könnte der Effekt zwischen den verschiedenen Materialien erklärt werden. Wie bereits in Kapitel 8.4.1 angesprochen wurde, scheint die en bloc Darbietungsmodalität einer intraindividuell invarianten Sequenz den Einsatz von Strategien zusätzlich zu forcieren.

8.4.3 DISKUSSION ZUR VERARBEITUNGSZEIT IM LANGZEITGEDÄCHNISSPEICHER

Zur Verarbeitungszeit im Langzeitgedächtnisspeicher waren zwei Hypothesen formuliert worden. Einerseits wurde bei bestehender Möglichkeit zur Artikulation kein Einfluß der Experimentalbedingungen erwartet und andererseits wurde bei Verhinderung der Artikulation eine Verlängerung der Verarbeitungszeit im Langzeitgedächtnisspeicher prognostiziert. Beide Annahmen sind nicht durchweg vereinbar mit den empirischen Ergebnissen, die im folgenden diskutiert werden.

Der Effekt der linguistischen Gruppe Die Ergebnisse zum Effekt der linguistischen Gruppe auf die Verarbeitungszeit im Langzeitgedächtnisspeicher sind heterogen. Für die deutsche und die chinesische Kohorte ergaben sich Zeitangaben für die Verarbeitungszeit, die auch gut mit

den Angaben von Schiffer (2000) übereinstimmen. Bei der spanischen Kohorte dagegen zeigte sich eine signifikante Verlängerung. Dies ist insofern erstaunlich, als daß auch für die spanische Gruppe die Möglichkeit zur Artikulation gegeben war. Der Effekt kann auch nicht aufgrund von Operationalisierungsproblemen oder Materialunterschieden entstanden sein. Eine Aussage, ob die Verarbeitungszeit bei Spaniern generell verlängert ist oder ob es sich um einen Effekt der Stichprobe handelt, kann aus diesen Ergebnissen nicht abgeleitet werden, da nur ein einziges Experiment mit dieser Kohorte durchgeführt wurde.

Der Effekt Artikulatorischer Unterdrückung Die Verarbeitungszeit im Langzeitgedächtnisspeicher sollte bei einer artikulatorischen Zusatzaufgabe verlängert sein. Diese Annahme konnte für chinesische Probanden nicht bestätigt werden. Für deutsche Probanden scheint dagegen die Verhinderung der Artikulation eine hinreichende Bedingung zur Verlängerung der Verarbeitungszeit im Langzeitgedächtnisspeicher zu sein (Kapitel 3.3). Der Effekt der linguistischen Gruppe kann in Zusammenhang mit den unterschiedlichen Schriftsystemen gebracht werden. Während im Deutschen eine Lautschrift etabliert ist und nahezu jedem Graphem ein Phonem zugeordnet ist, wird im Chinesischen eine ideographische Schrift eingesetzt: Meist wird ein Wort durch ein einziges, aus verschiedenen Radikalen bestehendes, Symbol ausgedrückt. Faktum ist, daß diese Schrift eher visuell aufgebaut und in einem wesentlich geringeren Umfang phonetisch angelegt ist. Lass (1997) verweist für Kurzzeitgedächtnisleistungen auf die Relevanz der Schrift. Probanden mit deutscher Muttersprache sind an die phonetische Präsentation gewöhnt, die eine phonologische Kodierung nahelegt, oder nach Lass et alii (2000b, S. 85) „Somit stützen diese Ergebnisse die eingangs diskutierte Tendenz von Probanden, wenn denn möglich, über die Phonologische Schleife und nicht über den Notizblock zu memorieren.“ Zur Erklärung der vorliegenden Befunde könnten zwei Vermutungen aufgestellt werden: Erstens könnten Chinesen wegen ihrer Schrift bezüglich der Kodierung flexibler sein. Nach dem Arbeitsgedächtnismodell von Baddeley ist keine Wirkung einer artikulatorischen Zusatzaufgabe auf visuell verarbeitetes Material zu erwarten und somit auch keine Verlängerung der Verarbeitungszeit im Langzeitgedächtnisspeicher. Zweitens könnte die Nutzung der limitierten Kapazitäten bei Chinesen effizienter sein wegen des verstärkten Einsatzes doppelter Kodierung. Für diese Annahme spricht, daß Chinesen auch bei figuralem, verbalisierbarem Material höhere Gedächtnisleistungen zeigen als deutsche Probanden (Lass et al., 2000b). Die Verhinderung der Artikulation wäre also nur dann eine hinreichende Bedingung zur Verlängerung der Verarbeitungszeit, wenn eine alternative, beispielsweise visuelle, Kodierung unmöglich ist oder von den Probanden nicht genutzt wird.

Der Materialeffekt Der Vergleich zwischen Parameter b der Experimente mit phonologisch un-, respektive ähnlichen Worten zeigt deskriptiv eine Verlängerung der Verarbeitungszeit, der Quotient der beiden materialspezifischen Konstanten der mTTH unterscheidet sich nicht signifikant. Wie in Vorexperiment 8 sichergestellt wurde, kann es sich bei dem proportionalen Zuwachs der Parameter a und b der mTTH bei steigendem Grad phonologischer Ähnlichkeit nicht um einen Effekt der semantischen Ähnlichkeit zwischen den Wortpaaren handeln. Während die Erhöhung von Parameter a bei Phonologischer Ähnlichkeit erwartet wurde, ist die Verlängerung der Verarbeitungszeit bei klangähnlichen Items erstaunlich, denn innerhalb des Arbeitsgedächtnis-Modells werden die Effekte Irrelevanter Sprache und Phonologischer Ähnlichkeit als äquivalente Methoden zur Erzeugung von Interferenz im Phonologischen Speicher interpretiert. Bei der experimentellen Bedingung Irrelevanter Sprache aber wurde keine Verlängerung der Verarbeitungszeit nachgewiesen (Hamm, 1998). Eine Erklärung wäre durch einen Vergleich mit dem Paradigma des lexikalischen Zugriffs abzuleiten. In der Sprachproduktionsforschung wird die Zeit des lexikalischen Zugriffs auf etwas 600 Millisekunden datiert (Levelt, Roelofs & Meyer, 1999). Der hier erhaltene Wert der Verarbeitungszeit im Langzeitgedächtnisspeicher liegt etwa 150 Millisekunden höher. Dies könnte in Zusammenhang mit der Untersuchungsmethode gesehen werden: Die Messung basiert auf Reproduktionsleistungen. Falls zwischen der Verarbeitungszeit im Langzeitgedächtnisspeicher – als Aktivierung von Langzeitgedächtnisrepräsentationen – und der Zeit des lexikalischen Zugriffs eine Korrelation nachgewiesen werden könnte, so wäre der aufgedeckte Unterschied zwischen den Bedingungen Irrelevanter Sprache und Phonologischer Ähnlichkeit vor dem Hintergrund einer unmittelbaren – und nicht zeitverzögert erfolgenden – Kommunikation, respektive von Interferenzen auf lexikalischer Ebene verständlich.

Der Effekt des Präsentationsmodus Die Verarbeitungszeit im Langzeitgedächtnisspeicher steigt, wenn Zahlen nicht wie gewohnt als Symbole, sondern als Worte dargeboten werden. Der Parameter der Verarbeitungszeit ist die einzige Variable für die ein Effekt des Präsentationsmodus nachzuweisen ist: er könnte durch die geringe Vertrautheit mit Zahlworten entstanden sein, wie auch der Lexikalitätseffekt. Ein Erklärungsansatz ergibt sich durch die Theorie von Engle (1996), der von einer interindividuell invarianten und automatischen Aktivierungszeit von Langzeitgedächtnisrepräsentationen berichtet, die mit circa 600 Millisekunden angegeben wird (Kapitel 2.2.3). Dieser Zeitangabe liegen wiederum im Gegensatz zu den hier ermittelten Werte der Verarbeitungszeit keine Reproduktionsleistungen zugrunde. Falls sich in Studien eine Korrelation des Parameters b mit dem Ansatz von Engle (1996) belegen läßt, würde der Effekt des Präsentationsmodus

als Phänomen des Bemühens um Umkodierung von Zahlworten in das gewohnte Format der Zahlensymbole entstehen (Kapitel 8.4.4).

8.4.4 ZUSAMMENFASSENDE DISKUSSION DER MTTH

Die mTTH hat sich in allen acht Experimenten bewährt, jedoch war auch die TTH in vier Experimenten gültig. Dieser Effekt wird auf den Versuchsaufbau (Sequenzlänge, Berechenbarkeit, en bloc Darbietung und Wahl der Darbietungszeit) zurückgeführt. Der materialspezifische Effekt bei der Äquivalenzhypothese wird erklärt durch spezifische Eigenschaften von Zahlenmaterial, wie etwa dem Bestehen schnell aktivierbarer und reichhaltiger Wissensseinheiten im Langzeitgedächtnisspeicher oder der semantischen Nähe, respektive der leichten Gruppierungsmöglichkeit bei Ziffernmaterial. Die Auswertungen zur Langzeitgedächtnisspanne sprechen ebenfalls für den Einsatz von Mnemostrategien bei Zahlen. Für die Strategie des „Umkodierens“ von Zahlworten in Zahlensymbole spricht der Vergleich der Langzeitgedächtnisspanne dieser beiden Experimente (Experiment 1 & 8). Der Wert der Langzeitgedächtnisspanne bei Zahlworten gleicht bei kurzer Präsentationsphase numerisch eher dem Wert des Experiments mit phonologisch unähnlichen Worten. Bei einer Darbietungszeit von zwanzig Sekunden dagegen erhöht sich die Langzeitgedächtnisspanne von Zahlworten und nähert sich dem analogen Wert von Zahlensymbolen an. Dieser Effekt kommt meines Erachtens deswegen zustande, weil es den Probanden bei einer zehn-sekündigen Darbietungszeit der zehn alphabetischen Ziffernitems nicht gelingt, das Zahlenmaterial umzukodieren und von den erlernten Fertigkeiten zu profitieren. Zwanzig Sekunden Präsentation dagegen scheint auszureichen, um aus dem gewohnten Umgang mit Zahlen Nutzen zu ziehen. Die Tatsache, daß die Langzeitgedächtnisspanne von Zahlworten bei einer Darbietungszeit von dreißig Sekunden die Spanne von Zahlensymbolen übertrifft, kann als Epiphänomen des zusätzlichen Bemühens um Umkodierung des Stimulusmaterials interpretiert werden. Folgt man dieser Argumentationsweise, so bietet sich folgende Möglichkeit zur Interpretation an:

Es könnte ein Zeitpunkt t_j existieren, ab dem Strategien effektiv eingesetzt werden können. Die konkrete Ausprägung von t_j könnte in Abhängigkeit der Materialart ebenso variieren wie die Effektivität der Strategien. Die Annahme, daß der Zeitpunkt t_j materialabhängig nicht konstant bleibt, könnte einen weiteren Effekt der Versuchsanordnung aufdecken: Es kann vermutet werden, daß bei Wortmaterialien nur Darbietungszeiten realisiert wurden, die den Zeitpunkt t_j unterschreiten. Bei Zahlworten dagegen scheinen zehn Sekunden als Darbietungszeit vor t_j zu liegen, zwanzig und dreißig Sekunden danach. Für Zahlensymbole ergibt sich ein anderes Bild: Bereits bei einer Darbietungszeit von zehn Sekunden ist die

Langzeitgedächtnisspanne extrem hoch (10.28 bit). Dies läßt vermuten, daß bereits dieses Präsentationsintervall den Zeitpunkt t_j überschreitet. Unter der Prämisse der Existenz von t_j läßt sich die Überschätzung der Kurzzeitgedächtnisspanne in bit durch die Asymptote als Wirkung leistungssteigernder Techniken sehen, die durch den Versuchsaufbau forciert werden. Man kann annehmen – neben der eingangs formulierten Möglichkeit, daß die Präsentationszeiten zu lang waren -, daß man die Effektivitätssteigerung der Strategien nicht in diesem Ausmaß nachweisen hätte können, wenn die Darbietungszeiten durchweg länger gewesen wären und somit der Nachteil der massierten Exposition gegenüber der verteilten auftreten hätte können.

8.5 ZUSAMMENFASSUNG UND AUSBLICK

Im folgenden werden die fünf wichtigsten Ergebnisse dieser Arbeit kurz zusammengefaßt und mit einem Ausblick auf weiterführende Studien verbunden.

Transkulturelle Untersuchungen Die Daten legen nahe, daß der Faktor „Sprache“ mit weiteren Faktoren, wie Effizienz der Transformationsprozesse, Schrift und flexibler Kodierung konfundiert ist und daß aus diesem Grund sprachabhängige Unterschiede nicht eindeutig auf abweichende Artikulationszeiten attribuiert werden können. Einfluß scheinen diese Faktoren besonders auf die Reproduktionszeit zu nehmen, da sich in diesem Parameter die Unterschiede zwischen den Sprachgruppen akkumulieren.

Zahlenmaterial Zahlenmaterial scheint kein geeignetes Untersuchungsmaterial für Aussagen über theoretische Implikationen des Arbeitsgedächtnis-Modells zu sein. Ziffern verweisen auf viele und stark assoziierte Wissenseinheiten im Langzeitspeicher und bieten den Vorteil einer doppelten Kodierbarkeit. Zudem treten sie meist nicht isoliert auf, sondern in Kombinationen, so daß eine große Vertrautheit im Umgang mit Strategien besteht. Das bedeutet, daß bei Bearbeitung und Erhaltung von Ziffern im Gedächtnis alle drei Systeme des Arbeitsgedächtnis-Modells involviert sein können. Spezifische Aussagen über einzelne Systemeigenschaften sind unter dieser Voraussetzung nicht zu leisten.

Die Invarianzhypothese von Baddeley Die empirischen Beobachtungen sprechen für die Annahme einer interindividuell variierenden, materialunabhängig invarianten Persistenz der Gedächtnisspur. Weitere Auswertungen belegen die Bedeutsamkeit des Langzeitgedächtnisspeichers auf die Kurzzeitgedächtnisspanne und die Relevanz von Prozessen, die während der Wiedergabe ablaufen.

Die Invarianzhypothese von Doshier und Ma Die Hypothese von Doshier und Ma verspricht wegen der verbesserten Prognosegüte, der bleibenden Aussagekraft trotz Ausparialisierung der Artikulationszeit und des größeren Anwendungsbereich einen Fortschritt gegenüber der Invarianzhypothese von Baddeley, doch zeigen sich auch hier zwei kritische Befunde. Erstens, es bestehen systematische Unterschiede zwischen den Sprachgruppen. Zweitens, die Auswertungen zu dem Ansatz von Cowan sprechen - folgt man der hier vorgeschlagenen Interpretation – für einen Einfluß der Zentralen Exekutive auf die Kurzzeitgedächtnisspanne.

Für weiterführende Untersuchungen zu den Invarianzhypothesen des Kurzzeitgedächtnisses bietet sich der Ansatz von Hulme et alii (1999) an, der nicht nur die globale Reproduktionszeit pro Item berücksichtigt, sondern eine Aufspaltung dieser Zeit in Präparationsintervall, Artikulationszeit und Pausenintervalle vorsieht. Ebenso sollte dem Ansatz von Engle (1996) nachgegangen werden und eine Verbindung zwischen der Höhe der Arbeitsgedächtnisspanne, den Gedächtnisleistungen und den Pausenintervallen hergestellt werden. Vermutet werden könnte erstens, daß mit einem größerem Umfang des Itempools verstärkt Anforderungen an Inhibitionsmechanismen gestellt werden und deswegen die Pausenlängen steigen, und zweitens, daß Probanden mit hoher Arbeitsgedächtnisspanne längere Pausenzeiten benötigen, dafür aber höhere Gedächtnisspannenwerte erzielen, weil während der Interstimulusintervalle ein verstärktes Auftreten intentionaler Suchprozesse vermutet wird.

Die mTTH Die Gültigkeit der mTTH wurde in allen Experimenten nachgewiesen. Die Verarbeitungszeit im Langzeitgedächtnisspeicher wird durch die Faktoren „Lexikalität“, „Phonologisch Ähnlichkeit“ und „Präsentationsmodus“ beeinflusst. Bei Prüfung der mTTH bestand das Problem, daß in vier Experimenten für Parameter a das theoretische Minimum von Eins nicht erreicht wurde. Für Zahlenmaterial konnte zudem die Annahme der Äquivalenzhypothese nicht akzeptiert werden. Dies wird auf die Effekte der doppelten Kodierung und der Möglichkeit zur Gruppierung zurückgeführt. Weitere Studien zur mTTH und ihrer Abgrenzung zur TTH sollten mit dem vorgeschlagenen, modifizierten Versuchsdesign des Supraspannen-Experimentes durchgeführt werden. Die Existenz eines Zeitpunktes t_j , ab dem Strategien effizient eingesetzt werden können, könnte durch eine Wiederholung von Experiment 8 (Stille, Deutsch, Zahlen, Schrift) und zusätzlicher Belastung der Zentralen Exekutive überprüft werden.

Summa summarum hat die Prüfung der Invarianzhypothesen die in Kapitel 3 zitierten Erwartungen an quantitativ-mathematischen Modelle (Kintsch et al., 1999, S. 437) erfüllt: Die Notwendigkeit weiterer Forschung („draws attention to gaps in the research evidence“, Kintsch et al., 1999, S. 437) im Bereich des Kurzzeitgedächtnisses wurde in Bezug auf das Präparationsintervall und die Pausenzeiten aus den Invarianzhypothesen von Baddeley sowie von Doshier und Ma abgeleitet, während die Befunde zur mTTH bezüglich der Integration von Langzeitgedächtnis-Repräsentationen und Gedächtnisstrategien eine weitere Exploration der Zentralen Exekutive nahelegen. Ferner ergaben sich aus der mTTH auch „some surprises and new discoveries“ (Kintsch et al., 1999, S. 437), wie beispielsweise die unerwartet fehlende Interaktion zwischen linguistischer Gruppe und Artikulatorischer Unterdrückung, die bestehenden Effekte des Wortmaterials und des Präsentationsmodus sowie die fehlende Analogie zwischen den Effekten Irrelevanter Sprache und Phonologischer Ähnlichkeit.

LITERATURVERZEICHNIS

- Anderson, J.R. (1983). A Spreading Activation Theory of Memory. *Journal of Verbal Learning and Verbal Behavior*, 22, 261-295.
- Atkinson, R. & Shiffrin, R.M. (1968). Human Memory: A Proposed System and its Control Processes. In K.W. Spence & J.T. Spence (Eds.), *The Psychology of Learning and Motivation. Advances in Research and Theory, Vol.2* (pp. 89-195). New York: Academic Press.
- Avons, S.E., Wright, K.L. & Pammer, K. (1994). The Word-length Effect in Probed and Serial Recall. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 47A, 207-231.
- Baayen, R.H., Piepenbrock, R. & Gulikers, L. (1995). *The Celex Lexical Data Base (CD-Rom)*. Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.
- Baddeley, A.D. (1966). Short-term Memory for Word Sequences as a Function of Acoustic, Semantic, and Formal Similarity. *Quarterly Journal of Experimental Psychology*, 18, 362-365.
- Baddeley, A. (1986). *Working Memory*. New York: Oxford University Press.
- Baddeley, A. (1991). *Human memory. Theory and Practice*. Hove: Erlbaum.
- Baddeley, A.D. (1992). Is Working Memory Working? The Fifteenth Bartlett Lecture. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 44A, 1-31.
- Baddeley, A.D. (1996). Exploring the Central Executive. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49A, 5-28.
- Baddeley, A.D. (1997). *Human Memory: Theory and Practice*. Hove: Psychology Press.
- Baddeley, A.D. (1998). The Central Executive: A Concept and Some Misconceptions. *Journal of the International Neuropsychological Society*, 4, 523-526.
- Baddeley, A.D. (1999). *Essentials of Human Memory*. Hove: Psychology Press.
- Baddeley, A.D. & Hitch, G.J. (1974). Working Memory. In G.H. Bower (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory, Vol. 8* (pp. 47-89.), New York: Academic Press.
- Baddeley, A.D. & Liebermann, K. (1980). Spatial Working Memory. In R.S. Nickerson (Ed.), *Attention and Performance, VIII* (pp. 521-537). Hillsdale: Erlbaum.
- Baddeley, A.D. & Wilson, B. (1985). Phonological Coding and Short-term Memory in Patients without Speech. *Journal of Memory and Language*, 24, 490-502.
- Baddeley, A.D. & Logie, R.H. (1999). Working Memory: The Multiple-Component Model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 28-61). Cambridge: Cambridge University Press.

- Baddeley, A.D. & Thomson, N. & Buchanan, M. (1975). Word Length and the Structure of Short-term Memory. *Journal of Verbal Learning and Verbal Behavior*, 14, 575-589.
- Baddeley, A.D., Lewis, V. & Vallar, G. (1984). Exploring the Articulatory Loop. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 36A, 233-252.
- Baddeley, A., Gathercole, S. & Papagno, C. (1998). The Phonological Loop as a Language Learning Device. *Psychological Review*, 105, 158-173.
- Baddeley, A.D., Bressi, S., Della Sala, S., Logie, R.H. & Spinnler, H. (1991). The Decline of Working Memory in Alzheimer's Disease: A Longitudinal Study. *Brain*, 114, 2521-2542.
- Banbury, S. & Berry, D.C. (1998). Disruption of Office-related Tasks by Speech and Office Noise. *British Journal of Psychology*, 89, 499-517.
- Baschek, I.-L., Bredenkamp, J., Oehrle, B. & Wippich, W. (1977). Bestimmung der Bildhaftigkeit (I), Konkretetheit (C) und der Bedeutungshaltigkeit (M) von 800 Substantiven. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 24, 353-396.
- Beaman, C.P. & Jones, D.M. (1998). Irrelevant Sound Disrupts Order Information in Free Recall as in Serial Recall. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 51A, 615-636.
- Belleville, S., Peretz, I. & Arguin, M. (1992). Contribution of Articulatory Rehearsal to Short-term Memory: Evidence from a Case of Selective Disruption. *Brain and Language*, 43, 713-746.
- Bosshardt, H.-G. & Laug, C. (1995). Zusammenhänge zwischen Wortlänge, Lexikalität und Sprechgeschwindigkeit. *Zeitschrift für Experimentelle Psychologie*, 42, 576-593.
- Bourassa, D.C. & Besner, D. (1994). Beyond the Articulatory Loop: A Semantic Contribution to Serial Order Recall of Subspan Lists. *Psychonomic Bulletin and Review*, 1, 122-125.
- Bredenkamp, J. (1975). Die modifizierte Invarianzhypothese und die Hypothese einer konstanten Langzeit-Gedächtnisspanne. *Psychologische Beiträge*, 17, 483-496.
- Bredenkamp, J. (1993). Die Verknüpfung verschiedener Invarianzhypothesen im Bereich der Gedächtnispsychologie. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 40, 368-385.
- Bredenkamp, J. & Hager, W. (1979). Experimentelle Befunde zur modifizierten Invarianzhypothese und zur Hypothese einer konstanten Langzeitgedächtnisspanne. *Psychologische Beiträge*, 21, 382-400.
- Bredenkamp, J. & Klein, K.-M. (1998). Experimental Tests of a Model Connecting Three Invariance Hypotheses on Learning and Memory Process. *Zeitschrift für Psychologie*, 206, 107-124.
- Bredenkamp, J. & Hamm, S. (1999). Status Quo der Ergebnisse der Sternberg-Untersuchungen. *Bericht zum DFG-Projekt Arbeitsgedächtnis*. Unveröffentlichtes Manuskript.

- Bredenkamp, J. & Hamm, S. (2001). Further Experimental Tests of Invariance Hypotheses on Learning and Memory Processes. *Zeitschrift für Psychologie*, 209, Heft 3.,
- Bredenkamp, J., Klein, K.-M., van Hayn, S. & Vaterrodt, B. (1988). Gedächtnispsychologische Untersuchungen eines Rechenkünstlers. *Sprache und Kognition*, 7, 69-83.
- Bridges, A.M. & Jones, D.M. (1996). Word Dose in the Disruption of Serial Recall by Irrelevant Speech: Phonological Confusions or Changing State? *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49A, 919-939.
- Broadbent, D.E. (1983). Recent Advances in Understanding Performance in Noise. In G. Rossi (Ed.), *Noise as a Public Health Problem: Proceedings of the Fourth International Congress* (pp. 719-738). Milan: Edizioni Techniche a cura del Centro Ricerche e Studi Amplifon.
- Broadbent, D.E. & Broadbent, M.H. (1981). Articulatory Suppression and the Grouping of Successive Stimuli. *Psychological Research*, 43, 57-67.
- Brown, G.D.A. & Hulme, C. (1995). Modeling Item Length Effects in Memory Span: No Rehearsal Needed?. *Journal of Memory and Language*, 34, 594-621.
- Buchner, A., Irmen, L. & Erdfelder, E. (1996). On the Irrelevance of Semantic Information for the „Irrelevant Speech” Effect. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49A, 765-779.
- Bugelski, B.R. & Kidd, E. & Segmen, J. (1968). Image as a Mediator in One-trial Paired-Associate Learning. *Journal of Experimental Psychology*, 76, 69-73.
- Burani, C., Vallar, G. & Bottini, G. (1991). Articulatory Coding and Phonological Judgements on Written Words and Pictures: The Role of the Phonological Output Buffer. *European Journal of Cognitive Psychology*, 3, 379-398.
- Caplan, D., Rochon, E. & Waters, G. (1992). Articulatory and Phonological Determinants of Word-length Effects in Span Tasks. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 45A, 177-192.
- Cavanagh, J. P. (1972). Relation between the Immediate Memory Span and the Memory Search Rate. *Psychological Review*, 79, 525-530.
- Chincotta, D. & Hoosain, R. (1995). Reading Rate, Articulatory Suppression and Bilingual Digit Span. *European Journal of Cognitive Psychology*, 7, 201-211.
- Chincotta, D. & Underwood, G. (1997a). Digit Span and Articulatory Suppression: A Cross-linguistic Comparison. *European Journal of Cognitive Psychology*, 9, 98-96.
- Chincotta, D. & Underwood, G. (1997b). Bilingual Memory Span Advantage for Arabic Numerals over Digit Words. *British Journal of Psychology*, 88, 295-310.
- Colle, H.A. & Welsh, A. (1976). Acoustic Masking in Primary Memory. *Journal of Verbal Learning and Verbal Behavior*, 15, 17-31.

- Colle, H.A. (1980). Auditory Encoding in Visual Short-term Recall: Effects of Noise Intensity and Spatial Location. *Journal of Verbal Learning and Verbal Behavior*, 19, 722-735.
- Conrad, R. (1964). Acoustic Confusions in Immediate Memory. *British Journal of Psychology*, 55, 75-84.
- Conrad, R. & Hull, A.J. (1964). Information, Acoustic Confusion and Memory Span. *British Journal of Psychology*, 55, 429-432.
- Conway, A.R.A. & Engle, R.W. (1994). Working Memory and Retrieval: A Resource-dependent Inhibition Model. *Journal of Experimental Psychology: General*, 123, 354-373.
- Cooper, E.H. & Pantle, A.J. (1967). The Total-time Hypothesis in Verbal Learning. *Psychological Bulletin*, 68, 221-234.
- Cowan, N. (1992). Verbal Memory Span and the Timing of Spoken Recall. *Journal of Memory and Language*, 31, 668-684.
- Cowan, N. (1994). Mechanisms of Verbal Short-term Memory. *Current Directions in Psychological Science*, 3, 185-189.
- Cowan, N. (1999). An Embedded-Processes Model of Working Memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). Cambridge: Cambridge University Press.
- Cowan, N., Wood, N.L., Nugent, L.D. & Treisman, M. (1997). There are Two Word-length Effects in Verbal Short-Term Memory: Opposed Effects of Duration and Complexity. *Psychological Science*, 8, 290-295.
- Cowan, N., Day, L., Saults, J.S., Keller, T.A., Johnson, T. & Flores, L. (1992). The Role of Verbal Output Time in the Effects of Word Length on Immediate Memory. *Journal of Memory and Language*, 31, 1-17.
- Cowan, N., Keller, T.A., Hulme, C., Roodenrys, S., McDouglas, S. & Rack, J. (1994). Verbal Memory Span in Children: Speech Timing Clues to the Mechanisms Underlying Age and Word Length Effects. *Journal of Memory and Language*, 33, 234-250.
- Craik, F.I. & Lockhart, R.S. (1972). Levels of Processing: A Framework for Memory Research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Cronbach, L.J. (1951). Coefficient Alpha and the Internal Structure of Tests. *Psychometrika*, 16, 297-334.
- D'Esposito, M., Detre, J.A., Alsop, D.C., Shin, R.K., Atlas, S. & Grossmann, M. (1995). The Neural Basis of the Central Executive System of Working Memory. *Nature*, 378, 279-281.
- Daneman, M. & Carpenter, P.A. (1980). Individual Differences in Working Memory and Reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 450-466.
- De Renzi, E., Faglioni, P. & Previdi, P. (1977). Spatial Memory and Hemispheric Locus of Lesion. *Cortex*, 13, 424-433.
- Dilger, S. (2000). *Arbeitsgedächtnis und Versprecher*. Bonn: Dissertation. [WWW

- document]. http://hss.ulb.uni-bonn.de:90/ulb_bonn/diss_online/phil_fak/2000/dilger_stefan
- Dosher, B.A. & Ma, J.J. (1998). Output Loss or Rehearsal Loop? Output-Time versus Pronunciation-Time Limits in Immediate Recall for Forgetting-Matched Materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 316-335.
- Ebbinghaus, H. (1966). *Über das Gedächtnis. Untersuchungen zur experimentellen Psychologie*. Amsterdam: Bonset.
- Ellermeier, W. & Zimmer, K. (1997). Individual Differences in Susceptibility to the „Irrelevant Speech Effect“. *Journal of the Acoustical Society of America*, 102, 2191-2199.
- Ellis, N.R. & Hennis, R.A. (1980). A Bilingual Word-length Effect: Implications for Intelligence Testing and the Relative Ease of Mental Calculation in Welsh and English. *British Journal of Psychology*, 71, 43-51.
- Engle, R. (1996). Working Memory and Retrieval: An Inhibition-Resource Approach. In J.T.E. Richardson, R.W. Engle, L. Hasher, R.H. Logie, E.R. Stoltzfus & R.T. Zacks (Eds.), *Working Memory and Human Cognition* (pp. 89-119). New York: Oxford University Press.
- Engle, R.W., Kane, M.J. & Tuholsky, S.W. (1999). Individual Differences in Working Memory Capacity and What They Tell Us About Controlled Attention, General Fluid Intelligence, and Functions of the Prefrontal Cortex. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 102-134). Cambridge: Cambridge University Press
- Erdfelder, E. (1994). Erzeugung und Verwendung empirischer Daten. In T. Herrman & W. Tack (Eds.), *Methodologische Grundlagen der Psychologie*. Enzyklopädie der Psychologie, Themenbereich B: Methodologie und Methoden, Serie I: Forschungsmethoden der Psychologie, Band 1: Methodologische Grundlagen der Psychologie (pp. 47-98). Göttingen: Hogrefe.
- Erdfelder, E. & Bredenkamp, J. (1994). Hypothesenprüfung. In T. Herrman & W. Tack (Eds.), *Methodologische Grundlagen der Psychologie*. Enzyklopädie der Psychologie, Themenbereich B: Methodologie und Methoden, Serie I: Forschungsmethoden der Psychologie, Band 1: Methodologische Grundlagen der Psychologie (pp.604-649). Göttingen: Hogrefe.
- Erdfelder, E. & Faul, F. & Buchner, A. (1996). GPOWER: A General Power Analysis Program. *Behavior Research Methods, Instruments, and Computers*, 28, 1-11.
- Ericsson, K.A. & Delaney, P.F. (1999). Long-term Working Memory as an Alternative to Capacity Models of Working Memory in Everyday Skilled Performance. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 257-297). Cambridge: Cambridge University Press.

- Ericsson, K.A., Chase, W.G. & Faloon, S. (1980). Acquisition of a Memory Skill. *Science*, 208, 1181-1182.
- Fankam, T. (1999). *Lernen und Arbeitsgedächtnis*. Universität Bonn: unveröffentlichte Diplomarbeit.
- Farah, M.J., Hammond, K.M., Levine, D.N. & Clavario, R. (1988). Visual and Spatial Mental Imagery: Dissociable Systems of Representation. *Cognitive Psychology*, 20, 439-462.
- Farrand, P. & Jones, D. (1996). Direction of Report in Spatial and Verbal Serial Short-term Memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49A, 140-158.
- Fodor, A. (1985). *The Modularity of Mind*. Cambridge: MIT.
- Frankish, C. (1985). Modality-specific Grouping Effects in Short-term Memory. *Journal of Memory and Language*, 24, 200-209.
- Fuster, J.M. (1998). Distributed Memory for Both Short and Long Term. *Neurobiology of Learning and Memory*, 70, 268-274.
- Gadenne, V. (1994). Theoriebewertung. In T. Herrman & W. Tack (Eds.), *Methodologische Grundlagen der Psychologie*. Enzyklopädie der Psychologie, Themenbereich B: Methodologie und Methoden, Serie I: Forschungsmethoden der Psychologie, Band 1: Methodologische Grundlagen der Psychologie (pp. 295-342). Göttingen: Hogrefe.
- Garai, J.E. & Scheinfeld, A. (1968). Sex Differences in Mental and Behavioral Traits. *Genetic Psychology Monographs*, 77, 169-229.
- Gathercole, S.E. & Baddeley, A.D. (1993). *Working Memory and Language*. Hove: Erlbaum.
- Gathercole, S.E. & Hitch, G.J. (1993). Developmental Changes in Short-term Memory: A Revised Working Memory Perspective. In A.F. Collins, S.E. Gathercole, M.A. Conway & P.E. Morris (Eds.), *Theories of Memory* (pp. 189-209). Hove: Erlbaum.
- Grainger, J. & Jacobs, A.M. (1998). On Localist Connectionism and Psychological Science. In J. Grainger & A.M. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 1-38). Mahwah: Erlbaum.
- Hamm, S. (1998). *Irrelevante Sprache und die Invarianzhypothesen der Lern- und Gedächtnispsychologie*. Universität Bonn: unveröffentlichte Diplomarbeit.
- Hanley, J.R., Young, A.W. & Pearson, N.A. (1991). Impairment of the Visuo-spatial Sketch Pad. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 43A, 101-126.
- Hasselhorn, M. (1988). Wie und warum verändert sich die Gedächtnisspanne über die Lebensspanne?. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 20, 322-337.
- Hempel, C.G. (1977). *Aspekte wissenschaftlicher Erklärung*. Berlin: De Gruyter.

- Henry, L.A. (1991). Development of Auditory Memory Span: The Role of Rehearsal. *British Journal of Developmental Psychology*, 9, 493-551.
- Hoosain, R. (1982). Correlation Between Pronunciation Speed and Digit Span Size. *Perceptual and Motor Skills*, 55, 1128.
- Hulme, C., Maughan, S. & Brown, G.D. (1991). Memory for Familiar and Unfamiliar Words: Evidence for a Long-term Memory Contribution to Short-term Memory Span. *Journal of Memory and Language*, 30, 685-701.
- Hulme, C., Roodenrys, S., Brown, G. & Mercer, R. (1995). The Role of Long-term Memory Mechanisms in Memory Span. *British Journal of Psychology*, 86, 527-536.
- Hulme, C., Newton, P., Cowan, N., Stuart, G. & Brown, G. (1999). Think before you Speak: Pauses, Memory Search, and Trace Redintegration Processes in Verbal Memory Span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 447-463.
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G.D.A., Martin, S. & Stuart, G. (1997). Word-Frequency Effects on Short-term Memory Tasks: Evidence for a Redintegration Process in Immediate Serial Recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 1217-1232.
- Isaacs, E.B. & Vargha-Kadem, F. (1989). Differential Course of Development of Spatial and Verbal Memory Span: A Normative Study. *British Journal of Developmental Psychology*, 7, 377-380.
- Jacoby, L.L. (1991). A Process-Dissociation Framework: Separating Automatic from Intentional Uses of Memory. *Journal of Memory and Language*, 30, 513-541.
- Jones, D. (1993). Objects, Streams, and Threads of Auditory Attention. In A.D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, Awareness, and Control: A Tribute to Donald Broadbent* (pp. 87-104). Oxford: Clarendon Press.
- Jones, D.M. & Macken, W.J. (1995a). Auditory Babble and Cognitive Efficiency: The Role of Number of Voices and their Location. *Journal of Experimental Psychology: Applied*, 1, 216-226.
- Jones, D.M. & Macken, W.J. (1995b). Organizational Factors in the Effect of Irrelevant Speech: The Role of Spatial Location and Timing. *Memory and Cognition*, 23, 192-200.
- Jones, D.M. & Macken, W.J. (1995c). Phonological Similarity in the Irrelevant Speech Effect: Within- or Between-stream Similarity?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 103-133.
- Jones, D., Alford, D., Bridges, A., Tremblay, S. & Macken, B. (1999). Organizational Factors in Selective Attention: The Interplay of Acoustic Distinctiveness and Auditory Streaming in the Irrelevant Sound Effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 464-473.
- Jonides, J., Reuter-Lorenz, P.A., Smith, E.E., Awh, E., Barnes, L.L., Drain, M., Glass, J.,

- Lauber, E.J., Patalano, A.L. & Schumacher, E.H. (1996). Verbal and Spatial Working Memory in Humans. In D.L. Medin (Ed.), *The Psychology of Learning and Motivation, Vol 35* (pp. 43-88). London: Academic Press.
- Kintsch, W. (1970). *Learning, Memory, and Conceptual Processes*. New York: Wiley.
- Kintsch, W., Healy, A.F., Hegarty, M., Pennington, B.F. & Salthouse, T.A. (1999). Models of Working Memory: Eight Questions and Some General Issues. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 412-441). Cambridge: Cambridge University Press.
- Klatte, M. (1996). *Struktur und Prozesse des Arbeitsgedächtnisses: Theoretische Modelle und experimentelle Untersuchungen*. Oldenburg: Dissertation.
- Klatte, M. & Hellbrück, J. (1993). Der „Irrelevant Speech Effect“: Wirkungen von Hintergrundschall auf das Arbeitsgedächtnis. *Zeitschrift für Lärmbekämpfung*, 40, 91-98.
- Klauer, K.C. & Stegmaier, R. (1997). Interference in Immediate Spatial Memory: Shifts of Spatial Attention or Central-Executive Involvement. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 50A, 79-99.
- Klein, K.-M. (1995). *Experimentelle Untersuchungen zu zwei Invarianzhypothesen des Kurzzeitgedächtnisses*. Bonn: Pace.
- Lass, U. (1997). Einfluß von Sprache auf Verarbeitungsprozesse im Kurzzeitgedächtnis – eine vergleichende Untersuchung mit deutschen und chinesischen Probanden. In G. Lüer & U. Lass (Eds.), *Erinnern und Behalten. Wege zur Erforschung des menschlichen Gedächtnisses* (pp. 244-268). Göttingen: Vandenhoeck & Ruprecht.
- Lass, U., Yunqiu, F., Guopeng, C., Becker, D. & Lüer, G. (1999). Is Memory for Shapes Subject to Language-Specific Effects? An Experimental Study of Memory Span in German and Chinese Subjects. *Sprache und Kognition*, 18, 136-145.
- Lass, U., Lüer, G., Becker, D., Yunqiu, F. & Guopeng, C. (2000a). *Encoding and Retrieval Components Affecting Memory Span: Articulation Rate, Memory Search and Trace Redintegration*. Unveröffentlichtes Manuskript.
- Lass, U., Lüer, G., Becker, D., Yunqiu, F., Guopeng, C. & Zhongming, W. (2000b). Kurzzeitgedächtnisleistungen deutscher und chinesischer Probanden mit verbalen und figuralen Items: zur Funktion von Phonologischer Schleife und visuell-räumlichem Notizblock. *Zeitschrift für Experimentelle Psychologie*, 47, 77-88.
- LeCompte, D.C. (1995). An Irrelevant Speech Effect with Repeated and Continuous Background Speech. *Psychonomic Bulletin and Review*, 2, 391-397.
- LeCompte, D.C. & Shaibe, D.M. (1997). On the Irrelevance of Phonological Similarity to the Irrelevant Speech Effect. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 50A, 100-118.
- LeCompte, D.C., Neely, C.B. & Wilson, J.R. (1997). Irrelevant Speech and Irrelevant Tones:

- The Relative Importance of Speech to the Irrelevant Speech Effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 472-483.
- Lehto, J. (1996). Are Executive Function Tests Dependent on Working Memory Capacity?. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49A, 29-50.
- Levelt, W.J.M., Roelofs, A. & Meyer, A.S. (1999). A Theory of Lexical Access in Speech Production. *Behavioral and Brain Sciences*, 22, 1-75.
- Logie, R.H. (1986). Visuo-spatial Processing in Working Memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 38A, 229-247.
- Logie, R.H. (1995). *Visuo-spatial Working Memory*. Hove: Erlbaum.
- Lovatt, P., Avons, S.E. & Masterson, J. (2000). The Word-length Effect and Disyllabic Words. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 53A, 1-22.
- Lüer, G., Becker, D., Lass, U., Yunqiu, F., Guopeng, C. & Zhongming, W. (1998). Memory Span in German and Chinese: Evidence for the Phonological Loop. *European Psychologist*, 3, 102-112.
- MacDonald, M.C., Just, M.A. & Carpenter, P.A. (1992). Working Memory Constraints on the Processing of Syntactic Ambiguity. *Cognitive Psychology*, 24, 56-98.
- McNemar, Q. (1962). *Psychological Statistics*. New York: Wiley.
- Meiser, T. (1997). *Arbeitsgedächtnis und Changing-State-Hypothese*. Heidelberg: Dissertation.
- Meyer, D.E. & Kieras, D.E. (1997a). A Computational Theory of Executive Cognitive Processes and Multiple-task Performance. Part 1: Basic Mechanisms. *Psychological Review*, 104, 3-26.
- Meyer, D.E. & Kieras, D.E. (1997b). A Computational Theory of Executive Cognitive Processes and Multiple-task Performance. Part 2: Accounts of Psychological Refractory-period Phenomena. *Psychological Review*, 104, 749-791.
- Miles, C., Jones, D.M. & Madden, C.A. (1991). Locus of the Irrelevant Speech Effect in Short-term Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 17, 578-584.
- Miller, G.A. (1956). The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information. *Psychological Review*, 63, 81-97.
- Monsell, S. (1987) On the Relation between Lexical Input and Output Pathways for Speech. In A. Allport, D.G. MacKay, W. Prinz & E. Scheerer (Eds.), *Language Perception and Production* (pp. 273-311). London: Academic Press.
- Morris, N. (1987). Exploring the Visuo-spatial Scratch Pad. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 39A, 409-430.

- Murray, D.F. (1968). Articulation and Acoustic Confusability in Short-term Memory. *Journal of Experimental Psychology*, 78, 679-684.
- Naveh-Benjamin, M. & Ayres, T.J. (1986). Digit Span, Reading Rate, and Linguistic Relativity. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 38A, 739-751.
- Neath, I. & Nairne, J.S. (1995). Word-length Effects in Immediate Memory: Overwriting Trace Decay Theory. *Psychonomic Bulletin and Review*, 2, 429-441.
- Nevelsky, P.B. (1970). Comparative Investigation of the Short- and Long-term Memory Span. In K.H. Pribram & D.E. Broadbent (Eds.), *Biology of Memory* (pp. 21-28). New York: Academic Press.
- Norman, D.A. & Shallice, T. (1986). Attention to Action: Willed and Automatic Control of Behavior. In R.J. Davidson, G.E. Schwartz & D. Shapiro (Eds.), *Consciousness and Self-Regulation: Advances in Research and Theory, Vol. 4* (pp. 1-18), New York: Plenum Press.
- Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. New York: Oxford University Press.
- Parkin, A.J. (1998). The Central Executive Does Not Exist. *Journal of the International Neuropsychological Society*, 4, 518-522.
- Pechmann, T. & Mohr, G. (1992). Interference in Memory for Tonal Pitch: Implications for a Working-Memory Model. *Memory and Cognition*, 20, 314-320.
- Phillips, W.A. & Christie, D.F. (1977). Interference with Visualization. *Quarterly Journal of Experimental Psychology*, 29, 637-650.
- Platon (zitiert nach 1981). *Theaitet griechisch - deutsch*. Stuttgart: Reclam.
- Popper, K.R. (1969). *Logik der Forschung*. Tübingen: Mohr.
- Puckett, J.M. & Kausler, D.H. (1984). Individual Differences and Models of Memory Span: A Role for Memory Search Rate?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 72-82.
- Quinn, J.G. (1988). Interference Effects in the Visuo-spatial Sketchpad. In M. Denis, J. Engelkamp & J.T.E. Richardson (Eds.), *Cognitive and Neuropsychological Approaches to Mental Imagery* (pp. 181-189). Amsterdam: Nijhoff.
- Quinn, J.G. & McConnell, J. (1996b). Irrelevant Pictures in Visual Working Memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49A, 200-215.
- Quinn, J.G. & McConnell, J. (1996a). Indications of the Functional Distinction Between the Components of Visual Working Memory. *Psychologische Beiträge*, 38, 355-367.
- Reisberg, D., Rappaport, I. & O'Shaughnessy, M. (1984). Limits of Working Memory: The Digit Digit-Span. *Journal of Experimental Psychology: Learning, Memory, and Cognition*,

- 10, 203-221.
- Romani, C. (1992). Are there Distinct Input and Output Buffers? Evidence from an Aphasic Patient with an Impaired Output Buffer. *Language and Cognitive Processes*, 7, 131-162.
- Ruchkin, D.S., Johnson, R., Grafman, J., Canoune, H. & Ritter, W. (1992). Distinctions and Similarities among Working Memory Processes: An Event-related Potential Study. *Cognitive Brain Research*, 1, 53-66.
- Sainsbury R.M. (1994). *Paradoxien*. Stuttgart: Reclam.
- Salamé, P. & Baddeley, A.D. (1982). Disruption of Short-term Memory by Unattended Speech: Implications for the Structure of Working Memory. *Journal of Verbal Learning and Verbal Behavior*, 21, 150-164.
- Salamé, P. & Baddeley, A.D. (1986). Phonological Factors in STM: Similarity and the Unattended Speech Effect. *Bulletin of the Psychonomic Society*, 24, 263-265.
- Salamé, P. & Baddeley, A.D. (1989). Effects on Background Music on Phonological Short-term Memory. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 41A, 107-122.
- Schiffer, S. (2000). *Arbeitsgedächtnis und Lernen: Untersuchung an bildhaftem Material*. Universität Bonn: unveröffentlichte Diplomarbeit.
- Schweickert, R. (1993). A Multinomial Processing Tree Model for Degradation and Redintegration in Immediate Recall. *Memory and Cognition*, 21, 168-175.
- Schweickert, R. & Boruff, B. (1986). Short-term Memory Capacity: Magic Number or Magic Spell?. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 419-425.
- Schweickert, R., Guentert L. & Hersberger, L. (1990) Phonological Similarity, Pronunciation Rate, and Memory Span. *Psychological Science*, 1, 74-77.
- Seitz, K. (1998). *Mentales Multiplizieren: Modalitätsspezifische Prozesse im Arbeitsgedächtnis*. Hamburg: Kovac.
- Seitz, K., Zoelch, C. & Schumann-Hengsteler, R. (2000). *Arbeitsgedächtnis und Rechnen*. Unveröffentlichtes Manuskript.
- Service, E. (1998). The Effect of Word Length on Immediate Serial Recall Depends on Phonological Complexity, not Articulatory Duration. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 51A, 283-304.
- Shallice, T. & Burgess, P. (1996). The Domain of Supervisory Processes and Temporal Organization of Behaviour. *Philosophical Transactions of the Royal Society of London B*, 351, 1405-1411.
- Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.
- Smith, E.E. & Jonides, J. (1995). Working Memory in Humans: Neuropsychological

- Evidence. In M. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 1009-1020.). Cambridge: MIT.
- Smyth, M.M. & Scholey, K.A. (1994) Interference in Immediate Spatial Memory. *Memory and Cognition*, 22, 1-13.
- Smyth, M.M. & Scholey, K.A. (1996). The Relationship between Articulation Time and Memory Performance in Verbal and Visuospatial Tasks. *British Journal of Psychology*, 87, 179-191.
- Smyth, M.M. (1996). Interference with Rehearsal in Spatial Working Memory in the Absence of Eye Movements. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49A, 940-949.
- Spitzbarth, M. (2000). *Der "Irrelevant-Sound-Effekt": Wirkungen auditiver Gruppierung*. Eichstätt: Dissertation.
- Sternberg, S. (1966). High-speed Scanning in Human Memory. *Science*, 153, 652-654.
- Stigler, J.W., Lee, S. & Stevenson, H.W. (1986). Digit Memory in Chinese and English: Evidence for a Temporally Limited Store. *Cognition*, 23, 1-20.
- Tattersall, A.J. & Broadbent, D.E. (1991). Output Buffer Storage and the Modality of Recall. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 43A, 1-18.
- Thurstone, L.L. (1919). The Learning Curve Equation. *Psychological Review Monograph Supplement*, 26, 3.
- Tremblay, S. & Jones, D.M. (1998). Role of Habituation in the Irrelevant Sound Effect: Evidence from the Effects of Token Set Size and Rate of Transition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 659-671.
- Tresch, M.C., Sinnamon, H.M. & Seamon, J.G. (1993). Double Dissociation of Spatial and Object Visual Memory: Evidence from Selective Interference in Intact Human Subjects. *Neuropsychologia*, 31, 211-219.
- Tukey, J.W. (1977). *Exploratory Data-Analysis*. Reading: Addison-Wesley.
- Turner, M.L. & Engle, R.W. (1989). Is Working Memory Capacity Task Dependent?. *Journal of Memory and Language*, 28, 127-154.
- Vallar, G. & Baddeley, A.D. (1984). Fractionation of Working Memory: Neuropsychological Evidence for a Phonological Short-term Store. *Journal of Verbal Learning and Verbal Behavior*, 23, 151-161.
- Vallar, G. & Cappa, S.F. (1987). Articulation and Verbal Short-term Memory: Evidence from Anarthria. *Cognitive Neuropsychology*, 4, 55-77.
- Vandierendonck, A., De Vooght, G. & Van der Goten, K. (1998a). Does Random Time Interval Generation Interfere with Working Memory Executive Functions?. *European Journal of Cognitive Psychology*, 10, 413-442.

- Vandierendonck, A., De Vooght, G. & Van der Goten, K. (1998b). Interfering with the Central Executive by Means of a Random Interval Repetition Task. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 51A, 197-218.
- Walker, I. & Hulme, C. (1999). Concrete Words are Easier to Recall than Abstract Words: Evidence for a Semantic Contribution to Short-term Serial Recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 1256-1271.
- Waters, G.S., Rochon, E. & Caplan, D. (1992). The Role of High-Level Speech Planning in Rehearsal: Evidence from Patients with Apraxia of Speech. *Journal of Memory and Language*, 31, 54-73.
- Wechsler, D. (1956). *Die Messung der Intelligenz Erwachsener*. Bern: Huber.
- Westermann, R. (2000). *Wissenschaftstheorie und Experimentalmethodik*. Göttingen: Hogrefe.
- Westmeyer, H. (1987). Zum Problem empirischen Gehalts psychologischer Theorien. in J. Brandtstädter (Ed.), *Struktur und Erfahrung in der psychologischen Forschung*. Berlin: de Gruyter.
- Westmeyer, H. (1998). Psychologische Methoden und soziale Konventionen. In K.C. Klauer & H. Westmeyer (Eds.), *Psychologische Methoden und soziale Prozesse* (pp. 250-266). Lengerich: Papst.
- Westmeyer, H., Hannemann, J., Nell, V., Völkel, U. & Winkelmann, K. (1987). Eine Monotheory-Multimethod Analyse: Plädoyer für einen deduktivistischen Multiplismus. *Diagnostica*, 33, 227-242.
- Zhang, W. & Feng, L. (1990). The Visual Recognition and Capacity of STM for Chinese Disyllabic Words. *Acta Psychologica Sinica*, 22, 383-390.
- Zimmer, H.D. (1993). Modalitätsspezifische Systeme der Repräsentation und Verarbeitung von Information. *Zeitschrift für Psychologie*, 201, 203-235.

ANHANG I

Instruktionen zu den Hauptexperimenten

Die Instruktion für deutsche Probanden wurden im Wortlaut nicht verändert. Allein das Stimulusmaterial, respektive der Umfang der Sequenzlänge im Supraspannen-Experiment, wurden je nach Experimentreihe variiert. Im folgenden werden die Instruktionen für das Experiment mit phonologisch ähnlichen Items als Beispiel dargestellt.

Diese Instruktionen übersetzte ein diplomierter Übersetzer in das Spanische. Für das Chinesische wurde die Instruktion erst in das Englische und dann in das Chinesische übertragen. Zudem wurde überprüft, ob die chinesische Version äquivalent zu dem deutschen Original ist.

Einleitung, Bildschirm I

Lieber Proband,
zu Beginn möchte ich Ihnen kurz die Versuche erläutern, die in der nächsten halben Stunde auf Sie zukommen. Das Untersuchungsmaterial besteht aus folgenden zehn Worten

WALD WELT WILD KILT BILD
ZELT FELD HELD GOLD GELD

In einem ersten Versuch interessiert es uns, wieviel Zeit Sie benötigen, um Wortsequenzen laut vorzulesen. In den beiden nächsten Untersuchungen möchten wir etwas über Ihr Gedächtnis erfahren: erstens wieviel Lerndurchgänge Sie benötigen, um sich zehn Worte zu merken und zweitens wieviel Worte sie unmittelbar korrekt wiedergeben können.

Bei genauer Beachtung der Instruktionen, die sie jeweils unmittelbar vor dem Start der Untersuchung bekommen, und mit Konzentration werden Sie die Versuche gut bewältigen können.

Einleitung, Bildschirm II

Ihre Daten unterliegen selbstverständlich dem Datenschutz und können nicht von Dritten eingesehen werden. Bitte haben Sie aber Verständnis dafür, daß auf Angaben zum Vornamen, Alter, Semesterzahl und Geschlecht nicht verzichtet werden kann.

Ganz zum Schluß noch eine Bitte: Versuchen Sie bitte, sich genau an die Instruktionen zu halten und eine möglichst gleichbleibende, hohe Leistung zu halten. Vielen Dank!!

Instruktion zum Experiment zur Messung der Artikulationszeit, Bildschirm I

Im folgenden sehen sie Wortreihen auf dem Bildschirm. Ihre Aufgabe besteht nun darin, diese Folgen so schnell als möglich laut vorzulesen.

Vor jedem Durchgang wird Ihnen das Wort ACHTUNG gefolgt von drei akustischen Signalen dargeboten. Nach Präsentation des dritten Tones erscheint die Wortfolge auf dem Bildschirm. Bitte beginnen Sie sofort und ohne Verzögerung mit dem Vorlesen.

Instruktion zum Experiment zur Messung der Artikulationszeit, Bildschirm II

[Beispiel]

Bitte versuchen Sie

- erstens mit einer Atemlänge auszukommen
- zweitens sich nicht zu versprechen
- drittens jedes Element deutlich auszusprechen und
- viertens so schnell wie möglich zu lesen.

Falls Sie keine weiteren Fragen haben, können wir beginnen!

Instruktion zum Subspannen-Experiment, Bildschirm I

Im folgenden *werden Ihnen verschieden lange Wortreihen auf dem Bildschirm dargeboten. Ihre Aufgabe ist es, die Sequenz vollständig und in der richtigen Reihenfolge wiederzugeben. Vor jedem Durchgang wird das Wort ACHTUNG eingeblendet. Nach jedem Durchgang werden Sie nach einer kurzen Pause zur Wiedergabe aufgefordert. Bitte warten Sie auf jeden Fall das Signal WIEDERGABE ab, bevor Sie anfangen zu sprechen.

Instruktion zum Subspannen-Experiment, Bildschirm II

[Beispiel]

Sprechen Sie bitte laut und deutlich.

Falls Sie keine weiteren Fragen haben, können wir beginnen!

Instruktion zum Surpraspannen-Experiment, Bildschirm I

Im folgenden wird Ihnen eine Zufallsfolge von 10 Worten für 10 Sekunden dargeboten. Ihre Aufgabe ist es, die jeweils dargebotene Sequenz vollständig und in der richtigen Reihenfolge mündlich wiederzugeben. Vor der Darbietung erscheint das Wort ACHTUNG, danach werden Sie durch das Signal WIEDERGABE zur mündlichen Reproduktion aufgefordert.

Instruktion zum Surpraspannen-Experiment, Bildschirm II

Sprechen Sie bitte laut und deutlich. Der Versuch ist beendet, wenn Sie die Reihe zweimal hintereinander korrekt reproduziert haben. Die Sequenz bleibt über alle Lerndurchgänge hinweg unverändert.

Für die Experimente mit Artikulatorischer Unterdrückung wurde folgende Ergänzung in jeder Instruktion vorgenommen.

In folgenden Versuchen sollen Sie - zusätzlich zu den Aufgaben - bitte immer im Gleichklang mit einem Computerton das Wort TAU aussprechen.

[Beispiel]

ANHANG II

Ergebnisse zu Vorexperiment 5. 6 und 7

Tabelle 32

Item	Bildhaftigkeit	Bedeutungsgehalt	Konkretheit	Frequenz
BACH	5.85	4.70	5.65	12.00
BAHN	5.70	5.35	5.50	68.00
BALL	6.80	5.75	6.80	61.00
BANK	5.35	4.65	5.70	91.00
BART	6.25	3.70	6.45	12.00
BAUM	6.70	6.10	6.90	64.00
BEIL	6.00	3.95	5.90	3.00
BERG	6.52	5.41	6.52	68.00
BIER	6.65	5.30	6.65	46.00
BILD	5.85	4.80	5.50	297.00
BILD*	5.88	4.81	5.42	
BLUT	6.60	5.30	6.73	44.00
BUCH	6.40	4.93	6.45	183.00
CHEF	6.12	4.04	5.52	56.00
COLT	5.80	3.65	6.20	14.00
DARM	3.55	3.75	5.55	3.00
DORF	5.50	5.00	5.25	96.00
FACH	3.00	2.85	3.70	23.00
FASS	6.52	4.50	6.70	13.00
FELD	6.00	4.65	4.90	100.00
FELL	6.00	5.00	6.30	6.00
FELS	6.00	4.55	6.45	12.00
FILZ	4.00	2.70	5.00	1.00
FLUT	5.72	4.52	5.67	14.00
FORM	4.52	3.67	4.12	228.00
GELD	6.20	4.80	5.25	202.00
GIFT	5.28	4.26	5.79	3.00
GOLD	6.20	4.45	5.00	48.00
GOLD*	6.28	4.67	6.47	
GOTT	2.25	5.15	1.75	174.00

Fortsetzung von Tabelle 32

Item	Bildhaftigkeit	Bedeutungsgehalt	Konkretheit	Frequenz
GRAB	6.25	5.25	5.90	44.00
GRAB*	6.48	4.93	6.36	
GRAD	1.50	2.40	2.65	129.00
GRAF	3.70	3.75	4.35	87.00
GRAL	3.00	3.05	3.20	0.00
GRAM	1.75	3.45	2.30	1.00
GRAS	6.25	4.45	6.00	24.00
GRAS*	6.40	5.00	6.55	
GRAT	2.45	2.10	3.60	2.00
HALS	6.05	4.30	6.55	44.00
HAUT	6.16	4.33	6.45	49.00
HEER	5.88	4.63	5.73	18.00
HELD	3.85	3.55	3.65	44.00
HEMD	5.95	4.75	6.40	25.00
HOLZ	6.05	5.40	6.15	46.00
HULD	1.30	1.80	1.40	1.00
KILT	4.50	3.20	5.95	0.00
KIND	6.32	5.44	6.22	421.00
KRAM	3.45	4.15	3.85	1.00
KRAN	5.95	3.80	6.15	8.00
KRUG	6.05	4.30	6.50	5.00
KULT	1.60	4.45	1.85	2.00
KUSS	6.52	5.48	6.39	12.00
LIED	2.70	4.80	3.85	49.00
LUFT	5.20	4.56	5.82	109.00
MALZ	2.30	2.35	4.55	1.00
MAST	5.05	3.30	5.30	3.00
MEER	6.64	6.11	6.58	70.00
MILZ	2.90	3.50	5.55	1.00
MOOS	5.40	3.40	5.75	4.00
MORD	5.92	4.93	5.70	42.00
NERV	2.65	3.90	4.00	27.00
NETZ	5.75	3.60	5.55	19.00
NORD	2.45	3.95	3.00	61.00

Fortsetzung von Tabelle 32

Item	Bildhaftigkeit	Bedeutungsgehalt	Konkretheit	Frequenz
PELZ	5.90	4.15	6.00	4.00
PELZ*	6.16	4.67	6.21	
PILZ	6.45	4.40	6.50	5.00
PLAN	5.08	3.67	4.63	205.00
PULS	1.85	3.20	3.50	5.00
PULT	5.25	3.35	6.20	1.00
SALZ	5.50	4.20	5.85	11.00
SEIL	5.84	4.44	6.52	9.00
SOLD	2.25	3.00	3.75	1.00
STIL	3.92	3.63	3.45	47.00
TAKT	3.88	3.52	4.06	12.00
TIER	6.24	5.62	5.88	128.00
VOLT	1.40	3.70	2.70	5.00
WALD	6.60	5.90	6.15	81.00
WALD*	6.68	5.81	6.45	
WALL	4.25	2.55	5.30	15.00
WELT	4.80	4.45	3.45	987.00
WELT*	4.80	4.48	3.73	
WERT	3.56	3.44	2.48	154.00
WILD	4.35	4.50	4.90	0.00
ZEIT	3.76	4.56	2.88	104.00
ZELT	6.25	4.20	6.25	10.00
ZIEL	5.20	3.67	4.03	259.00

Ergebnisse zu Vorexperiment 8 und 9

Tabelle 33

Wortpaar	Phonolog. Ähnlichkeit	Semant. Ähnlichkeit
BANK - BAUM	1.50	1.50
BANK - CHEF	1.13	1.94
BANK - FORM	1.00	1.25
BANK - KIND	1.13	1.13
BANK - WERT	1.00	2.25

Fortsetzung von Tabelle 33

Wortpaar	Phonolog. Ähnlichkeit	Semant. Ähnlichkeit
BAUM - BUCH	1.38	1.38
BAUM - KIND	1.00	1.25
BAUM - PLAN	1.00	1.06
BAUM - WERT	1.00	1.38
BILD - FELD	1.81	1.25
BILD - GELD	2.25	1.31
BILD - GOLD	1.75	1.44
BILD - HELD	2.00	1.31
BILD - WILD	2.88	1.13
BUCH - BANK	1.31	1.50
BUCH - KIND	1.00	1.31
BUCH - LUFT	1.25	1.06
BUCH - PLAN	1.06	1.38
BUCH - ZIEL	1.00	1.13
CHEF - BAUM	1.00	1.13
CHEF - BUCH	1.00	1.50
CHEF - LUFT	1.12	1.06
CHEF - PLAN	1.06	1.75
CHEF - ZIEL	1.00	1.69
FELD - GELD	2.88	1.06
FELD - HELD	2.94	1.38
FELD - KILT	1.81	1.06
FELD - WALD	1.69	2.56
FELD - ZELT	2.81	1.69
FOMR - PLAN	1.00	1.31
FORM - BAUM	1.13	1.44
FORM - BUCH	1.00	1.19
FORM - CHEF	1.00	1.13
FORM - ZIEL	1.00	1.13
GELD - GOLD	2.38	2.94
GELD - HELD	2.94	1.44
GELD - WILD	1.81	1.13
GELD - ZELT	2.94	1.00
GOLD - FELD	1.81	1.13
GOLD - HELD	1.75	1.19

Fortsetzung Tabelle 33

Wortpaar	Phonolog. Ähnlichkeit	Semant. Ähnlichkeit
GOLD – WELT	1.63	1.13
GOLD – WILD	1.63	1.06
HELD – KILT	1.63	1.19
HELD – WELT	2.94	1.56
HELD – WILD	1.88	1.19
HELD – ZELT	2.88	1.00
KILT – BILD	2.75	1.19
KILT – GELD	1.63	1.00
KILT – GOLD	1.50	1.00
KILT – WALD	1.56	1.13
KILT – WELT	2.06	1.00
KIND – CHEF	1.00	1.19
KIND – FORM	1.00	1.19
KIND – LUFT	1.00	1.13
KIND – PLAN	1.00	1.31
KIND – WERT	1.13	1.31
LUFT – BANK	1.00	1.13
LUFT – BAUM	1.00	1.56
LUFT – FORM	1.06	1.00
LUFT – WERT	1.00	1.13
PLAN – BANK	1.06	1.44
PLAN – LUFT	1.06	1.06
PLAN – WERT	1.00	1.44
PLAN – ZIEL	1.06	2.44
WALD – BILD	1.56	1.31
WALD – GELD	1.88	1.13
WALD – GOLD	1.63	1.13
WALD – HELD	1.69	1.50
WALD – WELT	2.38	1.56
WELT – BILD	1.69	1.63
WELT – FELD	3.00	1.31
WELT – GELD	2.94	1.63
WELT – ZELT	3.00	1.13
WERT – BUCH	1.06	1.75
WERT – CHEF	1.00	1.31

Fortsetzung von Tabelle 33

Wortpaar	Phonolog. Ähnlichkeit	Semant. Ähnlichkeit
WERT – FORM	1.13	1.13
WERT – ZIEL	1.00	1.75
WILD – FELD	2.06	2.25
WILD – KILT	3.00	1.06
WILD – WALD	2.44	2.38
WILD – WELT	2.50	1.38
WILD – ZELT	1.75	1.38
ZELT – BILD	1.69	1.06
ZELT – GOLD	1.56	1.00
ZELT – KILT	2.00	1.13
ZELT - WALD	1.81	1.88
ZIEL - BANK	1.13	1.19
ZIEL - BAUM	1.00	1.19
ZIEL - KIND	1.06	1.25
ZIEL - LUFT	1.06	1.06

* Kontroll-Items aus Baschek et alii (1977)