

Identifikation und Modifikation leistungsbeeinträchtigender Lärmparameter

**Inaugural-Dissertation zur Erlangung des Doktorgrades der Philosophisch-
Pädagogischen Fakultät**

der

Katholischen Universität Eichstätt-Ingolstadt

vorgelegt von

Dipl.-Psych. Alexandra Feil, Pappenheim

2005

Referent: Univ.-Prof. Dr. Jürgen Hellbrück

Koreferentin: Univ.-Prof. Dr. R. Schumann-Hengsteler

**Tag der letzten mündlichen Prüfung: 19. Oktober 2006
(gemäß § 15 Abs. 1)**

Danken möchte ich vor allem für die exzellente und freundschaftliche Betreuung bei Prof. Dr. J. Hellbrück.

Für die vielen aufmunternden und fachlichen Gespräche, die angenehme Arbeitsatmosphäre und die Hilfe bei technischen Problemen danke ich meinen Kollegen Dipl.-Psych. Andreas Liebl, Dr. phil. Dipl.-Psych. Sabine Schlittmeier, Dipl.-Phys. Matthias Vormann und Dipl.-Psych. Marlis Wegner, sowie meinen studentischen Hilfskräften cand. Psych. Alexander Assfalg, Dipl.-Psych. Andrea Binser, Dipl.-Psych. Christiane Hauser, Dipl.-Psych. Sabine Heinrich, Dipl.-Psych. Markus Nagler, cand. Psych. Katrin Scheuren und Dipl.-Psych. Andrea Settele für ihr Engagement und ihre Mitarbeit. Außerdem danke ich meiner Freundin Julia Hofmann für die sprachliche Durchsicht der Arbeit und meinem Mann Raphael Feil für die geduldige und verständnisvolle Unterstützung.

Die berichteten Experimente wurden im Rahmen des Verbundprojektes *Leiser Verkehr* des Bundesministeriums für Bildung und Forschung und des TÜV Rheinland (Laufzeit 2002-2008) erhoben.

Gewidmet meinen Eltern
Euer Vorbild hat mich hierher gebracht
Danke

Zusammenfassung

Lärmschutz besteht meist aus einer globalen Forderung nach Pegelreduktion. Dies stellt jedoch nicht immer die effektivste Modifikationsmöglichkeit für Störschalle dar. Aus der kognitionspsychologischen Forschung und ihren Modellen können Hypothesen abgeleitet werden, welche Spezifikationen oder Variationen im Schallsignal entscheidend für eine Störwirkung auf geistige Leistung sind. Es werden neun Experimente vorgestellt. Diese haben zum Ziel mit Hilfe kognitionspsychologischer Modelle die Auswirkungen der Variation von Schallparametern wie Zeitstruktur, Lautheit und Frequenzzusammensetzung auf einige grundlegende kognitive Funktionen zu spezifizieren. Damit sollten über eine reine Pegelreduktion hinausgehende Modifikationsmöglichkeiten von Störschallen im Hinblick auf fundierten Schallschutz ableitbar werden. Die Auswirkung der spektral-temporalen Struktur eines Schallsignals zeigt sich in der Forschung zum so genannten *Irrelevant Sound Effect*, einem gut untersuchten leistungsbeeinträchtigenden Effekt spektral-temporal stark variabler Schalle (unter anderem Sprache) auf das Memorieren serieller Informationen (Klatte, Kilcher, & Hellbrück, 1995). Dieser Effekt wird als präattentiv angenommen und im Rahmen des Working Memory Modells (Baddeley, 2000a) als eine Interferenz des Schalls mit den gemerkten Inhalten in einem Subsystem dieses Modells, der so genannten *phonologischen Schleife*, erklärt. Die Auswirkungen von hoher Lautheit und bestimmter Frequenzzusammensetzung von Schallen hingegen werden meist über eine Steigerung des Arousal erklärt. Eine Steigerung des Arousal kann, vermittelt über eine steigende Selektivität (Easterbrook, 1959) und Ablenkbarkeit des Aufmerksamkeitsfokus (Kahneman, 1973), zu einer Beeinträchtigung der Lösung von Aufgaben führen, die die so genannte *zentrale Exekutive*, das überwachende und kontrollierende Subsystem des Working Memory Modells erfordern.

In Fragestellung I (3 Experimente) werden diese beiden unterschiedlichen Wirkungen von Parametern der Hintergrundschalle auf drei Aufgaben untersucht, die selektiv Ansprüche an die beiden genannten Systeme des Arbeitsgedächtnisses stellen. Aufgaben, die vor allem die phonologische Schleife fordern (Serial Recall und KLTM, ein modifizierter Konzentrations-Leistungstest (Düker & Lienert, 1965)), werden erwartungsgemäß durch spektral-temporal stark variable Schalle (Sprachschall) gestört. Ein lautes und tieffrequent dominiertes Straßenverkehrslärmgeräusch ($L(A)_{eq}=70$ dB(A)) beeinträchtigt hingegen Aufgaben, welche die Funktion der zentralen Exekutive beanspruchen (Stroop Test (Stroop, 1935) und KLTM). In Fragestellung II (4 Experimente) werden Modifikationen der Verkehrsschalle untersucht (Reduktion in der Lautheit und/oder im tieffrequenten Bereich). Hierbei kann mit Einschränkungen eine geringere Störwirkung durch eine kombinierte Modifikation beider Parameter gezeigt werden (KLTM; Stroop Test; Grammatical Reasoning, das vergleichbar zum KLTM beide Systeme beansprucht). In einem weiteren Experiment (KLTM) wird ein möglicher Einfluss der Darbietungsart (Lautsprecher vs. Kopfhörer) ausgeschlossen. In Fragestellung III (2 Experimente) wird die Relevanz der Semantik für die Störwirkung des Sprachschalls untersucht, sowie Schienen- statt Straßenverkehrslärm verwendet (KLTM; Grammatical Reasoning). Es lässt sich eine teilweise Unabhängigkeit des Sprachschalleffektes von seiner Semantik zeigen (nur im KLTM). Für den Schienenverkehrsschall ergeben sich jedoch keine Effekte in derselben Größenordnung wie für den Straßenverkehr. Die Ergebnisse werden im Hinblick auf Schallschutzmaßnahmen in beruflichen Umwelten sowie eventueller Implikationen für die modelltheoretische Basis diskutiert.

Zusammenfassung**Gliederung**

1 Einleitung und theoretischer Hintergrund	1
1.1 Der aktuelle Stand der Lärmwirkungsforschung und daraus ableitbare Schallschutzmaßnahmen	3
1.1.1 <i>Lästigkeitsbeeinflussende Lärmparameter und ihre Relevanz im Lärmschutz</i>	3
1.1.2 <i>Auswirkungen von Lärm auf kognitive Leistungen</i>	7
1.2 Ein Modell des Arbeitsgedächtnisses	9
1.3 Interferenzen in der phonologischen Schleife zwischen Schall und bestimmten Aufgaben - der Irrelevant Sound Effect	12
1.4 Lärmwirkungen auf die zentrale Exekutive – Arousal im Rahmen des Aufmerksamkeitskonzepts	15
1.5 Die Verwendung von Rechentests in der Lärmwirkungsforschung	20
1.6 Das Grammatical Reasoning und seine Verwendung in der Lärmwirkungsforschung	25
1.7 Der Stroop Test – Theorie und Befunde	28
2 Fragestellung und Hypothesen	33
3 Versuchsdurchführung, -auswertung und Diskussion der Ergebnisse	35
3.1 Fragestellung I: Vergleich der Wirkung von Sprachschall und Verkehrsschallen auf Aufgaben, welche die phonologische Schleife oder die zentrale Exekutive fordern	35
3.1.1 Explizierung der Hypothesen	35
3.1.2 Allgemeines zur Auswertung	37
3.1.3 Serial Recall Aufgabe	38
3.1.3.1 <i>Aufgabenstellung</i>	38
3.1.3.2 <i>Versuchsplan, Versuchspersonen und Versuchsablauf</i>	39
3.1.3.3 <i>Ergebnisse</i>	39
3.1.3.4 <i>Diskussion</i>	42
3.1.4 Modifikation des Konzentrations-Leistungs-Tests KLTM	44
3.1.4.1 <i>Aufgabenstellung</i>	44
3.1.4.2 <i>Versuchsplan, Versuchspersonen und Ablauf</i>	45
3.1.4.3 <i>Ergebnisse</i>	45
3.1.4.4 <i>Diskussion</i>	50

3.1.5 Der Stroop Test	53
3.1.5.1 <i>Aufgabenstellung</i>	53
3.1.5.2 <i>Versuchsplan, Probanden und Ablauf</i>	53
3.1.5.3 <i>Ergebnisse</i>	54
3.1.5.4 <i>Diskussion</i>	57
3.1.6 Diskussion der Ergebnisse zu Fragestellung I	60
3.2 Fragestellung II: Auswirkungen einer Pegelreduktion im tiefen Frequenzbereich bei Verkehrsschallen	61
3.2.1 Explizierung der Hypothesen	61
3.2.2 Der KLTM mit pegelreduzierten Verkehrsschallen	64
3.2.2.1 <i>Aufgabenstellung</i>	65
3.2.2.2 <i>Versuchsplan, Versuchspersonen und Versuchsablauf</i>	65
3.2.2.3 <i>Ergebnisse</i>	65
3.2.2.4 <i>Diskussion</i>	69
3.2.3 KLTM mit Darbietung der Hintergrundschalle über Lautsprecher und Subwoofer	72
3.2.3.1 <i>Aufgabenstellung</i>	72
3.2.3.2 <i>Versuchsplanung, Probanden und Ablauf</i>	72
3.2.3.3 <i>Ergebnisse</i>	73
3.2.3.4 <i>Diskussion</i>	78
3.2.4 Der Stroop Test mit pegelreduzierten Schallen	79
3.2.4.1 <i>Aufgabenstellung</i>	79
3.2.4.2 <i>Versuchplan, Probanden und Ablauf</i>	79
3.2.4.3 <i>Ergebnisse</i>	79
3.2.4.4 <i>Diskussion</i>	82
3.2.5 Das Grammatical Reasoning	85
3.2.5.1 <i>Aufgabenstellung</i>	85
3.2.5.2 <i>Versuchsplan, Probanden und Versuchsablauf</i>	86
3.2.5.3 <i>Ergebnisse</i>	87
3.2.5.4 <i>Diskussion</i>	90
3.2.6 Diskussion der Ergebnisse zu Fragestellung II	93
3.3 Fragestellung III: Relevanz der Semantik für die Wirkung des Sprachschalls	95
3.3.1 Explizierung der Hypothesen	95
3.3.2 KLTM unter unverständlichem Sprachschall und Schienenverkehrslärm	97
3.3.2.1 <i>Aufgabenstellung</i>	97
3.3.2.2 <i>Versuchsplan, Versuchspersonen und Versuchsablauf</i>	97
3.3.2.3 <i>Ergebnisse</i>	98
3.3.2.4 <i>Diskussion</i>	100
3.3.3 Grammatical Reasoning mit unverständlichem Sprachschall und Schienenverkehrsgeräuschen	103
3.3.3.1 <i>Aufgabenstellung</i>	103

Gliederung

3.3.3.2 <i>Versuchsplan, Versuchspersonen und Ablauf</i>	103
3.3.3.3 <i>Ergebnisse</i>	103
3.3.3.4 <i>Diskussion</i>	106
3.3.4 Diskussion der Ergebnisse zu Fragestellung III	109
4 Zusammenfassende Diskussion	111
Literaturverzeichnis	117
Anhang	

1 Einleitung und theoretischer Hintergrund

„Mach mal bitte das Fenster zu, bei dem Lärm kann ich mich nicht konzentrieren!“ oder „Sei still, ich muss lernen!“ – diesen oder ähnlichen Aussagen begegnet man häufig im Zusammenhang mit Lärm und Leistung. Die meisten Personen fühlen sich durch Lärm im weitesten Sinne beeinträchtigt. Insbesondere, wenn geistig gearbeitet werden soll. Aber wirkt unerwünschter Schall tatsächlich beeinträchtigend? Und wenn ja, wie und auf welche kognitiven Funktionen? Hieran schließt sich die Frage an, ob sich eine etwaige störende Wirkung irrelevanten Schalls durch Schallschutzmaßnahmen, wie bauliche oder technische Veränderungen beeinflussen ließe.

Eine der am heftigsten diskutierten Lärmquellen in der öffentlichen Meinung stellt sicherlich der Verkehrslärm dar. Zudem wird dieser vermutlich, trotz immer leiserer Motoren, durch das steigende Verkehrsaufkommen in den nächsten Jahren weiter anwachsen. Der Güterfernverkehr auf Schiene und Straße wird ebenso wie der Luftverkehr im Jahre 2010 voraussichtlich um 80%, der innerstädtische Verkehr um 50% gegenüber 1995 zugenommen haben (European Commission, 1997).

Dass Verkehrslärm lästig ist dürfte außer Frage stehen, aber wie steht es um eine objektivierbare Messung seiner Störwirkung auf basale kognitive Funktionen? Im Vergleich zum gut untersuchten *Irrelevant Sound Effect (ISE)*, der eine spezifische Störwirkung eines irrelevanten spektral-temporal stark variablen oder sprachlichen Schalls auf eine Aufgabe, die serielles Memorieren fordert, beschreibt, soll in dieser anwendungsorientierten Arbeit untersucht werden, welche Komponenten des Verkehrslärms die entscheidenden bei einer eventuellen beeinträchtigenden Wirkung auf grundlegende kognitive Funktionen sind. Anschließend sollen Möglichkeiten gesucht werden, diese beeinträchtigende Wirkung durch eine Modifikation der Schalle zu reduzieren. Bisher wurde Schallschutz meist auf eine Reduktion des Mittelungspegels in Dezibel (A-Filterung) beschränkt. In dieser Arbeit soll anhand der Begründung der Aufgabenstellungen über kognitionspsychologische Modelle untersucht werden, welche Modifikationen (z.B. Beeinflussung der spektralen Zusammensetzung, Zeitstruktur oder Lautheit) sich tatsächlich als relevant für eine Leistungsbeeinträchtigung erweisen.

Im theoretischen Teil der vorliegenden Arbeit wird zunächst auf den aktuellen Stand der Lärmwirkungsforschung und daraus ableitbarer Schallschutzmaßnahmen eingegangen. Anschließend wird ein kognitionspsychologisches Modell des Arbeitsgedächtnisses vorgestellt, das Working Memory Modell von Baddeley (2000a), das als Basis einer theoriegeleiteten Erforschung von Lärmwirkungen benutzt werden soll. Vor allem die Subsysteme *phonologische Schleife* und *zentrale Exekutive* sind hier von Interesse. Als Wirkungsbeispiel der spektral-temporalen Struktur von Schallen in einem dieser Subsysteme wird anschließend der bereits erwähnte Irrelevant Sound Effect kurz dargelegt. In weiteren theoretischen Abschnitten werden Erklärungsansätze für eine leistungsbeeinträchtigende Wirkung von lauten und tieffrequenten Verkehrsschallen über eine Steigerung des Arousal erörtert. Danach werden Forschungen zu Aufgaben vorgestellt, mit deren Hilfe einige grundlegende kognitive Funktionen wie Aufmerksamkeit, Bearbeitung von Doppelaufgaben, schlussfolgerndes Denken sowie Inhibitionsfähigkeit getestet werden können, um die Funktionen der phonologischen Schleife und der zentralen Exekutive zu operationalisieren. Diese kognitiven Funktionen sind Grundlage für komplexere

geistige Tätigkeiten und damit essentiell für beinahe alle Bereiche geistiger Arbeit. Aus diesem Grund sollen diese Funktionen exemplarisch auf ihre Sensitivität sowohl gegenüber spektral-temporal stark variablen als auch gegenüber lauten und tieffrequenten Schallen untersucht werden.

Zielsetzung der im empirischen Teil berichteten Experimente ist es zu erforschen, ob die aus der kognitionspsychologischen Forschungstradition abgeleitete beeinträchtigende Wirkung von Sprachschall auf die phonologische Schleife und von lauten, tieffrequenten Schallen (Verkehrsschallen) auf die zentrale Exekutive experimentell während der Bearbeitung von Aufgaben mit verschiedenen Anforderungen an das Arbeitsgedächtnis gezeigt werden kann. Weiterhin wird ermittelt, welche Modifikation der Verkehrsgeräusche zu einer Verringerung einer beeinträchtigenden Wirkung führt, um Empfehlungen für einen sinnvollen Schallschutz ableiten zu können. Aus diesem Grund wird für den Sprachschall anschließend untersucht, ob sein leistungsbeeinträchtigender Effekt in den verwendeten Tests unabhängig von seiner Semantik nachgewiesen werden kann und somit in seiner spektral-temporalen Struktur begründet liegt. Die Ergebnisse der durchgeführten Experimente werden in einer Zusammenfassung auch im Hinblick auf die modelltheoretische Fundierung diskutiert. Die Arbeit schließt mit Anregungen für weitere Untersuchungen.

1.1 Der aktuelle Stand der Lärmwirkungsforschung und daraus ableitbare Schallschutzmaßnahmen

1.1.1 Lästigkeitsbeeinflussende Lärmparameter und ihre Relevanz im Lärmschutz

Bereits Anfang des 20. Jahrhunderts beschäftigten sich Forscher mit Lärm, Lärmmessung, Auswirkungen von Lärm und Schallschutz im weitesten Sinne. Auf die Auswirkungen von Schall mit einer Lautstärke von über 100 dB wie beispielsweise bleibende Schäden am Ohr, die Maskierung von wichtigen Reizen sowie physiologische Veränderungen, wie steigenden Blutdruck, Verlangsamung der Magenkontraktion und Steigerung des Herzschlages wurde bereits damals hingewiesen (Bartlett, 1934). An psychologischen Auswirkungen wurden Leistungsbeeinträchtigungen, Klagen über den Lärm und Versuche, ihn zu reduzieren, genannt. Als Gegenmaßnahmen wurden bauliche und administrative Maßnahmen zur Reduktion vor allem der Verkehrsgeräusche vorgeschlagen, z.B. eine Lautstärkenbegrenzung von Hupen oder Regelungen, in welchen Stadtteilen (Krankenhausnähe) nicht gehupt werden darf. Ein bereits damals als wichtig erachtetes Merkmal der Lärmbelastung ist die Frequenzlage. Tonhöhen in der Lage der menschlichen Stimme seien am wenigsten störend. Geräusche in höherer oder tieferer Tonlage seien als belästigender erlebt worden (Laird & Coye, 1929). Hieraus wurden zu Beginn des letzten Jahrhunderts Lärmschutzempfehlungen in der Arbeitsumgebung abgeleitet, in denen die Wichtigkeit einer Modifikation des Frequenzspektrums von Lärm hin zu den eher sprachähnlichen Bereichen betont wurde. Wie diese Modifikation vorzunehmen sei, wurde allerdings nicht beschrieben. Das Problem eines spezifizierten Lärmschutzes ist folglich nicht neu, hat aber auch nichts an seiner Relevanz eingebüßt.

Im Folgenden sollen zunächst die aktuellen Rechtsgrundlagen für den Lärmschutz in Deutschland vorgestellt sowie Lärmgrenzwerte kritisch diskutiert werden. Anhand des Lästigkeitsurteils wird anschließend ausgearbeitet, dass nicht nur der Schallpegel, sondern auch personenbezogene und situative Faktoren eine wichtige Rolle bei der Beurteilung von Lärm spielen. Weiterhin werden Parameter des Schalls dargelegt, die bei Belästigungsreaktionen auf Verkehrslärm wichtig zu sein scheinen, wie z.B. die Art des Verkehrsmittels, die Frequenzzusammensetzung sowie die zeitliche Struktur des Geräusches.

Inzwischen ist die Rechtsgrundlage für den Lärmschutz in Deutschland – historisch gewachsen – vielfältig und zersplittert geregelt. Durch eine Vorgabe von Grenzwerten in Dezibel (mit A-Filterung) ist sie nur noch auf eine globale Pegelreduktion bezogen. Andere Faktoren des Schalls wie seine Zeitstruktur oder Frequenzzusammensetzung fanden jedoch zunehmend Beachtung. Die Grenzwertregelungen liegen vor allem im *Bundes-Immissions-Schutz-Gesetz* (BImSchG) vor, in seinen ergänzenden Verordnungen und Verwaltungsvorschriften, aber auch zum Beispiel im *Bauplanungsrecht* (Baugesetzbuch und Baunutzungsverordnung) und den *Regelungen zum privaten Nachbarrecht* (Bürgerliches Gesetzbuch). Hinsichtlich konkreter Richtwerte wird meist auf drei Regelwerke zurückgegriffen: Erstens die *Technische Anleitung Lärm* von 1968, revidiert 1998 (TA Lärm, 1998), die Anweisungen zur Genehmigung und Überwachung von Anlagen für Behörden

enthält, zweitens auf die *VDI Richtlinie 2058* (Vereinigung deutscher Ingenieure, Richtlinie Nr. 2058, Blatt 1), die oft als eine Zusammenfassung des lärmschutzfachlichen Sachverständes gesehen wird und deshalb viel Einfluss im Schallschutz erlangt hat. Das dritte Regelwerk, das hier zu nennen ist, ist die *DIN-Norm 18005 des Deutschen Institutes für Normung*, die zwar nicht die Autorität einer staatlichen Regelung besitzt, deren Einhaltung aber häufig von Bauaufsichtsbehörden verlangt wird.

Der Schutz vor Verkehrslärm ist eigens geregelt. So wird dem Straßenverkehrslärm durch das Straßenrecht und das Verkehrsrecht Rechnung getragen. Die Grenzwerte für Verkehrslärm sind in der *Verkehrslärmschutzverordnung* (16.BImSchG) geregelt und liegen je nach Gebiet (Krankenhäuser, Wohngebiet oder Gewerbegebiet) und Tageszeit (tags vs. nachts) zwischen 47 dB(A) (nachts, Krankenhaus oder Altenheim) und 69 dB(A) (Gewerbegebiet, tagsüber). In reinen Wohngebieten schreibt der Gesetzgeber tagsüber einen Mittelungspegel¹ von 59 dB(A) und nachts von 49 dB(A) vor (Infodienst Lärmschutz in Kur- und Erholungsorten, 1992). Für Bahnlärm gelten im Allgemeinen nicht so strenge Bestimmungen wie für Straßenlärm. Hier kann eine Entschädigung erst ab Mittelungspegeln von 60-65 dB(A) (nachts) bis 75 dB(A) (tags) genehmigt werden. Fluglärm wird in einem eigenen Gesetz bedacht, dem *Gesetz zum Schutz gegen Fluglärm*, und allgemeiner im *Flugverkehrsrecht*. Vor allem gelten hier Mindesthöhen von 300 Metern für den Überflug von Ortschaften, und das Gebot, keinen vermeidbaren Lärm zu erzeugen.

Lärmgrenzwerte müssen allerdings kritisch gesehen werden, da sie Verallgemeinerungen darstellen und eigentlich noch stärker an bestimmte Personengruppen (z.B. schwer Kranke: Krankenhäusernähe) angepasst werden sollten (Jansen, 1987). Außerdem werden bei Grenzwertangaben, wie oben bereits andiskutiert, zumeist nicht die Frequenzzusammensetzung und die zeitliche Struktur des Geräusches bedacht, die aber, wie im Folgenden gezeigt werden soll, Belästigungs- und Belastungsreaktionen beeinflussen. Die Beschränkung auf einen reinen Grenzwert in Dezibel (mit A-Filterung) soll im weiteren Verlauf der Arbeit kritisch hinterfragt und differenziert betrachtet werden. Tatsächlich könnte es effizienter sein, spezifische Merkmale in den Schallen zu eliminieren, anstatt *nur* global eine Senkung des mittleren Schalldruckpegels anzustreben. Welche Parameter des Schalls hierfür in Frage kämen, soll im Weiteren vorgestellt werden.

Lärmgrenzwerte werden eingesetzt, um einer Beeinträchtigung durch Lärm vorzubeugen. Neben direkten auralen Effekten, wie Schwerhörigkeit, Tinnitus oder temporaler Schwellenverschiebung, ist ein hauptsächlich extraauraler Effekt von Lärm die Belästigung (Hellbrück & Ellermeier, 2004). Das Lästigkeitsurteil über Lärm ist in Abhängigkeit von der Lautstärke des Schalls zu sehen (Osada, 1991), hängt aber auch von der individuellen Lärmempfindlichkeit (Ellermeier, Eigenstetter, & Zimmer, 2001) und der aktuellen Stimmung ab. Es kann somit nicht als eindimensional angesehen werden (Västfjäll, 2002). Zu den individuellen Faktoren zählt sowohl die eben angesprochene Lärmempfindlichkeit als auch die wahrgenommene Kontrolle über die Geräusche, z.B. durch die Möglichkeit das Fenster zu schließen (Hatfield et al., 2002). Zu den situativen Bedingungen, die Einfluss auf das Lästigkeitsurteil haben, gehört z.B. die ausgeführte Tätigkeit. Vor allem bei Tätigkeiten im Freien und bei Kommunikation (Felscher-Suhr et al., 1996) wird

¹ Beim Mittelungspegel handelt es sich um den Schallpegel, der dem energetischen Mittelungswert bezogen auf eine bestimmte Beurteilungszeit entspricht. Vgl. auch Seite 16.

Verkehrslärm als belästigend eingestuft und führt auf Dauer zu einer Reduktion des allgemeinen Wohlbefindens (Ouis, 2001).

Das Belästigungserleben ist damit sowohl abhängig von der Schallquelle als auch von personenbezogenen und situativen Faktoren (Guski, 1993). Die Schallquelle wirkt sich bei der Belästigung durch Verkehrslärm ebenfalls aus: Fluglärm wird im Allgemeinen als lästiger erlebt als Straßenverkehrslärm (Taylor, 1995), Bahnverkehr hingegen schneidet in der subjektiven Beurteilung am besten ab. Die Unterschiede zwischen Straßen- und Schienenverkehrslärm liegen bei gleichem Mittelungspegel vor allem in der Pausenstruktur. Bei Straßenlärm gibt es ungleich mehr Pausen als bei Schienenlärm, diese sind bei letzterem im Schnitt 20 min lang, verglichen mit einer Länge von etwa einer Sekunde beim Straßenverkehr (Möhler, 1988). Allgemein wird dem Schienenverkehr vom Gesetzgeber ein Bonus von 5 dB(A) in Bezug auf das ausgelöste Belästigungserleben eingeräumt (Working Group Railway Noise on the European Commission, 2003), der im subjektiven Urteil sogar noch höher ausfallen könnte, wenn durch zusätzlichen Straßenverkehrslärm die Pausenstruktur der Schiene verschwimmen würde. Dies scheint vor allem an der besseren Vorhersagbarkeit der Schienengeräusche durch ihren regelmäßigen Fahrplan und an ihrem gleichmäßigen An- und Abschwelen zu liegen (Möhler, 1988). Bei Flugzeuggeräuschen muss im Vergleich zu Straßenverkehrslärm eher von einem Malus von ca. 5 dB(A) ausgegangen werden. Diese Geräusche werden im Vergleich als lästiger beurteilt (Working Group Railway Noise on the European Commission, 2003). Ein Teil dieser Unterschiede ist wohl durch die Lokalisation des Flugverkehrslärms zu erklären, da auch Schienenverkehrsgeräusche als ähnlich lästig beurteilt werden, wenn sie von oben zu kommen scheinen. Eine Lärmquelle von oben scheint ein höheres Gefahrenpotential zu vermitteln (Fastl, Fruhmann, & Ache, 2004). Mit diesen Boni und Mali wurde versucht, im Schallschutz anderen Aspekten als dem reinen Pegel (in dB(A)) Rechnung zu tragen, in diesem Fall der Geräuschquelle, woraus allerdings wieder nur eine Grenzwertverschiebung resultierte. Der Schwerpunkt soll in dieser Arbeit wegen der großen Brisanz aufgrund der höchsten Wachstumsraten (European Commission, 1997) auf dem Straßenverkehrslärm liegen.

Zusammenfassend kann man somit *nicht* folgern, dass das Lästigkeitsurteil nur abhängig vom Pegel des Schalls (in dB(A)) ist, worauf in gesetzlichen Vorgaben durch Grenzwertangaben rekuriert wird. Beachtet werden müssen hingegen auch die Schallquelle und ihre Bewertung als Besonderheit des Schallsignals sowie individuelle und situative Faktoren der exponierten Personen. Da diese Einflüsse allerdings einer gesetzlich vorgegebenen Maßnahme zur Veränderung wenig zugänglich sind, sollen im Folgenden weitere Parameter von Schallen vorgestellt werden, wie z. B. eine spezifische Frequenzzusammensetzung und hohe zeitliche Struktur (Pausenstruktur) von Lärm. Mit Hilfe einer Überarbeitung der Schallquelle könnten diese modifiziert werden. Um eine Konfundierung von Effekten zu vermeiden müssen personelle und situative Faktoren jedoch bei einer empirischen Überprüfung des Erfolges einer Modifikation von Störschallen kontrolliert werden.

Die Bebauung und Art der Verkehrsführung hat durch die dadurch hervorgerufene Veränderung im Schallsignal einen Einfluss auf die subjektive Beurteilung von Verkehrslärm. Schallaufnahmen von Kreisverkehren (83 dB(A)) mit einer gleichmäßigen zeitlichen Struktur riefen in einer Untersuchung (Notbohm, Gärtner, & Schwarze, 2003) deutlich schwächere aversive und

aktivierende Reaktionen sowohl physiologischer als auch subjektiver Parameter hervor als Aufnahmen von Ampelkreuzungen in derselben Lautstärke. Letztere weisen eine wesentlich stärkere zeitliche Struktur im Sinne von häufigen kurzen Pausen auf. Zusätzlich zur Geräuschquelle und der im Folgenden diskutierten Frequenzzusammensetzung scheint die zeitliche Struktur der Geräusche über die Wirkung auf den Menschen und die daraus folgende Beurteilung der Lästigkeit zu entscheiden. Auf einen Störschall mit einer sehr ausgeprägten Pausenstruktur, nämlich irrelevanten Sprachschall, und seiner beeinträchtigenden Auswirkung auf bestimmte kognitive Leistungen soll in Kapitel 1.3 bei der Vorstellung des so genannten *Irrelevant Sound Effect* noch näher eingegangen werden.

Auch die Frequenzzusammensetzung von Störschallen spielt eine wichtige Rolle für ihre Bewertung. Durch die weniger wirkungsvollen Abschirmungsmöglichkeiten von tieffrequenten Schallen und ihren hohen Anteil in Verkehrsgeräuschen jeder Quelle kann man das Lästigkeitsurteil über diese Schalle zu einem großen Teil mit ihrer Lastigkeit in diesem Frequenzbereich in Zusammenhang bringen (Berglund, Hassmen, & Job, 1996). Motorgeräusche und Mündungsgeräusche der Abgasanlage stellen die dominanten Geräuschdimensionen bei Straßenverkehrslärm dar, insbesondere in der Anfahrtsituation und Kfz-Vorbeifahrten in einer Geschwindigkeit von 20-50 km/h. Rollgeräusche spielen bei dieser Geschwindigkeit noch eine vernachlässigbare Rolle. Die spektrale Zusammensetzung der Mündungs- und Motorgeräusche hängt mit der Motordrehzahl und der Anzahl der Zylinder zusammen. Die tieffrequenten Oberwellen der Kurbelwellenumdrehung sind hier im Spektrum dominant. Je nach Drehzahl und Anzahl der Zylinder bewegt sich, bei einer Anfahrtsituation an der Ampel, der Hauptteil der abgestrahlten Geräusche in einem Bereich zwischen 50 und 500 Hz (Bisping, 2005; Klingenberg, 1991). Eine mögliche Modifikation von solchen Emissionen durch die Fahrzeugindustrie bestünde in einer Einflussnahme auf das Motormanagement oder der schalldämmenden Auskleidung des Motorraums. Hierbei könnten die tiefen Frequenzen unter 500 Hz bei Pkws um einige Dezibel gesenkt werden (Klingenberg, 1991). Kombinierte man Ausgleichswellen am Motor mit einer Verringerung der Schallausbreitung über Getriebe, Karosserie, Motorlager sowie Ansaug- und Abgastrakt, so wäre laut dem Autor eine Absenkung in diesem Bereich um einige weitere Dezibel technisch umsetzbar (Bisping, 2005). Dies führe trotz eines unveränderten A-bewerteten Schalldruckpegels zu einer geringeren subjektiven Lautheit (sone) und sei damit ein ökonomisches Mittel, die Lärmbelastung zu reduzieren. Ob dies aber auch die negativen Auswirkungen auf z.B. kognitive Leistungen beeinflusst, ist eine noch offene Frage, die in den folgenden Kapiteln wieder aufgegriffen wird.

Schienenverkehrslärm wird ähnlich wie Straßenverkehrslärm von tiefen Frequenzen dominiert. Hierbei sind im Wesentlichen drei Signalquellen involviert, nämlich tieffrequente Schwingungen der Räder, die Rad-Schiene Interaktion und Schienen-Schwellenschwingungen (Working Group Railway Noise on the European Commission, 2003). Die Spektralanalyse der Schienengeräusche lässt, anders als bei Kfz-Geräuschen, leider keine Differenzierung einzelner spektraler Komponenten bzw. Quellen zu. Eine sinnvolle Umgestaltung wäre folglich auch hier, den tiefen Frequenzbereich durch Modifikationen der Schienen und Räder im Pegel abzusenken (Working Group Railway Noise on the European Commission, 2003). Die Frage, ob sich hierdurch eine Reduktion der vermuteten negativen Auswirkungen auf kognitive Leistungen erreichen ließe, soll ebenfalls im weiteren Verlauf der Arbeit zu

beantworten versucht werden. Dies wäre gerade im Zusammenhang mit Grenzwertangaben in A-bewertetem Schalldruckpegel interessant, da sich dieses Maß durch eine solche Modifikation wenig ändern würde, obwohl eventuell negative Effekte des Lärms reduziert werden könnten.

Somit können als lästigkeitsbeeinflussende Verkehrslärmparameter, neben personenbezogenen und situativen Faktoren der exponierten Personen, vor allem die Art des Verkehrsmittels, die Pausenstruktur des Geräusches und seine tieffrequente Lastigkeit gelten. Die Auswirkungen dieser Lärmparameter auf kognitive Leistungen sollen im folgenden Kapitel kurz im Überblick vorgestellt werden, ehe der beabsichtigte Bogen von der Lärmwirkungsforschung zur Kognitionspsychologie gespannt wird.

1.1.2 Auswirkungen von Lärm auf kognitive Leistungen

Die Auswirkungen von Lärm auf kognitive Leistungen wurden in vielfältiger Variation untersucht, wobei die Untersuchungen zu Verkehrslärm, die in der Literatur zu finden sind, eher unverbunden nebeneinander stehende Ergebnisse erbrachten. Hier soll nur ein kurzer Überblick gegeben werden; in den folgenden Kapiteln wird dann ausführlicher auf spezifische Effekte, kognitionspsychologische Ansätze zur Erklärung der Lärmwirkungen und einzelne Aufgabentypen, die sich zu ihrer Untersuchung eignen eingegangen.

Leistungsbeeinträchtigende Wirkungen von Lärm wurden in Abhängigkeit von der Lautstärke des Schalls z.B. auf Vigilanzaufgaben (Broadbent, 1953) und Rechenaufgaben (Broadbent, 1958a) gefunden. Eine höhere Lautstärke (Breitbandrauschen mit 100 dB vs. 70 dB) führte hierbei zu einer stärkeren Beeinträchtigung der Leistung in Form einer Erhöhung der Fehlerraten. Mehr Falschantworten findet der Autor zudem bei einer Wahlreaktionsaufgabe während der Darbietung eines hochfrequenten Schalls mit einer Lautstärke von 100 dB im Vergleich zum selben Schall mit 80 dB. Im Vergleich zu einem tieffrequenten Schall mit 80 dB jedoch keinen Einfluss hat ein hochfrequenter Schall mit 90 dB oder ein tieffrequenter Schall mit 90 bzw. 100 dB (Broadbent, 1957). Die Lautstärken die der Autor verwendete würden aus heutiger Sicht allerdings als extrem bezeichnet werden und werden inzwischen aus ethischen Gründen Probanden nicht mehr ungefiltert dargeboten. Bei diesen Experimenten muss zudem der Vergleich zu einer so genannten *Ruhebedingung* kritisch gesehen werden, welche aus einem Schall mit einem Mittelungspegel von 80 dB bestand. Bereits diese Lautstärke hat erheblichen Einfluss auf die Leistung. Dies wird im Folgenden erläutert.

Aus einer Übersichtsarbeit lässt sich folgern, dass schon in einem Pegelbereich zwischen 50 und 85 dB beeinträchtigende Auswirkungen auf Versuchsaufgaben und Arbeitstätigkeiten nachweisbar sind. Dies ist vor allem bei einer hoher Komplexität der Aufgabe, großer Schwierigkeit oder gesteigerten Anforderungen an feinmotorische Koordination sowie bei starken Ansprüchen an Aufmerksamkeit und Konzentration oder beim Aufnehmen und Speichern von Informationen der Fall (Sust, 1987).

Folglich scheint bei einer Beeinträchtigung kognitiver Funktionen sowohl die Lautstärke als auch die Frequenzzusammensetzung eine Rolle zu spielen.

Arbeitsgedächtnisleistungen und Lästigkeitsbewertung werden durch diskontinuierlichen Lärm beeinflusst, was mittels einer erhöhten Ablenkbarkeit erklärt wird (Hörmann & Osterkamp, 1966a,

1966b). Andere Autoren betonen die aufmerksamkeitsdistrahierende Wirkung von diskontinuierlichem Schall, die zu einer Erschöpfung der für die Aufgaben nötigen Ressourcen führt (Schönpflug, 1993). Eine Verschiebung des Speed-Accuracy-Trade-Off hin zu einer schnelleren, aber ungenaueren Arbeitsweise, sowie eine gewisse Inflexibilität im Denken bei der Lösung von Aufgaben unter Hintergrundschall, werden als Folge von Belästigungs- und Ärgerreaktionen gesehen (Dornic, 1990). Die Einschätzung der Belästigung und die Leistungsbeeinträchtigung scheinen in manchen Untersuchungen nicht trennbar zu sein. Vielmehr beeinflussen sie sich wohl gegenseitig, und zwar insofern, dass die Beeinträchtigung zu steigen scheint, wenn die subjektive Lästigkeit zunimmt, und umgekehrt (Dornic, 1990). Dies bedeutet, dass bei einer Messung der Leistung unter Lärm immer auch die erlebte Belästigung mit erhoben werden sollte, um beide Parameter zu kontrollieren und eine Konfundierung zu verhindern. Bereits oben wurde dies für die personenbezogenen und situativen Variablen, die ihrerseits das Lästigkeitserleben beeinflussen, angeregt.

Aus diesen Erläuterungen kann man ableiten, dass Lärm vor allem in hoher Lautstärke zu Belästigung und Leistungseinbußen führen kann. Ebenso spielen die Frequenzzusammensetzung, die zeitliche Struktur des Geräusches und Personvariablen, wie die individuelle Lärmempfindlichkeit, die ausgeführte Tätigkeit oder auch der Umgang mit dem Lärm eine wichtige Rolle. Eine Spezifikation, welche kognitiven Leistungen durch was für Aspekte eines Lärmsignals beeinträchtigt werden können, fehlt in der aktuellen Literatur jedoch noch. Durch Einbeziehung kognitionspsychologischer Modelle soll abgeleitet werden, welche kognitiven Funktionen sich als empfindlich auf spezifische Parameter von Lärm erweisen könnten. Hieran schließt sich eine Operationalisierung und empirische Überprüfung dieser vermuteten Beeinträchtigungen an. Aus diesen Ergebnissen sollen dann Empfehlungen zum Schallschutz vor allem für geistige Tätigkeiten in Büros oder Schulen ausgesprochen werden, die über eine reine globale Pegelreduktion (in dB(A)) hinausgehen.

Zu diesem Zweck wird im Folgenden ein Modell des Arbeitsgedächtnisses vorgestellt, das sich seit drei Jahrzehnten von hohem heuristischem Wert erwiesen hat und das einen kognitionspsychologischen Zugang zur Lärmwirkungsforschung bieten kann – das *Working Memory Modell* von Baddeley (2000a).

Nach der Vorstellung des Working Memory Modells von Baddeley im folgenden Abschnitt (Kapitel 1.2), wird die eben nur knapp angerissene Forschung zu den Wirkmechanismen von Lärm auf kognitive Leistungen im Rahmen zweier Subsystem dieses Modells (phonologische Schleife: Kapitel 1.3; zentrale Exekutive: Abschnitt 1.4) diskutiert. Hierauf wird dann in den Abschnitten 1.5 bis 1.7 auf Aufgaben eingegangen, die Schlussfolgerungsprozesse, Kopfrechnen und selektive Aufmerksamkeit erfordern. Laut der modelltheoretischen Basis sollten sich diese Aufgaben selektiv sensitiv auf die Wirkung verschiedener Hintergrundschalle zeigen.

1.2 Ein Modell des Arbeitsgedächtnisses

Im Folgenden soll ein Arbeitsgedächtnismodell vorgestellt werden, das als Ausgangspunkt für viele andere Theorien gedient hat. Das *Working Memory Modell* von Baddeley (2000a) weist eine übersichtliche, gut operationalisierbare Struktur mit vier Subsystemen auf. Es gilt als heuristisch wertvoll, wird jedoch auch teilweise als zu vereinfachend kritisiert und ist bereits vom Autor selbst überarbeitet worden (Baddeley, 2000a). Für das hier vorliegende Anliegen einer kognitionspsychologischen Grundlage für die Lärmwirkungsforschung jedoch ist es weiterhin von hohem praktischem Wert.

Das *Working Memory Modell* von Baddeley und Hitch ist eines der bekanntesten Arbeitsgedächtnismodelle. Es wurde erstmals 1974 veröffentlicht (Baddeley & Hitch, 1974) und seitdem häufig bewertet, getestet und überarbeitet (Baddeley, 2000a; Baddeley, 1986; Baddeley & Logie, 1999). Aus den ehemals drei Subsystemen des Modells, der *zentralen Exekutive*, dem *visuell-räumlichen Notizblock* und der *phonologischen Schleife* wurde ein gut evaluiertes Modell. Dieses wurde inzwischen innerhalb seiner Komponenten erweitert, indem jüngst ein *episodischer Speicher* und Verbindungen zum Langzeitgedächtnis in das Modell integriert wurden (Baddeley, 2000a, vgl. Abbildung 1).

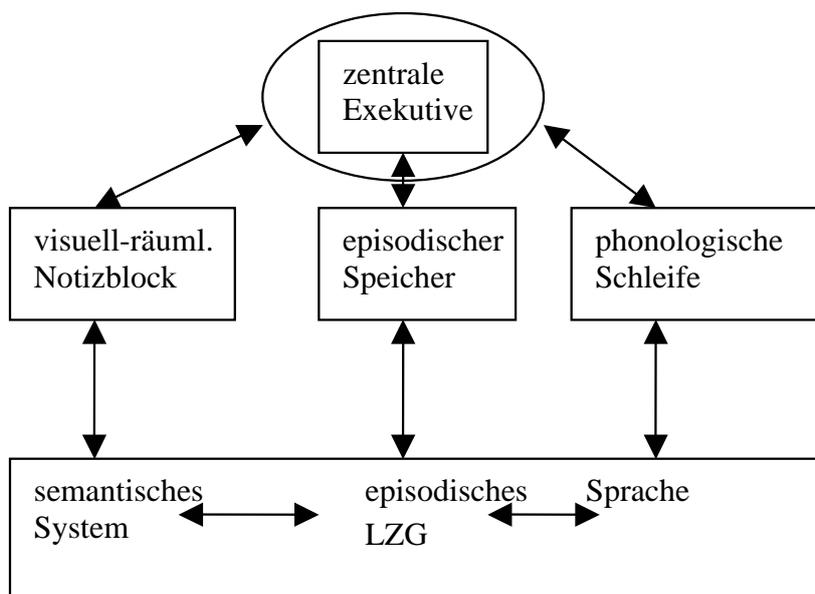


Abbildung 1: Das Working Memory Modell verändert nach Baddeley (2000a)

Der übergeordneten Kontrollinstanz des Modells, der zentralen Exekutive, werden laut dem Autor (Baddeley, 1996) Planungs-, Kontroll- und Integrationsaufgaben zugeschrieben. Diese Instanz geht in ihrer Konzeption auf das *Supervisory Attentional System* von Shallice und Burgess (1993) zurück. Diesem System wurde von den Autoren vor allem die Aufgabe zugeteilt zu entscheiden und zu planen, Fehler zu korrigieren, überlernte automatisierte Reaktionen falls nötig zu hemmen und einzugreifen, wenn neue Reaktionen gefordert sind oder eine Gefahr besteht.

Die Aufgaben, die der zentralen Exekutive im Working Memory Modell zugewiesen werden, kann man grob in drei Bereiche aufteilen: Aufgaben der Kontrolle und Korrektur, Integrationsleistungen und Planungsaufgaben.

Zu den Kontroll- und Korrekturaufgaben gehört die selektive Aufmerksamkeit. Diese bezeichnet die Fähigkeit, die Aufmerksamkeit auf bestimmte Merkmale einer Aufgabe zu fokussieren oder auf andere Merkmale zu verlagern und Unwichtiges *auszublenden*. Dies kann bedeuten, dass die Verarbeitung von relevanten Stimuli durch störende irrelevante Stimuli beeinträchtigt wird, da die zentrale Exekutive Kapazität für die Zurückweisung der irrelevanten Reize aufbringen muss. Dieser Effekt wird größer, wenn relevante wie irrelevante Reize in derselben Modalität dargeboten werden (Baddeley, 1996). Ebenso nimmt die Beeinträchtigung zu, je ähnlicher sich die irrelevanten und die relevanten Stimuli sind. Irrelevante Stimuli, die *in den Fokus der Aufmerksamkeit* gelangen, weil sie etwas mit den relevanten Reizen gemeinsam haben, müssen teilweise verarbeitet werden, bevor sie als irrelevant erkannt und zurückgewiesen werden können. Damit fungiert die zentrale Exekutive als Aufmerksamkeitskontrolle (Baddeley, 1993).

Eine weitere Kontroll- und Korrekturaufgabe der zentralen Exekutive ist die Hemmung von Automatismen. Eine klassische Aufgabe hierfür ist der so genannte Stroop Test (Stroop, 1935). In dieser Aufgabe muss der Proband die Farbe von farbig gedruckten Farbworten benennen. Die Leistung verlangsamt sich hierbei, sobald das Wort in einer anderen Farbe gedruckt ist, als die, die seiner Bedeutung entspräche. Hierbei muss laut dem Autor zunächst der automatisierte Prozess des Lesens gehemmt werden, was aber zu einem Zeitverlust bei der Antwort führt.

Der zweite Aufgabenbereich der zentralen Exekutive, die Integrationsleistung, lässt sich vor allem in der Koordination der anderen Subsysteme (visuell-räumlicher Notizblock, phonologische Schleife und episodischer Speicher) finden sowie im Zugriff auf Inhalte des Langzeitgedächtnisses und damit auch auf jene semantischer Strukturen (Baddeley, 1993). Dieser Zugriff wird jedoch durch die Arbeitsgedächtnisspanne auf fünf bis sieben Items limitiert.

Der dritte genannte Aufgabenbereich der zentralen Exekutive umfasst vor allem Planungsaufgaben, wie z. B. Spielkarten in diverse Kategorien zu sortieren (Baddeley, 1986).

Da die zentrale Exekutive als Kontroll-, Planungs- und Integrationsinstanz konzipiert ist, ergeben sich Schwierigkeiten der Operationalisierung und der spezifischen Hypothesentestung mit Aufgaben, die *nur* einer Kontrolle, Planung oder Integration bedürfen, aber keine Anforderungen an die Hilffsysteme stellen. Aus diesem Grund kann von den Autoren bisher leider wenig über die genaue Organisation dieser Instanz und eine eventuelle Unterteilung in verschiedene Bereiche ausgesagt werden (Baddeley & Logie, 1999).

Weitere Strukturen im Working Memory Modell sind, wie oben angesprochen, die phonologische Schleife und der visuell-räumliche Notizblock. Diese beiden Subsysteme sind darauf spezialisiert, Material, das in einem verbalen oder visuell-räumlichen Kode vorliegt, zu verarbeiten und kurzzeitig zu speichern. Im Falle der phonologischen Schleife erfolgt dies mit Hilfe eines *subvokalen Rehearsalprozesses*, der Informationen im so genannten *phonologischen Speicher* aufrechterhält. Aufgenommen werden verbale Informationen in dieses System direkt über den *visuellen* oder *auditiven Input*. Analog verhält es sich im visuell-räumlichen Notizblock – hier wird ein so genannter

Inner Scribe Prozess angenommen, mit dem visuell-räumliche Informationen im *visuell-räumlichen Speicher* aufrechterhalten werden.

Die vierte Komponente, der *episodische Speicher*, wurde in dem Modell als ein Speichersystem hinzugenommen, in dem – kontrolliert von der zentralen Exekutive – Informationen aus den beiden anderen Systemen und dem Langzeitgedächtnis bewusst integriert und manipuliert werden können (Baddeley, 2000a).

Das Working Memory Modell von Baddeley wurde in der Literatur auch kritisiert (z.B. Jones, Alford, Macken, Banbury, & Tremblay, 2000; Neath, 1999; Neath, Surprenant, & LeCompte, 1998; Surprenant, LeCompte, & Neath, 2000). Es soll hier nicht als vollständig oder unangreifbar verstanden werden. Es gibt sicherlich detaillierter ausgearbeitete Modelle (z.B. Cowan, 1999) als das hier vorgestellte. Modelle sollten aber nicht als Abbilder der Wirklichkeit verstanden werden, sondern vielmehr als Wegweiser für die Forschung (heuristische Funktion). Als solches ist ein übersichtlich strukturiertes Modell hilfreich, wenn die Fragestellung selbst nicht sehr ins Detail gehen kann. Natürlich darf das Modell in sich nicht fasch sein. Mit seiner klaren und übersichtlichen Struktur sowie seiner Verwendung bei der Erforschung spezifischer Lärmeffekte (Wirkungen irrelevanten Sprachschalls, siehe Kapitel 1.3) kann meiner Meinung nach das Working Memory Modell einen soliden Rahmen für eine theoriegeleitete Untersuchungsplanung zu Lärmwirkungen auf kognitive Prozesse bilden. Aus diesem Grunde sollen im Folgenden Hypothesen zu Lärmwirkungen auf die kognitive Leistungsfähigkeit aus diesem Modell abgeleitet werden.

Nachfolgend werden anhand zweier Subsysteme des Working Memory Modells, der phonologischen Schleife (Kapitel 1.3) und der zentralen Exekutive (Kapitel 1.4) mögliche Mechanismen für Lärmwirkungen vorgestellt. Zugleich werden, soweit aus der Literatur ersichtlich, die jeweils vermuteten störungsrelevanten Merkmale des Schalls spezifiziert.

1.3 Interferenzen in der phonologischen Schleife zwischen Schall und bestimmten Aufgaben - der Irrelevant Sound Effect

Eine Möglichkeit der Erklärung von Leistungsbeeinträchtigung durch Hintergrundschalle besteht im Working Memory Modell in einer Interferenz zwischen spektral-temporal stark variablen Schallen und dem seriellen Memorieren in der phonologischen Schleife. Anhand des *Irrelevant Sound Effect (ISE)* wird dies dargestellt.

Der so genannte *Irrelevant Sound Effect* findet erstmals 1976 Erwähnung in der Literatur (Colle & Welsh, 1976). Zu diesem Zeitpunkt wurde dieser Effekt allerdings noch nicht als *Irrelevant Sound Effect* bezeichnet. Die Autoren berichten in mehreren Experimenten von einer leistungsbeeinträchtigenden Wirkung von fremdsprachigem Hintergrundschall auf eine Aufgabe, die das serielle Erinnern von unverbundenen Ziffern oder Konsonanten (so genannte *Serial Recall Aufgabe*; meist 6 bis 9 Items) erfordert. Zunächst wurde dieser Effekt von Baddeley und Salamé untersucht (z. B. Baddeley & Salamé, 1986; Salamé & Baddeley, 1982). Einige Jahre später erweckte er das Interesse einer Forschergruppe um Dylan M. Jones (z. B. Jones, Madden, & Miles, 1992; Jones, Beaman, & Macken, 1996) und wurde in neuerer Zeit von vielen weiteren Forschergruppen aufgegriffen (z. B. Buchner, Irmen & Erdfelder, 1996; Ellermeier & Hellbrück, 1998; Hellbrück, Kuwano & Namba, 1996; LeCompte, Neely & Wilson, 1997; Neath, 2000).

Zur Erklärung des Irrelevant Sound Effect gibt es demgemäß verschiedene Theorien und Modelle. Im Working Memory Modell von Baddeley (2000a) wird er als eine Interferenz im phonologischen Speicher (Abbildung 2) zwischen artikulierten visuell aufgenommenen bzw. gehörten Items und einem irrelevanten Störgeräusch erklärt.

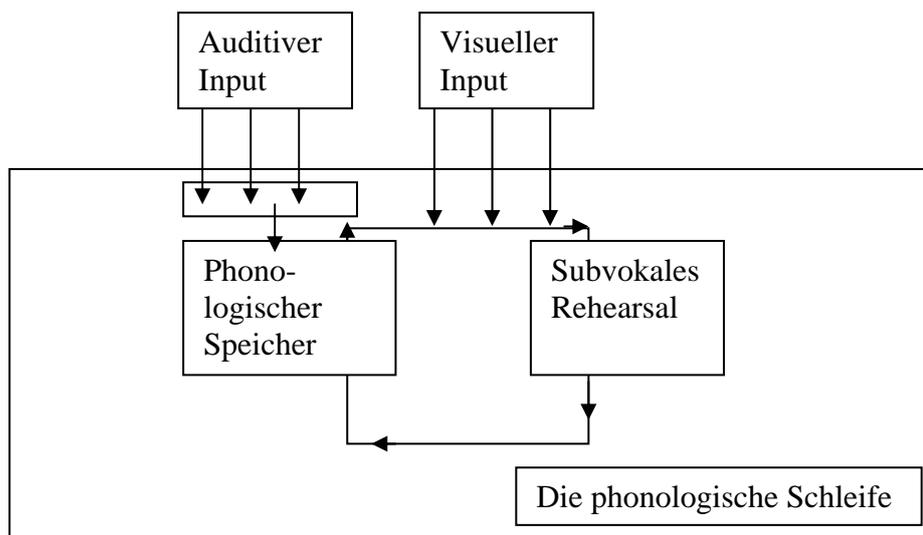


Abbildung 2: Die phonologische Schleife (verändert nach Baddeley & Hitch, 1974)

Zunächst betonte der Autor die Relevanz der phonologischen Ähnlichkeit des irrelevanten Hintergrundprechens mit den zu memorierenden Items. Wie sich jedoch in weiteren Untersuchungen

herausstellte, ist das nicht das wichtigste Merkmal (vgl. für eine Diskussion Baddeley, 2000b). Das irrelevante Hintergrundsignal erhält über den *auditiven Input* Zugang zum phonologischen Speicher. Die Serial Recall Aufgabe besteht aus der Präsentation von bis zu neun unverbundenen Items (Buchstaben oder Zahlen), die unmittelbar oder nach einem Retentionsintervall in der korrekten Reihenfolge wiederzugeben sind. Die für die Aufgabe relevante Information wird von den Probanden in der phonologischen Schleife durch den Prozess des subvokalen Rehearsal aufrechterhalten, da diese Aufgabe laut modelltheoretischer Grundlage selektiv die Funktion dieses Subsystems fordert. Hierbei kann es zu einer Interferenz mit dem Störschall kommen. Diese Interferenz wird als *präattentiv* angenommen, da die Probanden den Sprachschall im Arbeitsgedächtnis nicht *ausblenden* oder daran *habituierten* können (Hellbrück, Kuwano & Namba, 1996). Dies bedeutet, dass dieser Effekt nicht auf eine durch Habituation sich abschwächende Orientierungsreaktion, die durch den Schall ausgelöst wird, zurückzuführen ist.

Im Working Memory Modell ist die Erklärung des ISE in eine Theorie des Arbeitsgedächtnisses eingebettet. Weitere Modelle zur Erklärung des ISE gibt es noch von Jones und Kollegen, die mit ihrem *O-OER Modell* eine Spezifikation der zugrunde liegenden Prozesse vorlegten (Jones et al., 1996), wodurch dieses jedoch einen wesentlich geringeren Generalitätsanspruch erheben kann als das Working Memory Modell. In eine ähnliche Richtung gehen die Ansätze von Neath und Kollegen mit ihrem *Feature Modell* (Neath, 2000), das ebenfalls Teil eines umfassenderen Modells des Arbeitsgedächtnisses sein könnte (Baddeley, 2000b). Der Forschergruppe um Jones ist es vor allem zu verdanken, dass die zum ISE nötigen Bedingungen spezifiziert wurden, so etwa die Relevanz des *Changing State Characters*, einer temporal-spektralen Strukturierung des Schallsignals, auf die im Folgenden noch näher eingegangen wird (Jones et al., 1992). Das Modell von Neath und Kollegen stellte hingegen eine mathematische Annäherung an den ISE dar, anhand dessen auch Simulationen möglich wurden (Neath, 2000). Nachfolgend soll aber weiter der Bezug zum Working Memory Modell von Baddeley (2000a) gewahrt bleiben, da dieses das umfassendste der erwähnten Modelle ist und in ihm die Erklärung des ISE nur einen Teilaspekt darstellt. Dieses Modell kann somit auch für andere Mechanismen als den ISE den modelltheoretischen Zugang zu Lärmwirkungen erleichtern.

Beim *Irrelevant Speech Effect* oder *Irrelevant Sound Effect*, wie er seit Beginn der 1990er Jahre genannt wird, wurden (wie bereits erwähnt) diverse Manipulationen des Schalls untersucht, um zu identifizieren, welche Parameter relevant für das Hervorrufen einer Störwirkung sind.

Da bereits in der ersten Untersuchung (Colle & Welsh, 1976) eine Interferenzwirkung fremdsprachigen Schalls nachgewiesen werden konnte, wurde der Aspekt der Verständlichkeit des Sprachschalls mehrfach zum Untersuchungsgegenstand. Es zeigte sich aber generell eine Unabhängigkeit des Effekts von der Bedeutungshaltigkeit des Schalls (Klatte et al., 1995). Selbst die semantische Ähnlichkeit des Hintergrundschalls mit dem zu lernenden Material hat keinen Einfluss auf die Leistung, was sich in einer ähnlichen Störwirkung von Zahlen- und Buchstabenreihen als Hintergrundschalle beim Lernen von Ziffernfolgen zeigen lässt (Buchner, Irmen, & Erdfelder, 1996).

Ein weiteres Merkmal der Schalle, das sich als unerheblich für die Höhe des ISE erwiesen hat, ist die Lautstärke des irrelevanten Schalls (Tremblay & Jones, 1999). Es zeigen sich keine

Unterschiede in der Leistungsbeeinträchtigung durch Hintergrundsprache mit 60 dB(A) und 75 dB(A) (Ellermeier & Hellbrück, 1998).

Die wichtigste Rolle spielt beim ISE wohl die Zeitstruktur des Hintergrundschalls. In diesem Punkt, der intensiv von Dylan M. Jones untersucht wurde (vgl. Jones et al., 1992; Jones et al., 2000), herrscht inzwischen Einigkeit unter allen Forschergruppen. Die Störwirkung kann durch eine Glättung der Zeitstruktur reduziert werden, was z. B. anhand einer sukzessiven Verrauschung eines Sprachsignals erreicht werden kann (Ellermeier & Hellbrück, 1998) oder durch Überlagerung mehrerer Stimmen (*Irrelevant Babble Effect*). Die beeinträchtigende Wirkung nimmt auch durch eine graduelle Entfernung der Stimme vom Hörer ab, was zu Unverständlichkeit und Verschwommenheit führt (Kilcher & Hellbrück, 1993), oder allgemein durch eine Abnahme der physikalischen Variabilität im Schall, z.B. durch eine schrittweise Filterung. Diese Effekte sind jedoch unabhängig von der Sprachartigkeit des Schalls (ähnliche Ergebnisse für Sprache und Tonfolgen) (Jones et al., 2000). Auch instrumentale oder vokale Stakkatomusik (Klatte et al., 1995) sowie Folgen verschiedener Töne mit Pausenstruktur und Frequenzmodulation zeigen eine Störwirkung auf das serielle Erinnern. Es geschieht jedoch nicht in dem Ausmaß wie bei Buchstabenfolgen oder bedeutungsloser Sprache (LeCompte et al., 1997; Salamé, 1990). Das kritische Element im Schall scheint der so genannte *Changing State Charakter* zu sein, womit eine Abfolge von distinkten, aufeinander folgenden, aber voneinander unterscheidbaren Wahrnehmungseinheiten bezeichnet wird (Jones et al., 1992; LeCompte, 1995; Tremblay, Macken, & Jones, 2001). Ein wiederholter Ton, ein sich wiederholendes Wort oder ein gleichmäßiges Geräusch haben demnach keine Störwirkung.

Andere Autoren testeten verschiedene Gedächtnis- und Aufmerksamkeitsaufgaben unter Sprachschall und einem der zeitlichen Verlaufsform nach ähnlichen Verkehrsschall und fanden eine Interferenz beider Schalle mit dem Erinnern eines Textes, dem Zugriff auf das semantische Gedächtnis und der Aufmerksamkeit (Hygge, Boman, & Enmarker, 2003). Auch bei diesen Experimenten ließ sich eine Störwirkung auf das Rehearsal und die Speicherung von Informationen nachweisen. Dies bedeutet, dass auch Verkehrslärm mit starker zeitlicher Strukturierung die Leistung der phonologischen Schleife beeinträchtigen kann. Laut der eben vorgestellten Erklärung des ISE im Working Memory Modell dürften alle Aufgaben, die die phonologische Schleife beanspruchen, sensitiv gegenüber der Wirkung eines irrelevanten Sprachschalls beziehungsweise eines spektral-temporal stark variablen Schalls (*Changing State Charakter*) sein, der in diesem Subsystem mit den memorierten Inhalten interferieren kann.

Wie jedoch kann eine mögliche beeinträchtigende Wirkung eines lauten Schalls (z.B. eines gleichmäßigen Verkehrsschalls) auf andere kognitive Funktionen erklärt werden (wie sie in Kapitel 1.1. angesprochen wurden), die nicht das Memorieren sequentieller Informationen erfordern, sondern beispielsweise Aufmerksamkeitsprozesse? In der Konzeption des Working Memory Modells bleibt die Struktur der zentralen Exekutive, die für diese Prozesse verantwortlich ist, immer noch etwas unklar in ihrem Aufbau. Eine theoretische Konzeption, die Schallwirkungen auf Aufmerksamkeitsprozesse erklären kann, liegt in der Forschungstradition zum so genannten *Arousal* vor. Die wichtigsten Erklärungsansätze und Befunde zu dieser Konzeption im Zusammenhang mit Aufmerksamkeitsprozessen sollen im folgenden Kapitel dargestellt werden.

1.4 Lärmwirkungen auf die zentrale Exekutive - Arousal im Rahmen des Aufmerksamkeitskonzepts

In der Entwicklung des empirischen Aufmerksamkeitsbegriffs kann man nach Neumann (1996) drei Komponenten unterscheiden. Die Aufmerksamkeit wird als begrenzt (Kapazitätsaspekt) und selektiv angesehen (Selektionsaspekt) und sie kann willentlich ausgerichtet werden (Tätigkeitsaspekt). Auf die Komponente der willentlichen Ausrichtung und damit den Tätigkeitsaspekt wurde in neueren Ansätzen wieder vermehrt Gewicht gelegt. Dies ist unter anderem auch im Supervisory Attentional System (Shallice & Burgess, 1993) der Fall, auf welches sich Baddeley (2000a) in seinem Working Memory Modell mit der Konzeption der zentralen Exekutive (siehe oben) bezieht. In Bezug auf den Kapazitätsaspekt, die Begrenztheit der Aufmerksamkeit, gibt es Forschungen zum so genannten *Arousal* und seinem Zusammenhang mit Aufmerksamkeit und Leistung, die häufig auch die Selektivität der Aufmerksamkeit berücksichtigen. Einige der diesbezüglichen wichtigsten Theorien sollen vorgestellt werden, nachdem der Begriff des *Arousal* – auch in seiner physiologischen Bedeutung – allgemein eingeführt wurde.

Der Begriff *Arousal* soll nachfolgend im Sinne eines unspezifischen kortikalen Erregungsniveaus verwendet werden. Bereits 1949 wurde dieses Konzept von einer Forschergruppe um Lindsley (Lindsley, Bowden, & Magoun, 1949) in Bezug auf die *Formatio Reticularis* und das so genannte *unspezifische aufsteigende retikuläre Aktivierungssystem (ARAS)* untersucht. Dieses System besteht aus einem Fasergeflecht, das für den Schlaf-Wach-Rhythmus und ein allgemeines tonisches Erregungsniveau kortikaler Neuronenverbände sorgt (Birbaumer & Schmidt, 1996). Die *Formatio Reticularis* hat vor allem die Aufgabe, kurz dauernde Erregungsphasen und tonische Wachheit zu generieren, Einfluss auf die tonische Kontrolle der Muskulatur zu nehmen und die Aufnahme, Leitung und Integration aller sensorischen Signale zu kontrollieren (Birbaumer, 1975).

Lindsley und Kollegen stellten in Tierexperimenten fest, dass sensorische Stimulationen über dieses Fasergeflecht zu einer Zunahme der allgemeinen kortikalen neuronalen Feuerrate führen (Lindsley et al., 1949). Aus der Erforschung der kortikalen Verarbeitung von Geräuschen weiß man, dass vor allem tieffrequenter breitbandiger Schall zu einer großflächigen Erregung bereits der Basilarmembran im Innenohr führt. Dabei wird eine steigende Lautstärke durch eine Erhöhung der Anzahl und der Feuerungsrate der Neuronenverbände der Cochlea kodiert (Hellbrück & Ellermeier, 2004). Hieraus resultiert offenbar auch eine Zunahme der allgemeinen kortikalen Erregung, was als steigendes *Arousal* bezeichnet wird (Lindsley et al., 1949).

Zudem scheint auch tieffrequenter Schall zu einer stärkeren Erhöhung des *Arousal*s zu führen als mittel- oder höherfrequenter Schall. Dies konnte in Untersuchungen zu Änderungen der Hirnströme (gemessen über EEG-Ableitungen) während des Schlafs gezeigt werden. Geräusche mit 50 oder 250 Hz führten in einer Untersuchung im Schlaf bei Probanden zu einer stärkeren Erregungsbildung (gemessen über das EEG) als ein Vergleichsgeräusch (in zuvor subjektiv eingestellter gleicher Lautheit) mit 1 kHz (80 dB) (LeVere, Morlock, Thomas, & Hart, 1974). Tieffrequente Geräusche werden mit gleichem Schallpegel (in dB(A) oder auch in dB) im Allgemeinen als lästiger beurteilt als höherfrequente (Berglund et al., 1996). Als tieffrequent wird der Bereich bis ca.

250 Hz oder auch 500 Hz bezeichnet. Da tieffrequenter Schall zudem schlechter abgeschirmt werden kann, stellt dieser eine durch die A-Filter-Messung häufig unterschätzte Belästigungsquelle dar.

Aufgrund der Wichtigkeit von Variablen wie Lautstärke oder Frequenzzusammensetzung von Lärm im Zusammenhang mit dem Arousal soll nun kurz auf die Messung der Lautstärke von Lärm und der Wahrnehmung von Frequenzen eingegangen werden.

Um dem großen Dynamikbereich des menschlichen Ohrs gerecht zu werden, wird zur Darstellung des gehörbezogenen Schalldrucks eine logarithmische Skala verwendet. Die logarithmische Transformation des Schalldrucks bezeichnet man als Schall(druck)pegel. Die Schallpegel-Skala hat die Einheit *Dezibel* (dB), wobei 0 dB einem gerade hörbaren Schalldruck und 120 dB in etwa der Schmerzgrenze entspricht. Bei der Bestimmung der Lautstärke von Verkehrsgläuschen über einen bestimmten Zeitraum wird im Allgemeinen der äquivalente Dauerschallpegel (L_{eq}) angegeben. Dieser stellt den über diesen Zeitraum auf Energiebasis gemittelten Schallpegel dar und wird im Allgemeinen *Mittelungspegel* genannt (Hellbrück & Ellermeier, 2004).

Zwischen 20 und 20000 Hertz sind Geräusche für den Menschen wahrnehmbar. Zu berücksichtigen ist, dass die Lautstärkeempfindung frequenzabhängig ist. Dies kommt in den *Kurven gleicher Lautstärke (Isophone)* zum Ausdruck. Um dieser Frequenzabhängigkeit des Gehörs Rechnung zu tragen, wird der Schallpegel meist frequenzgewichtet gemessen. Es gibt drei Filter, die hierfür verwendet werden. Sie werden als A, B und C bezeichnet und senken sowohl im tiefen Bereich (A hier am meisten, B tendenziell weniger und C am wenigsten) als auch im hochfrequenten Bereich den Schalldruckpegel ab, wie es den Isophonen annähernd entspricht (A-Filter bei niedrigen, B bei mittleren, C bei hohen Pegeln). Vereinbarungsgemäß benutzt man bei der Lärmmessung das A-Filter (Hellbrück & Ellermeier, 2004).

Doch wie hängt das Arousalniveau mit der Leistungsfähigkeit der zentralen Exekutive zusammen? Es wurde vielfach versucht, eine Antwort auf diese Frage zu finden. Einige der zum Teil auch widersprüchlichen Ergebnisse hierzu und ihre Einbettung in theoretische Konzepte sollen jetzt vorgestellt werden.

Bereits zu Beginn des letzten Jahrhunderts wurde nach den Forschungen von Yerkes und Dodson (1908) das so genannte *Yerkes-Dodson Gesetz* formuliert. Es besagt, dass die Leistung durch steigendes Arousal bis zu einem gewissen Niveau profitiert, dann aber bei weiterer Steigerung abfällt, und zwar umso eher, je komplexer die verlangte Leistung ist. Dieser Zusammenhang wurde als umgekehrt U-förmig postuliert (Yerkes & Dodson, 1908), und konnte auch in empirischen Untersuchungen belegt werden. Es konnte nachgewiesen werden, dass die Leistung in komplexen Aufgaben, wie postuliert, durch physikalische (Lärm, Hitze), kognitive und emotionale Stressoren (Motivation, Konkurrenz) erschwert werden kann, d. h. es kommt zu vermehrten Fehlern oder längerer Bearbeitungszeit. In einfachen Aufgaben hingegen lässt sich eine Erleichterung zeigen, die Leistung wird durch steigendes Arousal eher besser (Van Gemmert & Van Galen, 1997). Arousal lässt sich im Sinne des Yerkes-Dodson Gesetzes als eine einheitliche, unspezifische Energiekomponente verstehen, die bei einem Anstieg eine erhöhte Erregung bewirkt, was sich wiederum auf die Leistung

auswirkt (vgl. auch Hockey, 1979). Ein weiterer Effekt von gesteigertem Arousal durch Lärm sei eine Verschiebung des *Speed-Accuracy-Trade-Off*. Die Probanden würden zwar schneller arbeiten aber mehr Fehler machen (Kahneman, 1973).

Von Broadbent (1958b) wird postuliert, dass es zur Fehlreaktion eines Aufmerksamkeitsfilters im Sinne einer *Verengung (narrowing)* der Aufmerksamkeit kommt. Hierdurch lässt sich jedoch nicht erklären, wie es unter Lärm zu einer gesteigerten Leistung in einer Primäraufgabe (Tracking) und gleichzeitig zu einer Verschlechterung in einer Sekundäraufgabe (Beobachtung von peripher verteilten Lämpchen) kommen kann (Hockey, 1978). In der Sekundäraufgabe war bei den Experimenten von Hockey (1978) sogar zu beobachten, dass die Probanden bei den zentraler im Gesichtsfeld liegenden Stimuli eine Leistungsverbesserung unter Lärm (100 dB(A)) im Gegensatz zur Vergleichsbedingung (70 dB(A)) zeigten, bei jenen in der Peripherie aber eine Verschlechterung. Dies spricht stärker für eine konstante Energiemenge, die durch steigendes Arousal umverteilt wird, wie es beispielsweise von Easterbrook (1959) postuliert wird. Eine Steigerung des Arousal würde hierbei zu einer erhöhten Selektivität und Einengung der Aufmerksamkeit führen. Dies bewirke eine Verbesserung in einfachen Aufgaben, führe aber bei anspruchsvollen Aufgaben zu einem Einbruch der Leistung, da diese häufig eine geteilte Aufmerksamkeit auf diverse Reize erfordern (Easterbrook, 1959). Die Komplexität der Aufgabe wird definiert durch die Anzahl der Hinweisreize, die gleichzeitig verwendet werden müssen, um das von der Aufgabe geforderte Leistungsniveau zu erreichen (van der Molen, 1996). Der Zusammenhang zwischen Leistung und Erregungsniveau über eine Einengung der Aufmerksamkeit konnte anhand von Doppelaufgaben belegt werden, bei denen sich die Leistung in der Zweitaufgabe bei Hinzukommen von Stressoren wie Lärm verschlechtert. Dass die Leistung nicht in beiden Aufgaben leidet, wird von Easterbrook (1959) dahingehend interpretiert, dass die Selektivität der Aufmerksamkeit zunimmt und nur noch auf eine Aufgabe gerichtet werden kann. Dieser Zusammenhang zwischen Leistung und Arousal wäre nicht U-förmig, sondern linear (Hockey, 1979) und konnte auch in Untersuchungen bezüglich der Wirkung von zu geringer Erregung in Form von Müdigkeit und steigendem Arousal durch Lärmbelastung bei unterschiedlicher Aufgabenanforderung gezeigt werden. Smith (1985) konnte bei einer Aufgabe mit einer Folge von großen Buchstaben, die aus kleinen zusammengesetzt waren (z.B. ein *L*, dessen beide Linien aus vielen kleinen *M* bestehen), unter Breitbandrauschen (85 dB) versus einer Vergleichsbedingung (60 dB) eine bessere Erinnerungsleistung für die kleinen Buchstaben und schlechteres Memorieren der großen Buchstaben finden, was sich ebenfalls über die Erhöhung der Selektivität des Aufmerksamkeitsfokus erklären lässt. Entsprechendes konnte auch in einer funktionellen Magnetresonanztomographie-Untersuchung bestätigt werden. Lärm (80 dB(A)) führte zu einer Verengung der Aufmerksamkeit. Dies wurde gefolgert aus dem Unterschied in der Neuronenfeuerrate zwischen Aufgaben, die einen weiten Fokus der Aufmerksamkeit verlangen, und solchen, die eine Fokussierung fordern (Tracy et al., 2000).

Zudem scheint die Speicherkapazität des Kurzzeitgedächtnisses mit zunehmendem Arousal zu sinken, wohingegen die für eine Aufgabe zur Verfügung stehenden Aufmerksamkeitsressourcen gleichzeitig ansteigen (von-Hecker, Dutke, & Sedek, 2000). Dies konnte durch geringere Leistung im *Sternberg Test* und bessere Leistung in einem Test zur *geteilten Aufmerksamkeit* (Erkennen von bestimmten Zielreizen in der Peripherie des Gesichtsfeldes, während gleichzeitig ein sich bewegender Balken innerhalb gewisser Grenzen gehalten werden musste) bei höherem Arousalniveau (erfasst

über einen Fragebogen) gezeigt werden. Dies erklären die Autoren damit, dass in einem Zustand hohen Arousal die zur Verfügung stehenden Ressourcen zugunsten der aktuellen Verarbeitung von Reizen verteilt werden (zugunsten der zentralen Exekutive im Working Memory Modell von Baddeley) und zuungunsten der Speicherung von Informationen (z. B. in der phonologischen Schleife). Auch dies kann wieder mittels einer Erhöhung der Selektivität erklärt werden, da die wichtigsten Prozesse in der Erregungssituation bevorzugt werden. Die bei hoher Erregung gleichzeitig auftretende erhöhte Ablenkbarkeit erklärt Kahneman (1973) damit, dass der Aufmerksamkeitsfokus durch steigendes Arousal zwar schmaler würde, sich aber schneller bewegen könne. Dies lasse sich in Experimenten mit Drogen zeigen, wo unter Stimulantien bei gleichzeitiger Einengung des Gesichtsfeldes verstärkte Sakkaden der Augen auftreten.

Zusammenfassend kann man die Wirkung von Lärm auf die Arbeit der zentralen Exekutive vor allem dahingehend eingrenzen, dass es bei einer Arousalerhöhung (z. B. durch tieffrequenten, lauten Schall) zu einer Steigerung der Schnelligkeit auf Kosten der Genauigkeit kommt, die Selektivität der Aufmerksamkeit und der Antworten zunimmt, die Arbeitsgedächtniskapazität eingeschränkt wird (Hockey, 1979) und Probanden eher nicht statt falsch antworten, was sich auch wieder über eine höhere Selektivität erklären ließe (Hockey, 1978). Weiterhin scheint Einigkeit darüber vorzuliegen, dass schwierige Aufgaben unter Lärm erschwert werden, leichte jedoch besser bewältigt werden können. Dies wird sowohl durch die steigende Selektivität des Aufmerksamkeitsfokus, der im Working Memory Modell durch die zentrale Exekutive gesteuert wird, als auch durch seine verstärkte Ablenkbarkeit erklärt. Das Arousal scheint kein einheitliches Konzept zu sein, da es auch Hinweise auf ein sehr allgemeines Aktivierungsniveau des Körpers und ein eher emotional gefärbtes Arousal gibt (Boucsein & Ottmann, 1996). Wichtig für die eingangs gestellte Frage nach den Wirkungen von Lärm auf kognitive Leistungen ist allerdings nur das tätigkeitsbezogene Arousal. Relevant würde das emotionale Arousal dann werden, wenn es zu einem Leistungsabfall unter Lärm kommt, da ablenkende Gedanken oder andere kognitive Reaktionen auf den Stress, z.B. Ärger über die Lärmexposition, hinzukommen (vgl. Näätänen, 1975). Dies kann somit bei der Leistungsbeeinträchtigung unter Lärm eine wichtige Moderatorvariable darstellen.

Es wurden bisher zwei mögliche Wirkungswege von Lärm auf kognitive Leistungen im Working Memory Modell dargestellt. In Kapitel 1.3 wurde die störende Wirkung von spektral-temporal stark variablem Schall auf die phonologische Schleife dargelegt. Weiterhin wurde mit dem Konzept des Arousal eine mögliche Leistungsbeeinträchtigung durch laute tieffrequente Schalle im System der zentralen Exekutive erläutert.

Um die Angemessenheit dieser theoretischen Konzepte für die Realität von Schallwirkungen zu überprüfen, sollen im Folgenden Aufgaben vorgestellt werden, die zu ihrer Operationalisierung dienen können. Ein Test, der selektiv die phonologische Schleife beansprucht, ist der in Kapitel 1.3 bereits beschriebene Serial Recall. Der so genannte Stroop Test hingegen, auf den in Kapitel 1.7 eingegangen wird, erfordert zu seiner Bearbeitung in erster Linie die Funktion der zentralen Exekutive. Aufgaben, die beide Subsysteme des Working Memory Modells ansprechen, sind der Konzentrations-Leistungstest (Düker & Lienert, 1965), ein komplexer Test, zu dessen Lösung unter anderem Kopfrechnen notwendig ist, und das grammatikalische Schlussfolgern, was in Kapitel 1.5 und 1.6

erklärt wird. Werden diese Aufgaben nun unter Sprachschall oder lautem, tieffrequentem Lärm durchgeführt, ließe sich die Angemessenheit der bisher beschriebenen theoretischen Konzepte empirisch überprüfen.

1.5 Die Verwendung von Rechentests in der Lärmwirkungsforschung

Zunächst sollen einige Effekte Erwähnung finden, die beim Kopfrechnen bedacht werden müssen, um das Schwierigkeitsniveau entsprechender Tests einzuschätzen. Einer dieser Effekte ist der *Problem-Size Effect*. Für die Berechnung größerer Zahlen wird eine längere Zeit in Anspruch genommen, da mehr Rechenoperationen nötig sind (Donlan, 1998). Ebenso verlängert sich die Lösungszeit, wenn Lösungen in umgekehrter Reihenfolge der Berechnung aufgeschrieben werden sollen (zuerst die Hunderter, dann die Zehner und Einser). Dies erfordert das mentale Speichern der berechneten Stellen. Ähnliche Ergebnisse werden durch Erhöhung der Anzahl von Überträgen oder durch Verkürzung der visuellen Präsentation der Aufgabe erzielt – alles zusätzliche Anforderungen an das Arbeitsgedächtnis (Hitch, 1978). Diese Effekte haben Einfluss auf die Belastung des Arbeitsgedächtnisses bzw. der phonologischen Schleife, was bei den im Anschluss zitierten Untersuchungen zur Leistung unter Hintergrundschall von Wichtigkeit ist.

Man kann, wie nachfolgend erläutert wird, davon ausgehen, dass ein Kopfrechentest Anforderungen vor allem an zwei Subsysteme des Arbeitsgedächtnisses stellt, nämlich an die phonologische Schleife und die zentrale Exekutive.

Dass die phonologische Schleife beim Kopfrechnen eine wichtige Rolle spielt, ließ sich vor allem anhand von Doppelaufgaben zeigen. Die Additionsleistung lässt sich durch artikulatorische Unterdrückung (Aussprechen irrelevanter Wörter oder *Zählen bis vier*) und Darbietung irrelevanter auditiver Ziffernfolgen reduzieren, was sich in einem Anstieg der Fehlerzahlen zeigt (Biewald & Schumann-Hengsteler, 1996; Logie, Gilhooly, & Wynn, 1994). Durch seine serielle Struktur hat vor allem die mentale Addition große Implikationen für dieses Gedächtnissubsystem. Auch Untersuchungen zur Beteiligung des visuell-räumlichen Notizblocks und der phonologischen Schleife beim Kopfrechnen konnten durch entweder visuelle oder phonologische Ähnlichkeit zweier Summanden nur einen Störeinfluss der phonologischen Ähnlichkeit belegen. Das wiederum spricht für eine Beteiligung der phonologischen Schleife beim Bearbeiten der Aufgabe (Noel, Desert, Aubrun, & Seron, 2001). Hierzu existieren jedoch auch gegenteilige Befunde. Thomas, Zoelch, Seitz-Stein und Schumann-Hengsteler (in press) konnten keine beeinträchtigende Wirkung von irrelevanter Hintergrundsprache auf Additionsaufgaben bei Grundschulkindern der dritten und vierten Klasse nachweisen. Dies würde gegen eine Beteiligung des eben genannten Subsystems beim Kopfrechnen sprechen. Bei ausreichender Automatisierung des Kopfrechnens scheint ab einem gewissen Alter der Kinder die phonologische Schleife nicht mehr an dieser Aufgabe beteiligt zu sein. Dies ist jedoch nur dann der Fall, wenn die Aufgabe fortwährend visuell präsentiert wird, was – wie oben erwähnt – Behaltensprozesse tatsächlich überflüssig macht.

Durch geforderte Handbewegungen oder gleichzeitige Präsentation irrelevanter Bilder ließ sich das Rechnen in keiner Untersuchung beeinflussen. Man kann somit einheitlich auf eine geringe Anforderung an den visuell-räumlichen Notizblock schließen (Biewald & Schumann-Hengsteler, 1996; Logie et al., 1994).

Wie Untersuchungen der Forschergruppe um Logie zeigten, verlangen Additionsaufgaben auch zentral exekutive Prozesse (Logie et al., 1994). Dies konnte aufgrund einer hohen Interferenz bei gleichzeitiger Lösung von auditiv oder visuell präsentierten Rechenaufgaben und der Produktion

zufälliger Zahlenreihen geschlussfolgert werden. Bei Multiplikationen sind die Anforderungen an die zentrale Exekutive etwas geringer, da diese im kleinen Einmaleins eher automatisiert werden (Biewald & Schumann-Hengsteler, 1997; Seitz & Schumann-Hengsteler, 2000). So konnten Biewald und Schumann-Hengsteler (1997) mit Hilfe des *dual-task* Paradigmas eine Beeinträchtigung der Additions- und mit Einschränkungen auch der Multiplikationsleistung von Kindern durch gleichzeitiges Produzieren von zufälligen Zahlenreihen (*Random Generation Aufgabe*) im Vergleich zu *Tapping* (mit einem Finger auf die Tischplatte klopfen) finden.

Es existieren in der Literatur keine Untersuchungen von Kopfrechnen unter Sprachschall *und* lautem, tieffrequentem Lärm (z.B. Verkehrslärm), jedoch finden sich einige Untersuchungen, die einen Zusammenhang zwischen dem Arousalniveau, der Lärmbelastung und der Leistung in Kopfrechentests nahe legen. Zudem gibt es Untersuchungen, die leistungsbeeinträchtigende Wirkungen von Sprachschall auf Kopfrechentests gefunden haben. Von beiden sollen einige Ergebnisse kritisch skizziert werden.

Sprachschall scheint Rechenaufgaben unabhängig von seinem semantischen Gehalt beeinträchtigen zu können. Die bereits zitierte Forschergruppe um Logie (1994; vgl. auch Logie & Baddeley, 1987) bot während der Bearbeitung einer visuellen Additionsaufgabe irrelevanten Sprachschall in Form von Ziffernfolgen dar und konnte eine Beeinträchtigung der Rechenleistung in Form einer Erhöhung der Fehlerzahlen zeigen. Diese Erhöhung wies eine ähnliche Effektgröße auf wie die Erhöhung der Fehlerzahlen durch artikulatorische Unterdrückung (gleichzeitiges wiederholtes Aussprechen eines irrelevanten Wortes). Daraus lässt sich schließen, dass die Ziffern des Sprachschalls nicht auf semantischer, sondern rein phonologischer Ebene interferierten. Eine Darbietung von Bürogeräuschen mit oder ohne Sprache und von einer reinen Sprachschallaufnahme (alle 65 dB(A)) während der Bearbeitung eines Kopfrechentests (Additionsaufgaben) ergab eine Leistungsbeeinträchtigung (erhöhte Fehlerzahlen bei unveränderter Bearbeitungszeit) unter allen Schallbedingungen, die von der Bedeutungshaltigkeit des Sprachschalls (Ziffern oder unverständliche Sprache) unbeeinflusst war (Banbury & Berry, 1998).

Kopfrechnen kann somit als sensitiv gegenüber gleichzeitig dargebotenem irrelevantem Sprachschall gelten. Dieser Effekt ist unabhängig von dessen semantischem Gehalt, was darauf hindeutet, dass er auf die spektral-temporale Struktur der Sprache rückführbar ist. Beachtet werden sollte hier allerdings die Stärke der Belastung des Arbeitsgedächtnisses durch die Gestaltung der Aufgaben. Laut der modelltheoretischen Zuordnung der Schalleffekte zur phonologischen Schleife kann das erheblichen Einfluss auf eine Störwirkung spektral-temporal stark variabler Schalle auf Kopfrechentests haben. Dies könnte auch die Divergenz zu den oben bereits erwähnten gegenteiligen Befunden der Forschergruppe um Thomas (Thomas et al., in press) erklären. In ihren Ergebnissen zeigte sich *kein* negativer Einfluss von Hintergrundsprache auf die Kopfrechenleistung. Dies ist jedoch nur dann bei älteren Grundschulkindern durch zunehmende Automatisierung des Rechnens der Fall, wenn keine Behaltensprozesse zur Lösung der Aufgaben notwendig sind. Die Rechnungen wurden in der zitierten Untersuchung die ganze Zeit über visuell präsentiert.

Empirisch konnte weiterhin aufgezeigt werden, dass Kopfrechnen durch gleichzeitige Darbietung irrelevanter lauter nicht-sprachlicher Schalle gestört werden kann. In einer Studie über die Wirkung von Verkehrslärm auf kurze Additions- und Subtraktionsaufgaben in visueller und auditiver

Variante bei Kindern lässt sich eine Störwirkung ab 70 phon² finden. Ab dieser Grenze kommt es zu einer Erhöhung der Fehlerzahlen und der Bearbeitungszeiten sowie zu einer Veränderung des Hautwiderstands, was vom Autor als Erhöhung des Erregungsniveaus interpretiert wird (Maruyama, 1964). Zwischen 65 und 70 phon scheint hierbei eine kritische Grenze der Lautstärke des Verkehrslärms zu liegen.

Die Art des lauten Geräusches scheint nicht relevant zu sein, in einer anderen Untersuchung stellten sich ein 750 Hz-Ton sowie ein Rauschen, bei dem sich Stimmen und Musik überlagern (Radiostörgeräusch) in einer Lautstärke von 90 dB in einem Kopfrechentest (Additionen und Subtraktionen) als leistungsbeeinträchtigend heraus (Fitger & Rettig, 1985). Die Leistungsbeeinflussung unter Lärm scheint nur von der Lautstärke abhängig zu sein. Auch im sehr lauten Bereich ergeben sich noch Unterschiede in der Leistung. Bei einem Vergleich der Subtraktionsleistung unter einem Schall in einer Lautstärke von 100 dB³ mit der unter einem von 70 dB fand Broadbent (1958a) eine Erhöhung der Fehlerzahlen unter dem lauterem Schall. Dies scheint aber nur bei Kopfrechenaufgaben der Fall zu sein, deren Rechenoperationen aufeinander aufbauen oder die durch eine kurze Darbietungsdauer anspruchsvoller werden (Jones & Broadbent, 1991). Da diese beiden Faktoren allerdings (wie eingangs betont) die Ansprüche an das Subsystem der phonologischen Schleife und nicht der zentralen Exekutive erhöhen sollten, ist eine Erklärung im Lichte der bisher vorgestellten Modelle für diese Befunde schwierig. Insgesamt muss aber eine Vergleichsbedingung von 70 dB als *Ruhebedingung* sehr kritisch gesehen werden, da der oben zitierte Autor in diesem Lautstärkebereich bereits beeinträchtigende Lärmefekte berichtete.

Ein Test, der zusätzlich zu den – teilweise umstrittenen (Thomas et al., in press) – Anforderungen des Kopfrechnens an sich die phonologische Schleife und die zentrale Exekutive stark beansprucht, ist der von Düker und Lienert 1965 entwickelte *Konzentrations-Leistungs-Test (KLT)*, der sich daher besonders für den Nachweis von Schallwirkungen auf die Subsysteme des Working Memory Modells eignet. Als Paper-Pencil Version testet diese Aufgabe laut den Autoren die Daueraufmerksamkeit. Diese Aufgabe stellt durch die Notwendigkeit, Zwischenergebnisse zu memorieren, zusätzlich zum Kopfrechnen hohe Anforderungen an die phonologische Schleife. Ebenso wird die zentrale Exekutive durch das Vergleichen von Ergebnissen und das Fällen einer Entscheidung über das weitere Vorgehen über die Beanspruchung durch das Kopfrechnen hinaus gefordert. Dies wird durch eine Serie von Rechenaufgaben erreicht, die anhand der so genannten *Düker-Regel* verrechnet werden müssen. Hierbei werden zweimal jeweils drei positive oder negative Ziffern addiert, deren memorierte Ergebnisse anschließend zusammengezählt werden, wenn das Ergebnis der zweiten Teilaufgabe größer ist als das der ersten. Im umgekehrten Fall sind die Ergebnisse der beiden Teilaufgaben zu subtrahieren. Im Originaltest wird die Anzahl der innerhalb von 30 Minuten richtig gelösten Aufgaben bewertet (Düker & Lienert, 1965). Da beim KLT zwei Aufgaben zur selben Zeit ausgeführt werden müssen (*dual tasking*), nämlich das Rechnen und das Merken der Zwischenergebnisse, steigen die Anforderungen der Aufgabe an die zentrale Exekutive noch zusätzlich. Im Working Memory Modell von Baddeley (1996) wird eine verschiedene Subsysteme

² Die Einheit phon bedeutet, dass das Verkehrsgeräusch subjektiv so laut ist wie ein 1 kHz-Ton mit 70 dB

³ 100 dB – ein so extrem lauter Schall wäre heute nicht mehr erlaubt!

fordernde Doppeltätigkeit von der zentralen Exekutive überwacht, da diese die Aufgabe hat, die Tätigkeit der Hilfssysteme zu kontrollieren und Fehler zu korrigieren.

Die Leistung im Konzentrations-Leistungs-Test wurde bereits unter Hintergrundschall untersucht. Heller (1988) modifizierte die Originalversion am PC (*Rechnen nach Düker, RND*), so dass die Versuchsperson nicht mehr die – im Original vorhandene – größer/ kleiner Grenze von *zehn* zur Vereinfachung der Rechenregel nutzen konnte. Diese Vereinfachung war möglich, da von den beiden Zwischenergebnissen immer eines größer als *zehn* war und eines kleiner. In der Originalversion wurde diese Regel von manchen Probanden erkannt, was dazu führte, dass sie, wenn das erste Zwischenergebnis größer bzw. kleiner als *zehn* war, sofort subtrahierend bzw. addierend weiterrechnen konnten. Somit war es möglich, das Memorieren des Zwischenergebnisses zu umgehen (Heller, 1988). Bei seiner Version fand Heller (1988) anders als beim originalen KLT eine leistungsbeeinträchtigende Wirkung von lautem Hintergrundrauschen (78 dB).

Eine Untersuchung dieses Tests unter Sprachschall liegt in der Literatur bisher noch nicht vor. Hierfür, bzw. für eine vergleichende Testung beider Schallarten würde sich dieser Test allerdings besonders anbieten, da er hohe Anforderungen an beide Subsysteme des Working Memory Modells, die phonologische Schleife und die zentrale Exekutive stellt. Selbst wenn man nicht sicher davon ausgehen kann, dass das Kopfrechnen an sich die phonologische Schleife (siehe gegenteilige Befunde von Thomas et al., in press) fordert, muss man in diesem Test durch das Memorieren der Zwischenergebnisse darauf schließen, dass dieses System erheblich belastet wird. Ebenso stark fordert dieser Test zusätzlich zum Kopfrechnen die zentrale Exekutive durch die Anwendung der Rechenregel und die Überwachung der Doppelaufgabentätigkeit.

Was sich in der Literatur immer wieder findet, sind Hinweise auf Moderatorvariablen von Schalleinflüssen auf Tests dieser Art. Der Persönlichkeitsfaktor *Introversion* versus *Extraversion* wurde in manchen Untersuchungen als wichtige Einflussgröße bei der Leistung unter lauten Hintergrundschallen herausgestellt. Hiermit wird die Voraktivierung des Individuums bezeichnet, die Einfluss auf die Wirkung einer Arousalsteigerung durch Lärm haben kann. Die Konzepte bezeichnen laut Eysenck (1967) ein Kontinuum der Reagibilität auf Außenreize bzw. der Grundaktivierung. Extravertierte besitzen laut dieser Theorie eine niedrigere Grundaktivierung, was dazu führt, dass sie stärkere Außenreize noch als angenehm empfinden können als introvertierte Menschen (Amelang & Ullwer, 1991). Extravertierte werden in Übereinstimmung mit dieser Theorie von lautem Verkehrslärm (88 dB(A)) beim Subtrahieren nicht so sehr gestört wie Introvertierte, die schneller ermüden (Belojevic, Slepcevic, & Jacolvjevic, 2001). Eine Interaktion zwischen Hintergrundlärm und Extraversion in einem Kopfrechentest zeigt sich auch im Trend bei anderen Autoren, wobei die Introvertierten vom Schall erwartungsgemäß mehr beeinträchtigt werden (Furnham & Strbac, 2002). Lärm in Erholungspausen (Rauschen, 85 dB) zeigt ebenfalls in Abhängigkeit von der Voraktivierung eine Wirkung auf anschließendes Kopfrechnen, bei Niedrigaktivierten führt er zu einer Leistungssteigerung, bei Hochaktivierten zu einem Leistungseinbruch (Lohmann, 1985).

Ebenso auf ein Aktivierungskontinuum bezieht sich das Konzept der *reaktiven Anspannungssteigerung*, das im Zusammenhang mit der Leistung im eben beschriebenen KLT untersucht wurde (Beckmann, 1999; Düker, 1963; Rohloff & Gollwitzer, 1999). Es bezeichnet das

Phänomen, dass sich durch zusätzliche Stressoren wie Hintergrundschall, Alkohol, Ermüdung oder Medikamente die Leistung der Probanden anfänglich etwas bessert, bevor es mit steigender Belastung zu einem Leistungseinbruch kommt. Dies erklärt sich theoretisch dadurch, dass für die Aufrechterhaltung des Leistungsniveaus trotz des zusätzlichen Hindernisses die Anspannung soweit gesteigert wird, dass es kurzzeitig sogar zu einer Leistungsverbesserung kommen kann. Erst bei weiter steigenden Anforderungen ist eine Steigerung der Anstrengung nicht mehr möglich und die Leistung verschlechtert sich. Dieses Befundmuster kann allerdings ebenso über das oben vorgestellte Konzept des Arousal (Kapitel 1.4) erklärt werden, da es sich bei diesem Test um komplexe Anforderungen handelt, die durch steigende Selektivität des Aufmerksamkeitsfokus initial besser gelöst werden können. Bei weiter steigender Selektivität und erhöhter Ablenkbarkeit der Aufmerksamkeit durch steigende Erregung können allerdings nicht mehr alle Informationen integriert werden. Leistungseinbußen wären die Folge davon (Easterbrook, 1959; Kahneman, 1973).

Allgemein lässt sich aus diesen Forschungsergebnissen zu Moderatorvariablen der Schluss ziehen, dass die Voraktivierung der Probanden beachtet werden muss, wenn mit Hilfe des KLT leistungsbeeinträchtigende Wirkungen von Hintergrundlärm untersucht werden sollen, die über das Arousalniveau vermittelt werden. Diese Aktivierung, die beispielsweise anhand des Konzepts der Extraversion oder der Leistungsmotivation (und dadurch einer reaktiven Anspannungssteigerung) erfasst werden kann, scheint einen Einfluss auf die Wirkung einer hinzukommenden Arousalsteigerung durch Lärm zu haben. Ähnliches wurde bereits im Zusammenhang mit der Beeinflussung des Belästigungserlebens durch personenbezogene Faktoren in Kapitel 1.1. angesprochen.

Zusammenfassend kann man sagen, dass Kopfrechentests sowohl zentral exekutive Anforderungen stellen als auch die phonologische Schleife fordern können. Ansprüche an letztgenannte Struktur scheinen Kopfrechentests vor allem dann zu stellen, wenn die Ziffern nur kurz präsentiert werden, wenn Zählen zur Lösung der Aufgabe genutzt werden kann oder Zwischenergebnisse memoriert werden müssen. In diesen Fällen lässt sich eine verstärkte Interferenz mit Aufgaben feststellen, die ebenfalls die phonologische Schleife benötigen, wie artikulatorischer Unterdrückung (DeStefano & LeFevre, 2004). Besonders hohe Anforderungen an die phonologische Schleife und die zentrale Exekutive stellt durch die geforderte Doppelaufgabenbearbeitung und das Memorieren von Zwischenergebnissen der so genannte KLT (Düker & Lienert, 1965).

Eine Aufgabe, die ebenso wie der Konzentrations-Leistungs-Test Anforderungen an die zentrale Exekutive stellt (Baddeley 1968), ist das *grammatikalische Schlussfolgern*. Zusätzlich wird auch bei dieser Aufgabe eine Beteiligung der phonologischen Schleife vermutet. Empirische Ergebnisse, die hierzu in der Literatur vorliegen, sollen im nächsten Kapitel vorgestellt werden.

1.6 Das Grammatical Reasoning und seine Verwendung in der Lärmwirkungsforschung

Das Grammatical Reasoning besteht in der Originalform (Baddeley, 1968) aus 32 Aufgaben. Es wird vom Autor als ein Test vorgestellt, mit dem die Auswirkungen von Stressoren auf *höhere mentale Prozesse* getestet werden können. Die Teilaufgaben ergeben sich aus der Kombination der Verben *precede* und *follow*, deren Aktiv- und Passivform, der vorhandenen oder nicht vorhandenen Verneinung, der Antworten *AB* oder *BA* und der Übereinstimmung des Satzes (z.B. *A does not precede B*) mit der Antwortzeile (z.B. *AB*). Die Versuchsperson wird aufgefordert, die Antwortzeile anhand des Satzes zu verifizieren und mit Zustimmung oder Ablehnung zu reagieren. Der Autor findet in eigenen Untersuchungen bei einem nur geringen Übungsfortschritt eine hohe Reliabilität der Leistung in diesem Test. Diese Aufgabe scheint durch erhöhte Anstrengung besser lösbar zu sein. Aus Fragebögen zum aktuellen Arousalniveau und der aufgewendeten Anstrengung wurde geschlossen, dass Probanden deshalb ohne erkennbare Leistungseinbußen gleichzeitig Auto fahren können, weil sie ihre Anstrengung entsprechend den Anforderungen erhöhen (Matthews, Sparkes, & Bygrave, 1996). Um eine Belastungsgrenze bei Probanden zu testen, kann bei diesem Test folglich die Schwierigkeit noch gesteigert werden.

Dieser Test wurde in leicht veränderter Form in die so genannte *AGARD STRES Battery* (**A**dvisory **G**roup for **A**erospace **R**esearch **D**evelopment, **S**tandardized **T**ests for **R**esearch with **E**nvironmental **S**tressors) aufgenommen. Diese wurde entwickelt, um ein standardisiertes Verfahren zur Erfassung der Auswirkungen von Stressoren auf geistige Leistungen vorzulegen. Die Batterie enthält verschiedene Tests, die sowohl gemeinsam als auch einzeln getestet werden können, was die Handhabung sehr flexibel gestaltet (Advisory Group for Aerospace Research and Development, 1989). In der Testbatterie wurde das Grammatical Reasoning aus Gründen der Sprachunabhängigkeit modifiziert, was bedeutet, dass die Verben *follow* und *precede* durch *before* und *after* ersetzt wurden. Die daraus folgende Vereinfachung wurde durch eine Erhöhung der Zahl an Statements pro Aufgabe ausgeglichen. Die hieraus resultierende Aufgabe lautet folgendermaßen: Betrachtet werden sollen zwei Aussagesätze, in denen die Reihenfolge von je zwei Symbolen mit den Worten *before* und *after* angegeben wird (z.B. * before &; & after #). Diese Sätze sollen mit der dritten Zeile (z.B. * & #) verglichen werden. Ist die Reihenfolge der Symbole in beiden Sätzen korrekt oder in beiden falsch angegeben, so ist die Antworttaste die Taste mit der Aufschrift *Identisch*. Ist einer der beiden Sätze verglichen mit der Symbolzeile korrekt und der andere falsch, so ist mit der Taste *Verschieden* zu antworten.

Um eine optische Vereinfachung der Lösung zu verhindern, gilt eine Reihenfolge auch dann als korrekt, wenn die fraglichen Symbole durch das dritte getrennt sind (Salamé, 1993). Aufgrund der Bearbeitungsdauer der Aufgaben kann man bereits generalisierende Aussagen zur aufgetretenen Belastung treffen. Dies lässt sich an einer multivariaten Analyse der bedingungsbezogenen Messgenauigkeit von acht Tests dieser Testbatterie erkennen (Schütte & Nickel, 2002). Hierbei darf man jedoch nicht zu dem Schluss kommen, dass man über diesen Indikator die psychisch-mentale Anstrengung vollständig erfassen könnte. Die Testbatterie erweist sich als besonders geeignet, ihr Ziel – Leistung unter Stress zu messen – zu erreichen. Die AGARD STRES Batterie wurde unter Umgebungsbedingungen wie simulierten Hochgebirgsbedingungen (Kennedy, Dunlap, Banderet,

Smith, & et al., 1989), erhöhtem Kohlenstoffdioxidgehalt der Luft (Manzey & Lorenz, 1998) oder bei Aufenthalt im Space-Lab (Manzey, Lorenz, & Poljakov, 1998) eingesetzt, um die Auswirkungen der hier jeweils wirksamen Stressoren auf die kognitive Leistungsfähigkeit einzuschätzen.

Die Prozesse, die der Lösung von Schlussfolgerungen je nach Formulierung des Problems zugrunde liegen, wurden anhand von Experimenten zu Interferenzen zwischen anschaulichem Denken und visuellem sowie räumlichem Vorstellungsvermögen untersucht (Knauff & Strube, 2002). Hierbei zeigte sich, dass Aufgaben des Typs: *Das rote Rechteck liegt links vom grünen Dreieck, das grüne Dreieck links vom blauen Kreis, liegt der blaue Kreis rechts vom roten Rechteck?* zwar mit räumlichen Arbeitsgedächtnisaufgaben (Positionsvergleich von Tönen oder Objekten auf dem Bildschirm) interferieren, nicht aber mit rein visuellen Aufgaben (visueller Helligkeitsvergleich von Objekten am Bildschirm). Logische Schlussfolgerungen dieses Typs scheinen eher auf räumlicher Vorstellung zu beruhen als auf rein visueller. Syllogistische Schlussfolgerungen (*Alle Hunde sind Säugetiere. Alle Schäferhunde sind Hunde. Sind alle Schäferhunde Säugetiere?*) stellen dagegen vor allem Anforderungen an die zentrale Exekutive (Gilhooly, Logie, Wetherick, & Wynn, 1993) und interferieren nicht mit artikulatorischer Unterdrückung oder visuell-räumlichem Tapping, jedoch mit der Produktion zufälliger Reihenfolgen von Zahlen oder Buchstaben (*Random Generation*).

Eine nachweisbare Interferenz mit Aufgaben, welche die phonologische Schleife beanspruchen, findet sich, wenn der oben beschriebene grammatikalische Schlussfolgerungstest in der Version von Baddeley (1968) während der Bearbeitung von anderen Aufgaben getestet wird (Baddeley, 1976; Hitch & Baddeley, 1976). Vor allem beim gleichzeitigen Memorieren von Zahlenreihen, in geringerem Ausmaß bei reiner artikulatorischer Unterdrückung (Wiederholung eines Wortes) oder beim Memorieren einzelner Zahlen ergeben sich Leistungseinbußen in der Hauptaufgabe. Diese Interferenzen sind abhängig von der Schwierigkeit der Aufgabe und treten am stärksten bei der schwersten Bedingung auf, d.h. bei passiven, verneinten Sätzen, die falsch sind.

In einem Vergleich von propositionalem (z. B. entweder-oder, weder-noch) und räumlichem Schlussfolgern (rechts von, links von) unter diversen Zweitaufgaben (*Random Generation, Artikulation, räumliches Tapping*) konnten andere Autoren wiederum zeigen, dass beide Arten des Schlussfolgerns von einer zentral exekutiven (*Random Generation*) und einer die phonologische Schleife beanspruchenden Zweitaufgabe (*Artikulation*) gestört werden (Klauer, Stegmaier, & Meiser, 1997). Eine räumlich-visuelle Aufgabe hingegen erschwerte nur das räumliche Schlussfolgern.

Hieraus kann der Schluss gezogen werden, dass bei Schlussfolgerungsprozessen im Allgemeinen vor allem zentral exekutive Strukturen sowie solche der phonologischen Schleife (z.B. beim Grammatical Reasoning von Baddeley) benötigt werden. Der visuell räumliche Notizblock scheint nur bei eindeutig räumlich-visuellem Schlussfolgern beteiligt zu sein.

Beide oben beschriebenen Versionen des Grammatical Reasoning – die der AGARD STRES Battery und die von Baddeley – wurden in der Literatur verwendet, um die Wirkung von irrelevanten Schallen auf Schlussfolgerungsprozesse zu testen. Das Grammatical Reasoning von Baddeley (1968) wurde von Kjellberg und Skoeldstroem (1991) unter irrelevanter Hintergrundsprache (65 dB(A)) sowie unter einem Breitbandrauschen in zwei Lautstärken (44 vs. 65 dB(A)) verwendet. Hierbei waren zwar

keine Leistungsunterschiede zu finden, jedoch ein höheres Lästigkeitsurteil über die beiden lautereren Schalle bei dieser Aufgabe im Vergleich zu einfachen Reaktionszeitaufgaben und einem Farb-Wort-Test. Die Urteile über den Sprachschall und das laute Rauschen unterschieden sich für das Grammatical Reasoning nicht.

In einer anderen Untersuchung ließ eine Autorengruppe ihre Probanden die Lautstärke eines irrelevanten Hintergrundgeräusches je nach Belästigungserleben selbst auf ein gerade noch erträgliches Maß einstellen (Landstroem, Kjellberg, & Bystroem, 1995). Dabei zeigte sich, dass ein 1000 Hz-Ton grundsätzlich leiser gestellt wurde als ein Breitbandrauschen. Seine Lautstärke wurde von den Versuchspersonen darüber hinaus noch weiter reduziert, wenn das Grammatical Reasoning anstatt eines einfachen Reaktionszeittests zu absolvieren war. Dies war auch dann der Fall, wenn das Geräusch diskontinuierlich war. Hieraus kann vermutet werden, dass der irrelevante Hintergrundschall die Probanden nicht nur mehr störte, wenn die Aufgabe anspruchsvoller wurde, sondern auch, wenn sich die Frequenzzusammensetzung oder die zeitliche Struktur des Schalls änderte.

In einem räumlich-visuell anspruchsvollen Schlussfolgerungstest (*Ist der Kreis rechts vom Dreieck größer als das Viereck links vom Kreis?*) konnten Dornic und Laaksonen (1988) die Schallwirkung eines lauten weißen Rauschens, dessen Lautstärke leider nicht näher benannt wird, in Form einer Verschlechterung der Leistung in der zweiten Hälfte des Tests feststellen. Die Autoren erklären dieses Ergebnis mit der Auswirkung des Schalls auf das Arousalssystem (Dornic & Laaksonen, 1988).

Die Befundlage in der Literatur zur Leistung im Grammatical Reasoning unter Störschallen ist leider eher gering. Diese Aufgabe scheint sensitiv auf Sprachschalle sowie laute Schalle zu sein, was sich allerdings meist nur im Lästigkeitsurteil über den Schall zeigt. Das Grammatical Reasoning scheint in den vorgestellten Versionen von den Probanden selbst unter zusätzlicher Belastung noch durch Erhöhung der Anstrengung lösbar zu sein (vgl. Matthews et al., 1996). Somit mag diese eher dürftige Befundlage an der zu einfachen Konzeption der Aufgabe liegen. Eine Leistungsver schlechterung unter lautem Schall findet sich nämlich wie berichtet in einem räumlich-visuellen Schlussfolgerungstest, was darauf hinweist, dass Schlussfolgern an sich sensitiv auf Schallwirkungen wäre.

Zusammenfassend kann das Grammatical Reasoning als eine gut untersuchte Aufgabe gelten, die vor allem Anforderungen an die zentrale Exekutive und die phonologische Schleife stellt.

Eine Aufgabe, die speziell zentral exekutive Anforderungen stellt ist der *Stroop Test*. Auf diesen Test, seine Ansprüche an das Arbeitsgedächtnis und Untersuchungsergebnisse zu Schallwirkungen soll im Folgenden eingegangen werden.

1.7 Der Stroop Test – Theorie und Befunde

In der ursprünglichen Version dieses Tests, die von J. R. Stroop 1935 publiziert wurde, besteht die Aufgabe der Versuchsperson darin, farbige Tafeln zu bearbeiten. Hiervon gibt es vier verschiedene: eine auf der Farbworte vorzulesen sind, eine bei der Farben zu benennen sind, und zwei auf denen farbige Farbworte abgebildet sind, wobei die Versuchsperson einmal die Farbe und das andere Mal das Wort berichten soll. Das als *Stroop Effekt* bezeichnete Phänomen zeigt sich in einer längeren Reaktionszeit für das Benennen der Farben der Worte als für das Aussprechen der farbigen Worte. Später wurde allgemein eher zu einer Einzelpräsentation der Worte übergegangen, was sich als valider als die Listenpräsentation erwiesen hat (Siegrist, 1995). Meist wird nur noch mit der Anweisung gearbeitet, die Farbe von Farbworten zu benennen, die mit der Wortbedeutung entweder übereinstimmt oder nicht. Hierbei zeigt sich der Stroop Effekt als Verlängerung der Antwortzeit für die *inkongruenten* (Farbe nicht identisch mit der Wortbedeutung) im Vergleich zu den *kongruenten* Worten (Farbe identisch mit der Wortbedeutung). Als Antwort war im ursprünglichen Test (Stroop, 1935) das Aussprechen vorgesehen. Wie sich in späteren Untersuchungen, die die Antwortmodalitäten variierten, zeigte, trug dies erheblich zur Verlängerung der Reaktionszeiten bei (Pritchatt, 1968). Bei einer Antwort über Tastendruck mit farblich markierten Tasten reduziert sich der Stroop Effekt (Verlängerung der Reaktionszeit) erheblich. Sobald die Tasten aber mit Farbworten kodiert werden, steigt der Stroop Effekt wieder auf dasselbe Niveau an wie bei einer verbalen Antwort (Glaser & Glaser, 1993). Außerdem wurde von Pritchatt (1968) anhand von Assoziationen zwischen sinnlosen Silben und Farben bzw. zwischen Farben und sinnlosen Silben der Versuch unternommen, den Stroop Effekt unabhängig von bekannten Worten nachzuweisen. Je nach Richtung der gelernten Assoziation kam es im anschließenden entsprechend modifizierten Stroop Test zu einer unterschiedlichen Reaktionszeitbeeinflussung. Je nachdem, welche Antwort direkt mit dem Stimulus verbunden ist, kommt es folglich zu einem Stroop Effekt. Sogar farbverbundene Worte (Gras – grün) können die Benennung der Farbe, in der sie gedruckt sind, verlangsamen (Klein, 1964).

Bezüglich einer theoretischen Erklärung des Stroop Effektes sprach Stroop selbst (Stroop, 1935) in seiner ersten Veröffentlichung von der später so genannten *Automatismushypothese*. Diese nimmt an, dass die Verlängerung der Reaktionszeiten beim Benennen der Farben dadurch hervorgerufen wird, dass das Lesen der Worte automatisiert abläuft und inhibiert werden muss, um die davon abweichende Farbe zu nennen. Im Rahmen des Arbeitsgedächtnismodells von Baddeley gesehen, scheint dieser Test somit zentral exekutive Ansprüche zu stellen, da die Inhibition von automatischen Reaktionen in diesem Modell explizit als Aufgabe dieses Systems genannt wird (vgl. z.B. Random Generation Aufgabe). Allerdings ist die Automatismushypothese nicht unumstritten. In Untersuchungen konnte Monahan (2001) zeigen, dass die Reduktion des Stroop Effekts, die sich bei Kolorierung eines einzelnen Buchstaben anstelle des ganzen Wortes ergibt, nicht auf eine Reduktion des automatischen Lesens zurückzuführen ist. Stattdessen kann man auf eine verlangsamte Farbverarbeitung schließen, da sich eine sogar noch langsamere Beantwortung bei Kolorierung einzelner Teile bei kongruenten Worten, geometrischen Figuren und Balken zeigen lässt (Monahan, 2001).

Alternativ wird häufig die Hypothese der *Antwortkonkurrenz* genannt (Klein, 1964; MacLeod, 1991). Ihr zufolge entsteht die längere Bearbeitungszeit bei inkongruenten Worten im Stroop Test dadurch, dass zwei alternative Antworten um Auswahl konkurrieren – das Wort der Farbe und das Wort an sich, die beide auf phonologischer Ebene präsent sind. Als Beleg hierfür zieht der Autor eigene Ergebnisse heran (siehe oben), bei denen er eine Interferenz auch zwischen farbverbundenen Worten und ihrer Farbe nachweisen konnte (*Gras* in blau, *Himmel* in grün). In einem Stroop Test mit den Worten *Yes* und *No* und *Right* und *Wrong* findet sich ein Interferenz-Effekt bei der Feststellung, ob die beiden Worte in derselben (Antwort *Yes*) oder in verschiedener Farbe (Antwort *No*) geschrieben waren (Dalrymple-Alford & Azkoul, 1972). Die Autoren schließen hieraus ebenfalls, dass die Interferenz nicht auf perzeptueller Stufe, sondern auf der Stufe der Antwortauswahl bzw. Antwortwettstreites stattfindet. Da die Kontrolle der Subsysteme und die Integration von Informationen aus dem Langzeitgedächtnis im Working Memory Modell eine Aufgabe der zentralen Exekutive ist, kann man auch aus dieser Erklärung des Phänomens auf eine Beteiligung dieser Instanz beim Stroop Test schließen.

Allerdings scheint die Interferenz nicht erst bei der Antwortauswahl, sondern bereits im *semantischen System* aufzutreten, wie spätere Untersuchungen ergaben (Luo, 1999). Durch Verlängerung der Intervalle zwischen zu vergleichenden Stimuli gelang es Luo, die Interferenz beim Vergleich der Farben eines Balkens und eines Farbwortes (visuelle Entscheidung) zu erhöhen und die Interferenz beim Vergleich der Farbe des Balkens mit der Bedeutung des Farbwortes (Entscheidung über Bedeutung) abzuschwächen. Da laut Forschungsergebnissen die Verarbeitung der Bedeutung eines Wortes schneller ist als die Verarbeitung seiner Farbe, ist der Nachweis einer Interferenz zwischen der Bedeutung des Wortes und seiner Farbe (Ist die Bedeutung des Wortes identisch mit der Farbe des Balkens?) im Vergleich mit der Farbe eines Balkens (Antwort: identisch oder nicht) überraschend. Dies kann nur dann der Fall sein, wenn die Verarbeitung der Oberflächenfarben des Wortes und des Balkens im semantischen System zeitgleich abläuft und es somit zu einer Interferenz zwischen den Farben bereits hier und nicht erst bei der Antwortselektion kommt. In dieselbe Richtung weisen auch Ergebnisse von Untersuchungen zum Stimulusmaterial, in denen sich zeigte, dass farbkongruente Worte (d.h. Worte, deren Bedeutung als Farbe im Test auch vorkommt) zu längeren Reaktionszeiten führen als farbdistante Worte (Worte, deren entsprechende Farbe im Test nicht vorkommt) (Effler, 1978). Da allerdings auch der Zugriff auf das semantische System im Working Memory Modell von der zentralen Exekutive geleistet wird, bezieht auch diese Erklärung die Funktion dieser Instanz mit ein.

Eine vierte Erklärung des Phänomens stammt von Harvey (1984). Diese entwickelte er anhand eines Stroop Tests, bei dem Hinweisreize anzeigen, auf welches Merkmal beim folgenden Item geachtet werden muss. Seine Ergebnisse legen nahe, dass die klassische Stroop-Interferenz auch darauf zurückzuführen sein könnte, dass die Versuchspersonen die Anweisung nicht den ganzen Versuch über befolgen können. Bei kongruenten Items kann nämlich auch ebenso auf das Wort geantwortet werden. Weniger wahrscheinlich ist nach seinen Daten die Erklärung, dass die Probanden ihre Aufmerksamkeit nicht auf das gesuchte Merkmal (die Farbe) fixieren können (Harvey, 1984). Bereits die Anwesenheit inkongruenter Worte hat einen verlangsamenden Einfluss auf die Verarbeitung von kongruenten Worten. Die Probanden bereiten sich auf inkongruente Worte vor,

unterdrücken aber erst die Aussprache des Wortes, wenn tatsächlich ein inkongruentes Wort erscheint. Dies führt dann zum Stroop Effekt (Monahan, 2001). Schwerer wird die Aufgabe noch, wenn der Prozentsatz der inkongruenten Items gesenkt wird, da die Anweisung auf die Farbe statt auf das Wort zu antworten, über lange Zeiten im Experiment nicht benötigt wird, aber im Arbeitsgedächtnis des Probanden aktiv gehalten werden muss (Engle, 2002). In einer Untersuchung der Leistung von Probanden bei der Präsentation von Stroop Tests, bei denen bestimmte Farben nur inkongruent und andere nur kongruent auftraten, ließen sich itemspezifische Strategiewechsel finden (Jacoby, Lindsay, & Hessels, 2003). Hieraus lässt sich ebenfalls ableiten, dass die Anweisung *auf die Farbe zu antworten*, nur dann befolgt wird, wenn es notwendig ist. Durch eine Änderung der Anzahl inkongruenter Worte kann somit die Interferenz beeinflusst werden.

Wichtig in diesem Zusammenhang ist, dass alle vier Hypothesen eine Beteiligung der zentralen Exekutive nach Baddeley (2000a) implizieren. Die erste Erklärung des automatischen Lesens bezieht sich mit ihrer Annahme der Inhibition von automatisierten Reaktionen im Working Memory Modell auf diese zentrale Kontrollinstanz. Auch die beiden Theorien, die als Erklärung auf eine Konkurrenz der beiden möglichen Antworten im Stadium der Antwortselektion bzw. bereits im semantischen System rekurren, kann man der zentralen Exekutive zuschreiben, da diese den Abruf aus dem semantischen System kontrolliert (Baddeley & Logie, 1999) und bei ihr darüber hinaus Kontroll- und Entscheidungsaufgaben lokalisiert werden. Die vierte Erklärung des Stroop Phänomens bezieht sich auf die exekutive Aufmerksamkeit, die nötig sei, um auf das richtige Merkmal der Reize zu fokussieren. Laut Baddeley sei diese ebenfalls von der zentralen Kontrollinstanz gewährleistet (Baddeley & Logie, 1999). Es herrscht folglich Einigkeit über die Beteiligung der Instanz der zentralen Exekutive bei Variationen des Stroop Test.

Auch beim Stroop Test werden im Folgenden, wie bei Rechentests und dem Grammatical Reasoning, in der Literatur vorhandene Ergebnisse zu einer Bearbeitung unter Hintergrundschall vorgestellt.

Die einzigen Durchführungen des Stroop Tests mit irrelevanter Hintergrundsprache sind Untersuchungen, in denen auditive Farbwoorte bei visueller Farbbenennung dargeboten wurden. Hierbei wurde von einigen Autoren eine Interferenz festgestellt (Cowan, 1989; Cowan & Barron, 1987), von anderen hingegen nicht (Miles & Jones, 1989; Miles, Madden, & Jones, 1989). Die Darbietung von Buchstaben oder Musik erzeugte bei denselben Autoren keine Interferenz (Cowan & Barron, 1987). In weiterführenden Experimenten explizierten die Autoren (Elliott, Cowan, & Valle-Inclan, 1998) ihre Schlussfolgerungen dahingehend, dass es beim Interferenzeffekt zwischen gehörten Farbworten und dem Benennen einer visuell präsentierten Farbe auf die zeitliche Darbietung beider ankommt. Nur bei gleichzeitiger Präsentation kommt es zu einem Stroop Effekt. Bei einem um 500 ms früheren Einsetzen des auditiv präsentierten Wortes dagegen zeigt sich keine Interferenz. Dies erklären die Autoren damit, dass im Gegensatz zum ersten Fall, das Wort im Zweiten bereits verarbeitet ist, bevor der Stimulus dargeboten wird. Weitergehende Analysen desselben Autorenteam (Elliott & Cowan, 2001) konnten verdeutlichen, dass Habituation auf die eben beschriebene Interferenzanordnung erleichternd wirken kann. Das geschieht allerdings nur, wenn die

Störitems bei der Habituation ignoriert werden, also dieselbe Aufgabenrelevanz besitzen wie im anschließenden Versuch. Aufgrund der Interferenz der Items, die nur bei zeitgleicher Darbietung auftritt, und einer möglichen Habituation an den Störschall, kann hier allerdings nicht auf eine Wirkung des spektral-temporal stark variablen Schalls auf die phonologische Schleife geschlossen werden (vgl. Kapitel 1.3).

Zur leistungsbeeinflussenden Wirkung eines lauten irrelevanten Hintergrundschalls, der über eine Steigerung des Arousal die Arbeit der zentralen Exekutive und damit den Stroop Test beeinträchtigen könnte, liegen in der Literatur widersprüchliche Ergebnisse vor.

Schon 1985 finden Smith und Broadbent einen verlangsamenden Einfluss eines Rauschens (85 dB(C)) im Vergleich zu Ruhe (55 dB(C)). Dieser liegt jedoch nur auf die Lesebedingung im klassischen Stroop Test vor, nicht auf die Interferenzbedingung, d.h. das Benennen der Farbe von farbigen Farbwörtern. Dieser Effekt tritt auch nur dann ein, wenn die Probanden dem Lärm bereits vor der Aufgabenbearbeitung ausgesetzt waren (Broadbent, 1981; Smith & Broadbent, 1985).

Eine Erleichterung der Aufgabe durch Schall hingegen wird von Houston und Jones (1967) berichtet. Ihren Ergebnissen zufolge kann es unter lautem Hintergrundschall zu einer schnelleren Bearbeitung in der Interferenzbedingung des Stroop Tests in der originalen Paper-Pencil Version und zu einer Verlangsamung in den anderen – einfacheren – Bedingungen (dem Benennen von farbigen Sternen) kommen. Man kann hieraus einen erleichternden Einfluss des Schalls auf die Interferenzneigung (im Sinne einer nicht stattfindenden Verlangsamung) ableiten (Houston & Jones, 1967). Leider wird von den Autoren der Hintergrundschall nicht genau beschrieben. Dieser sei eine Zusammenstellung bekannter Geräusche (natürliche und künstliche wie Zugfahrten, tropfendes Wasser und elektronische Musik) gewesen, deren Lautstärke allerdings nicht berichtet wird. Die Probanden seien instruiert worden, diesen Schall zu ignorieren. In einer weiteren Untersuchung kommt Houston zu dem Schluss, dass diese Erleichterung der Inhibition durch das Ignorieren des Schalls (hier wird er mit einer Lautstärke von 78 dB berichtet) nicht von der Schwierigkeit der untersuchten Aufgaben beeinflusst ist (Houston, 1969). Der Autor interpretiert sein Befundmuster dahingehend, dass die Inhibition durch das Ignorieren des Schalls erleichtert werde, da die Probanden bereits etwas inhibierten. Der Befund sei nicht über Arousal erklärbar, da die Probanden keine erhöhten Pulsschlagraten aufwiesen. Dies hätte auf einen Zusammenhang zum Arousalniveau hingedeutet.

Dornic (1982) testete die Auswirkungen einer ebenfalls leistungserleichternden sinkenden Interferenzneigung durch steigende Selektivität der Aufmerksamkeit in Abhängigkeit von der Interferenzanfälligkeit der Probanden. Diese variierte er anhand der Vertrautheit mit einer Zweitsprache, in der die Farbwörter geschrieben waren. Je vertrauter den Probanden die Sprache war, desto eher neigten sie zu einer Interferenz und desto mehr profitierten sie in seinen Experimenten von Hintergrundlärm (weißes Rauschen mit 85 dB versus Ruhe von 55 dB). Der Autor lässt in seinen Schlussfolgerungen offen, ob dieser Effekt eher durch eine steigende Selektivität der Aufmerksamkeit oder eine Beeinträchtigung der semantischen Verarbeitung zu erklären sei (Dornic, 1982).

Allerdings scheint die Dauer der Lärmexposition ein wichtiges Kriterium zu sein. Andere Autoren fanden unter Hintergrundschall einen kumulativen leistungsbeeinträchtigenden Effekt des

Lärms auf die Interferenzneigung der Probanden (Hartley & Adams, 1974). Verglichen wurden von den Autoren die Effekte eines Breitbandrauschens mit 70 dB und einem solchen mit 95 dB. Nach einer kurzen Darbietung des lauten Schalls wurde die Leistung initial besser, was sich aber nach einer längeren Exposition in erhöhte Interferenzneigung und damit längere Bearbeitungszeiten wandelte. Dies wird von den Autoren über eine Erhöhung des Arousal erklärt.

Im Sinne der oben dargelegten Theorien zum Konzept des Arousal könnte man diese Ergebnisse folgendermaßen einordnen: Die erhöhte Selektivität des Aufmerksamkeitsfokus durch gesteigertes Arousal erleichtert die Aufgabe sich nur auf das Merkmal der Farbe zu konzentrieren, was die Interferenz im semantischen System (Luo, 1999) zu reduzieren hilft und initial zu einer verbesserten Leistung führen könnte (Easterbrook, 1959). Wenn allerdings bei weiter steigendem Arousal zusätzlich eine verstärkte Ablenkbarkeit der Aufmerksamkeit hinzukommt, kann die Aufmerksamkeit nicht mehr auf dem Merkmal *Farbe* gehalten werden. Die Verarbeitung des Wortes im semantischen System rückt damit wieder mehr in den Vordergrund, wodurch eine stärkere Interferenz auftreten kann (Kahneman, 1973). Interessant wäre, ab welcher Lautstärke oder Darbietungsdauer es zu einer Verschlechterung statt zu einer Verbesserung kommen kann und durch welche anderen Geräuscheigenschaften die Leistung beeinflusst werden kann. Diese Fragen können aus den in der Literatur vorliegenden empirischen Ergebnissen nicht beantwortet werden.

Nachfolgend sollen Fragestellung und Hypothesen der vorliegenden Arbeit aufbauend auf den dargelegten theoretischen Hintergründen vorgestellt werden.

2 Fragestellung und Hypothesen

Im Folgenden wird eine Serie von Experimenten vorgestellt, die der Beantwortung folgender Fragen dient:

1) Wie wirken sich Straßenverkehrsgeräusche im Vergleich zu Sprachschall auf die Leistung in den folgenden Tests aus?

- Serial Recall
- KLTM (modifizierter KLT)
- Stroop Test

Unter Bezug auf die oben dargelegten theoretischen Zusammenhänge ergeben sich folgende Hypothesen:

- Im Serial Recall wird eine Leistungsbeeinträchtigung unter Sprachschall im Vergleich zu einer Ruhebedingung erwartet, nicht jedoch unter lautem Verkehrslärm.
- Im Stroop Test wird eine Leistungsbeeinträchtigung unter lautem Verkehrslärm im Vergleich zu einer Ruhebedingung erwartet, nicht jedoch unter Sprachschall.
- Im KLTM wird erwartet, dass sowohl durch lauten Verkehrslärm als auch durch Sprachschall eine Leistungsbeeinträchtigung im Vergleich zu einer Ruhebedingung hervorgerufen wird.

2) Kann durch eine Pegelreduktion von Verkehrsschallen im tiefen Frequenzbereich ihre leistungsbeeinträchtigende Wirkung verkleinert werden?

Nach den theoretischen Grundlagen, die in Kapitel 1.4 dargelegt wurden, lässt sich folgende Hypothese aufstellen:

- Eine Pegelreduktion im tiefen Frequenzbereich führt im Stroop Test, beim KLTM und dem Grammatical Reasoning durch sinkende Lautheit *und* weniger stark vertretene tiefe Frequenzen zu einer geringeren Leistungseinbuße als der unmodifizierte Schall.

3) Findet sich der Effekt des Sprachschalls auf den KLTM und das Grammatical Reasoning unabhängig von seiner Semantik?

Aus den oben dargelegten Zusammenhängen wird folgende Hypothese generiert:

- Beim KLTM und dem Grammatical Reasoning führt ein semantisch unverständlicher Sprachschall zu einer Leistungsbeeinträchtigung gegenüber der Ruhebedingung, die in einer vergleichbaren Größenordnung wie die durch deutschsprachigen Hintergrundschall liegt.

Im empirischen Teil werden die Experimente nach folgendem Schema vorgestellt (Tabelle 1: bei jedem Experiment lag als Vergleich eine Ruhebedingung vor).

Tabelle 1: Überblick über die Experimente

Fragestellung	Test	Sprachschall $L_{eq}=60$ dB(A)	Verkehrsschalle c/h = PKW/Stunde; oder Schienenverkehr (diverse Züge)	Absenkung Straße <500 Hz um 12 dB; 2000 c/h; $L_{eq}=70$ dB(A) Schiene: <220 Hz um 12 dB; $L_{eq}=70$ dB(A)
Fragestellung I Schallwirkung direkte Interferenz vs. über Arousal	<i>Serial Recall</i>	Verständlich	100 c/h- 60 dB(A) 2000 c/h- 50 dB(A) 2000 c/h- 60 dB(A) 2000 c/h- 70 dB(A)	
	<i>KLTM</i>	Verständlich	100 c/h- 60 dB(A) 2000 c/h- 50 dB(A) 2000 c/h- 60 dB(A) 2000 c/h- 70 dB(A)	
	<i>Stroop Test</i>	Verständlich	100 c/h- 60 dB(A) 2000 c/h- 50 dB(A) 2000 c/h- 60 dB(A) 2000 c/h- 70 dB(A)	
Fragestellung II Pegelreduktion der tiefen Frequenzen	<i>KLTM</i>	Verständlich	2000 c/h- 70 dB(A)	- in Lautheit und tieffrequent reduziert - nur tieffrequent red. - nur in Lautheit reduziert
	<i>KLTM (Subwoofer)</i>	Verständlich	2000 c/h- 70 dB(A) (Lautsprecher mit Subwoofer)	in Lautheit und tieffrequent reduziert
	<i>Stroop Test</i>	Verständlich	2000 c/h- 70 dB(A)	- in Lautheit und tieffrequent reduziert - nur tieffrequent red. - nur in Lautheit reduziert
	<i>Grammatical Reasoning</i>	Verständlich	2000 c/h- 70 dB(A)	in Lautheit und tieffrequent reduziert
Fragestellung III Relevanz der Semantik	<i>KLTM</i>	Unverständlich (Japanisch)	Schiene 70 dB(A) (Lautsprecher und Subwoofer)	Schiene-12 (in Lautheit und tieffrequent reduziert)
	<i>Grammatical Reasoning</i>	Unverständlich (Japanisch)	Schiene 70 dB(A)	Schiene-12 (in Lautheit und tieffrequent reduziert)

3 Versuchsdurchführung, -auswertung und Diskussion der Ergebnisse

3.1 Fragestellung I: Vergleich der Wirkung von Sprachschall und Verkehrsschallen auf Aufgaben, welche die phonologische Schleife oder die zentrale Exekutive fordern

Wie eben skizziert werden nun die Ergebnisse des Serial Recall, des KLTM und des Stroop Tests unter Beschallung vorgestellt. Die verwendeten Schalle⁴ setzen sich zusammen aus vier Verkehrslärmaufnahmen, einem Sprachschall ($L_{eq}=60$ dB(A)) und einer Ruhebedingung. Die Verkehrslärmaufnahmen bestehen aus einer Aufzeichnung von 100 Pkw-Vorbeifahrten pro Stunde (dadurch eher variable Zeitstruktur; $L_{eq}=60$ dB(A)) und drei mit 2000 Pkw-Vorbeifahrten pro Stunde (gleichmäßiger Zeitverlauf; L_{eq} ansteigend: 50, 60 und 70 dB(A)).

3.1.1 Explizierung der Hypothesen

Es wird erwartet, dass es durch die spektral-temporal stark variablen Schalle (Sprachschall und Verkehrslärm mit 100 Pkw-Vorbeifahrten pro Stunde) zu einer Erhöhung der Fehlerraten im KLTM und dem Serial Recall kommt. Da laut der vorgestellten modelltheoretischen Fundierung in der phonologischen Schleife eine Interferenz der memorierten Items mit dem Störschall vorliegt, sollte kein Speed-Accuracy-Trade-Off auftreten. Wenn die memorierten Inhalte durch Interferenz *zerstört* werden, ist es nicht sinnvoll, mehr Zeit für die Beantwortung der Aufgabe aufzuwenden.

Beim Stroop Test ist mit keiner Beeinträchtigung durch zeitlich stark variable Schalle in einer mittleren Lautstärke (60 dB(A)) zu rechnen. Erwartet wird, dass der Verkehrslärm mit einem Mittelungspegel von 70 dB(A) im Vergleich zur Ruhebedingung in beiden Aufgaben (KLTM und Stroop Test) zu einer deutlichen Leistungsbeeinträchtigung führt (vgl. Maruyama, 1964). Die identischen Schalle mit 60 dB(A) bzw. 50 dB(A) hingegen sollten die Leistung nur noch wenig bzw. gar nicht mehr beeinträchtigen. Der leistungsbeeinträchtigende Effekt von lautem Verkehrslärm auf den Stroop Test und den KLTM sollte sich in einer Erhöhung der Fehlerraten bei der Aufgabenbearbeitung nachweisen lassen. Um einem Speed-Accuracy-Trade-Off vorzubeugen werden die Probanden instruiert, *so schnell wie möglich, aber ohne Fehler* zu arbeiten. Dies wird im Folgenden durch die Auswertung der Reaktionszeiten kontrolliert.

Auf die Serial Recall Aufgabe wird keine negative Wirkung eines lauten gleichmäßigen Schalls erwartet.

Bei allen Aufgaben wird mit einem Übungsfortschritt gerechnet, da alle verwendeten Aufgabenstellungen anspruchsvoll sind. Dies wird entweder durch zeitgleiche Bearbeitung mehrerer Teilaufgaben (KLTM), durch eine starke Belastung der Gedächtnisspanne (Serial Recall: 9 Items; zehnstündiges Retentionsintervall) oder durch hohen Zeitdruck erreicht.

⁴ Vgl. für eine Darstellung der Pausenstruktur der Schalle: exemplarische Amplitudenzeitverläufe für einen fünfminütigen Abschnitt der Schalle in Anhang A

Um eine Überlagerung der erwarteten Schalleffekte mit dem Übungsfortschritt zu vermeiden, werden die Schallbedingungen den Probanden in einem *within-Design* in interindividuell komplett ausbalancierter Reihenfolge dargeboten. Zudem werden die Daten im Folgenden mit Hilfe von Trendanalysen der Fehlerraten und Reaktionszeiten immer auch im Hinblick auf einen Übungseffekt ausgewertet. In allen Tests wird eine Leistungsverbesserung im Sinne sinkender Fehlerraten und abnehmender Reaktionszeiten durch Übung erwartet. Dies dürfte sich anhand eines fallenden Trends in den Daten über die Zeit hinweg nachweisen lassen.

Als Kontrollvariablen werden bei jeweils einem Teil der in den folgenden Kapiteln berichteten Experimente die Leistungsmotivation, Extraversion und Lärmempfindlichkeit der Probanden erhoben (vgl. Kapitel 1.1 und 1.5). Extrem hochmotivierte, lärmempfindliche oder extravertierte Probanden sollten im Gegensatz zu Probanden, die eine extrem niedrige Ausprägung auf diesen Dimensionen angeben, empfindlicher auf Hintergrundschall reagieren, schneller an eine Leistungsgrenze stoßen und somit unter den Schallbedingungen mehr Fehler begehen. Diese Hypothese soll im Folgenden – sofern Extremgruppen in den Daten vorliegen – anhand von Kovarianzanalysen der Fehlerraten mit der Ausprägung dieser potentiellen Moderatorvariable überprüft werden. Ob Extremgruppen vorliegen, wird anhand des Streuungsmaßes (Standardfehler) der Daten der Kontrollvariable geprüft.

Weiterhin wird damit gerechnet, dass leistungsbeeinträchtigende Schalle mehr subjektive Störung verursachen als die Ruhebedingung. Ebenso erwartet man eine signifikant niedrigere Beurteilung der Konzentrationsmöglichkeit und eine subjektiv höhere Beurteilung der aufgewandten Anstrengung und Schwierigkeit der Aufgabe unter leistungsbeeinträchtigenden Schallen im Gegensatz zur Ruhebedingung. Deshalb werden subjektive Beurteilungen der Lästigkeit und Störwirkung der Hintergrundschalle (*Wie lästig/störend fanden Sie die Hintergrundgeräusche?*) sowie der Konzentrationsmöglichkeiten (*Wie gut konnten Sie sich konzentrieren?*), der Schwierigkeit der Aufgabe (*Wie schwer ist Ihnen dieser Test bisher gefallen?*) und der aufgebrauchten Anstrengung (*Als wie anstrengend empfanden Sie die Aufgabe?*) vorgenommen. Nachfolgend werden diese Beurteilungsdimensionen als *Störung*, *Lästigkeit*, *Konzentration*, *Schwierigkeit* und *Anstrengung* bezeichnet. Diese Daten werden anhand von schriftlichen Befragungen (vgl. exemplarisch Anhang B) auf einer fünfstufigen verbal kodierten Skala (*außerordentlich* - *ziemlich* - *mittelmäßig* - *kaum* - *gar nicht*) nach jeder Schallbedingung erhoben und varianzanalytisch ausgewertet.

3.1.2 Allgemeines zur Auswertung

Bei allen in dieser Arbeit dargestellten varianzanalytischen Auswertungen wurde die Voraussetzung der Varianzhomogenität mit Hilfe des *Mauchly Tests auf Sphärizität* geprüft. Wenn dieser auf dem 5%-Niveau signifikant war und der kritische Referenzwert unterschritten wurde ($\epsilon < .75$), wurde der *F-Test* anschließend nach *Greenhouse-Geisser* korrigiert. In der Ergebnisdarstellung wird dies über den Zusatz *Greenhouse-Geisser Korrektur* deutlich gemacht. Es werden gegebenenfalls die korrigierten Werte dargestellt. Im Falle von *Signifikanzen (5% bzw. 1%-Niveau)* wird zusätzlich die geschätzte Effektstärke (das *partielle η^2*) berichtet. Dieses Maß ist unabhängig von der Anzahl der Faktoren.

Zur *Post-hoc Auswertung* wurde das Verfahren der *false discovery rate* gewählt (Benjamini & Hochberg, 1995, 2000; Benjamini & Yekutieli, 2001), das selbst bei einer großen Anzahl von Vergleichen noch valide Signifikanzentscheidungen hervorbringt. Bei diesem Verfahren wird eine Serie von t-Tests berechnet, die anhand ihrer errechneten *p-Werte* aufsteigend sortiert und jeweils mit einem korrigierten Alphawert verglichen werden. Im Folgenden werden diese Werte mit *(BH)* gekennzeichnet. Da diese Korrektur für eine große Anzahl an Post-hoc Tests konzipiert wurde, wird sie vor allem für mehr als vier Vergleiche angewandt.

Bei nicht metrisch skalierten Daten wurde zur varianzanalytischen Auswertung der nonparametrische *Friedman-Test* gewählt. Post hoc wurden diese Daten mit *Wilcoxon-Tests* und der eben beschriebenen *false discovery rate* Korrektur ausgewertet.

Die Mittelwerte und Standardfehler (bzw. Median und Quartile) aller im Folgenden berichteten Leistungsdaten sind in Anhang I tabellarisch dargestellt.

3.1.3 Serial Recall Aufgabe

Die bei den folgenden Experimenten verwendeten Verkehrslärmaufnahmen variierten in Verkehrsdichte (100 vs. 2000 Durchfahrten in der Stunde; vgl. für die Pausenstruktur Anhang A: exemplarische Amplitudenzeitverläufe) und Mittelungspegel (50, 60, 70 dB(A)). Sie wurden im Vergleich zu Hintergrundsprache (Mittelungspegel: 60 dB(A)) und Ruhe untersucht (Tabelle 2; im Folgenden werden die Bezeichnungen der Schalle aus der Tabelle verwendet). Die Schalle setzen sich aus Aufnahmen von Ampelsituationen und Kfz-Vorbeifahrten in einer Geschwindigkeit zwischen 20 und 50 km/h zusammen. Die Aufnahmen wurden mit der Software Artemis 6.0.200 (Head Acoustics GmbH) bearbeitet, um die angegebenen Mittelungspegel herzustellen; die Schalle mit 2000 Durchfahrten pro Stunde unterscheiden sich nur im künstlich veränderten Mittelungspegel.

Tabelle 2: Beschreibung der Schalle

Name	Beschreibung	Dauer	L_{eq} über 12 min in dB(A)
Ruhe			35 dB(A)
Sprache	Helmut Griem spricht Briefe von Georg Büchner	12 min	60 dB(A)
100 – 60	100 PKW Vorbeifahrten pro Stunde	12 min	60 dB(A)
2000 – 50	2000 PKW Vorbeifahrten pro Stunde	12 min	50 dB(A)
2000 – 60	2000 PKW Vorbeifahrten pro Stunde	12 min	60 dB(A)
2000 – 70	2000 PKW Vorbeifahrten pro Stunde	12 min	70 dB(A)

3.1.3.1 Aufgabenstellung

Beim Serial Recall Paradigma, dessen Aufbau bereits in Kapitel 1.3 beschrieben wurde, werden die Ziffern 1 bis 9 auf dem Bildschirm (Schriftgröße: 56 p; Schriftart: Chicago, Darbietungsdauer: 700 ms, Inter Trial Intervall: 300 ms) in randomisierter Reihenfolge präsentiert. Die Aufgabe der Versuchsperson besteht darin, diese Ziffern nach einem Retentionsintervall von 10 Sekunden in der korrekten Reihenfolge auf einem auf dem Bildschirm erscheinenden Zahlenblock mithilfe der Maus abzuklicken. Hierzu werden in randomisierter Reihenfolge die Ziffern 1 bis 9 in einer 3x3 Matrix dargestellt (*Reconstruction of serial order*; Schriftgröße 22 p). Diese Leistungserfassung erwies sich in einer Untersuchung von Klatt, Lee und Hellbrück (2002) als ebenso aussagekräftig wie die handschriftliche Wiedergabe. Eine Korrektur ist nicht möglich, da jede angeklickte Ziffer vom Bildschirm verschwindet.

3.1.3.2 Versuchsplan, Versuchspersonen und Versuchsablauf

Die Serial Recall Aufgabe wurde anhand eines einfaktoriellen varianzanalytischen Designs mit Messwiederholung auf dem sechsfach gestuften Faktor *Schallbedingung* durchgeführt.

An dem Versuch nahmen 18 freiwillige Versuchspersonen (überwiegend Psychologiestudenten) teil, die sich auf einen Aushang hin meldeten und von denen 4 männlich und 14 weiblich waren. Das Alter der Teilnehmer bewegte sich zwischen 20 und 39 Jahren (Median 23 Jahre). Ein Teil von ihnen wurde im Rahmen ihres Studiums mit eineinhalb Versuchspersonenstunden entlohnt, der Rest erhielt als Aufwandsentschädigung 7.50 €. Alle Probanden gaben an, über normales Hörvermögen zu verfügen.

Die Versuchspersonen bearbeiteten nach einer schriftlichen Instruktion (Anhang C) und dem Ausfüllen eines Lärmempfindlichkeitsfragebogens (NoiSeQ, Schütte & Marks, 2004) zunächst neun Übungsaufgaben unter Ruhe. Anschließend wurden im Einzelversuch vor einem Macintosh I-Book pro Schallbedingung 20 Aufgaben vorgegeben, insgesamt folglich 120 (Aufgabe programmiert in Psyscope: J.D. Cohen, Mac Whinney, Flatt, & Provost, 1993). Nach 10 Aufgaben folgte eine zehnssekündige Pause, zwischen den Schallbedingungen jeweils eine Einminütige. Die Schalldarbietung erfolgte über Kopfhörer (Sennheiser HD 600; Sony Compact Disc Player CDP-103). Die Lautstärke der Schalle wurde mit Hilfe eines künstlichen Ohres (Artificial Ear Type 4153 Brüel & Kjaer) kontrolliert. Die Schalle wurden in einer interindividuell balancierten Reihenfolge dargeboten. Der gesamte Versuch fand in einer schallarmen Kammer (Industrial Acoustics Company GmbH) in einem Versuchsraum der Katholischen Universität Eichstätt-Ingolstadt statt. Nach jedem Durchgang füllten die Probanden eine kurze Zwischenbefragung (vgl. Anhang B) aus, in der sie um eine Einschätzung der Störung und Lästigkeit des Hintergrundschalls sowie der ihnen möglichen Konzentration, der Schwierigkeit der Aufgabe und ihrer aufgewandten Anstrengung gebeten wurden.

3.1.3.3 Ergebnisse

Die Auswertung der prozentualen Fehlerraten in dieser Aufgabe unter dem Faktor *Schallbedingung* ergibt folgendes, in Abbildung 3 dargestelltes Bild.

Eine Varianzanalyse (mit Messwiederholungen) verdeutlicht einen signifikanten Einfluss des Faktors *Schallbedingung* (Greenhouse-Geisser Korrektur: $F(2.9; 48.6)=5.9$; $p<.01$; partielles $\eta^2=0.26$). Die Post-hoc Auswertung mit den gerichteten Hypothesen, dass die Fehlerraten unter den zeitlich stark variablen Schallen (*Sprache* und *100-60*) gegenüber der Ruhebedingung zunehmen, zeigt, dass höhere Fehlerraten unter *Sprache* als unter der Ruhebedingung vorliegen (einseitig: Ruhe: $T(17)=4.7$; $p<.01$ (BH=.01)). Die Leistung unter dem Hintergrundschall *100-60* jedoch unterscheidet sich nicht signifikant von *Ruhe* ($T(17)=1.1$; $p=.14$ (BH=.02)). Wie vermutet, beeinträchtigen die lauten, aber temporal gleichmäßigen Verkehrshintergrundschalle *2000-70* ($T(17)=1.1$; $p=.14$ (BH=.03)) und *2000-60* ($T(17)=0.6$; $p=.26$ (BH=.05)) die Leistung im Vergleich zur Ruhebedingung nicht. Gleiches gilt für den leisesten Schall *2000-50* ($T(17)=1.1$; $p=.15$ (BH=.04)).

Die Ergebnisse des Lärmempfindlichkeitsfragebogens (NoiseQ: vorläufige Normwerte (Schütte & Marks, 2004): $M=43$; $SD=18$; gemessen: $M=45.7$; $SE=3.1$) weisen aufgrund der nicht auffallend hohen Streuung im Vergleich zu den vorläufigen Normwerten nicht auf Extremgruppen hin, weswegen auf eine Aufnahme dieser Variable als Kovariate in die Auswertung verzichtet wurde.

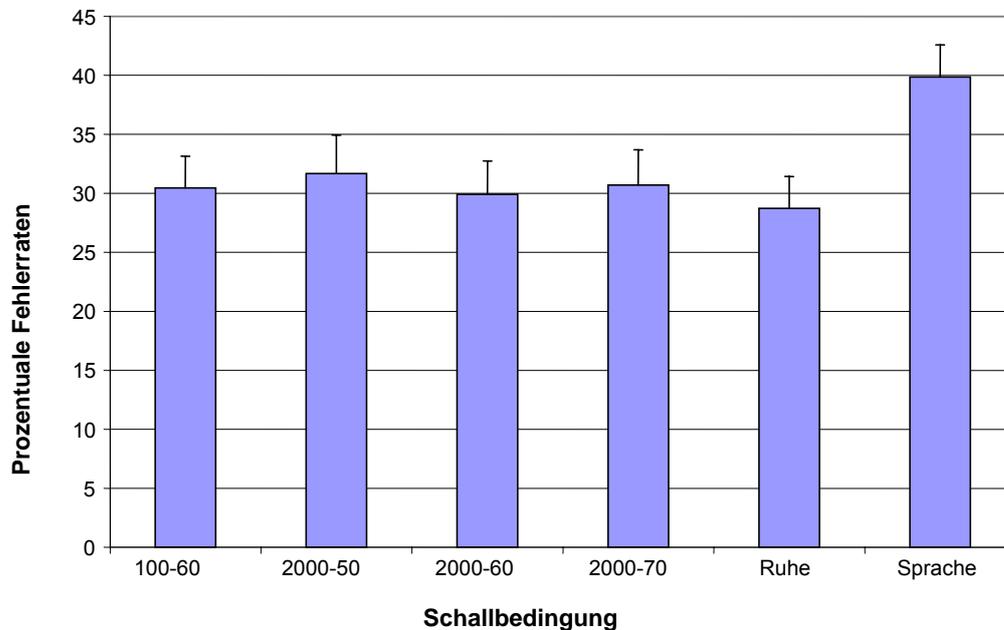


Abbildung 3: Prozentuale Fehlerraten unter dem Faktor *Schallbedingung* im Serial Recall (Mittelwert und Standardfehler)

Um einen potentiellen Übungseffekt zu kontrollieren, wurden die Fehlerraten unabhängig von der Schalldarbietung über die Versuchsblöcke gemittelt. Im Folgenden wird dieser Faktor als *Versuchsabfolge* benannt. Es gibt allerdings keine Anzeichen für einen Übungseffekt ($F(5; 85)=0.4$; $p=.81$). Eine varianzanalytische Auswertung der von den Probanden für die Eingabe der Ziffern über den Zahlenblock benötigten Zeit ergibt zudem wie vermutet keine Hinweise auf einen Speed-Accuracy-Trade-Off. Die Lösungszeiten werden durch den Faktor *Schallbedingung* nicht nachweisbar beeinflusst (Greenhouse-Geisser Korrektur: $F(3; 51.1)=1.8$; $p=.15$). Auch die Lösungszeiten wurden ungeachtet der Art des im Einzelnen dargebotenen Schalls nach dem Faktor *Versuchsabfolge* ausgewertet. Auch hier deutet nichts auf einen Übungseffekt hin (Greenhouse-Geisser Korrektur: $F(2.4; 41.4)=1.7$; $p=.19$).

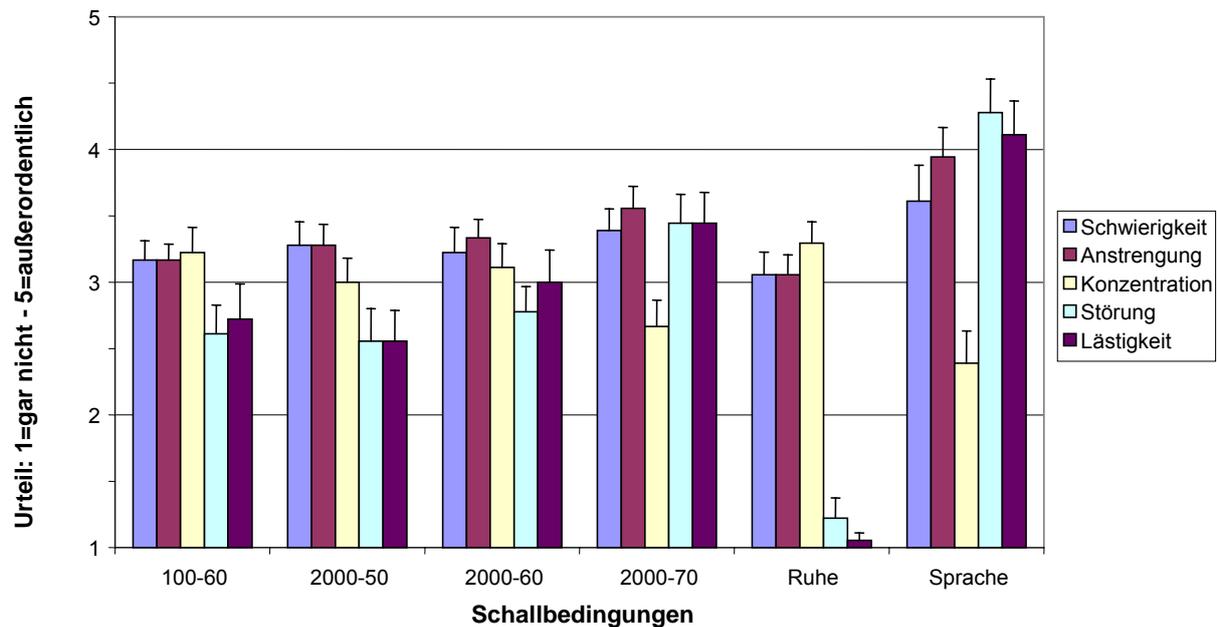


Abbildung 4: Gemittelte Urteile in der Zwischenbefragung im Serial Recall (Mittelwert und Standardfehler), Skala von gar nicht (1) bis außerordentlich (5)

Eine Auswertung der Urteile der Versuchspersonen über die Items der Zwischenbefragung ergibt das in Abbildung 4 dargestellte Bild. Die Probanden wurden jeweils direkt im Anschluss an die Bearbeitung der Aufgabe unter einer Schallbedingung um die Beantwortung von fünf Fragen (Störwirkung und Lästigkeit der Schalle, Konzentrationsmöglichkeit, Anstrengung und Schwierigkeit der Aufgabe) gebeten. Hierbei kam eine fünfstufige verbal kodierte Skala mit den Polen *gar nicht* bis *außerordentlich* zum Einsatz (vgl. Anhang B). Die Daten wurden getrennt für die einzelnen Fragen (*Störung*, *Lästigkeit*, *Schwierigkeit*, *Anstrengung* und *Konzentration*) ausgewertet. Im Falle von signifikanten Unterschieden in der Beantwortung der Fragen durch den Faktor *Schallbedingung* wurde eine Post-hoc Auswertung (einseitige Testung: zur Begründung vgl. Kapitel 3.1.1) hinsichtlich der Urteile über die Schalle im Vergleich zur Ruhebedingung vorgenommen.

Für die Frage nach der empfundenen Schwierigkeit der Aufgabe ergibt sich in einer einfaktoriellen Varianzanalyse ein signifikanter Haupteffekt des Faktors *Schallbedingung* ($F(5; 85)=2.9$; $p=.02$; partielles $\eta^2=0.15$). Eine Post-hoc Auswertung nach dem oben beschriebenen Schema von Benjamini und Hochberg erbrachte hypothesenkonform eine signifikant geringere Beurteilung der Schwierigkeit der Aufgabe unter *Ruhe* als unter der Schallbedingung *Sprache* ($T(17)=-2.6$; $p=.01$ ($BH=.01$)). Genau wie bei den Fehlerraten gilt dies aber nicht für den zeitlich stark variablen Verkehrsschall *100-60* ($T(17)=-0.8$; $p=.21$ ($BH=.05$)); ebenso ist keine erhöhte empfundene Schwierigkeit für die gleichmäßigen Verkehrsschalle in mittlerer Lautstärke nachweisbar (2000-60: $T(17)=-1.1$; $p=.13$ ($BH=.04$); 2000-50: $T(17)=-1.7$; $p=.05$ ($BH=.03$)). Allerdings wird die Aufgabe unter

dem lautesten Verkehrsschall 2000-70 als schwieriger empfunden als unter *Ruhe* ($T(17)=-2.4$; $p=.02$ (BH=.02)).

Die Einschätzung der aufgewandten Anstrengung für die Bearbeitung der Aufgabe unterscheidet sich ebenfalls je nach Ausprägung des Faktors *Schallbedingung* signifikant ($F(5; 85)=5.9$; $p<.01$; partielles $\eta^2=0.26$). Die Post-hoc Auswertung zeigt, dass hier die Unterschiede in der Beurteilung der Ruhebedingung im Vergleich zur Hintergrundsprache ($T(17)=-4.2$; $p<.01$ (BH=.01)) und zur lautesten Verkehrsschallbedingung 2000-70 ($T(17)=-2.7$; $p<.01$ (BH=.02)) begründet liegen. Keine nachweisbaren Unterschiede in der Beurteilung der Anstrengung ergeben sich zwischen den restlichen drei Verkehrslärmbedingungen und der Ruhebedingung (100-60: $T(17)=0.8$; $p=.21$; (BH=.05); 2000-50: $T(17)=1.5$; $p=.08$; (BH=.04); 2000-60: $T(17)=1.8$; $p=.05$ (BH=.03)). Sowohl unter dem Sprachhintergrundschall als auch unter dem lautesten Verkehrsschall 2000-70 wurde die Aufgabe somit subjektiv als anstrengender erlebt.

Das Urteil über die *Konzentration* fällt wie das über die *Anstrengung* in Abhängigkeit von der *Schallbedingung* unterschiedlich aus (Greenhouse-Geisser Korrektur: $F(2.9; 46.2)=3.8$; $p=.02$; partielles $\eta^2=0.19$). Eine Post-hoc Auswertung verdeutlicht eine bessere Beurteilung der Konzentrationsmöglichkeit unter Ruhe als unter dem Sprachschall ($T(17)=2.7$; $p<.01$ (BH=.01)) sowie dem Verkehrsschall 2000-70 ($T(17)=2.4$; $p=.01$ (BH=.02)). Unter den anderen Verkehrsschallen jedoch wird die Konzentrationsmöglichkeit nicht als signifikant schlechter als unter Ruhe eingeschätzt (100-60: $T(17)=-0.2$; $p=.42$ (BH=.05); 200-50: $T(17)=-1.0$; $p=.16$ (BH=.03); 2000-60: $T(17)=-0.7$; $p=.24$ (BH=.04)).

Auch in Bezug auf die Beurteilung der Störwirkung der Hintergrundschalle durch die Probanden zeigt sich ein statistisch bedeutsamer Einfluss der *Schallbedingung* ($F(5; 85)=35.7$; $p<.01$; partielles $\eta^2=0.68$). Die Ruhebedingung wird als weniger störend erlebt als alle Schallbedingungen (Post-hoc Testung: Sprache: $T(17)=-10.7$; $p<.01$ (BH=.01); 2000-70: $T(17)=-9.4$; $p<.01$ (BH=.02); 2000-60: $T(17)=-6.7$; $p<.01$ (BH=.03); 2000-50: $T(17)=-4.6$; $p<.01$ (BH=.05); 100-60: $T(17)=-5.1$; $p<.01$ (BH=.04)).

Bezogen auf die *Lästigkeit* ergibt sich ebenfalls ein signifikanter Haupteffekt ($F(5; 85)=42.6$; $p<.01$; partielles $\eta^2=0.72$) des Faktors *Schallbedingung*. Die Post-hoc Auswertung legt nahe, dass alle Schalle als signifikant lästiger beurteilt werden als die Ruhebedingung, die Sprache aber am lästigsten von allen (Sprache: $T(17)=-11.7$; $p<.01$ (BH=.01); 2000-70: $T(17)=-10.4$; $p<.01$ (BH=.02); 2000-60: $T(17)=-7.8$; $p<.01$ (BH=.03); 2000-50: $T(17)=-6.5$; $p<.01$ (BH=.04); 100-60: $T(17)=-6.0$; $p<.01$ (BH=.05)).

3.1.3.4 Diskussion

Die Serial Recall Aufgabe wird, wie in Kapitel 1.3 beschrieben, meist zum Nachweis des Irrelevant Sound Effect herangezogen. Auch in der vorgestellten Durchführung dieser Aufgabe lässt sich die postulierte störende Wirkung des Sprachschalls durch eine Erhöhung der Fehlerraten aufzeigen. Zudem wird der Sprachschall von den Probanden im Vergleich zur Ruhebedingung als schwieriger, anstrengender, ungünstiger für die Konzentration, störender und lästiger beurteilt. Wie

vermutet finden sich keine Hinweise auf einen Speed-Accuracy-Trade-Off. Der zeitlich stark variable Verkehrsschall (100 Durchfahrten pro Stunde) scheint einen zu geringen *Changing State Charakter* aufzuweisen, um eine Störwirkung hervorzurufen. In Anhang A sind die Amplitudenzeitverläufe in exemplarischen Ausschnitten aller verwendeten Schalle dargestellt. Zu sehen ist die stärkere Pausenstruktur dieses Schalls (Abb. 36) im Vergleich zum Verkehrsschall mit 2000 Durchfahrten (Abb. 37). Diese temporale Struktur muss aber verglichen mit der des Sprachschallsignals (Abb. 39) als eher gering ausgeprägt gelten.

Der Mittelungspegel der gleichmäßigen Hintergrundschalle (50, 60 oder 70 dB(A)) ist, wie in Kapitel 1.3 für den ISE postuliert, für die Leistung in diesem Test irrelevant. Allerdings wird der lauteste Verkehrsschall (2000-70), obwohl er die Leistung nicht beeinträchtigt, von den Probanden als störender und lästiger als die Ruhebedingung beurteilt. Dieser Schall scheint zudem im Vergleich zur Ruhebedingung die Aufgabe subjektiv zu erschweren, anstrengender zu machen und die Konzentration zu beeinträchtigen. Diese negativen Auswirkungen des Schalls finden sich jedoch nicht in den Fehlerraten wieder.

In Kapitel 1.4 wurde die Überlegung von Näätänen (1975) dargelegt, dass die leistungsbeeinträchtigende Wirkung von Lärm besser durch die von Schall ausgelöste Ablenkung durch Gedanken erklärbar sei als durch eine Arousalsteigerung. Hierfür ergibt sich im vorliegenden Test wenig Evidenz. Die Probanden empfinden den Verkehrslärm zwar als störend, lästig, konzentrationshinderlich und er erwecke den Eindruck die Aufgabe zu erschweren und anstrengender zu machen, was sich jedoch nicht in einem Niederschlag auf die Leistungsdaten finden lässt. Die individuelle Lärmempfindlichkeit (erhoben mit dem NoiSeQ) variiert unter den Probanden nicht genügend, um einen Extremgruppenvergleich zu rechtfertigen.

3.1.4 Modifikation des Konzentrations-Leistungs-Tests – KLTM

Im Folgenden soll eine modifizierte Version des Konzentrations-Leistungs-Tests (Düker & Lienert, 1965) vorgestellt werden. Als Hintergrundbedingungen dienten hier ebenfalls die in Tabelle 2 beschriebenen sechs Schalle.

3.1.4.1 Aufgabenstellung

Die Aufgabe im verwendeten Test (als *KLTM* bezeichnet, wobei *M* für *Modifikation* steht) besteht in Anlehnung an den Konzentrations-Leistungs-Test (KLT, Düker & Lienert, 1965) aus Kopfrechenaufgaben, die sich aus *drei* dreiteiligen Additions- bzw. Subtraktionsaufgaben zusammensetzen und nach folgender *Regel* zu einem Gesamtergebnis zu verrechnen sind: Ist das Ergebnis der ersten Aufgabe größer als das der dritten, so sind die Ergebnisse zu subtrahieren, andernfalls zu addieren. Das Ergebnis der zweiten Aufgabe wird nach derselben Regel mit dem errechneten Zwischenergebnis verglichen und verrechnet. Nachfolgend ist beispielhaft eine Aufgabe dargestellt.

(1) $6+3-5= 4$

(2) $3-1+7= 9$

(3) $4-1-1= 2$

(4) Zwischenergebnis: Verrechnung von (1) und (3) nach der *Regel*= $4-2= 2$

(5) Endergebnis: Verrechnung von (2) und (4) nach der *Regel*= $9-2= 7$

Dieses auf Psyscope (J.D. Cohen et al., 1993) programmierte Experiment besteht aus der Präsentation der Summanden der Aufgaben in der Mitte des Bildschirms. Die Summanden setzen sich aus den Ziffern 1 bis 9 mit einem Operanden (+/-) zusammen (Schriftgröße: 50 p; Schriftart: Chicago). Die Ziffern werden einzeln sukzessive präsentiert, um die Anforderungen an das Arbeitsgedächtnis zu erhöhen. Diese Erhöhung wird zusätzlich durch eine verkürzte Darbietung erreicht, wie bereits im Abschnitt 1.5 beschrieben wurde. Um die Anforderungen an das Arbeitsgedächtnis zu steigern werden weiterhin gegenüber dem KLT drei statt zwei Rechenaufgaben vorgegeben. Die Probanden müssen folglich in dieser Modifikation des KLT mehr Zwischenergebnisse für längere Zeit memorieren, während sie die Berechnung der restlichen Aufgabe durchführen.

Die Geschwindigkeit der Ziffernpräsentation bestimmt die Versuchsperson durch Drücken der Leertaste selbst (300 ms nach Betätigung der Taste wird die nächste Zahl präsentiert). Vor jeder der Teilaufgaben weist ein sich verkleinerndes Quadrat (2x 500 ms) in der Mitte des Bildschirms auf die beginnende Teilaufgabe hin (100 ms danach wird die erste Zahl präsentiert). Zu Beginn einer neuen Aufgabe gibt es einen Signalton (Inter Trial Intervall: 500 ms). Das errechnete Endergebnis müssen die Probanden am Ende jeder Aufgabe über die Tastatur eingeben. Eine Korrektur ist möglich, bis sie die Eingabe mit *Return* bestätigen. Erst nach der Bestätigung der Lösung endet die Lösungszeiterfassung. Bei der Eingabe muss zuerst die Zehnerstelle und anschließend die Einerstelle

eingetippt werden, was ebenfalls eine Erhöhung der Anforderungen an die Arbeitsgedächtniskapazität darstellt (Donlan, 1998). Aufgrund der gesteigerten Anforderungen an das Arbeitsgedächtnis wurden die Anzahl der Überträge und die Größe der errechneten Zahlen randomisiert, um eine möglichst homogene Schwierigkeitsverteilung der einzelnen Aufgaben zu erreichen (vgl. Kapitel 1.5).

3.1.4.2 *Versuchsplan, Versuchspersonen und Ablauf*

Der Versuchsplan lag als einfaktorielles varianzanalytisches Design mit Messwiederholung auf dem erneut sechsfach gestuften Faktor *Schallbedingung* vor.

An dem insgesamt dreistündigen Versuch nahmen 18 freiwillige Probanden teil (5 männlich, 13 weiblich), überwiegend Psychologiestudenten im Alter zwischen 19 und 31 Jahren (Median 23 Jahre). Alle hatten sich auf einen Aushang hin gemeldet. Alle gaben an, ein normales Hörvermögen zu besitzen. Ein Großteil von ihnen wurde im Rahmen des Studiums mit drei Versuchspersonenstunden entlohnt, der Rest erhielt 15 € als Aufwandsentschädigung.

Jede Versuchsperson bearbeitete im Einzelversuch an zwei Testterminen vor einem Apple Macintosh-I-Book unter jedem der sechs Schalle (vgl. Tabelle 2) für eine Dauer von 15 Minuten so viele Aufgaben wie möglich. Die Probanden bearbeiteten beim ersten Termin nach einer Instruktion (Anhang D), einem Leistungsmotivationstest (LMI; Schuler & Prochaska, 2000) und einem Extraversionsfragebogen (EPI; Eggert, 1974) zunächst zehn Übungsaufgaben unter Ruhe. Anschließend wurde ihnen dreimal jeweils eine der Schallbedingungen für 15 Minuten präsentiert. Die Schalle wurden über Kopfhörer (Sennheiser HD 600; Sony Compact Disc Player CDP-103) dargeboten. Die Reihenfolge der Schalldarbietungen wurde zwischen den Probanden ausbalanciert. Anschließend an jede Bedingung füllten sie die oben beschriebene Zwischenbefragung aus (vgl. Anhang B). Am zweiten Termin (nach einer Woche) wurde wieder dreimal je 15 Minuten unter den restlichen drei Schallen gearbeitet. Der Versuch fand in einem ruhigen Versuchsraum statt.

3.1.4.3 *Ergebnisse*

Da sich beim KLTM mehrere Maße zur Auswertung anbieten, wurden zunächst die Anzahl der bearbeiteten und der richtig gelösten Aufgaben sowie die absolute Fehleranzahl ausgewertet (Abbildung 5).

Eine varianzanalytische Auswertung (einfaktoriell mit Messwiederholungen) der Anzahl der bearbeiteten und richtig gelösten Aufgaben pro 15-Minuten-Abschnitt lässt keinen Schluss auf einen bedeutsamen Einfluss des Faktors *Schallbedingung* zu (gesamt bearbeitet: $F(5; 85)=0.4$; $p=.84$; richtig gelöst: $F(5; 85)=1.4$; $p=.24$). Auf die Anzahl der Fehler hingegen hat die *Schallbedingung* einen signifikanten Einfluss ($F(5; 85)=2.3$; $p=.05$; partielles $\eta^2=0.12$). Anhand eines gerichteten t-Tests lässt sich wie vermutet ein statistisch bedeutender Unterschied in der absoluten Anzahl der Fehler zwischen den Schallbedingungen *Sprache* und *Ruhe* nachweisen. Unter Sprachschall werden signifikant mehr Fehler gemacht ($T(17)=1.9$; $p(\text{einseitig})=.04$). Auf eine Durchführung aller anderen Vergleiche wurde aufgrund der geringen Unterschiede verzichtet.

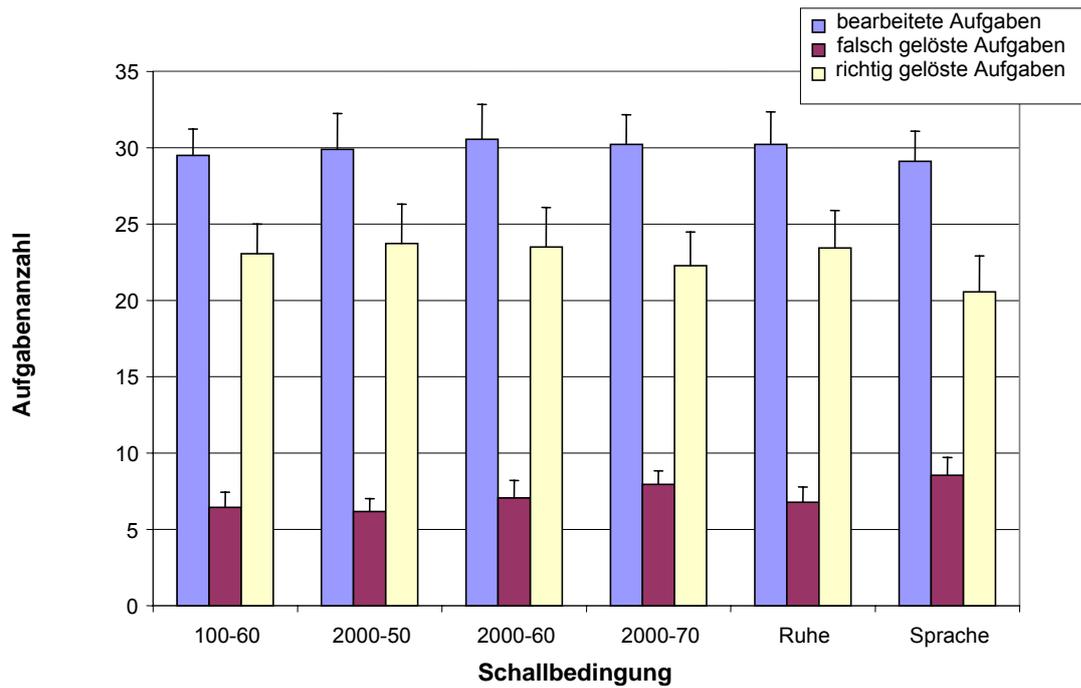


Abbildung 5: Anzahl der bearbeiteten, richtig und falsch gelösten Aufgaben im KLTM (Mittelwert und Standardfehler)

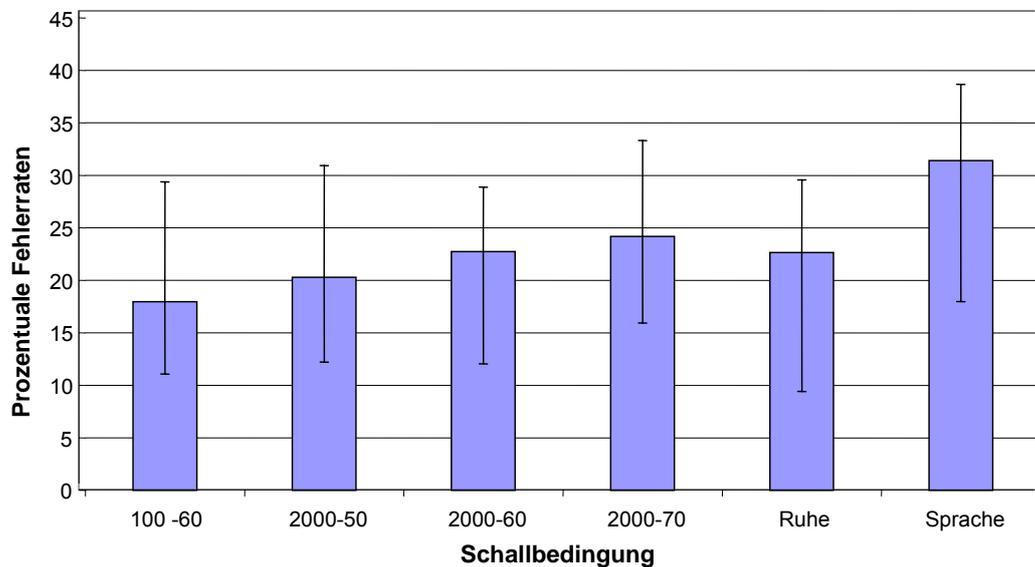


Abbildung 6: Prozentuale Fehlerraten unter den Schallbedingungen im KLTM (Median sowie 1. und 3. Quartil)

Um sowohl die Güte als auch die Geschwindigkeit der Arbeit in der Auswertung zu berücksichtigen, wurden im Folgenden die prozentualen Fehlerraten ausgewertet (Abbildung 6; zu beachten ist, dass die großen Streuungswerte auf einem Darstellungseffekt beruhen, da der Median

mit Quartilen dargestellt ist). Da bei diesen Daten aber aufgrund unterschiedlicher Grundhäufigkeiten nicht von einer metrischen Skalierung ausgegangen werden kann, wurde eine nonparametrische Varianzanalyse (nach *Friedman*) vorgenommen. Im KLTM war die Bearbeitungszeit festgelegt, weshalb ein prozentualer Fehlerwert für sehr verschiedene Leistungsgüte stehen kann. Ein Proband der sehr wenige Aufgaben löst und davon nur wenige falsch macht, würde denselben Wert erhalten wie einer, der viele löst, diese aber prozentual nur in der selben Güte. Die nonparametrische Auswertung erbringt einen signifikanten Haupteffekt des Faktors *Schallbedingung* ($\chi^2=11.0$; $df=5$; $p=.05$).

Hierauf wurden zwei nonparametrische Einzelvergleiche mit der Ruhebedingung durchgeführt, und zwar mit den Schallbedingungen *Sprache* und *2000-70* (*Wilcoxon Test*; einseitige Testung wegen gerichteter Hypothesen, zur theoretischen Begründung vgl. Kapitel 3.1.1). Hierbei konnte ein bedeutsamer Unterschied zwischen der Leistung unter *Sprache* und *Ruhe* ($Z=-2.1$; $p=.02$) gezeigt werden, jedoch nicht zwischen dem lauten Verkehrsschall und der Ruhebedingung (2000-70: $Z=0.7$; $p=.23$). Dessen ungeachtet unterscheidet sich die Leistung unter den Verkehrsschallen *2000-70* und *2000-50* ($Z=-2.0$; $p=.02$) statistisch signifikant voneinander. Unter dem leiseren Verkehrsschall war die Leistung besser. Ein Unterschied zwischen den beiden Verkehrsschallbedingungen unterschiedlicher temporaler Struktur mit 60 dB(A) (100 versus 2000 Durchfahrten) ist nicht nachweisbar ($Z = -0.851$; $p=.20$). Da nur vier gezielte Post-hoc Vergleiche durchgeführt wurden, fand keine Korrektur nach Benjamini-Hochberg statt (vgl. Roberts & Russo, 1999).

Auf individueller Ebene zeigt sich bei 14 von 18 Probanden eine Verschlechterung der Leistung unter dem Sprachschall im Vergleich zur Ruhebedingung, unter dem lautesten Verkehrsschall immerhin noch bei 12 von 18 Probanden.

Bezüglich der ansteigenden Lautstärke des Verkehrslärms von 50 dB(A) über 60 dB(A) zu 70 dB(A) scheint ein ansteigender Trend vorzuliegen, je lauter der Schall wird, desto mehr Fehler werden gemacht (vgl. Abbildung 6).

Die erfassten Kontrollvariablen Leistungsmotivation (Prozentränge (Schuler & Prochaska, 2000): $M=39.3$; $SE=5.4$) und Extraversion (Normwerte (Eggert, 1974): $M=14.1$; $SD=3.9$; gemessen: $M=14.6$; $SE=0.9$) weisen im Vergleich zu den Normwerten keine Streuung in einer Größenordnung auf, die einen Hinweis auf Extremgruppen gibt, weswegen auf eine weitere kovarianzanalytische Auswertung verzichtet wurde.

Die varianzanalytische Auswertung der mittleren Lösungszeiten (in ms) pro Aufgabe im KLTM ergibt keinen nachweisbaren Einfluss des Faktors *Schallbedingung* ($F(5; 85)=0.6$; $p=.71$). Somit ergeben sich wie vermutet keine Hinweise auf einen Speed-Accuracy-Trade-Off.

Des Weiteren wurden alle oben genannten Maße, die Gesamtzahl der bearbeiteten Aufgaben und deren richtig bzw. falsch gelöst, die prozentualen Fehlerraten und die Reaktionszeiten im Hinblick auf einen Übungseffekt ausgewertet. Hierzu wurden die Daten (wie beim Serial Recall beschrieben) nach der zeitlichen Abfolge im Experimentalverlauf ausgewertet, wobei der jeweils

dargebotene Schall vernachlässigt wurde. Diese Auswertungen deuten auf einen sehr großen Übungseffekt hin.

Drei einfaktorielle Varianzanalysen mit Messwiederholungen ergeben einen statistisch bedeutsamen Effekt des Faktors *Versuchsabfolge* für die Gesamtzahl der bearbeiteten Aufgaben (Greenhouse-Geisser Korrektur: $F(2.9; 49.1)=27.3$; $p<.01$; partielles $\eta^2=0.62$) und die Anzahl der richtigen Lösungen ($F(5; 85)=19.9$; $p<.01$; partielles $\eta^2=0.54$). Die Anzahl der Fehler hingegen zeigt keine nachweisbare Änderung durch Übung ($F(5; 85)=0.9$; $p=.50$). Eine Trendanalyse bestätigt diese Ergebnisse (vgl. Kapitel 3.1.1 zur theoretischen Begründung der Trendanalyse). Es findet sich ein hochsignifikant steigender Trend für die Anzahl der richtig gelösten ($L_{emp}=1541>L_{tab}(18; 6; 1\%)=1401$) und die Gesamtzahl der bearbeiteten Aufgaben ($L_{emp}=1559>L_{tab}(18; 6; 1\%)=1401$), was auf eine Leistungsverbesserung durch Übung hindeutet.

Eine Analyse der prozentualen Fehlerraten, die ebenfalls mit dem Faktor *Versuchsabfolge* durchgeführt wurde, ergibt einen signifikanten Übungseffekt (Friedmantest: $\chi^2=15.5$; $df=5$; $p=.01$). Erst im zweiten Teil des Tests scheint sich die Leistung der Probanden zu stabilisieren. Eine Trendanalyse bestätigt dieses Ergebnis, es lässt sich ein hochsignifikant fallender Trend in den Daten nachweisen (Trendtest nach Page: $L_{emp}=1447>L_{tab}(18; 6; 1\%)=1401$).

Ähnliche Ergebnisse gehen aus einer Auswertung der mittleren Lösungszeiten pro Aufgabe hervor. Hier zeigt sich ein statistisch bedeutsamer Effekt des Faktors *Versuchsabfolge* (Greenhouse-Geisser Korrektur: $F(2.6; 44.9)=32.0$; $p<.01$; partielles $\eta^2=0.65$). Ein hochsignifikant fallender Trend lässt sich ebenfalls in den Daten nachweisen ($L_{emp}=1569>L_{tab}(18; 6; 1\%)=1401$), was erneut für zunehmende Übung spricht.

Auch beim KLTM wurden die Daten der Zwischenbefragungen ausgewertet, um der Art der Beeinträchtigung durch die Hintergrundschalle auf den Grund zu gehen.

In Abbildung 7 sind die Urteile aus den Zwischenbefragungen dargestellt. Gefragt wurde nach der Anstrengung durch die Aufgabe, der Schwierigkeit des Rechnens, der Konzentrationsmöglichkeit und der Lästigkeit der Schalle (in dieser Version der Zwischenbefragung nicht nach der Störwirkung). Die Probanden konnten auch hier ihr Urteil auf einer verbal skalierten fünfstufigen Skala von *gar nicht* bis *außerordentlich* abgeben.

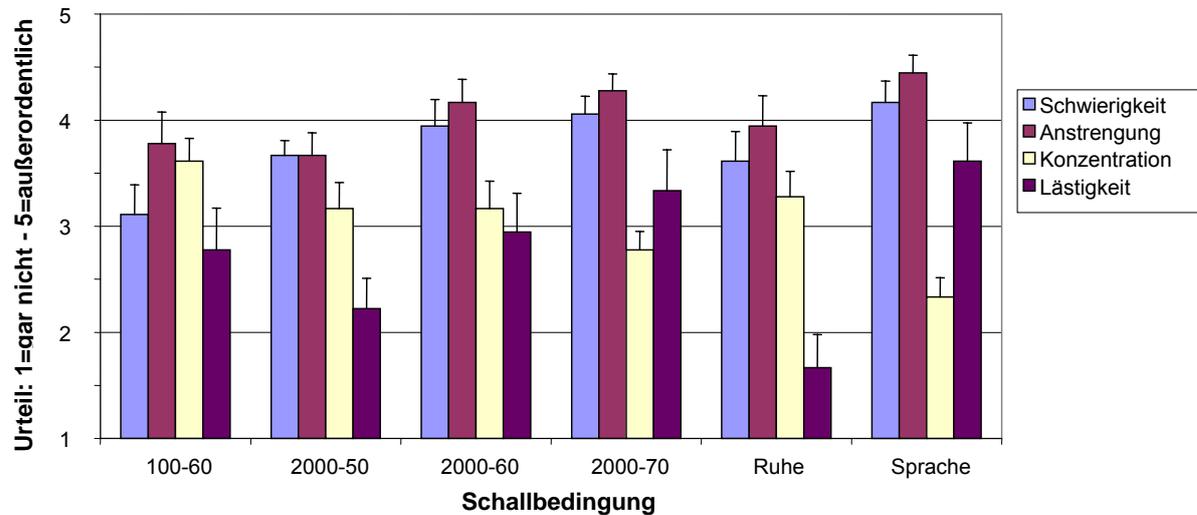


Abbildung 7: Urteile der Probanden in den Zwischenbefragungen im KLTM unter den verschiedenen Schallbedingungen, getrennt nach den Urteilkriterien (*Schwierigkeit*, *Anstrengung*, *Konzentration* und *Lästigkeit*) auf einer fünfstufigen Skala (1=gar nicht - 5=außerordentlich) (Mittelwert und Standardfehler)

Die Daten wurden varianzanalytisch für die einzelnen Fragen (*Lästigkeit*, *Schwierigkeit*, *Anstrengung* und *Konzentration*) ausgewertet. Im Falle von signifikanten Unterschieden durch den Faktor *Schallbedingung* wurden die Urteile über die Schalle im Vergleich zur Ruhebedingung post hoc ausgewertet (äquivalent zur Auswertung der Fehlerraten; einseitige Testung aufgrund gerichteter Hypothesen, vgl. Kapitel 3.1.1). Die einfaktorielle varianzanalytische Testung ergibt statistisch bedeutsame Unterschiede durch den Faktor *Schallbedingung* für jede der Urteilsdimensionen (Schwierigkeit: $F(5; 85)=5.1$; $p<.01$; partielles $\eta^2=0.38$; Anstrengung: $F(5; 85)=2.4$; $p=.05$; partielles $\eta^2=0.25$; Konzentration: $F(5; 85)=5.3$; $p<.01$; partielles $\eta^2=0.38$; Lästigkeit: $F(5; 85)=6.8$; $p<.01$; partielles $\eta^2=0.41$).

Die subjektiv erlebte Schwierigkeit der Aufgabe unterscheidet sich in der Post-hoc Auswertung signifikant zwischen *Ruhe* und den Schallen *Sprache* ($T(17)=-2.1$; $p=.02$ (BH=.02)) und *100-60* ($T(17)=-2.3$; $p=.01$ (BH=.01)). Durch die Korrektur nicht mehr signifikant ist der Unterschied zwischen dem lautesten Verkehrslärm *2000-70* und *Ruhe* ($T(17)=-1.8$; $p=.04$ (BH=.03)). Ebenfalls die Signifikanz verfehlen die Urteile über die anderen beiden Verkehrsschalle mit 2000 Durchfahrten pro Stunde im Vergleich zur Ruhebedingung (2000-60: $T(17)=1.2$; $p=.12$; (BH=.04); 2000-50: $T(17)=0.2$; $p=.43$ (BH=.05)).

Die *Anstrengung* ist durch die Schalle im Vergleich zur Ruhebedingung nicht nachweisbar beeinflusst (*Sprache*: $T(17)=1.6$; $p=.06$ (BH=.01); 2000-70: $T(17)=1.3$; $p=.10$ (BH=.02); 2000-60: $T(17)=0.7$; $p=.26$ (BH=.04); 2000-50: $T(17)=-1.0$; $p=.18$ (BH=.03); 100-60: $T(17)=-0.5$; $p=.33$ (BH=.05)).

Die Urteile über die *Konzentration* sind unter den Schallbedingungen *Sprache* und *2000-70* im Vergleich zur Ruhebedingung signifikant niedriger. Unter beiden Schallen scheint es subjektiv schwerer gewesen zu sein, sich zu konzentrieren (*Sprache*: $T(17)=-5.0$; $p<.01$ ($BH=.01$); *2000-70*: $T(17)=-2.0$; $p=.02$ ($BH=.02$)). Unter allen anderen Schallbedingungen unterscheidet sich das Urteil über die *Konzentration* nicht von dem unter der Ruhebedingung (*2000-60*: $T(17)=-0.4$; $p=.34$ ($BH=.05$); *2000-50*: $T(17)=-0.5$; $p=.31$ ($BH=.04$); *100-60*: $T(17)=1.1$; $p=.14$ ($BH=.03$)).

Signifikant lästiger als *Ruhe* werden von den Probanden die Schalle *Sprache* ($T(17)=4.6$; $p<.01$ ($BH=.01$)), *2000-70* ($T(17)=3.8$; $p<.01$ ($BH=.02$)), *2000-60* ($T(17)=2.7$; $p<.01$ ($BH=.03$)) und *100-60* ($T(17)=2.4$; $p=.02$ ($BH=.04$)) eingeschätzt, nicht jedoch der leiseste Verkehrsschall *2000-50* ($T(17)=1.4$; $p=.09$ ($BH=.05$)).

3.1.4.4 Diskussion

Es konnte ein signifikanter Einfluss des Sprachschalls auf den KLTM gezeigt werden. Die Modifikationen, die am Vorbild des KLT (Düker & Lienert, 1965) vorgenommen wurden, haben die Belastung der beiden Subsysteme des Arbeitsgedächtnisses gesteigert, wie in Kapitel 1.5 angesprochen wurde. Dies spiegelt sich in den vergleichsweise hohen Fehlerraten (17-32%) wieder.

Es finden sich ansteigende Fehlerzahlen und sinkende Lösungsraten ohne eine Verlängerung der Lösungszeiten unter dem sprachlichen Hintergrundschall im Vergleich zur Ruhebedingung. Die Leistung unter dem lautesten Verkehrslärm ($L_{eq}=70$ dB(A)) unterscheidet sich zwar nicht, wie vermutet von der unter der Ruhebedingung, allerdings ergibt sich ein signifikanter Unterschied im Vergleich zur Leistung unter dem leisesten Verkehrslärm ($L_{eq}=50$ dB(A)). Somit konnte die Hypothese, dass beide Parameter (Lautstärke und spektral-temporale Struktur) bei dieser Aufgabe eine entscheidende Rolle spielen, zumindest teilweise bestätigt werden. Die fraglich beeinträchtigende Wirkung des lauten Verkehrslärms soll in weiteren Experimenten (siehe Fragestellung II) genauer spezifiziert werden.

Ein Einfluss der temporalen Variation des Verkehrsschalls (Vergleich von *100-60* mit *2000-60*) konnte nicht hinreichend genau spezifiziert werden. Jedoch muss auch in diesem Experiment davon ausgegangen werden, dass die spektral-temporale Variation im Verkehrsschall mit *100* Durchfahrten gegenüber dem Sprachhintergrundschall eher gering ausgeprägt ist (vgl. Pegelzeitverläufe Anhang A).

Die Leistungsbeeinträchtigung durch den Sprachhintergrundschall jedoch gibt Hinweise darauf, dass der ISE auch in anderen Aufgaben als dem Serial Recall nachgewiesen werden kann.

Beim KLTM zeigte sich ein starker Übungseffekt, der unabhängig vom Schallhintergrund zu einer signifikanten Verbesserung der Leistung führt. Je länger die Probanden arbeiten, desto mehr verkürzen sich die Lösungszeiten, steigen die Lösungsraten an und sinken die Fehlerzahlen. Dieser Übungseffekt verfälscht zwar aufgrund der balancierten Schalldarbietung die schallbezogene Auswertung der Leistungsdaten nicht, kann jedoch zu einer Abschwächung ihrer Effektgröße führen.

Anhand der Befragungsdaten können die Schalleffekte weiter erklärt werden. Beim KLTM scheint der negative Einfluss der Schalle mit durch ihre Lästigkeit bedingt zu sein. Der Verkehrslärm mit 70 dB(A) wird in der Zwischenbefragung sogar als lästiger eingestuft als der Sprachschall.

Die entstehende Beeinträchtigung durch den Hintergrundschall hängt im subjektiven Eindruck der Probanden wohl weiterhin mit der durch ihn gestörten Konzentration zusammen, was in der Zwischenbefragung bei den Schallbedingungen *Sprache* und *2000-70* betont wird.

Die subjektiv berichtete Leistungsmotivation bzw. Extraversion weist in diesem Test nicht auf Extremgruppen unter den Probanden hin. Aus diesem Grunde konnten auch die in Abschnitt 1.5 ausführlich berichteten Interaktionseffekte zwischen diesen Merkmalen und der Lärmwirkung vor allem auf Tests, die Kopfrechnen erfordern, nicht nachgeprüft werden (Belojevic, Slepcevic, & Jacolvjevic, 2001; Furnham & Strbac, 2002; Lohmann, 1985).

In Kapitel 1.5 wurde auch auf das Konzept der *reaktiven Anspannungssteigerung* als Erklärung der Wirkung eines lauten Verkehrsschalls auf kognitive Prozesse hingewiesen (Beckmann, 1999; Düker, 1963; Rohloff & Gollwitzer, 1999). Um die Schallwirkung im Verlauf des Experiments in diesem Sinne zu spezifizieren, wurde die Leistung der Versuchspersonen nach Schallabschnitten getrennt betrachtet. Dies bedeutet, dass die richtig gelösten Aufgaben unter den ersten, zweiten, dritten und letzten Vierteln einer jeden Schalldarbietung separat berechnet wurden. Hierbei sollte sichtbar werden, wie sich der Leistungsverlauf während der jeweils fünfzehnminütigen Beschallung darstellt. Ein Leistungseinbruch zu Beginn mit einer anschließenden Leistungssteigerung, die über das Niveau unter Ruhe hinaussteigt, aber zum Ende wieder absinkt, würde auf reaktive Anspannungssteigerung hindeuten.

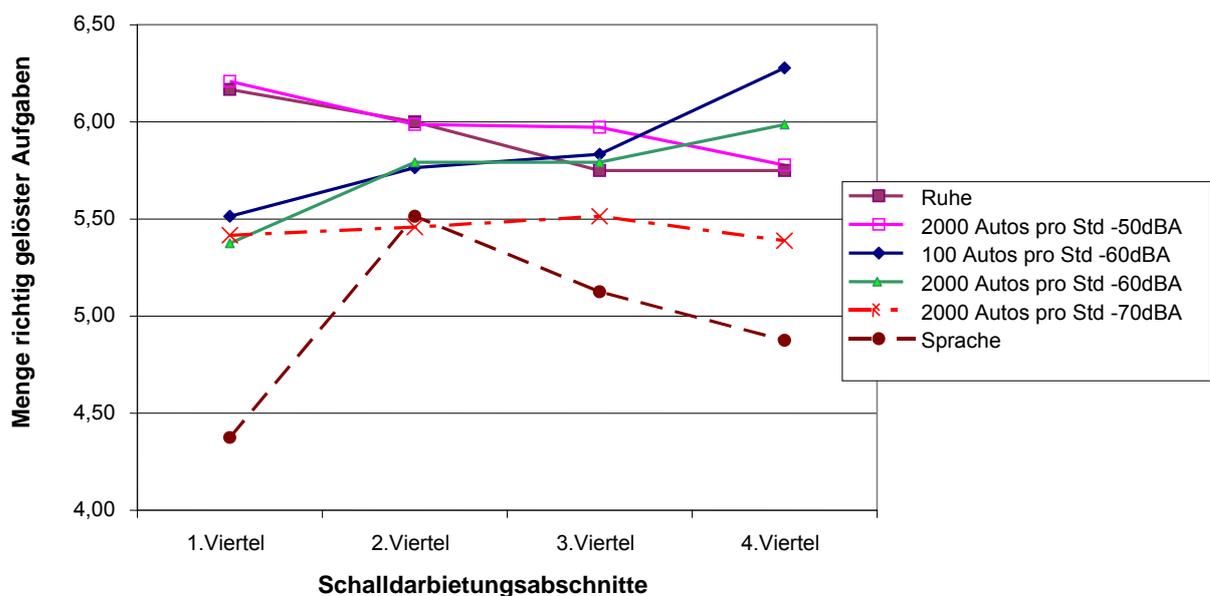


Abbildung 8: Verlauf der Leistung unter jedem Schall im KLTM: richtig gelöste Aufgaben pro Abschnittsviertel (Mittelwert über alle Probanden)

In Abbildung 8 wird deutlich, dass unter den beiden *leisen* Schallen *Ruhe* und *2000-50* ein anderer Verlauf der Leistung vorliegt als unter den beiden Schallen mit 60 dB(A). Im Gegensatz zum Abwärtstrend der Erstgenannten kommt es bei diesen im Verlauf eher zu einer Verbesserung der Leistung. Die Lösungswahrscheinlichkeit der Aufgabe unter dem Verkehrslärm mit 70 dB(A) liegt unterhalb der anderen Verlaufslinien. Die Durchführung des Tests scheint durch diesen Schall gleichmäßig negativ beeinflusst zu werden. Auffällig ist hier der Verlauf der Leistung unter dem Sprachschall. Die Versuchspersonen scheinen anfangs relativ stark durch den Schall beeinträchtigt zu sein, was kurzzeitig kompensiert werden kann, aber ab der Hälfte des Abschnitts wieder zu einem Leistungsabfall führt. Diesen Verlauf könnte man als Hinweis auf reaktive Anspannungssteigerung interpretieren. Allerdings postuliert dieses Konzept, dass die Leistung sogar über den Wert der unter Ruhe erhoben werden kann, steigt. Dies ist in den vorliegenden Daten nicht der Fall, weswegen das Konzept der reaktiven Anspannungssteigerung zur Erklärung der vorliegenden Daten als unpassend gelten muss.

Auch Hinweise auf eine Orientierungsreaktion in Folge des Schalls oder auf Habituationsprozesse lassen sich in Abbildung 8 bezogen auf die Verkehrsschalle nicht erkennen. Auch der Verlauf der Leistung unter dem Sprachschall kann hiermit nicht erklärt werden, da eine Orientierungsreaktion sofort im ersten Abschnitt stattfinden müsste, ab da müsste ein Anstieg der Leistung durch Habituation feststellbar sein. Der empirische Leistungsverlauf unter Sprachschall ist nun gerade umgekehrt, es kommt zu einer zunehmenden Verschlechterung nach einer kurzfristigen Leistungsverbesserung im mittleren Teil.

3.1.5 Der Stroop Test

Im Folgenden wird die Wirkung von Hintergrundgeräuschen auf den Stroop Test beschrieben. Auch hier wurden die in Tabelle 2 beschriebenen Verkehrsgeräusche variierend in Verkehrsdichte und Lautstärke im Vergleich zur Ruhebedingung und der Sprachschallaufnahme eingesetzt.

3.1.5.1 Aufgabenstellung

Die Aufgabe im Stroop Test besteht darin, auf die Farbe eines Farbwortes mit einem entsprechenden vorher festgelegten und geübten Tastendruck zu reagieren, was durch Inkongruenzen zwischen der Bedeutung des Wortes und seiner Farbe erschwert wird (z.B. das Wort *blau* in roter Farbe). Die Tastenbelegung ist verbal kodiert, nicht farblich (so muss auf *gelb* mit der Taste *g* geantwortet werden, auf *rot* mit *r*, auf *blau* mit *l* und auf *grün* mit *ü*). Die Worte werden auf dem Bildschirm eines Pentium I PC zentriert präsentiert (Schriftgröße 48 p, Dauer: 2000 ms, Inter Trial Intervall: 1000 ms; Fixationskreuz vor Stimulus: 500 ms Dauer, 500 ms Pause).

3.1.5.2 Versuchsplan, Probanden und Ablauf

Der Stroop Test wurde anhand eines zweifaktoriellen varianzanalytischen Designs mit Messwiederholung auf beiden Faktoren *Kongruenz* (zweifach gestuft) und *Schallbedingung* (sechsfach gestuft) durchgeführt.

An dem zweistündigen Versuch nahmen 24 freiwillige Probanden (subjektiv normales Hörvermögen; 7 männlich, 17 weiblich) im Alter zwischen 19 und 31 Jahren (Median 23 Jahre) teil, die sich auf einen Aushang hin gemeldet hatten. Sie wurden entweder im Rahmen ihres Studiums mit zwei Versuchspersonenstunden entlohnt oder erhielten 10 € als Aufwandsentschädigung.

Jede Versuchsperson absolvierte im Einzelversuch an zwei Testterminen vor einem Pentium I PC unter jedem der Schalle (vgl. Tabelle 2) 192 Aufgaben des auf ERTS⁵ programmierten Stroop Tests. Zunächst bearbeitete jede Versuchsperson beim ersten Termin nach einer schriftlichen Instruktion (Anhang E), einem Leistungsmotivationsfragebogen (LMI; Schuler & Prochaska, 2000) und einer Extraversionskala (EPI; Eggert, 1974) 24 Übungsaufgaben unter Ruhe. Anschließend wurden den Probanden dreimal 192 Aufgaben unter drei in der Reihenfolge ausbalancierten Schallen vorgegeben, nach denen sie jedes Mal in den dreiminütigen Pausen eine kurze Zwischenbefragung ausfüllten (vgl. Anhang B, auch hier ohne das Item *Störung*, wie beim KLTM). Beim zweiten Termin nach einer Woche waren je 192 Aufgaben unter den restlichen drei Schallen zu lösen. Die Schalle wurden über Kopfhörer (Sennheiser, HD 600, Sony Digital Audio Tape Deck DTC-2E700) in einer interindividuell ausbalancierten Reihenfolge präsentiert. Der Test fand in einem ruhigen Versuchsraum statt.

⁵ Dr. Beringer, BeriSoft Cooperation

3.1.5.3 Ergebnisse

Die Auswertung der Daten erfolgte zunächst anhand einer zweifaktoriellen Varianzanalyse mit Messwiederholungen über die gemessenen Fehlerraten mit den Faktoren *Schallbedingung* und *Kongruenz* (Abbildung 9).

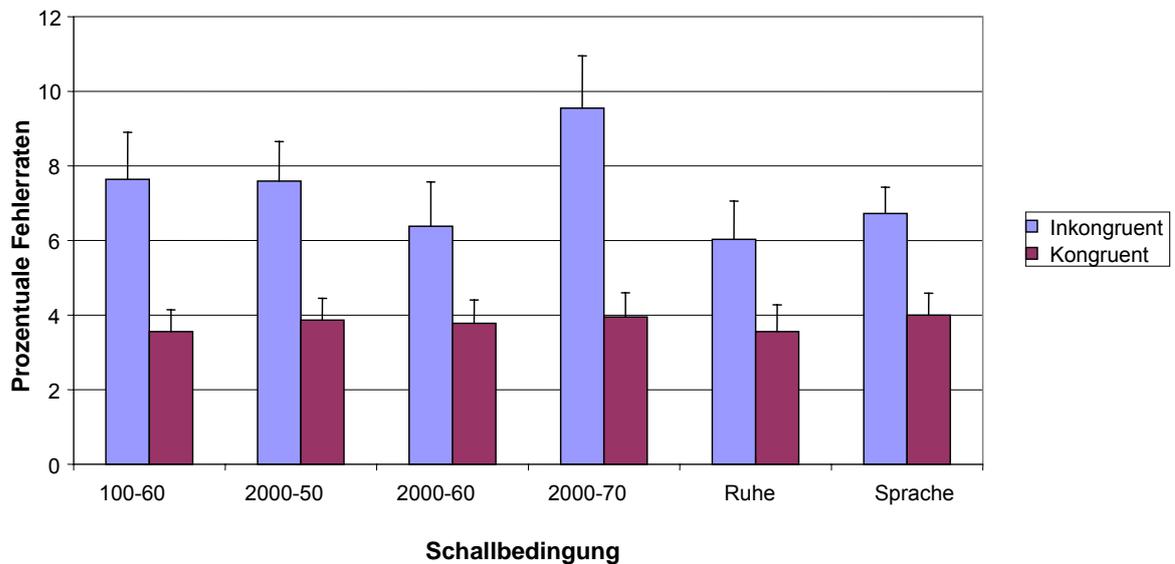


Abbildung 9: Prozentuale Fehlerraten im Stroop Test getrennt nach inkongruenten und kongruenten Items unter den sechs Schallbedingungen (Mittelwert und Standardfehler)

Es ergibt sich ein signifikanter Haupteffekt für den Faktor *Schallbedingung* ($F(5; 115)=2.5$; $p=.03$; partielles $\eta^2=0.10$), eine signifikante Wechselwirkung ($F(5; 115)=3.1$; $p=.01$; partielles $\eta^2=0.12$) und ein hochsignifikanter Haupteffekt der *Kongruenz* ($F(1; 23)=26.5$; $p<.01$; partielles $\eta^2=0.54$).

Die hierauf durchgeführten einfaktoriellen Varianzanalysen an dem nach dem Faktor *Kongruenz* aufgeteilten Datensatz zeigen nur für die inkongruenten Items einen hochsignifikanten Haupteffekt des Faktors *Schallbedingung* ($F(5; 115)=4.3$; $p<.01$; partielles $\eta^2=0.16$), jedoch nicht für die kongruenten ($F(5; 115)=0.2$; $p=.98$).

Eine Post-hoc Testung mit einer Serie von gepaarten t-Tests und Benjamini – Hochberg Korrektur (einseitige Testung aufgrund gerichteter Hypothesen, vgl. Kapitel 3.1.1 zur theoretischen Begründung) erbringt bei fünf geplanten Vergleichen einen statistisch bedeutsamen Unterschied. Geplant wurden bei inkongruenten Items Vergleiche der Leistung unter der Ruhebedingung mit der unter allen Schallen. Ein signifikanter Unterschied ergab sich zwischen dem lautesten Verkehrsschall (2000-70) und der Ruhebedingung (Ruhe: $T(23)=2.8$; $p<.01$ (BH=.01)). Durch alle anderen Schalle hingegen lässt sich statistisch keine Leistungsbeeinträchtigung im Vergleich zur Ruhebedingung nachweisen (Sprache: $T(23)=0.9$; $p=.20$ (BH=.04); 2000-60: $T(23)=0.4$; $p=.35$ (BH=.05); 2000-50:

$T(23)=2.1$; $p=.03$ ($BH=.02$); 100-60: $T(23)=1.9$; $p=.04$ ($BH=.03$)). Dies bedeutet, dass die Leistung im Stroop Test nur durch den lautesten Verkehrsschall signifikant beeinträchtigt wird.

Die Streuungswerte der Leistungsmotivation (Prozentränge (Schuler & Prochaska, 2000): $M=47.5$; $SE=6.4$) und der Extraversion (Normwerte nach Eggert, 1974: $M=14.1$; $SD=3.9$; gemessen: $M=14.2$; $SE=0.8$) weisen im Vergleich zu den Normwerten wieder nicht auf das Vorliegen von Extremgruppen hin, weswegen auf eine weiterführende Kovarianzanalyse verzichtet wurde.

Um einen Speed-Accuracy-Trade-Off auszuschließen, wurden die mittleren Lösungszeiten in einer zweifaktoriellen Varianzanalyse mit Messwiederholungen über die Faktoren *Schallbedingung* und *Kongruenz* ausgewertet (Abbildung 10).

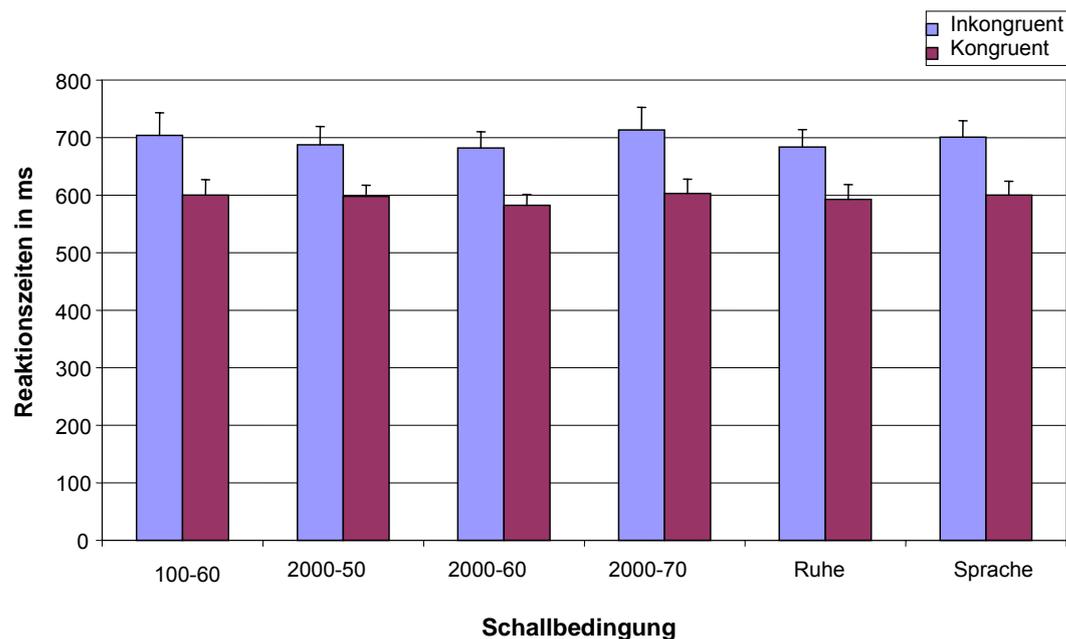


Abbildung 10: Reaktionszeiten in ms im Stroop Test unter den Schallbedingungen für kongruente und inkongruente Items (Mittelwert und Standardfehler)

Es zeigt sich nur der erwartete signifikante Haupteffekt ($F(1; 23)=60.8$; $p<.01$; partielles $\eta^2=0.73$) des Faktors *Kongruenz* – auf kongruente Items wird um ungefähr 100 ms schneller geantwortet als auf inkongruente. Der Schall hat demnach keine statistisch nachweisbare Wirkung auf die Lösungszeiten im Stroop Test (Haupteffekt *Schallbedingung*: $F(5; 115)=0.6$; $p=.71$; Interaktion: Greenhouse-Geisser Korrektur: $F(2.1; 48.1)=0.9$; $p=.48$).

Des Weiteren wurden die Daten hinsichtlich des Vorhandenseins eines Übungseffekts ausgewertet. Hierzu wurden wie oben beschrieben, die Daten nach der zeitlichen Abfolge im Experimentalverlauf zusammengefasst.

Für die prozentualen Fehlerraten ergibt sich in einer zweifaktoriellen Varianzanalyse mit den beiden Faktoren *Kongruenz* und *Versuchsabfolge* (jeweils mit Messwiederholungen) nur ein hochsignifikanter Haupteffekt für die *Kongruenz* ($F(1; 23)=26.5$; $p<.01$; partielles $\eta^2=0.54$), aber kein

Hinweis auf einen Übungseffekt ($F(5; 115)=2.3; p=.06$) und keine bedeutsame Interaktion ($F(5; 115)=0.2; p=.95$).

Stärker für einen Übungsfortschritt sprechen die Ergebnisse der Auswertung der Lösungszeiten. Eine zweifaktorielle Varianzanalyse (Faktoren: *Kongruenz* und *Versuchsabfolge*) mit Messwiederholungen zeigt einen hochsignifikanten Effekt der *Versuchsabfolge* (Greenhouse-Geisser Korrektur: $F(1.9; 73.9)=15.5; p<.01$; partielles $\eta^2=0.40$), den erwarteten hochsignifikanten Haupteffekt der *Kongruenz* ($F(1; 23)=60.3; p<.01$; partielles $\eta^2=0.72$) und eine signifikante Wechselwirkung (Greenhouse-Geisser Korrektur: $F(2.5; 56.7)=3.7; p=.02$; partielles $\eta^2=0.14$). Eine Trendanalyse dieser Daten ergibt für kongruente und inkongruente Items einen hochsignifikant fallenden Trend (Trendtest nach Page: inkongruent: $L_{emp}=2001>L_{tab}(24; 6; 1\%)=1851$; kongruent: $L_{emp}=2034>L_{tab}(24; 6; 1\%)=1851$).

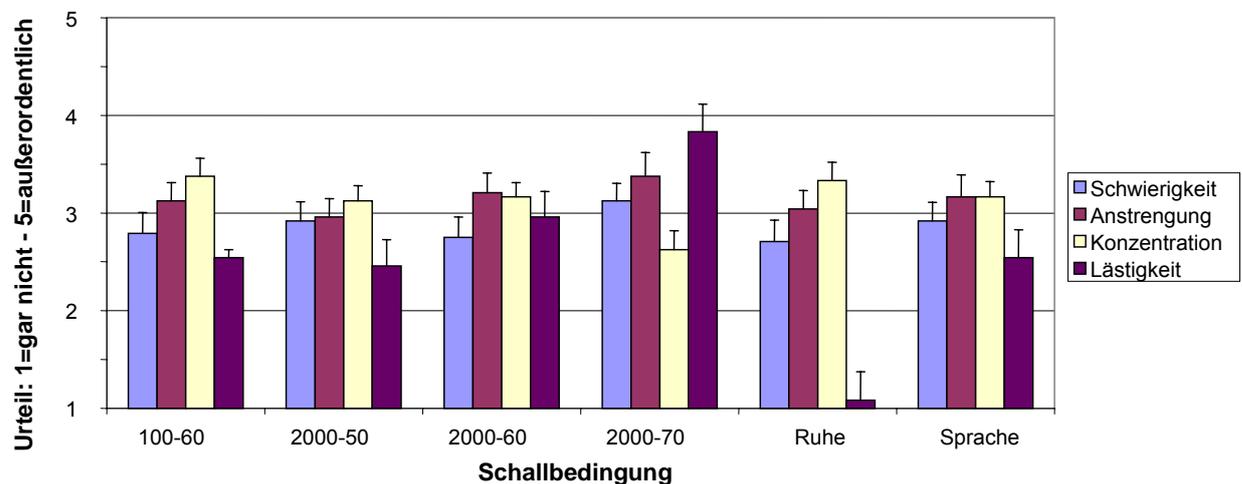


Abbildung 11: Urteile der Probanden in der Zwischenbefragung im Stroop Test über die einzelnen Schalle anhand der fünfstufigen Skalen Schwierigkeit, Anstrengung, Konzentration und Lästigkeit (Mittelwert und Standardfehler)

Um mehr Informationen über die Art der Leistungsbeeinträchtigung durch die hohe Lautstärke des Verkehrsschalls zu erhalten, wurden die Daten der Zwischenbefragung ausgewertet. Abbildung 11 verdeutlicht, dass die Probanden den Verkehrsschall mit 70 dB(A) eher negativ beurteilen. Zusätzlich zur hohen Lästigkeit sinkt offenbar die Konzentration, wie der niedrige Wert (2.6, *eher weniger möglich*) unter dem lautesten Verkehrsschall zeigt.

Eine Auswertung der einzelnen Items legt in vier einfaktoriellen Varianzanalysen statistisch bedeutsame Haupteffekte für den Faktor *Schallbedingung* bei den Items *Konzentration* ($F(5; 115)=3.4; p<.01$; partielles $\eta^2=0.13$) und *Lästigkeit* offen ($F(5; 115)=17.4; p<.01$; partielles $\eta^2=0.43$), nicht jedoch für die *Schwierigkeit* ($F(5; 115)=1.0; p=.41$) und die *Anstrengung* ($F(5; 115)=1.0; p=.44$).

Eine Post-hoc Auswertung (einseitige Testung, zur theoretischen Begründung vgl. Kapitel 3.1.1) der Urteile über die Konzentrationsmöglichkeit unter den Schallen ergibt eine signifikant höhere subjektiv eingeschätzte Konzentrationsbeeinträchtigung unter dem Schall 2000-70 als unter der Ruhebedingung ($T(23)=3.2$; $p<.01$ (BH=.01)). Die anderen Schalle beeinträchtigen die Konzentration im Vergleich zur Ruhebedingung nicht nachweisbar (Sprache: $T(23)=0.9$; $p=.19$ (BH=.04); 2000-60: $T(23)=0.9$; $p=.19$ (BH=.03); 2000-50: $T(23)=1.1$; $p=.14$ (BH=.02); 100-60: $T(23)=0.2$; $p=.42$ (BH=.05)).

Die subjektiv empfundene *Lästigkeit* der Schalle ist unter dem lautesten Verkehrsschall 2000-70 nach der Post-hoc Auswertung (siehe oben) signifikant höher unter als unter der Ruhebedingung ($T(23)=-8.3$; $p<.01$ (BH=.01)). Jedoch werden auch alle anderen Verkehrsschalle und der Sprachschall im Vergleich zur Ruhebedingung als signifikant lästiger eingeschätzt (2000-60: $T(23)=5.8$; $p<.01$ (BH=.02); 2000-50: $T(23)=4.4$; $p<.01$ (BH=.05); 100-60: $T(23)=5.0$; $p<.01$ (BH=.04); Sprache: $T(23)=5.1$; $p<.01$ (BH=.03)).

3.1.5.4 Diskussion

Die Schwierigkeit im Stroop Test besteht darin, die Konkurrenz zwischen der Benennung von Wort und Farbe zu lösen, eine Aufgabe, die – wie in Kapitel 1.7 dargelegt wurde – der zentralen Exekutive zugeschrieben wird. Hypothesenkonform führt in diesem Test der Verkehrslärm in hoher Lautstärke (2000 Durchfahrten; $L_{eq}=70$ dB(A)) bei inkongruenten Items zu einer Zunahme von Fehlreaktionen ohne Beeinflussung der Bearbeitungszeiten. Dieser Effekt tritt hypothesenkonform nur bei inkongruenten Items auf, da nur diese eine Interferenz hervorrufen. Der Sprachschall, der entsprechend dem Working Memory Modell direkten Zugang zur phonologischen Schleife hat, interferiert dort nicht mit der Aufgabe und wirkt somit wie erwartet auch nicht leistungsbeeinträchtigend.

Die störende Wirkung des lauten Hintergrundschalls scheint in engem Zusammenhang mit seiner empfundenen Lästigkeit zu stehen, was sich in den Urteilen der Probanden in der Zwischenbefragung niederschlägt. Hier schneidet der lauteste Verkehrslärm unter dem Kriterium *Lästigkeit* am schlechtesten ab. Das zweite auffallende Urteil der Zwischenbefragung betrifft die *Konzentration*, die im Vergleich zur *Schwierigkeit* des Tests und dessen verursachter *Anstrengung* durch hohe Schalldruckpegel wesentlich mehr beeinflusst zu sein scheint. Die Probanden beurteilen die Konzentration unter dem lautesten Verkehrsschall als relativ eingeschränkt.

Auch dieses Befundmuster scheint, wie sich in einer Betrachtung des Leistungsverlaufs zeigt, nicht über das Konzept der reaktiven Anspannungssteigerung (Düker, 1963) erklärbar zu sein. Eine Visualisierung des Leistungsverlaufs, anhand dessen die Wirkung der Schalle klarer werden kann, ist in Abbildung 12 dargestellt.

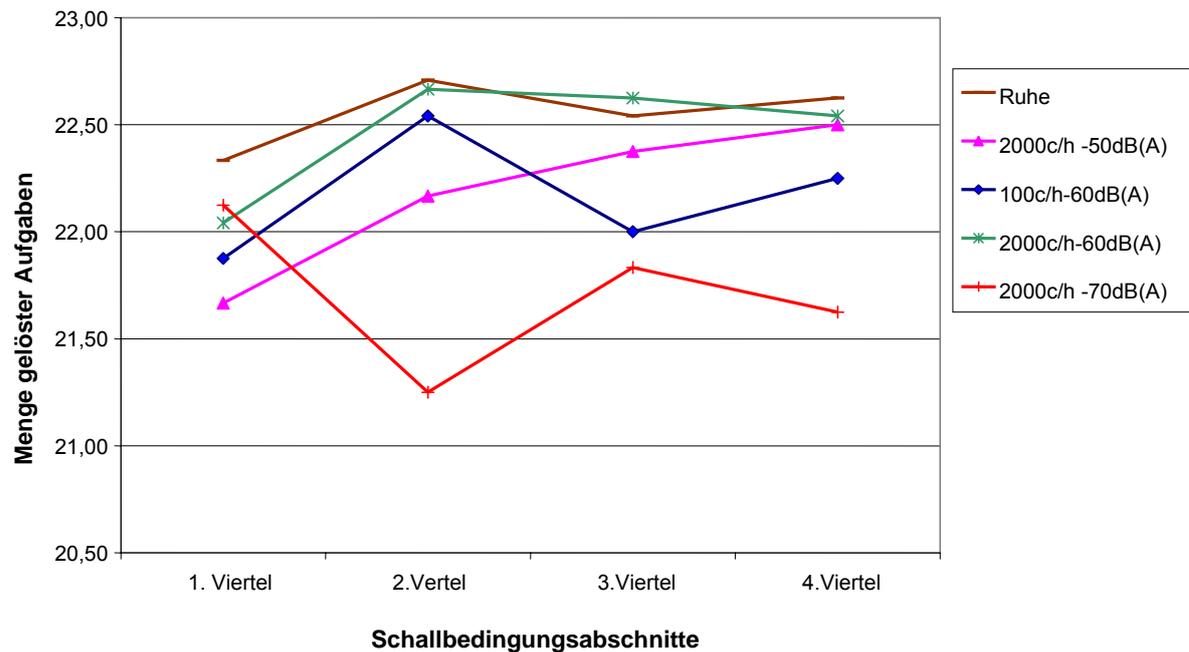


Abbildung 12: Leistungsverlauf im Stroop Test – Mittelwert gelöster Aufgaben über alle Probanden pro Viertel der Schallbedingung (nur inkongruente Items für den Vergleich der Ruhebedingung mit den Verkehrsschallbedingungen; c/h= Autos pro Stunde)

In Abbildung 12 ist die Anzahl der richtig gelösten Items nach Vierteln der Schallabschnitte abgetragen. Obwohl die absoluten Unterschiede in den Fehlerhäufigkeiten hier sehr gering sind (ein bis zwei Aufgaben) zeigt sich, dass der Leistungsverlauf unter dem lautesten Verkehrslärm bei inkongruenten Items (nur diese für den Vergleich der Verkehrsschalle mit der Ruhebedingung sind dargestellt) den anderen Verläufen beinahe entgegengesetzt ist. Die Leistung scheint vor allem im zweiten Viertel einzubrechen und sich nicht mehr merklich zu verbessern. Der Schall mit einem Mittelungspegel von 70 dB(A) scheint auf die Probanden vor allem bei längerer Dauer beeinträchtigend zu wirken. Auch hier zeigt sich folglich kein Hinweis auf eine reaktive Anspannungssteigerung. In der Argumentationstradition dieses Konzepts wäre die Verlaufslinie für den lauten Verkehrsschall genau spiegelverkehrt zu erwarten gewesen. Auch auf eine Orientierungsreaktion und anschließende Habituation lässt sich bei diesem Schall nicht rückschließen, eher noch bei seinen leiseren Varianten (2000-50 und 2000-60), unter denen es im Laufe der Exposition zu einer sukzessiven Leistungssteigerung kommt, und somit ein Habitationsprozess vorliegen könnte.

Ein Zusammenhang der subjektiv berichteten Leistungsmotivation und Extraversion mit den Leistungsmessungen ist auch in diesem Test nicht nachweisbar, da wieder keine Extremgruppenausprägungen auf diesen Variablen vorliegen.

Die leistungsbeeinträchtigende Wirkung des Verkehrsschalls mit 2000 Durchfahrten und einem Mittelungspegel von 70 dB(A) scheint somit, da sich weder für reaktive Anspannungssteigerung

noch für einen Zusammenhang von individuellen Merkmalen der Probanden (Leistungsmotivation, Extraversion) mit der Leistung empirische Evidenz ergibt, auf einer Steigerung des Arousal zu beruhen.

3.1.6 Diskussion der Ergebnisse zu Fragestellung I

Wie vermutet ließ sich eine störende Wirkung des sprachlichen Hintergrundschalls auf die Leistung in den beiden Aufgaben finden, die laut theoretischer Erklärung (Kapitel 1.3 und 1.5) die phonologische Schleife des Working Memory Modells von Baddeley (2000a) beanspruchen – die Serial Recall Aufgabe und den KLTM. Die erwartete Störwirkung des zeitlich stärker variablen Verkehrsschalls (100 Durchfahrten pro Stunde) findet sich nicht. Hier scheint das Merkmal des *Changing State* im Schallsignal nicht genügend ausgeprägt gewesen zu sein (vgl. Pegelzeitverläufe Anhang A). Somit gelang ein Nachweis des ISE auch für eine andere Aufgabe als den Serial Recall, jedoch nur für Sprachschall. Die Leistung im Stroop Test, der Anforderungen an die zentrale Exekutive stellt, wird nicht durch den Sprachschall, jedoch wie der KLTM durch den lautesten Verkehrslärm beeinträchtigt.

Somit konnten Belege für alle diesbezüglichen Hypothesen gesammelt werden. Abgeleitet aus den theoretischen Grundlagen, die in den Kapiteln 1.3 und 1.4 dargelegt wurden, war genau dieses Ergebnismuster erwartet worden.

Die erhobenen Kontrollvariablen Leistungsmotivation, Extraversion und die individuelle Lärmempfindlichkeit variieren jeweils nicht genügend um einen Extremgruppenvergleich zu erlauben. Die Ablenkung durch störende und lästige Gedanken (Näätänen, 1975) scheint nicht im Vordergrund zu stehen, insbesondere nicht bei Tests, die selektiv die phonologische Schleife beanspruchen, wie der Serial Recall. In diesem Test werden die Verkehrsschalle zwar als lästig und störend beurteilt, was aber in keinem nachweisbaren Zusammenhang zur Leistung steht. Obwohl im Stroop Test und im KLTM die Urteile über die Schalle als störend, lästig und konzentrationshinderlich mit den Leistungsdaten gut korrespondieren, kann daraus nicht auf einen kausalen Zusammenhang zwischen Kognitionen über einen Schall und der Leistung geschlossen werden. Die Probanden scheinen allerdings die Leistungsbeeinträchtigung in diesen beiden Tests realistisch einzuschätzen und dies in der Befragung wiederzugeben

Für den Schallschutz in Büro- oder Schulnähe ergeben sich bereits aus diesen Experimenten Hinweise auf eine Gefährdung einer geistig anspruchsvollen Arbeit unter lauten Verkehrsschallen, was Fragen nach Implikationen für Schallschutzmaßnahmen aufwirft. Dieses Thema soll vor allem in Kapitel 4 wieder aufgegriffen werden.

Um die gefundenen Ergebnisse bezüglich der beeinträchtigenden Wirkung des lauten Verkehrsschalls weiter zu spezifizieren, war das nächste Ziel zu überprüfen inwieweit eine Modifikation der modelltheoretisch abgeleiteten relevanten Parameter (hohe subjektive Lautheit und viel tiefe Frequenzanteile) diese Wirkung beeinflussen würde. Hieraus können anschließend ganz gezielte Maßnahmen zur Veränderung störender Schallquellen abgeleitet werden, um diese nicht auf eine globale Pegelreduktion der Schalle beschränken zu müssen.

3.2 Fragestellung II: Auswirkungen einer Pegelreduktion im tiefen Frequenzbereich bei Verkehrsschallen

Im Folgenden wird die Wirkung von Straßenverkehrslärm, der unterhalb einer Frequenz von 500 Hz um bis zu 12 dB abgesenkt wurde auf die Leistung im Stroop Test, im KLTM und im Grammatical Reasoning untersucht (vgl. Anhang G: Filterfunktion der Straßenverkehrsschalle). Eine Pegelsenkung von Verkehrsgeräuschen um bis zu sechs Dezibel im tieffrequenten Bereich (unter 500 Hz) ist eine durch Motormanagement und schalldämmende Auskleidung des Motorraums umsetzbare Reduktion der Lärmbelastung, die bei relativ unverändertem A-Pegel zu einer geringeren subjektiven Lautheit (in sone) führt (Klingenberg, 1991). Eine Pegelreduktion in dieser Größenordnung zu untersuchen, würde vermutlich aufgrund der Fehlervarianz zu keiner signifikanten Änderung der Leistungsbeeinflussung führen. Dieser geringe Unterschied könnte unter experimentellen Bedingungen von manchen Probanden noch kompensiert werden. Mit einer Pegelreduktion um zwölf Dezibel soll überprüft werden, ob eine solche Maßnahme prinzipiell sinnvoll und empfehlenswert ist. Wenn dies der Fall ist, wäre eine technisch realisierbare Reduktion der Geräusche um sechs Dezibel bereits eine erhebliche Erleichterung für exponierte Personen, und eine Klangmodifikation der Geräusche der richtige Weg, die leistungsbeeinträchtigende Wirkung von Verkehrsschallen zu modifizieren.

Die Schalle wurden in fast allen Experimenten aus Gründen der besseren Kontrollierbarkeit eventueller Störgeräusche in den Versuchsräumen über Kopfhörer dargeboten. Um einen möglichen Effekt der abgeschwächten Übertragung sehr tiefer Frequenzen über den Kopfhörer auszuschließen, wurde ein zusätzliches Experiment zur Evaluation der Unterschiede zwischen Kopfhörerdarbietung und einer Darbietung über eine Lautsprecheranlage mit Subwoofer geplant. Hierfür wird die Durchführung des KLTM mit der Lautsprecheranlage wiederholt und die Daten der beiden Experimente anhand der Kovariate der Übungsdurchgänge verglichen. Dies stellt zugleich eine Replikation der Ergebnisse dieses Tests dar.

3.2.1 Explizierung der Hypothesen

Es wird erwartet, dass sich der leistungsbeeinträchtigende Effekt des unmodifizierten Verkehrsschalls im Stroop Test und im KLTM wie in Fragestellung I in einer Erhöhung der Fehlerraten niederschlägt. Anhand einer tieffrequenten Pegelreduktion sollte der Schall sowohl durch die geringer werdenden tiefen Frequenzanteile als auch durch seine sinkende Lautheit (in sone) sein störendes Potential im Vergleich zur Ruhebedingung zum großen Teil verlieren. Von einem Speed-Accuracy-Trade-Off ist auch hier nicht auszugehen. Beide Tests werden mit der Instruktion durchgeführt, so *schnell wie möglich, aber ohne Fehler* zu arbeiten. Trotzdem sollten nachfolgend auch bei diesen beiden Tests jeweils die Reaktionszeiten ausgewertet werden, um diese Annahme zu kontrollieren. Da in Fragestellung I die erwarteten starken Übungseffekte gefunden wurden erfolgte die Darbietung der Schallbedingungen wieder in interindividuell balancierter Reihenfolge, um eine Überlagerung von Übungs- und Schalleffekten zu vermindern. Beim KLTM wird den Schallbedingungen zusätzlich ein

Übungsdurchgang in Ruhe vorangestellt, um die Verbesserung durch Übung während des Experiments geringer zu halten. Das Vorliegen von Übungseffekten soll in der Auswertung wieder durch Trendanalysen nachgeprüft werden.

Da Kopfhörer tiefe Frequenzen in einer geringeren Intensität übertragen, wird im KLTM bei dieser Darbietungsart im Vergleich zur Darbietung der Geräusche über die Lautsprecheranlage eine eventuell geringere Störwirkung erwartet (vgl. Berglund, Hassmen, & Job, 1996). Allerdings nimmt die Eindringlichkeit bei einer In-Kopf-Lokalisation der Geräusche zu, was durch den Kopfhörer begünstigt wird. Aufgrund dieser Überlegungen wird vorhergesagt, dass sich durch die geänderte Darbietungsart der Schalle kein Unterschied in der Leistung unter den verschiedenen Schallen ergibt. Diese dritte Durchführung des KLTM sollte somit eine Replikation der Ergebnisse des zuvor vorgestellten Experiments ergeben.

Das *Grammatical Reasoning* – ausführlich in Kapitel 1.6 beschrieben –, das die kognitive Funktion des schlussfolgernden Denkens testet, stellt gemäß des Working Memory Modells ähnliche Anforderungen an Subsysteme des Arbeitsgedächtnisses wie der KLTM. Somit wird bei einer Durchführung dieses Tests unter Hintergrundschall erwartet, dass sich wie beim KLTM durch den Sprachschall ein leistungsbeeinträchtigender Effekt im Vergleich zur Ruhebedingung zeigen lässt. Der unmodifizierte Verkehrsschall mit einem Mittelungspegel von 70 dB(A) sollte ebenfalls gegenüber der Ruhebedingung die Leistung behindern. Diese Leistungsbeeinträchtigung wird nach einer tieffrequenten Pegelreduktion des Verkehrsschalls nicht mehr erwartet. Es ist davon auszugehen, dass sich die Leistungsbeeinträchtigungen wie beim KLTM in einer Steigerung der Fehlerraten niederschlagen. Ein Speed-Accuracy-Trade-Off wird nicht erwartet, da auch hier die Anweisung galt *so schnell wie möglich, aber ohne Fehler zu arbeiten*. Diese Aufgabe wird wie der KLTM in gesteigerter Schwierigkeit gegenüber der Originalversion (Baddeley, 1968) vorgegeben, um die Anforderungen an das Arbeitsgedächtnis zu erhöhen. Aus diesem Grund wird ein Übungseffekt im Sinne sinkender Fehlerraten und abnehmender Reaktionszeiten über die Zeit hinweg erwartet, der in der Auswertung durch Trendtests kontrolliert werden soll. Um eine Konfundierung mit den erwarteten Schalleffekten zu vermindern werden die Schallbedingungen in interindividuell balancierter Reihenfolge vorgegeben.

Erwartet wird zudem erneut eine Kovarianz der individuellen Lärmempfindlichkeit, die als Kontrollvariable erhoben wurde (außer im Replikationsexperiment mit dem KLTM unter Lautsprecherdarbietung), mit der leistungsbeeinträchtigenden Wirkung der Hintergrundschalle. Lärmempfindliche Probanden sollten sensibler auf Hintergrundschall reagieren und somit unter den unmodifizierten Verkehrsschallbedingungen mehr Fehler begehen. Dies soll im Folgenden anhand von Kovarianzanalysen der Fehlerraten mit dieser potentiellen Moderatorvariable überprüft werden. Wie in Fragestellung I wird diese Variable allerdings nur bei vorliegenden Extremgruppen, d. h. hoher Streubreite (kontrolliert über den Standardfehler im Vergleich zu den Normwerten) als Kovariate in die Auswertung einbezogen.

Es wird zudem ein höheres Lästigkeits- und Störungsurteil über leistungsbeeinträchtigende Schalle im Vergleich zur Ruhebedingung erwartet. Die Konzentrationsbeeinträchtigung, die Anstrengung und die Aufgabenschwierigkeit sollten erneut unter leistungsbeeinträchtigenden Schallen als stärker ausgeprägt beurteilt werden als unter Ruhe (gerichtete Hypothesen). Zu diesem Zweck werden abermals in Zwischenbefragungen die subjektive Beurteilung der Lästigkeit und Störwirkung von Hintergrundschallen, sowie der Konzentrationsmöglichkeit, der Schwierigkeit der Aufgabe und der aufgebrauchten Anstrengung erhoben.

3.2.2 Der KLTM mit pegelreduzierten Verkehrsschallen

Da das erste Experiment mit dem KLTM einen signifikant leistungsbeeinträchtigenden Effekt des Verkehrslärms ab einer bestimmten Lautstärke ergeben hatte, wurde nun dieser Verkehrsschall in Frequenzzusammensetzung, Lautheit (sone) und Lautstärke (dB(A)) variiert. Die Filterung der Geräusche unter 500 Hz wurde mit der *Fast Fourier Transformation–Methode* der schnellen Faltung bei einer zeitlichen Überlappung von 50% und Zeitfenstern von 8192 Stützstellen durchgeführt⁶ (Abbildung der Filterfunktion siehe Anhang G).

Die in Tabelle 3 aufgelisteten Variationen des Schalls wurden vorgenommen, um Hinweise darauf zu erhalten, ob die tieffrequenten Anteile oder die Lautheit des Schalls die entscheidenden Faktoren der Störwirkung sind. Bei der oben beschriebenen Absenkung des Schalls unter 500 Hz um 12 dB variieren beide Parameter *Lautheit* und *Frequenzzusammensetzung* (Brummigkeit). Um eine Aussage darüber treffen zu können, welcher von beiden Parametern entscheidender bei der Störwirkung ist oder ob sich die Wirkung wie vermutet additiv zusammensetzt, wurden folgende Variationen erzeugt: Beide Schalle wurden sowohl im originalen Lautheitsverhältnis zueinander als auch in der Lautheit (in sone) des jeweils anderen Schalls dargeboten, woraus die in Tabelle 3 dargestellten vier Varianten entstehen. Es entstand folglich ein *brummiger*, lauter Schall (2000-70), ein *nicht brummiger*, leiser Schall (2000-70-12), ein *brummiger*, leiser Schall (2000-70 abgesenkt) und ein *nicht brummiger*, leiser Schall (2000-70-12 erhöht). Diese wurden mit der Software Artemis 6.0.200, Head Acoustics GmbH hergestellt. Ihre Darbietungslautstärke wurde mit einem Artificial Ear Type 4153 von Brüel & Kjaer überprüft (exemplarische Pegelzeitverläufe der Schalle 2000-70 und 2000-70-12 vgl. Anhang A). Im Folgenden werden die Bezeichnungen aus der Tabelle verwendet.

Als fünfter Schall neben der Kontrollbedingung *Ruhe* fungierte erneut die bereits verwendete zwanzigminütige Sprachaufnahme (Helmut Griem spricht Briefe von Georg Büchner; Tabelle 3).

Tabelle 3: Beschreibung der Schalle

Name	Beschreibung	Dauer	L_{eq} über 20 min in dB(A)
Ruhe			35 dB(A)
Sprache	Helmut Griem spricht Briefe von Georg Büchner	20 min	60 dB(A) – 15 sone
2000 – 70	2000 Vorbeifahrten pro Stunde	20 min	70 dB(A) – 28 sone
2000 – 70 –12	2000 Vorbeifahrten pro Stunde, unter 500 Hz um 12 dB abgesenkt	20 min	70 dB(A) – 22 sone
2000 – 70 abgesenkt	2000 Vorbeifahrten pro Stunde, insgesamt abgesenkt auf die Lautheit des –12 dB Schalls	20 min	66 dB(A) – 22 sone
2000 – 70 –12 erhöht	2000 Vorbeifahrten pro Stunde, unter 500 Hz um 12 dB abgesenkt, auf die Lautheit des Originalschalls erhöht	20 min	73 dB(A) – 28 sone

⁶ Prof. Bisping, SASS acoustic research & design GmbH

3.2.2.1 Aufgabenstellung

Die Aufgabe im KLTM wurde wie in Fragestellung I verwendet. Allerdings wurden die Rechenphasen von 15 auf 10 Minuten verkürzt, so dass der gesamte Versuch in einer Sitzung durchlaufen werden konnte.

3.2.2.2 Versuchsplan, Versuchspersonen und Versuchsablauf

Der KLTM wurde mit einem einfaktoriellen varianzanalytischen Design mit Messwiederholung auf dem sechsfach gestuften Faktor *Schallbedingung* durchgeführt.

An dem zweistündigen Versuch nahmen 18 freiwillige Probanden (2 männlich, 16 weiblich) teil, die sich auf einen Aushang hin gemeldet hatten. Ihr Alter lag zwischen 19 und 39 Jahren (Median 23 Jahre). Die meisten von ihnen waren Psychologiestudenten, die im Rahmen ihres Studiums mit zwei Versuchspersonenstunden entlohnt wurden. Die übrigen erhielten 10 € als Aufwandsentschädigung. Alle Probanden berichteten über normales Hörvermögen. Der Versuch fand in einem ruhigen Versuchsraum der Universität statt.

Die Versuchspersonen bearbeiteten nach einer schriftlichen Instruktion (vgl. Anhang D) und einem Lärmempfindlichkeitsfragebogen (NoiSeQ, Schütte & Marks, 2004) zehn Übungsaufgaben unter Ruhe. Anschließend lösten die Probanden unter jedem der in balancierter Reihenfolge dargebotenen Schalle jeweils 10 Minuten lang so viele Aufgaben, wie ihnen möglich war (Darbietung der Schalle über Kopfhörer: Sennheisser HD 600; Sony Compact Disc Player CDP-103). Nach jedem Durchgang füllten sie die bereits wiederholt beschriebene kurze Zwischenbefragung (vgl. Anhang B) aus. Der erste Durchgang erfolgte aufgrund des großen Übungsfortschritts bei der ersten Durchführung (Fragestellung I) immer unter Ruhe und wurde als Übungsdurchgang nicht in die Auswertung aufgenommen.

3.2.2.3 Ergebnisse

Eine Auswertung der Anzahl der bearbeiteten, richtig und falsch gelösten Aufgaben ergibt folgendes, in Abbildung 13 dargestelltes Bild. Eine varianzanalytische Auswertung kann für die Menge der bearbeiteten Aufgaben keinen Effekt des Faktors *Schallbedingung* belegen ($F(5; 85)=0.5$; $p=.77$), ebenso wenig für die Menge der richtig gelösten Aufgaben ($F(5; 85)=1.0$; $p=.40$) und die Anzahl der Fehler in diesem Messwiederholungsdesign ($F(5; 85)=2.2$; $p=.07$).

Im Hinblick auf einen Übungseffekt verdeutlicht eine einfaktorielle Varianzanalyse mit Messwiederholungen der bearbeiteten Aufgaben im Versuchsverlauf signifikante Unterschiede für den Faktor *Versuchsabfolge* ($F(5; 85)=16.0$; $p<.01$; partielles $\eta^2=0.48$), ebenso für die richtig gelösten Aufgaben ($F(5; 85)=6.8$; $p<.01$; partielles $\eta^2=0.29$). Nur für die Anzahl der Fehler pro Versuchsbedingung zeigt sich kein Hinweis auf eine Veränderung durch den Faktor *Versuchsabfolge* ($F(5; 85)=0.8$; $p=.57$).

Eine Trendanalyse nach Page (Begründung für die Trendanalyse vgl. Kapitel 3.2.1) im Hinblick auf einen Übungseffekt ergibt sowohl für die Anzahl der bearbeiteten als auch für die der richtig gelösten Aufgaben statistisch bedeutsame Hinweise auf einen steigenden Trend in den Daten (Menge der bearbeiteten Aufgaben: $L_{emp}=1500 > L_{tab}(18; 6; 1\%)=1401$; Menge der richtig gelösten Aufgaben: $L_{emp}=1460,5 > L_{tab}(18; 6; 1\%)=1401$) und somit auf einen Übungsfortschritt. Dies bedeutet, dass sowohl die Anzahl der bearbeiteten, als auch die der richtig gelösten Aufgaben im Versuchsverlauf zunimmt.

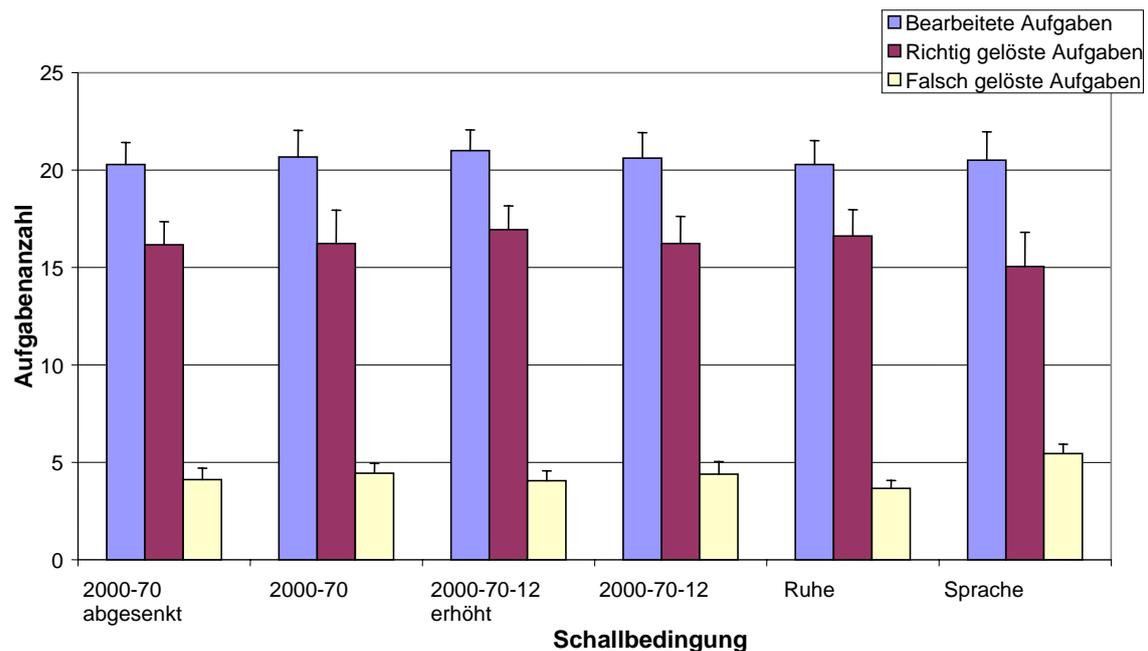


Abbildung 13: Anzahl bearbeiteter, richtig und falsch gelöster Aufgaben im KLTM (Mittelwert und Standardfehler)

Da durch unterschiedliche Grundhäufigkeiten beim Maß der prozentualen Fehlerraten, welches im Anschluss ausgewertet wurde, erneut nicht von einer metrischen Skalierung ausgegangen werden kann, wurde eine nonparametrische Auswertung mit dem Faktor *Schallbedingung* durchgeführt (Abbildung 14: zu beachten ist, dass die großen Fehlerbalken durch die Darstellung mit Median und Quartilen bedingt sind). Diese ergibt insgesamt keine signifikanten Unterschiede (Friedman Test: $N=18$; $\chi^2=6.9$; $df=5$; $p=.23$). Trotzdem wurden die aus den Hypothesen abgeleiteten, geplanten Einzelvergleiche durchgeführt. Es wurden allerdings aufgrund der nicht signifikanten Varianzanalyse nur vier Vergleiche gerechnet: die Leistung unter der Ruhebedingung im Vergleich zu den Schallen *Sprache*, *2000-70* und der Modifikation *2000-70-12* (mit Variation *beider* Parameter), und der Vergleich dieser beiden Verkehrsschalle. Da nur vier geplante Post-hoc Vergleiche durchgeführt wurden erfolgte keine Korrektur nach Benjamini-Hochberg (vgl. Roberts & Russo, 1999). Es ließen sich zwei statistisch bedeutsame bzw. knapp signifikante Unterschiede zeigen, und zwar zwischen der Leistung unter *Ruhe* und *Sprache* (Wilcoxon Test: $Z=-2.77$; $p<.01$; einseitig) sowie

zwischen *Ruhe* und dem originalen Verkehrsschall mit 70 dB(A) (2000-70) (Wilcoxon Test: $Z=-1.55$; $p=.06$; einseitig). Zwischen *Ruhe* und 2000-70-12 (Wilcoxon Test: $Z=-0.72$; $p=.24$; einseitig) und zwischen den beiden Verkehrsschallen (Wilcoxon Test: $Z=-0.50$; $p=.31$; einseitig) ließ sich kein signifikanter Unterschied nachweisen. Das Ergebnis repliziert die in Fragestellung I gefundenen Effekte. Alle anderen Vergleiche wurden aufgrund der geringen Differenz zur Ruhebedingung und des nicht signifikanten Ergebnisses der Varianzanalyse nicht durchgeführt (die Leistung unter Ruhe verglichen mit der unter den anderen Varianten der Verkehrsschalle). Auf individueller Ebene zeigt sich auch hier bei 14 von 18 Probanden wie in Fragestellung I eine Verschlechterung unter dem Sprachschall im Vergleich zur Ruhebedingung, sowie bei 13 von 18 Probanden unter dem lauten unmodifizierten Verkehrsschall (2000-70).

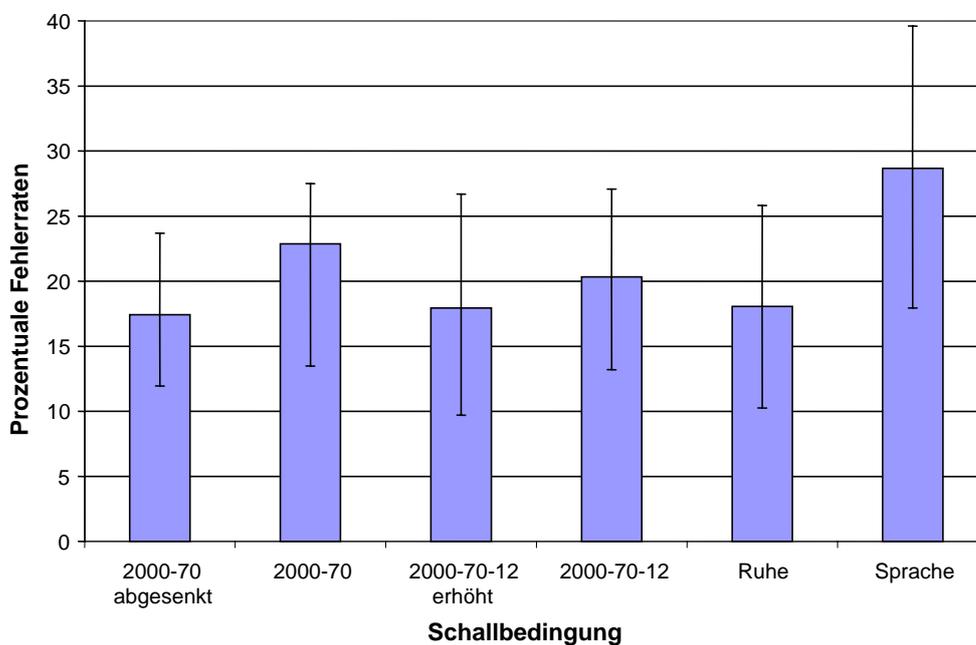


Abbildung 14: Prozentuale Fehlerraten im KLTM, Faktor *Schallbedingung* (Median sowie 1. und 3. Quartil); vgl. Tabelle 3 bzgl. der Bezeichnungen der Schalle

Die individuelle Lärmempfindlichkeit der Probanden (erhoben über den NoiSeQ: vorläufige Normwerte (Schütte & Marks, 2004): $M=43$; $SD=18$; gemessen: $M=54.1$; $SE=2.6$) streut erneut nicht genügend (im Vergleich zu den vorläufigen Normwerten), um auf das Vorliegen von Extremgruppen schließen zu können, weswegen auf eine weiterführende kovarianzanalytische Auswertung verzichtet wurde.

Ein Trendtest nach Page über die Fehlerraten unter Ruhe, dem unter 500 Hz um 12 dB abgesenkten Verkehrsschall (2000-70-12) und dem unmodifizierten Schall (2000-70) ergibt einen signifikant ansteigenden Trend ($L_{emp}=226 > L_{tab}(18; 3; 6\%)=226$). Die Leistung unter dem modifizierten Verkehrsschall liegt somit, wie in den Hypothesen postuliert, im Bereich zwischen jener unter Ruhe und der unter dem unmodifizierten Schall.

Es lassen sich zwar aufgrund der geringen Differenz keine signifikanten Unterschiede zwischen den modifizierten Verkehrsschallen und dem Originalschall nachweisen, aber auch kein signifikanter Unterschied mehr zur Ruhebedingung.

Hinweise auf einen Übungseffekt zeigen sich bei einer nonparametrischen Auswertung der prozentualen Fehler hinsichtlich eines Einflusses des Faktors *Versuchsabfolge* nicht (Friedman Test: $N=18$; $\chi^2=6.8$; $df=5$; $p=.24$).

Eine Auswertung der Lösungszeiten im KLTM verdeutlicht in einer einfaktoriellen Varianzanalyse mit Messwiederholung keine signifikanten Unterschiede durch den Faktor *Schallbedingung* ($F(5; 85)=1.2$; $p=.31$). Bei einer Auswertung gemittelt über die Abschnitte des Experiments ungeachtet der dargebotenen Schallbedingung ergeben sich jedoch Hinweise auf einen signifikanten Übungseffekt durch den Faktor *Versuchsabfolge* ($F(5; 85)=12.3$; $p<.01$; partielles $\eta^2=0.42$). Ein Trendtest nach Page (theoretische Begründung der Trendanalyse vgl. Kapitel 3.2.1) bestätigt diesen Übungseffekt durch einen signifikant fallenden Trend in den Lösungszeiten im Versuchsverlauf unabhängig von der Ausprägung des Faktors *Schallbedingung* ($L_{emp}=1551 > L_{tab}(18; 6; 1\%)=1401$).

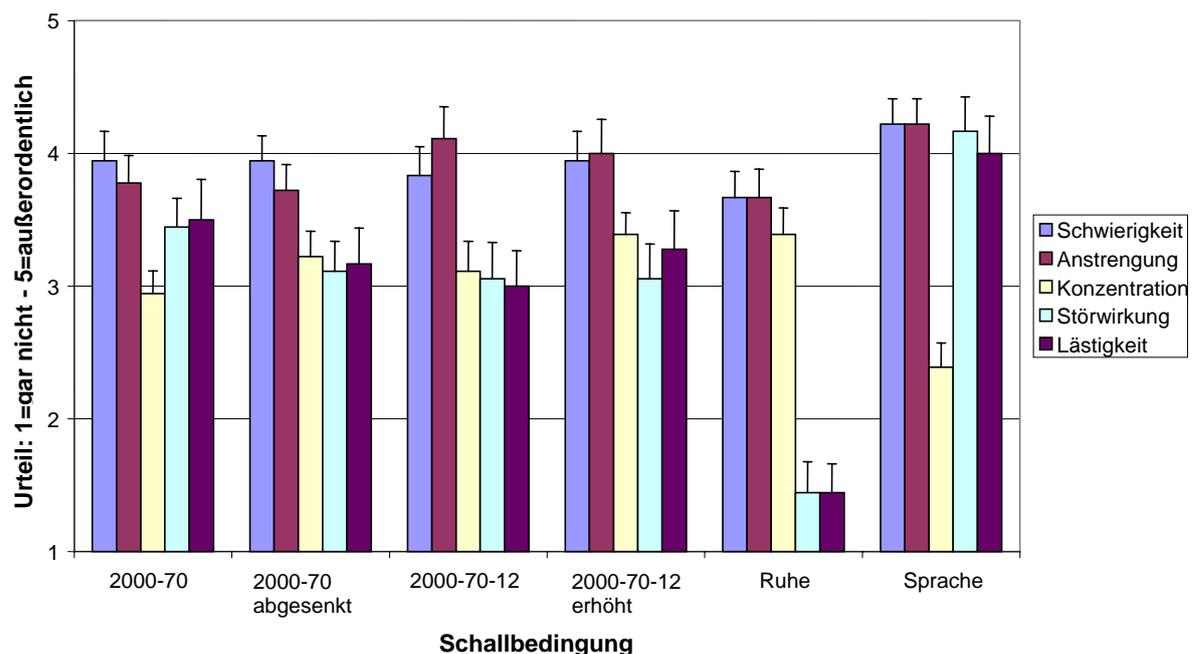


Abbildung 15: Urteile in den Zwischenbefragungen im KLTM, angetragen nach Items und Schallen, Ratingskala von gar nicht (1) bis außerordentlich (5); Mittelwert und Standardfehler

Die Auswertung der Beurteilungen (auf der oben beschriebenen fünfstufigen verbal kodierten Skala) der Items *Schwierigkeit*, *Anstrengung*, *Konzentration*, *Störung* und *Lästigkeit* in der Zwischenbefragung ergibt das in Abbildung 15 dargestellte Meinungsbild.

Eine varianzanalytische Auswertung getrennt nach den Fragestellungen legt eine je nach Ausprägung des Faktors *Schallbedingung* signifikant unterschiedliche Bewertung der Aufgabe in den

Dimensionen *Anstrengung* ($F(5; 85)=2.7$; $p=.02$; partielles $\eta^2=0.14$), *Konzentration* ($F(5; 85)=6.8$; $p<.01$; partielles $\eta^2=0.27$), *Störung* ($F(5; 85)=22.5$; $p<.01$; partielles $\eta^2=0.57$) und *Lästigkeit* nahe (Greenhouse-Geisser Korrektur: $F(2.6; 44.5)=18.7$; $p<.01$; partielles $\eta^2=0.52$), nicht jedoch für die subjektiv eingeschätzte *Schwierigkeit* (Greenhouse-Geisser Korrektur: $F(2.7; 46.1)=2.5$; $p=.08$).

Eine Post-hoc Auswertung (*false discovery rate* – Korrektur; einseitige Testung aufgrund von gerichteten Hypothesen; vgl. Kapitel 3.2.1) der Daten verdeutlicht ein signifikant höheres Urteil über die *Anstrengung* unter *Sprache* als unter *Ruhe* ($T(17)=2.6$; $p=.01$ (BH=.01)). Kein anderer Vergleich der Einschätzung von *Anstrengung* unter den Schallbedingungen mit derjenigen unter der Ruhebedingung erreicht die Signifikanz (2000-70: $T(17)=0.7$; $p=.25$ (BH=.04); 2000-70-12: $T(17)=2.2$; $p=.02$ (BH=.02); 2000-70abgesenkt: $T(17)=0.4$; $p=.36$ (BH=.05); 2000-70-12erhöht: $T(17)=1.6$; $p=.07$ (BH=.03)).

Eine entsprechende Auswertung der Urteile über die subjektive Konzentrationsmöglichkeit ergibt signifikante Unterschiede erneut nur zwischen *Sprache* und *Ruhe* ($T(17)=-4.1$; $p<.01$ (BH=.01)). Die Beurteilung, dass die Konzentration unter der Ruhebedingung besser als unter dem unmodifizierten Verkehrsschall 2000-70 sei, verfehlt durch die Korrektur knapp die Signifikanz ($T(17)=1.8$; $p=.04$ (BH=.02)). Die Konzentration scheint den Probanden unter Ruhe besser möglich zu sein. Die Beurteilung der Konzentrationsmöglichkeit unter den drei modifizierten Verkehrsschallen unterscheidet sich nicht signifikant von jener unter der Ruhebedingung (2000-70-12: $T(17)=-1.1$; $p=.15$ (BH=.03); 2000-70abgesenkt: $T(17)=-0.6$; $p=.26$ (BH=.04); 2000-70-12erhöht: $T(17)=0.0$; $p=.99$ (BH=.05)).

Die Einschätzung der Störwirkung der Hintergrundschalle ist in Einzelvergleichen signifikant in ihrem Ausmaß zwischen Ruhe und allen Schallen verschieden (Sprache: $T(17)=-7.9$; $p<.01$ (BH=.008); 2000-70: $T(17)=-5.8$; $p<.01$ (BH=.017); 2000-70 abgesenkt: $T(17)=-5.7$; $p<.01$ (BH=.025); 2000-70-12: $T(17)=-4.6$; $p<.01$ (BH=.042); 2000-70-12 erhöht: $T(17)=4.6$; $p<.01$ (BH=.03)), sowie zwischen den Schallen 2000-70 und 2000-70-12 ($T(17)=1.8$; $p=.05$ (BH=.05)). Alle Schalle werden als signifikant störender eingeschätzt als die Ruhebedingung. Der originale Verkehrsschall wird zudem als störender beurteilt als seine Absenkung (um 12 dB unter 500 Hz).

Für die subjektiv beurteilte Lästigkeit der Schalle ergibt sich ein ähnliches Bild – auch hier zeigt sich, dass alle Schalle als signifikant lästiger eingeschätzt werden als die Ruhebedingung (Sprache: $T(17)=8.7$; $p<.01$ (BH=.008); 2000-70: $T(17)=5.4$; $p<.01$ (BH=.017); 2000-70 abgesenkt: $T(17)=5.1$; $p<.01$ (BH=.025); 2000-70-12: $T(17)=4.4$; $p<.01$ (BH=.042); 2000-70-12 erhöht: $T(17)=4.6$; $p<.01$ (BH=.03)) sowie der originale Verkehrsschall 2000-70 als lästiger als der modifizierte 2000-70-12 Schall ($T(17)=2.7$; $p=.02$ (BH=.05)).

3.2.2.4 Diskussion

In dieser Durchführung des KLTM unter Hintergrundschall ließen sich die Ergebnisse der Fragestellung I replizieren. Auch diesmal zeigte sich eine Leistungsver schlechterung gegenüber der Ruhebedingung in einem statistisch bedeutsamen Rahmen unter dem Sprachschall. Hier liegt nun auch ein signifikanter Unterschied zwischen dem lautesten Verkehrslärm mit einem Mittelungspegel

von 70 dB(A) bzw. 28 sone und der Ruhebedingung vor, nachdem dieser Unterschied in Fragestellung I nur im Vergleich zum leisesten Verkehrslärm ($L_{eq}=50$ dB(A)) nachweisbar war.

Der im tieffrequenten Bereich pegelreduzierte Verkehrsschall führt wie vermutet zu keiner signifikanten Leistungsverschlechterung gegenüber der Ruhebedingung. Ob dies allerdings ein Effekt der gesunkenen Lautheit oder der geringeren tieffrequenten Anteile ist, lässt sich aufgrund der geringen Unterschiede zwischen den Fehlerraten unter diesen Bedingungen nicht im Detail entscheiden. Es ist davon auszugehen, dass beide Variablen – Lautheit des Schalls und Frequenzzusammensetzung (vgl. Berglund et al., 1996; LeVere et al., 1974) – in diesem Test in Kombination zu einer Reduktion der leistungsbeeinträchtigenden Wirkung beitragen.

Eine Reduktion der Geräusche in den Parametern Lautheit und tieffrequente Anteile, wie die hier vorgenommene Pegelabsenkung um 12 dB im Bereich unter 500 Hz, ist somit eine Erfolg versprechende Maßnahme, um diese Beeinträchtigung zu reduzieren.

Trotz des vorausgehenden Übungsdurchgangs kommt es auch bei dieser Durchführung des KLTM wieder zu einem signifikanten Übungseffekt. Die Wichtigkeit dieses Ruhedurchgangs muss damit nochmals betont werden, um die Gefahr einer Konfundierung der Schalleffekte mit der Verbesserung durch Übung zu minimieren.

In der Zwischenbefragung zeigt sich ein ähnliches Bild wie bei der Auswertung der Leistungsdaten: die Lästigkeit und Störwirkung der Schalle *Verkehr original* und *tieffrequent abgesenkt* unterscheiden sich statistisch nachweisbar. Ohne tieffrequente Absenkung wird der Schall als störender und lästiger eingeschätzt. Ebenso werden der Sprachschall und der originale Verkehrsschall im Vergleich zur Ruhebedingung als störender und lästiger beurteilt. Weiterhin interessant in diesem Zusammenhang ist das Urteil der Probanden über die mögliche Konzentration in den einzelnen Bedingungen, da diese unter der Hintergrundsprachbedingung und der Präsentation des unmodifizierten Verkehrsschalls als geringer eingestuft wird als unter Ruhe.

Es gibt keine Hinweise darauf, dass die individuelle Lärmempfindlichkeit für die Schalleffekte als Erklärung dienen kann, da sich erneut keine Extremgruppen bilden lassen.

Auch hier spricht das Befundmuster gegen die Erklärung der Leistungseinbußen durch den lauten Verkehrslärm über das Konzept der *reaktiven Anspannungssteigerung* (Düker, 1963). Eine grafische Darstellung der Anzahl der gelösten Aufgaben im Verlauf jedes Durchgangs (Abbildung 16) gibt keine Hinweise auf eine besondere Wirkung einer der Schalle im Zeitverlauf und bestätigt nochmals die oben berichtete Auswertung. Auch auf eine Orientierungsreaktion oder Habituationsprozesse lässt sich anhand des Leistungsverlaufs nicht schließen. In Abbildung 16 ist nicht die prozentuale Fehlerrate, sondern die Anzahl gelöster Aufgaben dargestellt.

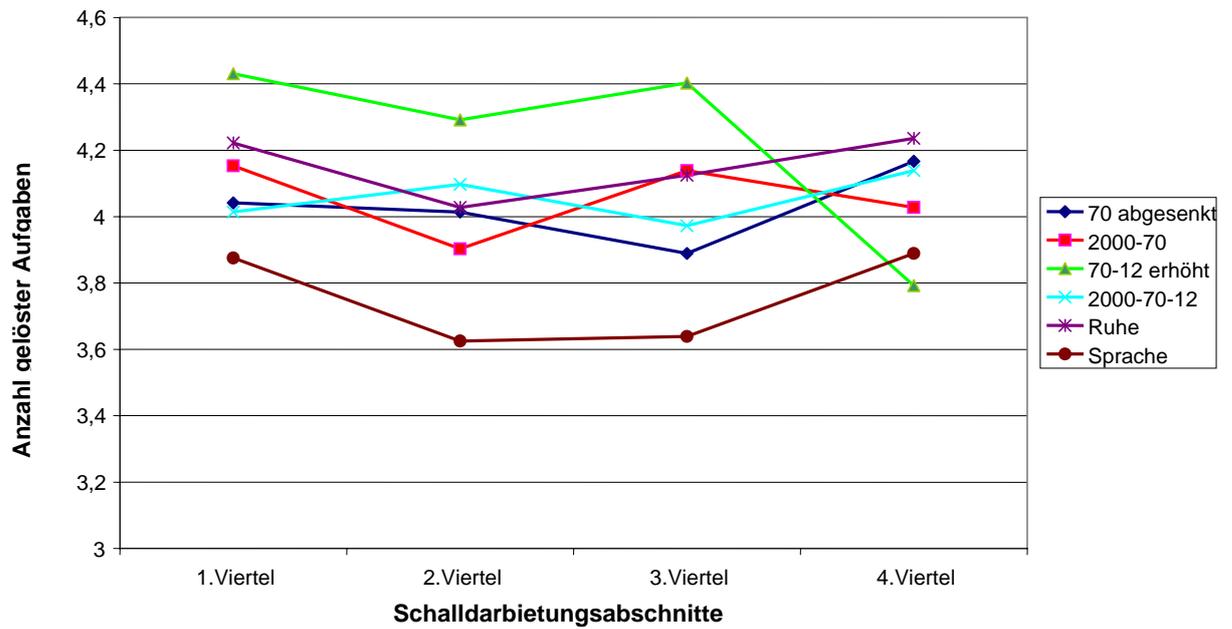


Abbildung 16: Leistung im KLTM: Verlauf der Anzahl gelöster Aufgaben pro Schall (Mittelwert über alle Probanden)

3.2.3 KLTM mit Darbietung der Hintergrundschalle über Lautsprecher und Subwoofer

Um den Einfluss des Schalldarbietungsmodus zu überprüfen, wurde der in Fragestellung I und oben beschriebene KLTM (Experimentalplan wie oben, Kapitel 3.2.2) mit einer geringeren Anzahl an Schallen (nur Ruhe, Sprache, 2000-70 und 2000-70-12, vgl. Tabelle 3) unter Einsatz einer Lautsprecheranlage (in Standardstereoanstellung: Genelec 1030A BI - Amplified Monitoring System) mit Subwoofer (Genelec 7070A Active Subwoofer) wiederholt. Bei der überwiegenden Anzahl der durchgeführten Experimente wurden die Hintergrundschalle aufgrund der besseren Kontrollierbarkeit von Störgeräuschen über Kopfhörer dargeboten. Da allerdings durch den Kopfhörer die tiefen Frequenzen mit einer geringeren Gewichtung übertragen werden, sollte geprüft werden, ob die Daten mit Lautsprecherdarbietung vergleichbar ausfallen.

3.2.3.1 Aufgabenstellung

Die Aufgabenstellung entsprach ganz der in Kapitel 3.2.2 beschriebenen. Unter jeder der Schallbedingungen mussten die Probanden innerhalb von zehn Minuten so viele Aufgaben wie möglich lösen.

3.2.3.2 Versuchsplanung, Probanden und Ablauf

Der KLTM wurde wie oben anhand eines einfaktoriellen varianzanalytischen Plans mit Messwiederholung auf dem Faktor *Schallbedingung* durchgeführt.

An diesem eineinhalbstündigen Experiment nahmen 20 Probanden teil, von denen 6 männlich und 14 weiblich waren und deren Alter zwischen 20 und 39 Jahren (Median 22.5 Jahre) lag. Sie wurden im Rahmen ihres Psychologiestudiums mit 1.5 Versuchspersonenstunden oder 7.50 € entlohnt. Alle meldeten sich auf einen Aushang hin und gaben an, über normales Hörvermögen zu verfügen.

Die Lautstärke der Schalle wurde mit einem Brüel & Kjaer Schallpegelmessgerät (Type 2231) kontrolliert; die Testdurchführung erfolgte über einen Pentium IV PC in einem ruhigen Versuchsraum der Katholischen Universität Eichstätt-Ingolstadt.

Nach einer schriftlichen Instruktion (vgl. Anhang D) absolvierten die Probanden zunächst zehn Übungsaufgaben unter Ruhe. Anschließend lösten sie unter jeder der Schallbedingungen jeweils zehn Minuten lang so viele Aufgaben wie möglich. Die Schalle wurden in interindividuell balancierter Reihenfolge über die beschriebene Lautsprecheranlage präsentiert. Der erste Durchgang fand immer in Ruhe statt und zählte als Übungsdurchgang; in den einminütigen Pausen füllten die Probanden jeweils die kurze Zwischenbefragung (vgl. Anhang B) aus.

3.2.3.3 Ergebnisse

Eine varianzanalytische Auswertung der Anzahl an bearbeiteten Aufgaben mit dem Faktor *Schallbedingung* ergibt keinen statistisch bedeutsamen Haupteffekt ($F(3; 57)=2.1$; $p=.11$), ebenso eine gleichsinnige Auswertung der Anzahl an falsch gelösten Aufgaben pro Schallbedingung ($F(3; 57)=1.8$; $p=.15$; Abbildung 17). Die Anzahl der richtig gelösten Aufgaben jedoch unterscheidet sich signifikant je nach *Schallbedingung* ($F(3; 57)=5.1$; $p<.01$; partielles $\eta^2=0.21$). Eine Post-hoc Auswertung (false discovery rate; einseitige Testung: zur theoretischen Begründung vgl. Kapitel 3.2.1) gibt Hinweise auf signifikant höhere Anzahlen richtig gelöster Aufgaben unter der Ruhebedingung als unter Sprachhintergrundschall ($T(19)=3.6$; $p<.01$ (BH=.01)) sowie unter dem unmodifizierten Verkehrsschall 2000-70 ($T(19)=3.7$; $p<.01$ (BH=.03)) und dem 2000-70-12 Schall ($T(19)=2.8$; $p<.01$ (BH=.04)). Der Vergleich der Verkehrsschalle miteinander jedoch verfehlt die Signifikanz ($T(19)=0.6$; $p=.29$ (BH=.05)).

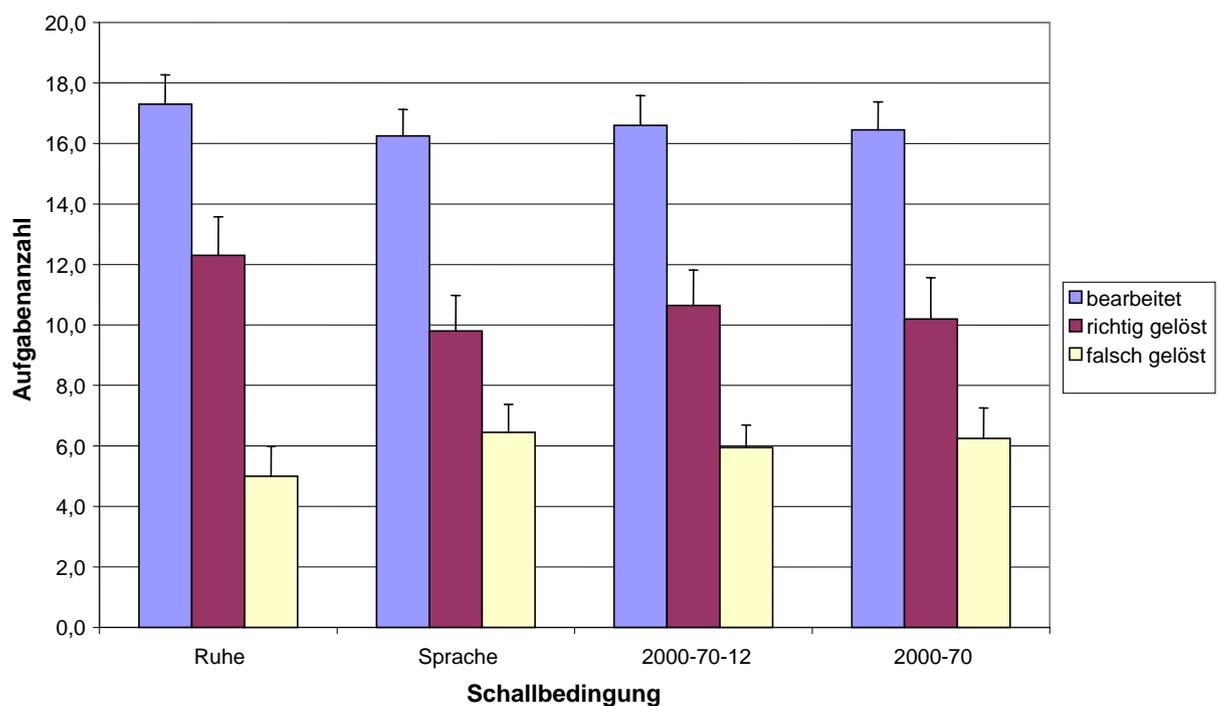


Abbildung 17: Anzahl bearbeiteter, richtig und falsch gelöster Aufgaben unter den Schallbedingungen im KLTM (Mittelwert und Standardfehler)

Eine einfaktorielle varianzanalytische Auswertung der Menge an bearbeiteten Aufgaben über die Versuchsabschnitte ergibt einen signifikanten Effekt der *Versuchsabfolge* ($F(3; 57)=11.2$; $p<.01$; partielles $\eta^2=0.37$); in einer Trendanalyse nach Page (zur theoretischen Begründung vgl. Kapitel 3.2.1) zeigt sich ein signifikant steigender Trend, was auf einen Übungseffekt hinweist ($L_{emp}=555 > L_{tab}(20; 4; 1\%)=531$). Im Verlauf des Versuches wurden somit zunehmend mehr Aufgaben pro Schall bearbeitet.

Eine gleichsinnige Auswertung der Anzahl richtig und falsch gelöster Aufgaben hingegen legt keine signifikanten Effekte nahe (Richtig: $F(3; 57)=1.8$; $p=.16$; Falsch: $F(3; 57)=0.5$; $p=.67$).

Eine nonparametrische Auswertung (aufgrund ungleicher Grundhäufigkeiten, vgl. Kapitel 3.2.2) der prozentualen Fehlerraten unter dem Faktor *Schallbedingung* verdeutlicht im Friedman Test einen signifikanten Effekt ($N=20$; $df=3$; $\chi^2=10.8$; $p=.01$; Abbildung 18). In einer nonparametrischen Post-hoc Auswertung (Wilcoxon Tests, einseitige Testung: zur Begründung vgl. Kapitel 3.2.1) zeigt sich, dass die Unterschiede in der geringeren Fehlerrate unter *Ruhe* als unter *Sprache* ($Z=2.6$; $p<.01$ ($BH=.03$)) und unter *2000-70* ($Z=2.9$; $p<.01$ ($BH=.01$)) liegen. Die Leistung unter dem Schall *2000-70-12* unterscheidet sich nicht signifikant von *Ruhe* ($Z=1.7$; $p=.05$ ($BH=.04$)), ebenso wenig ist ein Unterschied zwischen den beiden Verkehrsschällen nachweisbar ($Z=0.7$; $p=.25$ ($BH=.05$)). Auf individueller Ebene betrachtet zeigt sich, dass 16 von 20 Probanden unter dem Sprachschall mehr Fehler unterliefen als unter der Ruhebedingung. Der lauteste Verkehrsschall führt in diesem Vergleich bei 14 von 20 Probanden zu einer Erhöhung der Fehlerraten.

Eine nonparametrische Auswertung der prozentualen Fehlerraten mit dem Faktor *Versuchsabfolge* erbringt keine Hinweise auf einen Übungseffekt ($N=20$; $df=3$; $\chi^2=1.1$; $p=.77$).

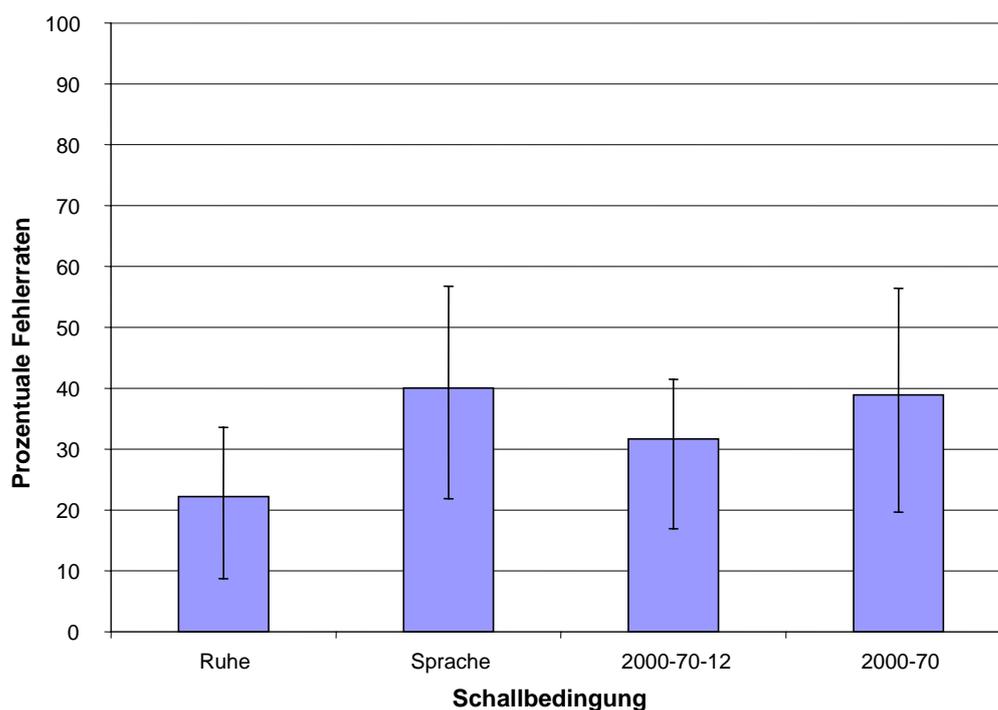


Abbildung 18: Prozentuale Fehlerraten im KLTM unter den Schallbedingungen (Median sowie 1. und 3. Quartil)

Eine varianzanalytische Auswertung der mittleren Lösungszeiten verdeutlicht keinen signifikanten Einfluss der *Schallbedingung* ($F(3; 57)=0.5$; $p=.65$). Gemittelt über die Abschnitte des Experiments ausgewertet, ergibt sich allerdings ein signifikanter Haupteffekt der *Versuchsabfolge* und

somit ein Hinweis auf einen Übungseffekt ($F(3; 57)=8.9$; $p<.01$; partielles $\eta^2=0.32$). Eine Trendanalyse nach Page legt einen signifikanten fallenden Trend ($L_{emp}=557>L_{tab}(20; 4; 1\%)=531$) und somit einen Übungseffekt nahe. Im Versuchsverlauf wurde von den Probanden somit immer weniger Zeit für die Lösung einzelner Aufgaben benötigt.

Um die Leistungen der Probanden unter der Kopfhörerdarbietung (Kapitel 3.2.2: im Folgenden bezeichnet als: *KLTM(Kopfhörer)*) und der Schallpräsentation über die Lautsprecheranlage mit Subwoofer (im Folgenden bezeichnet als: *KLTM(Subwoofer)*) zu vergleichen, wurde die Leistung der Probandengruppen im Übungsdurchgang unter der Ruhebedingung als Kovariate in die Rechnung eingeführt (vgl. Anhang I, Tabelle 11 für Mittelwerte und Standardfehler dieser Leistungsdaten). Verglichen werden sollten die Gesamtzahl der bearbeiteten und der gelösten Aufgaben, da sich aus diesen beiden Maßen auch das oben bereits häufiger berichtete Maß der prozentualen Fehlerraten zusammensetzt. Die Leistung im Übungsdurchgang korreliert in beiden Gruppen signifikant mit den Ergebnissen unter den Testschallbedingungen (es wurden nur die gemeinsam verwendeten Schallbedingungen *Ruhe*, *Sprache*, *2000-70* und *2000-70-12* verglichen: Tabelle 4).

Tabelle 4: Korrelationen (und Signifikanz) der Leistung im Übungsdurchgang mit der unter den Schallbedingungen im *KLTM(Kopfhörer)* und *KLTM(Subwoofer)* für die Anzahl der bearbeiteten und richtig gelösten Aufgaben

Korrelation des Übungsdurchgangs	KLTM(Kopfhörer) Bearbeitet	KLTM(Subwoofer) Bearbeitet	KLTM(Kopfhörer) Richtig gelöst	KLTM(Subwoofer) Richtig gelöst
Ruhe	r=.92 p<.01	r=.87 p<.01	r=.82 p<.01	r=.81 p<.01
Sprache	r=.92 p<.01	r=.82 p<.01	r=.88 p<.01	r=.78 p<.01
2000-70-12	r=.88 p<.01	r=.86 p<.01	r=.78 p<.01	r=.73 p<.01
2000-70	r=.91 p<.01	r=.88 p<.01	r=.88 p<.01	r=.86 p<.01

In Abbildung 19 sind die prozentualen Fehlerraten unter den vier Schallen sowie dem Übungsdurchgang unter *Ruhe* für den *KLTM(Kopfhörer)* und den *KLTM(Subwoofer)* dargestellt (da hier beide anschließend ausgewerteten Maße einfließen). Eine zweifaktorielle kovarianzanalytische Auswertung der Anzahl bearbeiteter Aufgaben unter den Schallen mit dem *Between-Faktor* der *Darbietungsgruppe* und der Kovariate der *Leistung im Übungsdurchgang* ergibt keinen signifikanten Haupteffekt der *Schallbedingung* ($F(3; 105)=0.2$; $p=.87$) aber auch keinen signifikanten Haupteffekt der *Darbietungsgruppe* ($F(1; 35)=0.6$; $p=.46$) sowie keine statistisch nachweisbare Interaktion ($F(3; 105)=1.0$; $p=.39$). Die Kovariate der *Leistung im Übungsdurchgang* und somit die *Gruppengüte* hingegen zeigt sich als statistisch bedeutsam ($F(1; 35)=184.3$; $p<.01$; partielles $\eta^2=0.84$), ihre Interaktion mit dem Faktor *Schallbedingung* jedoch nicht ($F(3; 105)=0.1$; $p=.94$).

Eine gleichsinnige Auswertung der Anzahl an richtig gelösten Aufgaben verdeutlicht, wie oben bei den Einzelauswertungen, einen signifikanten Haupteffekt der *Schallbedingung* ($F(3; 105)=4.7$; $p<.01$; partielles $\eta^2=0.12$), jedoch wieder keinen der *Darbietungsgruppen* ($F(1; 35)=2.1$; $p=.16$) und

keine statistisch nachweisbare Interaktion mit der *Schallbedingung* ($F(3; 105)=0.7; p=.56$). Die *Leistung im Übungsdurchgang* als Baseline der Leistung erweist sich auch hier als eine statistisch bedeutsame Kovariate ($F(1; 35)=110.5; p<.01; \text{partielles } \eta^2=0.76$), die auch bedeutsam mit der *Schallbedingung* interagiert ($F(3; 105)=3.0; p=.03; \text{partielles } \eta^2=0.08$).

Diese Ergebnisse verdeutlichen, was auch in Abbildung 19 in der Zusammenfassung der beiden Befunde zu sehen ist. Es scheint einen Gruppenunterschied zu geben, der allerdings nicht mit der Schallwirkung interagiert, sondern zu einem gleichmäßigen Leistungsunterschied unter allen Schallbedingungen führt. Es zeigt sich folglich kein nachweisbarer Zusammenhang der Leistung zu den beiden Darbietungsbedingungen der Schalle. Ein wichtiger Einflussfaktor scheint hingegen die Baseline der Leistungsfähigkeit der Probanden zu sein, die – erfasst über den Übungsdurchgang – zu divergenten Leistungen und im Fall der Anzahl richtig gelöster Aufgaben auch zu einer unterschiedlichen leistungsbeeinflussenden Wirkung der Schallbedingungen zu führen scheint.

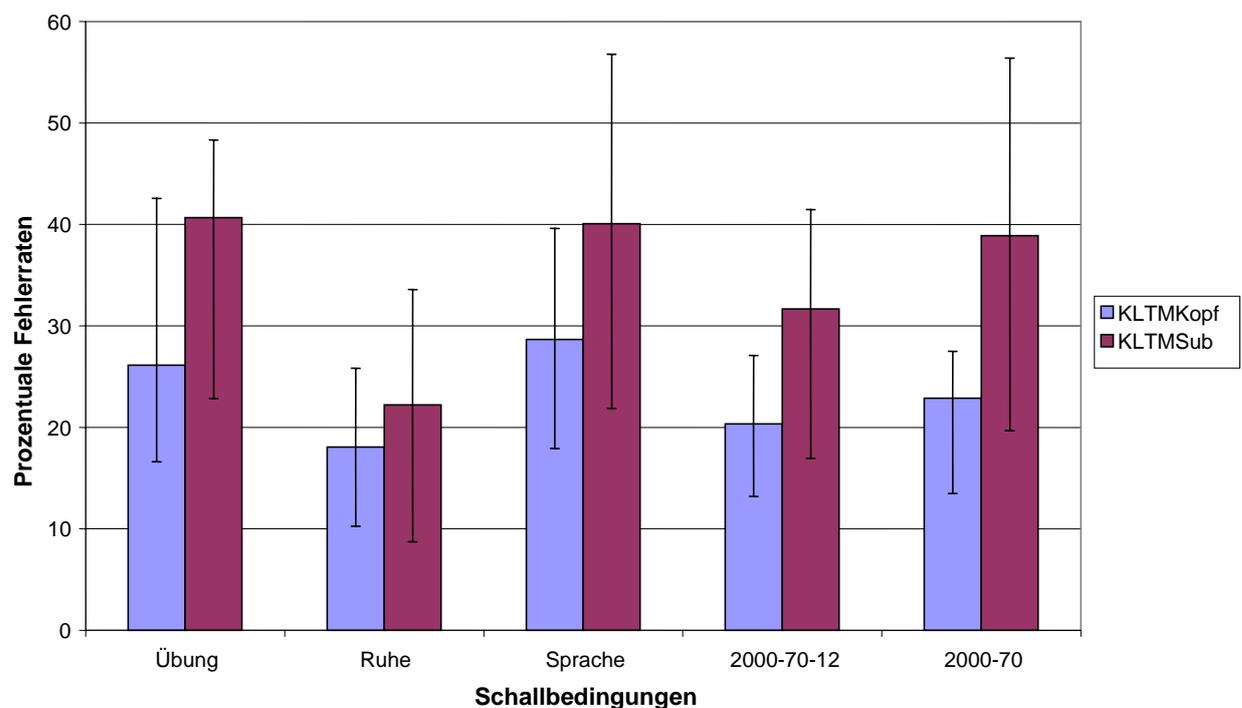


Abbildung 19: Vergleich *KLTM(Kopfhörer)* und *KLTM(Subwoofer)*: Fehlerraten in Prozent unter den Schallbedingungen und im Übungsdurchgang (Median sowie 1. und 3. Quartil; KLTMKopf = KLTM(Kopfhörer); KLTMSub = KLTM(Subwoofer))

Eine Auswertung der Zwischenbefragungsdaten (Abbildung 20) zeigt ein ähnliches Bild wie beim KLTM in Kapitel 3.2.2.

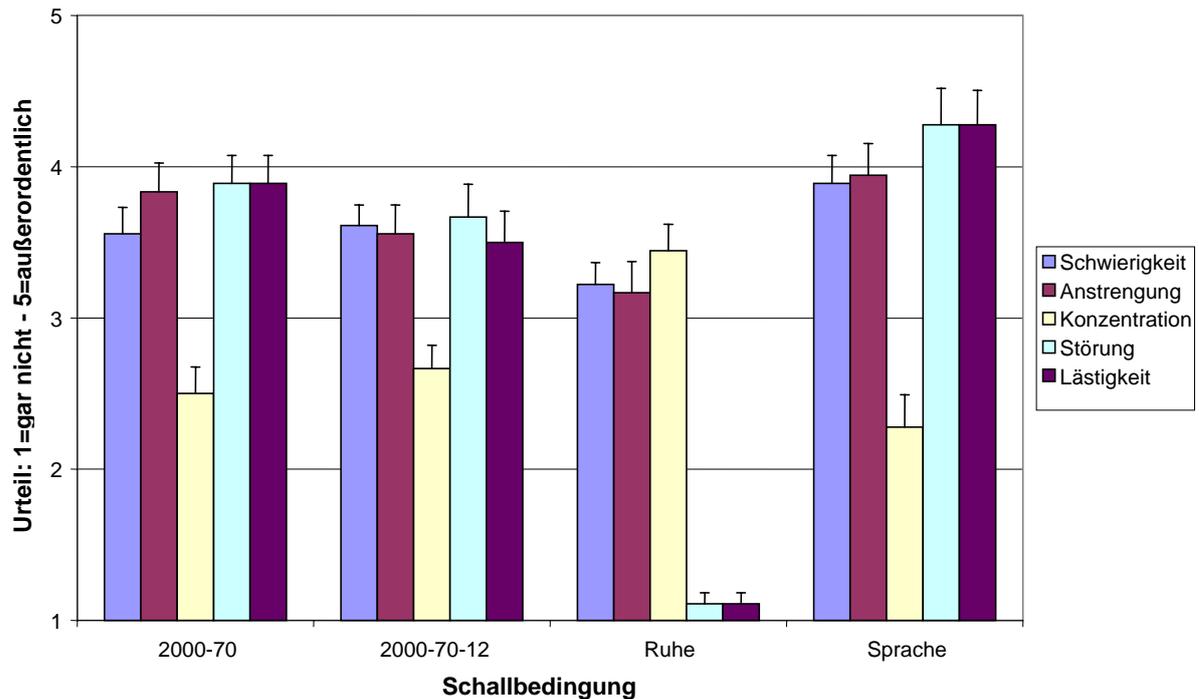


Abbildung 20: Ratingurteile in der Zwischenbefragung im KLTM (Mittelwert und Standardfehler), von 1=gar nicht bis 5=außerordentlich

Eine einfaktorielle varianzanalytische Auswertung verdeutlicht einen signifikanten Schalleffekt auf die empfundene Schwierigkeit der Aufgabe ($F(3; 57)=3.4$; $p=.02$; partielles $\eta^2=0.15$). In einer Post-hoc Auswertung (im Folgenden immer einseitig getestet, zur theoretischen Begründung vgl. Kapitel 3.2.1) zeigen sich signifikante Unterschiede zwischen der Bewertung der Ruhebedingung und allen anderen Bedingungen. Unter Ruhe wurde die Aufgabe als am wenigsten schwierig empfunden (Sprache: $T(19)=-2.6$; $p<.01$ (BH=.02); 2000-70: $T(19)=-1.8$; $p=.04$ (BH=.05); 2000-70-12: $T(19)=-2.3$; $p=.02$ (BH=.03)).

Auch die Frage nach der *Anstrengung* zeigt sich signifikant durch die *Schallbedingung* beeinflusst ($F(3; 57)=5.0$; $p<.01$; partielles $\eta^2=0.21$). Alle Schallbedingungen werden in diesem Test als anstrengender beurteilt als die Ruhebedingung (Sprache: $T(19)=-3.4$; $p<.01$ (BH=.02); 2000-70: $T(19)=-2.2$; $p=.02$ (BH=.03); 2000-70-12: $T(19)=-2.0$; $p=.03$ (BH=.05)).

Die subjektiv mögliche *Konzentration* scheint ebenfalls bedeutsam durch die *Schallbedingung* beeinflusst zu werden ($F(3; 57)=9.2$; $p<.01$; partielles $\eta^2=0.33$). Auch hier zeigt sich, dass die Ruhebedingung als weniger beeinträchtigend für die Konzentration empfunden wird als alle anderen Schallbedingungen (Post-hoc Auswertung: Sprache: $T(19)=4.3$; $p<.01$ (BH=.02); 2000-70: $T(19)=4.2$; $p<.01$ (BH=.03); 2000-70-12: $T(19)=3.3$; $p<.01$ (BH=.05)).

Der Effekt des Faktors *Schallbedingung* auf die Einschätzung der *Störung* erweist sich in einer Varianzanalyse als bedeutsam ($F(3; 57)=74.6$; $p<.01$; partielles $\eta^2=0.80$). Die Ruhebedingung wird im Vergleich zu den Schallen als weniger störend bewertet (Post-hoc Auswertung: Sprache: $T(19)=-12.4$; $p<.01$ (BH=.02); 2000-70: $T(19)=-15.0$; $p<.01$ (BH=.03); 2000-70-12: $T(19)=-10.4$; $p<.01$ (BH=.05)).

Ähnlich sieht es beim Urteil über die *Lästigkeit* der Schalle aus. Hier ergibt sich ebenfalls ein signifikanter Effekt der *Schallbedingung* ($F(3; 57)=83.4$; $p<.01$; partielles $\eta^2=0.81$); in der Post-hoc Auswertung ist auch dieser Effekt auf die Beurteilung der Ruhebedingung als weniger lästig als alle drei Schalle zurückzuführen (Sprache: $T(19)=-12.9$; $p<.01$ (BH=.02); 2000-70: $T(19)=-13.5$; $p<.01$ (BH=.03); 2000-70-12: $T(19)=-10.4$; $p<.01$ (BH=.05)).

3.2.3.4 Diskussion

Die Ergebnisse im *KLTM(Subwoofer)* decken sich mit den oben berichteten Ergebnissen des *KLTM(Kopfhörer)*. Auch hier bewirken nur der Sprachschall und der unmodifizierte Verkehrsschall (2000-70) im Vergleich zur Ruhebedingung eine signifikante Leistungsbeeinträchtigung in den prozentualen Fehlerraten, der tieffrequent abgesenkte Verkehrsschall jedoch nicht mehr. Vor allem die vergleichende Kovarianzanalyse legt nahe, dass die Gruppenunterschiede, die sich in den Daten zeigen, nicht nachweisbar etwas mit der Art der Schalldarbietung, sondern mit von vornherein bestehenden Gruppenunterschieden zu tun hatten.

Die Ergebnisse der Zwischenbefragung decken sich ebenfalls mit den oben berichteten Ergebnissen des *KLTM(Kopfhörer)*, weswegen hier nicht nochmals extra darauf eingegangen werden soll.

Allerdings kann diese Durchführung des KLTM unter Hintergrundschall aufs Neue belegen, dass die in Fragestellung I und oben berichteten Befunde replizierbar sind. Es ließ sich erneut zeigen, dass dieser Test sowohl auf Sprachschall als auch auf lauten Verkehrsschall sensitiv ist. Die tieffrequente Pegelreduktion des Verkehrslärms als probates Mittel, die Lärmbelastung zu reduzieren, ließ sich auch bei einer Schalldarbietung über Lautsprecher verdeutlichen. In dieser Durchführung des KLTM waren die Effekte sogar am stärksten nachweisbar. In Fragestellung I unterschied sich die Leistung unter dem lautesten Verkehrsschall nur signifikant von der unter dem leisesten, nicht jedoch von der unter Ruhe. In Kapitel 3.2.2 ergaben sich trotz einer nicht signifikanten Varianzanalyse signifikante Unterschiede in der Leistung zwischen dem lautesten Verkehrsschall und der Ruhebedingung, nicht jedoch durch den abgesenkten Verkehrsschall. In dieser dritten Durchführung nun konnte dieser Effekt mit einem signifikanten Ergebnis der Varianzanalyse repliziert werden, die Absenkung der Verkehrsschalle im tieffrequenten Bereich bewirkt tatsächlich eine geringer beeinträchtigende Wirkung in diesem Test, der Verkehrsschall mit einem Mittelungspegel von 70 dB(A) stört nachweislich die Leistung. Stabil zeigte sich zudem der leistungsbeeinträchtigende Effekt des Sprachschalls gegenüber der Ruhebedingung über alle drei Experimente hinweg.

3.2.4 Der Stroop Test mit pegelreduzierten Schallen

Nachdem in Fragestellung I beim Stroop Test ein leistungsbeeinträchtigender Effekt des Verkehrsschalls mit hoher Lautstärke gefunden werden konnte, wurde nun überprüft, ob eine tieffrequente Pegelreduktion dieses Schalls wie beim KLTM eine Minderung des Störeinflusses bewirkt. Hierfür wurden dieselben Verkehrsschalle wie beim KLTM verwendet (vgl. Tabelle 3 bzw. Anhang A für exemplarische Amplitudenzeitverläufe des Schalls 2000-70-12 und 2000-70).

3.2.4.1 Aufgabenstellung

Die Aufgabe im Stroop Test blieb wie bei Fragestellung I beschrieben.

3.2.4.2 Versuchsplan, Probanden und Ablauf

Der Stroop Test wurde nach einem zweifaktoriellen varianzanalytischen Design mit Messwiederholungen auf beiden Faktoren (*Kongruenz* und *Schallbedingung*) durchgeführt.

An dem zweistündigen Versuch nahmen 24 freiwillige Probanden (3 männlich, 21 weiblich) im Alter zwischen 19 und 30 Jahren (Median 21 Jahre) teil, die im Rahmen ihres Studiums entweder mit zwei Versuchspersonenstunden oder mit 10 € Aufwandsentschädigung entlohnt wurden. Alle hatten sich auf einen Aushang hin gemeldet und berichteten über normales Hörvermögen.

Jede Versuchsperson bearbeitete im Einzelversuch vor einem Pentium I PC nach einer schriftlichen Instruktion (vgl. Anhang E) und einem Lärmempfindlichkeitsfragebogen (NoiSeQ) 24 Übungsaufgaben unter Ruhe. Anschließend wurden ihnen unter jedem der Schalle 168 Aufgaben (jeweils gefolgt von dreiminütigen Pausen) des auf ERTS (Dr. Beringer, BeriSoft Cooperation) programmierten Stroop Tests vorgegeben. In den Pausen wurde wiederum die kurze Zwischenbefragung (vgl. Anhang B) ausgefüllt. Die Schalle wurden in einer interindividuell ausbalancierten Reihenfolge über Kopfhörer (Sennheiser, HD 600; Sony Digital Tape Deck DTC-2E700) dargeboten. Der Versuch fand in einem ruhigen Experimentalraum der Katholischen Universität Eichstätt-Ingolstadt statt.

3.2.4.3 Ergebnisse

Die Auswertung der prozentualen Fehlerraten unter dem Faktor *Schallbedingung* ergibt folgendes, in Abbildung 21 dargestelltes Bild.

Eine zweifaktorielle Varianzanalyse mit Messwiederholungen deckt einen signifikanten Einfluss des Faktors *Schallbedingung* (Greenhouse-Geisser Korrektur: $F(3.4; 77.3)=3.9$; $p=.01$; partielles $\eta^2=0.15$), des Faktors *Kongruenz* ($F(1; 23)=27.2$; $p<.01$; partielles $\eta^2=0.54$) und eine signifikante Interaktion der beiden ($F(5; 115)=2.7$; $p=.02$; partielles $\eta^2=0.11$) auf.

Eine hierauf folgende getrennte Auswertung der kongruenten und inkongruenten Items ergibt für die inkongruenten Items einen signifikanten Einfluss des Faktors *Schallbedingung* (einfaktorielle

ANOVA, $F(5; 115)=4.0$; $p<.01$; partielles $\eta^2=0.15$). Anhand geplanter Einzelvergleiche mit Benjamini – Hochberg Korrektur (*false discovery rate*; einseitige Testung wegen gerichteter Hypothesen, vgl. Kapitel 3.2.1 zur theoretischen Begründung) ergeben sich signifikant niedrigere Fehlerraten unter der Ruhebedingung als unter den Schallen *2000-70* ($T(23)=-2.9$; $p<.01$ (BH=.02)), *2000-70-12* ($T(23)=-2.9$; $p<.01$ (BH=.01)) und *2000-70-12 erhöht* ($T(23)=-2.8$; $p<.01$ (BH=.03)). Die Leistung unter den Schallen *Sprache* und *2000-70 abgesenkt* unterscheidet sich jedoch nicht von der Leistung unter Ruhe (*Sprache*: $T(23)=-1.8$; $p=.05$ (BH=.04); *2000-70 abgesenkt*: $T(23)=0.2$; $p=.41$ (BH=.05)).

Die ermittelte individuelle subjektive Lärmempfindlichkeit (NoiseQ: vorläufige Normwerte (Schütte & Marks, 2004): $M=43$; $SD=18$; gemessen: $M=52.5$; $SE=3.0$) variiert erneut nicht genügend, um einen Extremgruppenvergleich nahe zu legen, weswegen auf eine weiterführende kovarianzanalytische Auswertung verzichtet wurde.

Die varianzanalytische Auswertung der kongruenten Items ergibt keinen signifikanten Einfluss der *Schallbedingung* (Greenhouse-Geisser Korrektur: $F(3.2; 73.6)=2.3$; $p=.08$).

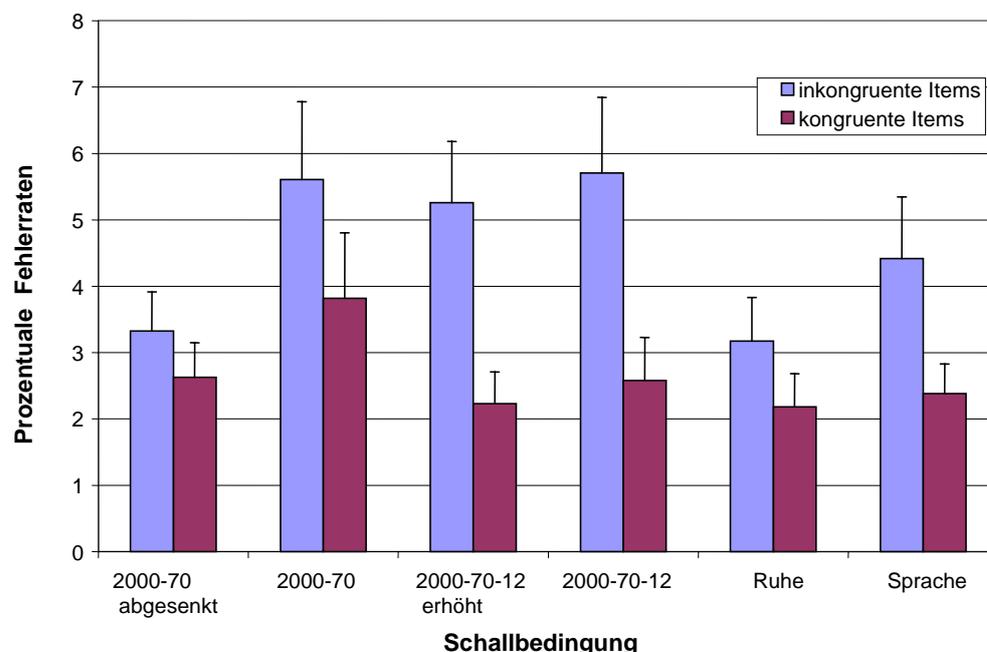


Abbildung 21: Prozentuale Fehlerraten im Stroop Test: Faktoren *Schallbedingung* und *Kongruenz* (Mittelwert und Standardfehler)

Eine im Hinblick auf einen Übungseffekt vorgenommene varianzanalytische Betrachtung der prozentualen Fehlerraten im Stroop Test erbringt außer dem erwarteten Haupteffekt des Faktors *Kongruenz* ($F(1; 23)=27.3$; $p<.01$; partielles $\eta^2=0.54$) keinen signifikanten Haupteffekt des Faktors *Versuchsabfolge* (Greenhouse-Geisser Korrektur: $F(3; 69.3)=0.8$; $p=.52$) und keine bedeutsame Wechselwirkung (Greenhouse-Geisser Korrektur: $F(3.4; 77.8)=0.2$; $p=.91$).

Eine Auswertung der Lösungszeiten, welche die Probanden zum Beantworten der Aufgaben benötigten, lässt nicht auf einen Speed-Accuracy-Trade-Off schließen. In einer zweifaktoriellen

ANOVA mit Messwiederholungen ist wieder nur der Haupteffekt der *Kongruenz* signifikant ($F(1; 23)=112.1$; $p<.01$; partielles $\eta^2=0.83$), nicht aber der der *Schallbedingung* ($F(5; 115)=0.5$; $p=.81$) oder deren Interaktion ($F(5; 115)=1.2$; $p=.33$).

Die Auswertung der Lösungszeiten, die über die Abschnitte des Experiments gemittelt wurden, ergibt wie vermutet Hinweise auf einen Übungseffekt. Eine zweifaktorielle Varianzanalyse belegt nicht nur statistisch den Stroop Effekt (Haupteffekt *Kongruenz*: $F(1; 23)=112.1$; $p<.01$; partielles $\eta^2=0.83$), sondern auch einen signifikanten Effekt des Faktors *Versuchsabfolge* (Greenhouse-Geisser Korrektur: $F(3.4; 78.3)=8.1$; $p<.01$; partielles $\eta^2=0.26$), und zwar bei den kongruenten wie den inkongruenten Items (keine bedeutsame Interaktion: Greenhouse-Geisser Korrektur: $F(3.7; 84.3)=0.8$; $p=.50$). Eine hierauf folgende Trendanalyse nach Page (theoretische Begründung für die Trendanalyse vgl. Kapitel 3.2.1) legt Hinweise auf einen Übungseffekt bei beiden Itemarten nahe (hochsignifikant fallender Trend: inkongruent: $L_{emp}=1931>L_{tab}(24; 6; 1\%)=1851$; kongruent: $L_{emp}=1954>L_{tab}(24; 6; 1\%)=1851$).

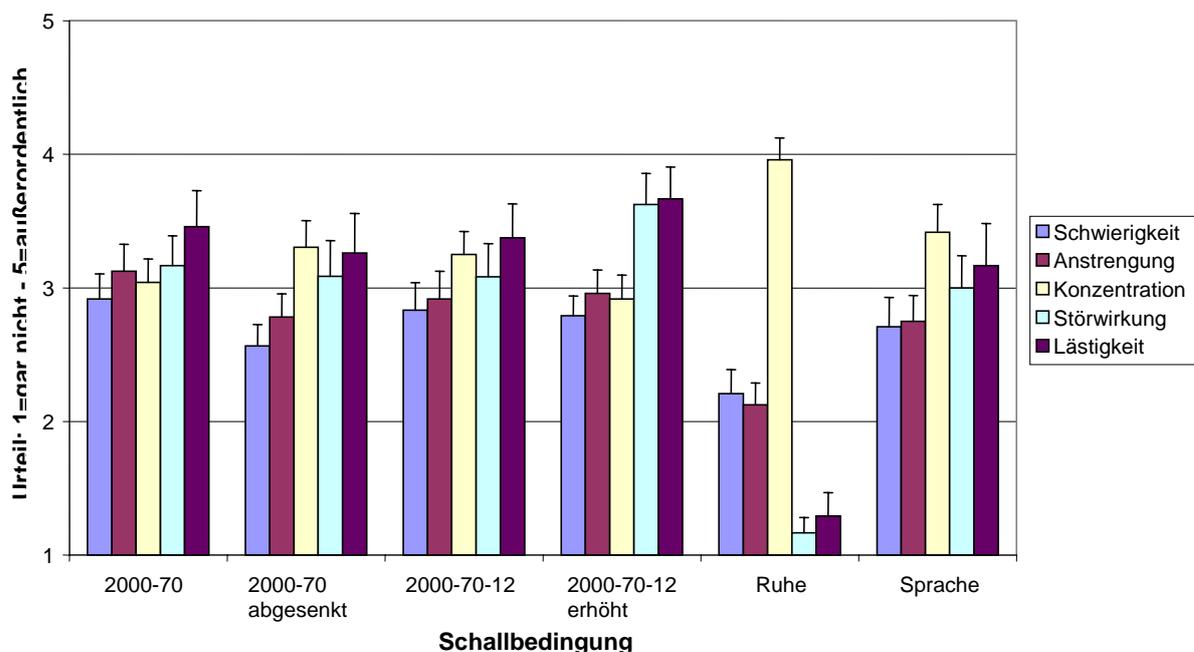


Abbildung 22: Auswertung der Urteile in der Zwischenbefragung im Stroop Test, Ratingskala von 1 bis 5 (1=gar nicht bis 5=außerordentlich); Mittelwert und Standardfehler

Eine Auswertung der Zwischenbefragungsurteile ergibt das in Abbildung 22 dargestellte Bild. Die Probanden beurteilten wieder jeweils nach jedem Aufgabenabschnitt die Schwierigkeit der Aufgabe, die benötigte Anstrengung, die Konzentrationsmöglichkeit, die Störwirkung des Schalls und seine Lästigkeit. Die Daten wurden varianzanalytisch getrennt nach den Fragen ausgewertet. Für alle Items ergeben sich statistisch bedeutsam unterschiedliche Urteile abhängig von der Ausprägung des Faktors *Schallbedingung* (Schwierigkeit: $F(5; 110)=3.5$; $p=.01$; partielles $\eta^2=0.14$; Anstrengung: $F(5;$

110)=5.9; $p < .01$; partielles $\eta^2 = 0.21$; Konzentration: $F(5; 110) = 6.3$; $p < .01$; partielles $\eta^2 = 0.22$; Störung: $F(5; 110) = 23.1$; $p < .01$; partielles $\eta^2 = 0.51$; Lästigkeit: Greenhouse-Geisser Korrektur: $F(3.3; 73.6) = 15.8$; $p < .01$; partielles $\eta^2 = 0.42$). Die Post-hoc Auswertung erfolgte mit gerichteten Hypothesen, wie in Kapitel 3.2.1 begründet wurde.

Die Post-hoc Testung mit Benjamini-Hochberg Korrektur für *Schwierigkeit* weist darauf hin, dass die Aufgabe unter allen Schallen als signifikant schwieriger als unter der Ruhebedingung beurteilt wird (Sprache: $T(23) = 2.1$; $p = .02$ (BH=.05); 2000-70: $T(23) = 3.6$; $p < .01$ (BH=.02); 2000-70 abgesenkt: $T(23) = 2.8$; $p < .01$ (BH=.04); 2000-70-12: $T(23) = 3.7$; $p < .01$ (BH=.01); 2000-70-12 erhöht: $T(23) = 2.8$; $p < .01$ (BH=.03)).

In Einzelvergleichen der Urteile über die *Anstrengung* konnten ebenfalls Unterschiede durch die Schallbedingungen dargelegt werden. Unter allen Schallen war mehr Anstrengung zur Lösung der Aufgabe nötig als unter der Ruhebedingung (Sprache: $T(23) = 2.9$; $p < .01$ (BH=.04); 2000-70: $T(23) = 4.4$; $p < .01$ (BH=.01); 2000-70 abgesenkt: $T(22) = 2.7$; $p < .01$ (BH=.05); 2000-70-12: $T(23) = 4.2$; $p < .01$ (BH=.02); 2000-70-12 erhöht: $T(23) = 3.6$; $p < .01$ (BH=.03)).

Weiterhin zeigt sich eine signifikant bessere Einschätzung der Konzentrationsmöglichkeit unter Ruhe als unter allen Schallen (Sprache: $T(23) = 2.5$; $p = .01$ (BH=.05); 2000-70: $T(23) = 5.4$; $p < .01$ (BH=.02); 2000-70 abgesenkt: $T(22) = 3.2$; $p < .01$ (BH=.04); 2000-70-12: $T(23) = 3.3$; $p < .01$ (BH=.03); 2000-70-12 erhöht: $T(23) = 5.9$; $p < .01$ (BH=.01)).

Die subjektive Störwirkung aller Schalle wurde von den Probanden ebenfalls als signifikant höher als die der Ruhebedingung beurteilt (Sprache: $T(23) = 6.9$; $p < .01$ (BH=.04); 2000-70: $T(23) = 7.4$; $p < .01$ (BH=.02); 2000-70 abgesenkt: $T(22) = 5.9$; $p < .01$ (BH=.05); 2000-70-12: $T(23) = 7.0$; $p < .01$ (BH=.03); 2000-70-12 erhöht: $T(23) = 9.9$; $p < .01$ (BH=.01)).

Wie die Post-hoc Auswertung des Items Lästigkeit ergibt, werden zudem alle Schalle von den Versuchspersonen als signifikant lästiger empfunden als die Ruhebedingung (Sprache: $T(23) = 5.1$; $p < .01$ (BH=.05); 2000-70: $T(23) = 6.1$; $p < .01$ (BH=.03); 2000-70 abgesenkt: $T(22) = 5.3$; $p < .01$ (BH=.04); 2000-70-12: $T(23) = 6.4$; $p < .01$ (BH=.02); 2000-70-12 erhöht: $T(23) = 8.6$; $p < .01$ (BH=.01)).

3.2.4.4 Diskussion

In diesem Test bestehen entgegen den Hypothesen keine Hinweise darauf, dass die subjektive Lautheit (in sone) der Hintergrundschalle das relevante Merkmal des Störeinflusses ist. Bei einer Variation der Schalle in Lautheit und Frequenzzusammensetzung zeigt sich nur ein Effekt des Schalldruckpegels (Mittelungspegel in Dezibel), da alle Verkehrsschalle mit mehr als 69 dB(A) im $L(A)_{eq}$ einen signifikant störenden Einfluss auf die Interferenzleistung (inkongruente Items) aufweisen. Nur der frequenzunabhängig reduzierte Schall ($L(A)_{eq} = 66$ dB(A)) besitzt keinen negativen Einfluss auf die Leistung im Vergleich zur Ruhebedingung.

Im Hinblick auf den Effekt des Sprachschalls auf diesen, vor allem das System der zentralen Exekutive fordernden Test, bestätigt sich das Ergebnis aus der ersten Durchführung. Dieser Schall hat keine nachweisbare leistungsbeeinträchtigende Wirkung.

Beim Stroop Test ist somit davon auszugehen, dass nicht die Frequenzzusammensetzung oder die Lautheit des Schalls entscheidend sind, sondern einzig sein Schalldruckpegel. Bei dieser Schlussfolgerung muss allerdings kritisch auf die insgesamt geringen Fehlerzahlen (drei bis sechs Prozent) hingewiesen werden, was dieses Ergebnis als Bodeneffekt erscheinen lässt. Die Unterscheidung von subjektiver Lautheit und Schalldruckpegel ist hier wichtig, da das Maß der Lautheit (Sone-Skala; N; Zwicker & Fastl, 1986) einer Verhältnisskalierung der Lautheitsempfindung entspricht. Der Schalldruckpegel hingegen wird in Dezibel gemessen (siehe Kapitel 1.4). Durch die A-Bewertung wird das Signal im tieffrequenten Bereich relativ stark abgesenkt, was aber nicht der menschlichen Wahrnehmung bei hohen Lautstärken angemessen ist. Aus diesem Grund wurde erwartet, dass eine höhere Lautheit, die für eine höhere Lautheitsempfindung steht, auch zu einer stärkeren Leistungsbeeinträchtigung führt als ein Anstieg des Schalldruckpegels.

Eine Auswertung der Daten des Stroop Test im Hinblick auf die Aufklärung des Leistungsverlaufs ergibt nach einer Auftrennung in jeweils vier Schallabschnitte das in Abbildung 23 dargestellte Bild (nur inkongruente Items, Vergleich der Verkehrsschalle zur Ruhebedingung). Hier sieht man deutlich den unterschiedlich leistungsstarken Beginn der Arbeit unter dem jeweiligen Schall. Zum Ende der Abschnitte hin gleicht sich die Leistung an, der Unterschied reduziert sich, was für Habituationsprozesse sprechen würde. Da sich dieser Unterschied jedoch bei den inkongruenten Items in der Größenordnung von 19 bis 20,5 richtig gelösten Aufgaben (bei maximal möglichen 21 richtigen Aufgaben pro Schallabschnitt) bewegt, verdeutlicht dies noch einmal den Verdacht auf einen Bodeneffekt und damit den eingeschränkten Aussagewert *dieser* Ergebnisse.

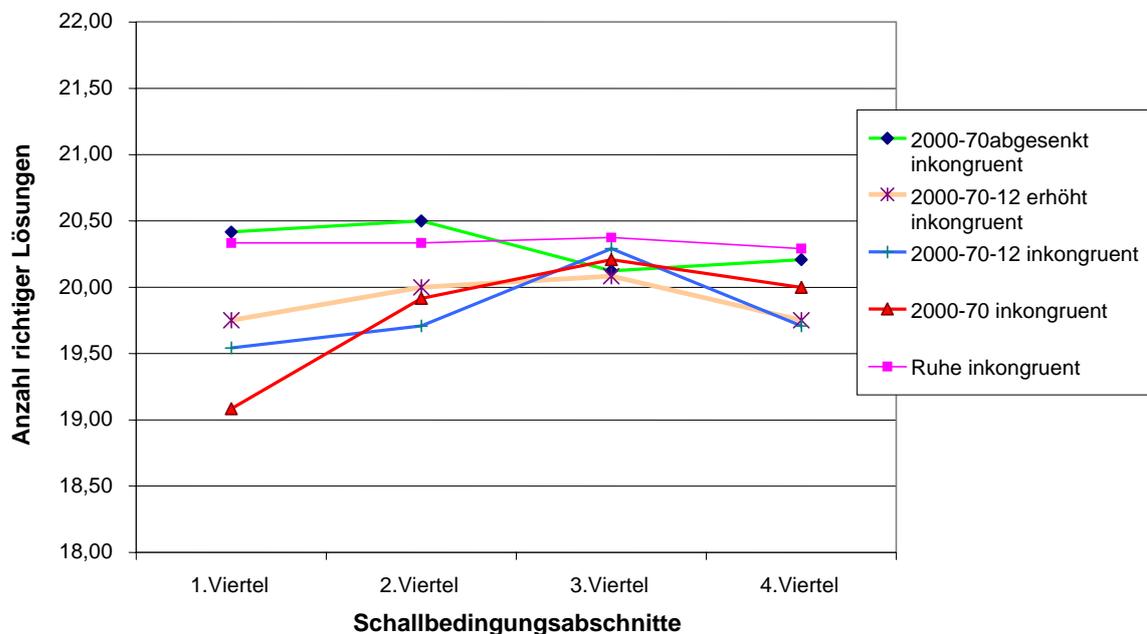


Abbildung 23: Leistungsverlauf im Stroop Test über die Schallabschnitte (Mittelwert der richtig gelösten Aufgaben; nur inkongruente Items, Vergleich der Verkehrsschalle mit der Ruhebedingung)

In der subjektiven Bewertung der Probanden zeigt sich, dass alle Schalle die Aufgabe erschweren und anstrengender machen, wobei ein Unterschied zwischen den beiden *nicht* frequenzmodifizierten Geräuschen zu sehen ist. Der *abgesenkte* Schall (66 dB(A)) beeinträchtigt die Aufgabe subjektiv etwas weniger als der *Originale*, was sich so auch in der Leistung wieder findet. Zudem wirken alle Geräusche im Vergleich zur Ruhebedingung lästig und erschweren die Konzentration. Die Störwirkung der Hintergrundschalle im Vergleich zur Ruhe liegt auf der Hand, jedoch wird von den Probanden zwischen den Verkehrsschallen differenziert. Im direkten Vergleich der jeweils nur in der Lautstärke unterschiedlichen Schalle mit *derselben* Frequenzzusammensetzung ergibt sich ein deskriptiv höheres Störungsurteil über den jeweils *lauteren* der beiden Schalle (vgl. Abbildung 22).

Der Zusammenhang der Lärmempfindlichkeit mit der Leistung unter Schall konnte erneut nicht untersucht werden, da aufgrund zu geringer Varianz kein Extremgruppenvergleich möglich war.

Auch bei diesem Test kommt es zu einem nachweisbaren Übungseffekt, was für die hohen Ansprüche der Aufgabe an das geforderte Subsystem, die zentrale Exekutive, spricht. Da aber die Schalle in interindividuell balancierter Reihenfolge dargeboten wurden sind die Interpretationen der Ergebnisse trotz der Übungseffekte aussagekräftig und fallen eher zu konservativ aus.

3.2.5 Das Grammatical Reasoning

Um die Spezifikation von leistungsbeeinträchtigenden Lärmparametern an einer weiteren Aufgabe, die ähnliche Anforderungen an das Arbeitsgedächtnis stellt wie der KLTM zu überprüfen, soll im Folgenden die Durchführung eines Grammatical Reasoning Tests mit vier der bisher beschriebenen Schalle vorgestellt werden.

Es wurden neben einer Ruhebedingung und dem bereits wiederholt beschriebenen Sprachschall nur zwei Verkehrsgeräusche im Vergleich verwendet. Da sich anhand der Ergebnisse der bisher durchgeführten Tests die Lautheit und die tieffrequente Lastigkeit als die kritischsten Merkmale der Verkehrsschalle erwiesen hatten, wurde zum einen der Schall mit 2000 Durchfahrten und einem Mittelungspegel von 70 dB(A) gewählt, sowie im direkten Vergleich dazu die Absenkung im tieffrequenten Bereich (unter 500 Hz) um 12 dB, hier bezeichnet mit 2000-70-12 (Tabelle 5).

Tabelle 5: Beschreibung der Schalle

Name	Beschreibung	Dauer	L_{eq} über 20 min in dB(A)
Ruhe			35dB(A)
Sprache	Helmut Griem spricht Briefe von Georg Büchner	20 min	60 dB(A) – 15 sone
2000 – 70	2000 Vorbeifahrten pro Stunde	20 min	70 dB(A) – 28 sone
2000 – 70-12	2000 Vorbeifahrten pro Stunde unter 500 Hz um 12 dB abgesenkt	20 min	70 dB(A) – 22 sone

3.2.5.1 Aufgabenstellung

Um einen Grammatical Reasoning Test in Anlehnung an das Original von Baddeley (1968) zu entwickeln, wurden Aufgaben nach diesem englischen Vorbild konstruiert – mit den Verben *vorausgehen* und *nachfolgen*, mit oder ohne Verneinung, im Aktiv oder Passiv (nur *nachfolgen*) und als korrekte oder falsche Aussage. Als Vergleichsobjekte wurden die drei Symbole aus dem Grammatical Reasoning der AGARD STRES Battery (Advisory Group for Aerospace Research and Development, 1989) verwendet (&, *, #). Weiterhin wurde aus der AGARD Aufgabe die Anordnung mit zwei Aussagesätzen und einer Symbolfolge von drei Zeichen übernommen sowie die Antwortregel, dass es eine Taste gibt (*Identisch*), die betätigt werden muss, wenn beide Sätze falsch oder beide richtig sind und eine andere (*Verschieden*), wenn einer falsch und einer richtig ist (Beispielaufgabe folgend). Es wurde ebenfalls beibehalten, dass eine korrekte Reihenfolge auch dann vorliegt, wenn die zwei fragten Symbole durch das dritte getrennt stehen (Salamé, 1993).

- | | |
|--------------------------------|-------------|
| (1) * geht & voraus | (1) Richtig |
| (2) # wird nicht gefolgt von & | (2) Falsch |
| (3) # * & | |

Antworttaste: *Verschieden*

Die Anforderungen, die diese Aufgabe an das Arbeitsgedächtnis stellt, sollten auf diese Weise im Vergleich zur Version von Baddeley (1968) erhöht werden.

Um eine gleichmäßige Schwierigkeitsverteilung zu erhalten, wurden die einzelnen Aussagesätze nach ihrer Schwierigkeit beurteilt und danach zu Gruppen ähnlicher Schwierigkeit zusammengefasst. Dies geschah folgendermaßen: Ein passiv verneinter Satz (z.B. *& wird nicht von # gefolgt*) erhielt die Schwierigkeitsstufe 3; ein passiver **oder** verneinter Satz (z.B. *# wird von & gefolgt oder * geht & nicht voraus*) erhielt die Schwierigkeitsstufe 2 und ein aktiver, positiver Satz (z.B. *& geht * voraus*) die Schwierigkeitsstufe 1. Anschließend wurden Sätze mit Schwierigkeitsstufe 1 und 3 kombiniert und 2 mit 2, um annähernd gleich schwere Aufgaben zu erhalten. Weiterhin wurde bei der Auswahl der Aufgaben darauf geachtet, dass richtige und falsche Lösungen gleich verteilt waren, sowie dass alle Symbole ebenso wie alle Symbolanordnungen gleich häufig vorhanden waren. Zudem gab es keine Kombination zweier gleicher Aussagen zu einer Aufgabe und es wurde ausbalanciert, dass jede Aussagemöglichkeit gleich häufig an erster oder zweiter Stelle dargeboten wurde.

Aus dem Pool der möglichen Aufgaben wurden anhand dieser Kriterien 50 Aufgaben für die Verwendung im Versuch ausgewählt. Aus Gründen der Vergleichbarkeit der Schallbedingungen wurde dieser Aufgabenpool für alle Versuchsbedingungen verwendet.

Die Präsentation der Aufgaben erfolgte auf drei Zeilen verteilt, zentriert auf dem Bildschirm eines Pentium I PC (Schriftgröße 16 p). Die Aufgabe wurde auf ERTS programmiert (Dr. Beringer, BeriSoft Cooperation). Im Übungsdurchgang betrug die Darbietungsdauer maximal 30 Sekunden, danach erschien die Aufforderung das nächste Mal schneller zu antworten. Außerdem erhielten die Probanden während der Übung eine Rückmeldung über die Korrektheit ihrer Antworten. In den Versuchsdurchgängen wurde die Zeitbegrenzung für die einzelnen Aufgaben ohne Wissen der Probanden aufgehoben, das Inter-Trial-Intervall betrug 500 ms.

3.2.5.2 *Versuchsplan, Probanden und Versuchsablauf*

Es liegt ein einfaktorielles varianzanalytisches Design mit Messwiederholung auf dem vierfach gestuften Faktor *Schallbedingung* vor.

An dem zweistündigen Versuch nahmen 20 Versuchspersonen im Alter zwischen 19 und 36 Jahren (Median 21.5 Jahre) teil (4 männlich, 16 weiblich; alle berichteten über normales Hörvermögen), die dafür mit zwei Versuchspersonenstunden im Rahmen ihres Psychologiestudiums oder 10 € Aufwandsentschädigung entlohnt wurden. Alle meldeten sich freiwillig auf einen Aushang hin.

Jede Versuchsperson bearbeitete im Einzelversuch in einem Versuchsraum der Katholischen Universität Eichstätt-Ingolstadt vor einem Pentium I PC unter jeder Schallbedingung 50 Aufgaben. Die Schalle wurden über einen Kopfhörer (Sennheiser HD 600; Sony Digital Audio Tape Deck DTC-2E700; Lautstärke gemessen mit einem Artificial Ear Type 4153 von Brüel & Kjaer) in interindividuell ausbalancierter Reihenfolge dargeboten. Die Versuchspersonen bearbeiteten nach einer schriftlichen Instruktion (Anhang F) und einem Lärmempfindlichkeitsfragebogen (NoiSeQ, Schütte & Marks, 2004) 25 Übungsaufgaben unter Ruhe. Anschließend wurden sie über den Versuchsstart informiert. Die

ersten 50 Aufgaben lösten sie jedoch immer unter Ruhe, was als Übungsphase nicht in die Auswertung aufgenommen wurde. Erst danach bearbeiteten sie unter je einem der balanciert dargebotenen Schalle 50 Aufgaben und füllten jedes Mal in der zweiminütigen Pause die bereits beschriebene kurze Zwischenbefragung (vgl. Anhang B) aus.

3.2.5.3 Ergebnisse

Eine einfaktorielle varianzanalytische Auswertung mit Messwiederholungen der prozentualen Fehlerraten (Abbildung 24) ergibt keine signifikanten Unterschiede ($F(3; 57)=1.7$; $p=.17$) in der Leistung durch den Faktor *Schallbedingung*.

Eine vergleichbare Auswertung der prozentualen Fehlerraten im Hinblick auf einen Effekt des Faktors *Versuchsabfolge* legt ebenso wenig Hinweise auf einen Übungseffekt nahe (Greenhouse-Geisser Korrektur: $F(1.6; 29.7)=0.3$; $p=.69$).

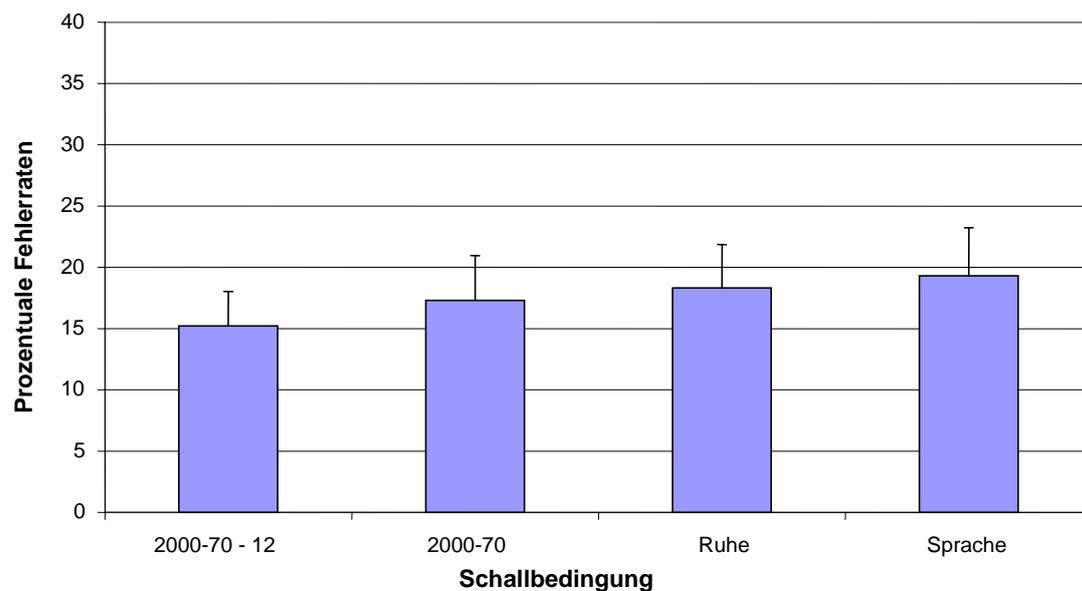


Abbildung 24: Prozentuale Fehlerraten im Grammatical Reasoning unter dem Faktor *Schallbedingung* (Mittelwert und Standardfehler)

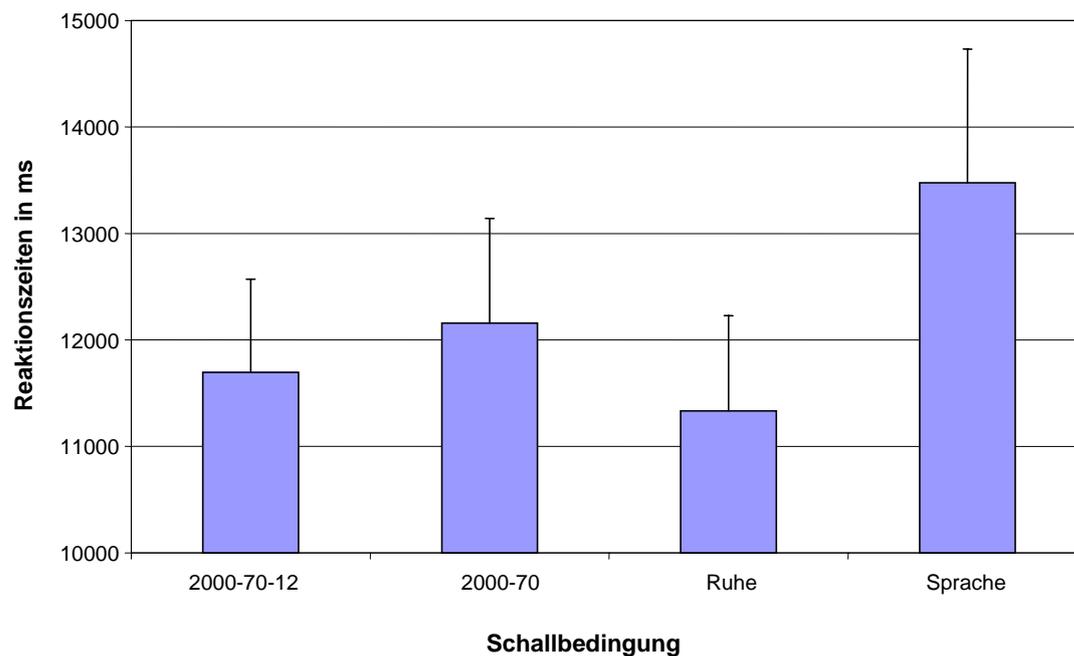


Abbildung 25: Mittlere Reaktionszeiten im Grammatical Reasoning unter den Schallbedingungen (Mittelwert und Standardfehler)

Ein anderes unerwartetes Bild entsteht, wenn die Reaktionszeiten unter den Hintergrundschallen varianzanalytisch ausgewertet werden (Abbildung 25). Hier zeigen sich statistisch bedeutsame Unterschiede durch den Faktor *Schallbedingung* (Greenhouse-Geisser Korrektur: $F(1.6; 30.2)=7.1$; $p=.01$; partielles $\eta^2=0.27$). Es liegt folglich ein Speed-Accuracy-Trade-Off vor. Dies bedeutet, die Probanden benötigten unter Lärm mehr Zeit um die Leistung zu halten. Eine Post-hoc Testung ergibt bei vier Einzelvergleichen mit gepaarten t-Tests (einseitige Testung aufgrund gerichteter Hypothesen einer *Verschlechterung* der Leistung unter den Schallen im Vergleich zur Ruhebedingung) und *Benjamini – Hochberg false discovery rate Korrektur* signifikant längere Reaktionszeiten unter der Schallbedingung Sprache als unter der Ruhebedingung ($T(23)=3.2$; $p<.01$ (BH=.01)). Der unmodifizierte Verkehrsschall 2000–70 verlängert ebenso die Reaktionszeiten im Vergleich zur Ruhe ($T(23)=2.5$; $p=.01$ (BH=.03)). Der Vergleich zwischen den beiden Verkehrsschallen verfehlt mit 6% nur ganz knapp die Signifikanz ($T(19)=1.6$; $p=.06$ (BH=.04)). Unter dem unmodifizierten 2000-70 Schall wird etwas mehr Zeit für die Lösung benötigt. Der modifizierte Schall 2000-70-12 verlangsamt die Lösungszeit im Vergleich zur Ruhebedingung nicht signifikant ($T(19)=1,5$; $p=.08$ (BH=.05)).

Die Daten der individuellen Lärmempfindlichkeit (NoiseQ: vorläufige Normwerte (Schütte & Marks, 2004): $M=43$; $SD=18$; gemessen: $M=51.7$; $SE=3.3$) bringen erneut aufgrund geringer Varianz keinen Hinweis auf das Vorliegen von Extremgruppen, weswegen auf eine weiterführende kovarianzanalytische Auswertung verzichtet wurde.

Ungeachtet der dargebotenen Schallbedingung ergibt eine Auswertung der Reaktionszeiten gemittelt über die Abschnitte des Experiments keine Hinweise auf einen Übungseffekt (Greenhouse-Geisser Korrektur: $F(2.1; 40.1)=0.8$; $p=.50$).

In der Zwischenbefragung wurden die Probanden abermals gebeten, ein Urteil auf einer verbal kodierten fünfstufigen Skala zu den Dimensionen *Schwierigkeit*, *Anstrengung*, *Konzentration*, *Störung* und *Lästigkeit* des Schalls abzugeben. Die Daten zeigen in der Beurteilung der Schwierigkeit der Aufgabe und der Anstrengung, die zu ihrer Lösung notwendig war, nur kleine Unterschiede durch den Faktor *Schallbedingung* (Abbildung 26; Auswertung mit gerichteten Hypothesen, zur Begründung vgl. Kapitel 3.2.1). Insgesamt wird die Aufgabe als eher anstrengend und schwierig beurteilt, wenn auch am meisten unter der Hintergrundsprache.

Eine einfaktorielle varianzanalytische Auswertung ergibt keine signifikanten Unterschiede durch den Faktor *Schallbedingung* in der Beurteilung der *Anstrengung* und der *Schwierigkeit* (Greenhouse-Geisser Korrektur: Schwierigkeit: $F(2.2; 41.7)=1.1$; $p=.35$; Anstrengung: $F(1.6; 31.8)=2.4$; $p=.12$). Interessant ist die Beurteilung der unter den Hintergrundschallen subjektiv möglichen Konzentration, da diese von den Probanden unter dem Sprachschall und dem unmodifizierten Verkehrsschall als eher weniger möglich beurteilt wird, unter Ruhe und dem abgesenkten Verkehrsschall hingegen als mittelmäßig bis eher gut möglich (auch in der Varianzanalyse zeigen sich signifikante Unterschiede durch den Faktor *Schallbedingung*: $F(3; 57)=8.2$; $p<.01$; partielles $\eta^2=0.30$). Zwischen *Störung* und *Lästigkeit* differenzieren die Probanden nur wenig. Die Urteile spiegeln sehr deutlich die Leistungsdaten wieder, mit Sprache als dem am meisten störenden Schall, gefolgt von dem originalen Verkehrsschall und dem abgesenkten Verkehrslärm (Varianzanalyse: Störung: $F(3; 57)=58.1$; $p<.01$; partielles $\eta^2=0.75$; Lästigkeit: $F(3; 57)=62.9$; $p<.01$; partielles $\eta^2=0.77$).

Eine Post-hoc Auswertung mit t-Tests und *Benjamini-Hochberg Korrektur* legt nahe, dass die *Konzentration* von den Probanden unter *Ruhe* als besser möglich eingeschätzt wird als unter *Sprache* ($T(19)=4.7$; $p<.01$ (BH=.013)) und unter dem unmodifizierten Verkehrsschall *2000-70* ($T(19)=2.5$; $p=.01$ (BH=.03)). Der Unterschied zwischen den beiden Verkehrsschallen *2000-70* und *2000-70-12* verfehlt durch die Korrektur nur ganz knapp die Signifikanz ($T(19)=-1.8$; $p=.05$ (BH=.04)). Subjektiv scheint es den Probanden leichter gefallen zu sein, sich unter dem modifizierten Verkehrsschall zu konzentrieren. Nicht signifikant unterscheidet sich bei den Probanden die subjektive *Konzentration* unter dem modifizierten Verkehrsschall *2000-70-12* von der unter der Ruhebedingung ($T(19)=-1.4$; $p=.08$ (BH=.05)).

Die Störwirkung aller Schalle wird, wie die Post-hoc Auswertung verdeutlicht, als signifikant höher eingeschätzt als die der Ruhebedingung (Sprache: $T(19)=-11.9$; $p<.01$ (BH=.017); 2000-70: $T(19)=-8.5$; $p<.01$ (BH=.03); 2000-70-12: $T(19)=-8.0$; $p<.01$ (BH=.05)).

Weiterhin zeigt sich, dass alle drei Schalle als lästiger empfunden wurden als die Ruhebedingung (Sprache: $T(19)=11.9$; $p<.01$ (BH=.017); 2000-70: $T(19)=9.9$; $p<.01$ (BH=.03); 2000-70-12: $T(19)=9.7$; $p<.01$ (BH=.05)).

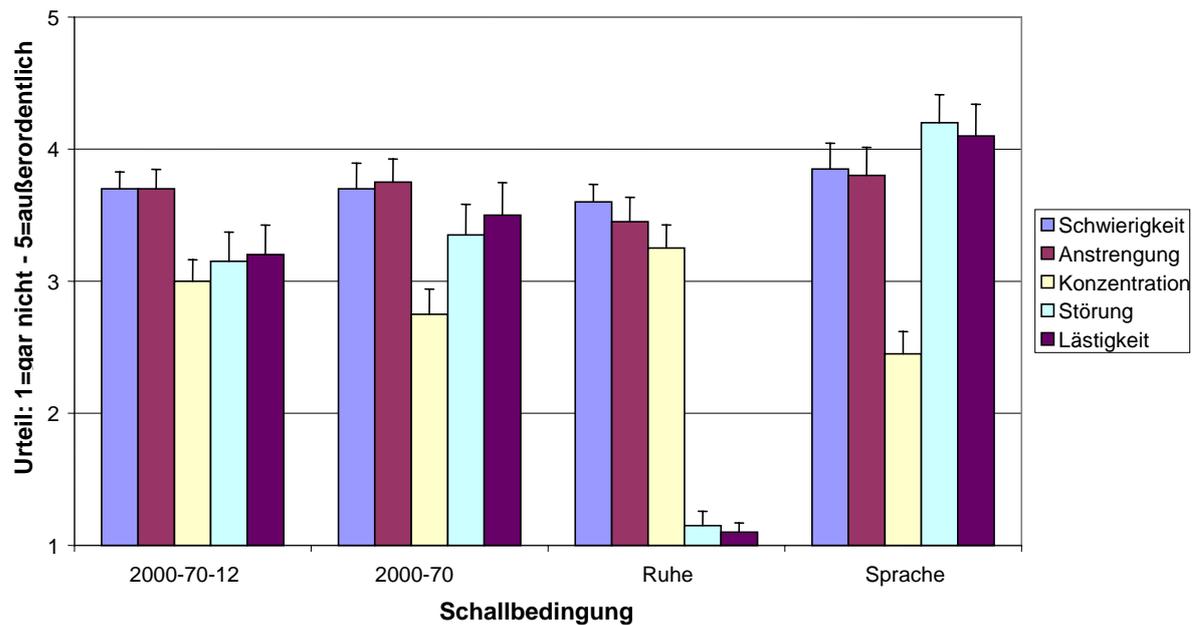


Abbildung 26: Zwischenbefragungsurteile im Grammatical Reasoning (Mittelwert und Standardfehler) auf einer Skala von 1(gar nicht) bis 5 (außerordentlich)

3.2.5.4 Diskussion

In dieser anspruchsvollen Version des grammatikalischen Schlussfolgerns, in der erst eine Übersetzung der z. T. negativen und passiven Formulierungen in eine Reihenfolge der Symbole erfolgen muss, wirken sich sowohl der in der Lautheit höhere Verkehrslärm mit mehr tiefen Frequenzanteilen als auch der Sprachschall verlangsamernd auf die Leistung aus. In dieser Aufgabe verlagerte sich die Leistungsbeeinträchtigung auf die Schnelligkeit statt auf die Genauigkeit der Aufgabenbearbeitung. Eventuell war in der Instruktion die Komponente des Zeitdrucks nicht genügend betont worden, so dass die Probanden die Beeinträchtigung der Leistung durch den Schall über eine Verlängerung der Antwortzeiten auszugleichen versuchten. Dies ist beim Grammatical Reasoning im Gegensatz zu den anderen vorgestellten Aufgaben auch eine zielführende Taktik, da in diesem Test die Aufgaben so lange zu sehen sind, bis die Probanden eine Antworttaste drücken. Es liegt kein Retentionsintervall vor wie beim Serial Recall, die Reize verschwinden nicht sukzessive wie im KLTM und die Darbietung ist nicht an sich zeitlich limitiert wie im Stroop Test.

Die tieffrequente Absenkung des Verkehrslärms bewirkt bei gleichzeitiger Reduktion der subjektiven Lautheit (in *sona*) eine signifikante Beschleunigung der Leistung im Vergleich zum unbearbeiteten Schall. Dieser Unterschied tritt vor allem im ersten und dritten Viertel der Bearbeitungszeit auf, wie eine Betrachtung des Leistungsverlaufs (Abbildung 27) ergibt. Dies bestätigt die Ergebnisse der Auswertung der Reaktionszeiten, die sich je nach Hintergrundschall anders gestalten. Auch hier ergeben sich jedoch keine Hinweise auf eine Passung des Konzepts der

reaktiven Anspannungssteigerung (Düker, 1963). Keine der Leistungen unter Hintergrundschall steigt nach einem anfänglichen Leistungseinbruch über das Niveau unter Ruhe an, was Hinweise auf einen Kompensationsvorgang geben würde. Da die Reaktionszeiten in der Ruhebedingung am konstantesten über den Abschnitt hinweg abnehmen, kann zudem wohl eher von Übungsfortschritten als von Habituationsprozessen an die Schalle ausgegangen werden.

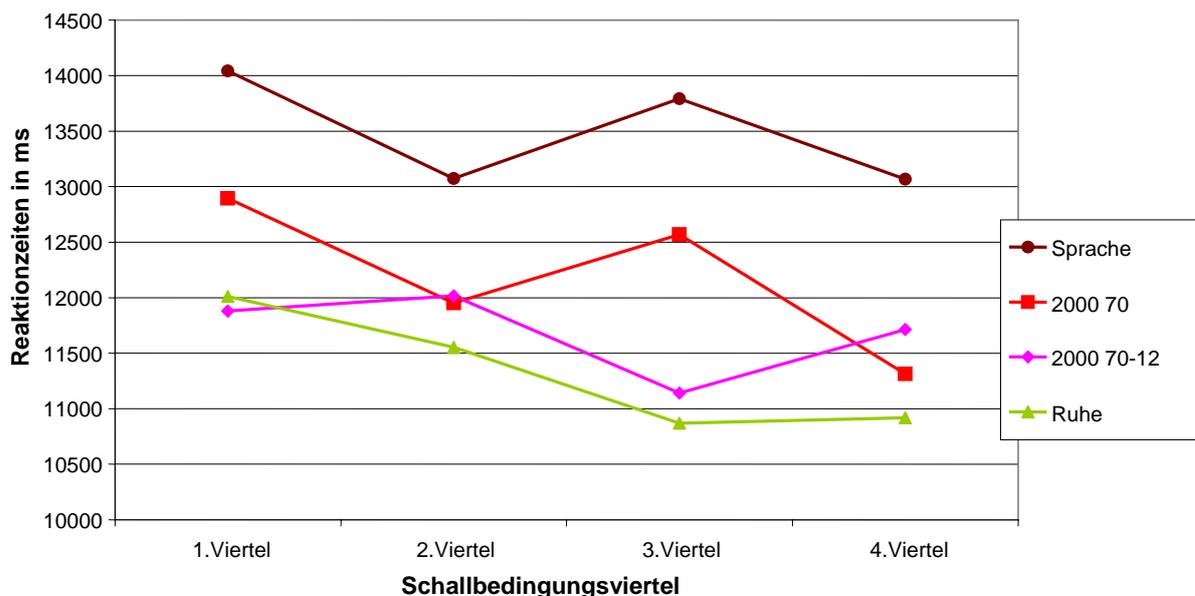


Abbildung 27: Leistungsverlauf im Laufe eines Schallabschnittes, Reaktionszeiten in ms im Grammatical Reasoning (Mittelwert über die Versuchspersonen)

Aus den Daten der Zwischenbefragung kann man schließen, dass der Effekt der Schalle *Sprache* und *2000-70* vor allem in der beeinträchtigten Konzentration liegt, obwohl alle drei Schalle als störend und lästig beurteilt werden. Da aber die individuelle Lärmempfindlichkeit wiederum aufgrund zu geringer Varianz nicht in die Rechnung einbezogen werden kann, spricht auch dieses Befundmuster nicht für eine reine Erklärung über ablenkende Gedanken oder Ärgerkognitionen durch den Schall.

Somit gibt es auch in diesem Test Hinweise darauf, dass wie beim KLTM eine Wirkung des Sprachschalls in moderater Lautstärke ($L_{eq}=60$ dB(A)) auf die phonologische Schleife vorliegt. Folglich ergeben sich Hinweise, dass der ISE bei einer weiteren Aufgabe als dem Serial Recall nachzuweisen sein könnte.

Die erwartete leistungsbeeinträchtigende Wirkung des lauten Verkehrsschalls ($L_{eq}=70$ dB(A)) lässt sich ebenfalls finden. Nach der Absenkung im tiefen Frequenzbereich (um 12 dB unter 500 Hz), welche zugleich die Lautheit absinken lässt (2000-70-12), findet sich die leistungsbeeinträchtigende Wirkung des Verkehrsschalls wie vermutet nicht mehr (Berglund, Hassmen, & Job, 1996).

Auch hier zeigt sich somit, dass eine Modifikation von Verkehrsschallen nach Klingenberg (1991) eine adäquate Maßnahme wäre, um kognitive Arbeit unter Verkehrslärmexposition zu erleichtern. Auf weitere praktische Implikationen soll in Kapitel 4 in der zusammenfassenden Diskussion rekuriert werden.

3.2.6 Diskussion der Ergebnisse zu Fragestellung II

Die Leistungsbeeinträchtigung im KLTM unter den modifizierten Verkehrsschallen entspricht den Erwartungen. Es ist davon auszugehen, dass sowohl die subjektive Lautheit der Schalle als auch deren tiefe Frequenzanteile relevant für das Leistungsniveau sind. Eine Reduktion dieser Parameter bewirkt, dass keine Leistungsbeeinträchtigung mehr nachweisbar ist. Eine Unterscheidung der Wirkung der Parameter war in der vorliegenden Größenordnung der Fehlerraten und aufgrund der nicht signifikanten Varianzanalyse leider nicht möglich.

Eine Replikation der Ergebnisse des KLTM und damit auch der Daten der Fragestellung I gelang mit einer Darbietung der Hintergrundschalle über Lautsprecher. Die kovarianzanalytische Auswertung ergibt keine nachweisbaren Unterschiede durch die Art der Schalldarbietung. Die Verwendung von Kopfhörern zur besseren Kontrolle von Störgeräuschen kann somit als probates Mittel gelten, um selbst tieffrequente Schalleffekte zu untersuchen. Der leistungsbeeinträchtigende Effekt des lautesten Verkehrsschalls (Mittelungspegel von 70 dB(A)) gegenüber der Ruhebedingung ließ sich in dieser dritten Durchführung des KLTM signifikant nachweisen. In Fragestellung I lag dieser Effekt noch nur im Vergleich zum leisesten Verkehrsschall vor, in der zweiten Durchführung mit der Einschränkung einer nicht signifikanten Varianzanalyse. Auch der positive leistungsbeeinflussende Effekt der tieffrequenten Pegelreduktion des Verkehrsschalls konnte mit der dritten Durchführung des KLTM sogleich repliziert werden. Eine signifikante Leistungsbeeinträchtigung durch den Sprachschall ließ sich zudem stabil in allen drei Testdurchführungen (Fragestellung I und II) zeigen. Der KLTM in der vorliegenden Version kann somit als sehr gut geeignet für eine Erforschung der Auswirkungen von Sprachschall und Verkehrsschall auf kognitive Leistungen angesehen werden.

Die Ergebnisse des Stroop Tests scheinen in eine andere Richtung zu deuten: signifikante Unterschiede in den Fehlerraten der Interferenzbedingung ergeben sich beim Vergleich der Schalldruckpegel (in dB(A)) der Schalle. Da allerdings nicht der Schalldruckpegel (in dB(A)) der Geräusche, sondern ihre subjektive Lautstärke (Lautheit in sone) dem menschlichen Empfinden entspricht, ist dieses Ergebnis überraschend. Eine so spezifische leistungsbeeinträchtigende Wirkung des Schallpegels mit einer Grenze bei einem Mittelungspegel von 69 dB(A) war nicht zu erwarten (Hellbrück & Ellermeier, 2004; Zwicker & Fastl, 1986). Eine mögliche Erklärung für diesen Widerspruch ergibt sich, wenn die Höhe der Fehlerraten betrachtet wird. Die signifikanten Unterschiede finden sich hier zwischen 3,2 und 5,7% Fehler, was für nur ein bis zwei falsch gelöste Aufgaben steht. Da diese Fehlerzahlen auffallend gering sind, kann man hier von einem Bodeneffekt ausgehen, was die Aussagekraft *dieser* Daten erheblich schmälert.

Anhand der Grammatical Reasoning Aufgabe konnten vergleichbare Schallwirkungen aufgezeigt werden wie sie beim KLTM gefunden wurden. Allerdings lag ein Speed-Accuracy-Trade-Off vor, da sich die Schalleffekte in einer Verlängerung der Reaktionszeiten zeigten, nicht jedoch in einer Erhöhung der Fehlerraten. Beim Grammatical Reasoning ergaben sich ebenso wie beim KLTM Hinweise auf die vorhergesagte leistungsbeeinträchtigende Wirkung des Sprachschalls (in einer moderaten Lautstärke) und des unmodifizierten Verkehrsschalls (in hoher Lautstärke). Auch hier

bewirkte die Pegelreduktion des Verkehrsschalls im tieffrequenten Bereich eine Verringerung der Störwirkung.

Eine Modifikation der Pkw-Geräusche durch Motormanagement und schalldämmende Auskleidung des Motorraums, wie die von Klingenberg (1991) vorgeschlagene, kann somit als probates Mittel gelten, die störenden Einflüsse von Straßenverkehrslärm zu reduzieren. Die in den Experimenten verwendete Reduktion um zwölf Dezibel unter 500 Hz kann im Vergleich zu den bisher technisch umsetzbaren bis zu sechs Dezibel (Klingenberg, 1991; Bisping, 2005) als richtungweisend gelten. In einer so kurzen Bearbeitungszeit und einer konstruierten Leistungssituation mit nur 18, 20 bzw. 24 Probanden wie in den vorgestellten Tests ist eine geringere Leistungsbeeinträchtigung durch eine Modifikation der Verkehrsschalle um zwölf Dezibel nachweisbar. Es kann somit davon ausgegangen werden, dass bereits eine Reduktion um sechs Dezibel für kognitiv arbeitende Büroangestellte, die unter realen Bedingungen und oft unter größerem Stress stundenlang und das jeden Tag arbeiten müssen eine erhebliche Erleichterung darstellen würde. Somit kann aus diesen Ergebnissen als Empfehlung für die Gesetzgebung bzw. die Fahrzeugindustrie abgeleitet werden, dass eine Reduktion der tieffrequenten Anteile von Verkehrsschallen eine sinnvolle Möglichkeit wäre, die beeinträchtigende Wirkung auf geistige Arbeiten, die zentral exekutiver Kontrolle unterliegen, zu vermindern.

3.3 Fragestellung III: Relevanz der Semantik für die Wirkung des Sprachschalls

Wie in Kapitel 2 beschrieben, werden nachfolgend der KLTM und das Grammatical Reasoning unter semantisch unverständlichem japanischem Hintergrundsprachschall getestet. Zusätzlich wurde in diesen Experimenten statt Straßen- Schienenverkehrslärm verwendet. Auch der Schienenverkehr wurde wie der Straßenverkehrsschall im tiefen Frequenzbereich abgesenkt, in diesem Fall um 12 dB unter 220 Hz (vgl. Anhang H: Filterfunktion), da in diesem Bereich bei Schienenverkehrslärm eine Möglichkeit der Umgestaltung durch Modifikation der Schienen und Räder läge (Working Group Railway Noise on the European Commission, 2003).

Die *Lärmempfindlichkeit* wird weiterhin als Kontrollvariable erhoben, obwohl nach den bisherigen Daten keine großen Unterschiede zu erwarten sind.

Um einen möglichen Einfluss des Darbietungsmodus auch beim Schienenverkehrslärm zu kontrollieren, wird der KLTM exemplarisch mit der Lautsprecherdarbietung durchgeführt, das Grammatical Reasoning jedoch mit einer Darbietung über Kopfhörer.

3.3.1 Explizierung der Hypothesen

Im KLTM und dem Grammatical Reasoning wird ein Anstieg der Fehlerraten unter dem unverständlichen Sprachschall erwartet. Es ist entsprechen den ISE-Ergebnissen davon auszugehen, dass sich diese Leistungsverschlechterung in einer Größenordnung bewegt, die mit der unter deutschem Schall (Fragestellung I und II) vergleichbar ist. Ein Speed-Accuracy-Trade-Off wird aufgrund der vermuteten direkt interferierenden Wirkungsweise des Schalls wiederum nicht erwartet, wie bereits in Fragestellung I vorgestellt. Erwartet wird weiterhin, dass der Schienenverkehrslärm bei einer hohen Lautstärke (Mittelungspegel von 70 dB(A)) zu einer Leistungsbeeinträchtigung in einer ähnlichen Größenordnung wie der Straßenverkehrslärm führt. Ein Speed-Accuracy-Trade-Off wird nicht erwartet, da die Aufgaben unter Zeitdruck durchgeführt werden. Beim Grammatical Reasoning werden die Probanden aufgrund des Speed-Accuracy-Trade-Off, der in der ersten Durchführung (Fragestellung II) aufgetreten ist, nach den Übungsdurchgängen zusätzlich mündlich instruiert, schneller zu arbeiten. Auch in diesem Test sollte sich somit kein Trade-Off mehr finden lassen. Für die tieffrequent abgesenkten Schienenverkehrsgeräusche (-12 dB unter 220 Hz) sollte keine Leistungsbeeinträchtigung nachweisbar sein.

Da erneut eine Verbesserung der Leistung im Verlaufe der Versuche erwartet wird, erfolgt abermals eine Kontrolle der Übungseffekte in der Auswertung mittels Trendanalysen sowie eine Darbietung der Schallbedingungen in interindividuell balancierter Reihenfolge, um eine Konfundierung der Schalleffekte mit einer Verbesserung der Leistung durch Übung zu vermindern.

In Zwischenbefragungen werden wieder subjektive Urteile der Probanden über die *Schwierigkeit*, die *Anstrengung*, die *Konzentration* sowie die *Störung* und *Lästigkeit* des Hintergrundschalls erhoben. Es wird erneut ein höheres Lästigkeits- und Störungsurteil über die Schallbedingungen erwartet, die als leistungsbeeinträchtigend in den jeweiligen Tests vermutet

werden. Diese sollten zudem im Vergleich zur Ruhebedingung subjektiv die Konzentration behindern und die Aufgabe erschweren sowie anstrengender machen.

Die individuelle Lärmempfindlichkeit, die als Kontrollvariable erneut erhoben wird, soll erneut nur dann als Kovariate in die Rechnung eingeführt werden, wenn eine genügend große Varianz in den Daten auf Extremgruppen schließen lässt.

3.3.2 KLTM unter unverständlichem Sprachschall und Schienenverkehrslärm

Die semantisch unverständliche fünfzehnminütige Sprachaufnahme bestand aus dem Mitschnitt einer Vorlesung in japanischer Sprache (männlicher Sprecher: vgl. Tabelle 6). Der verwendete Schienenverkehrslärm bestand aus Aufnahmen von Durchfahrten unterschiedlicher Zugtypen (ICE, Güterzug, RE), die zum Zweck einer besseren Vergleichbarkeit mit den Straßengeräuschen mit kurzen Pausen (1-2 Sekunden) aneinandergereiht wurden⁷. Um auch bei diesen Geräuschen die Auswirkung einer tieffrequenten Absenkung zu überprüfen, wurde dieses Signal anschließend unter 220 Hz um 12 dB abgesenkt (Filterfunktion siehe Anhang H; exemplarische Pegelzeitverläufe für beide Schienenverkehrsschalle siehe Anhang A). In diesem Bereich liegen die besten Möglichkeiten einer Reduktion der Lärmbelastung durch eine Modifikation der Rad-Schiene-Interaktion (Bisping, 2005; Working Group Railway Noise on the European Commission, 2003).

Tabelle 6: Beschreibung der Schalle

Name	Beschreibung	Dauer	L_{eq} über 20 min in dB(A)
Ruhe			35 dB(A)
Japanische Sprache	Japanischer Sprecher	15 min	60 dB(A) – 13 sone
Schiene original	div. Züge, 70 dB(A) L_{eq}	15 min	70 dB(A) – 26 sone
Schiene - 12	div. Züge, 70 dB(A) L_{eq} , unter 220 Hz um 12 dB abgesenkt	15 min	70 dB(A) – 25 sone

3.3.2.1 Aufgabenstellung

Die Aufgabe war der in Fragestellung I und II beschriebenen sehr ähnlich. Jede Versuchsperson bearbeitete im Einzelversuch vor einem I-Book unter jedem Hintergrundschall aber diesmal festgelegt je 24 Aufgaben, die in Schwierigkeit, Anzahl der Additionen/Subtraktionen und notwendigen Zehnersprüngen in den Rechnungen ausbalanciert waren.

3.3.2.2 Versuchsplan, Versuchspersonen und Versuchsablauf

Der KLTM wurde erneut nach einem einfaktoriellen varianzanalytischen Plan mit Messwiederholung auf dem diesmal vierfach gestuften Faktor *Schallbedingung* durchgeführt.

An dem zweistündigen Versuch nahmen 20 freiwillige Probanden (6 männlich, 14 weiblich) im Alter zwischen 19 und 30 Jahren (Median 22 Jahre) teil, von denen ein Großteil Psychologiestudenten war. Alle hatten sich auf einen Aushang hin gemeldet und gaben an, über ein normales Hörvermögen zu verfügen. Ein Teil von ihnen wurde im Rahmen ihres Studiums mit zwei Versuchspersonenstunden entlohnt, die Übrigen erhielten 10 € als Aufwandsentschädigung.

Die Versuchspersonen bearbeiteten nach einer schriftlichen Instruktion (vgl. Anhang D) und einem Lärmempfindlichkeitsfragebogen (NoiSeQ, Schütte & Marks, 2004) zunächst zehn

⁷ erstellt von Prof. Bisping, SASS acoustic research & design GmbH im Rahmen des Verbundprojektes *Leiser Verkehr* des BMBF und des TÜV Rheinland, Laufzeit 2002-2008

Übungsaufgaben unter Ruhe. Der erste Durchgang nach den Übungsaufgaben fand ebenfalls unter Ruhe statt und wurde als zusätzlicher Übungsdurchgang nicht ausgewertet. Anschließend bearbeiteten sie unter allen Schallbedingungen jeweils 24 Aufgaben. Nach jedem Abschnitt füllten sie die bereits wiederholt beschriebene kurze Zwischenbefragung (vgl. Anhang B) aus. Die Hintergrundschalle wurden über Lautsprecher (Standardstereoanstellung: Genelec 1030A; Subwoofer: Genelec 7070A; Darbietung über Pentium IV PC) in einer interindividuell balancierten Reihenfolge dargeboten.

3.3.2.3 Ergebnisse

Eine einfaktorielle varianzanalytische Auswertung der prozentualen Fehlerraten (Abbildung 28) ergibt einen statistisch bedeutsamen Einfluss des Faktors *Schallbedingung* ($F(3; 57)=7.3$; $p<.01$; partielles $\eta^2 =0.28$). In Post-hoc Einzelvergleichen zeigt sich, dass die Unterschiede vor allem in der signifikant schlechteren Leistung unter dem Sprachschall relativ zur Ruhebedingung liegen (einseitige Testung: Ruhe-Japanisch: $T(19)=-3.7$; $p<.01$ (BH=.01)). Im Einzelnen betrachtet zeigen 15 von 20 Versuchspersonen unter dem Sprachschall eine geringere Leistung als unter Ruhe. Die Vergleiche der Fehlerraten unter den Schienenverkehrsschallen mit jener unter der Ruhebedingung verfehlen die Signifikanz (einseitige Testung: Ruhe-Schiene original: $T(19)=-0.2$; $p=.44$ (BH=.05); Ruhe – Schiene–12: $T(19)=0.4$; $p=.36$ (BH=.03)), ebenso wie der Vergleich der prozentualen Fehler unter den Schienenverkehrsschallen (Schiene original – Schiene-12: $T(19)=0.2$; $p=.42$ (BH=.04)). Nur 9 von 20 Probanden zeigen eine Leistungsverschlechterung unter dem originalen Schienengeräusch im Vergleich zur Ruhebedingung.

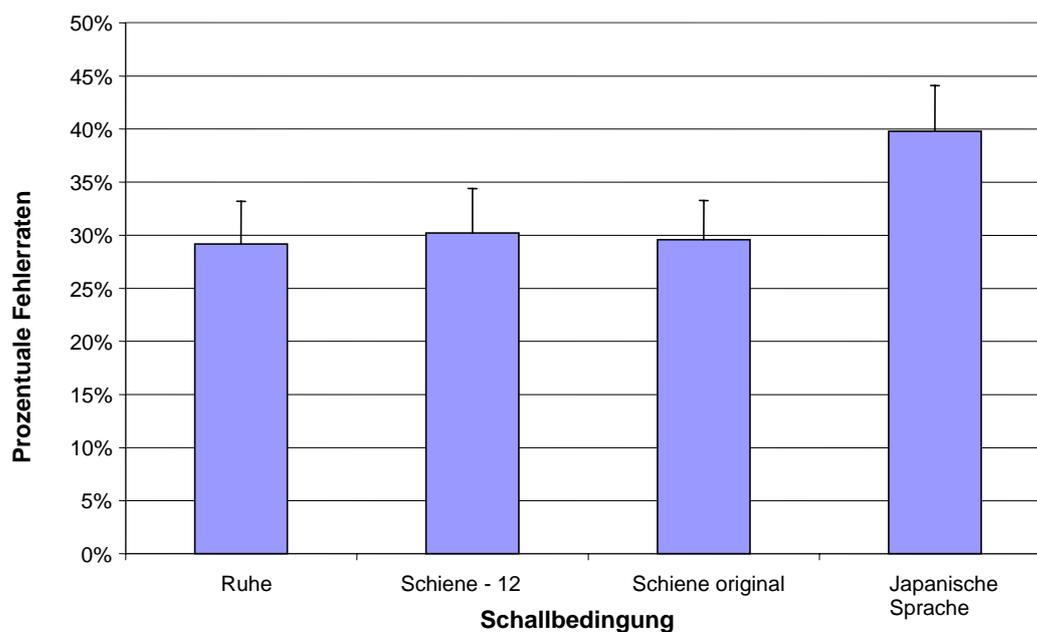


Abbildung 28: Auswertung der prozentualen Fehlerraten im KLTM; Faktor *Schallbedingung* (Mittelwert und Standardfehler)

Eine Auswertung der Fehlerraten gemittelt über die Abschnitte des Experiments im Hinblick auf einen Übungseffekt ergibt in einer varianzanalytischen Auswertung keinen signifikanten Unterschied durch den Faktor *Versuchsabfolge* ($F(3; 57)=2.0$; $p=.13$). Es liegt somit kein Hinweis auf einen Übungseffekt vor.

Bei der Auswertung der mittleren Lösungszeit pro Aufgabe zeigt sich kein bedeutsamer Unterschied durch den Faktor *Schallbedingung* (einfaktorielle ANOVA: $F(3; 57)=0.3$; $p=.82$). Es gibt folglich auch keine Evidenz für einen Speed-Accuracy-Trade-Off.

Eine varianzanalytische Auswertung der mittleren Lösungszeiten für die Darbietungsbedingungen im Hinblick auf einen Übungseffekt legt signifikante Unterschiede in der Arbeitsgeschwindigkeit in Abhängigkeit vom Faktor *Versuchsabfolge* nahe (Greenhouse-Geisser Korrektur: $F(2.1; 39.7)=9.7$; $p<.01$; partielles $\eta^2=0.34$). Ein hierauf folgender Trendtest (theoretische Begründung für den Trendtest vgl. Kapitel 3.3.1) offenbart einen hochsignifikant fallenden Trend mit Abnahme der mittleren Lösungszeiten im Versuchsverlauf und somit einen Übungseffekt (Trendtest nach Page: $L_{emp}(20; 4)=543 > L_{tab}(20; 4; 1\%)=531$).

Die individuelle Lärmempfindlichkeit der Probanden (erhoben über den NoiSeQ: vorläufige Normwerte (Schütte & Marks, 2004): $M=43$; $SD=18$; gemessen: $M=51.1$; $SE=1.9$) deutet wiederum nicht auf das Vorliegen von Extremgruppen hin, da eine zu geringe Varianz im Vergleich zu den Normwerten vorliegt. Aus diesem Grund wurde erneut auf eine weiterführende kovarianzanalytische Auswertung verzichtet.

Eine Auswertung der Zwischenbefragungsurteile untermauert die der Leistungsdaten (Abbildung 29; Post-hoc Testung mit gerichteten Hypothesen, zur theoretischen Begründung vgl. Kapitel 3.3.1). Aus der varianzanalytischen Auswertung der Urteile über die einzelnen Items gehen für die *Schwierigkeit* signifikante Unterschiede in der Bewertung abhängig vom Faktor *Schallbedingung* hervor ($F(3; 57)=3.5$; $p=.02$; partielles $\eta^2=0.16$). Eine Post-hoc Einzeltestung zeigt, dass die Aufgabe unter dem japanischen Hintergrundsfall als schwerer als unter Ruhe beurteilt wird ($T(19)=-3.2$; $p<.01$ (BH=.02)). Durch die Schienenverkehrsschalle dagegen wird die Aufgabe im Vergleich zur Ruhebedingung subjektiv nicht nachweisbar erschwert (Schiene-12: $T(19)=1.6$; $p=.06$ (BH=.03); Schiene original: $T(19)=0.9$; $p=.19$ (BH=.05)).

Auch die *Anstrengung* unterscheidet sich signifikant je nach Ausprägung der *Schallbedingung* ($F(3; 57)=4.5$; $p<.01$; partielles $\eta^2=0.19$). In der Post-hoc Testung liegen Unterschiede zwischen den Schallbedingungen *Ruhe* und *Sprache* vor – unter Sprachschall wird die Aufgabe als signifikant anstrengender eingeschätzt ($T(19)=-4.1$; $p<.01$ (BH=.02)). Wiederum werden die Schienenverkehrsschalle im Vergleich zur Ruhebedingung nicht signifikant anders beurteilt (Schiene original: $T(19)=1.1$; $p=.15$ (BH=.03); Schiene-12: $T(19)=0.9$; $p=.20$ (BH=.05)).

Die Konzentrationsmöglichkeit wird, wie sich in einer einfaktoriellen Varianzanalyse mit Messwiederholungen zeigt, ebenfalls je nach Ausprägung des Faktors *Schallbedingung* verschiedenartig eingeschätzt (Greenhouse-Geisser Korrektur: $F(1.8; 34.8)=4.8$; $p=.02$; partielles $\eta^2=0.20$). Auch hier liegt der Unterschied, wie in der Post-hoc Auswertung deutlich wird, in der unterschiedlichen Bewertung von Ruhebedingung und Sprachhintergrundsfall – unter Sprachschall

scheint den Probanden die Konzentration schwerer zu fallen ($T(19)=3.5$; $p<.01$ (BH=.02)). Erneut wird nicht nachweislich zwischen den Schienenverkehrsschall und der Ruhebedingung differenziert (Schiene original: $T(19)=-1.4$; $p=.10$ (BH=.03); Schiene-12: $T(19)=1.0$; $p=.17$ (BH=.05)).

Für die subjektiv empfundene Störung durch den Schall ergibt sich varianzanalytisch ein signifikanter Haupteffekt für die *Schallbedingung* ($F(3; 57)=60.8$; $p<.01$; partielles $\eta^2=0.76$). In der Post-hoc Testung zeigen sich Beurteilungsunterschiede zwischen Ruhe und den drei anderen Schallbedingungen. Alle Schallbedingungen werden von den Probanden als signifikant störender bewertet als die Ruhebedingung (Japanisch: $T(19)=-15.2$; $p<.01$ (BH=.02); Schiene original: $T(19)=-7.8$; $p<.01$ (BH=.03); Schiene-12: $T(19)=-7.4$; $p<.01$ (BH=.05)).

Auch das Lästigkeitsurteil ist durch den Schall beeinflusst ($F(1.9; 37.0)=81.5$; $p<.01$; partielles $\eta^2=0.81$). Die drei Schalle werden statistisch bedeutsam als lästiger eingeschätzt als die Ruhebedingung (post hoc; Japanisch: $T(19)=-14.4$; $p<.01$ (BH=.02); Schiene original: $T(19)=-9.4$; $p<.01$ (BH=.03); Schiene -12: $T(19)=-9.1$; $p<.01$ (BH=.05)).

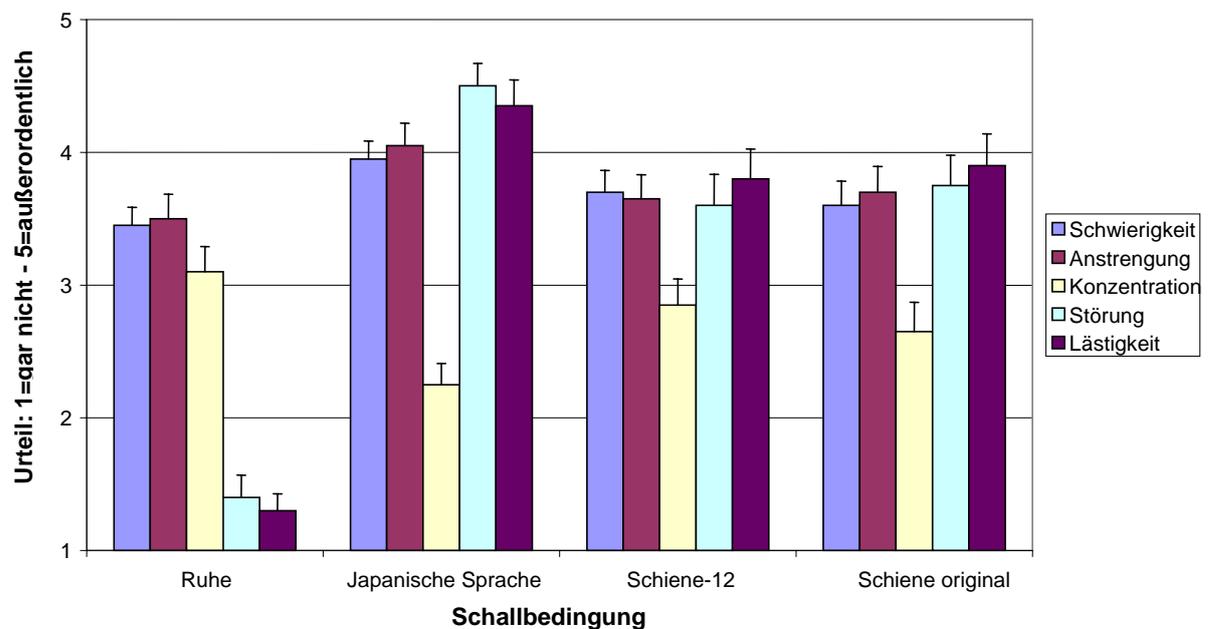


Abbildung 29: Urteile in der Zwischenbefragung im KLTM auf einer Ratingskala von 1=gar nicht bis 5=außerordentlich (Mittelwert und Standardfehler)

3.3.2.4 Diskussion

In dieser Durchführung des KLTM konnte gezeigt werden, dass der leistungsbeeinträchtigende Effekt der Hintergrundsprache wie vermutet unabhängig von ihrer Semantik nachweisbar ist. Im Vergleich zu den vorherigen Testdurchführungen weist der japanische Hintergrundschall eine leistungsbeeinträchtigende Wirkung in einer vergleichbaren Größenordnung wie der deutsche Sprecher (Verschlechterung zur Ruhebedingung um jeweils ca. zehn Prozentpunkte)

auf. Dies ist ein wichtiger Hinweis darauf, dass der KLTM ein geeigneter Test ist, um die Wirkung von spektral-temporal stark variablem Schall auf die Leistung der phonologischen Schleife zu testen. Zugleich kann dies als Hinweis darauf gelten, dass sich der ISE auf diesen Test übertragen lässt und nicht nur anhand der Aufgabe des Serial Recall nachweisbar ist.

Eine Betrachtung der Anzahl richtig gelöster Aufgaben im Verlauf jeder Schallbedingung, aufgeteilt über die Abschnittviertel, ergibt das in Abbildung 30 dargestellte Bild. Die unterhalb der anderen verlaufende Linie für den Hintergrundsprachschaall untermauert die berichteten Unterschiede zwischen der Leistung unter Sprache und allen anderen Schallen. Wiederum ergeben sich keine Hinweise auf eine Passung des Konzepts der reaktiven Anspannungssteigerung, da sich die Leistung über den Versuchsabschnitt hinweg nur wenig ändert und es zu keiner Leistungsbeeinflussung durch eine Anpassung der Anstrengung zu kommen scheint, geschweige denn die Leistung unter einem Schall deutlich über jene unter Ruhe hinausgeht. Auch für eine Orientierungsreaktion oder Habituationsprozesse ergeben sich keine Anhaltspunkte.

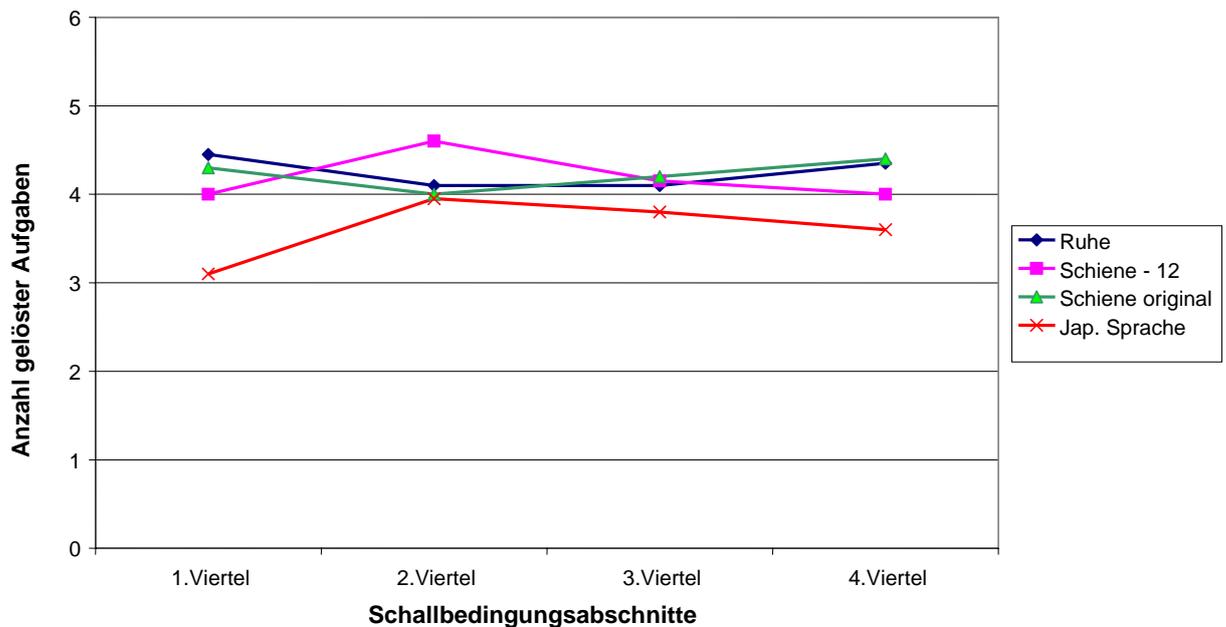


Abbildung 30: Leistungsverlauf im KLTM über die Viertel der Bedingungen getrennt nach Schallen (Mittelwert und Standardfehler); Anzahl gelöster Aufgaben, maximal möglich waren 6 Aufgaben

In diesem Test konnte im Mittel kein leistungsbeeinträchtigender Effekt der Schienenverkehrsgeräusche nachgewiesen werden. Dies kann bei genauerer Betrachtung an einem Mittelungsartefakt liegen, da sich in einem groben deskriptiven Vergleich des Leistungsverlaufs mit dem Pegelzeitverlauf des Schienenverkehrsschalls Leistungseinbrüche an Stellen der lautesten Ereignisse im Schall abzeichnen. In Abbildung 31 sieht man diesen Zusammenhang der Leistung unter den Verkehrsschallen zum Pegelzeitverlauf des Schalldruckpegels. Leistungseinbrüche scheinen im Mittel zeitgleich mit den lautesten Schallereignissen (Linie: *Schiene original*)

stattzufinden. In der Abbildung ist graphisch ein Vergleich des Pegelzeitverlaufs des unmodifizierten Schienenverkehrsgeräuschs (*Schiene original*) mit dem Leistungsverlauf (operationalisiert über die gemessenen Lösungswahrscheinlichkeiten der einzelnen Aufgaben) entsprechend der mittleren Bearbeitungszeiten aller Probanden dargestellt.

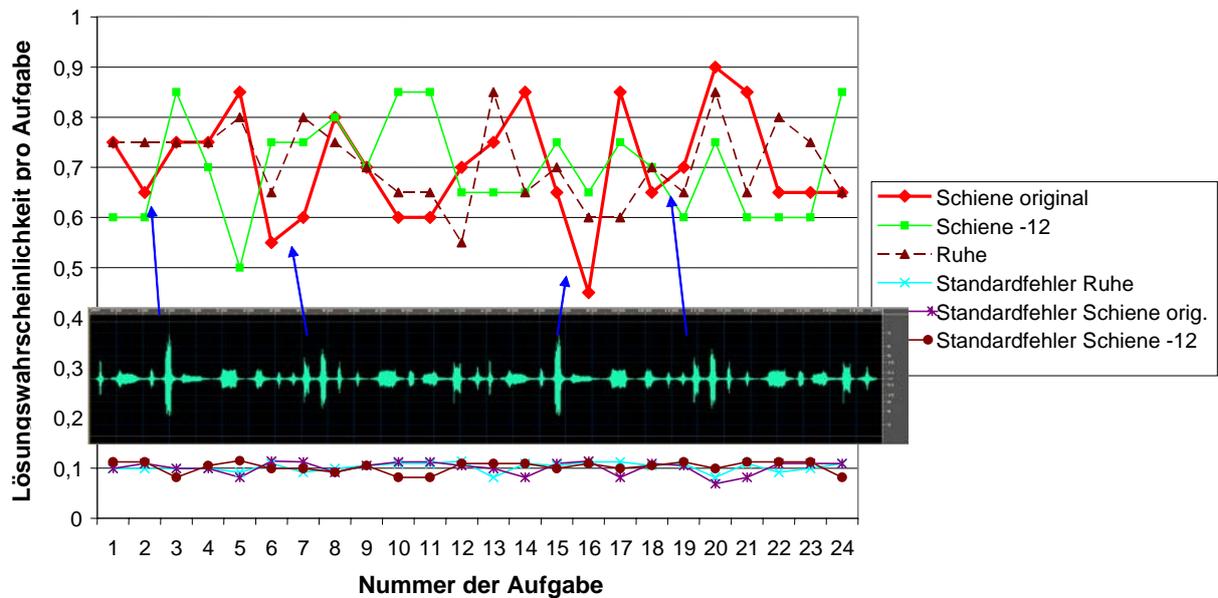


Abbildung 31: Lösungswahrscheinlichkeit und Standardfehler für jede Aufgabe im KLTM im Vergleich zum im Mittel zeitgleichen Amplitudenzeitverlauf des Schallsignals *Schiene original*

Die Kontrollvariable der Lärmempfindlichkeit konnte erneut nicht in die Auswertung einbezogen werden. Da aber in der Zwischenbefragung alle Schalle im Vergleich zur Ruhebedingung als störender und lästiger beurteilt werden, kann *nicht* davon ausgegangen werden, dass der Schienenverkehrslärm *keinen* Einfluss auf die Leistung hat. Der Sprachschall erschwert die Aufgabe aber zusätzlich noch subjektiv und macht sie anstrengender. Ebenso wird die Konzentration subjektiv durch den Sprachschall beeinträchtigt. Aus diesem Experiment ergibt sich somit der Schluss, dass Schienenverkehrslärm auf diesen Test wohl eine störende Wirkung entfalten kann. Dies könnte dann nachgewiesen werden, wenn die Geräusche selektiv an einzelne Aufgaben gekoppelt wären bzw. wenn die kurzfristigen leistungsbeeinträchtigenden Effekte nach einer Zugdurchfahrt im Schallsignal ausgewertet werden könnten. Besonders an hochanspruchsvollen Arbeitsplätzen, an denen bereits *ein* Fehler massive Folgen haben könnte, kann dies bei einer angrenzenden Zugstrecke zu Problemen führen. Während der direkten Belastung durch eine Zugvorbeifahrt droht die Leistung einzubrechen, wenn durch die Aufgabe zentrale exekutive Strukturen gefordert sind.

3.3.3 Grammatical Reasoning mit unverständlichem Sprachschall und Schienenverkehrsgeräuschen

3.3.3.1 Aufgabenstellung

In der ersten Durchführung des Grammatical Reasoning lag ein Speed-Accuracy-Trade-Off vor. Um ein erneutes Auftreten dieses Trade-Off zu verhindern wurde die Instruktion dahingehend modifiziert, dass der Zeitdruck erhöht wurde. Den Probanden wurde nach dem Übungsdurchgang mündlich rückgemeldet, sie sollten bitte im Folgenden schneller arbeiten. Der Test wurde mit den in Tabelle 6 beschriebenen Schallen durchgeführt.

3.3.3.2 Versuchplan, Versuchspersonen und Ablauf

Das Grammatical Reasoning wurde wieder anhand eines einfaktoriellen varianzanalytischen Designs mit Messwiederholung auf dem vierfach gestuften Faktor *Schallbedingung* durchgeführt.

An dem zweistündigen Versuch nahmen 20 Versuchspersonen im Alter zwischen 20 und 38 Jahren (Median 22 Jahre) teil (4 männlich, 16 weiblich), die dafür mit zwei Versuchspersonenstunden im Rahmen ihres Psychologiestudiums oder 10 € entlohnt wurden. Alle meldeten sich freiwillig auf einen Aushang hin und gaben an, über normales Hörvermögen zu verfügen.

Die Versuchspersonen bearbeiteten im Einzelversuch vor einem Pentium I PC in einem Versuchsraum der Universität nach einer schriftlichen Instruktion (Anhang F) und einem Lärmempfindlichkeitsfragebogen (NoiSeQ, Schütte & Marks, 2004) 25 Übungsaufgaben unter Ruhe. Anschließend wurden ihnen unter jedem der Schalle 50 Aufgaben vorgegeben, wonach sie die kurze Zwischenbefragung (vgl. Anhang B) ausfüllten. Die Schalle wurden in einer interindividuell ausbalancierten Reihenfolge über einen Kopfhörer (Sennheiser HD 600; Sony Digital Audio Tape Deck DTC-2E700) dargeboten. Der erste Durchgang fand immer unter Ruhe statt und wurde als Übungsdurchgang nicht ausgewertet.

3.3.3.3 Ergebnisse

Eine einfaktorielle varianzanalytische Auswertung der prozentualen Fehlerraten ergibt trotz der Modifikation der Instruktion keinen signifikanten Haupteffekt ($F(3; 57)=1.0$; $p=.41$) des Faktors *Schallbedingung* (Abbildung 32).

Da die Probanden instruiert worden waren schneller zu arbeiten, müsste sich wenn überhaupt ein Schalleffekt in den Fehlerraten niederschlagen. Aus diesem Grunde wurden trotzdem vier aus den Hypothesen abgeleitete geplante Post-hoc Vergleiche (einseitige Testung, zur theoretischen Begründung vgl. Kapitel 3.3.1) vorgenommen. Diese wurden nicht nach Benjamini-Hochberg korrigiert (Roberts & Russo, 1999). Der Vergleich der Leistung unter der Ruhebedingung mit der unter dem unmodifizierten Schienenverkehrsschall ergibt einen signifikanten Unterschied in der Leistung ($T(19)=1.7$; $p=.05$). Unter dem Schienenverkehrslärm wurden mehr Fehler begangen. Ein Unterschied

in der Leistung zwischen dem Sprachhintergrundschall und der Ruhebedingung lässt sich nicht zeigen ($T(19)=1.0$; $p=.17$). Die Vergleiche zwischen den beiden Verkehrsschällen ($T(19)=1.3$; $p=.11$) sowie zwischen dem modifizierten Schienenlärm und der Ruhebedingung verfehlen ebenfalls die Signifikanz ($T(19)=-0.14$; $p=.45$).

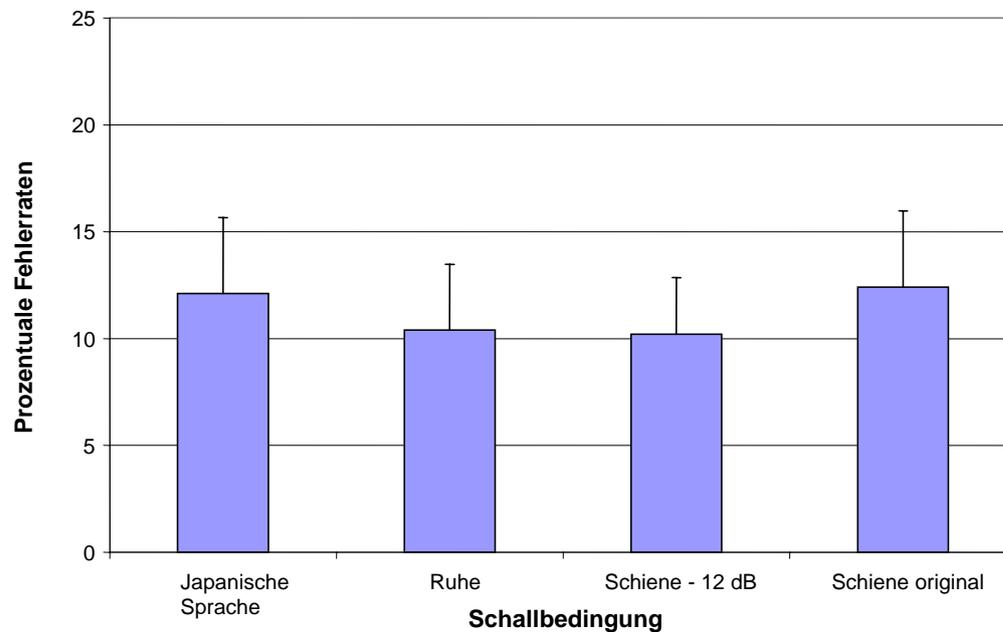


Abbildung 32: Prozentuale Fehlerraten im Grammatical Reasoning unter den Schallbedingungen (Mittelwert und Standardfehler)

Eine varianzanalytische Auswertung der Fehlerraten ergibt einen signifikanten Effekt des Faktors *Versuchsabfolge* ($F(3; 57)=3.1$; $p=.03$; partielles $\eta^2=0.14$). Ein Trendtest nach Page (zur theoretischen Begründung des Trendtests vgl. Kapitel 3.3.1) verdeutlicht einen signifikant fallenden Trend in den Daten. Folglich liegt wieder ein Übungsfortschritt vor ($L_{emp}(20; 4)=530.5 > L_{tab}(20; 4; 5\%)=522$).

Hinsichtlich der Reaktionszeiten bringt eine varianzanalytische Auswertung für den Faktor *Schallbedingung* keinen signifikanten Effekt hervor (einfaktorielle VA: $F(3; 57)=0.6$; $p=.65$). Es zeigt sich auch kein Unterschied in der Antwortlatenz, wenn die Daten gemittelt über die Abschnitte des Experiments nach dem Faktor *Versuchsabfolge* ausgewertet werden ($F(3; 57)=1.8$; $p=.16$).

Die Lärmempfindlichkeit der Probanden (erhoben über den NoiSeQ: vorläufige Normwerte (Schütte & Marks, 2004): $M=43$; $SD=18$; gemessen: $M=47.4$; $SE=2.6$) variiert erneut nicht genügend, um einen Hinweis auf Extremgruppen zu geben, weshalb wiederum auf eine weiterführende kovarianzanalytische Auswertung verzichtet wurde.

Eine deskriptive Auswertung der Urteile der Zwischenbefragungen zeigt die Abbildung 33. Eine varianzanalytische einfaktorielle Auswertung der Daten ergibt für die *Schwierigkeit* in der subjektiven Beurteilung keine Unterschiede durch den Faktor *Schallbedingung* ($F(3; 57)=0.4$; $p=.75$).

Die subjektive Beurteilung der *Anstrengung* unterscheidet sich ebenfalls nicht statistisch bedeutsam durch den Faktor *Schallbedingung* ($F(3; 57)=1.8; p=.15$).

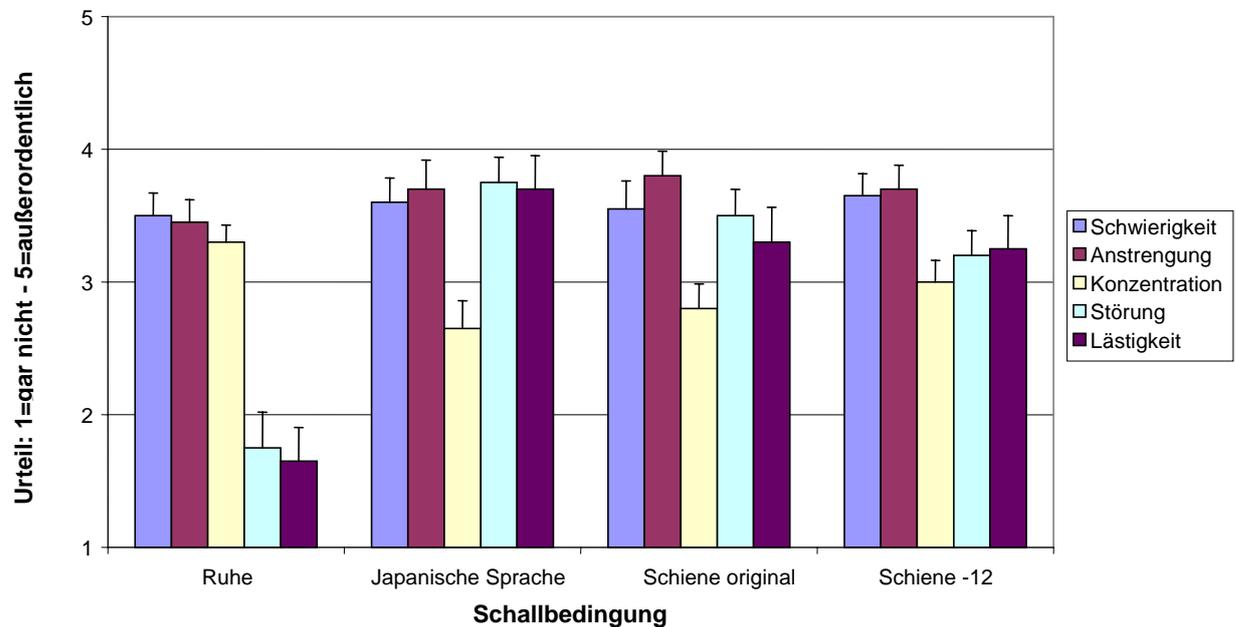


Abbildung 33: Urteile der Zwischenbefragung im Grammatical Reasoning (Ratingskala 1-5; gar nicht bis außerordentlich); Mittelwert und Standardfehler

Die *Konzentration* hingegen zeigt sich durch den Faktor *Schallbedingung* signifikant beeinflusst ($F(3; 57)=5.3; p<.01$; partielles $\eta^2=0.22$). Eine Post-hoc Auswertung der Urteile (alle folgenden Post-hoc Testungen wurden einseitig durchgeführt, zur theoretischen Begründung vgl. Kapitel 3.3.1) über die einzelnen Schalle legt signifikante Unterschiede zwischen der Beurteilung der Ruhebedingung und der aller Schalle nahe (Japanisch: $T(19)=3.6; p<.01$ (BH=.02); Schiene original: $T(19)=2.9; p=.01$ (BH=.03); Schiene -12: $T(19)=2.3; p=.02$ (BH=.05)). Die Konzentrationsmöglichkeit ist somit unter allen drei Schallen im Vergleich zur Ruhebedingung subjektiv eingeschränkt.

Die subjektive *Störung* durch die Schalle wird ebenfalls durch die *Schallbedingung* beeinflusst ($F(3; 57)=21.6; p<.01$; partielles $\eta^2=0.51$). In einer Post-hoc Auswertung mit t-Tests und *Benjamini-Hochberg Korrektur* liegen diese Unterschiede in Form einer größeren subjektiven Störwirkung aller Schalle im Vergleich zur Ruhebedingung vor (Japanisch: $T(19)=-6.2; p<.01$ (BH=.02); Schiene original: $T(19)=-5.7; p<.01$ (BH=.03); Schiene -12: $T(19)=-5.1; p<.01$ (BH=.05)).

Die Auswertung der Lästigkeitsurteile offenbart ebenfalls einen signifikanten Effekt der *Schallbedingung* ($F(3; 57)=15.8; p<.01$; partielles $\eta^2=0.45$). In der Post-hoc Auswertung zeigt sich, dass alle Schalle als lästiger bewertet werden als die Ruhebedingung (Japanisch: $T(19)=-6.8; p<.01$ (BH=.02); Schiene original: $T(19)=-4.5; p<.01$ (BH=.05); Schiene -12: $T(19)=-5.4; p<.01$ (BH=.03)).

3.3.3.4 Diskussion

Aus der vorliegenden Testung lässt sich leider nicht, wie in Kapitel 2 postuliert, die Schlussfolgerung ziehen, dass der leistungsbeeinträchtigende Effekt des Sprachschalls nicht rein von der Semantik des Schalls abhängig ist, sondern von der spektral-temporalen Zusammensetzung des Signals. Der japanische Hintergrundschall führt nicht wie deutsche Sprache im Mittel zu einer Beeinträchtigung der Leistung, nur 8 von 20 Probanden schaffen es nicht, die Leistung aufrechtzuerhalten.

Für den Schienenverkehrslärm jedoch findet sich in diesem Test ein ähnliches Befundmuster wie für den Straßenverkehrschall in Fragestellung II. Auch bei Schienenverkehrsgeräuschen führt die tieffrequente Pegelreduktion im Grammatical Reasoning zu einer geringeren Leistungsbeeinträchtigung im Vergleich zur Ruhebedingung als das unbearbeitete Signal. Dass hier die Effekte in den Fehlerraten und nicht in den Reaktionszeiten ausgewertet wurden liegt an einer feinen Modifikation der mündlichen Instruktion. Den Probanden wurde nach dem Übungsdurchgang in dieser Testdurchführung rückgemeldet, im Folgenden schneller zu arbeiten. Das Ziel, einen erneuten Speed-Accuracy-Trade-Off zu verhindern, kann als erreicht gelten, da die Lösungszeiten diesmal nicht signifikant durch die Schallbedingung beeinflusst wurden. Da aber auch die Varianzanalyse der Fehlerraten die Signifikanz verfehlt, ist mit diesen Ergebnissen sehr vorsichtig umzugehen.

Im Leistungsverlauf über die Abschnitte hinweg zeigt sich, dass der negative Effekt des Sprachschalls und des unbearbeiteten Schienenverkehrsgeräusches vor allem im ersten bzw. den ersten beiden Vierteln jedes Abschnitts zu liegen scheint (Abbildung 34). Auch hier lässt sich erneut nicht auf eine Passung des Konzepts der *reaktiven Anspannungssteigerung* schließen. Es kommt unter dem Lärm nicht zu dem hierfür charakteristischen Anstieg der Leistung sogar über das Niveau unter der Ruhebedingung hinaus, was als Überkompensation der Auswirkungen des Lärms auf die Leistung begründet werden würde. Die anfänglich beeinträchtigte Leistung unter Sprachschall mit der anschließenden Angleichung an das Niveau unter Ruhe könnte für die Effekte einer Orientierungsreaktion mit anschließender Habituation sprechen. Da jedoch auch die Leistung unter Ruhe über den Abschnitt hinweg ansteigt, kann insgesamt von einem Übungsfortschritt ausgegangen werden. Somit kann wieder auf die hohe Komplexität dieser Aufgabe rückgeschlossen werden.

Die Lärmempfindlichkeit konnte aufgrund zu geringer Varianz erneut nicht als Kontrollvariable in die Auswertung einbezogen werden.

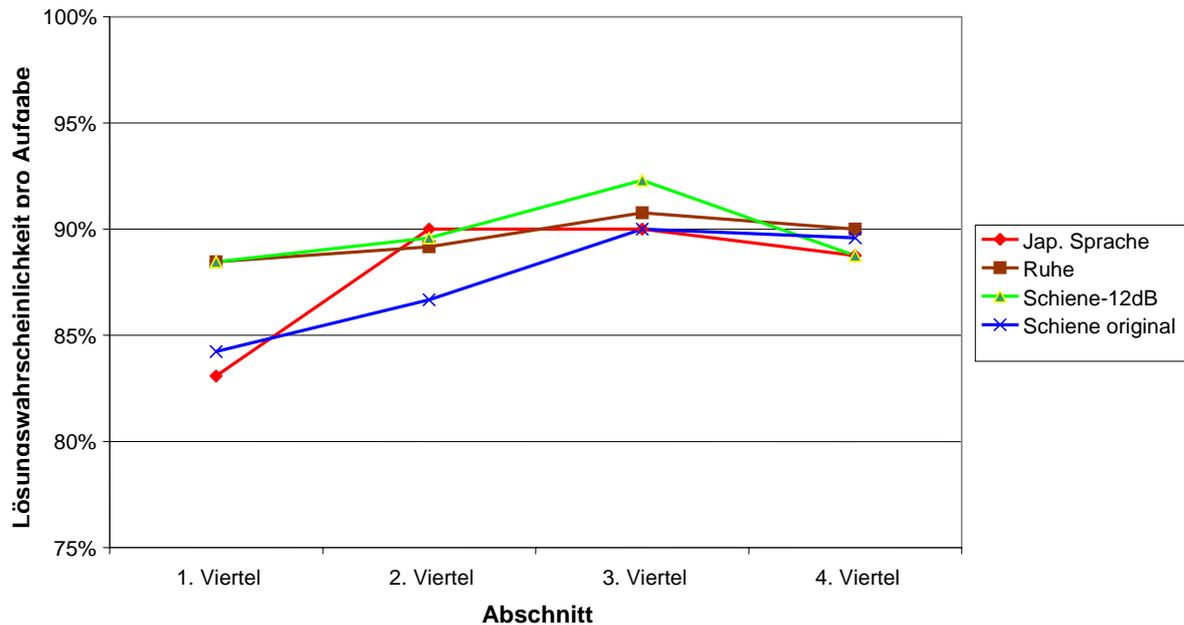


Abbildung 34: Fehlerverlauf im Grammatical Reasoning, getrennt nach Abschnittsvierteln (Mittelwert über alle Probanden)

Obwohl die Leistungsver schlechterung durch den unbearbeiteten Schienenverkehrs schall im Vergleich zur Ruhebedingung statistisch bedeutsam ist, mutet sie mit zwei Prozentpunkten relativ gering an, was aber durch eine Auswertung des Leistungsverlaufs relativiert werden kann. Wie schon beim KLTM, kommt es bei einzelnen lautereren Ereignissen im Schall zu größeren Leistungseinbrüchen. Durch Mittelungsartefakte resultiert allerdings nur eine geringe gemittelte Leistungsver schlechterung. Dies zeigt sich bei einer deskriptiven Betrachtung des Pegelzeitverlaufs des unmodifizierten Schienenverkehrsgeräusches (Linie *Schiene original*) im Vergleich zum Zeitverlauf der Leistung anhand der von allen Probanden benötigten mittleren Bearbeitungszeit (Abbildung 35: rein deskriptive Auswertung). Abhängig von der Aufgabe am Arbeitsplatz kann allerdings bereits *eine* Fehlreaktion während einer Zugvorbeifahrt fatale Folgen nach sich ziehen, weswegen aufgrund dieser Daten nicht von einer geringeren Störf Wirkung der Schienenverkehrsgeräusche im Vergleich zu Straßenverkehrslärm ausgegangen werden kann (wie bereits beim KLTM andiskutiert).

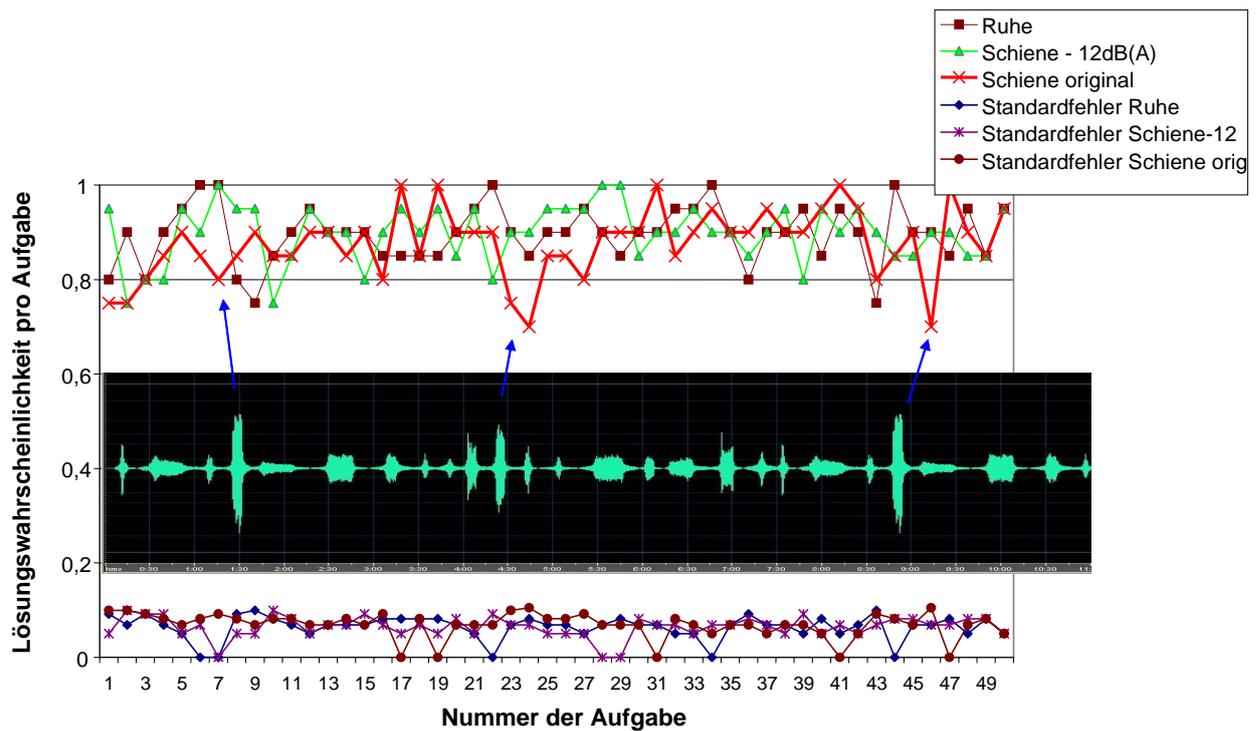


Abbildung 35: Leistungsverlauf (Lösungswahrscheinlichkeit pro Aufgabe, sowie Standardfehler) im Zusammenhang mit den Zugdurchfahrten (Amplitudenzeitverlauf von *Schiene original*) im Grammatical Reasoning (rein deskriptive Auswertung)

Im subjektiven Urteil finden sich diese Effekte wieder – unter allen Schallen wird die Konzentration als fast unmöglich beurteilt. Ebenso werden alle Schalle als lästig und störend im Vergleich zur Ruhebedingung eingeschätzt.

Im Vergleich der exemplarisch beim KLTM für die Schienenverkehrsgeräusche verwendeten Lautsprecheranlage zur Kopfhörerpräsentation im eben vorgestellten Grammatical Reasoning zeigt sich kein auffälliger qualitativer Unterschied. In beiden Tests war ein leistungsbeeinträchtigender Einfluss der Verkehrsschalle nur schwer nachweisbar.

3.3.4 Diskussion der Ergebnisse zu Fragestellung III

Im KLTM bleibt der beeinträchtigende Effekt des Sprachschalls auf die prozentualen Fehlerraten auch unter japanischer Sprache in einer ähnlichen Größenordnung bestehen wie unter einem deutschen Hintergrundsprecher. Es kommt zu einer Verschlechterung der Leistung um jeweils ca. zehn Prozentpunkte im Vergleich zur Ruhebedingung. Beim Grammatical Reasoning jedoch lässt sich kein Einfluss der japanischen Hintergrundsprache auf die Leistung zeigen.

Nur im KLTM konnte somit der leistungsbeeinträchtigende Effekt der Hintergrundsprache unabhängig von ihrem semantischen Gehalt belegt werden. Damit kann diese Störwirkung, wie es auch in der Forschung zum ISE bei der Serial Recall Aufgabe postuliert wird (Klatte, Kilcher, & Hellbrück, 1995), im KLTM auf die spektral-temporale Struktur der Schalle, ihren Changing State Charakter zurückgeführt werden.

Eine Übertragung der Ergebnisse der Fragestellungen I und II auf die zweite wichtige Verkehrsschallquelle neben dem Straßenverkehrslärm, nämlich Schienenverkehrsgeräusche, gelang nur teilweise. Unter Schienenverkehrslärm lässt sich keine Störwirkung in ähnlicher Größenordnung wie unter den Straßenverkehrsschallen finden. Im Grammatical Reasoning ergibt sich nur eine geringe Erhöhung der Fehlerraten (um zwei Prozentpunkte) unter dem nicht modifizierten Schienenverkehrsschall im Vergleich zur Ruhebedingung (und das bei nicht signifikanter Varianzanalyse). Im KLTM ließ sich im Mittel kein signifikanter Einfluss der Verkehrsschalle nachweisen. Die tieffrequente Absenkung des Schienenverkehrssignals bewirkte in beiden Tests keine Erhöhung der Fehlerraten. Unter diesem Schall wurden jeweils ähnlich viele Fehler gemacht wie unter der Ruhebedingung.

Die geringe Übertragbarkeit der leistungsbeeinträchtigenden Wirkung des Straßenverkehrslärms auf Schienenverkehrsschalle kann anhand des Schallpegelverlaufs der Signale (exemplarische Pegelzeitverläufe siehe Anhang A) erklärt werden. Um die beiden Verkehrsschallarten vergleichbar zu machen, wurden Zugdurchfahrten ohne größere Pausen aneinander geschnitten, was aber trotzdem zu einer stärkeren Pausenstruktur bzw. größeren Pegelschwankungen im Zeitverlauf als beim Straßensignal führte. Bei einer Betrachtung des Leistungsverlaufs im KLTM und dem Grammatical Reasoning im Vergleich zum Schallpegelverlauf des unmodifizierten Schienenverkehrsgeräuschs zeigen sich – bereits in einer rein deskriptiven graphischen Auswertung – Hinweise darauf, dass selektive Leistungseinbrüche zeitgleich zur Durchfahrt der lautesten Züge vorhanden sind. Durch die Mittelung über den gesamten Verlauf ist diese kurzfristige Leistungsbeeinflussung allerdings statistisch nicht nachweisbar.

Die Schienenverkehrsschalle wurden bei der Bearbeitung des KLTM über die Lautsprecheranlage dargeboten, während die Probanden beim Grammatical Reasoning mit Kopfhörern arbeiteten. Es zeigt sich durch die Änderung der Darbietung kein bedeutsamer Unterschied. Es kommt beim KLTM nicht zu einer auffällig höheren Störwirkung durch die Schalle als beim Grammatical Reasoning. Obwohl durch den Kopfhörer sehr tiefe Frequenzanteile etwas schwächer übertragen werden, sind Störschalle besser kontrollierbar. Die Darbietung über Kopfhörer

kann somit als adäquate Methode gelten, die Kontrollierbarkeit aber auch die Originalität der Präsentation von Störschallen während kognitiver Tests zu gewährleisten.

4 Zusammenfassende Diskussion

Gegenstand der vorliegenden Arbeit ist die Spezifikation leistungsbeeinträchtigender Parameter von Hintergrundlärm um gezielte, über eine reine Pegelreduktion hinausgehende, Modifikationsmaßnahmen für Störschalle ableiten zu können. Die Vorhersagen hierfür wurden kognitionspsychologisch begründet.

Die am häufigsten diskutierten Umweltlärmquellen sind sicherlich Verkehrsgeräusche. Unterschiede und Besonderheiten zwischen den Lärmquellen Straße, Schiene und Flugverkehr wurden ebenso dargestellt, wie Probleme bei der Messung von Lärm und der Beurteilung seiner Auswirkungen, sowie gesetzliche Vorgaben von Grenzwerten. Die Lästigkeitsbewertung von Verkehrslärm ist sowohl von der Geräuschquelle selbst, als auch von personenbezogenen und situativen Faktoren abhängig. Zudem sind Faktoren der Zeitstruktur (Kreisverkehr versus Ampelkreuzung) und der Frequenzzusammensetzung neben der Lautstärke des Geräusches von Bedeutung für die Beurteilung der Lästigkeit eines Verkehrsschalls. Im Allgemeinen werden Geräusche als unangenehmer empfunden, die vor allem tiefe Frequenzen enthalten im Gegensatz zu mittel- oder etwas höherfrequenten. Zudem ist dieser Frequenzbereich schwer abzuschirmen.

Es ist davon auszugehen, dass sich Lärm auf kognitive Leistungen auswirkt; hierbei muss aber eine Interaktion mit der Lästigkeit des Geräusches bedacht werden. Eine Lärmquelle, die im Zusammenhang mit kognitiven Funktionen sehr gut untersucht ist, ist irrelevante Hintergrundsprache. Der so genannte *Irrelevant Sound Effect* wurde in diesem Zusammenhang ausführlich vorgestellt. Er bezeichnet die Störwirkung eines spektral-temporal stark variablen (*Changing State*) Hintergrundschalls auf das serielle Memorieren von Inhalten. Forschungsergebnisse deuten darauf hin, dass dieser Effekt nicht durch Habituation beeinflussbar, nicht von der Verständlichkeit der Sprache oder ihrer Lautstärke abhängig und nicht auf Sprachschalle beschränkt ist. Spektral-temporal stark variable Schalle wirken beeinträchtigend auf ein Speichersystem des Arbeitsgedächtnisses. Der Irrelevant Sound Effect wurde als mögliche Erklärung von Schallwirkungen auf ein Subsystem eines Arbeitsgedächtnismodells vorgestellt, welches zuvor dargelegt wurde: das Working Memory Modell von Baddeley (2000a). Besonderer Wert wurde bei der Darstellung des Working Memory Modells auf die beiden Systeme der phonologischen Schleife und der zentralen Exekutive gelegt. Modelltheoretisch wird der Irrelevant Sound Effect über eine Interferenz der gehörten irrelevanten Sprache (oder eines anderen spektral-temporal stark variablen Schalls) mit den zu memorierenden Items in der phonologischen Schleife erklärt, die dort über einen Rehearsalprozess aufrechterhalten werden.

Das andere hier interessante Subsystem des Working Memory Modells, die zentrale Exekutive, hat laut Baddeley die Aufgaben Hilffsysteme zu koordinieren, Automatismen zu hemmen, Aufmerksamkeit selektiv zu steuern, Inhalte des Langzeitgedächtnisses zu aktivieren und zu integrieren sowie allgemein Planungs-, Integrations- und Kontrollaufgaben zu übernehmen. Um mögliche Schallwirkungen auf dieses System zu erklären, wird häufig auf das Konzept des Arousal zurückgegriffen, das anschließend vorgestellt wurde.

Der Aufmerksamkeitsbegriff hängt mit dem Arousal und der Funktion der zentralen Exekutive im Working Memory Modell zusammen. Bei der Aufmerksamkeit können drei Funktionen unterschieden werden (Neumann, 1996): der Kapazitätsaspekt, die Selektionsfunktion und der Tätigkeitsaspekt. Letzterer, nämlich die willentliche Steuerung der Aufmerksamkeit, ist bei Baddeley als Funktion der zentralen Exekutive beschrieben. Die Aspekte der Kapazität und der Selektivität hingegen hängen mit dem Arousal zusammen, wie in Kapitel 1.4 dargelegt wurde. Arousal, welches als unspezifisches kortikales Erregungsniveau verstanden wird, kann durch Lärm erhöht werden. Die Auswirkungen und die Erklärungen dieser Erhöhung differieren. Die Aufmerksamkeit wird in ihrer Kapazität als begrenzt gesehen und es kann durch Lärm zu einer Umverteilung der vorhandenen Ressourcen kommen. Es ist davon auszugehen, dass sich unter steigendem Arousal die Selektivität des Aufmerksamkeitsfokus erhöht (Easterbrook, 1959) und zudem seine Ablenkbarkeit steigt (Kahneman, 1973). Zu bedenken ist jedoch eine mögliche Konfundierung der Leistungsbeeinflussung durch Lärm direkt über das Arousal, mit einer über die Lästigkeit des Schalls und hierüber entstehender Ärgerreaktionen (Näätänen, 1975).

Aus den vorgestellten Theorien zum Arbeitsgedächtnis und zum Arousal Konzept ist abzuleiten, dass eine Exposition mit einem lauten Verkehrsschall mit dominanten tiefen Frequenzanteilen, vermittelt über eine Erhöhung des Arousal zu einer höheren Selektivität und Ablenkbarkeit des Aufmerksamkeitsfokus und damit zu einer Leistungseinschränkung der zentralen Exekutive führen kann. Ein spektral-temporal stark variabler Schall (mit Changing State Charakter: z.B. Sprachschall oder ein Verkehrsschall mit starker Pausenstruktur) hingegen sollte unabhängig von seiner Lautstärke eine Störung des Memorierens von Inhalten in der phonologischen Schleife hervorrufen.

Diese Hypothesen empirisch zu prüfen, um spezifische Lärmmodifikationen ableiten zu können, war Ziel des empirischen Teils. Hierfür geeignete Tests wurden im Anschluss dargelegt. Zunächst wurden Forschungen zu zwei Testkategorien, die Anforderungen an die phonologische Schleife und an die zentrale Exekutive stellen, vorgestellt: Grammatikalisches Schlussfolgern und Kopfrechentests, zu denen der nachfolgend verwendete KLT zählt. Dieser stellt jedoch durch seine komplexe Struktur, die geforderte Speicherung und Integration verschiedener Informationen sowie Anwendung der Rechenregel Anforderungen an die beiden Subsysteme, die über reine Kopfrechentests weit hinausgehen. Im Anschluss wurden Forschungsergebnisse zum Stroop Test berichtet, welcher Anforderungen an das System der zentralen Exekutive stellt. Forschungsergebnisse zu Schalleffekten auf diese Tests sind in der Literatur leider nur spärlich vorhanden und häufig ohne genügende Kontrolle von Moderatorvariablen wie der Extraversion, der Lärmempfindlichkeit, Lästigkeitsurteilen oder der Leistungsmotivation entstanden. Zudem deuten manche Ergebnisse in unterschiedliche Richtungen, was auch an einer Variation der Aufgabenschwierigkeit und den daraus resultierenden unterschiedlichen Anforderungen an die Subsysteme des Arbeitsgedächtnisses liegen mag. Ein Test der nur die phonologische Schleife beansprucht, war bereits im Zusammenhang mit dem Irrelevant Sound Effect vorgestellt worden: der so genannte Serial Recall.

Die kognitionspsychologischen Erkenntnisse zu diesen Testkategorien wurden sehr ausführlich vorgestellt. Es ist wichtig, genau ableiten zu können, welche Subsysteme des Working Memory Modells ein Test erfordert, um eine spezifische leistungsbeeinflussende Wirkung eines bestimmten Schallparameters vorhersagen zu können.

Aus diesen theoretischen Grundlagen wurden drei Fragestellungen bezüglich der Auswirkungen von Hintergrundschallen auf kognitive Leistungen abgeleitet und auf ihre empirische Belegbarkeit hin getestet. Im Folgenden werden die wichtigsten Ergebnisse der Fragestellungen im Hinblick auf die Untersuchungsfragen diskutiert, bevor abschließend modelltheoretische Implikationen erwogen werden.

Die differenzierte Störwirkung der beiden Schallarten Verkehrslärm und Sprachschall, die sich in Fragestellung I zeigen ließ, war aus den theoretischen Grundlagen abgeleitet worden. Sprachschall sollte laut dem Working Memory Modell von Baddeley die Arbeit der phonologischen Schleife beeinträchtigen (Serial Recall, KLTM). Lauter tieffrequenter Verkehrsschall sollte über eine Arousalerhöhung zu einer verstärkten Ablenkbarkeit und Einengung des Aufmerksamkeitsfokus führen und damit die Arbeit der zentralen Exekutive behindern (KLTM, Stroop Test). Für beides gelang es empirische Belege zu finden (beim KLTM ließen sich signifikante Unterschiede in der Leistung unter dem lautesten Verkehrslärm und unter Ruhe in der zweiten und dritten Durchführung in Fragestellung II zeigen). Die selektive Störwirkung der Verkehrsschalle ab einer gewissen Lautstärke auf zentral exekutive Funktionen, nicht jedoch auf Tätigkeiten der phonologischen Schleife, hätten zu einer Konfusion in der Interpretation der Daten führen können. Wären die kognitionspsychologischen Grundlagen außer Acht gelassen worden, wäre unerklärlicherweise nur auf gewisse Aufgaben eine beeinträchtigende Wirkung von Verkehrsschall nachweisbar gewesen. Weiß man jedoch um die modelltheoretischen Hintergründe, ist es nicht verwunderlich keine störende Wirkung eines gleichmäßigen, aber lauten Schalls auf eine Aufgabe zu finden, die selektiv die phonologische Schleife beansprucht. Hieraus lässt sich folgern, dass eine kognitionspsychologisch basierte Lärmwirkungsforschung für mehr Klarheit und Ordnung in der Erforschung von leistungsbeeinträchtigenden Lärmeffekten sorgen kann, woraus dann spezifische Empfehlungen zur Modifikation störender Schalle ableitbar sein sollten. Modifikationsableitungen dieser Art wurden in der zweiten Fragestellung überprüft.

Aus der kognitionspsychologischen Theorie zu Arousal wurde abgeleitet, dass eine tieffrequente Pegelreduktion von Verkehrsschallen in Tests, welche die zentrale Exekutive fordern, zu einer Senkung der Leistungsbeeinträchtigung führen dürfte (KLTM, Stroop Test, Grammatical Reasoning). Diese Strategie erwies sich in der zweiten Fragestellung als zielführend. Durch die Simulationen einer Modifikation von Motorengeräuschen und ihrer empirischen Testung können Empfehlungen an die Fahrzeugindustrie abgeleitet werden. Aus dem vorliegenden Fall kann geschlussfolgert werden, dass es für den Schutz kognitiv arbeitender exponierter Personen sinnvoller ist, die Emissionen in spezifischer Richtung zu modifizieren, anstatt nur global zu versuchen den Pegel zu reduzieren. Würde dieser nämlich A-bewertet gemessen werden, könnte eine solche Reduktion, wie die empirischen Daten nun nahe legen, eventuell *keine* adäquate Verringerung der

beeinträchtigenden Wirkung erbringen, da die tiefen Frequenzanteile und die subjektive Lautheit des Geräusches entscheidende Faktoren sind. Im vorliegenden Fall war der A-bewertete Mittelungspegel des originalen und des in Lautheit (sone) und tieffrequenten Anteilen reduzierten Geräusches sogar identisch mit jeweils 70 dB(A). Da nun von diesen beiden Geräuschen aber nur das mit dem hohen tieffrequenten Anteil eine beeinträchtigende Wirkung erzielte, verdeutlicht, dass eine reine Pegelreduktion kein adäquates Modifikationsmittel bei dieser Art kognitiver Arbeit ist.

In den vorgestellten Experimenten erwies sich eine Pegelsenkung der Verkehrslärmgeräusche um zwölf Dezibel unter 500 Hz als probates Mittel, um die Leistungsbeeinträchtigung in Tests, die zentral exekutive Ansprüche stellen (KLTM, Grammatical Reasoning), zu reduzieren. Aus diesem Grund kann folgende Empfehlung an die Fahrzeugindustrie gegeben werden: eine Modifikation von Autogeräuschen um sechs Dezibel unter 500 Hz durch schalldämmende Auskleidung des Motorraums und ein verbessertes Motormanagement (wie Klingenberg (1991) vorschlägt) wäre bereits ein guter Schritt auf dem Weg zu einer Reduktion der Störwirkungen von Straßenverkehrsgeräuschen.

Der Nachweis einer geringeren Beeinträchtigung durch die Reduktion um zwölf Dezibel gelang trotz einer so kurzen Arbeitszeit und der künstlichen Situation des Experiments, bei einer eher kleinen Anzahl an Probanden (N=18-24). Daher kann davon ausgegangen werden, dass eine Pegelabsenkung um sechs Dezibel in realistischen Arbeitsumwelten und für lange Expositionszeiten bereits zu einer sinkenden Beeinträchtigung kognitiver Arbeit durch Verkehrslärm beitragen würde.

Darüber hinaus scheinen die gefundenen Effekte unabhängig von der Darbietungsart (über Kopfhörer oder Lautsprecheranlage mit Subwoofer) zu sein, wie die Replikation des KLTM mit einer Lautsprecheranlage in dieser Fragestellung zeigt. Aus den Ergebnissen lässt sich zudem der Schluss ziehen, dass die gefundenen Effekte des KLTM stabil nachweisbar sind. In dieser dritten Durchführung des KLTM findet sich wiederum der leistungsbeeinträchtigende Effekt des Sprachschalls gegenüber der Ruhebedingung, aber auch ein solcher des lautesten Verkehrsschalls. Eine beeinträchtigende Wirkung des abgesenkten Verkehrsschalls auf die Leistung hingegen ist wiederum nicht nachweisbar.

Aus den Ergebnissen der dritten Fragestellung kann man folgern, dass eine Übertragung des ISE auf andere Tests als den Serial Recall wohl möglich ist. Auch bei dem vorgestellten KLTM konnte eine störende Wirkung des Sprachschalls unabhängig von seiner semantischen Verständlichkeit nachgewiesen werden. Um dieser Ausweitung des ISE auf andere Aufgabenstellungen weiter nach zu gehen, müsste in zukünftigen Arbeiten noch untersucht werden, ob sich eine Störwirkung auch durch Stakkatomusik oder Tonfolgen nachweisen ließe.

Die Auswirkungen von Schienenverkehrslärm auf kognitive Leistungen konnten anhand der vorliegenden Experimente noch nicht so detailliert spezifiziert werden, wie dies für Straßenverkehrslärm gelang (KLTM, Grammatical Reasoning). Da der Schienenverkehr aber neben Straßenverkehr sicherlich die bedeutendste Lärmquelle darstellt, sollten auch hieraus Schlussfolgerungen gezogen werden. Durch die stärkere Pausenstruktur der Geräusche konnte – wahrscheinlich bedingt durch ein Mittelungsartefakt – keine signifikante Beeinträchtigung in Tests, die zentral exekutive Ansprüche stellen, nachgewiesen werden. Eine Betrachtung des Leistungsverlaufs während der Bearbeitung legt aber nahe, dass eine Beeinträchtigung durchaus vorliegen könnte,

allerdings von kurzfristigerer Dauer ist als beim gleichmäßigeren Straßenverkehrsgeräusch. Um den Nachweis einer störenden Wirkung, welche aus den vorliegenden Daten nur vermutet werden kann, sowie der positiven Auswirkungen der tieffrequenten Reduktion des Signals erbringen zu können, ist noch weitere Forschung notwendig. Ebenso empfiehlt es sich zur dritten Verkehrslärmquelle, dem Flugverkehrslärm, der in dieser Arbeit vorerst unberücksichtigt blieb.

Praktisch wäre aus diesen Ergebnissen der Schluss zu ziehen, dass für jede geistige Tätigkeit spezifische Modifikationen von Hintergrundschallen verfolgt werden müssten, was für eine Umsetzung im Alltag allerdings nicht praktikabel ist. Wenn jedoch beide vorgestellten Arten der Leistungsbeeinträchtigung (Wirkung spektral-temporal stark variabler Schalle auf die phonologische Schleife und lauter tieffrequenter Schalle auf die zentrale Exekutive) beachtet werden, und somit das Ziel verfolgt wird, sowohl die Lautheit und die tieffrequenten Anteile als auch die spektral-temporale Struktur von Störgeräuschen zu reduzieren, wäre kognitiv arbeitenden lärmexponierten Personen in Lern- und Arbeitsumwelten bereits viel geholfen.

In allen Fragestellungen wurden Variablen, die sich in der Literatur als Moderatoren von Lärmwirkungen herausgestellt hatten, wie Extraversion (Amelang & Ullwer, 1991; Standing, Lynn, & Moxness, 1990), Lärmempfindlichkeit (Ellermeier, Eigenstetter, & Zimmer, 2001) und Leistungsmotivation (Schuler & Prochaska, 2000) mit erhoben. Diese variierten allerdings in keinem der Tests ausreichend, um auf Extremgruppenbildung hinzuweisen und wurden aus diesem Grunde auch nicht als Kovariablen in die Auswertungen aufgenommen. Als wichtig erwiesen sich hingegen die subjektiven Urteile der Probanden über die Schalle, die empfundene Schwierigkeit der Aufgabe sowie die Konzentrationsmöglichkeiten. Zum Teil finden sich in diesen Daten die Leistungsunterschiede wieder. Teilweise zeigen sich jedoch auch Divergenzen insofern, als dass der laute Verkehrslärm meist als störender und lästiger sowie als hinderlicher für die Konzentration eingestuft wurde als seine leiseren Varianten. Das zeigte sich allerdings so teilweise – vielleicht noch – nicht in den Leistungsdaten.

Aus den vorhandenen Daten kann insgesamt folgender Schluss abgeleitet werden: die Erklärung der Leistungsbeeinflussung durch eine Arousalsteigerung über die oben genannten Mechanismen der Beeinflussung des Aufmerksamkeitsfokus (Easterbrook, 1959; Kahneman, 1973) ist geeignet zur Erklärung der Schalleffekte. Da die Belästigung und somit die ablenkenden Kognitionen nicht immer eng mit der Leistungsbeeinträchtigung zusammenhängen, kann nicht davon ausgegangen werden, dass die Leistung *nur* sekundär über ablenkende Ärgerkognitionen bezüglich der Lärmbelästigung beeinträchtigt wird, wie von Näätänen (1975) vermutet.

In Bezug auf das in Kapitel 1.2 vorgestellte Arbeitsgedächtnismodell lässt sich diskutieren, dass die Versuchsplanung anhand des Working Memory Modells von Baddeley (2000a) zu Ergebnissen führte, aus denen konkrete Schallschutzempfehlungen ableitbar sind. Die Bedeutsamkeit der spektral-temporalen Struktur konnte anhand dieses Modells vorhergesagt werden, was für den Bereich der Wirkungen der Lautheit und der tiefen Frequenzanteile von Lärm nicht gilt. Die Konzeption der zentralen Exekutive ist hierfür zu unklar in ihrem Aufbau und müsste um die Energiekomponente

des Arousal erweitert werden. Somit erscheint die in dieser Arbeit verwendete Verknüpfung des Working Memory Modells von Baddeley (2000a) mit einer Energiekomponente im System der zentralen Exekutive, angelehnt an Arousal-Konzepte in der Tradition von Easterbrook (1959) und Kahneman (1973), als sinnvolle Variante einer modelltheoretischen Fundierung kognitionspsychologischer Lärmwirkungsforschung.

Zusammenfassend kann man folgern, dass es gelang nachzuweisen, dass irrelevanter Sprachschall (beim KLTM sogar unabhängig von seiner Semantik), wie aus der modelltheoretischen Fundierung abgeleitet, negative Auswirkungen auf die Leistung in die phonologische Schleife fordernden Aufgaben hat (Serial Recall, KLTM und Grammatical Reasoning). Zudem konnte ein leistungsbeeinträchtigender Effekt von Verkehrslärm einer bestimmten Lautstärke (Mittelungspegel von 70 dB(A)) auf Aufgaben erbracht werden, die zentral exekutive Anforderungen stellen (Stroop Test, KLTM und Grammatical Reasoning). Dies kann über eine Erhöhung des Arousal erklärt werden. Die andere mögliche Deutung dieser Leistungseinbußen mit dem Konzept der reaktiven Anspannungssteigerung (Düker, 1963) ist, wie eine Auswertung des jeweiligen Leistungsverlaufs zeigte, nicht passend, und musste wiederholt verworfen werden. Einer genaueren Untersuchung in weiterführenden Arbeiten bedürften sicherlich noch die z. T. fraglich vorhandenen Orientierungsreaktions- bzw. Habituationseffekte. Bei der Betrachtung des Leistungsverlaufs unter den jeweiligen Schallabschnitten ergaben sich teilweise Hinweise auf solche Prozesse, teilweise aber auch nicht. Zu spezifizieren, wann, bei welchen Aufgaben und unter welchen Bedingungen es zu einer Habituation an die Schallwirkungen kommen kann, wäre eine sehr spannende Fragestellung für zukünftige Arbeiten.

Die leistungsbeeinträchtigende Wirkung von Sprachschall und Verkehrslärm unterscheidet sich somit, wie theoretisch abgeleitet, voneinander, und ist stabil in bestimmten Tests nachweisbar (vgl. gelungene Replikationen der Ergebnisse des KLTM in Fragestellung II und III). Hier wäre weitere Forschung interessant, um genauer die Unterschiede in den divergenten Wirkmechanismen herauszuarbeiten, und jeweils sensible Tests besser voneinander abgrenzen zu können.

Die Abhängigkeit der Leistungsbeeinträchtigungen von akustischen Eigenschaften wie dem Anteil an tiefen Frequenzen im Schallsignal ermöglicht gezielte technische Maßnahmen, beispielsweise Änderungen am Motormanagement oder Schalldämmungsmaßnahmen (Klingenberg, 1991). Wenn damit eine tieffrequente Pegelreduktion von Automobilgeräuschen erreicht werden kann, wäre dies bereits eine erhebliche Entlastung für kognitiv arbeitende lärmexponierte Personen. Selbst wenn sich die gefundenen Effekte zumeist in einer geringen Größenordnung bewegten, können in gewissen Arbeitsumfeldern bereits ein oder zwei Fehler während einer Zug- oder LKW-Vorbeifahrt zu kostspieligen Leistungseinbußen führen; man denke hierbei z. B. an kontrollierende oder überwachende Tätigkeiten. Auch in diesem Bereich wäre somit noch weitere Forschung sinnvoll, um die leistungsbeeinträchtigende Wirkung von stark tieffrequenten Geräuschen weiter zu spezifizieren, und genaue hierfür sensitive Aufgabenbereiche zu definieren.

Literaturverzeichnis

- Advisory Group for Aerospace Research and Development. (1989). *Human performance assessment methods*. Neuilly-sur-Seine, France: Author.
- Amelang, M., & Ullwer, U. (1991). Ansatz und Ergebnisse einer (fast) umfassenden Überprüfung von Eysencks Extraversionstheorie. *Psychologische Beiträge*, 33(1-2), 23-46.
- Baddeley, A. (1996). Exploring the central executive. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 49A(1), 5-28.
- Baddeley, A. (2000a). The episodic buffer: a new component of working memory? *Trend in Cognitive Sciences*, 4(11), 417-423.
- Baddeley, A. (2000b). The phonological loop and irrelevant speech effect: Some comments on Neath (2000). *Psychonomic Bulletin & Review*, 7(3), 544-549.
- Baddeley, A., & Salamé, P. (1986). The Unattended Speech Effect: Perception or Memory? *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12(4), 525-529.
- Baddeley, A. D. (1968). A three-minute reasoning test based on grammatical transformation. *Psychonomic Science*, 10, 341-342.
- Baddeley, A. D. (1976). *The psychology of memory*. New York: Basic Books.
- Baddeley, A. D. (1986). *Working memory*. New York: Clarendon Press/Oxford University Press.
- Baddeley, A. D. (1993). Working memory or working attention? In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent*. New York: Clarendon Press/Oxford University Press.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *Recent advances in learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.
- 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.
- Banbury, S., & Berry, D. C. (1998). Disruption of office-related tasks by speech and office noise. *British Journal of Psychology*, 89, 499-517.
- Bartlett, F. C. (1934). *The problem of noise*. London: Cambridge University Press.
- Beckmann, J. (1999). Leistungssteigerung durch leistungsbeeinträchtigende Bedingungen. In L. Tent (Ed.), *Heinrich Düker. Ein Leben für die Psychologie und für eine gerechte Gesellschaft* (Vol. 1, pp. 287-303). Lengerich: Papst.
- Belojevic, G., Slepcevic, V., & Jacolvjevic, B. (2001). Mental Performance in Noise: The Role of Introversion. *Journal of Environmental Psychology*, 21(2), 209-213.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, Methodological*, 57, 289-300.
- Benjamini, Y., & Hochberg, Y. (2000). On the Adaptive Control of the False Discovery Rate in Multiple Testing With Independent Statistics. *Journal of Educational and Behavioral Statistics*, 25(1), 60-83.

- Benjamini, Y., & Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics*, 29, 1165-1188.
- Berglund, B., Hassmen, P., & Job, S. R. F. (1996). Sources and effects of low frequency noise. *Journal of the Acoustical Society of America*, 99(5), 2985-3002.
- Biewald, C., & Schumann-Hengsteler, R. (1996). *Arbeitsgedächtnisprozesse bei mentaler Addition und Multiplikation: die phonologische Schleife* (Eichstätter Berichte zur Entwicklungs- und Pädagogischen Psychologie). Eichstätt: Universität, Entwicklungs- und Pädagogische Psychologie.
- Biewald, C., & Schumann-Hengsteler, R. (1997). *Die zentrale Exekutive bei mentaler Addition und Multiplikation* (Eichstätter Berichte zur Entwicklungs- und Pädagogischen Psychologie). Eichstätt: Universität, Entwicklungs- und Pädagogische Psychologie.
- Birbaumer, N. (1975). *Physiologische Psychologie*. Heidelberg: Springer.
- Birbaumer, N., & Schmidt, R. F. (1996). *Biologische Psychologie*. Heidelberg: Springer.
- Bisping, R. (2005). Psychometrische und akustische Analyse von Kraftfahrzeug-Geräuschen. In *Bericht zum 51. Arbeitswissenschaftlichen Kongress der Gesellschaft für Arbeitswissenschaft* (pp. 659-664). Dortmund: GfA-Press.
- Boucsein, W., & Ottmann, W. (1996). Psychophysiological stress effects from the combination of night-shift work and noise. *Biological Psychology*, 42, 301-322.
- Broadbent, D. E. (1953). Noise, paced performance and vigilance tasks. *British Journal of Psychology*, 44, 295-303.
- Broadbent, D. E. (1957). Effects of noises of high and low frequency on behaviour. *Ergonomics*, 1, 21-29.
- Broadbent, D. E. (1958a). Effect of noise on an "intellectual" task. *The Journal of the Acoustical Society of America*, 30(9), 824-827.
- Broadbent, D. E. (1958b). *Perception and communication*. London: Pergamon Press.
- Broadbent, D. E. (1981). Effects of moderate levels of noise on human performance. In J. V. Tobias & E. D. Schubert (Eds.), *Hearing Research and Theory* (pp. 165-187). New York: Academic Press.
- Buchner, A., Irmen, L., & Erdfelder, E. (1996). On the Irrelevance of Semantic Information for the "Irrelevant Speech" Effect. *The Quarterly Journal of Experimental Psychology*, 49A(3), 765-779.
- Cohen, J. D., Mac Whinney, B., Flatt, M., & Provost, J. (1993). Psyscope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments and Computers*, 25(2), 257-271.
- Colle, H. A., & Welsh, A. (1976). Acoustic Masking in Primary Memory. *Journal of Verbal Learning and Verbal Behavior*, 15, 17-31.
- Cowan, N. (1989). A reply to Miles, Madden and Jones (1989): Mistakes and other flaws in the challenge to the cross-modal Stroop effect. *Perception and Psychophysics*, 45(1), 82-84.
- 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., & Barron, A. (1987). Cross-modal, auditory-visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics*, *41*, 393-401.
- Dalrymple-Alford, E. C., & Azkoul, J. (1972). The locus of interference in the Stroop and related tasks. *Perception & Psychophysics*, *11*, 385-388.
- DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology*, *16*(3), 353-386.
- Donlan, C. (Ed.). (1998). *The development of mathematical skills*. Hove, England: Psychology Press/Taylor and Francis (UK).
- Dornic, S. (1982). Semantic interference as a function of arousal. *Reports from the Department of Psychology, U. Stockholm*, *590*.
- Dornic, S. (1990). Noise and information processing: Findings, trends, and issues. *Reports from the Department of Psychology, U. Stockholm*, *715*, 30.
- Dornic, S., & Laaksonen, T. (1988). A repetitive reasoning task based on simple spatial relations. *Perceptual and Motor Skills*, *66*(2), 589-590.
- Düker, H. (1963). Über reaktive Anspannungssteigerung. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *10*(1), 46-72.
- Düker, H., & Lienert, G. A. (1965). *Konzentrations-Leistungs-Test K-L-T (2. Auflage)*. Göttingen: Hogrefe.
- Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*, 183-201.
- Effler, M. (1978). Farbdistante versus farbkongruente Farbwörter. Ihr Einfluss auf die Benennungszeiten beim Stroop-Test. *Psychologische Beiträge*, *20*(3), 345-359.
- Eggert, D. (1974). *Eysenck-Persönlichkeits-Inventar: E-P-I: Handanweisung für die Durchführung und Auswertung*. Oxford, England: Verlag für Psychologie.
- Ellermeier, W., Eigenstetter, M., & Zimmer, K. (2001). Psychoacoustic correlates of individual noise sensitivity. *Acoustical Society of America*, *109*(4), 1464-1473.
- Ellermeier, W., & Hellbrück, J. (1998). Is level irrelevant in "irrelevant speech"? Effects of loudness, signal-to-noise ratio, and binaural unmasking. *Journal of Experimental Psychology Human Perception and Performance*, *24*(5), 1406-1414.
- Elliott, E. M., & Cowan, N. (2001). Habituation to auditory distractors in a cross-modal, color-word interference task. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *27*, 654-667.
- Elliott, E. M., Cowan, N., & Valle-Inclan, F. (1998). The nature of cross-modal, color-word interference effects. *Perception & Psychophysics*, *60*, 761-767.
- Engle, R. W. (2002). Working Memory Capacity as Executive Attention. *Psychological Science*, *11*(1), 19-23.
- European Commission. (1997). *Green Paper on Future Noise Policy*. COM 540 Final.
- Eysenck, H. J. (1967). *The biological basis of personality*. Springfield, Ill: Ch. Thomas.
- Fastl, H., Fruhmann, M., & Ache, S. (2004). Is the Railway Bonus influenced by the Directions of the Sound Sources? In C. Nocke & D. Cassereau (Eds.), *CFA/DAGA '04, Strasbourg, France*. Oldenburg: Isensee Verlag.

- Felscher-Suhr, U., Guski, R., Hunecke, M., Kastka, J., Paulsen, R., Schümer, R., et al. (1996). Eine methodologische Studie zur aktuellen Erfassung von Alltagstätigkeiten und deren Störungen durch Umweltlärm. *Zeitschrift für Lärmbekämpfung*, 43(3), 61-68.
- Fitger, C., & Rettig, K. (1985). Bewertung natürlicher Lärmformen und Auswirkungen auf Leistungen in einem Rechentest. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 32(3), 370-383.
- Furnham, A., & Strbac, L. (2002). Music is as distracting as noise: the differential distraction of background music and noise on the cognitive test performance of introverts and extraverts. *Ergonomics*, 45(3), 203-217.
- Gilhooly, K. J., Logie, R. H., Wetherick, N. E., & Wynn, V. (1993). Working memory and strategies in syllogistic-reasoning tasks. *Memory and Cognition*, 21(1), 115-124.
- Glaser, W. R., & Glaser, M. O. (1993). Colors as properties: Stroop-like effects between objects and their colors. In G. Strube & K. F. Wender (Eds.), *The cognitive psychology of knowledge* (Vol. 81, pp. 1-32). Amsterdam: North-Holland.
- Guski, R. (1993). *Psychische Auswirkungen von Umweltbelastungen* (No. Nr. 42). Bochum: Universität, Psychologisches Institut, Arbeitseinheit Kognitionspsychologie.
- Hartley, L. R., & Adams, R. G. (1974). Effect of noise on the stroop test. *Journal of Experimental Psychology*, 102(1), 62-66.
- Harvey, N. (1984). The Stroop effect: Failure to focus attention or failure to maintain focusing? *Quarterly Journal of experimental Psychology*, 36 A, 89-115.
- Hatfield, J., Job, R. F., Hede, A., Carter, N., Peploe, R., Taylor, R., et al. (2002). Human Response to Environmental Noise: The Role of Perceived Control. *International Journal of Behavioral Medicine*, 9(4), 341-350.
- Hellbrück, J., & Ellermeier, W. (2004). *Hören*. Göttingen: Hogrefe.
- Hellbrück, J., Kuwano, S., & Namba, S. (1996). Irrelevant background speech and human performance: Is there long-term habituation? *Journal of the Acoustic Society of Japan*, 17(5), 239-247.
- Heller, O. (1988). Eine neue Version des Rechnens nach Düker. *Archiv für Psychologie*, 140(3), 191-203.
- Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology*, 10, 302-323.
- Hitch, G. J., & Baddeley, A. (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology*, 28, 603-621.
- Hockey, G. R. J. (1978). Effects of noise on human work efficiency. In D. N. May (Ed.), *Handbook of Noise Assessment*. New York: Van Nostrand-Reinhold.
- Hockey, G. R. J. (1979). Stress and the cognitive components of skilled performance. In V. Hamilton & D. M. Warburton (Eds.), *Human stress and cognition* (pp. 141-177). Chichester, U.K.: Wiley.
- Hörmann, H., & Osterkamp, U. (1966a). Über den Einfluß von diskontinuierlichem Lärm auf die Organisation von Gedächtnisinhalten. *Zeitschrift für experimentelle und angewandte Psychologie*, 8, 265-273.

- Hörmann, H., & Osterkamp, U. (1966b). Über den Einfluß von kontinuierlichem Lärm auf die Organisation von Gedächtnisinhalten. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 13, 31-38.
- Houston, B. K. (1969). Noise, Task Difficulty, and Stroop Color-Word Performance. *Journal of Experimental Psychology*, 82(2), 403-404.
- Houston, B. K., & Jones, T. M. (1967). Distraction and Stroop Color-Word Performance. *Journal of Experimental Psychology*, 74(1), 54-56.
- Hygge, S., Boman, E., & Enmarker, I. (2003). The effects of road traffic noise and meaningful irrelevant speech on different memory. *Scandinavian Journal of Psychology*, 44(1), 13-22.
- Infodienst Lärmschutz in Kur- und Erholungsorten. (1992). Rechtsgrundlagen des Lärmschutzes. *Infodienst Lärmschutz in Kur- und Erholungsorten*, 2.
- Jacoby, L. L., Lindsay, D. S., & Hessels, S. (2003). Item-specific control of automatic processes: Stroop process dissociations. *Psychonomic Bulletin and Review*, 10(3), 638-644.
- Jansen, G. (1987). Verkehrslärmwirkungen bei besonderen Personengruppen. *Zeitschrift für Lärmbekämpfung*, 34, 152-156.
- Jones, D., Madden, C., & Miles, C. (1992). Privileged Access by Irrelevant Speech to Short-term Memory: The Role of Changing State. *The Quarterly Journal of Experimental Psychology*, 44A(4), 645-669.
- Jones, D. M., Alford, D., Macken, W., Banbury, S., & Tremblay, S. (2000). Interference from degraded auditory stimuli: Linear effects of changing state in the irrelevant sequence. *Journal of the Acoustical Society of America*, 108(3), 1082-1088.
- Jones, D. M., Beaman, P., & Macken, W. J. (1996). The object-oriented episodic record model. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 209-237). Hove: Psychology Press.
- Jones, D. M., & Broadbent, D. E. (1991). Human performance and noise. In C. M. Harris (Ed.), *Handbook of acoustical measurements and noise control* (pp. 24.21-24.24). New York: McGraw-Hill.
- Kahneman, D. (1973). *Attention and Effort*. New Jersey: Prentice-Hall.
- Kennedy, R. S., Dunlap, W. P., Banderet, L. E., Smith, M. G., & et al. (1989). Cognitive performance deficits in a simulated climb of Mount Everest: Operation Everest II. *Aviation, Space, and Environmental Medicine*, 60(2), 99-104.
- Kilcher, H., & Hellbrück, J. (1993). The irrelevant speech effect: Is binaural processing relevant or irrelevant, *Vortrag "Noise & Man '93"*. Nizza.
- Kjellberg, A., & Skoeldstroem, B. (1991). Noise annoyance during the performance of different nonauditory tasks. *Perceptual & Motor Skills*, 73(1), 39-49.
- Klatte, M., Kilcher, H., & Hellbrück, J. (1995). Wirkungen der zeitlichen Struktur von Hintergrundschall auf das Arbeitsgedächtnis und ihre theoretischen und praktischen Implikationen. *Zeitschrift für Experimentelle Psychologie*, 42(4), 517-544.
- Klatte, M., Lee, N., & Hellbrück, J. (2002). Effects of irrelevant speech and articulatory suppression on serial recall of heard and read materials. *Psychologische Beiträge*, 44(2), 166-186.
- Klauer, K. C., Stegmaier, R., & Meiser, T. (1997). Working Memory Involvement in Propositional and Spatial Reasoning. *Thinking and Reasoning*, 3(1), 9-47.

- Klein, G. S. (1964). Semantic power measured through the interference of words with colour-naming. *American Journal of Psychology*, 77, 576-588.
- Klingenberg, H. (1991). *Automobil- Meßtechnik, Band A: Akustik*. Berlin: Springer-Verlag.
- Knauff, M., & Strube, G. (2002). Anschauliches Denken und Arbeitsgedächtnis: Kognitive und kortikale Prozesse. *Psychologische Rundschau*, 53(2), 49-60.
- Laird, D. A., & Coye, K. (1929). Psychological measurements of annoyance as related to pitch and loudness. *Journal of the Acoustical Society of America*, 1, 158-163.
- Landstroem, U., Kjellberg, A., & Bystroem, M. (1995). Acceptable levels of tonal and broad-band repetitive and continuous sounds during the performance of nonauditory tasks. *Perceptual & Motor Skills*, 81(3), 803-816.
- LeCompte, D. C. (1995). An irrelevant speech effect with repeated and continuous background speech. *Psychonomic Bulletin & Review*, 2(3), 391-397.
- 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(2), 472-483.
- LeVere, T. E., Morlock, G. W., Thomas, L. P., & Hart, F. D. (1974). Arousal from sleep: the differential effects of frequencies equated for loudness. *Physiology & Behavior*, 12(4), 573-582.
- Lindsley, D. B., Bowden, J. W., & Magoun, H. W. (1949). Effect upon the EEG of acute injury to the brain stem activating system. *Electroencephalography & Clinical Neurophysiology*, 1, 475-486.
- Logie, R. H., & Baddeley, A. D. (1987). Cognitive Processes in Counting. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13(2), 310-326.
- Logie, R. H., Gilhooly, K. J., & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory and Cognition*, 22(4), 395-410.
- Lohmann, M. (1985). Über die Wirkung von Lärm in Erholungspausen. In D. Albert (Ed.), *Bericht über den 34. Kongress der Deutschen Gesellschaft für Psychologie in Wien 1984. Band II: Anwendungsbezogene Forschung* (pp. 857-858). Göttingen: Hogrefe.
- Luo, C. R. (1999). Semantic competition as the basis of Stroop interference: Evidence from color-word matching tasks. *Psychological Science*, 10(1), 35-40.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163-203.
- Manzey, D., & Lorenz, B. (1998). Effects of chronically elevated CO₂ on mental performance during 26 days of confinement. *Aviation, Space, and Environmental Medicine*, 69(5), 506-514.
- Manzey, D., Lorenz, B., & Poljakov, V. (1998). Mental performance in extreme environments: Results from a performance monitoring study during a 438-day spaceflight. *Ergonomics*, 41(4), 537-559.
- Maruyama, K. (1964). Studies on influences of train noise upon school children: VI: GSR during mental calculations. *Tohoku Psychologica Folia*, 23(1-2), 20-25.
- Matthews, G., Sparkes, T. J., & Bygrave, H. M. (1996). Attentional overload, stress, and simulated driving performance. *Human Performance*, 9(1), 77-101.
- Miles, C., & Jones, D. M. (1989). The fallacy of the cross-modal Stroop effect: A rejoinder to Cowan (1989). *Perception and Psychophysics*, 45(1), 85-86.

- Miles, C., Madden, C., & Jones, D. M. (1989). Cross-modal, auditory-visual stroop interference: A reply to Cowan and Barron (1987). *Perception and Psychophysics*, 45(1), 77-81.
- Möhler, U. (1988). Vergleich der Pausenstruktur von Schienenverkehrslärm und Straßenverkehrslärm. *Zeitschrift für Lärmbekämpfung*, 35, 10-15.
- Monahan, J. S. (2001). Colouring single Stroop elements: Reducing automaticity or slowing colour processing? *Journal of General Psychology*, 128(1).
- Näätänen, R. (1975). The inverted-U relationship between activation and performance: a critical review. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and Performance* (Vol. 5). New York: Academic Press.
- Neath, I. (1999). Modelling the disruptive effects of irrelevant speech on order information. *International Journal of Psychology*, 34(5/6), 410-418.
- Neath, I. (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review*, 7(3), 403-423.
- Neath, I., Surprenant, A. M., & LeCompte, D. C. (1998). Irrelevant speech eliminates the word length effect. *Memory & Cognition*, 26(2), 343-354.
- Neumann, O. (1996). Theorien der Aufmerksamkeit. In O. Neumann & A. F. Sanders (Eds.), *Aufmerksamkeit; Enzyklopädie der Psychologie, Themenbereich C, Theorie und Forschung, Serie II, Kognition, Band 2*. Göttingen: Hogrefe.
- Noel, M.-P., Desert, M., Aubrun, A., & Seron, X. (2001). Involvement of short-term memory in complex mental calculation. *Memory & Cognition*, 29(1), 34-42.
- Notbohm, G., Gärtner, C., & Schwarze, S. (2003). Psycho-physiologische Bewertung der Geräuschqualität von Straßenverkehrsgeräuschen. *Zeitschrift für Lärmbekämpfung*, 50(3), 75-81.
- Osada, Y. (1991). Comparison of community reactions to traffic noise. *Journal of Sound and Vibration*, 151(3), 479-486.
- Ouis, D. (2001). Annoyance from Road Traffic Noise: A Review. *Journal of Environmental Psychology*, 21, 101-120.
- Pritchatt. (1968). An investigation into some of the underlying associative verbal processes of the stroop colour effect. *Quarterly Journal of experimental Psychology*, 20, 351-359.
- Roberts, M. J., & Russo, R. (1999). *A student's guide to analysis of variance*. London: Routledge.
- Rohloff, U. B., & Gollwitzer, P. M. (1999). Reaktive Anspannungssteigerung und Geschwindigkeit in der Zielverfolgung. In L. Tent (Ed.), *Heinrich Düker. Ein Leben für die Psychologie und für eine gerechte Gesellschaft* (Vol. Band 1, pp. 305-327). Lengerich: Papst.
- Salamé, P. (1990). Exploring the Irrelevant Speech Effect: Musics, Speech-Like Noises and Visual Immediate Memory. In B. Berglund & T. Lindvall (Eds.), *Noise as a Public Health Problem* (Vol. 4, pp. 411-423). Stockholm: Swedish Council for Building Research.
- Salamé, P. (1993). The AGARD grammatical reasoning task: A defect and proposed solutions. *Ergonomics*, 36(12), 1457-1464.
- Salamé, P., & Baddeley, A. (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.

- Schönpflug, W. (1993). Continuous noise, discontinuous noise, and performance. In A. Schick (Ed.), *Contributions to psychological acoustics. Results of the Sixth Oldenburg Symposium on Psychological Acoustics* (pp. 615-626). Oldenburg: BIS Bibliotheks- und Informationssystem der Universität Oldenburg.
- Schuler, H., & Prochaska, M. (2000). Entwicklung und Konstruktvalidierung eines berufsbezogenen Leistungsmotivationstests. *Diagnostica*, 46(2), 61-72.
- Schütte, M., & Marks, A. (2004). Entwicklung des Dortmunder Lärmempfindlichkeits-Fragebogens (DoLe). In Gesellschaft für Arbeitswissenschaft (Ed.), *Arbeit + Gesundheit in effizienten Arbeitssystemen* (pp. 387-390). Dortmund: GfA Press.
- Schütte, M., & Nickel, P. (2002). Multivariate Analyse der bedingungsbezogenen Messgenauigkeit von Beanspruchungsindikatoren. *Zeitschrift für Arbeitswissenschaft*, 56(1-2), 55-66.
- Seitz, K., & Schumann-Hengsteler, R. (2000). Mental multiplication and working memory. *European Journal of Cognitive Psychology*, 12(4), 552-570.
- Shallice, T., & Burgess, P. (1993). Supervisory control of action and thought selection. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent*. New York, NY, US: Clarendon Press/Oxford University Press.
- Siegrist, M. (1995). Reliability of the Stroop test with single-stimulus presentation. *Perceptual and Motor Skills*, 81, 1295-1298.
- Smith, A. P. (1985). The effects of noise on the processing of global shape and local detail. *Psychological Research*, 47, 103-108.
- Smith, A. P., & Broadbent, D. E. (1985). The effects of noise on the naming of colours and reading of colour names. *Acta Psychologica*, 58, 275-285.
- Standing, L. G., Lynn, D., & Moxness, K. (1990). Effects of noise upon introverts and extroverts. *Bulletin of the Psychonomic Society*, 28(2), 138-140.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643-662.
- Surprenant, A. M., LeCompte, D. C., & Neath, I. (2000). Manipulations of irrelevant information: Suffix effects with articulatory suppression and irrelevant speech. *The Quarterly Journal of Experimental Psychology*, 53A(2), 325-348.
- Sust, C. (1987). *Geräusche mittlerer Intensität. Bestandsaufnahme ihrer Auswirkungen*. Bremerhaven: Wirtschaftsverlag NW.
- TA Lärm. (1998). *TA Lärm: Allgemeine Verwaltungsvorschrift über genehmigungsbedürftige Anlagen nach § 16 der Gewerbeordnung - Technische Anleitung zum Schutz gegen Lärm. Sechste AVwV vom 26.8.1998 zum Bundesimmissionsschutzgesetz*: GMBI. Nr. 26, S. 503.
- Taylor, M. (1995). Zur Lästigkeit von Verkehrslärm - Untersuchungen der McMaster Forschergruppe. *Zeitschrift für Lärmbekämpfung*, 42, 32-35.
- Thomas, J., Zoelch, C., Seitz-Stein, K., Schumann-Hengsteler, R. (in press). *Phonologische und zentral-exekutive Arbeitsgedächtnisprozesse bei der mentalen Addition und Multiplikation von Grundschulkindern*.
- Tracy, J., Mohamed, F., Faro, S., Tiver, R., Pinus, A., Bloomer, C., et al. (2000). The effect of autonomic arousal on attentional focus. *Neuroreport*, 11(18), 4037-4042.

- Tremblay, S., & Jones, D. M. (1999). Change of intensity fails to produce an irrelevant sound effect: Implications for the representation of unattended sound. *Journal of Experimental Psychology: Human Perception and Performance*, 25(4), 1005-1015.
- Tremblay, S., Macken, W., & Jones, D. M. (2001). The impact of broadband noise on serial memory: Changes in band-pass frequency increase disruption. *Memory*, 9(4/5/6), 323-331.
- van der Molen, M. W. (1996). Energetik und der Reaktionsprozeß. In O. Neumann & A. F. Sanders (Eds.), *Aufmerksamkeit; Enzyklopädie der Psychologie, Themenbereich C, Theorie und Forschung, Serie II, Kognition, Band 2*. Göttingen: Hogrefe.
- Van Gemmert, A. W., & Van Galen, G. P. (1997). Stress, neuromotor noise, and human performance: a theoretical perspective. *Journal of Experimental Psychology: Human Perception and Performance*, 23(5), 1299-1313.
- Västfjäll, D. (2002). Influences of Current Mood and Noise Sensitivity on Judgements of Noise Annoyance. *Journal of Psychology*, 136(4), 357-371.
- von-Hecker, U., Dutke, S., & Sedek, G. (2000). *Generative mental processes and cognitive resources. Integrative research on adaptation and control*. Dordrecht: Kluwer Academic Publishers.
- Working Group Railway Noise on the European Commission. (2003). *Position Paper on the European strategies and priorities for railway noise abatement*. Luxembourg: Office for Official Publications of the European Communities.
- Yerkes, R. M., & Dodson, J. D. (1908). The Relation of Strength of Stimulus to Rapidity of Habit Formation. *Journal of Comparative Neurology & Psychology*, 18, 459-482.
- Zwicker, E., & Fastl, H. (1986). Sinnvolle Lärmessung und Lärmgrenzwerte. *Zeitschrift für Lärmbekämpfung*, 33, 61-67.

Anhang

Anhang A

Amplitudenzeitverläufe der Hintergrundschalle (exemplarische fünfminütige Ausschnitte der verwendeten Signale)



Abbildung 36: Straßenverkehrsschall: 100 Durchfahrten pro Stunde – $L_{eq}=60$ dB(A)

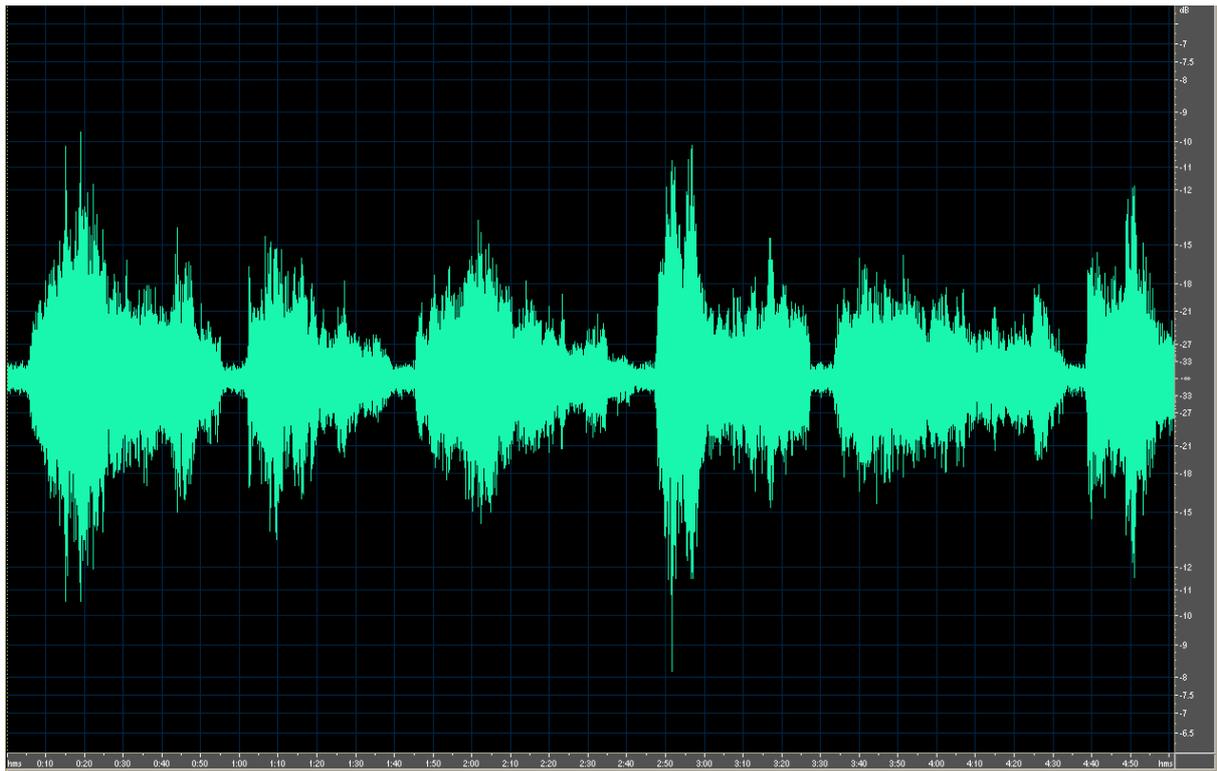


Abbildung 37: Straßenverkehrsschall: 2000 Durchfahrten pro Stunde – $L_{eq}=70$ dB(A)



Abbildung 38: Straßenverkehrsschall: 2000 Durchfahrten pro Stunde – $L_{eq}=70$ dB(A) – unter 500 Hz um 12 dB abgesenkt

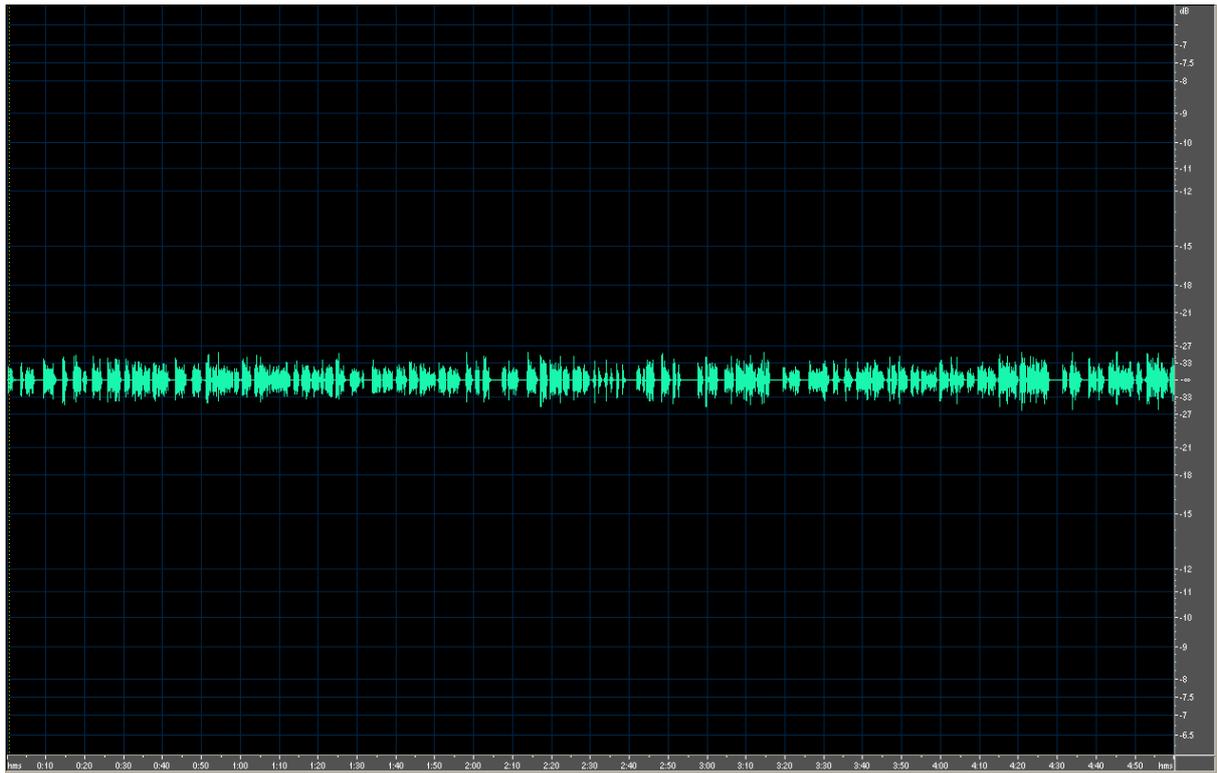


Abbildung 39: Deutscher Sprachschall: Helmut Griem spricht Briefe von Georg Büchner – $L_{eq}=60$ dB(A)

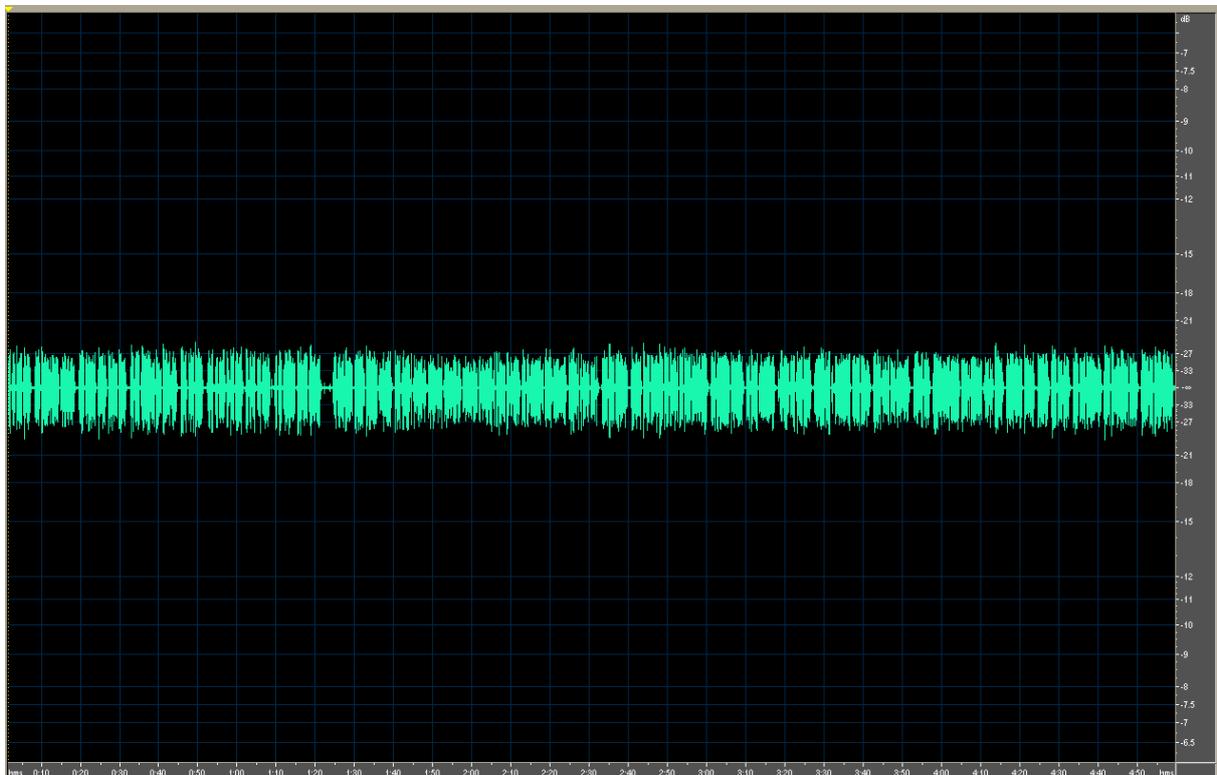


Abbildung 40: Japanischer Sprachschall: Mitschnitt einer Vorlesung – $L_{eq}=60$ dB(A)



Abbildung 41: Schienenverkehrslärm: diverse Zugdurchfahrten – $L_{eq}=70$ dB(A)



Abbildung 42: Schienenverkehrslärm: diverse Zugdurchfahrten – $L_{eq}=70$ dB(A); unter 220 Hz um 12 dB abgesenkt

Anhang B

Zwischenbefragung

Bitte kreuzen Sie an, was am ehesten auf Sie zutrifft:

Vp-Nr: _____

Bitte beantworten Sie folgende Fragen für den ersten Teil des Experiments:

1. Wie schwer ist Ihnen dieser Test bisher gefallen?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
2. Als wie anstrengend empfanden Sie die Aufgabe?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
3. Wie gut konnten Sie sich konzentrieren?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
4. Als wie störend empfanden Sie die Hintergrundgeräusche?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
5. Wie lästig fanden Sie die Hintergrundgeräusche?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht

Bitte beantworten Sie folgende Fragen für den zweiten Teil des Experiments:

1. Wie schwer ist Ihnen dieser Test bisher gefallen?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
2. Als wie anstrengend empfanden Sie die Aufgabe?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
3. Wie gut konnten Sie sich konzentrieren?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
4. Als wie störend empfanden Sie die Hintergrundgeräusche?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
5. Wie lästig fanden Sie die Hintergrundgeräusche?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht

Bitte beantworten Sie folgende Fragen für den dritten Teil des Experiments:

1. Wie schwer ist Ihnen dieser Test bisher gefallen?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
2. Als wie anstrengend empfanden Sie die Aufgabe?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
3. Wie gut konnten Sie sich konzentrieren?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
4. Als wie störend empfanden Sie die Hintergrundgeräusche?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht
5. Wie lästig fanden Sie die Hintergrundgeräusche?
Außerordentlich - ziemlich - mittelmäßig- kaum - gar nicht

Anhang C

Instruktion Serial Recall

Liebe Teilnehmerin, Lieber Teilnehmer!

Zunächst einmal vielen Dank für Ihre Teilnahme an diesem Versuch!

Es handelt sich um ein Experiment zum sprachlichen Kurzzeitgedächtnis.
Dabei gilt es, folgende Aufgabe am Computer zu bearbeiten:

- * Auf dem Bildschirm erscheinen die Ziffern 1 bis 9 nacheinander in zufälliger Reihenfolge. Ihre Aufgabe besteht darin, sich die genaue Reihenfolge der Ziffern so gut wie möglich einzuprägen.
- * Nach einer kurzen Pause erscheinen alle neun Ziffern gleichzeitig auf dem Bildschirm. Sie sollen nun mit der Maus die Ziffern in genau der Reihenfolge abklicken, in der die Ziffern präsentiert wurden.

Es ist wichtig, dass Sie so genau und so zügig wie möglich arbeiten.
Es müssen alle neun Ziffern abgeklickt werden, auch wenn Sie sich nicht sicher sind.
Eingabefehler können nicht korrigiert werden.

Vor dem eigentlichen Versuch gibt es neun Übungsaufgaben, damit Sie sich mit der Aufgabe vertraut machen können.

Der Versuch selbst dauert etwa eineinhalb Stunde und besteht aus 6 Blöcken mit je 20 Aufgaben.
Nach 10 Aufgaben wird jeder Block von einer 10-sekündigen Pause unterbrochen.
Nach jedem Block folgt eine 1-minütige Pause.

In den meisten Blöcken wird über einen Kopfhörer Hintergrundschall dargeboten.
Der Hintergrundschall ist für die Bearbeitung der Aufgabe völlig irrelevant und soll ignoriert werden!
Es werden später keine Fragen zu den Hintergrundschallen gestellt. Uns interessiert ausschließlich ihre Lernleistung.
Es ist äußerst wichtig, dass Sie so genau und so zügig wie möglich arbeiten.

Wenn Sie noch Fragen haben, wenden Sie sich damit jetzt an die Versuchsleiterin.
Ansonsten können Sie über Mausclick den ersten von insgesamt neun Übungsdurchgängen starten.
Nach den Übungsdurchgängen haben Sie nochmals die Möglichkeit Fragen zu stellen.

Anhang D

Instruktion KLTM

Liebe Versuchsperson,

vielen Dank für Ihre Teilnahme an diesem Experiment.

Es ist natürlich anonym, und Ihre Daten werden vertraulich behandelt.

es werden Ihnen im Folgenden kleine Rechenaufgaben (Plus/Minus) dargeboten, deren Lösung Sie bitte im Kopf errechnen und anschließend mit der Tastatur eingeben.

- Jede der Aufgaben besteht aus 3 Teilaufgaben:

Die Teilaufgaben beinhalten immer 3 Zahlen, die zu addieren oder subtrahieren sind, bspw. $4+6-3=?$.

Das Tempo der Zahlendarbietung kann selbst bestimmt werden - nach jeder Zahl wird durch das Drücken der Returntaste die nächste Zahl auf den Bildschirm geholt.

- Die Ergebnisse der Teilaufgaben sollen im Kopf errechnet und folgendermaßen verglichen werden:

Zuerst werden die Ergebnisse der 1. und 3. Teilaufgabe verglichen und nach folgender Regel verrechnet:

- Ist Ergebnis 1 größer als ($>$) Ergebnis 3, werden die Ergebnisse subtrahiert (Erg.1 minus Erg.3).

- Ist Ergebnis 1 kleiner als ($<$) Ergebnis 3, werden die Ergebnisse addiert (Erg.1 plus Erg.3).

Anschließend wird das Ergebnis der 2. Teilaufgabe mit dem eben errechneten Zwischenergebnis verglichen:

- Ist Ergebnis 2 größer als ($>$) das Zwischenergebnis von 1 und 3, werden die Ergebnisse subtrahiert (Erg.2 minus Zwischenerg.).

- Ist Ergebnis 2 kleiner als ($<$) das Zwischenergebnis von 1 und 3, werden die Ergebnisse addiert (Erg.2 plus Zwischenerg.).

- Das hieraus resultierende Endergebnis wird mit der Tastatur eingeben.

- Das folgende Rechenbeispiel dient noch einmal der Verdeutlichung des Rechenweges:

1. Teilaufgabe: $4 + 6 - 3 \rightarrow$ Ergebnis 1 = 7

2. Teilaufgabe: $8 + 9 + 5 \rightarrow$ Ergebnis 2 = 22

3. Teilaufgabe: $5 - 4 + 9 \rightarrow$ Ergebnis 3 = 10

\rightarrow Ergebnis 1 $<$ Ergebnis 3, also: $7 + 10 = 17$

\rightarrow Ergebnis 2 $>$ Zwischenergebnis, also: $22 - 17 = 5$

\rightarrow Endergebnis = 5

- Durch Drücken der Returntaste bestätigen Sie Ihre Eingabe, anschließend beginnt die nächste Aufgabe.

Haben Sie sich bei der Lösung vertippt, kann dies mit "Backspace" (\leftarrow) korrigiert werden.

Es werden verschiedene Geräusche eingespielt werden, die Sie bitte ignorieren.

Es werden später keine Fragen zu den Hintergrundschallern gestellt.

Der Versuch besteht aus 3 Blöcken mit je 15 Minuten Bearbeitungszeit, in denen sie bitte versuchen, so viele Rechenaufgaben wie möglich richtig zu lösen.

Hinweis: Es gibt keine negativen Lösungen, und nie sind Teilergebnisse, die direkt miteinander zu vergleichen sind, gleich groß.

Bitte arbeiten Sie konzentriert und zügig, es wird ausschließlich die Rechenleistung bewertet.

Bitte nicht artikulieren oder motorisch kodieren (flüstern oder die Finger zu Hilfe nehmen)!

Bitte arbeiten Sie so schnell wie möglich, aber ohne Fehler! Es ist für den Versuch wichtig, dass Sie sich anstrengen und Ihr Leistungsmaximum bringen.

NOCHMAL: STRENGEN SIE SICH BITTE AN, SO SCHNELL WIE MÖGLICH ZU RECHNEN, ES KOMMT AUF DIE MENGE GELÖSTER AUFGABEN PRO ABSCHNITT AN, WOBEI ALLERDINGS NUR RICHTIG GELÖSTE AUFGABEN ZÄHLEN!

Falls Sie jetzt noch Fragen haben, wenden Sie sich bitte an den Versuchsleiter, ansonsten beginnen mit Return 10 Übungsaufgaben.

Anhang E

Instruktionen Stroop

Stroop Farbinterferenz

Sie bekommen Farbworte gezeigt, die selbst in farbigen Buchstaben geschrieben sind. Ihre Aufgabe ist es, auf die Farbe zu antworten, mit der die Farbwörter geschrieben sind, nicht auf das Wort selbst. Bitte fixieren Sie hierzu immer zuerst das kurz aufleuchtende Fixationskreuz, und entscheiden hierauf über die Farbe des erscheinendes Wortes:

Hierfür können Sie mit folgenden Tasten antworten:

R für Rot; Ü für Grün; G für Gelb; L für Blau

Um möglichst schnell antworten zu können, legen Sie bitte den Zeigefinger Ihrer linken Hand auf die Taste G und Ihren linken Mittelfinger auf die Taste R. Mit der rechten Hand verfahren Sie ähnlich, und legen Ihren Zeigefinger auf die Taste L und Ihren Mittelfinger auf die Taste Ü.

Drücken Sie zum Weitermachen bitte eine der Antworttasten!

Zum Starten des Übungsdurchgangs, der aus 48 Aufgaben besteht, drücken Sie irgendeine der Antworttasten!

Anhang F

Instruktionen Grammatical Reasoning

Satzverifikations- Aufgabe

Die folgenden Aufgaben bestehen jeweils aus zwei Aussagesätzen und einer Anordnung von drei Symbolen

Jeder Aussagesatz beschreibt die Reihenfolge von zwei Symbolen.

Ihre Aufgabe ist es, zu beurteilen, ob die Aussagen mit der Symbolanordnung übereinstimmen.

Sind beide Aussagen zutreffend oder beide unzutreffend, drücken Sie die Taste für IDENTISCH. Ist ein Satz zutreffend, der andere aber nicht, dann drücken Sie die Taste für VERSCHIEDEN.

Ein Beispiel:

Aussagesatz 1: # wird nicht von & gefolgt

Aussagesatz 2: # geht * voraus

Symbolanordnung: *#&

Das # kommt in der Symbolanordnung vor dem &, wird also von ihm gefolgt, deshalb ist der erste Satz falsch. Das # kommt nicht vor dem *, geht ihm also nicht voraus, deshalb ist der zweite Satz ebenfalls falsch. Da somit beide Sätze falsch sind, ist die korrekte Antwort IDENTISCH.

Betrachten Sie nun folgende Aufgabe:

geht & nicht voraus

* wird von & gefolgt

&#*

Der # kommt nach dem &, geht ihm also nicht voraus, demnach ist der erste Satz richtig. Das & kommt nicht nach dem *, es wird also nicht vom & gefolgt, demnach ist der zweite Satz falsch. Die korrekte Antwort ist deshalb VERSCHIEDEN.

Beachten Sie, dass die zwei in einem Aussagesatz genannten Symbole in der Symbolanordnung nicht direkt benachbart stehen müssen. Sie können auch durch das dritte Symbol getrennt sein. Es zählt allein die relative Position in der Folge der Symbole!

Versuchen Sie, auf die Aufgaben so schnell wie möglich zu antworten, aber ohne Fehler.

SHL <== VERSCHIEDEN IDENTISCH ==> SHR

Sie haben nun Gelegenheit, die Aufgabe zu üben. Dabei bekommen Sie nach jeder Antwort eine Rückmeldung (richtig/ falsch).

Um eine neue Aufgabe aufzurufen, betätigen Sie nochmals eine Antworttaste.

Die Aufgabe wird nun in fünf Durchgängen von je etwa 15 Minuten Dauer bearbeitet. In jedem Durchgang wird Ihnen ein Hintergrundschall über Kopfhörer eingespielt werden, den Sie aber bitte ignorieren!

Nach jedem Durchgang haben Sie zwei Minuten Pause, bevor der nächste Durchgang beginnt.

Setzen Sie nun bitte den Kopfhörer auf!

Versuchen Sie, die Aufgaben so schnell wie möglich zu beantworten, aber ohne Fehler.

Es ist wichtig, dass Sie Ihr Leistungsmaximum bringen!

Nur ihre Leistung wird beurteilt, sonst nichts!

Anhang G

Filterfunktion der Verkehrsschalle (-12 dB unter 500 Hz)

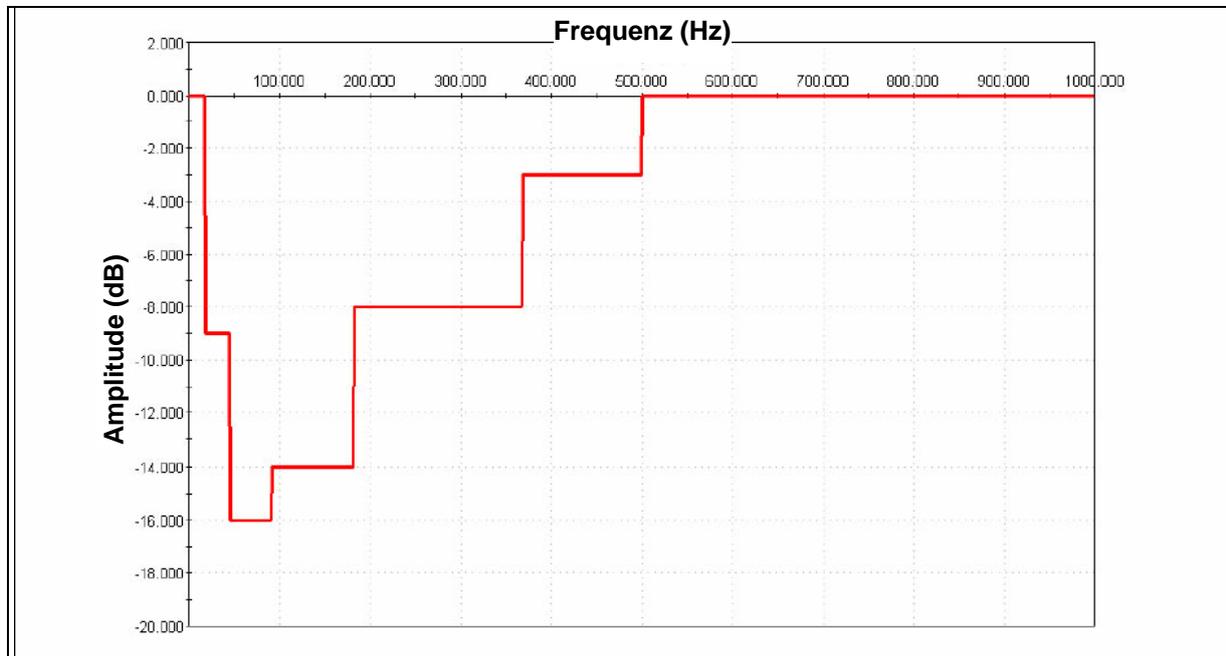


Abbildung 43: Filterfunktion der Straßenverkehrsschalle (-12 dB unter 500 Hz): Prof. Bisping; SASS acoustic research & design GmbH

Anhang H

Filterfunktion Schienenverkehr (-12 dB unter 220 Hz)

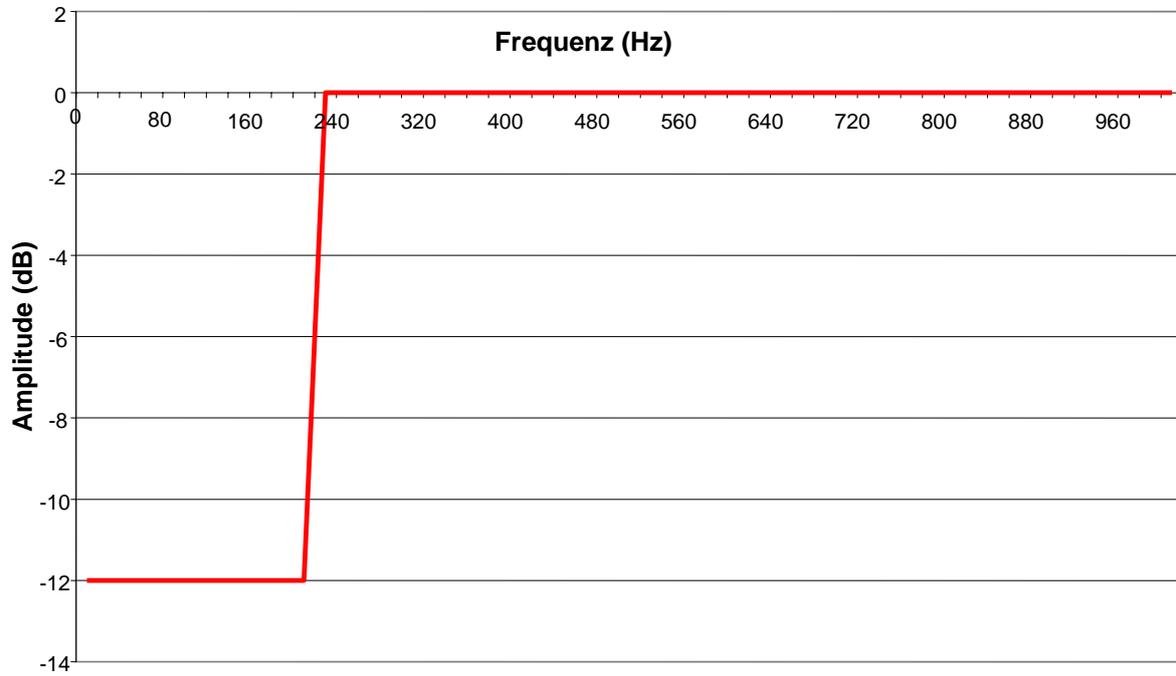


Abbildung 44: Filterfunktion der Schienenverkehrsschalle (-12 dB unter 220 Hz): Prof. Bisping; SASS acoustic research & design GmbH

Anhang I

Mittelwerte und Standardfehler (bzw. Median und Quartile) der Fehlerraten und Reaktionszeiten in den Experimenten

Fragestellung I

Tabelle 7: Serial Recall, Mittelwert und Standardfehler der Leistungsdaten

Schallbedingung		Ruhe	Sprache	100-60	2000-50	2000-60	2000-70
Fehlerraten in %	M	28.7	39.9	30.5	31.7	29.9	30.7
	SE	2.7	2.7	2.7	3.3	2.8	3.0
Reaktionszeiten ms	M	8758.6	8538.3	8448.9	8836.7	8390.3	8295.7
	SE	449.0	471.7	378.2	461.8	389.4	366.4

Tabelle 8: KLTM, Mittelwert und Standardfehler, bzw. Median und Quartile der Leistungsdaten

Schallbedingung		Ruhe	Sprache	100-60	2000-50	2000-60	2000-70
Bearbeitete Aufg. %	M	30.2	29.1	29.5	30.0	30.6	30.2
	SE	2.1	2.1	1.7	2.4	2.3	1.9
Richtig gel. Aufg. %	M	23.4	20.6	23.1	23.7	23.5	22.3
	SE	2.4	2.4	1.9	2.6	2.6	2.2
Falsch gel.. Aufg. %	M	6.8	8.6	6.4	6.2	7.1	7.9
	SE	1.0	1.2	1.0	0.9	1.2	1.0
Fehlerraten %	Median	22.7	31.4	18.0	20.3	22.7	24.2
	1.Quartil	13.2	13.4	6.9	8.1	10.7	8.3
	3.Quartil	6.9	7.3	11.4	10.6	6.1	9.1
Reaktionszeiten ms	M	25411.9	25607.2	25228.7	26931.0	25900.0	25006.7
	SE	2564.2	2525.6	2106.7	3144.8	2628.2	2080.2

Tabelle 9: Stroop Test, Mittelwert und Standardfehler der Leistungsdaten

Schallbedingung		Ruhe	Sprache	100-60	2000-50	2000-60	2000-70
Fehlerraten % Inkongruent	M	5.8	6.5	7.3	7.3	6.1	9.2
	SE	1.0	0.7	1.2	1.0	1.1	1.4
kongruent	M	3.4	3.8	3.4	3.7	3.6	3.8
	SE	0.7	0.6	0.6	0.6	0.6	0.6
Reaktionszeiten ms Inkongruent	M	683.8	701.1	704.0	687.7	682.3	713.5
	SE	30.3	28.6	39.3	31.6	28.3	39.1
kongruent	M	592.7	600.4	600.2	597.8	582.4	602.9
	SE	25.9	24.0	26.9	19.7	18.9	25.0

Fragestellung II

Tabelle 10: KLTM(Kopfhörer), Mittelwert und Standardfehler, Median und Quartile der Leistungsdaten

Schallbedingung		Ruhe	Sprache	2000-70	2000-70-12	2000-70 abgesenkt	2000-70 -12 erhöht
<i>Bearbeitete Aufg. %</i>	M	20.3	20.5	20.7	20.6	20.3	21.0
	SE	1.2	1.5	1.4	1.3	1.1	1.1
<i>Richtig gel. Aufg. %</i>	M	16.6	15.1	16.2	16.2	16.2	16.9
	SE	1.4	1.8	1.7	1.4	1.2	1.2
<i>Falsch gel. Aufg. %</i>	M	3.7	5.4	4.4	4.4	4.1	4.1
	SE	0.4	0.5	0.5	0.7	0.6	0.5
<i>Fehlerraten %</i>	Median	18.1	28.7	22.9	20.3	17.4	18.0
	1.Quartil	7.8	10.7	9.4	7.2	5.5	8.3
	3.Quartil	7.7	11.0	4.6	6.7	6.3	8.8
<i>Reaktionszeiten ms</i>	M	25065.3	25889.4	25122.7	25117.4	24942.3	23505.0
	SE	2385.2	3123.4	2742.8	3119.5	2427.0	2210.9

Tabelle 11: KLTM(Subwoofer), Mittelwert und Standardfehler, bzw. Median und Quartile der Leistungsdaten; Leistungsdaten im Übungsdurchgang unter Ruhe

Schallbedingung		Ruhe Übung	Ruhe	Sprache	2000-70	2000-70-12	Übungsdurchgang in Ruhe KLTM(Kopfhörer)
<i>Bearb. Aufg. %</i>	M	14.4	17.3	16.3	16.5	16.6	17.2
	SE	0.8	1.0	0.9	0.9	1.0	0.9
<i>Richtige. Aufg. %</i>	M	8.8	12.3	9.8	10.2	10.7	12.6
	SE	1.0	1.3	1.2	1.4	1.2	1.2
<i>Falsche Aufg. %</i>	M	5.6	5.0	6.5	6.3	6.0	4.6
	SE	0.8	1.0	0.9	1.0	0.7	0.6
<i>Fehlerraten %</i>	Median	40.7	22.2	40.1	38.9	31.7	26.1
	1.Quartil	17.8	13.5	18.2	19.2	14.7	9.5
	3.Quartil	7.7	11.4	16.7	17.5	9.8	16.4
<i>Reaktionszeiten in ms</i>	M	38499.8	31773.5	33212.2	32798.8	33400.7	30867.1
	SE	3299.1	3035.5	3012.2	2887.3	3403.8	2833.0

Tabelle 12: Stroop Test, Mittelwert und Standardfehler der Leistungsdaten

Schallbedingung		Ruhe	Sprache	2000-70	2000-70-12	2000-70 abgesenkt	2000-70 -12 erhöht
<i>Fehlerraten %</i> Inkongruent	M	3.2	4.4	5.6	5.7	3.3	5.3
	SE	0.7	0.9	1.2	1.1	0.6	0.9
kongruent	M	2.2	2.4	3.8	2.6	2.6	2.2
	SE	0.5	0.5	1.0	0.7	0.5	0.5
<i>Reaktionszeiten ms</i> Inkongruent	M	720.8	708.3	711.8	696.8	710.9	712.6
	SE	31.5	31.5	26.3	27.7	31.5	26.6
kongruent	M	620.4	628.1	613.1	606.2	618.0	611.7
	SE	24.4	26.7	22.3	19.7	22.6	20.4

Tabelle 13: Grammatical Reasoning, Mittelwert und Standardfehler der Leistungsdaten

Schallbedingung		Ruhe	Sprache	2000-70	2000-70- 12
<i>Fehlerraten %</i>	M	18.3	19.3	17.3	15.2
	SE	3.6	3.9	3.7	2.8
<i>Reaktionszeiten ms</i>	M	11333.8	13476.1	12159.9	11694.7
	SE	894.3	1255.3	982.5	877.3

Fragestellung III

Tabelle 14: KLTM, Mittelwert und Standardfehler der Leistungsdaten

Schallbedingung		Ruhe	Japanisch	Schiene	Schiene abgesenkt
<i>Fehlerraten %</i>	M	29.2	39.8	29.6	30.2
	SE	4.0	4.3	3.7	4.2
<i>Reaktionszeiten ms</i>	M	26987.9	27184.4	27402.6	26557.9
	SE	1621.0	1486.0	1858.2	1660.6

Tabelle 15: Grammatical Reasoning, Mittelwert und Standardfehler der Leistungsdaten

Schallbedingung		Ruhe	Japanisch	Schiene	Schiene abgesenkt
<i>Fehlerraten %</i>	M	10.4	12.1	12.4	10.2
	SE	3.1	3.6	3.6	2.6
<i>Reaktionszeiten ms</i>	M	2021.8	2593.6	2399.2	2356.1
	SE	841.1	674.4	754.1	726.4