

## Ein Algorithmus zur Vorhersage des Irrelevant Sound Effects

Tobias Weissgerber<sup>1</sup>, Sabine Schlittmeier<sup>2</sup>, Stefan Kerber<sup>1</sup>, Hugo Fastl<sup>1</sup>, Jürgen Hellbrück<sup>2</sup>

<sup>1</sup>AG Technische Akustik, MMK, TU München, Email: wei@mmk.ei.tum.de

<sup>2</sup>Arbeits-, Umwelt- und Gesundheitspsychologie, KU Eichstätt-Ingolstadt

### Einleitung

Die verbale Arbeitsgedächtnisleistung wird durch Hintergrundschalle, die deutliche temporal-spektrale Schwankungen aufweisen, signifikant gestört. Beispielsweise kann eine Ziffernfolge unter Sprache oder Staccato-Musik wesentlich schlechter behalten werden als in Ruhe, wohingegen Legato-Musik oder kontinuierliches Rauschen kaum bis gar nicht stört. Da die Hintergrundschalle keinen Bezug zu den Ziffernfolgen bzw. zur Aufgabe haben, wird dieser Effekt als Irrelevant Sound Effect (ISE) bezeichnet (z.B. in [1]). Dieses in der Kognitionspsychologie vielfach untersuchte Phänomen (für einen Überblick siehe z.B. [2]) konnte bislang jedoch nicht durch psychoakustische Maße vorhergesagt werden. Basierend auf der instrumentellen Messung der Hörempfindung "Schwankungsstärke" stellen wir im vorliegenden Beitrag einen Algorithmus vor, der die Ergebnisse von Experimenten zum ISE nachbildet.

### Experimente/Datenbank

Die Datenbank besteht aus den Ergebnissen von Verhaltensexperimenten, welche an der Katholischen Universität Eichstätt-Ingolstadt durchgeführt wurden. Als irrelevante Schalle wurden hier Sprache, Musik, Rosa Rauschen, Tiergeräusche, Tonsequenzen sowie Büro- und Verkehrslärm getestet.

In allen Experimenten wurde eine Serial Recall Aufgabe und damit die Standardaufgabe zur Erfassung der Arbeitsgedächtniskapazität bzw. des ISE eingesetzt. Dabei wurden die Ziffern von 1 bis 9 in randomisierter Reihenfolge unter verschiedenen Hintergrundschallbedingungen dargeboten und waren nach einem kurzen Retentionsintervall von 10 s in exakt der präsentierten Reihenfolge wiederzugeben. Jede Ziffer, die nicht genau an der Serienposition erinnert wurde, an der sie präsentiert worden war, zählte als Fehler. Unter jeder Schallbedingung wurden – je nach Experiment – 15-20 Trials bearbeitet (ein Trial umfasst Itempräsentation – Retentionsintervall – Recall). Für jede Versuchsperson wurden die Fehlerraten über alle Trials einer Schallbedingung gemittelt. Aus diesen Fehlerraten wurde dann über die gesamte Versuchspersonengruppe je Schallbedingung der Median bestimmt. Im Folgenden ist immer von solchen Fehlermedien (Prozent Fehler) die Rede. Es wurden Experimente sowohl für eine visuelle als auch eine auditive Darbietung der Ziffernfolge durchgeführt.

Die Hintergrundschalle wurden entweder über Lautsprecher oder über Kopfhörer bei definiertem Schalldruckpegel dargeboten. Bei Lautsprecherdarbietung wurde der Pegel mittels Brüel & Kjaer 2231 Schallpegelmessung oder via Noise-Book 2.0, HEAD acoustics GmbH, gemessen. Bei Kopfhörerdarbietung wurde der Pegel mittels eines künstlichen Ohrs Brüel & Kjaer 4153 gemessen. Die Hintergrundschalle wurden binaural dargeboten.

### Algorithmus und Diskussion

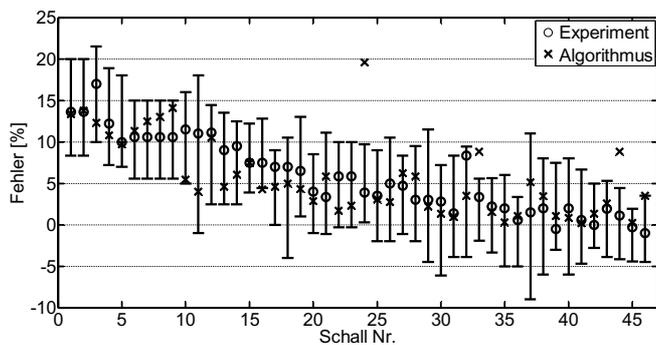
Die Experimente haben gezeigt, dass Sprachsignale die größte Störwirkung induzieren. Bei fließender Sprache werden ungefähr 4 Silben pro Sekunde gesprochen, was einer Schwankung der zeitlichen Hüllkurve von 4 Hz entspricht [3]. Bei dieser Modulationsfrequenz von 4 Hz erreicht die Empfindungsgröße "Schwankungsstärke" ihr Maximum. Die Schwankungsstärke beschreibt modulierte Schallsignale mit Modulationsfrequenzen kleiner als 20 Hz. Diese Änderungen können vom Gehör als Einzelereignisse erkannt werden und werden somit als zeitliche Lautheitsänderungen wahrgenommen. Da die Schwankungsstärke der übrigen Schalle ebenfalls mit der verursachten Störwirkung korreliert, ist eine Vorhersage des ISEs über die Schwankungsstärke für viele Schalle möglich.

Hierzu werden die Schalldateien analog zur Dauer eines Trials auf 30 Sekunden beschränkt und dem dargebotenen Schalldruckpegel angepasst. Die Schwankungsstärken der Schalldateien wurden mit der Software PAK® der Müller-BBM VibroAkustik Systeme GmbH bestimmt und der arithmetische Mittelwert der Schwankungsstärke über die Zeit berechnet. Als Referenzschall für den Algorithmus wurde ein Schall mit kleinem Interquartilbereich gewählt. Der Fehlerwert (ISE) und die Schwankungsstärke dieser Referenz dienen als Normierung für den Algorithmus. Schlittmeier et al. haben in [4] nachgewiesen, dass die Existenz des ISE unabhängig von der Präsentationsmodalität der zu erinnernden Items ist. Da die Stärke des ISE jedoch von der Präsentationsmodalität der Items abzuhängen scheint, wird im Folgenden für visuell bzw. auditiv präsentierte Items je ein eigener Algorithmus vorgestellt. Abgesehen von einem Faktor sind beide Algorithmen jedoch mathematisch identisch.

Die Datenbank für visuell präsentierte Items umfasst die Ergebnisse zu 46 Schallen. Als Referenzschall wurde Musik in Staccato-Spielweise bestimmt (entspricht Schall Nr. 15 in Abbildung 1). Dieser Schall hat eine Schwankungsstärke  $F$  von 0,68 vacil bei einem ISE von 7,5 Prozent. Damit kann der ISE für visuelle Items  $ISE_v$  eines beliebigen Schalls gemäß (1) abgeschätzt werden.

$$ISE_v = \frac{F}{0,68 \text{ vacil}} \cdot 7,5 \quad [\%] \quad (1)$$

Basierend auf Formel (1) wurde der ISE für alle 46 Schalle algorithmisch bestimmt. Algorithmisch und experimentell ermittelter Fehlerwert korrelieren signifikant miteinander (Spearman'sche Rangkorrelation:  $r_s=0,72$ ;  $p<0,01$ ).



**Abbildung 1:** Irrelevant Sound Effect der 46 Schalle bei visueller Itempräsentation bei experimenteller (Kreis) und algorithmischer (Kreuz) Bestimmung.

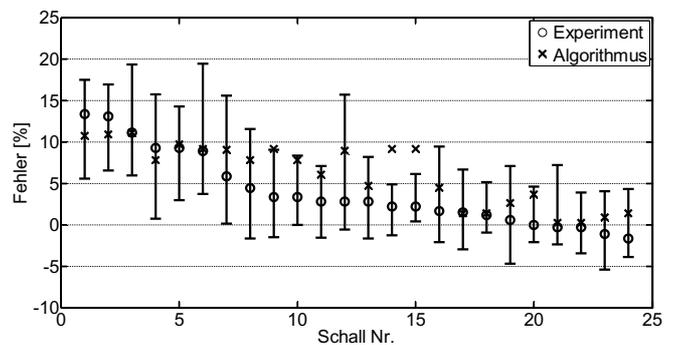
Der Vergleich von experimentellem und algorithmischem Ergebnis wird in Abbildung 1 dargestellt. Mit 43 von 46 vorgegebenen Schallen wird für 93,5 Prozent aller Schalldateien die Wirkung des ISEs innerhalb der Interquartilbereiche der vorliegenden Datenbankergebnisse korrekt vorhergesagt.

Die Schalle Nr. 24, 33 und 44, für die die algorithmische Vorhersage nicht zutrifft, sind synthetische Schalle und beinhalten regelmäßig wiederkehrende Elemente (Periodizität) im temporalen und/oder im spektralen Bereich. Schalle, bei denen sich ein und dieselbe Wahrnehmungseinheit ständig wiederholt, stören die Arbeitsgedächtnisleistung kaum oder gar nicht. Sie werden nach der kognitionspsychologischen und allgemein zur Erklärung des ISE anerkannten Changing-State Hypothese als "Steady-State" Schalle bezeichnet [5]. Ihre Periodizitäten werden vom menschlichen auditiv-perzeptiven System erkannt. Da der Algorithmus in seiner aktuellen Form aber ausschließlich die Schwankungsstärke eines Schalls berücksichtigt, überschätzt er die zu erwartende Störwirkung dieser künstlichen Steady-State Schalle deutlich.

Die Datenbank für auditiv präsentierte Items umfasst die Ergebnisse zu 24 Schallen. Als Referenzschall wurde ein Sprachsignal mit einer Schwankungsstärke  $F$  von 0,95 vacil bei einem ISE von 11,5 Prozent gewählt. Die Vorhersage des Irrelevant Sound Effects  $ISE_a$  bei einer auditiven Darbietung der Ziffernfolge eines beliebigen Schalls berechnet sich nach (2).

$$ISE_a = \frac{F}{0,95 \text{ vacil}} \cdot 11,5 \quad [\%] \quad (2)$$

Hier kann der Algorithmus für 87,5 Prozent aller Schalldateien die Wirkung des ISEs innerhalb der Interquartilbereiche der experimentellen Ergebnisse richtig bestimmen. Auch hier korrelieren die im Verhaltensexperiment gemessenen Leistungsbeeinträchtigungen unter dem jeweiligen Schall signifikant ( $r_s=0,87$ ;  $p<0,01$ ) mit der vom Algorithmus vorhergesagten Höhe des ISE. Die Gegenüberstellung von Experiment und Algorithmus ist in Abbildung 2 angegeben. Analog zum Algorithmus für visuell präsentierte Items handelt es sich bei den Schallen, für die die Berechnung des ISE



**Abbildung 2:** Irrelevant Sound Effect der 24 Schalle bei auditiver Itempräsentation bei experimenteller (Kreis) und algorithmischer (Kreuz) Bestimmung.

zu hohe Ergebnisse liefert (Schall Nr. 9, 14 und 15), um Steady-State Schalle.

## Zusammenfassung

Zusammenfassend ist festzuhalten, dass die vorgestellten Algorithmen die Störwirkung von Hintergrundschallen auf die Arbeitsgedächtnisleistung (Irrelevant Sound Effect) auf Grundlage der Schwankungsstärke zuverlässig vorhersagen können. Die Vorhersage gelingt nicht bei einigen wenigen synthetischen Schallen, bei denen ein und dieselbe Wahrnehmungseinheit wiederholt dargeboten wurde (Steady-State Schalle). Für die Evaluation von im Feld anzutreffenden Schallbedingungen wie Hintergrundsprache oder –musik und Bürolärm liefern die Algorithmen jedoch sehr gute Ergebnisse. Insgesamt trifft die Vorhersage für über 85% der Schalle innerhalb der wahrscheinlichen Abweichung der Experimentalmessung zu.

## Literatur

- [1] Banbury, S. P., Tremblay S., Macken, W. J. & Jones, D. M.: Auditory distraction and short-term memory Phenomena and practical implications. *Human Factors* 2001, 43(1), 12-29.
- [2] Schlittmeier, S.: Arbeitsgedächtnis und Hintergrundschall – Gibt es einen "Irrelevant Sound Effect" bei auditiv präsentierten Items? Logos Verlag, Berlin, 2005.
- [3] Fastl, H. & Zwicker, E.: Psychoacoustics – Facts and Models, 3rd Edition. Springer Verlag, Heidelberg, 2007.
- [4] Schlittmeier, S. J., Hellbrück, J. & Klatt, M.: Does irrelevant music cause an irrelevant sound effect for auditory items? *European Journal of Cognitive Psychology*, 2007 (DOI: 10.1080/09541440701427838).
- [5] Macken, W. J., Mosdell, N. & Jones, D. M.: Explaining the irrelevant sound effect. Temporal distinctiveness or changing state? *Journal of Experimental Psychology: Learning, Memory & Cognition* 1999, 25(3), 820-824.