

# Arbeitsgedächtnis, räumliche Quellentrennung und Sprachverständlichkeit von Hörgeschädigten mit Hörgeräten

Matthias Vormann<sup>1</sup>, Kirsten C. Wagoner<sup>1</sup>, Tobias Neher<sup>2</sup>, Rosa-Linde Fischer<sup>3</sup>

<sup>1</sup> Hörzentrum Oldenburg GmbH und Cluster of Excellence Hearing4All, Oldenburg

<sup>2</sup> Institute of Clinical Research, University of Southern Denmark, Odense, Denmark

<sup>3</sup> Sivantos GmbH, Erlangen

**Schlüsselwörter: Arbeitsgedächtnis, räumliche Quellentrennung und Sprachverständlichkeit**

## Einleitung

In einer Studie am Hörzentrum Oldenburg wurde der Einfluss der Akklimatisierung auf die Performanz und Präferenz von neuversorgten Hörgeräteträgern (Wagoner et al. 2018) untersucht. Als Performanzmaße wurden u.a. die Arbeitsgedächtnisleistung und die räumliche Quellentrennung erfasst. Diese Prinzipien sind für die Kommunikationsfähigkeit im Hinblick auf die Sprachverständlichkeit im Störgeräusch (Cocktail-Party Effekt) von entscheidender Bedeutung. Die dafür verwendeten Testverfahren sollen im Folgenden vorgestellt und auf ihre Trennschärfe und Wiederholbarkeit hin untersucht werden. Dafür werden die Daten der Kontrollgruppe aus dieser Studie herangezogen. Die Probanden waren während des gesamten Zeitraums mit ihren eigenen Hörsystemen versorgt und es wurden währenddessen keinerlei Änderungen an den Hörgeräten vorgenommen.

## Testverfahren

Die Arbeitsgedächtnisleistung (Baddeley und Hitch, 1974) als ein wesentlicher Faktor für die menschliche Sprachkommunikation, wurde mit einer deutschen Version des Listening Span Test (LST, Doherty und Desjardins 2015) erfasst, die auf ähnlichen Mechanismen beruht wie der etablierte Reading Span Test (RST, Akeroyd, 2008). Beim LST wird der Stimulus akustisch, d.h. mittels Sprachsignalen dargeboten, wodurch für Hörgeräteträger eine typischere Situation erfasst werden kann als bei der optischen Darbietung, wie sie beim RST erfolgt.

Für die Implementation des LST wurde das Sprachmaterial aus dem Basler Satztest (BASA, Tschopp und Züst 1994) verwendet. Dieser Test ist eine modifizierte deutsche Version des amerikanischen SPIN Tests (Speech Perception in Noise, Kalikow et al. 1977). Die 300 Sätze (20 Testlisten mit jeweils 15 Sätzen) wurden, über die Testlisten verteilt, zufällig von einer männlichen bzw. weiblichen Stimme aufgesprochen. Die Sätze bestehen aus sechs - neun Silben, an dessen Ende jeweils ein einsilbiges Substantiv steht. Für je 10 komplette Testlisten ist das Substantiv aus dem vorangehenden Kontext vorhersagbar (high predictable, HP – z.B. „Das Schaf auf der Weide frisst *Gras*.“), oder nicht vorhersagbar (low predictable, LP – z.B. „Was bedeutet dieser *Strich*?“). Diese Sätze werden bei einem SNR von 12 dB in einem Störgeräusch bei 65 dB, welches der Geräuschkulisse einer diffusen Cafeteria nachempfunden ist, den Probanden präsentiert<sup>1</sup>. Die Aufgabe der Probanden besteht darin, den Satz laut zu wiederholen und sich das Schlusswort zu merken. Dieser Vorgang wird in Blöcken zu je 4 oder 6 Sätzen durchgeführt, jeweils getrennt für LP / HP Sätze. Am Ende eines Blocks wird der Proband aufgefordert, die 4 oder 6 gemerkten Schlusswörter aus der Erinnerung heraus wiederzugeben. Um die Schwierigkeit konstant zu halten, wird innerhalb eines Blockes jeweils alle 4 Sekunden ein Satz dargeboten. Diese Taktung kann vom Probanden nicht beeinflusst werden. Vor dem eigentlichen Test wurde ein Übungsdurchgang durchgeführt.

Die Fähigkeit der Probanden zur akustischen Quellentrennung wird mittels einer deutschen Variante des „Stock-price monitoring“ (Getzmann et al., 2015) erfasst („Börsenszenario“). In der verwendeten Version führen die Probanden einen Oldenburger Satztest (OLSA) mit weiblicher, deutscher Sprecherin („Target“) bei einem fixen Sprachpegel von 71dB im Störgeräusch (gleiche Cafeteria wie zuvor) bei 65 dB durch. Synchron dazu wurden konkurrierende männliche Olsa-Sätze auf Spanisch bzw. Russisch (Hochmuth et al. 2013) bei einem Pegel von jeweils 71 dB aus anderen Richtungen dargeboten. Die drei verwendeten Richtungen relativ zum Proband waren -45°, 0°, 45°. Bei jeder Darbietung wurde die Zuordnung der drei Sprecher pseudo-zufällig auf die drei Richtungen vorgenommen, d.h. es standen sechs Permutationen aus Richtung und Sprecher zur Verfügung. Die Aufgabe der Probanden bestand darin, nur die *Zahl* des deutschen Olsa-Satzes zu erkennen und nachzusprechen. Vorab wurden zwei Durchgänge zur Übung durchgeführt.

<sup>1</sup> Das zu jedem Satz gehörende, spezifische Rauschen der originalen BASA Version wird hier nicht verwendet.

## Studiendesign und Probanden der Kontrollgruppe

Die Signale für beide Tests wurden im Vorfeld der Messungen in einem Aufbau mit 8 Lautsprechern generiert und mittels KEMAR aufgenommen. Die Aufnahmen wurden danach entsprechend des individuellen Hörverlustes für jeden einzelnen Probanden verstärkt und für die Wiedergabe mit Kopfhörer entzerrt. Die Tests wurden im Anschluss in einer Abhörkabine mit Sennheiser HDA 200 durchgeführt, vgl. Wagener et al. (2018).

Für die Akklimatisierungsstudie wurden die o.g. Tests dreimal mit einem Abstand von 6 Wochen durchgeführt (T1, T2, T3). Für die Kontrollgruppe wurde die Verstärkung der eigenen Hörgeräte unverändert übernommen und für alle drei Messzeitpunkte konstant gehalten. Die beschriebenen Tests wurden neben anderen Messungen an allen drei Terminen zweimal hintereinander durchgeführt (Test und Retest).

Die Kontrollgruppe bestand aus zehn versorgten Hörgeschädigten, davon fünf Frauen und fünf Männer. Ihr Alter lag zwischen 64-81 Jahren ( $\bar{X}$  73,4  $\pm$  6,0 Jahre). Ihr mittlerer Hörverlust (siehe Abb. 1) lag bei einem PTA (für 500, 1k, 2k, 4kHz) von 46,7  $\pm$  7,4 dB HL.

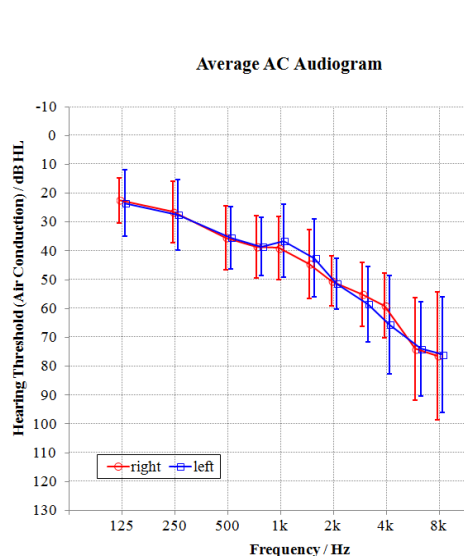


Abbildung 1: Gemittelter Hörverlust für die Kontrollgruppe (N=10).



Abbildung 2: Boxplots für die Anzahl der richtig erkannten Zahlen in den Olsa-Sätzen.

## Ausgewählte Ergebnisse Börsenszenario

Die Messergebnisse für das Börsenszenario zeigt Abb. 2. Für die drei Messzeitpunkte sind jeweils die richtig erkannten Zahlen aus den Olsa-Sätzen der Targetsprecherin für Test (grün) und Retest (rot) über die zehn Probanden als Boxplot zusammen gefasst. Eine zweifaktorielle ANOVA für messwiederholte Daten (3-stufiger Faktor Messzeitpunkt: T1, T2, T3 und 2-stufiger Faktor Test/Retest) ergibt mit  $F(1,9)=6.61$ ,  $p=0.030$  einen signifikanten Unterschied für den Faktor Test (MW=11.7, SE=1.1) und Retest (MW=12.7, SD=1.0), während der Faktor Messzeitpunkt keine signifikanten Unterschiede zeigt  $F(2,18)=0.54$ ,  $p=0.593$ .

Um auch die Vergleichbarkeit der individuellen Messdaten über die Messzeitpunkte bzw. für Test/Retest zu untersuchen, sind die zu vergleichenden Messungen jeweils gegeneinander aufgetragen, siehe Abb. 3. Auf der linken Seite ist die Bewertung des Retests über die Bewertung des Tests für jeden einzelnen Probanden und für alle Messzeitpunkte aufgetragen. Eine Pearson-Korrelation liefert für T1 ein Korrelationskoeffizient  $r = 0.86$  ( $p=0.002$ ), für T2  $r=0.83$  ( $p=0.003$ ) und für T3  $r=0.91$  ( $p<0.001$ ). Die Korrelationskoeffizienten sind also für alle drei Messzeitpunkte vergleichbar und zeigen eine hohe Korrelation, d.h. einen linearen Zusammenhang zwischen Test und Retest an.

Die rechte Seite der Abb. 3 verdeutlicht den Zusammenhang zwischen den drei Messzeitpunkten, aufgetragen für die individuellen Mittelwerte aus Test und Retest. Auch hier zeigt eine Pearson-Korrelation hohe Korrelationskoeffizienten, sodass ein linearer Zusammenhang der Ergebnisse zwischen den Messzeitpunkten angenommen werden kann: T2 vs. T1:  $r=0.80$ ,  $p=0.005$ ; T3 vs. T1:  $r=0.77$ ,  $p=0.009$ ; T3 vs. T2:  $r=0.89$ ,  $p<0.001$ .

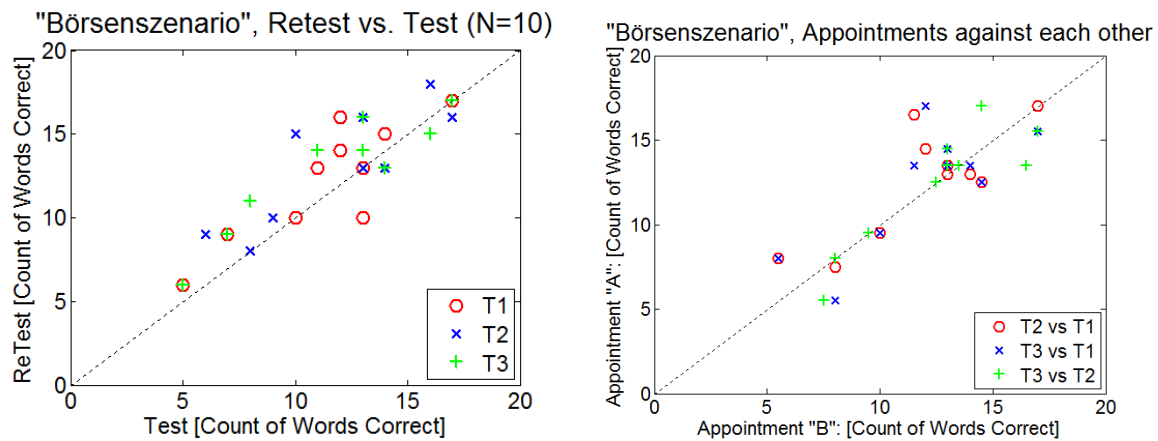


Abbildung 3: Individuelle Messdaten für den Vergleich Test-Retest an den drei Messzeitpunkten (links) und für den Vergleich der drei Messzeitpunkte (Test-Retest jeweils gemittelt) (rechts).

### Ausgewählte Ergebnisse LST

Die gemessenen Daten für den LST zeigen die Abb. 4 und Abb. 5. Im Folgenden werden ausschließlich die Werte für die erinnerten Wörter ausgewertet<sup>2</sup>. Ein Friedman-Test über die beiden Blöcke (4er, 6er), die drei Termine (T1, T2, T3) und die beiden Wiederholungen (Test/Retest) zeigt keine signifikanten Unterschiede für die beiden Satztypen HP und LP. Das bestätigt die Daten von Neher et al. (2018), der auch keine Unterschiede zwischen diesen beiden Satztypen finden konnte. Ebenso zeigt der Friedman-Test keinen signifikanten Unterschied zwischen den beiden Blöcken (4er und 6er) über die drei Termine, die beiden Wiederholungen und die Satztypen an<sup>3</sup>. Da es keine Unterschiede gibt, werden die Daten für beide Satztypen und beide Blockgrößen addiert und alle weiteren Auswertungen nur noch mit den so zusammen gefassten Daten durchgeführt. Der Wertebereich der richtig erinnerten Wörter liegt damit zwischen 0 und 20 Wörtern. Eine messwiederholte, 2-faktorielle ANOVA<sup>4</sup> auf diesen Daten zeigt keine signifikanten Mittelwertunterschiede zwischen Test/Retest oder zwischen den drei Terminen, siehe Abb. 4. Eine analoge Betrachtung der individuellen Daten wie für das Börsenszenario zeigt Abb. 5.

Die Pearson-Korrelation für Test-Retest Daten liefert für T1 ein Korrelationskoeffizient  $r$  von 0.26 ( $p=0.49$ ), für T2  $r=0.57$  ( $p=0.11$ ) und für T3  $r=0.30$  ( $p=0.44$ ). Die Korrelationskoeffizienten sind klein und es ist nicht von einem linearen Zusammenhang auszugehen (Abb. 5, links). Die rechte Seite der Abb. 5 verdeutlicht den Zusammenhang zwischen den drei Messzeitpunkten, aufgetragen für die individuellen Mittelwerte aus Test und Retest. Hier zeigt die Pearson-Korrelation höhere Korrelationskoeffizienten zwischen den Messzeitpunkten an: T2 vs. T1:  $r=0.51$ ,  $p=0.16$ ; T3 vs. T1:  $r=0.73$ ,  $p=0.025$ ; T3 vs. T2:  $r=0.58$ ,  $p=0.10$ .

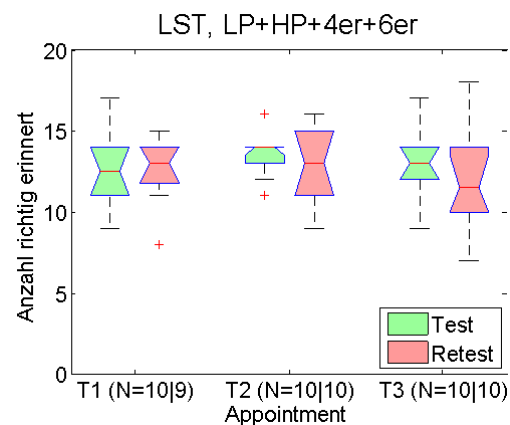


Abbildung 4: Boxplots für die Summe richtig erinnelter Wörter.

### Zusammenfassung

Das verwendete Messverfahren „Börsenszenario“ liefert für die Kontrollgruppe konstante Ergebnisse für die Vergleiche zwischen Test – Retest und ist über 12 Wochen stabil bezüglich individueller Daten als auch deren Mittelwerte. Auftretende Unterschiede der Daten für die Experimentalgruppen lassen sich anhand dieser Ergebnisse daher auf das Treatment der Probanden zurückführen.

<sup>2</sup> Es wurden, wie beabsichtigt, nahezu allen Wörter richtig verstanden. Dadurch war es möglich, die Merkleistung zu erheben, ohne die nicht verstandenen Wörter berücksichtigen zu müssen.

<sup>3</sup> Die Unterschiede zwischen den 4er- und 6-er Gruppen sind nur dann nicht unterschiedlich, wenn die absoluten Zahlen betrachtet werden (relative Betrachtungsweise wäre signifikant unterschiedlich)

<sup>4</sup> ohne einen Probanden, weil für diesen Datensatz eine Messung fehlt, d.h. N=9.

Für das Verfahren „Listening Span Test“ unterscheiden sich die Mittelwerte über die Probanden der Kontrollgruppe ebenfalls nicht, d.h. auf gruppenebene kann ein Treatment untersucht werden, wenn eine ausreichende Gruppenstärke gewährleistet wird. Die individuellen Daten streuen stärker als für das Börsenszenario, sowohl für Test/Retest an einem Messzeitpunkt als auch über die drei Messzeitpunkte hinweg. Individuelle Veränderungen bzgl. eines Treatments sind daher nicht aussagekräftig. Wie bereits in anderen Studien gefunden, zeigen für die Kontrollgruppe die HP und LP Sätze keine Unterschiede, ebenso wenig wie die 4er bzw. 6er Blöcke bei absoluter Betrachtungsweise.

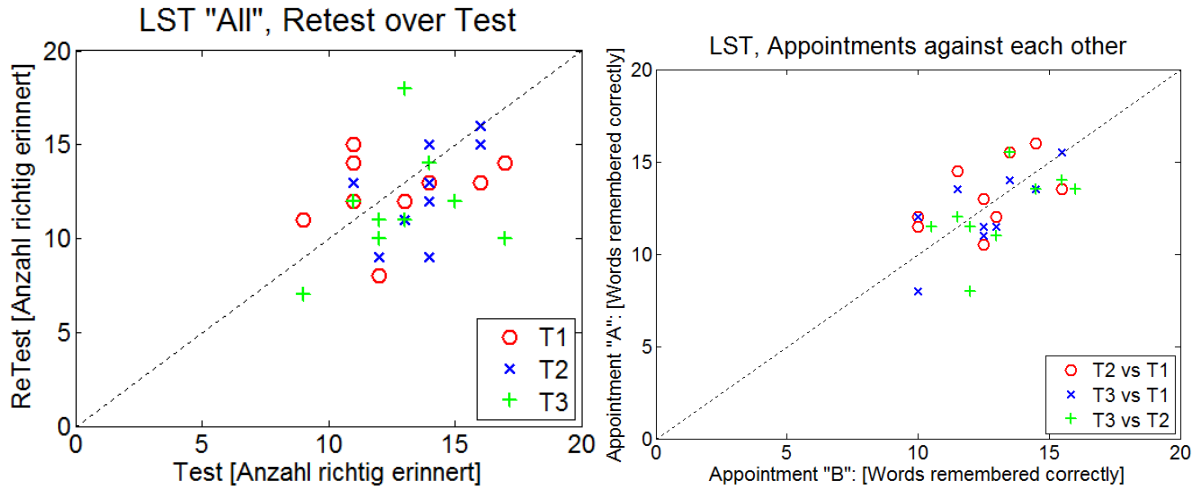


Abbildung 5: Individuelle Daten. Links: Vergleich Test-Retest an den drei Messzeitpunkten, rechts: Vergleich der drei Messzeitpunkte (Mittelwerte über Test-Retest).

## Literatur

- Akeroyd M.A. (2008) Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int J Audiol.* 47, 53-71.
- Baddeley A.D. und Hitch G. (1974) *Working memory*. In: G. H. Bower (Hrsg.): *The psychology of learning and motivation: Advances in research and theory*. Vol. 8, Academic Press, New York 1974, S. 47–89.
- Doherty K.A. und Desjardins J.L. (2015) The benefit of amplification on auditory working memory function in middle-aged and young-older hearing impaired adults. *Front Psychol.* 2015; 6: 721.
- Getzmann S., Wascher E. und Falkenstein M. (2015) What does successful speech-in-noise perception in aging depend on? Electrophysiological correlates of high and low performance in older adults. *Neuropsychologia* 70, 43-57.
- Hochmuth S., Jürgens T., Brand T. und Kollmeier B. (2013) Talker- and language-specific effects on speech intelligibility in noise assessed with bilingual talkers: Which language is more robust against noise and reverberation? *International Journal of Audiology*, Volume 54, 23-34
- Kalikow D.N., Stevens K.N. und Elliott L.L. (1977) Development of a test of speech intelligibility in noise using sentence materials with controlled predictability. *J Acoust Soc Am* 61, 1337-1351.
- Neher T., Wagener K.C. und Fischer R.-L. (2018) Hearing aid noise suppression and working memory function, *International Journal of Audiology*, 57:5, 335-344
- Tschopp K. und Züst H.J. (1994) Performance of normally hearing and hearing -impaired listeners using a German version of the SPIN test. *Scand Audiol* 23, 241-247
- Wagener K.C., Vormann, M., Fischer R.L. und Neher T. (2018) Akklimatisierung, Präferenz und Performanz. 21. Jahrestagung der Dt. Ges. für Audiologie (DGA), in Bearbeitung