

Evaluation des Ziffern-Tripel-Tests über Kopfhörer und Telefon

Kirsten Carola Wagener¹, Timo Bräcker², Thomas Brand³, Birger Kollmeier³

¹Hörzentrum Oldenburg GmbH, ²Institut für Hörtechnik und Audiologie, Fachhochschule Oldenburg/Ostfriesland/Wilhelmshaven, ³Medizinische Physik, Carl von Ossietzky Universität Oldenburg

Einleitung

Auf der DGA 2005 wurde die deutsche Version des Zahlen-Tripel-Tests für Screening-Sprachverständlichkeitstests im Störgeräusch vorgestellt (Wagener et al., 2006, niederländisches Vorbild: Smits et al. 2004). Da der Test nur Zahlen-Tripel (zusammengesetzt aus einsilbigen Ziffern) verwendet (z.B. 0-9-8), ist dieser Test für die Durchführung über ein Telefon geeignet (Antworten werden über die Telefontastatur gegeben). Ein Screening-Hörtest über das Telefon kann anonym von zu Hause aus durchgeführt werden und bringt für den Anrufer eine Einschätzung seiner Hörfähigkeit aus unabhängiger Quelle.

Wegen des stark eingeschränkten Wortmaterials (9 verschiedene Ziffern) ist er auch als klinischer Sprachtest einsetzbar, der auch von Personen mit eingeschränktem Wortschatz durchgeführt werden kann.

Dieser Beitrag beschreibt die Evaluation des Tests mit Normalhörenden, bei denen zum einen die Vergleichbarkeit der Testlisten untersucht wurde und zum anderen die Referenzfunktionen erhoben wurden (Bräcker, 2005).

Sprachmaterial und Optimierung des Tests

Es werden alle neun Ziffern von 0 bis 9 ohne die zweisilbige 7 im Wortmaterial verwendet. Diese Ziffern ergeben zu Tripeln zusammengesetzt (z.B. 0-9-1) das Sprachmaterial des Tests. Der Test wird im Störgeräusch durchgeführt, so dass der Einfluss des Absolutpegels des jeweiligen Telefons minimiert wird. Die Details des Sprachmaterials sind in Wagener et al. (2006) beschrieben.

Der Test ist zur Bestimmung der Sprachverständlichkeitsschwelle im Störgeräusch gedacht (L_{mid} : Signal-Rausch-Verhältnis, bei dem 50% der Sprache verstanden wird mit Berücksichtigung der Ratewahrscheinlichkeit, vgl. Gl. 1). Für eine hohe statistische Genauigkeit und Reproduzierbarkeit der Ergebnisse ist für diese Messungen eine steile Verständlichkeitsfunktion des Sprachtests wichtig. Der Zahlen-Tripel-Tests wurde dahingehend durch Messungen mit normalhörenden Probanden und einen darauf basierenden Verständlichkeitsausgleich einzelner Ziffern optimiert (Wagener et al., 2006).

$$SV(L) = 1/A + \frac{A-1}{A(1 + e^{\frac{4A}{A-1}s(L_{mid}-L)})} \quad (\text{Gleichung 1})$$

- SV: Sprachverständlichkeit
- L: Signal-Rausch-Verhältnis der Darbietung
- A: Anzahl der Antwortalternativen (hier 10 für alle Ziffern auf dem Telefon)
- L_{mid} : Mittelpunkts-SNR, entspricht Sprachverständlichkeitsschwelle mit Berücksichtigung der Ratewahrscheinlichkeit (bei zehn Antwortalternativen 55%)
- s: Steigung am L_{mid}

Die Optimierung brachte eine Erhöhung der Steigung s von 12,2%/dB vor der Optimierung auf 17,6%/dB (Kopfhörerdarbietung) bzw. 17,2%/dB (Telefonbandfilterung) nach der Optimierung. Der L_{mid} betrug nach der Optimierung -10,3 dB SNR bei Darbietung über Kopfhörer und -10,6 dB SNR für die telefonbandgefilterten Signale (Wagener et al., 2006). Da während der Optimierung die Verständlichkeitsfunktionen für die einzelnen Ziffern ausgewertet wurden, wurde bei der Auswertung Ziffern-Scoring verwendet, d.h. jede Ziffer im Tripel wurde als richtig oder falsch gewertet. Für die geplante Anwendung als Screening-Test über Telefon ist es jedoch besser, Tripel-Scoring zu verwenden, d.h. für ein korrektes Tripel müssen alle drei Ziffern richtig sein. Bei der Verwendung von Ziffern-Scoring wäre es zu entscheiden, ob die Antwort ‚2-3‘ bei der Darbietung ‚2-1-3‘ als zwei richtige Ziffern zu werten ist (erste und dritte Position richtig) oder nur als eine richtige Ziffer (nur erste Position). Daher wurden in der hier vorgestellten Evaluation des Tests beide Auswerte-Methoden verwendet (Ziffern-Scoring für den Vergleich mit den Optimierungsdaten, Tripel-Scoring für die Erhebung der Referenzfunktionen).

Die Optimierung des Tests wurde nicht über ein reales Telefon durchgeführt, sondern mit telefonbandpassgefilterten Signalen, die über Kopfhörer dargeboten wurden. In der Evaluation des Tests wurden reale Telefone verwendet.

Evaluation des Tests

Der Ziffern-Tripel-Test wurde sowohl breitbandig über Kopfhörer als auch über reale Telefone im Hinblick auf die folgenden Fragestellungen evaluiert:

- 1) Sind die Testlisten hinsichtlich ihrer Verständlichkeit äquivalent?
- 2) Wie sind die Referenzfunktionen über Kopfhörer und Telefon?
- 3) Ist das Ergebnis abhängig vom Telefon?
- 4) Ist die vorhergehende Ziffer in einem Tripel wichtig für die Verständlichkeit einer gegebenen Ziffer (gibt es Kontexteffekte)?

Evaluationsmessungen

Der Test besteht aus sechs Testlisten mit jeweils 27 Ziffern-Tripeln. Jede Liste wurde zweimal pro Person gemessen. Alle Listen wurden an drei verschiedenen Signal-Rausch-Verhältnissen SNRs gemessen. Die Reihenfolge der Listen und der SNRs war randomisiert.

Jedes Tripel wurde durch die Wörter „Die Ziffern...“ angekündigt. Das Störgeräusch begann 500 ms vor jedem Sprachsignal und endete 500 ms nach jedem Sprachsignal.

Die Auswertung erfolgte sowohl mit Ziffern- als auch mit Tripel-Scoring. An die Verständlichkeitsdaten aller ProbandInnen wurde die logistische Modellfunktion (Gl. 1) angepasst, um die listenspezifischen Verständlichkeitsfunktionen zu erhalten.

Messungen über Kopfhörer

Die Messungen erfolgten über einen Kopfhörer (Sennheiser HDA200), der mit einem FIR Filter (Ordnung 800) freifeldentzerrt wurde über die Oldenburger Messprogramme (HörTech gGmbH). Die Messungen wurden geschlossen durchgeführt, d.h. die Antwort wurde von den ProbandInnen bei den Kopfhörmessungen auf einem Touchscreen eingegeben, der eine Telefontastatur darstellte. Der Darbietungspegel des Störgeräuschs betrug immer 65 dB SPL.

Es wurde an den drei SNRs -12; -10.5; -9 dB SNR gemessen, so dass Verständlichkeiten unterhalb und oberhalb von 50% erfasst wurden.

15 normalhörende Probanden nahmen an den Messungen teil.

Messungen über Telefon

Für die Messungen über Telefon wurde eine Telefonanbindung in die Oldenburger Messprogramme integriert: nach Auswahl der Testliste und des SNR durch den Versuchsleiter hat das Messprogramm auf den Anruf gewartet. Die Darbietung des Tests und die Antwortgabe per Telefon wurden über eine ISDN-Einsteckkarte (AVM Fritz! Card) realisiert.

Die Telefontastatur wurde als Eingabe-Medium verwendet. Bei den Messungen über Telefon konnte der

Darbietungspegel des Störgeräuschs nicht kontrolliert werden. Es wurde darauf geachtet, dass das Störgeräusch gut hörbar war. Wagener und Brand (2005) zeigten, dass der dargebotene SNR wichtig für die Sprachverständlichkeit ist, nicht der Darbietungspegel des Störgeräuschs. Dies gilt solange das Störgeräusch mit einer entsprechenden Lautheit zwischen ‚leise‘ und ‚laut‘ dargeboten wird.

Vormessungen über reale Telefone zeigten, dass im Gegensatz zur Optimierung mit telefonbandpassgefilterten Signalen die Verständlichkeiten mit realen Telefonen zu höheren SNRs verschoben sind. Daher wurden die Messungen an andern SNRs durchgeführt als die Kopfhörmessungen. Es wurde an den drei SNRs -11; -8.5; -6 dB SNR gemessen, so dass Verständlichkeiten unterhalb und oberhalb von 50% erfasst wurden. Vor den eigentlichen Messungen wurde jeweils eine Trainingsmessung mit 9 Ziffern-Tripeln durchgeführt.

Es wurden Messungen mit einem Referenztelefon im Labor (Telecom Octophon F) mit 15 normalhörenden Probanden durchgeführt. Zudem wurden mit zehn normalhörenden Probanden Messungen mit deren eigenen Telefonen zu Hause (externe Telefone) durchgeführt. Dabei wurden sowohl Telefone mit als auch ohne Schnur verwendet, jedoch keine Mobiltelefone.

Ergebnisse und Diskussion

Tabelle 1 zeigt die Ergebnisse der Evaluationsmessungen sowohl für Ziffern-Scoring (Zeile 1 und 2) als auch für Tripel-Scoring (Zeile 2).

	Kopfhörer		Referenztelefon	
	$L_{\text{mid}} \pm \sigma_{\text{mid}}$ [dB SNR]	s [%/dB]	$L_{\text{mid}} \pm \sigma_{\text{mid}}$ [dB SNR]	s [%/dB]
Ziffern-Scoring, Listen gepoolt	$-11 \pm 0,6$	15	$-8,7 \pm 0,6$	12
Ziffern-Scoring, Probanden gepoolt	$-11 \pm 0,2$	14,5	$-8,7 \pm 0,2$	11,8
Tripel-Scoring, Probanden gepoolt	$-9,3 \pm 0,2$	19,6	$-6,4 \pm 0,1$	16,1

Tabelle 1: Ergebnisse der Evaluationsmessungen. Die Werte sind jeweils für die Kopfhörmessungen (linker Teil) und für die Telefonmessungen (rechter Teil) aufgeführt. Gegeben sind die gemittelten L_{mid} Werte mit den dazugehörigen Standardabweichungen σ_{mid} sowie die mittleren Steigungen der Verständlichkeitsfunktionen s . Gemittelt wurde entweder über die Probanden (1. Zeile) oder über Testlisten (2. und 3. Zeile).

Die Unterschiede zwischen den Probanden (ausgedrückt durch die Standardabweichung des L_{mid} bei gepoolten Testlisten) sind für Kopfhörer- und Referenz-Telefondarbietung drei Mal so groß wie die Unterschiede zwischen den Testlisten (Standardabweichung des L_{mid} bei gepoolten Probanden). Auch die Varianzanalyse der Verständlichkeiten an den drei gemessenen SNR zeigt, dass die Testlisten vergleichbar sind.

Kopfhörmessungen:

$$\hat{F}_{-12\text{dB}} = 0,84; \hat{F}_{-10,5\text{dB}} = 1,40; \hat{F}_{-9\text{dB}} = 0,46;$$

Referenz-Telefonmessungen:

$$\hat{F}_{-11\text{dB}} = 0,52; \hat{F}_{-8,5\text{dB}} = 1,02; \hat{F}_{-6\text{dB}} = 0,35 \quad \text{jeweils}$$

kleiner als $F_{(5;54;5\%)} \approx 2,3$.

Wie erwartet zeigen sich beim Tripel-Scoring ein höherer L_{mid} und eine höhere Steigung der Verständlichkeitsfunktion.

Abbildung 1 zeigt die Verständlichkeitsfunktionen des Ziffern-Tripel-Tests bei Durchführung über das Referenztelefon (links) und über die externen Telefone (rechts). Die einzelnen Kurven zeigen die Verständlichkeitsfunktionen der einzelnen Probanden mit Tripel-Scoring. Die zugehörigen Parameter der mittleren Verständlichkeitsfunktion geben die Referenzwerte für Normalhörende an. Es zeigte sich, dass die Ergebnisse unabhängig vom verwendeten Telefon sind.

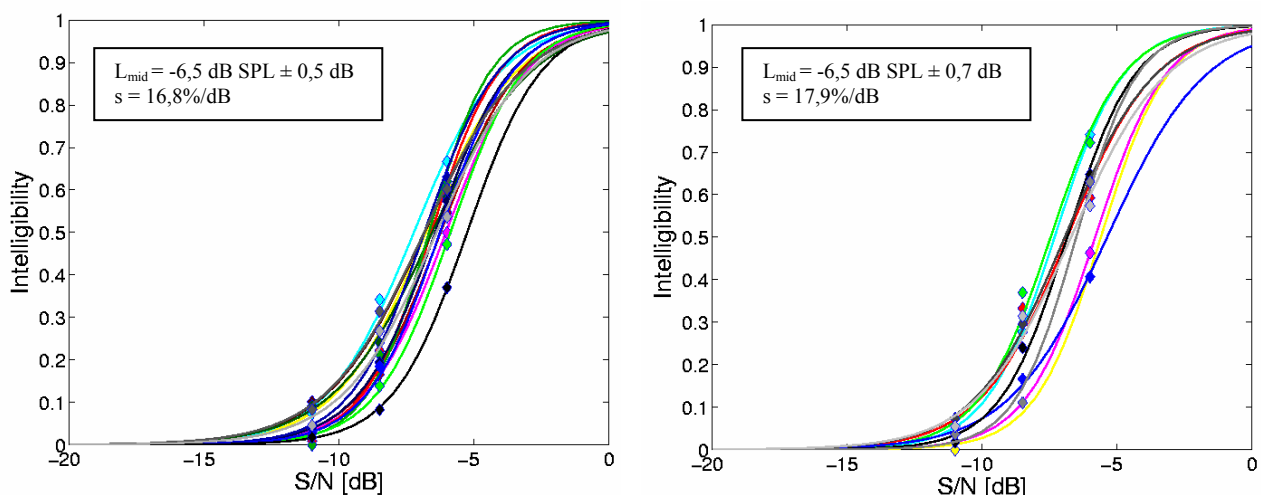


Abbildung 1: Verständlichkeitsfunktionen des Ziffern-Tripel-Tests bei Durchführung über das Referenztelefon (links) und über die externen Telefone (rechts). Die einzelnen Kurven zeigen die Verständlichkeitsfunktionen der einzelnen Probanden mit Tripel-Scoring.

Es zeigte sich zudem, dass es keinen signifikanten Einfluss der vorhergehenden oder nachfolgenden Ziffer auf die Verständlichkeit einer gegebenen Ziffer gibt. Das bedeutet, dass die einzelnen Ziffern-Tripel online während des Tests aus den einzelnen Ziffern zusammengesetzt werden können und somit die Anzahl der Testlisten nicht auf sechs beschränkt sein muss.

In den Optimierungsmessungen verhielten sich die breitbandige Version und die telefonbandpassgefilterte Version des Tests bzgl. der Verständlichkeit gleich. Über reale Telefone verschiebt sich jedoch die Verständlichkeitsfunktion zu höheren SNRs. Dies kann dadurch erklärt werden, dass die Übertragungsfunktion der realen Telefone im Frequenzbereich von 200-1000 Hz nicht mit dem Telefonbandpass übereinstimmt.

Im Vergleich zu anderen deutschen Sprachverständlichkeitstests im Störgeräusch (Göttinger und Oldenburger Satztest, Kollmeier und Wesselkamp, 1997 bzw. Wagener et al, 1999) und zum niederländischen Telefentest von Smits et al (2004) liegt der L_{mid} für Kopfhörermessungen um ca. 4 dB niedriger. Die Steigung der Verständlichkeitsfunktionen ist vergleichbar.

Zusammenfassung und Ausblick

In diesem Beitrag wurde die Evaluation des deutschen Ziffern-Tripel-Tests vorgestellt. Dieser Test ist zum einen für Screening-Messungen der Sprachverständlichkeit im Störgeräusch über Telefon geeignet. Zum anderen ist er auch als Sprachverständlichkeitstest bei eingeschränktem deutschen Wortschatz einsetzbar, da er nur aus den Ziffern von 0 bis 9 besteht. Durch das geschlossene Antwortdesign kann der Test auch in der passenden Sprache für Nichtmuttersprachler eingesetzt werden (die Versuchsleitung muss die Testsprache nicht unbedingt beherrschen). Die Messdauer einer Testliste mit 27 Tripeln (81 Ziffern) beträgt 3-4 min.

Die Vergleichbarkeit der Testlisten ist sehr hoch. Es zeigte sich kein Einfluss des verwendeten Telefons auf die Sprachverständlichkeit. Dies gilt für die verwendeten Telefone mit und ohne Schnur, Mobiltelefone wurden nicht untersucht.

Danksagung

Dieser Beitrag wurde gefördert von der EU FP6-004171 HearCom.

Literatur

- Bräcker T (2005) *Evaluation des deutschen Ziffern-Tripel-Tests*. Diplomarbeit. Institut für Hörtechnik und Audiologie. Fachhochschule Oldenburg/Ostfriesland/Wilhelmshaven
- Kollmeier B, Wesselkamp M (1997) *Development and evaluation of a German sentence test for objective and subjective speech intelligibility assessment*. J. Acoust. Soc. Am. 102, 2412-2421.
- Smits C, Kapteyn TS, Houtgast T (2004) *Development and validation of an automatic speech-in-noise screening test by telephone*. Int. J. Audiol. 43, 15-28.
- Wagener K, Brand T, Kollmeier B (1999) *Entwicklung und Evaluation eines Satztests für die deutsche Sprache I-III: Design, Optimierung und Evaluation des Oldenburger Satztests*. ZfA 38 (1-3), 4-15, 44-56, 86-95.
- Wagener KC, Brand T (2005) *Sentence intelligibility in noise for listeners with normal hearing and hearing impairment: Influence of measurement procedure and masking parameters*. Intern J Audiol 44 (3), 144-157.
- Wagener KC, Eeenboom F, Brand T, Kollmeier B (2006) *Ziffern-Tripel-Test: Sprachverständlichkeitstest über das Telefon*. DGA 8. Jahrestagung Göttingen 2005, TagungsCD, ISBN = 3-9809869-4-2.