

Ziffern-Tripel-Test: Sprachverständlichkeitstest über das Telefon

Kirsten Carola Wagener, Frauke Eeenboom, Thomas Brand, Birger Kollmeier

Medizinische Physik, Carl von Ossietzky Universität Oldenburg, Email: kirsten.wagener@uni-oldenburg.de

Einleitung

Oftmals wird die Diagnose einer Schwerhörigkeit verzögert, weil der Besuch beim HNO-Arzt oder Hörgeräte-Akustiker zu aufwendig und evtl. auch zu verbindlich scheint. Daher ist die Motivation, einen Screening-Hörtest über das Telefon zu realisieren. Dieser kann anonym von zu Hause aus durchgeführt werden und bringt eine Einschätzung der Hörfähigkeit aus unabhängiger Quelle. Bei Sprachverständlichkeitstests werden dem Probanden meist Wörter oder Sätze vorgespielt, die entweder wiederholt oder aus einer Anzahl von Antwortalternativen ausgewählt werden sollen. Bei der Durchführung über das Telefon bietet sich die Telefontastatur als Eingabemedium der Antwort an. Daher wurde von Smits et al. (2004) ein niederländischer Sprachtest über das Telefon entwickelt, der als Sprachmaterial Zifferntripel bestehend aus einsilbigen Ziffern verwendet. Dieser Beitrag beschreibt die Übertragung dieses Tests in die deutsche Sprache und die Optimierung des Tests.

Sprachmaterial

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. Im niederländischen Vorbild wurde alle im Sprachmaterial vorkommende Tripel einzeln aufgesprochen. Im Gegensatz dazu wurden orientiert am Oldenburger Satztest (Wagener et al., 1999) die Tripel im deutschen Test aus einzelnen Ziffern zusammengesetzt: Durch die Aufnahme von neun verschiedenen Ziffern-Tripeln wurden alle neun Ziffern an jeder Position im Tripel (vorne, Mitte, hinten) genau einmal aufgesprochen. Um für die spätere Optimierung des Sprachmaterials mehr Material zur Verfügung zu haben, wurden zwei verschiedene Listen mit jeweils neun Ziffern-Tripel mehrfach aufgesprochen. Diese aufgesprochenen Tripel wurden auseinandergeschnitten. Ein Test-Tripel wird aus den geschnittenen einzelnen Ziffern zusammengesetzt, wobei die Stellung im Tripel immer mit der Stellung im Tripel bei der Aufnahme übereinstimmt (eine 3, die an erster Stelle im Tripel aufgesprochen wurde, wird immer für die erste Position im Tripel verwendet). Die Tripel werden so zu Testlisten zusammengestellt, dass jede Ziffer an jeder Position im Tripel genau dreimal pro Liste vorkommt. Dies führt zu einer Gesamtzahl von 27 Tripeln pro Liste. Im niederländischen Test besteht jede Testliste aus 23 Tripeln.

Aus dem deutschen Sprachmaterial wurde durch 30fache Überlagerung mit zufällig eingefügten Pausen variabler Länge das Störgeräusch generiert. Dieses zeigt

das gleiche Langzeitspektrum wie das verwendete Sprachmaterial und verdeckt dieses daher optimal. Im niederländischen Test wurde das Störgeräusch im Langzeitspektrum an das Sprachmaterial angepasst.

Optimierung des Sprachmaterials

Die Genauigkeit der Bestimmung der 50%-Schwelle ist proportional zur Steigung der Verständlichkeitsfunktion. Diese Steigung ist höher, je homogener die Verständlichkeiten der verwendeten Testwörter sind (probabilistisches Modell nach Kollmeier, 1990, siehe auch Wagener et al 1999). Daher wurden in Optimierungsmessungen die ziffernspezifischen Verständlichkeitsfunktionen ausgemessen (Abhängigkeit der Sprachverständlichkeit SV vom dargebotenen Signal-Rausch-Verhältnis L). Sind diese bekannt, kann durch Aussteuerungsänderungen einiger Ziffern (z.B. „schwer zu verstehende“ Ziffern etwas im Pegel anheben) die Homogenität der Verständlichkeiten gesteigert werden.

Die Optimierungsmessungen wurden zum einen über Kopfhörer mit dem breitbandigen Sprachmaterial (Bandbreite 22050 Hz) und über Kopfhörer mit telefonbandpassgefiltertem Sprachmaterial (300-3400 Hz, nach ITU 2003) durchgeführt.

Optimierungsmessungen

An den Messungen nahmen 12 normalhörende ProbandInnen teil. Sowohl für das breitbandige als auch für das telefonbandbegrenzte Sprachmaterial wurden zwei Versionen mit sechs verschiedene Testlisten erstellt. Die beiden Versionen unterschieden sich dadurch, dass sie zwei verschiedene Aufsprachen der Ziffern enthielten. Die einzelnen Listen wurden an sieben verschiedenen Signal-Rausch-Verhältnissen zwischen -18 und -2 dB SNR gemessen. Die Messung erfolgte über Kopfhörer (Sennheiser HDA200), die mit einem FIR Filter (Ordnung 800) freifeldentzerrt wurden über die Oldenburger Messprogramme (HörTech gGmbH). Jedes Tripel wurde durch die Wörter „Die Ziffern...“ angekündigt. Der Darbietungspegel des Störgeräuschs betrug immer 65 dB SPL. Das Störgeräusch begann 500 ms vor jedem Sprachsignal und endete 500 ms nach jedem Sprachsignal. Die Messungen wurden geschlossen durchgeführt, d.h. die Antwort wurde von den ProbandInnen auf einem Touchscreen eingegeben, der eine Telefontastatur darstellte. Es wurde digit scoring verwendet, d.h. die richtigen und falschen Ziffern wurden gezählt, somit war die Reihenfolge der Ziffern im Tripel bei der Antwort relevant. An die Verständlichkeitsdaten aller ProbandInnen wurde die logistische Modellfunktion (Gl. 1) angepasst, um die

zifferspezifischen Verständlichkeitsfunktionen zu erhalten.

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

(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

Auswahl und Verständlichkeitsangleich der Ziffern

Die oberste Zeile von Tabelle 1 zeigt die Ergebnisse der Optimierungsmessungen. L_{mid} entspricht dem Mittelpunkts-SNR der Verständlichkeitsfunktion unter Berücksichtigung der Ratewahrscheinlichkeit, gemittelt über alle Ziffern. σ_{Ziffer} zeigt die Standardabweichung der L_{mid} zwischen den einzelnen Ziffern. m_{Ziffer} gibt die mittlere Steigung der zifferspezifischen Verständlichkeitsfunktionen an. m_{Liste} ist die vorhergesagte Steigung des gesamten Tests (vorhergesagt mit dem probabilistischen Modell). Die Steigung des nicht optimierten Tests ist mit 12%/dB relativ gering. Daher wurde zunächst aus den beiden verschiedenen Aufspracheversionen jeder einzelnen Ziffer diejenige Aufsprache ausgewählt, deren zugehöriger L_{mid} Wert näher zum Gesamtmittel- L_{mid} liegt. Dadurch konnte die Steigung des Tests auf 14%/dB gesteigert werden. Um die Verständlichkeiten der einzelnen Ziffern noch stärker zu homogenisieren und somit eine steilere Gesamtverständlichkeitsfunktion zu erhalten, wurde die Auspegelung von einzelnen Ziffern verändert. Bei diesem Verständlichkeitsangleich wurden „schwer zu verstehende“ Ziffern, deren L_{mid} um mehr als 0,5 dB vom Gesamtmittel abwich, im Pegel angehoben. „Leicht zu verstehende“ Ziffern wurde analog im Pegel abgesenkt. Die Pegeländerungen betragen maximal 2 dB (breitbandig) bzw. 2,8 dB (telefonbandbegrenzt). Somit wurde sichergestellt, dass keine unnatürlichen Pegelsprünge innerhalb der Ziffern-Tripel auftreten. Durch diesen Verständlichkeitsangleich wurde die Varianz der L_{mid} -Verteilung der Ziffern reduziert ($\sigma_{Ziffer}=0,4$ dB), so dass die zu erwartende Steigung für den Test auf 17%/dB erhöht wurde.

Die zu erwartenden Verständlichkeitsfunktionen für den Test sind über Kopfhörer im breitbandigen und im telefonbandbegrenzten Fall sehr ähnlich (s. Tab. 1). Im Vergleich zu anderen deutschen Sprachverständlichkeitstests im Störgeräusch (Göttinger und Oldenburger Satztest, Kollmeier und Wesselkamp, 1997 bzw. Wagener et al, 1999) und zum niederländischen Telefontest von Smits et al (2004) liegt der L_{mid} um ca. 3 dB niedriger. Die Steigung der Verständlichkeitsfunktionen ist vergleichbar.

	breitbandig			telefonbandbegrenzt		
	$L_{mid} \pm \sigma_{mid}$ [dB]	m_{Ziffer} [%/dB]	m_{Liste} [%/dB]	$L_{mid} \pm \sigma_{mid}$ [dB]	m_{Ziffer} [%/dB]	m_{Liste} [%/dB]
Vor Optimierung	-10,2 ± 2,2	17,9	12,2	-10,5 ± 2,0	16,5	12,2
Zifferauswahl	-10,2 ± 1,5	17,6	14,1	-10,5 ± 1,4	17,2	14,2
Verständlichkeitsangleich	-10,3 ± 0,4	17,6	17,6	-10,6 ± 0,4	17,2	17,2

Tabelle 1: Ergebnisse der Optimierungsmessungen und Auswirkungen der Optimierungsstufen. Die Werte sind jeweils für das breitbandige Sprachmaterial (linker Teil) und für das telefonbandbegrenzte Sprachmaterial (rechter Teil) aufgeführt. Gegeben sind die über alle Ziffern gemittelten L_{mid} Werte mit den dazugehörigen Standardabweichungen σ_{mid} , die mittleren Steigungen der zifferspezifischen Verständlichkeitsfunktionen m_{Ziffer} sowie die aus dem probabilistischen Modell errechneten Vorhersagen für die Steigung der listenspezifischen Verständlichkeitsfunktionen m_{Liste} .

Zusammenfassung und Ausblick

In diesem Beitrag wurde die Realisation und Optimierung des deutschen Ziffern-Tripel-Tests vorgestellt. Dieser ist zum Einen für Screening-Messungen der Sprachverständlichkeit im Störgeräusch übers Telefon geeignet. Zum anderen ist er auch als Sprachverständlichkeitstest für ProbandInnen mit 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 (der/die VersuchsleiterIn muss die Testsprache nicht unbedingt beherrschen). Die Messdauer einer Testlisten mit 27 Tripeln (81 Ziffern) beträgt 3-4 min. Die Steigung der Verständlichkeitsfunktion des Tests ist vergleichbar zu der von Satztests im Störgeräusch.

Die aufgrund der Optimierung vorhergesagten Verständlichkeitsfunktionen werden in unabhängigen Evaluationsmessungen über Kopfhörer und über reale Telefone überprüft. Dabei wird die Homogenität der Testlisten überprüft sowie die Referenzfunktionen für Normalhörende bestimmt.

Danksagung

Dieser Beitrag wurde gefördert von der EU FP6-004171 HearCom

Literatur

- ITU (2003) ITU-T P.310: Transmission characteristics for Telephone band (300-3400 Hz) digital telephones.
- Kollmeier B (1990) Messmethodik, Modellierung und Verbesserung der Verständlichkeit von Sprache. Habilitation, Georg-August-Universität Göttingen
- 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 , 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 III: Design, Optimierung und Evaluation des Oldenburger Satztests. *ZfA* 38 (1-3), 415, 44-56, 86-95.