

MATCH - Ein neuer Kindersprachtest

Volker Schirkonyer^{1,2}; Andrea Bohnert³; Annerose Keilmann³; Thomas Janssen¹

¹ Experimentelle Audiologie, HNO- und Poliklinik, Klinikum rechts der Isar, TU München

² PATH medical GmbH, Germering

³ Universitätsmedizin Mainz, Klinik und Poliklinik für HNO und Kommunikationsstörung

Schlüsselwörter: Pädaudiologie, Sprachtest, SRT, Normwerte

Einleitung

Sprache ist das wichtigste Kommunikationsmittel, daher sind sprachaudiometrische Verfahren ein wichtiges Instrument zur Einschätzung der kommunikativen Relevanz einer Hörstörung. Beim Testen von sehr jungen Kindern gibt es einige Besonderheiten, die sich auf das Testdesign auswirken. In erster Linie ist das limitierte Vokabular zu nennen. Weiterhin sind offene Tests, im Sinne von korrektem Wiederholen der Worte, zu komplex und darüber hinaus von der Sprachproduktion des Kindes und dem Hörvermögen des Testers abhängig. Es wird auch berichtet, dass Kinder irritiert sein können, wenn kein Sprecher zu sehen ist und dass eine weibliche Sprecherstimme von Vorteil ist.

Der zurzeit bei jungen Kindern am häufigsten eingesetzte Sprachtest ist der Mainzer Kindersprachtest I (Biesalski *et al.*, 1974). Er besteht aus nur 10 Wörtern und enthält Ein-, Zweisilber sowie Lautmalereien. Zur Unterstützung bei der Durchführung sind Bildkarten verfügbar. Das begrenzte (8 Wörter, 2 Lautmalereien) und veraltete (z.B. Bahn) Vokabular und Bildmaterial sind die Hauptkritikpunkte des Tests.

Material und Methoden

Basierend auf aktuellen Untersuchungen zum Wortschatz von 21-24 Monate alten Kindern (Suchodoletz und Sachse, 2008) wurden 26 Nomen, die überwiegend gut bildlich darstellbar sind, für einen neuen Test ausgewählt. Der neue *Mainzer Audiometric Test for Children* (MATCH) wird als Picture-Pointing-Task mittels Touchscreen durchgeführt, so dass die Ergebnisse unabhängig von der Sprachproduktion des Kindes und der Hörleistung des Untersuchers sind.

Die Phonemverteilung des MATCH wurde mit aktuellen Erhebungen zur Phonemverteilung der Alltagssprache (Schiel, 2010) verglichen. Als Maß für die Übereinstimmung kam die Summe der absoluten Differenz zwischen der Häufigkeit der Test-Phoneme und den 50 häufigsten deutschen Phonemen zum Einsatz.

Die Aufnahmen wurden in einem Tonstudio mit einer professionellen Sprecherin gemacht. Das Langzeitspektrum der Aufnahmen wurde mit dem *International (Universal) Long Term Average Speech Spectrum (ILTASS)* (Byrne *et al.*, 1994) verglichen.

An insgesamt 111 Kindergarten-Kindern (213 Ohren) im Alter zwischen 2 Jahren und 11 Monaten und 6 Jahren und 9 Monaten wurde eine Normierungsstudie durchgeführt. 88 Kinder (157 Ohren) waren normalhörend, 38 (56 Ohren) schwerhörend. Da es Fälle mit einseitiger Schwerhörigkeit gab, ist die Summe aus schwerhörenden und normalhörenden Kindern größer als 111. Die Messungen fanden in Kindergärten in einem ruhigen Raum oder in der Audiologie des Schwerpunkts Kommunikationsstörungen der Universitätsklinik Mainz statt. Das Mess-Equipment bestand aus einem PATH medical Senti Desktop Messgerät und einem Sennheiser HDA-200 Kopfhörer mit schalldämmenden Peltorkapseln.

Als normalhörend wurden Kinder nach erfolgreichem Abschließen eines 30 dB HL Screenings bei den Frequenzen 500 Hz, 1 kHz, 2 kHz, 4 kHz und 6 kHz angesehen oder wenn alle Hörschwellen im dem Frequenzbereich 20 dB HL und besser waren.

Vor den sprachaudiometrischen Messungen wurde das Kind mit dem Messgerät und anhand von Bildtafeln mit dem Testmaterial vertraut gemacht. So wurde sichergestellt, dass sich die Prüf-Items im Wortschatz befinden und die Bilder richtig erkannt werden. Die Prüf-Items werden über freifeldentzerrte Kopfhörer monaural dargeboten. Auf dem Touchscreen des Messgerätes erscheinen die entsprechenden Bilder: das Prüf-Item und die drei alternativen Bild-Items. Die akustische Darbietung des Prüf-Items und die Darstellung der Bilder auf dem Touchscreen erfolgen durch Drücken auf eine „Zauberbox“ auf dem Bildschirm. Bei falschen Antworten wird der Pegel um 10 dB erhöht und bei richtigen Antworten um 4 dB verringert bis alle 26 Items abgearbeitet sind. Die hieraus abgeleitete Sprachverständnisschwelle ist der Mittelwert der letzten 20 Darbietungspegel. Wegen der

vorgenommenen Pegelfolge entspricht dies nach Kaernbach (2001) einer 71,4%-Sprachverständnisschwelle.

Zur Auswertung wurde eine Einteilung in 3 Altersgruppen vorgenommen: bis 4 Jahre 3 Monate (<4,25 Jahre), zwischen 4 Jahren 3 Monaten und 5 Jahren 6 Monaten (4,25 - 5,5 Jahre), über 5 Jahre 6 Monate (>5,5 Jahre).

Ergebnisse

Abb. 1 zeigt die relative Häufigkeit der für die Worterkennung wichtigsten deutschen Phoneme. Die schwarze Linie stellt die Phonemhäufigkeit alltäglicher Konversationen (BASat, (Schiel, 2010)) dar, die roten und grünen Balken veranschaulichen die Phonemhäufigkeiten des Mainzer Kindersprachtests I (Mainzer I) und des MATCH. Die Übereinstimmung zwischen MATCH und BASat, gemessen als Summe der absoluten Differenz der Häufigkeiten, ist wesentlich höher als beim Mainzer I (MATCH: 64,7%; Mainzer I: 113,3%). Der MATCH weist auch weniger Lücken in der Verteilung auf als der Mainzer I.

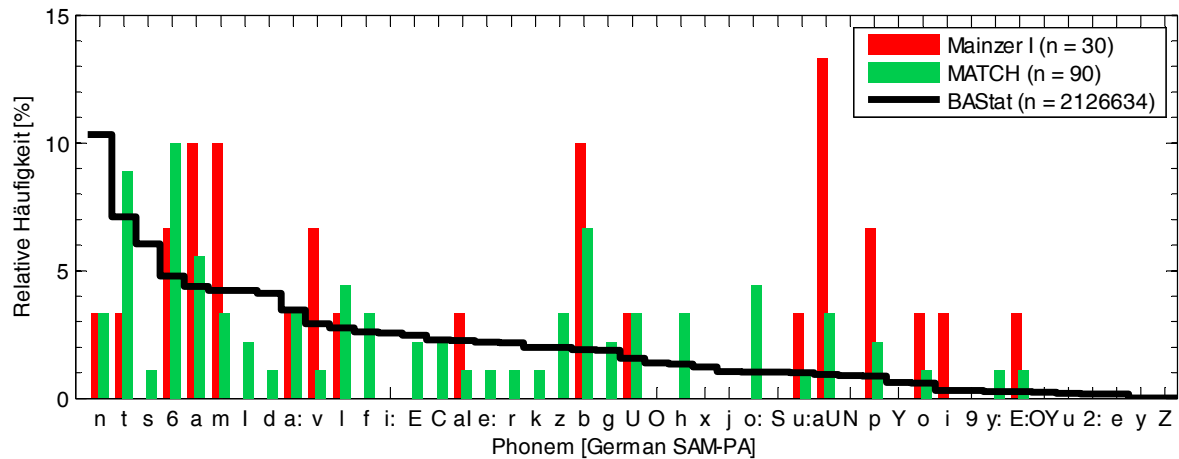


Abb. 1: Phonemstatistik der deutschen Spontansprache (BASat), des Mainzer Kindersprachtests I (Mainzer I) und des MATCH. Der MATCH zeigt eine wesentlich bessere Übereinstimmung mit der Spontansprache als der Mainzer I.

Die Langzeitspektren des Mainzer I (rot), MATCH (grün) und das mittlere Spektrum für Sprecherinnen (ILTASS W, grau) sind in Abb. 2 dargestellt. Die Signale wurden jeweils auf 70dB SPL Gesamtpegel normiert. Sowohl beim Mainzer I als auch beim MATCH kommen Frauenstimmen zum Einsatz. Der Vergleich des Langzeitspektrums der MATCH Aufnahmen mit dem ILTASS W zeigt bei nahezu allen Frequenzen zwischen 200Hz und 10kHz Abweichungen von höchstens 3dB (außer 1kHz: +4dB und 5kHz: -5.5dB). Der MATCH weist eine bessere Übereinstimmung mit dem ILTASS W auf als der Mainzer I

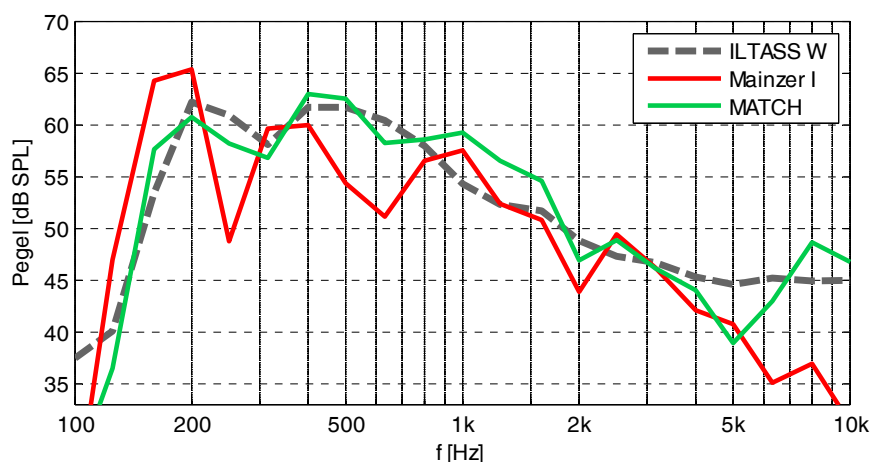


Abb. 2: Langzeitspektrum für Sprecherinnen (ILTASS W), den Mainzer Kindersprachtests I (Mainzer I) und den MATCH. In beiden Tests kommen Sprecherinnen zum Einsatz. Die Aufnahmen des MATCH zeigen eine wesentlich bessere Übereinstimmung mit dem ILTASS W als der Mainzer I.

Aus den Messdaten wurde für jede Altersgruppe im 2dB Raster zwischen 18dB und 50dB die Sprachverständlichkeit bestimmt. Diese Daten dienen zum Schätzen der altersspezifischen Diskriminationsfunktionen, die als logistische Funktion mit den Parametern Steigung (m) und Sprachverständlichkeitsschwelle für 50% Sprachverstehen (SVS) entsprechend Gleichung (1) modelliert wurden. Die Standardabweichung wurde mittels Bootstrapping mit 1000 Stichproben berechnet. Die ermittelten psychometrischen Diskriminationsfunktionen (Fit), die die Sprachverständlichkeit als Funktion des Darbietungspegels darstellen und der Normbereich ($\pm 2\sigma$) sind in Abb. 3 abgebildet. Schwarz gestrichelt ist die gemäß Einpegel-Algorithmus theoretisch gemessene Sprachverständlichkeit dargestellt. Sie beträgt 71,4 %. Die zur Ermittlung der Kurven verwendeten Werte (Quadrat, Kreis und Dreieck) haben die jeweilige Kurvenfarbe. Tabelle 1 zeigt die Eigenschaften der Diskriminationsfunktionen als Zahlenwerte. Insgesamt besteht eine gute Übereinstimmung des Modells der Diskriminationskurven mit den Messdaten. Die erzielten Normwerte liegen in der Größenordnung von bekannten Tests: Beispielsweise beträgt die Steigung der Diskriminationsfunktion beim Oldenburger Kinderreimtest ca. 5%/dB, die 50%-Schwelle beträgt ca. 19 – 22 dB für Kinder der 1. – 4. Klassenstufe (Wagner *et al.*, 2006). Die 50%-Sprachverständlichkeitsschwelle des Freiburger Einsilbertests beträgt 29,3 dB, die mittlere Steigung zwischen 40% und 60% beträgt 3,6 % / dB (DIN 45626-1, 1995).

$$f(L) = \frac{1}{1 + e^{4 \cdot m \cdot (SVS - L)}} \quad (1)$$

- L: Darbietungs-Schalldruckpegel [dB]
 m: Steigung im Wendepunkt (50% Sprachverständlichkeit) [%/dB]
 SVS: 50%-Sprachverständlichkeitsschwelle [dB]

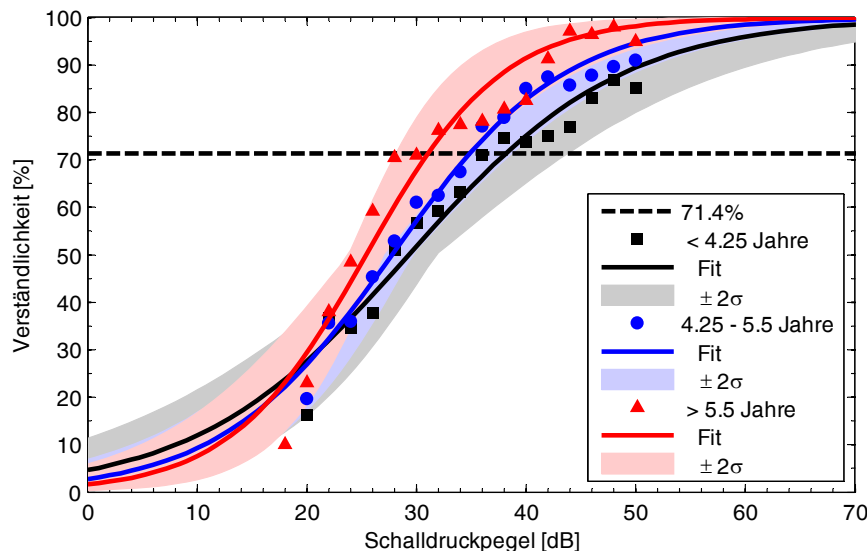


Abb. 3: Diskriminationsfunktionen mit Standardabweichung für alle Altersgruppen. Die gefitteten logistischen Funktionen zeigen eine gute Übereinstimmung mit den experimentellen Daten

Gruppe	unter 4,25 Jahre	4,25 - 5,5 Jahre	über 5,5 Jahre
Steigung bei 50% [% / dB]	2,7 ± 0,4	3,3 ± 0,4	4,1 ± 0,6
50% Schwelle [dB]	29,4 ± 1,2	27,8 ± 0,9	25,4 ± 0,8
Steigung bei 71,4% [% / dB]	2,2 ± 0,3	2,7 ± 0,3	3,4 ± 0,4
71,4% Schwelle [dB]	38,2 ± 1,0	34,9 ± 0,7	31,1 ± 1,0

Tabelle 1: Kennzahlen der gefitteten Diskriminationsfunktionen in Abhängigkeit des Alters.

Abb. 4 zeigt die Test-Retest Reliabilität als Streudiagramm der 2. Messung in Abhängigkeit der 1. Messung für Kinder aller getesteter Altersgruppen. Die 2. Messung fand im Allgemeinen 24 ± 1 h nach der 1. Messung statt. Die graue Linie ist die Winkelhalbierende auf der im Idealfall alle Punkte liegen, die Ausgleichsgerade ist blau dargestellt. Test und Re-Test stimmen gut überein bei einer Korrelation von 0,89.

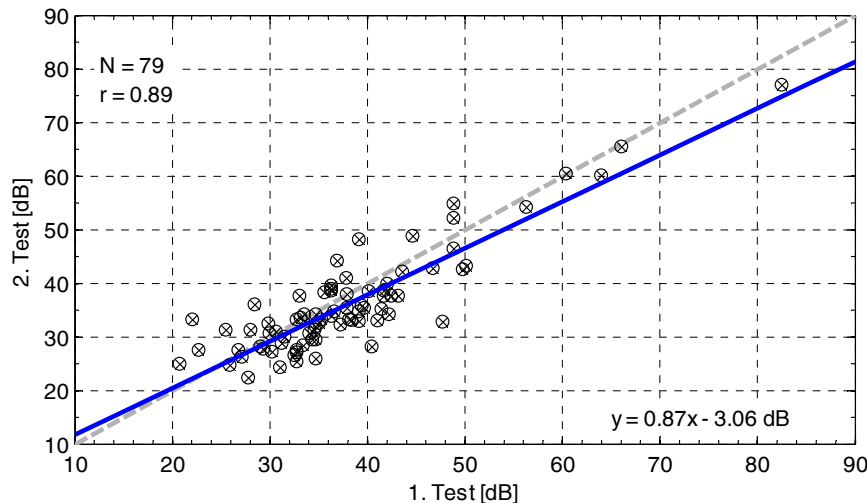


Abb. 4: Test-Retest Reliabilität. Gute Übereinstimmung von 1. Test und 2. Test.

Eine möglichst kurze Testdauer ist bei der Untersuchung von Kindern besonders wichtig, weil deren Aufmerksamkeitsspanne wesentlich kürzer ist als diejenige von Erwachsenen. Beim MATCH beträgt die mittlere Messdauer zwischen 3 und 4 Minuten pro Ohr, je nach Altersgruppe, wobei die Unterschiede zwischen den Gruppen nicht statistisch signifikant sind. Die Messdauern sind in Tabelle 2 aufgeführt.

Gruppe	unter 4,25 Jahre	4,25 – 5,5 Jahre	über 5,5 Jahre
Minimum - Maximum	1,6 - 6,9	1,7 - 6,5	1,6 - 5,9
Median (95% Konfidenzintervall)	4,0 (3,6 - 4,4)	3,4 (3,0 - 3,7)	3,1 (2,7 - 3,8)
Unteres - Oberes Quartil	3,1 - 4,8	2,5 - 4,3	2,4 - 4,0

Tabelle 2: Testdauern aller Kinder in Abhängigkeit des Alters in Minuten.

Zusammenfassung und Schlussfolgerungen

Ein neuer Kindersprachtests mit neuem Bild- und Wortmaterial, das auf aktuellen Untersuchungen zum Wortschatz von ca. 2 jährigen Kindern beruht, wurde erstellt. Das Langzeitspektrum und die Phonemverteilung des Sprachmaterials zeigen eine wesentlich bessere Übereinstimmung mit der dem durchschnittlichen Langzeit-Sprachspektrum bzw. der Phonemstatistik alltäglicher Sprache als der alte Mainzer I. Dadurch, dass der Test wie ein Spiel als Selbsttest durchführbar ist, erhöht sich die Konzentration und Motivation der Kinder. Die erzielten Normwerte liegen in der Größenordnung von bekannten Tests. Der Test weist eine hohe Test-Retest Reliabilität auf.

Literatur

- Biesalski, P., H. Leitner, E. Leitner, et al. (1974). Der Mainzer Kindersprachtest. Sprachaudiometrie im Vorschulalter. HNO 22(5): 160-161.
- Byrne, D., H. Dillon, K. Tran, et al. (1994). An international comparison of long-term average speech spectra. The Journal of the Acoustical Society of America 96(4): 2108-2120.
- DIN 45626-1 (1995). Tonträger mit Sprache für Gehörprüfung. Teil 1: Tonträger mit Wörtern nach DIN 45621-1 (Aufnahme 1969). Deutsches Institut für Normung e.V.
- Kaernbach, C. (2001). Adaptive threshold estimation with unforced-choice tasks. Percept Psychophys 63(8): 1377-1388.
- Schiel, F. (2010). BASat : New Statistical Resources at the Bavarian Archive for Speech Signals. Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta, European Language Resources Association (ELRA).
- Suchodoletz, W. v. & S. Sachse (2008). SBE-2-KT: Sprachbeurteilung durch Eltern - Kurztest für die U7, <http://www.kjp.med.uni-muenchen.de/sprachstoerungen/SBE-2-KT.php>.
- Wagener, K. C., T. Brand & B. Kollmeier (2006). Evaluation des Oldenburger Kinder-Reimtests in Ruhe und im Störgeräusch. HNO 54(3): 171-178.