

Identifizierung von Sprechermerkmalen durch Zuhörergruppen unterschiedlicher Altersstufen und Expertise

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Hohen Medizinischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität

Bonn

Johanna Maria Thiele

aus Bonn

2020

Angefertigt mit der Genehmigung
der Medizinischen Fakultät der Universität Bonn

1. Gutachter: Prof. Dr. med. Götz Schade
2. Gutachter: Prof. Dr. med. Rainer Ganschow

Tag der Mündlichen Prüfung: 28.10.2020

Aus der Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde
Direktor: Prof. Dr. med. Sebastian Strieth

Inhaltsverzeichnis

Abkürzungsverzeichnis	6
1. Einleitung	7
1.1 Physiologische Grundlagen der Stimmbildung	7
1.1.1 Kehlkopf und Phonation	7
1.1.2 Ansatzräume und Artikulation.....	9
1.2 Physiologische und akustische Veränderungen der Stimme.....	9
1.2.1 Stimmveränderungen in Abhängigkeit vom Alter.....	9
1.2.2 Stimmunterschiede in Abhängigkeit vom Geschlecht.....	11
1.2.3 Stimmunterschiede in Abhängigkeit von Körpergröße/-gewicht	13
1.3 Wissensstand und Fragestellung	14
2. Material und Methoden	18
2.1 Studiendesign.....	18
2.2 Erstellung des Materials	18
2.2.1 Erstellung der Fragebögen	18
2.2.2 Erstellung der Stimmaufnahmen	20
2.3 Auswahl der Studienteilnehmer	21
2.3.1 Auswahl der vorlesenden Personen.....	21
2.3.2 Auswahl der Probanden	23
2.4 Ablauf der Befragungen.....	25
2.5 Auswertung der Fragebögen	26
2.6 Statistische Auswertung der Ergebnisse	27

3.	Ergebnisse	29
3.1	Umgang mit fehlenden und ungültigen Werten	29
3.2	Allgemeine Daten der Probanden	29
3.3	Gesamtergebnis der Probanden	30
3.4	Ergebnisse nach Variablen.....	32
3.4.1	Alter	32
3.4.2	Körpergröße	37
3.4.3	BMI	40
3.4.4	Geschlecht	42
3.4.5	Akzent/Dialekt/Hochdeutsch.....	44
3.4.6	Rauchverhalten	48
3.4.7	Schulabschluss	50
3.4.8	Berufsausbildung.....	52
3.4.9	Beruf.....	54
3.4.10	Persönlichkeit	56
3.4.11	Äußere Erscheinung.....	57
3.4.12	Höreindruck der Stimme.....	58
4.	Diskussion	60
4.1	Erörterung der Ergebnisse in den einzelnen Variablen	60
4.1.1	Alter	60
4.1.2	Körperbau.....	62
4.1.3	Geschlecht	65
4.1.4	Akzent/Dialekt/Hochdeutsch.....	67
4.1.5	Rauchverhalten	70
4.1.6	Bildungsstand.....	72

4.1.7	Persönlichkeit	74
4.1.8	Äußere Erscheinung.....	74
4.1.9	Höreindruck der Stimme.....	75
4.2	Erörterung der Ergebnisse im Hinblick auf Alter und Expertise	76
4.3	Evaluation des Studiendesigns	77
4.4	Einordnung im Hinblick auf die Fragestellung	78
5.	Zusammenfassung.....	80
6.	Anhang	82
7.	Abbildungsverzeichnis	89
8.	Tabellenverzeichnis	91
9.	Literaturverzeichnis	93

Abkürzungsverzeichnis

Abb.	Abbildung
ANOVA	Varianzanalyse (Analysis of Variance)
BMI	Body-Mass-Index
cm	Zentimeter
Grs.	Grundschüler
Gymn.	Gymnasiasten
Hz	Hertz
kg	Kilogramm
m	Meter
M	Mittelwert
max.	maximal
min	Minute
mm	Millimeter
ms	Millisekunde
Nr.	Nummer
Phon.	Phoniater
s	Sekunde
s.	siehe
SD	Standardabweichung (Standard Deviation)
Stud.	Studierende
Tab.	Tabelle
WHO	World Health Organization

1. Einleitung

1.1 Physiologische Grundlagen der Stimmbildung

1.1.1 Kehlkopf und Phonation

Als Phonation wird die Stimmerzeugung im Kehlkopf bezeichnet. Voraussetzung für die Phonation sind Atmungssystem und Atmung, da durch den kontinuierlichen Luftstrom der Lunge die Stimmlippenschwingungen ermöglicht werden (Mead et al., 1968). Abb. 1 zeigt schematisch den Aufbau von Kehlkopf und Kehlkopfmuskulatur. Der Kehlkopf liegt oberhalb der Luftröhre und besteht aus einem Gerüst beweglich miteinander verbundener Knorpel. Zwischen Schildknorpel (*Cartilago thyroidea*) und Aryknorpeln (*Cartilagine arytaenoideae*) sind die Stimmbänder (*Ligamenta vocalia*) gespannt, die zwischen sich die Stimmritze (*Glottis*) einfassen (Geiger, 2019). Als Stimmlippen werden Stimmband, *Musculus vocalis* sowie das Schleimhautepithel bezeichnet. Die Stimmlippen lassen sich in Drittel einteilen, wobei die vorderen beiden Drittel dem ligamentären Anteil entsprechen und das hintere Drittel dem kartilaginären Stimmlippenanteil (Schade et al., 2004a). Bei der Atmung ist die Glottis durch Zug des *Musculus cricoarytaenoideus posterior* weit geöffnet. Während der Expiration bei der Stimmerzeugung hingegen nähern sich die Stimmbänder durch Kontraktion der *Musculi cricoarytaenoidei laterales*, *Musculi thyroarytaenoidei* und *Musculi arytaenoidei an* (Kreiman und Sidtis, 2011); zudem werden sie von den *Musculi vocales* und den *Musculi cricothyroidei* (nicht abgebildet) gespannt (Siebel, 2018). Durch die forcierte Ausatmung wird als nächstes ein Luftdruck unterhalb der Glottis erzeugt, der schließlich zu einer leichten Öffnung der Stimmlippen führt (Kreiman und Sidtis, 2011). Aufgrund mehrerer Faktoren werden anschließend die Stimmbänder wieder zusammengezogen: Der subglottische Druck sinkt aufgrund des beschriebenen Entweichens der Luft durch die Glottis wieder ab und die elastischen Stimmbänder nehmen ihre ursprüngliche Anordnung nahe der Mittellinie wieder ein (van den Berg, 1958; Kreiman und Sidtis, 2011). Gemäß des Bernoulli-Effekts bedingt die erhöhte Strömungsgeschwindigkeit der Luft einen Druckabfall zwischen den Stimmbändern, sodass diese sich wieder annähern müssen (Jiang et al., 2000; Siebel, 2018). Aus dem schnell ablaufenden Wechsel zwischen Glottisöffnung und -schluss re-

sultieren die periodischen Schwingungen der Stimmlippen bzw. der Stimmschall (Schulte und Seidner, 2015).

V.a. die unterschiedlichen Längen der Stimmbänder, welche nach Schade et al. (2004b) mittels Laservermessung präzise bestimmt werden können, führen zu verschiedenen Grundfrequenzen, bspw. von Männer- und Frauenstimmen; aber auch Änderungen der Stimmlippenspannung und des subglottalen Druckes beeinflussen diese (Jiang et al., 2000; Titze, 1989). Die sogenannte Grundfrequenz ist ein Parameter bezogen auf die Schwingungsrate der Stimmlippen und korreliert mit der wahrgenommenen Tonhöhe der Stimme eines Sprechers (aus Gründen der besseren Lesbarkeit wird im Folgenden durchgehend die grammatisch männliche Form geschlechtsneutral verwendet. Die entsprechenden Begriffe beziehen sich in gleicher Art und Weise auf beide Geschlechter.) (Baker et al., 2008; Kreiman und Sittis, 2011).

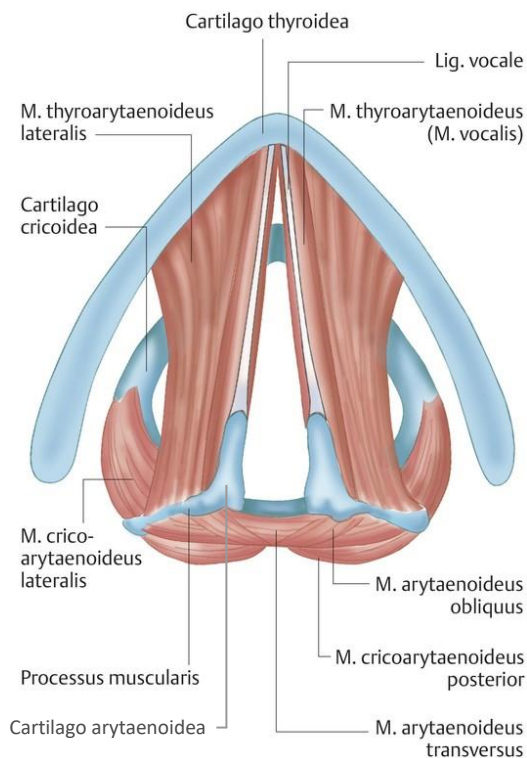


Abb. 1: Darstellung von Kehlkopf und inneren Kehlkopfmuskeln (nach Schüssler et al., 2017)

1.1.2 Ansatzräume und Artikulation

Der von den Stimmlippen erzeugte Klang wird bei der Artikulation in den Ansatzräumen, die aus Rachen-, Mund- und Nasenraum bestehen, ausgestaltet. Die Ansatzräume fungieren als Resonanzraum und es werden je nach ihrer Form, welche durch Bewegungen bspw. von Zunge, Lippen und Rachenmuskulatur veränderbar ist, verschiedene Resonanzfrequenzen erzeugt, die sogenannten Formanten (Kreiman und Sidtis, 2011). So entstehen, u.a. gekennzeichnet durch unterschiedliche Formantfrequenzen, die Sprachlaute. Vokale werden bei weiten Ansatzräumen und frei entweichendem Luftstrom gebildet, während sich die Ansatzräume bei der Produktion von Konsonanten verengen und den Luftstrom begrenzen (Geiger, 2019).

1.2 Physiologische und akustische Veränderungen der Stimme

Im Rahmen der vorliegenden Arbeit soll u.a. ergründet werden, ob Zuhörer physische Merkmale von Sprechern anhand des alleinigen Hörens ihrer Stimmen identifizieren können. Daher wird zunächst überblicksartig dargestellt, wie sich die menschliche Stimme abhängig von Alter, Geschlecht und Körperbau eines Sprechers verändert bzw. unterscheidet, und die akustischen Korrelate werden beschrieben. Die wichtigsten Aspekte werden aufgeführt, es besteht kein Anspruch auf Vollständigkeit.

1.2.1 Stimmveränderungen in Abhängigkeit vom Alter

Beim Neugeborenen ist das Atmungssystem noch nicht ausgereift und entwickelt sich in den ersten Lebensjahren in vielen Aspekten weiter; beispielhaft zu nennen sind Veränderungen hinsichtlich Struktur der Lunge (z.B. Zunahme der Lungengröße sowie von Zahl und Größe der Alveolen) und Atemmechanik (z.B. Zunahme der elastischen Rückstellkräfte der Lunge) (Boliek et al., 1996). Der Kehlkopf des Neugeborenen liegt im Vergleich zu dem des Erwachsenen höher im Hals (auf Höhe des 1.-3. Halswirbels, beim Erwachsenen auf Höhe des 6.-7. Halswirbels, Prakash und Johnny, 2015), weshalb der Resonanzraum in Form der supraglottischen Ansatzräume kleiner ist. Die Stimmlippenlänge beträgt nach Kurita (1988) lediglich 2,5-3 mm, woraus eine hohe Grundfrequenz von ca. 500 Hz resultiert (Kreiman und Sidtis, 2011). Die Gewebeschichten der Stimmlippen sind noch nicht differenziert und das Stimmband ist noch nicht ent-

wickelt (Hirano et al., 1981; Sapienza et al., 2004). Neben der Unreife der Strukturen von Atmungssystem und Kehlkopf sind die Mundbewegungen des Neugeborenen eher auf das Saugen ausgelegt, was eine Produktion von Sprachlauten weiter einschränkt.

Im Verlauf des Kindesalters ist die Lungenentwicklung erst mit dem 8.-10. Lebensjahr abgeschlossen (Schünke et al., 2018). Der Kehlkopf senkt sich zunehmend ab, Vokaltrakt und Stimmlippen wachsen mit dem restlichen Körper mit und beginnen sich zu differenzieren (Fried et al., 1982): So beginnt sich das Stimmband zwischen dem 1. und 4. Lebensjahr auszubilden, welches erst vor dem 20. Lebensjahr ausgereift ist (Hirano et al., 1981). Aufgrund der zunehmenden Größe des Kehlkopfes nimmt die Grundfrequenz mit steigendem Alter ab, etwa beim 8-jährigen Kind auf ca. 275 Hz (Kreiman und Sidtis, 2011). Selbiges gilt für die Formantfrequenzen, die ebenfalls bei Kindern höher sind als bei Erwachsenen (Vorperian und Kent, 2007); zu den unterschiedlichen Formantfrequenzen bei Mädchen und Jungen s. 1.2.2. Kinder müssen für die Erzeugung einer lauten Stimme mehr Atemarbeit anwenden, da sie im Vergleich zu Erwachsenen beim lauten Sprechen einen höheren subglottalen Druck erzeugen. Aus letzterem resultiert, wie in 1.1.1 dargelegt, eine höhere Grundfrequenz. Somit ist es für Sprecher im Kindesalter schwieriger, Tonhöhe und Lautstärke ihrer Stimme zu kontrollieren (Kreiman und Sidtis, 2011).

Bei Erwachsenen liegt die Grundfrequenz niedriger als bei den kindlichen Sprechern und unterscheidet sich deutlich zwischen Männern und Frauen; für die Werte der Grund- und auch Formantfrequenzen s. 1.2.2. Die Sprechgeschwindigkeit, die im Verlauf des Kindesalters zunimmt, ist schnell. Mit zunehmendem Alter wird die Atmung ineffektiver, weil der Brustkorb steifer wird und somit die Vitalkapazität der Lunge abnimmt (Sharma und Goodwin, 2006). Es sind mehr Atemzüge notwendig und die Sprechgeschwindigkeit sinkt. Im Rahmen der Muskeldegeneration kann es, v.a. bei Männern, zu einer Atrophie der inneren Kehlkopfmuskeln kommen, bspw. des *Musculus vocalis* (Sato und Tauchi, 1982). Dies führt insbesondere bei älteren Männern zu einer Abnahme der Masse der Stimmlippen und zu einer höheren Grundfrequenz. V.a. bei älteren Frauen nach der Menopause konnte in einigen Fällen ein Ödem der Stimmlippen beobachtet werden. Es kann eine Verdickung der Stimmlippen resultieren, was u.a. eine Erklärung für die niedrigeren Grundfrequenzen bei älteren Frauen im Vergleich zu jüngeren Frauen sein könnte (Hirano et al., 1989; Honjo und Isshiki, 1980). Die Stimmen von älteren Sprechern

werden in einigen Fällen u.a. als heiser oder schwächer bis atemlos beschrieben (Honjo und Isshiki, 1980; Kahane, 1987). Da der Kehlkopf mit zunehmendem Alter weiter absinkt, gibt es die Annahme, dass aufgrund der Verlängerung der supraglottischen Ansatzräume die Formantfrequenzen abnehmen (u.a. Linville, 2001). Allerdings scheinen die Veränderungen der Formantfrequenzen für die verschiedenen Formanten variabel und geschlechtsabhängig zu sein; so schlagen Eichhorn et al. (2018) vor, dass die festgestellten Veränderungen für den ersten und zweiten Formanten eher auf Effekte der Artikulation zurückzuführen sind als auf die Verlängerung des Vokaltraktes. Mit steigendem Alter gehen noch weitere Entwicklungen, mit variablen, teilweise nicht gänzlich geklärten Effekten auf die Stimme, einher, zu nennen u.a. degenerative histologische Veränderungen der Stimmlippen sowie die zunehmende Verknöcherung der Kehlkopfknorpel (Hatley et al., 1965). Generell kann sich im Alter im Rahmen neurodegenerativer Prozesse die Stimmstabilität verringern, was sich bspw. in einem Stimmzittern äußern kann (Martins et al., 2014).

Nach Shipp et al. (1992) scheinen Zuhörer die Sprecher v.a. anhand der unterschiedlichen Grundfrequenzen sowie Sprechgeschwindigkeiten mit Anzahl und Länge der Atempausen in verschiedene Altersgruppen (junge, mittelalte und ältere Erwachsene) einzuordnen.

1.2.2 Stimmunterschiede in Abhängigkeit vom Geschlecht

Bis zum Beginn der Pubertät verläuft die Entwicklung des Stimmapparates bei Jungen und Mädchen sehr ähnlich (s. 1.2.1). Die Gesamtlänge des Vokaltraktes ist geschlechtsabhängig nicht unterschiedlich (Vorperian et al., 2005) und die Grundfrequenzen von Jungen und Mädchen unterscheiden sich nicht signifikant voneinander (Sorenson, 1989). Allerdings wurde gezeigt, dass die Formantfrequenzen im Alter von 4 Jahren zwischen den Geschlechtern differieren (Perry et al., 2001) und bei Jungen im Alter von 7 bis 8 Jahren durchweg niedriger sind als bei gleichaltrigen Mädchen (Bennett, 1981). Die physiologischen Korrelate dieser Differenz sind nicht gänzlich geklärt; nach Bennett (1981) könnten auch Unterschiede in der Artikulation von Jungen und Mädchen, wodurch die Form der Ansatzräume veränderbar ist (s. 1.1.2), eine Rolle spielen.

Ab der Pubertät wachsen bei Mädchen Stimmtrakt und Kehlkopf proportional mit dem restlichen Körper mit (Kreiman und Sidtis, 2011). Bei Jungen ist die Größenzunahme von Rachen und Kehlkopf überproportional ausgeprägt; so beträgt etwa die Stimmlippenlänge nach Kurita (1988) beim erwachsenen Mann 15-21 mm, im Vergleich zu 10-15 mm bei der Frau. Hierbei handelt es sich um Werte von japanischen Männern und Frauen, je nach Ethnie der untersuchten Personen können die festgestellten Stimmlippenlängen abweichen. Kahane (1982) belegt, dass bei den untersuchten männlichen Kaukasiern die Stimmlippenlänge vom Kindes- zum Erwachsenenalter mehr als doppelt so stark zunimmt wie bei den weiblichen Kaukasiern. Der Stimmwechsel von der kindlichen Stimme zu der eines Erwachsenen macht sich v.a. bei Jungen bemerkbar: Neben dem stärkeren Abfall der mittleren Sprechstimmlage und ggfs. dem Stimmbruch mit raschen Wechseln zwischen hohem und tiefem Register bildet sich morphologisch durch das starke Wachstum der Stimmlippen und die Verkleinerung des Winkels des Schildknorpels (von ca. 120° auf ca. 90°) die Prominentia laryngea („Adamsapfel“) aus (Schutte und Seidner, 2015). Schließlich liegt die durchschnittliche Grundfrequenz des erwachsenen Mannes mit ca. 130 Hz knapp eine Oktave tiefer als die Grundfrequenz der Frau mit ca. 220 Hz (Hillenbrand und Clark, 2009). Die Länge des männlichen Vokaltraktes ist im Erwachsenenalter, anders als vor der Pubertät, signifikant höher als die des weiblichen Vokaltraktes (der männliche Vokaltrakt ist ca. 15 % länger als derjenige der Frauen, Fitch und Giedd, 1999; Klatt und Klatt, 1990). Daraus resultieren niedrigere Formantfrequenzen bei Männern im Vergleich zu Frauen, beschrieben werden um ca. 20 % niedrigere Formantfrequenzen von Männerstimmen (Kreiman und Sidtis, 2011). Ein weiterer Unterschied zwischen Männer- und Frauenstimmen besteht in Hinsicht auf das Schwingungsmuster ihrer Stimmlippen. U.a. treten bei Frauen häufiger unvollständige Glottisschlüsse auf, wodurch ihre Stimmen aufgrund der durch den entstehenden Spalt entweichenden Luft behauchter oder weicher klingen können (Biever und Bless, 1989; Södersten et al., 1995). Weiterhin stellten u.a. Haan und van Heuven (1999) fest, dass Frauen mit einer höheren Tonhöhenvariabilität sprechen als Männer, beim Vorlesen von sowohl Fragen als auch Aussagen; dies konnte durch Daly und Warren (2001) bestätigt werden. Es gibt Hinweise darauf, dass Frauen eine deutlichere Aussprache haben als Männer, welche häufig Vokale und Konsonanten reduzieren oder weglassen (Whiteside,

1996), sodass Frauen von Zuhörern leichter zu verstehen sein können (Hazan und Markham, 2004).

Zusammenfassend stellen nach Bachorowski und Owren (1999) Grund- und Formantfrequenzen die wichtigsten akustischen Korrelate zur Identifizierung des Geschlechts eines Sprechers dar.

1.2.3 Stimmunterschiede in Abhängigkeit von Körpergröße/-gewicht

Wie in 1.2.2 erläutert, nimmt mit dem Größenwachstum von Kindern die Grundfrequenz ihrer Stimme ab; bei Jungen stärker als bei Mädchen. Somit gibt es einen Zusammenhang zwischen Körpergröße/-gewicht und Grundfrequenz bzw. Tonhöhe beim Vergleich der Altersgruppen und Geschlechter: Die Grundfrequenz ist bei Kindern höher als bei Erwachsenen sowie bei Männern niedriger als bei Frauen (Pisanski und Rendall, 2011). Allerdings scheint *innerhalb* der Geschlechter- und Altersgruppen keine signifikante Korrelation vorzuliegen; dies belegten u.a. Smith et al. (2007) sowie Künzel (1989), welcher die Grundfrequenzen von Männer- und Frauenstimmen mit ihren jeweiligen Größen- und Gewichtsmaßen in Beziehung setzte. Dies lässt sich möglicherweise durch die Topografie des menschlichen Kehlkopfes begründen: Der Kehlkopf ist am Zungenbein beweglich aufgehängt und ansonsten in seinem Wachstum nicht eingeschränkt (Aumüller und Wennemuth, 2020). Dementsprechend nimmt er, wie in 1.2.2 ausgeführt, beim Jungen in der Pubertät auch unabhängig vom Wachstum des restlichen Körpers stark an Größe zu. Des Weiteren wurde festgestellt, dass Sprecher ihre Grundfrequenz respektive Tonhöhe beim Sprechen im Bereich einer Oktave variieren (Hudson und Holbrook, 1982).

Im Unterschied dazu wird die Länge des Vokaltraktes von Säugetieren, auch des Menschen, durch den knöchernen Schädel begrenzt und ist somit stärker vom Wachstum des restlichen Körpers abhängig. Nach Pisanski et al. (2014) korrelieren die Formantfrequenzen besser mit Körpergröße/-gewicht von Sprechern als die Grundfrequenzen. Jedoch sind die Ergebnisse der verschiedenen Studien heterogen und zeigen auch schwache oder fehlende Korrelationen; González (2004) etwa dokumentiert einen lediglich schwachen Zusammenhang zwischen Formantfrequenz und Körpergröße/-gewicht innerhalb der Gruppen von Frauen und Männern. Es ist hierbei zu berücksichtigen, dass

Sprecher die Form ihrer Ansatzräume durch Bewegungen der beteiligten anatomischen Strukturen verändern können (s. 1.1.2).

1.3 Wissensstand und Fragestellung

Häufig begegnet uns im Alltag die Situation, dass wir eine uns unbekannte Stimme hören, ohne den zugehörigen Sprecher zu sehen oder etwas über diesen zu wissen. Sei es bspw. am Telefon oder im Radio - unweigerlich machen wir uns eine Vorstellung über die Person, die zu dieser Stimme gehört. Dies kann physiologische Parameter wie Geschlecht, Alter und Körperbau betreffen, aber auch die geografische Herkunft, Bildungsgrad, Persönlichkeit usw. Relevanz erhält diese Thematik im Bereich der Forensik. In der Kriminaltechnik ist die forensische Sprechererkennung für die Aufklärung oder Verhinderung von Straftaten bedeutsam. Im Rahmen der Sprechererkennung werden vom Bundeskriminalamt Stimmenanalysen und Stimmenvergleiche durchgeführt, wofür die Stimmen auditiv sowie akustisch (z.B. durch Messung der Grund- und Formantfrequenzen) analysiert werden (Jessen, 2007). Eine Stimmenanalyse wird durchgeführt, wenn bspw. ein anonymes Anruf eines Erpressers vorliegt, aber kein Tatverdächtiger vorhanden ist; es wird ein Stimmenprofil erstellt, welches wichtige Informationen etwa über Geschlecht, Alter, regionale und soziale Herkunft des Sprechers liefern kann. Beim Vorliegen eines oder mehrerer Tatverdächtiger erfolgt ein Stimmenvergleich der Stimmproben des unbekanntes Sprechers und des Verdächtigen (Jessen, 2008). In Fällen, in denen keine Audio-Aufzeichnung vorliegt, aber ein Zeuge (oder Opfer) die Stimme des Straftäters gehört hat, kann dieser ggfs. entscheidend zur Aufklärung beitragen: Ist dem Zeugen die Stimme unbekannt und es gibt einen Verdächtigen, können ersterem im Rahmen einer auditiven Gegenüberstellung Stimmproben vorgespielt werden; hierbei handelt es sich um die Stimmprobe des Verdächtigen sowie Vergleichsstimmproben mit ähnlichen Stimm-/ Sprachcharakteristika (Broeders und van Amelsvoort, 1999). Möglicherweise gelingt dann mithilfe des Zeugen die Identifikation des Täters. In dem Fall, dass dem Zeugen die Stimme des Straftäters bspw. aus seinem Umfeld bekannt ist, kann er eine reguläre Zeugenaussage tätigen (Jessen, 2008).

Im Rahmen dieser Arbeit soll untersucht werden, inwieweit Zuhörer bestimmte Merkmale ihnen unbekannter Sprecher anhand des Hörens ihrer Stimmufnahmen identifizieren können. Im Folgenden soll ein erster Überblick über bestehende Forschungsliteratur hinsichtlich der zu untersuchenden Merkmale gegeben werden.

Die Wahrnehmung physischer Merkmale anhand der Stimme wurde in einigen Studien überprüft. In Bezug auf das *Alter* von Sprechern wurde zum einen getestet, wie gut Zuhörer dieses kategorisieren können, etwa wie präzise sie die Sprecher vorgegebenen Altersgruppen zuordnen können; u.a. sollten die Zuhörer ihnen unbekannte Sprecher als junge, mittelalte oder ältere Erwachsene einordnen (Linville und Fisher, 1985) sowie Altersdekaden im Bereich von 20-90 Jahren zuordnen (Shipp und Hollien, 1969). Zum anderen hatten die Zuhörer die Aufgabe, direkte Schätzungen des Sprecheralters vorzunehmen (u.a. Braun, 1996; Huntley et al., 1987; Ryan und Capadano, 1978). Hinsichtlich der *Körpergröße/-gewicht* von Sprechern wurde ebenfalls in der Mehrzahl der Studien untersucht, wie präzise die direkten Schätzungswerte der Zuhörer ausfallen, mit insgesamt inkonsistenten und eher schwach ausfallenden Ergebnissen (Gonzalez, 2003; van Dommelen und Moxness, 1995). Rendall et al. (2007) überprüften hingegen, wie gut Zuhörer die Sprechergrößen im Vergleich jeweils zweier Sprecher einordnen konnten. Im Unterschied zu den Ergebnissen bspw. hinsichtlich des Körperbaus gelingt den Zuhörern die Beurteilung des *Geschlechts* anhand der Stimme eines Sprechers offenbar besser (u.a. Lass et al., 1979; Whiteside, 1998). Bennett und Montero-Diaz (1982) testeten, inwieweit Kinder im Alter von 6-9 Jahren das Geschlecht von Sprechern identifizieren konnten.

Bezüglich der Identifizierung von *Akzent/Dialekt* (oder Hochdeutsch) der Sprecher wurde in jüngerer Zeit das Forschungsgebiet der perzeptuellen Dialektologie geprägt, welches die Sichtweisen von linguistischen Laien auf regionale Besonderheiten von Sprache beleuchtet (Preston, 2017; Hundt et al., 2017). In entsprechenden Studien sollten die Zuhörer die Sprecher meist einer vorgegebenen Auswahl von Regionen zuordnen, u.a. van Bezooijen und Gooskens (1999); als Beispiel für eine Untersuchung im deutschsprachigen Raum s. Hundt et al. (2017). Bei Clopper und Pisoni (2007) wurden keine Dialektkategorien vorgegeben, sondern die Zuhörer konnten die Dialekte frei klassifizieren, d.h. beliebig viele bzw. große Sprechergruppen bilden. In Studien zum Akzent wurde überprüft, ob Zuhörer zwischen muttersprachlichen Sprechern und solchen mit

ausländischem Akzent unterscheiden können und ob sie die Stärke des etwaigen Akzents korrekt einzuordnen vermögen (Cunningham-Andersson und Engstrand, 1989). Girard et al. (2008) testeten die Fähigkeiten von französischsprachigen Kindern, zwischen dem Dialekt ihrer Heimatregion und jeweils einem anderen französischen Dialekt bzw. einem ausländischen Akzent zu differenzieren.

Der Zusammenhang von *Rauchverhalten* und Stimmwahrnehmung wurde bislang wenig ergründet; bekannt ist, dass die Stimmen von Rauchern niedrigere Grundfrequenzen aufweisen als von Nichtrauchern, was sich durch ein Ödem und eine erhöhte Masse der Stimmlippen begründen lässt (Murphy und Doyle, 1987; Wallner, 1954). Braun und Rietveld (1995) untersuchten den Einfluss des Rauchens auf das geschätzte Alter von Sprechern.

Ob Zuhörer anhand von Stimm-/Sprachmerkmalen eines Sprechers Rückschlüsse auf dessen *Bildungsstand* (Schulabschluss, Berufsausbildung, Beruf) ziehen können, wurde bisher kaum untersucht. Jessen (2007) beschreibt als eines der Sprechermerkmale, welche in der forensischen Sprachanalyse bestimmt werden, den Soziolekt; nach Jessen umfasst dieser u.a. den Bildungsgrad sowie den Beruf eines Sprechers. Bspw. kann aufgrund der Verwendung gehobener Sprache oder einer komplexen Syntax möglicherweise auf einen Sprecher mit höherem Bildungsgrad geschlossen werden. Allerdings werden in der vorliegenden Arbeit nicht frei formulierte Äußerungen der Sprecher beurteilt, sondern ein vorgelesener phonetischer Mustertext. Rammstedt (2013) ermittelte die Lesekompetenzen von Personen mit unterschiedlichem Bildungsabschluss.

Schließlich werden mit *Persönlichkeit*, *Erscheinungsbild* und *Höreindruck* der Sprecher drei Parameter untersucht, deren Bewertung in besonderem Maße subjektiv gefärbt und daher schwierig ‚messbar‘ ist. Es ist naheliegend, dass dies die Durchführung von Studien in diesem Bereich erschwert. Hunt und Lin (1967) überprüften die Fähigkeit von Zuhörern, die Persönlichkeit von Sprechern anhand ihrer Stimmaufnahmen zu beurteilen; als Kriterium für die Genauigkeit der Einschätzungen der Zuhörer dienten Selbsteinschätzungen der Sprecher über ihre Persönlichkeit. In neueren Ansätzen wurde zur Persönlichkeitsmessung der Sprecher das 5-Faktoren-Modell, welches die 5 Hauptdimensionen der Persönlichkeit inkludiert, verwendet (Imhof, 2010). Hinsichtlich der äußeren Erscheinung von Sprechern lagen den Zuhörern in den wenigen vorhandenen Studien Fotografien der Sprecher vor, welche sie bspw. den Stimmaufnahmen der Sprecher

zuordnen sollten (McGlone und Hicks, 1979); dies stellt einen Unterschied zu der Aufgabe in vorliegender Arbeit dar, bei der die Zuhörer lediglich anhand der Stimmaufnahmen das unbekannte Erscheinungsbild der Sprecher abschätzen sollten. In Bezug auf den Höreindruck überprüften u.a. Deal und Oyer (1991), inwiefern die Einschätzungen von Zuhörern, welche Stimmen sie als angenehm empfanden, von Alter und Geschlecht der Sprecher beeinflusst wurden.

In dieser Arbeit werden unterschiedliche Zuhörergruppen miteinbezogen. So nahmen an vorliegender Studie neben erwachsenen Probanden auch Probanden im Kindes- und Jugendalter (Grundschüler und Gymnasiasten) teil. Gerade aus forensischer Perspektive, wie in diesem Kapitel ausgeführt, erscheint es relevant, herauszufinden, wie valide Zeugenaussagen - auch von Kindern - sind, die nur auf dem Höreindruck einer Stimme fußen. Jedoch sind wenige Studien, bis auf die bei den Kategorien Geschlecht und Akzent/Dialekt genannten, zu finden, die der Frage nachgehen, inwieweit Kinder und Jugendliche bereits ähnliche Informationen aus der Stimme einer Person gewinnen können wie Erwachsene. Außerdem wurden Fachärzte für Phoniatrie und Pädaudiologie als professionell geschulte Zuhörer befragt. In den meisten Studien zur Stimmerkennung handelt es sich bei den Probanden um Laien, nur vereinzelt wurden die von ihnen erzielten Ergebnisse mit denjenigen von geschulten Probanden verglichen (bspw. Braun und Rietveld, 1995).

Somit soll im Zuge dieser Arbeit v.a. folgenden Fragen nachgegangen werden: (i) Ist die Stimmerkennung (bzw. die dafür benötigten Fähigkeiten) altersabhängig, d.h. unterscheidet sich die Stimmerkennung von Kindern von derjenigen von Jugendlichen oder Erwachsenen? (ii) Erzielen die phonetisch geschulten Zuhörer bessere Ergebnisse in der Stimmerkennung als die Laienhörer?

Zur Beantwortung der Fragestellung wurden die Antworten der Probanden mithilfe eines Punktesystems bewertet und die erreichten Gesamtpunktzahlen der unterschiedlichen Hörergruppen ermittelt. Zudem werden die Ergebnisse der Hörergruppen in den einzelnen untersuchten Kategorien dargestellt.

2. Material und Methoden

2.1 Studiendesign

Es handelt sich bei der vorliegenden Arbeit um eine Querschnittsstudie. Den Probanden vierer unterschiedlicher Gruppen wurden mehrere Stimmufnahmen eines phonetischen Mustertextes vorgespielt und sie sollten einen Fragebogen zu den vermuteten Merkmalen der ihnen unbekanntem Sprecher ausfüllen. Der Text wurde zuvor durch unterschiedliche Sprecher vorgelesen und die Stimmufnahmen anschließend digitalisiert. Die Teilnahme der Probanden an der Studie erfolgte freiwillig und ohne Aufwandsentschädigung. Zuvor wurde die Zustimmung der Ethik-Kommission der Medizinischen Fakultät der Universität Bonn eingeholt.

2.2 Erstellung des Materials

2.2.1 Erstellung der Fragebögen

Die Fragebögen zur Befragung der Probanden wurden unter Beratung des Instituts für Medizinische Biometrie, Informatik und Epidemiologie des Universitätsklinikums Bonn erstellt (s. Anhang). Sie umfassen jeweils *eine* auszufüllende DIN-A4-Seite zu jedem Sprecher, bei 23 Sprechern mussten somit 23 Seiten durch den Probanden bearbeitet werden. Zu jedem Sprecher mussten dieselben 12 Fragen beantwortet werden, die sich auf die verschiedenen Merkmale der Sprecher beziehen. Die Fragen lassen sich, wie in 1.3 bereits ausgeführt, thematisch teilweise gruppieren: 4 Fragen beziehen sich auf physische Merkmale (Frage nach Geschlecht, Alter, Körperbau, Körpergröße des Sprechers). 3 Fragen betreffen den Bildungsstand (Frage nach Schulabschluss, Berufsausbildung, Beruf des Sprechers). Weiterhin wird nach Akzent/Dialekt, Rauchverhalten, Persönlichkeit und äußerer Erscheinung des Sprechers gefragt. Ihren subjektiven Hör-eindruck der Stimme des Sprechers sollten die Probanden ebenfalls benennen.

Die Formulierungen im Fragebogen, v.a. die Antwortmöglichkeiten betreffend, sind bewusst vereinfacht, um bestmögliche Verständlichkeit für alle Zuhörergruppen, insbesondere für die jüngeren, zu erreichen. Bei den möglichen Dialekten (Frage 3) ist mit dem umgangssprachlich verwendeten „Sächsisch“ die thüringisch-obersächsische Dialekt-

gruppe gemeint. Die Bezeichnung „Bayerisch“ ist ebenfalls nicht genau und unterscheidet nicht zwischen den in diesem Staatsgebiet gesprochenen Dialektgruppen (Bairisch, Ostfränkisch, Schwäbisch-Alemannisch). Die in der Frage nach dem Beruf (Frage 9) genannten Antwortmöglichkeiten sind stark schematisiert und versuchen eine grobe Bündelung der Vielzahl an Berufsfeldern; hier besteht kein Anspruch auf Vollständigkeit. 2 Beispiele sind der jeweiligen ‚Berufsgruppe‘ angefügt, um diese v.a. den jüngeren Zuhörern verständlicher zu machen.

Bei den Variablen Alter (Frage 2) und Körperlänge (Frage 5) wurden die Antwortmöglichkeiten in Klassen mit gleicher Klassenbreite von 10 aufgeteilt.

Der Körperbau des Sprechers (Frage 4) wurde mithilfe des BMI bestimmt. Dieser ist definiert als Quotient aus Körpermasse in Kilogramm und Körpergröße in Metern zum Quadrat. In Bezug auf den BMI von Erwachsenen wird in der Leitlinie der Deutschen Adipositas-Gesellschaft (2014) eine Adipositas-Klassifikation gemäß der Einteilung der World Health Organization (2000) vorgenommen. Laut WHO liegt bei einem BMI von unter $18,5 \text{ kg}/(\text{m})^2$ Untergewicht vor, Normalgewicht bei einem BMI zwischen $18,5$ und $24,9 \text{ kg}/(\text{m})^2$. Bei einem BMI von 25 bis $29,9 \text{ kg}/(\text{m})^2$ spricht man von Übergewicht, während bei einem BMI ab $30 \text{ kg}/(\text{m})^2$ Adipositas vorliegt. Diese wird weiterhin in 3 Schweregrade unterteilt (Adipositas Grad I: BMI 30 - $34,9 \text{ kg}/(\text{m})^2$, Adipositas Grad II: 35 - $39,9 \text{ kg}/(\text{m})^2$, Adipositas Grad III: $\geq 40 \text{ kg}/(\text{m})^2$).

Davon unterscheidet sich die Berechnung des BMI im Kindes- und Jugendalter. Den Leitlinien der Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter (2015) zufolge müssen Alter und Geschlecht miteinbezogen werden, sodass zur Beurteilung alters- und geschlechtsspezifische Perzentilen herangezogen werden. Die Erstellung von BMI-Perzentilen erfolgte u.a. durch Kromeyer-Hauschild et al. (2001). Darauf basierend schlägt die Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter folgende Einteilung vor: BMI-Perzentile < 10 : Untergewicht, BMI-Perzentile 10 - 90 : Normalgewicht, BMI-Perzentile > 90 - 97 : Übergewicht, BMI-Perzentile > 97 - $99,5$: Adipositas, BMI-Perzentile $> 99,5$: extreme Adipositas. Der BMI der Sprecher im Kindes- und Jugendalter wurde mithilfe des BMI-Rechners der Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter berechnet.

Zur Veranschaulichung insbesondere für die Probanden im Kindes- und Jugendalter wurden vereinfachte, schematische Zeichnungen der Körperformen angefertigt. Hierbei

entsprechen die Abbildungen in aufsteigender Reihenfolge grob der jeweiligen ‚Gewichtsklasse‘. Wie oben beschrieben, umfasst die BMI-Einteilung für Erwachsene insgesamt 6 Klassen, diejenige für Kinder und Jugendliche beinhaltet jedoch 5 Klassen. Daher wurden zur Vereinheitlichung bei den Erwachsenen die Adipositas-Grade II und III in Abb. 5 zusammengefasst. Dies ist auch zu der Auswahl der Sprecher passend: Lediglich ein Sprecher liegt mit seinem BMI ($40,1 \text{ kg/m}^2$) knapp im Bereich der Adipositas Grad III.

2.2.2 Erstellung der Stimmaufnahmen

Um die Daten bzw. Merkmale der vorlesenden Personen zu erfassen, sollten diese zunächst ein Informationsblatt über sich selbst ausfüllen (s. Anhang). Dies umfasste auch Selbsteinschätzungen hinsichtlich Persönlichkeit und äußerer Erscheinung für die Auswertung der Fragen 10 und 11 des Fragebogens.

Die Stimmaufnahmen wurden mithilfe eines portablen Audiorekorders (Philips Voice Tracer DVT2500) in einer geräuscharmen Umgebung aufgezeichnet und anschließend digitalisiert, sodass sie von einem Computer aus abgespielt werden konnten. Vorgelesen werden sollte der phonetische Mustertext „Der Nordwind und die Sonne“, welcher von der International Phonetic Association (1999) übersetzt wurde. Zuvor konnten die Sprecher sich den Text durchlesen, um sich damit vertraut zu machen. Dies war insbesondere für die Personen mit Dialekt von Bedeutung, da der Text in Hochdeutsch vorgelegt wurde. Anschließend sollten die Sprecher den Text von ca. 1 min Dauer so flüssig und natürlich wie möglich sowie in normaler Sprechlautstärke vorlesen. Die Aufnahme konnte wiederholt werden, bspw. wenn durch Fehler beim Vorlesen der Sprachfluss beeinträchtigt war bzw. ein Sprecher die Aufnahme spontan abbrach. Jedoch wurde grundsätzlich darauf geachtet, die individuelle Art des Vorlesens jedes Sprechers so wenig wie möglich zu beeinflussen, um den jeweiligen Eindruck bei den Probanden nicht zu verfälschen.

2.3 Auswahl der Studienteilnehmer

2.3.1 Auswahl der vorlesenden Personen

Bei der Auswahl der vorlesenden Personen war die Zielsetzung, eine heterogene Stichprobe mit einem möglichst breiten Spektrum an Merkmalen zu erhalten, um letztere adäquat abfragen zu können. Gleichzeitig musste die Anzahl der Stimmaufnahmen jedoch begrenzt werden, um den Studienablauf zeitlich im Rahmen zu halten und insbesondere die Konzentrationsfähigkeit der jüngeren Probanden nicht zu überfordern.

Im Rahmen dieser Studie wurden insgesamt 23 vorlesende Personen im Alter zwischen 9 und 91 Jahren, davon 13 weiblichen und 10 männlichen Geschlechts, ausgewählt. Einen Überblick über die Verteilung der Merkmale gibt Tab. 1.

Tab. 1: Überblick über die Merkmalsverteilung der vorlesenden Personen

Vorlesende Personen		Geschlecht		Anzahl
		männlich	weiblich	
Alter (in Jahren)	5-14	2	1	3
	15-24	0	1	1
	25-34	1	2	3
	35-44	2	2	4
	45-54	0	2	2
	55-64	3	2	5
	65-74	1	1	2
	75-84	1	0	1
	85-94	0	2	2
Akzent	Französisch	0	1	1
	Englisch/Amerikanisch	0	1	1
	Türkisch/Arabisch	1	0	1
Dialekt	Kein Dialekt (Hochdeutsch)	7	4	11
	Bayerisch	1	1	2
	Sächsisch	0	1	1
	Schwäbisch	1	1	2
	Kölsch	0	3	3
	Norddeutsch	0	1	1
Körperbau	Untergewicht	0	2	2
	Normalgewicht	4	8	12
	Übergewicht	2	1	3
	Adipositas Grad I	1	0	1
	Adipositas Grad II/III	3	2	5
Körperlänge (in cm)	131-140	1	0	1
	141-150	1	0	1
	151-160	0	6	6
	161-170	0	5	5
	171-180	6	2	8
	181-190	1	0	1
	191-200	1	0	1
Rauchverhalten	RaucherIn	2	3	5
	kein/e RaucherIn	8	10	18
Schulabschluss	Gymnasium/Gesamtschule	4	5	9
	Realschule	3	5	8
	Hauptschule	1	2	3
	kein Schulabschluss	0	0	0
	noch SchülerIn	2	1	3

Beruflicher Abschluss	Studium	5	2	7
	Ausbildung/Lehre	3	10	13
	keine Berufsausbildung	2	1	3
Beruf	ArbeiterIn	0	0	0
	Bürotätigkeit	0	6	6
	HandwerkerIn	3	4	7
	AkademikerIn	5	2	7
	noch nicht berufstätig	2	1	3
Persönlichkeit	nach innen gekehrt	5	5	10
	nach außen gekehrt	5	8	13
äußere Erscheinung	alltäglich/normal	10	9	19
	vornehm/elegant	0	4	4
Gesamtzahl		10	13	23

Die für die Studie ausgewählten Sprecher mit Akzent wuchsen in ihren jeweiligen Heimatländern (Marokko, USA, Frankreich) auf, bevor sie nach Deutschland kamen und die deutsche Sprache erlernten. Somit sprechen sie Deutsch mit ihrem jeweiligen muttersprachlichen bzw. fremdsprachigen Akzent. Insbesondere die Auswahl der Sprecher mit Dialekt war schwierig zu treffen, nicht nur aufgrund der logistischen Gegebenheiten, sondern v.a. aufgrund der Herausforderung beim Vorlesen für die Sprecher. Sie empfanden es oftmals als ungewohnt und anspruchsvoll, einen hochdeutschen Text in Mundart vorzulesen, da im Dialekt einzelne Wörter und teilweise sogar der Satzbau vom Hochdeutschen abweichen.

2.3.2 Auswahl der Probanden

Die Probanden der Studie lassen sich in 4 Hörergruppen einordnen. Es wurden die Schüler von vier 3. Grundschulklassen, drei 9. Klassen der gymnasialen Mittelstufe, Studierende sowie Fachärzte für Phoniatrie und Pädaudiologie einbezogen.

Zwei 3. Grundschulklassen à 11 und 14 Schüler der Waldschule Bonn-Venusberg nahmen an der Studie teil, ebenso wie zwei 3. Klassen à 11 und 12 Schüler der Gemeinschaftsgrundschule Drabenderhöhe (Wiehl). Somit nahmen insgesamt 48 Probanden aus der Gruppe der Grundschüler teil.

Im Beethoven-Gymnasium Bonn wurden drei 9. Klassen à 22, 19 sowie 29 Schüler und damit 70 Gymnasiasten befragt.

Die Studierenden wurden im Rahmen des Blockpraktikums der Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde/Chirurgie des Universitätsklinikums Bonn für die Studie gewonnen und es fanden mehrere Befragungen mit Gruppen von jeweils ca. 10 Studierenden statt, so konnten insgesamt 52 Studierende befragt werden.

Die Fachärzte für Phoniatrie und Pädaudiologie wurden auf der Jahrestagung der Deutschen Gesellschaft für Phoniatrie und Pädaudiologie in Innsbruck 2017 um ihre Teilnahme gebeten. Des Weiteren wurden Fragebögen mitsamt CDs der Stimmaufnahmen postalisch an einige Phoniater verschiedener Kliniken und Praxen in Deutschland sowie in Österreich und der deutschsprachigen Schweiz gesendet. Die Teilnehmerzahl der Phoniater belief sich auf 32.

Ursprünglich war geplant, mit Mitarbeitern des Bundeskriminalamtes noch eine 5. Gruppe in die Studie miteinzubeziehen. Die Mitarbeiter am Kriminaltechnischen Institut des Bundeskriminalamtes in Wiesbaden sind auf Sprechererkennung spezialisiert und verfügen über ein Spracherkennungssystem als Instrument der forensischen Phonetik. Die Ergebnisse der ausgewiesenen Experten für Stimmerkennung sollten mit denjenigen der Phoniater als Spezialisten für Stimme aus dem medizinischen Bereich und mit den Ergebnissen der phonetischen Laien verglichen werden.

Allerdings entschied sich das Sprechererkennungsteam gegen eine Teilnahme an der Studie, was durch Herrn Dr. Michael Jessen ausführlich begründet wurde. Zunächst würden er und seine Mitarbeiter die Stimmaufnahmen nicht nur auditiv analysieren, sondern im Rahmen der auditiv-akustischen Methode auch akustische Parameter messen wie bspw. Grund- und Formantfrequenzen. Bei der Erstellung eines Sprecherprofils würde auch Forschungsliteratur miteinbezogen werden. Diese Vorgehensweise sei nicht nur sehr zeitaufwendig und würde mehrere Tage beanspruchen, sondern unterscheide sich auch von derjenigen der anderen Hörergruppen, die die genannten Aspekte nicht einbeziehen. Somit wären die Ergebnisse nicht mit denen der anderen Hörergruppen vergleichbar.

Des Weiteren würden sie Aussagen in Form von Wahrscheinlichkeitsgraden treffen, abhängig von der Qualität des Materials, der Forschungslage sowie der eigenen Erfahrungen. Dazu sei die Beantwortung eines Fragebogens nicht geeignet. Zuletzt schlossen Herr Dr. Jessen und sein Team die Beantwortung von Fragen zu psychologischen Parametern aus, da sie diese Kategorie grundsätzlich nicht allein anhand

von Sprachaufnahmen beurteilen würden. Dies bezieht sich auf die Fragen nach der Persönlichkeit (Frage 10), äußeren Erscheinung (Frage 11) und dem Höreindruck (Frage 12) des Sprechers. Die Schwierigkeit der Beantwortung von Fragen zu psychologischen Parametern wurde allerdings hinsichtlich der Bewertung berücksichtigt (s. 2.5).

2.4 Ablauf der Befragungen

Im Vorfeld der Befragungen wurden Informationsschreiben und Einverständniserklärungen an die Probanden verteilt bzw. versendet (s. Anhang). Die Grundschüler erhielten darüber hinaus ein Informationsblatt zu der Studie in einfacher, kindgerechter Sprache (s. Anhang). Bei den Grundschülern und Gymnasiasten war die Einwilligung *beider* Elternteile erforderlich. Es wurde angeboten, dass die Studienteilnehmer auch selbst bei möglichen Fragen oder Unsicherheiten Kontakt aufnehmen könnten. Zu diesem Zwecke wurde eine E-Mail-Adresse angegeben.

Die Befragungen der Probanden liefen nach einem standardisierten Schema ab. Zu Beginn wurde gemeinsam mit den Grundschülern und Gymnasiasten, bei den Studierenden und Phoniatern in kürzerer Ausführung, der Fragebogen durchgesprochen, um ihnen einen ersten Überblick über die Fragen zu verschaffen. Den Grundschülern wurden ihnen unbekannte Begriffe erklärt und Rückfragen beantwortet. Anschließend wurden die Stimmufnahmen in einer vorher festgelegten, zufällig generierten Reihenfolge nacheinander von einem Computer abgespielt und mit Lautsprechern verstärkt. Währenddessen sollten die Probanden mit dem Ausfüllen der Fragebögen beginnen. Im Anschluss an jede Stimmufnahme blieb den Probanden ca. 1 min Zeit, um ihre Angaben zu der jeweiligen Person zu vervollständigen. Nach der Hälfte der Aufnahmen wurde bei den Schülern eine ca. 10-minütige Pause eingelegt, um ihre Konzentration auch bei den restlichen Aufnahmen sicherzustellen, und sie erhielten eine kleine Belohnung. Es wurde auf das vollständige Ausfüllen des Fragebogens und selbstständiges Arbeiten hingewiesen bzw. die Schüler wurden wie bei einer schriftlichen Leistungsüberprüfung auseinander gesetzt oder es wurden Sichtschutze aufgestellt.

Die Befragungen erfolgten anonym, die Probanden sollten lediglich ihr Alter und Geschlecht angeben.

2.5 Auswertung der Fragebögen

Die Antworten der Probanden wurden hinsichtlich des Gesamtergebnisses der Hörergruppen mithilfe eines Punktesystems ausgewertet. Hiervon ausgeschlossen waren die Fragen 10, 11 und 12. Die Antworten auf Frage 10 nach der Persönlichkeit und Frage 11 nach der äußeren Erscheinung beruhen auf einer Selbsteinschätzung des jeweiligen Sprechers und nicht auf einer objektiven Messung, sodass eine Bewertung der Antworten der Probanden als „richtig“ oder „falsch“ bzw. eine Punktevergabe wenig sinnvoll erscheinen. Auf Frage 12 nach dem von den Zuhörern empfundenen Höreindruck der Stimme des Sprechers gibt es keine vorgegebene Antwort.

Da einige Fragen aufgrund einer höheren Anzahl an möglichen anzukreuzenden Antworten schwieriger zu beantworten waren als andere (vgl. bspw. Frage 1 nach dem Geschlecht mit 2 möglichen Antworten und Frage 2 nach dem Alter mit 9 Antwortmöglichkeiten), wurden erstere Fragen mit einer höheren Punktzahl bewertet (2-3 Punkte) als die Fragen mit nur zwei Antwortmöglichkeiten (1 Punkt). Zudem wurde bei den Fragen mit > 2 Antwortmöglichkeiten die Stärke der Abweichung von der richtigen Antwort in der Punktevergabe berücksichtigt. Auf diese Weise konnten auch Tendenzen bewertet werden und ein Proband, der mit seiner Antwort nur wenig von der richtigen Antwort abwich, erhielt dennoch 1-2 Punkte. Zur Veranschaulichung zwei Beispiele:

- 1) Ein Proband hat bei Sprecher Nr. 2 die Frage 2 nach dem Alter mit „45-54“ beantwortet, tatsächlich ist Sprecher 2 jedoch 66 Jahre alt und somit wäre „65-74“ die korrekte Antwort gewesen. Aufgrund der Abweichung von 2 Klassenbreiten von der richtigen Antwort erhält der Proband 1 von max. 3 Punkten.
- 2) Frage 7 nach dem Schulabschluss von Sprecher Nr. 2 hat der Proband mit „Hauptschule“ beantwortet, der Sprecher hat jedoch einen Realschulabschluss erworben (mittlerer Schulabschluss). Die Antwortmöglichkeiten zu Frage 7 und auch Frage 8 nach dem beruflichen Abschluss sind in hierarchischer Reihenfolge nach der Höhe des Bildungsabschlusses geordnet. Somit hat der Proband statt der korrekten Antwort den nächst niedrigeren Schulabschluss vermutet und erhält 1 von max. 2 Punkten.

2.6 Statistische Auswertung der Ergebnisse

Zur Sammlung und Verarbeitung der erhobenen Daten diente das Statistikprogramm SPSS (Softwarefirma IBM, Version 25), sämtliche statistische Analysen wurden ebenfalls mithilfe von SPSS durchgeführt. Die Tabellen und Abbildungen wurden mittels SPSS und Microsoft Excel 2013 erstellt. Die statistische Auswertung erfolgte unter Beratung von Statistikern des Hochschulrechenzentrums der Universität Bonn.

Zur Darstellung des Gesamtergebnisses der Probanden (s. 3.3) sowie der Ergebnisse nach Variablen (s. 3.4) wurde zunächst eine deskriptive Statistik mit Erhebung von Häufigkeiten, Mittelwerten und Standardabweichungen durchgeführt. Bei den ordinalskalierten Variablen (Alters-, Körpergrößen-, BMI-Klassen) wurde zusätzlich der Spearman-Korrelationskoeffizient berechnet, welcher den linearen Zusammenhang zweier Variablen wiedergibt; dieser kann Werte von -1 bei einem perfekten negativen Zusammenhang bis +1 bei einem perfekten positiven Zusammenhang annehmen. Bei einem Wert von 0 liegt keine Korrelation vor. Zur grafischen Veranschaulichung der Verteilung der Werte, respektive der korrekten Antworten der Probanden, dienten Boxplot-Diagramme. Hier von ausgenommen sind die Items, bei denen die Antworten der Probanden nicht als „richtig“ oder „falsch“ bewertet werden können (s. 2.5). Die Box wird dabei durch das obere und untere Quartil begrenzt, ihre Höhe ist dem Interquartilsabstand entsprechend, und umfasst die mittleren 50 % der Werte. Die durchgehende Linie innerhalb der Box entspricht dem Median. Die Enden der Antennen stellen den jeweils größten bzw. kleinsten Wert dar; beim Vorliegen von Ausreißern oder Extremwerten enden die Antennen bei den Werten, die noch innerhalb des 1,5fachen Interquartilsabstandes liegen. Als Ausreißer sind diejenigen Werte definiert, deren Abstand vom oberen Quartil nach oben bzw. vom unteren Quartil nach unten mehr als das 1,5fache des Interquartilsabstandes beträgt; sie sind in den Boxplots in Form von einzelnen Punkten dargestellt. Extremwerte (durch Sterne gekennzeichnet) liegen mehr als 3 Interquartilsabstände oberhalb des oberen bzw. unterhalb des unteren Quartils.

Zur Prüfung der Ergebnisse auf signifikante Unterschiede wurde bei den *metrisch skalierten* Daten, wie bspw. den Häufigkeiten der korrekten Angaben, eine einfaktorische Varianzanalyse (ANOVA) durchgeführt. Als Voraussetzung dieses Tests konnte nach dem zentralen Grenzwertsatz von einer Normalverteilung der Daten ausgegangen werden, weil die Stichprobengröße der 4 Probandengruppen jeweils mindestens 30 beträgt

(s. 3.2). Varianzhomogenität war gemäß Levene-Test ($p < 0,001$) bei jeglichen Varianzanalysen nicht gegeben, sodass stets eine Welch-ANOVA angewendet wurde. Bei zwei statistischen Testungen lag eine andere Gruppierungsvariable als die Probandengruppe vor und die Stichprobengröße war kleiner als 30. In diesen beiden Fällen (s. 3.4.1 und 3.4.4) wurden die Daten mittels des Kolmogorov-Smirnov-Tests auf Normalverteilung geprüft und der Kruskal-Wallis-Test sowie der Mann-Whitney-U-Test angewendet. Für die *nominalskalierten* Variablen wurde nach Erstellung einer Kreuztabelle ein Chi-Quadrat-Test (χ^2) durchgeführt. Bei sämtlichen Chi-Quadrat-Tests waren die erwarteten Zellhäufigkeiten in den Kreuztabellen wie vorausgesetzt größer als 5. Für die statistischen Tests wurde ein Signifikanzniveau von $\alpha \leq 0,05$ festgelegt.

3. Ergebnisse

3.1 Umgang mit fehlenden und ungültigen Werten

In der Auswertung der Fragebögen wurden fehlende Werte, d.h. Fragen, bei denen die Probanden keine Antwortmöglichkeit ankreuzten, von ungültigen Werten unterschieden; in letzterem Fall handelte es sich vornehmlich um das Ankreuzen von mehr als einer Antwortmöglichkeit bei einer Frage.

Es wurde festgelegt, diejenigen Probanden aus der Studie auszuschließen, deren Antworten zu mindestens 15 % fehlende Werte oder ungültige Werte oder die Summe aus beiden darstellten. Bei einer hohen Anzahl fehlender oder ungültiger Werte fallen die Punktzahlen sehr niedrig aus und entsprechen in diesen Fällen vermutlich nicht der tatsächlichen Leistung in der Stimmerkennung der Zuhörer. Wahrscheinlicher ist, dass diese aufgrund mangelnder Konzentration oder Motivation bei vielen Fragen keine Antwortmöglichkeit ankreuzten oder die Aufgabenstellung nicht verstanden haben. Überdies entsprechen die Probanden mit mindestens 15 % Fehlwerten den Extremwerten im Boxplot der erreichten Punktzahl im Fragebogen, was auf den Zusammenhang zwischen Fehlwerten und schlechter Gesamtpunktzahl hinweist. Die Extremwerte im Boxplot befinden sich mehr als 3 Interquartilsabstände unterhalb des unteren Quartils, welches zusammen mit dem oberen Quartil die mittleren 50 % der Werte umfasst. Insgesamt wurden 5 Probanden aus der statistischen Auswertung ausgeschlossen.

3.2 Allgemeine Daten der Probanden

Nach Ausschluss der 5 Probanden mit mindestens 15 % Fehlwerten verblieben insgesamt 197 Probanden, davon 89 männlich und 108 weiblich, im Alter von 8 bis 74 Jahren. Darunter waren 48 Grundschüler zu verzeichnen, davon 25 Jungen und 23 Mädchen. Die Altersspanne betrug 8-11 Jahre, das Durchschnittsalter 8,52 Jahre.

Es handelte sich um 69 Gymnasiasten, 34 Jungen und 35 Mädchen, im Alter von 13 bis 16 Jahren mit einem Durchschnittsalter von 14,45 Jahren.

Die Anzahl der Studierenden belief sich auf 50, davon 19 männlichen und 31 weiblichen Geschlechts, im Alter von 21 bis 31 Jahren, durchschnittlich 23,66 Jahren.

Die Anzahl der Phoniater betrug 30 mit 11 Männern und 19 Frauen, im Alter von 32 bis 74 Jahren und mit einem durchschnittlichen Alter von 49,23 Jahren.

3.3 Gesamtergebnis der Probanden

Zunächst sollen die im Fragebogen erreichten Gesamtpunktzahlen der 4 Hörergruppen dargestellt werden, um die ‚Gesamtleistungen‘ in der Stimmerkennung miteinander vergleichen zu können (ausgenommen Fragen 10-12 des Fragebogens, s. 2.5).

Die durchschnittlich erreichten Punktzahlen in Prozent nahmen von Grundschulern ($M = 59,47$, $SD = 7,82$) zu Gymnasiasten ($M = 72,95$, $SD = 3,52$) zu. Die höchste durchschnittliche Punktzahl erreichten die Studierenden ($M = 75,65$, $SD = 3,13$), gefolgt von den Phoniatern ($M = 74,87$, $SD = 3,61$). Die höhere Streuung der Werte bei den Grundschulern lässt sich auch im Boxplot (Abb. 2) anhand des größeren Interquartilsabstandes entsprechend der Länge der Box erkennen.

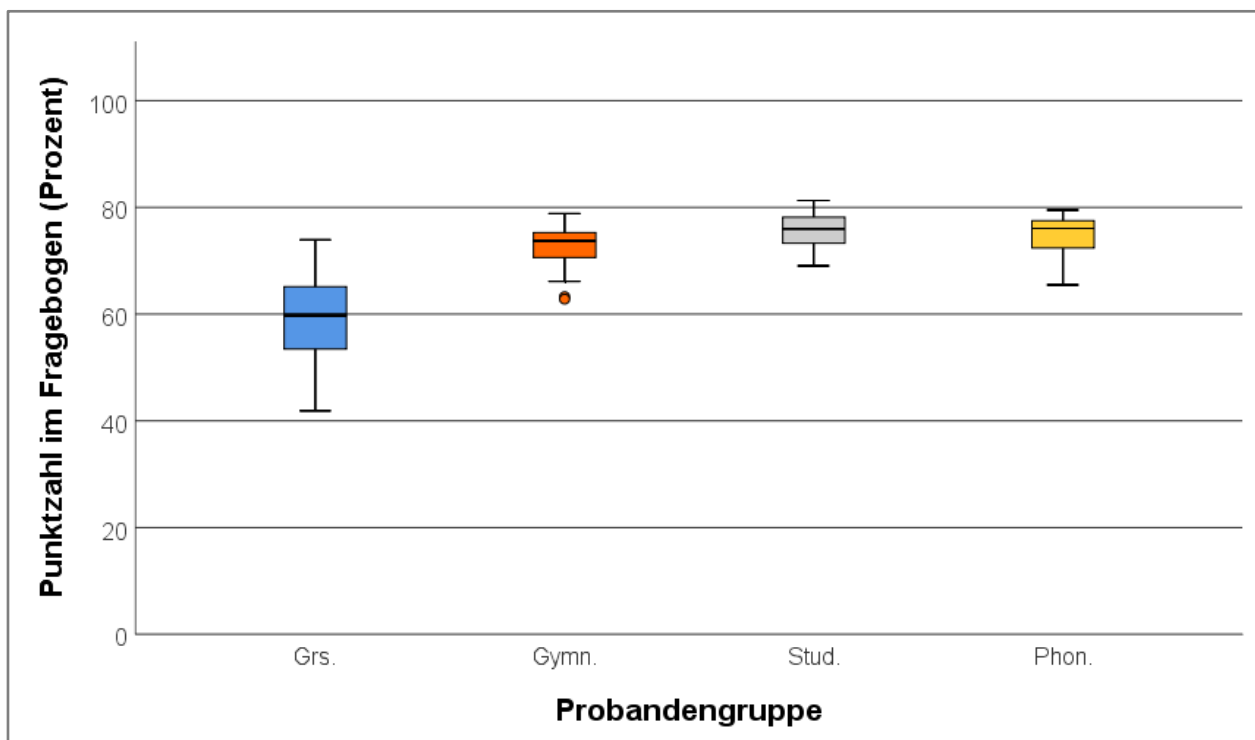


Abb. 2: Boxplot der Punktzahl im Fragebogen (Prozent) sortiert nach den Probandengruppen (Grs. = Grundschüler; Gymn. = Gymnasiasten; Stud. = Studierende; Phon. = Phoniater)

Es wurde eine einfaktorielle Varianzanalyse (ANOVA) durchgeführt, um zu untersuchen, ob es einen signifikanten Unterschied in der Fähigkeit der Stimmerkennung (gemessen durch die erreichte Punktzahl in Prozent im Fragebogen) abhängig von der Probandengruppe gab. Das Ergebnis der Welch-ANOVA war signifikant, Welch-Test $F(3, 89,09) = 60,51$, $p < 0,001$. Der Games-Howell Test zeigte einen signifikanten Unterschied ($p < 0,001$) zwischen folgenden Gruppen: Die durchschnittliche erreichte Punktzahl im Fragebogen nahm zu, von Grundschulern zu Gymnasiasten, von Grundschulern zu Studierenden und zu den Phoniatern sowie von Gymnasiasten zu Studierenden. Phoniater und Gymnasiasten ($p = 0,081$) und Phoniater und Studierende ($p = 0,758$) wiesen keine signifikant unterschiedlichen Punktzahlen auf.

Des Weiteren sollen die durchschnittlichen Anzahlen der fehlenden und ungültigen Werte erläutert werden. Es fällt auf, dass diese insgesamt zwar niedrig ausfallen, die Grundschüler aber mehr fehlende Werte als die anderen Probandengruppen aufweisen: Durchschnittlich haben sie 2,54 fehlende Werte pro Fragebogen und somit bei 1,23 % der Fragen keine Antwortmöglichkeit ausgewählt, während die anderen Zuhörergruppen in unter 1 % der Fragen keine Antwortmöglichkeit ankreuzten; die Phoniater führend mit 0,87 %. Bei den ungültigen Werten pro Fragebogen liegen die Grundschüler bei durchschnittlich 3,42, damit bei 1,65 % der Fragen. Gymnasiasten, Studierende und Phoniater bleiben hinsichtlich der ungültigen Antworten bei unter 0,3 % der Fragen. In Abb. 3 sind die Mittelwerte der ungültigen und fehlenden Werte der 4 Probandengruppen dargestellt.

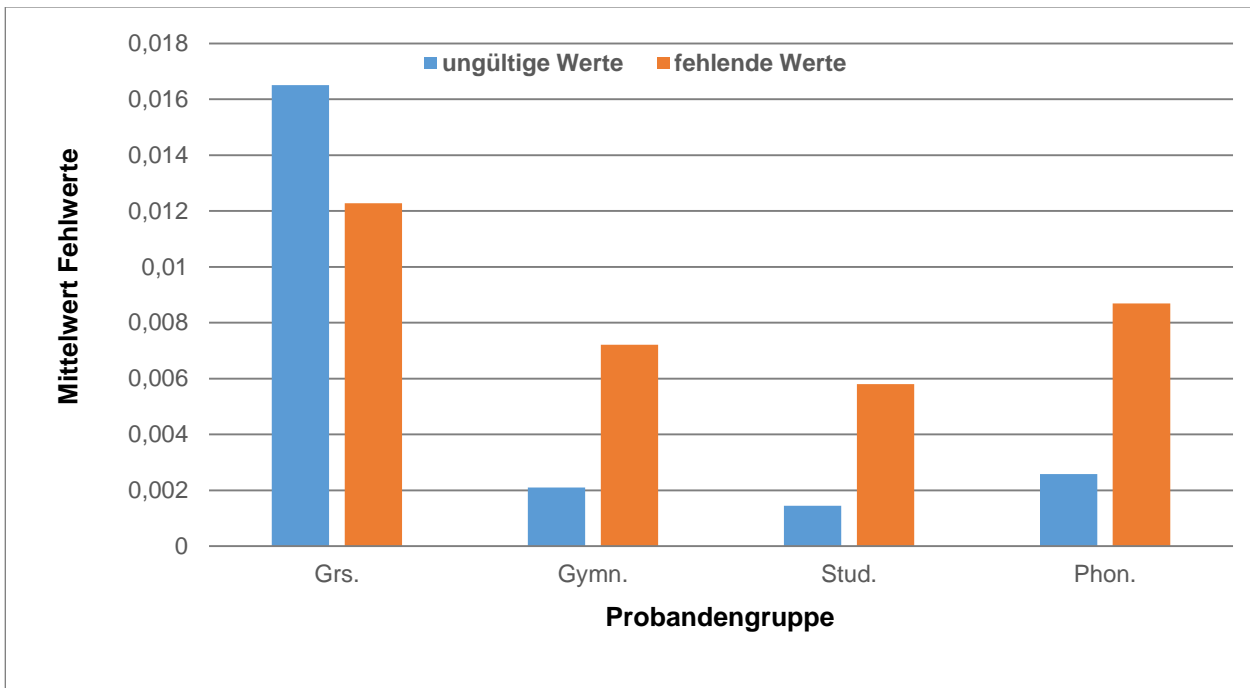


Abb. 3: Mittelwerte der ungültigen und fehlenden Werte sortiert nach den Probandengruppen

3.4 Ergebnisse nach Variablen

3.4.1 Alter

Die Probanden sollten die vermuteten Altersgruppen der Sprecher angeben (s. Frage 2 des Fragebogens). Insgesamt gab es 9 Antwortmöglichkeiten bzw. mögliche Altersdekaden von 5 bis 94 Jahren. Es wurde hinsichtlich der Altersgruppen der Spearman-Korrelationskoeffizient berechnet. Es ergab sich eine stark positive Korrelation mit $r = 0,772$ ($p < 0,001$) zwischen der vermuteten und der tatsächlichen Altersgruppe.

Für die grafische Darstellung in einem Streudiagramm (Abb. 4) wurden die Zentralwerte der tatsächlichen Altersgruppen ermittelt, diese stimmen mit den jeweiligen Mittelwerten überein; bspw. beträgt der Zentralwert der Altersgruppe 5-14 Jahre 9,5 Jahre, der Zentralwert der Altersgruppe 15-24 Jahre 19,5 Jahre usw. Sie wurden gegen die Mittelwerte der geschätzten Alterszuordnungen der 4 Zuhörergruppen aufgetragen. Es zeigt sich dem Koeffizienten entsprechend ein positiver linearer Zusammenhang. Insgesamt ist ersichtlich, dass die Zuhörer häufig das Alter jüngerer Sprecher überschätzten und das Alter älterer Sprecher unterschätzten. Insbesondere die Grundschüler (Altersklasse 8-11 Jahre) wichen in ihren Altersschätzungen entsprechend nach oben bzw. unten ab. Die

geringste Abweichung der Schätzungen von der korrekten Altersgruppe lag bei den Sprechern im Alter zwischen 45 und 54 Jahren vor.

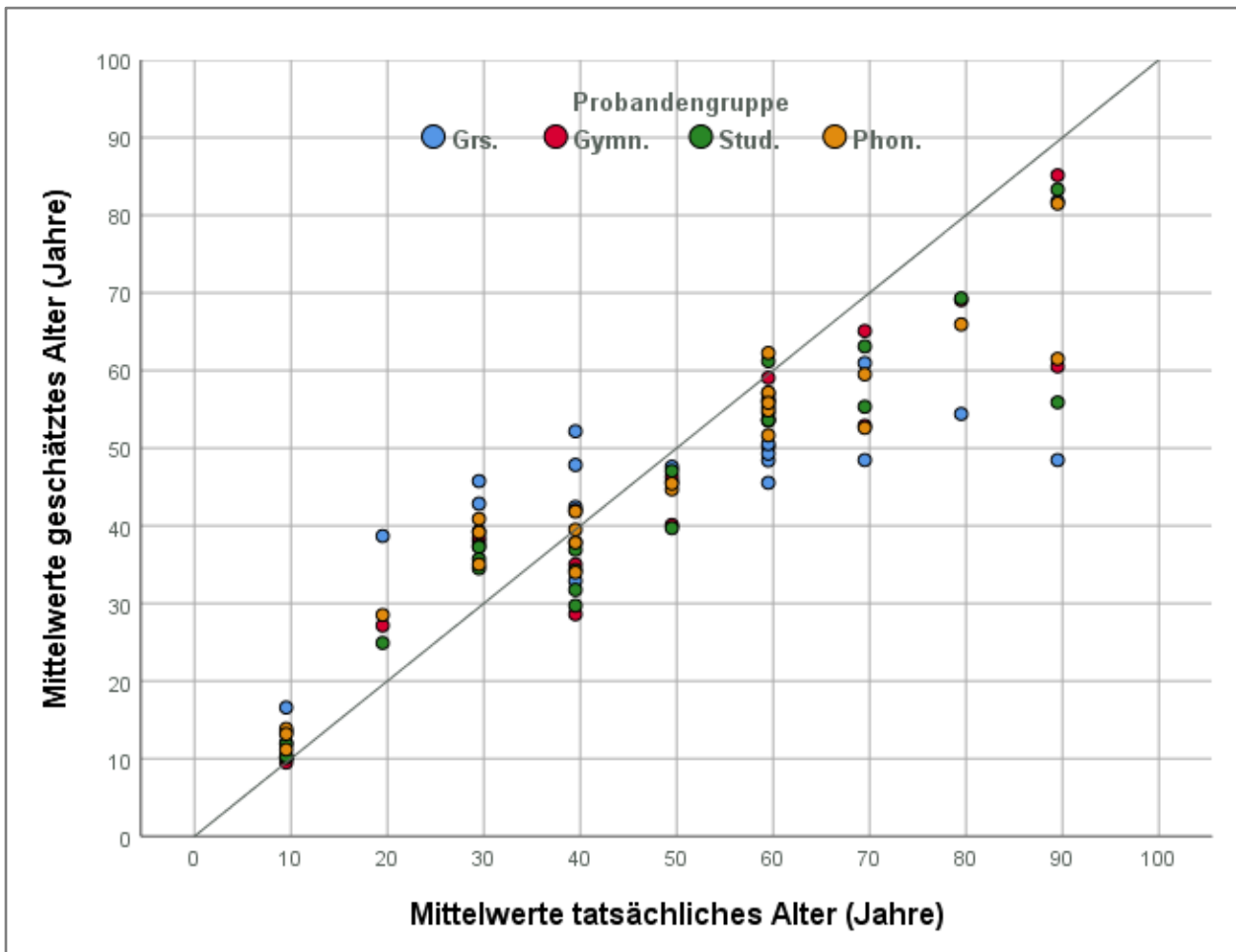


Abb. 4: Streudiagramm der Mittelwerte der geschätzten und tatsächlichen Altersgruppen der Sprecher. Die durchgezogene Linie stellt dar, wo geschätztes und korrektes Alter übereinstimmen.

Um die Beobachtung der Über- und Unterschätzung des Alters eingehender zu untersuchen, wurden die Antwortmöglichkeiten der Frage bzw. die möglichen Altersdekaden von 1 bis 9 durchnummeriert; für jeden Probanden wurde die Differenz von geschätzter und tatsächlicher Altersdekade berechnet. Anhand des Vorzeichens ist die Richtung der Abweichung von der korrekten Altersdekade erkennbar sowie anhand der Größe des Betrages die Stärke der Abweichung. Bspw. ergibt sich bei einer geschätzten Altersgruppe von 65-74 Jahren und der tatsächlichen Altersgruppe 75-84 Jahre eine Differenz von $7 - 8 = -1$, wobei das negative Vorzeichen die Unterschätzung des Alters angibt. In Tab. 2 sind die Mittelwerte und Standardabweichungen der Differenzen pro Probandengruppe und Altersgruppe der Sprecher aufgeführt sowie in Abb. 5 grafisch dargestellt.

Die Sprecher wurden in 3 breitere Altersgruppen (5-34, 35-64, 65-94 Jahre) aufgeteilt. Das Alter der Kinder/Jugendlichen und jungen Erwachsenen (5-34 Jahre) wurden von den Zuhörern überschätzt, während die älteren Sprecher (65-94 Jahre) jünger geschätzt wurden. Auch die Erwachsenen im mittleren Alter (35-64) wurden eher jünger geschätzt (negatives Vorzeichen), jedoch mit geringerer Abweichung von der korrekten Altersgruppe als die älteren Sprecher. Weiterhin bestätigt sich, dass die Grundschüler stärkere Abweichungen aufweisen sowie eine höhere Streuung der Werte.

Tab. 2: Mittelwerte \pm Standardabweichungen der Differenzen geschätztes Alter - tatsächliches Alter sortiert nach Altersgruppen der Sprecher (Jahre) und Probandengruppen (Grs. = Grundschüler; Gymn. = Gymnasiasten; Stud. = Studierende; Phon. = Phoniater)

Altersgruppe Sprecher (Jahre)	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
5-34	0,62\pm1,01	1,03 \pm 1,40	0,49 \pm 0,82	0,40 \pm 0,79	0,64 \pm 0,81
35-64	-0,40\pm1,21	-0,38 \pm 1,72	-0,39 \pm 1,05	-0,52 \pm 0,95	-0,26 \pm 0,93
65-94	-1,55\pm1,57	-2,07 \pm 2,09	-1,29 \pm 1,40	-1,41 \pm 1,31	-1,53 \pm 1,12

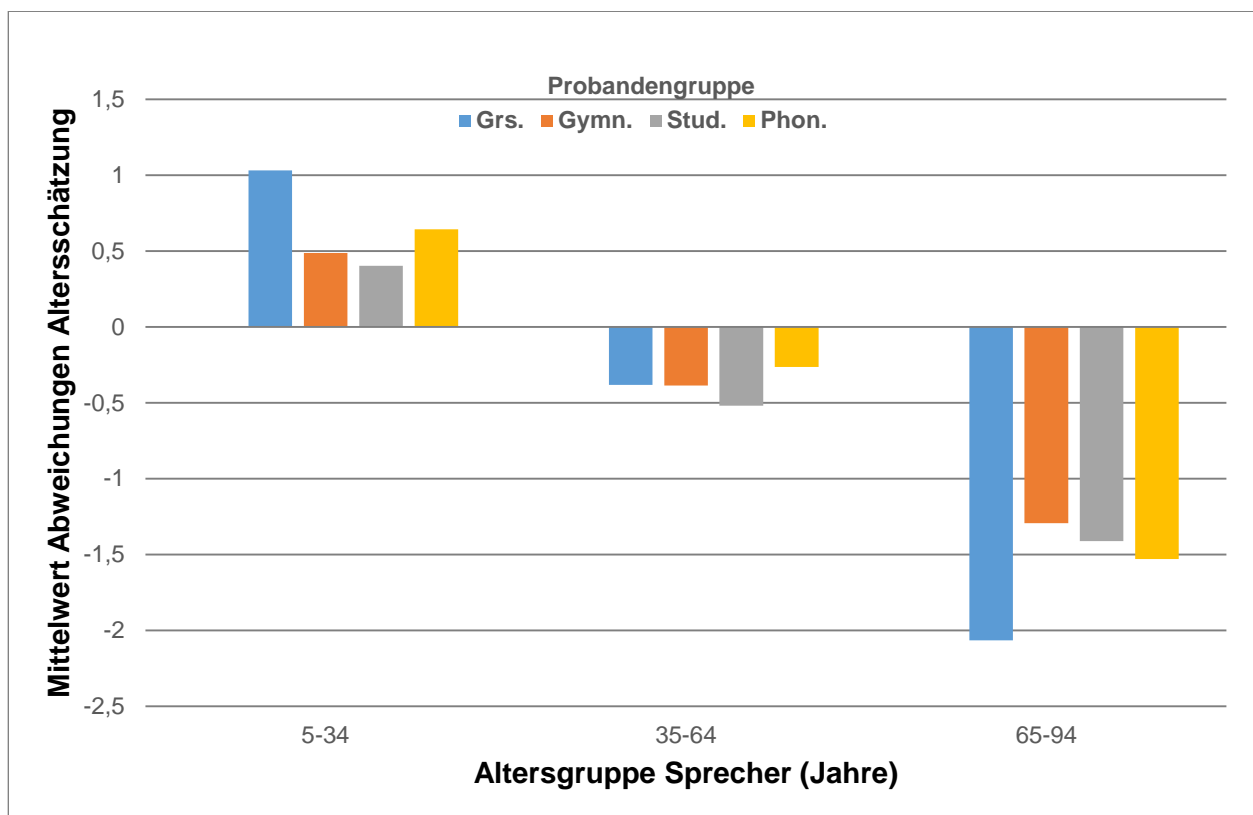


Abb. 5: Mittelwerte der Differenzen geschätzte Altersgruppe - tatsächliche Altersgruppe, sortiert nach Altersgruppen der Sprecher (Jahre) und Probandengruppen

Weiterhin sollen die Häufigkeiten der korrekten Alterszuordnungen dargestellt werden. Insgesamt ordneten die Probanden die Sprecher in 1609 von 4531 Fällen (35,51 %) ihrer korrekten Altersgruppe zu. Wenn ein Fehler von ± 1 Altersgruppe, d.h. eine Abweichung von ± 1 von der korrekten Altersgruppe, gestattet wurde, lagen in 3361 von 4531 Fällen (74,18 %) richtige Alterszuordnungen vor. Tab. 3 zeigt die Prozentzahlen der korrekten Antworten sowie diejenigen bei einem Fehler von ± 1 Altersgruppe und ± 2 Altersgruppen, welche die 4 Zuhörergruppen erreicht haben.

Tab. 3: Korrekte Alterszuordnungen sowie Alterszuordnungen bei einem Fehler von ± 1 und ± 2 Altersgruppe(n) (Prozent) sortiert nach den Probandengruppen

Alterszuordnung	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
korrekt	35,51	27,45	38,31	39,83	34,78
± 1 Altersgruppe	74,18	57,52	78,58	81,04	79,28
± 2 Altersgruppen	90,07	79,44	93,64	93,56	93,04

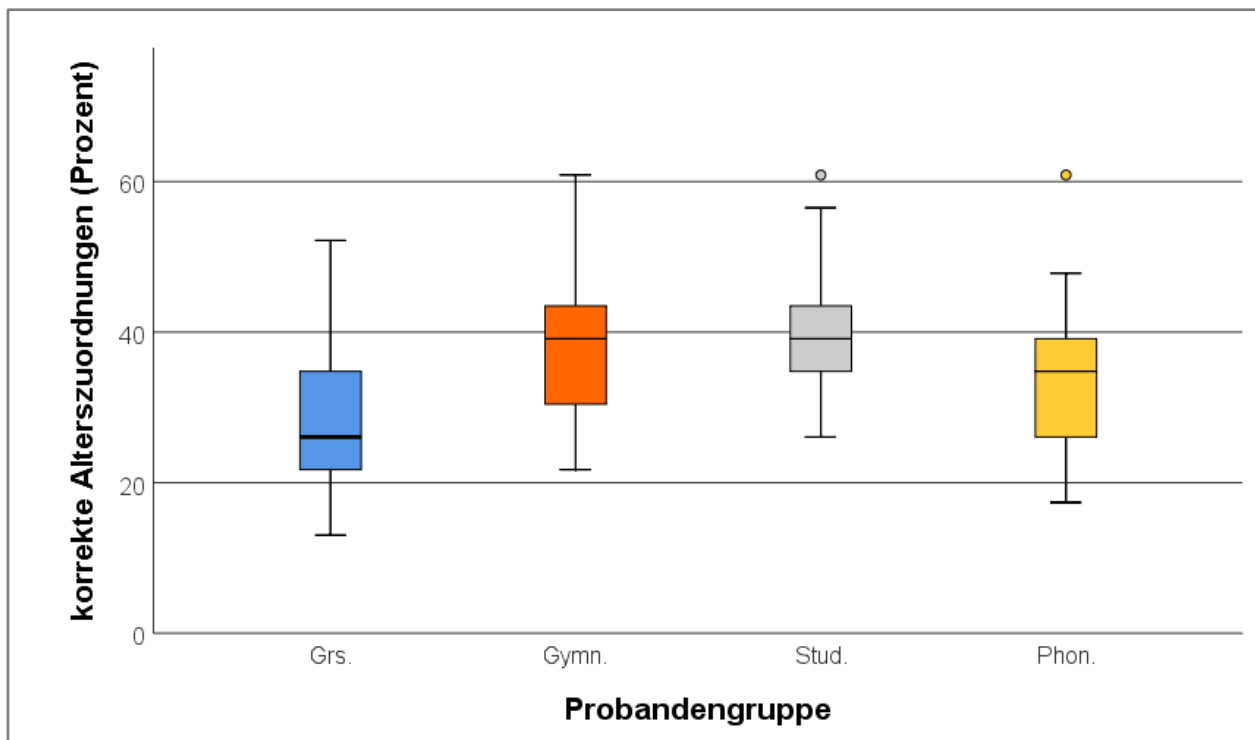


Abb. 6: Boxplot der korrekten Alterszuordnungen (Prozent) sortiert nach den Probandengruppen

Eine Welch-ANOVA ergab einen signifikanten Unterschied der korrekten Altersschätzungen für die 4 Zuhörergruppen, Welch-Test $F(3, 2169,32) = 16,69$, $p < 0,001$. Im Ga-

mes-Howell Test ergaben sich signifikante Unterschiede ($p < 0,05$) zwischen den Grundschulern und jeweils den 3 anderen Hörergruppen. Gymnasiasten, Studierende und Phoniater unterschieden sich untereinander nicht signifikant.

Die korrekten Altersschätzungen sollen im Folgenden auch in Abhängigkeit von der Altersgruppe der *Sprecher* dargelegt werden (Tab. 4, Abb. 7). Alle 4 Zuhörergruppen erzielten bei den Kindern/Jugendlichen und jungen Erwachsenen (5-34 Jahre) die besten Ergebnisse, mit steigendem Alter der Sprecher wurden schlechtere Ergebnisse erzielt. Der Kolmogorov-Smirnov-Test ergab keine Normalverteilung ($p < 0,001$), im daraufhin durchgeführten Kruskal-Wallis-Test war dieser Unterschied signifikant (Kruskal-Wallis $H(2) = 319,80$, $p < 0,001$), im Post-hoc-Test zwischen allen 3 Sprechergruppen.

Tab. 4: Korrekte Alterszuordnungen (Prozent) sortiert nach Altersgruppen der Sprecher (Jahre) und Probandengruppen

Altersgruppe Sprecher (Jahre)	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
5-34	54,24	40,48	58,80	63,71	50,00
35-64	29,63	22,16	30,70	32,55	34,24
65-94	22,23	20,83	26,38	22,40	14,67

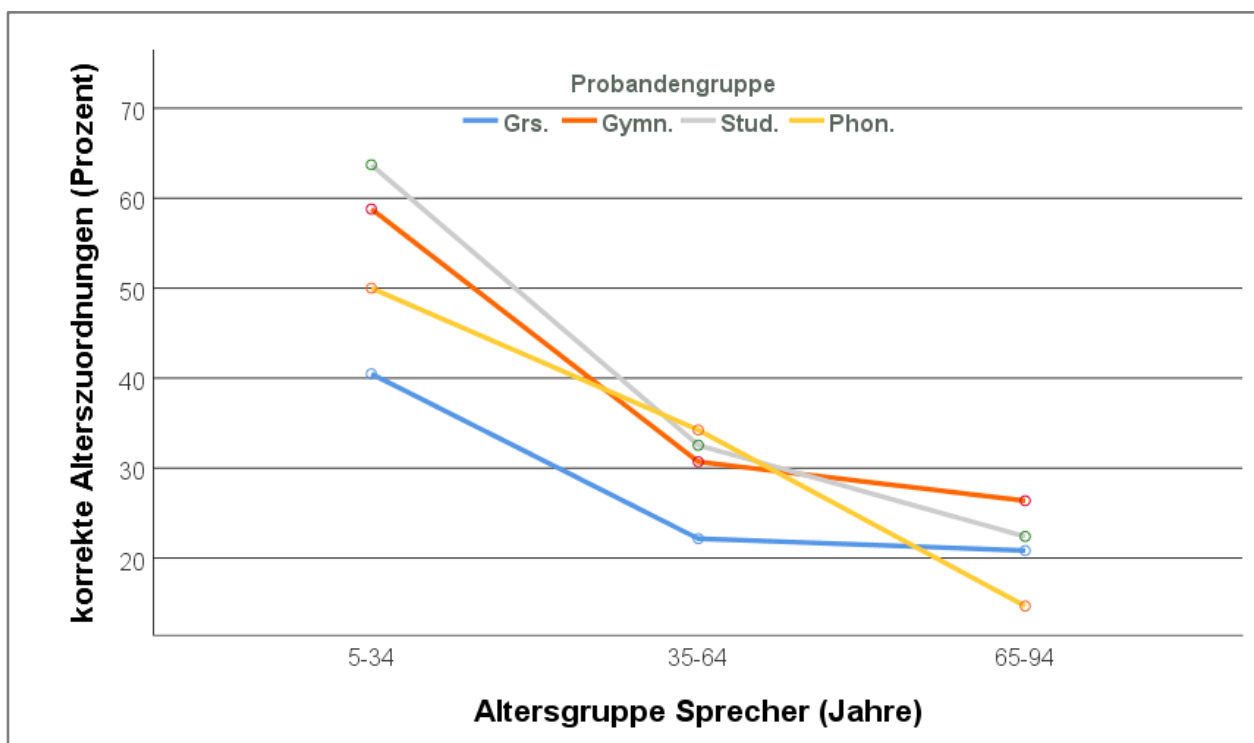


Abb. 7: Liniendiagramm der korrekten Alterszuordnungen (Prozent) sortiert nach Altersgruppen der Sprecher (Jahre) und Probandengruppen

3.4.2 Körpergröße

Die Frage nach der Körpergröße der Sprecher beinhaltete 7 Antwortmöglichkeiten, die jeweils in Zehnerschritte unterteilt, einen Wertebereich von 131 bis 200 cm umfassten (s. Frage 5 des Fragebogens). Der Spearman-Korrelationskoeffizient ergab mit $r = 0,617$ ($p < 0,001$) eine stark positive Korrelation zwischen der vermuteten und der tatsächlichen Größenkategorie.

Den Probanden gelang es in 1775 von 4531 Fällen (39,17 %), die Sprecher ihrer korrekten Größenkategorie zuzuordnen. Bei einem erlaubten Fehler von ± 1 Größenkategorie waren ihre Antworten in 3799 von 4531 Fällen (83,84 %) richtig. In Tab. 5 sind die entsprechenden Ergebnisse für die 4 Probandengruppen dargestellt. Im Boxplot (Abb. 8) fällt die höhere Streuung der Werte bei den Grundschulern insbesondere an der großen Spannweite mit einem Minimum von 0 % und einem Maximum von 56,52 % auf. 6 der 48 Grundschüler ordneten jeweils lediglich 1-2 der 23 Sprecher ihrer korrekten Größe zu (entspricht 4,35 % bzw. 8,70 %), ein Grundschüler wählte bei keinem Sprecher die korrekte Antwortmöglichkeit aus.

Tab. 5: Korrekte Größenzuordnungen sowie Größenzuordnungen bei einem Fehler von ± 1 und ± 2 Größenkategorie(n) (Prozent) sortiert nach den Probandengruppen

Größenzuordnung	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
korrekt	39,17	25,91	41,90	44,70	44,93
± 1 Größenkategorie	83,84	65,76	88,41	90,17	91,74
± 2 Größenkategorien	96,05	87,68	98,42	98,87	99,28

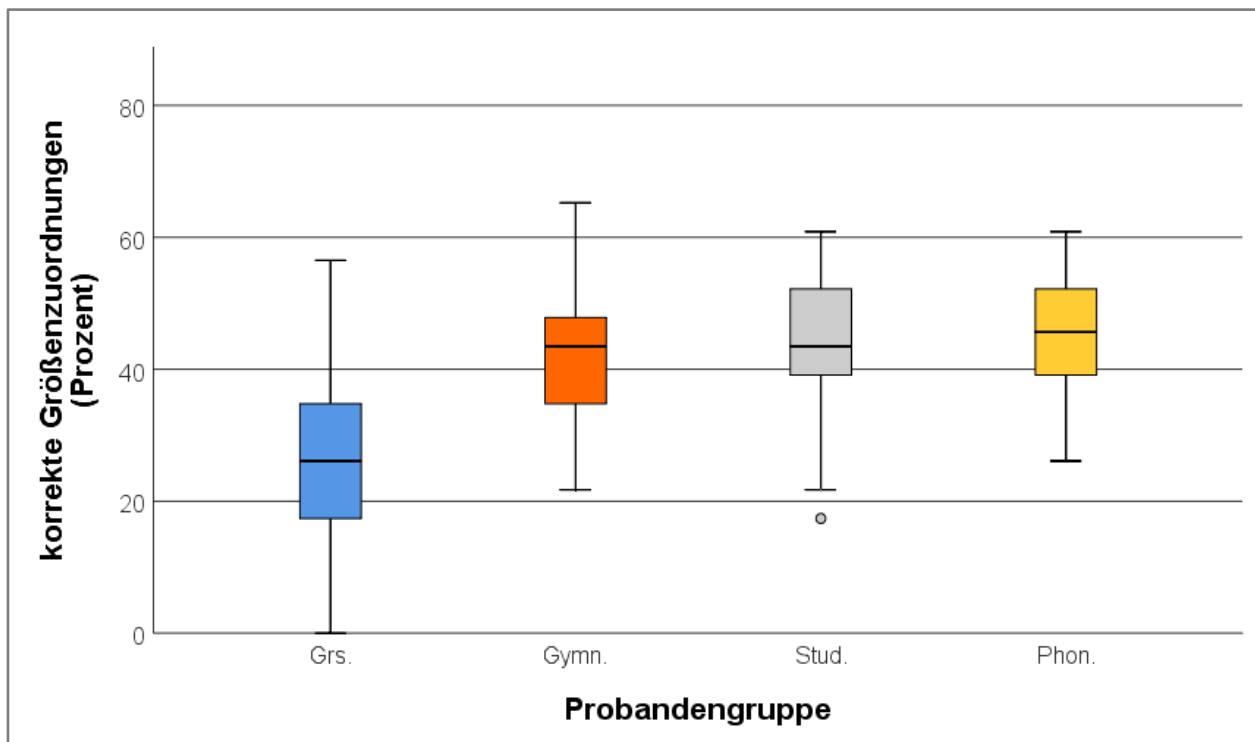


Abb. 8: Boxplot der korrekten Größenzuordnungen (Prozent) sortiert nach den Probandengruppen

In der Welch-ANOVA unterschieden sich die Ergebnisse der Alterszuordnungen signifikant (Welch-Test $F(3, 2159,80) = 42,63, p < 0,001$), nach Games-Howell Test zwischen den Grundschulern und jeweils den 3 anderen Zuhörergruppen ($p < 0,001$).

Die Sprecher lassen sich hinsichtlich ihres Alters in die Gruppen Kinder (Altersgruppe 5-14 Jahre) und Jugendliche/Erwachsene (15-94 Jahre) unterteilen. Bei beiden Gruppen waren die Zuhörer in ihren Größenschätzungen ähnlich erfolgreich mit 37,39 % bzw. 39,44 % korrekten Angaben (s. Tab. 6). Auch den Grundschulern gelang die grobe Einordnung der Größen der kindlichen und erwachsenen Sprecher.

Tab. 6: Korrekte Größenzuordnungen (Prozent) sortiert nach dem Alter der Sprecher (Jahre) und den Probandengruppen

Alter Sprecher (Jahre)	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
Kind (5-14)	37,39	34,72	42,03	36,00	33,33
Erwachsener (15-94)	39,44	24,58	41,88	46,00	46,67

Abb. 9 zeigt, dass innerhalb der Gruppe der erwachsenen Sprecher die besten Ergebnisse bei den mittleren Größenkategorien erzielt wurden: Bei den Frauen (Körpergrößen

zwischen 151 und 180 cm) in der Klasse 161-170 cm, bei den männlichen Sprechern (171-200 cm) in der Klasse 171-180 cm. Dementgegen wurde etwa der Sprecher mit der höchsten Körpergröße (191-200 cm) in lediglich 8 von 197 Fällen (4,06 %) richtig zugeordnet.

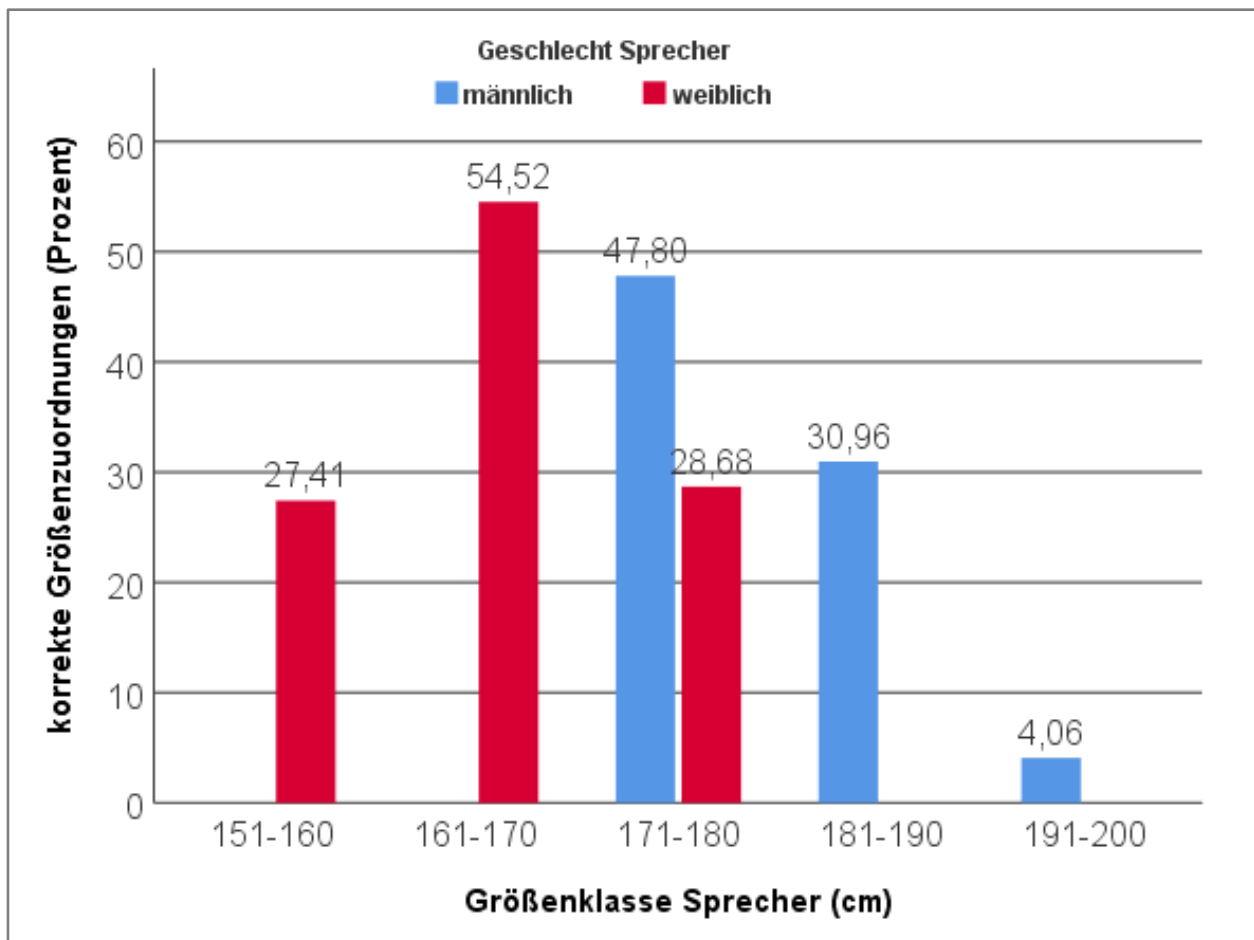


Abb. 9: korrekte Größenzuordnungen (Prozent) sortiert nach Größe (cm) und Geschlecht der Sprecher

Zur Darstellung der Abweichungen der Größenschätzungen wurde für jeden Probanden, analog zur Berechnung bei der Variable Alter (s. 3.4.1), die Differenz von geschätzter und tatsächlicher Größenkategorie ermittelt; die Mittelwerte und Standardabweichungen pro Zuhörergruppe sind in Tab. 7 aufgeführt. Die erwachsenen Sprecher wurden mit zunehmender Körpergröße (181-200 cm Körpergröße) stärker unterschätzt, d.h. für kleiner gehalten, wie am negativen Vorzeichen sowie der Stärke der Abweichung von 0 ersichtlich, während bei den kleineren Sprechern (151-160 cm Körpergröße) eine Überschätzung vorherrschte (positives Vorzeichen). Die geringsten Abweichungen von der

tatsächlichen Größe bestanden, Abb. 9 entsprechend, bei den Sprechern mit den mittleren Körpergrößen (161-180 cm). In Bezug auf die kindlichen Vorleser gelangen lediglich bei dem Sprecher mit der Körpergröße 141-150 cm präzise Größenschätzungen, da hier die Abweichung mit $-0,08$ nahe 0 liegt.

Tab. 7: Mittelwerte \pm Standardabweichungen der Differenzen geschätzte Größe - tatsächliche Größe sortiert nach Größenklassen der Sprecher (cm) und Probandengruppen

Größenklasse Sprecher (cm)	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
Erwachsene					
151-160	0,81\pm0,87	0,72 \pm 1,32	0,92 \pm 0,70	0,79 \pm 0,59	0,71 \pm 0,65
161-170	-0,08\pm0,84	-0,41 \pm 1,21	0,06 \pm 0,69	-0,01 \pm 0,59	-0,01 \pm 0,63
171-180	-0,27\pm1,04	-1,01 \pm 1,32	0,04 \pm 0,81	-0,08 \pm 0,79	-0,14 \pm 0,80
181-190	-0,94\pm1,13	-2,08 \pm 1,32	-0,55 \pm 0,85	-0,67 \pm 0,60	-0,47 \pm 0,73
191-200	-1,90\pm1,07	-2,74 \pm 1,44	-1,60 \pm 0,85	-1,76 \pm 0,69	-1,47 \pm 0,57
Kinder					
131-140	0,59\pm0,90	0,71 \pm 1,22	0,46 \pm 0,65	0,60 \pm 0,83	0,66 \pm 0,86
141-150	-0,08\pm0,96	-0,19 \pm 1,08	-0,01 \pm 0,90	-0,10 \pm 0,91	-0,03 \pm 1,00
151-160	-1,11\pm0,92	-1,30 \pm 0,72	-1,10 \pm 0,86	-1,18 \pm 0,96	-0,70 \pm 1,15

3.4.3 BMI

Die Probanden sollten den Körperbau der Sprecher abschätzen und schematischen Zeichnungen zuordnen (s. Frage 4 des Fragebogens). Die 5 Antwortmöglichkeiten umfassten Untergewicht (Abb. 1), Normalgewicht (Abb. 2), Übergewicht (Abb. 3), Adipositas (Abb. 4, entspricht Adipositas Grad I nach WHO) sowie starke Adipositas (Abb. 5, entspricht Adipositas Grad II und III nach WHO, s. 2.2.1). Für die BMI-Einteilung anhand von Perzentilen bei den kindlichen Sprechern s. 2.2.1. Der Spearman-Korrelationskoeffizient zeigte mit $r = 0,202$ ($p < 0,001$) eine schwach positive Korrelation zwischen der vermuteten und der tatsächlichen BMI-Klasse.

Die Probanden gaben in 1255 von 4531 Fällen (27,70 %) die BMI-Klassen der Sprecher korrekt an, bei einem erlaubten Fehler von ± 1 BMI-Klasse in 3182 von 4531 Fällen (70,23 %). Tab. 8 zeigt die Ergebnisse der 4 Zuhörergruppen. Im Boxplot (Abb. 10) fällt die höhere Streuung der Werte bei den Grundschulern anhand des größeren Interquartilsabstandes entsprechend der Länge der Box auf.

Tab. 8: Korrekte BMI-Zuordnungen sowie BMI-Zuordnungen bei einem Fehler von ± 1 und ± 2 BMI-Klasse(n) (Prozent) sortiert nach den Probandengruppen

BMI-Zuordnung	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
korrekt	27,70	25,09	28,67	29,13	27,25
± 1 BMI-Klasse	70,23	66,85	71,52	72,78	68,41
± 2 BMI-Klassen	90,42	86,96	90,42	93,13	91,45

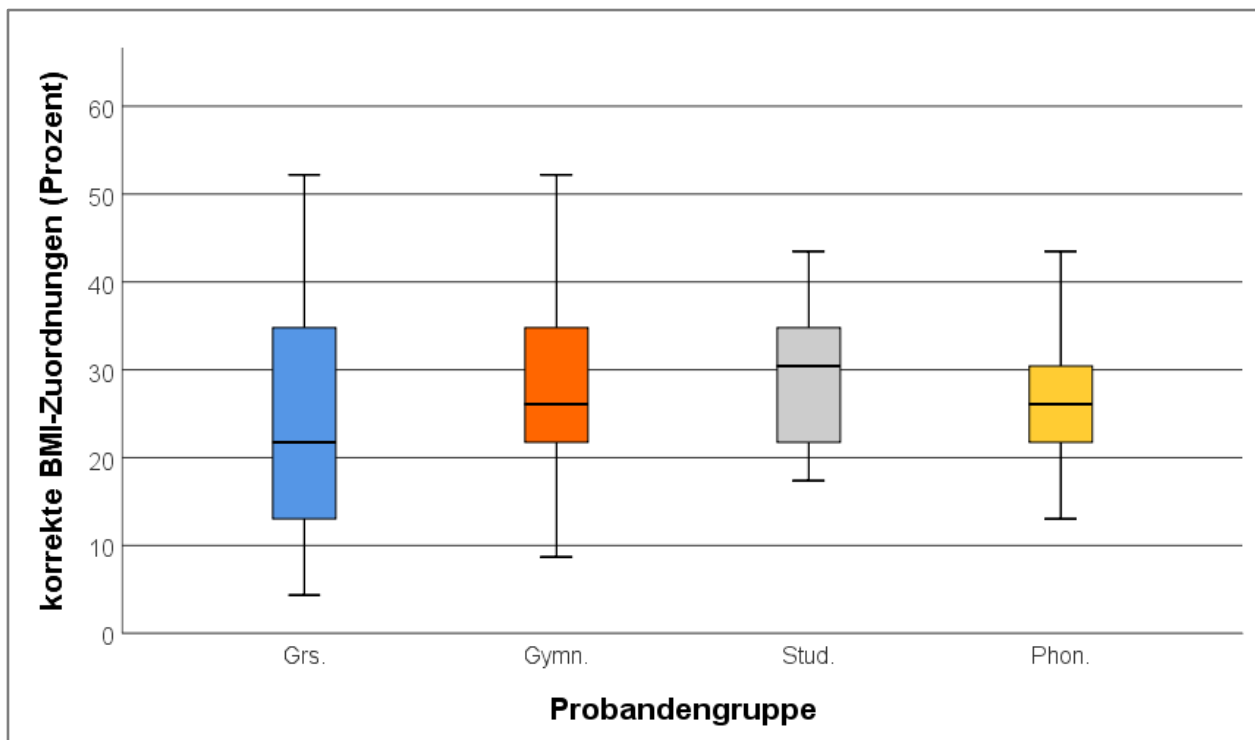


Abb. 10: Boxplot der korrekten BMI-Zuordnungen (Prozent) sortiert nach den Probandengruppen

In der Welch-ANOVA wie im Post-Hoc-Test unterschieden sich die korrekten BMI-Zuordnungen der 4 Zuhörergruppen nicht signifikant (Welch-Test $F(3, 2167,79) = 1,98$, $p = 0,116$).

Beim Vergleich der Ergebnisse in den einzelnen BMI-Klassen (Abb. 11) kann konstatiert werden, dass die meisten korrekten BMI-Schätzungen mit 35,62 % bzw. 43,65 % bei den Sprechern mit Normal- und Übergewicht vorlagen. Die Sprecher mit dem niedrigsten BMI (Untergewicht) und dem höchsten BMI (starke Adipositas) wurden dagegen am stärksten über- bzw. unterschätzt, wie in Tab. 9 anhand des Vorzeichens sowie der Stärke der Abweichung von 0 zu erkennen ist. Es wurde, wie bei den Variablen Alter

und Körpergröße, die Differenz von geschätzter und tatsächlicher BMI-Klasse berechnet (s. 3.4.1).

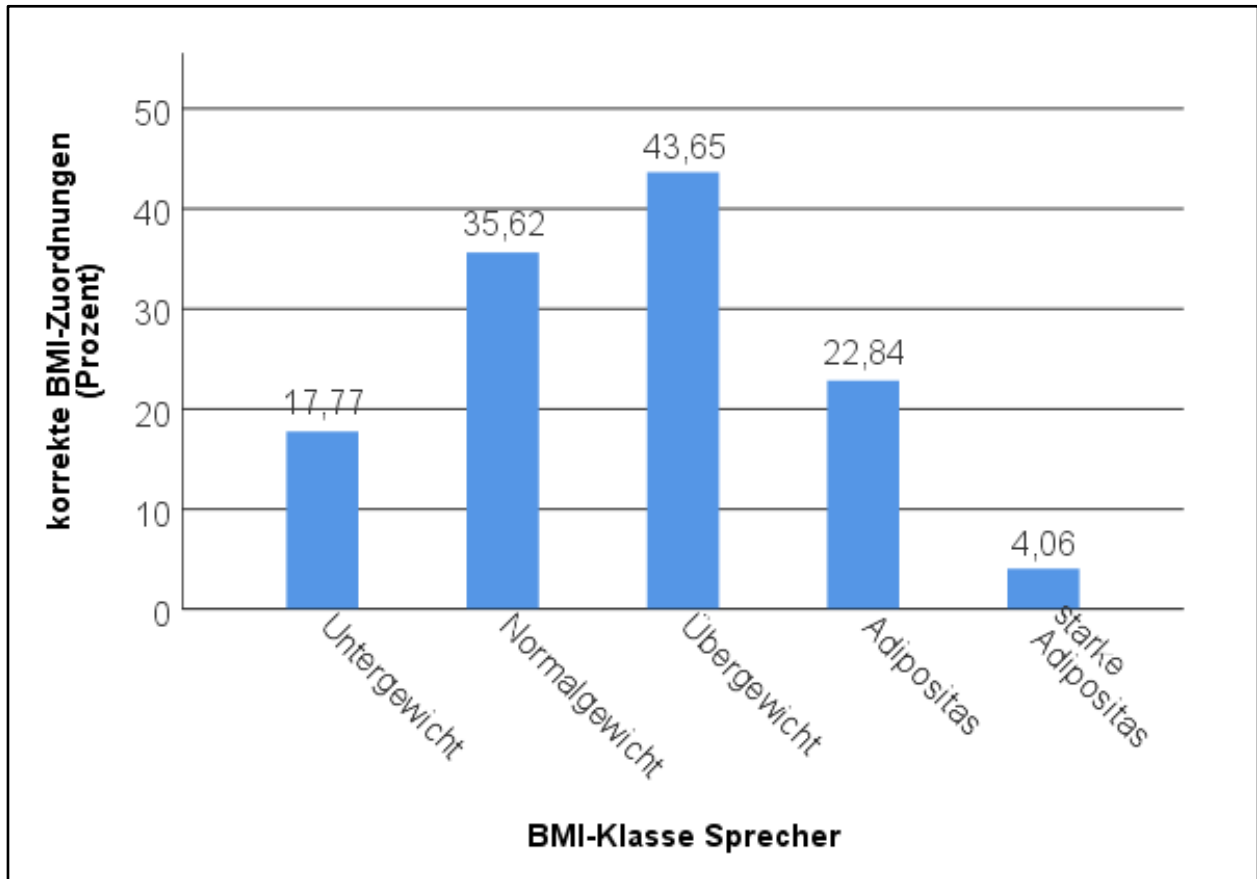


Abb. 11: Korrekte BMI-Zuordnungen (Prozent) sortiert nach BMI-Klasse der Sprecher

Tab. 9: Mittelwerte \pm Standardabweichungen der Differenzen geschätzter BMI - tatsächlicher BMI sortiert nach BMI-Klassen der Sprecher und Probandengruppen

BMI-Klasse Sprecher	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
Untergewicht	1,00\pm0,63	1,26 \pm 0,64	0,88 \pm 0,64	0,88 \pm 0,52	1,08 \pm 0,62
Normalgewicht	0,68\pm0,91	0,84 \pm 0,95	0,63 \pm 0,89	0,55 \pm 0,85	0,77 \pm 0,91
Übergewicht	-0,19\pm0,87	0,27 \pm 0,93	-0,31 \pm 0,79	-0,47 \pm 0,83	-0,18 \pm 0,71
Adipositas	-1,04\pm0,96	-0,89 \pm 0,89	-0,94 \pm 1,09	-1,24 \pm 0,85	-1,17 \pm 0,91
starke Adipositas	-2,13\pm0,98	-2,13 \pm 1,01	-2,18 \pm 0,99	-2,09 \pm 0,90	-2,08 \pm 1,02

3.4.4 Geschlecht

Die Probanden identifizierten in 4301 von 4531 Fällen (94,92 %) das Geschlecht der Sprecher korrekt. In Bezug auf die 4 Probandengruppen nahm die Prozentzahl der kor-

rekten Zuordnungen von Grundschulern (93,39 %) zu Gymnasiasten (95,15 %) zu, das höchste Ergebnis erreichten die Studierenden mit 96,61 %. Die Phoniater gaben in 94,06 % das Geschlecht richtig an. Im Boxplot (Abb. 12) ist erkennbar, dass sich die korrekten Angaben im oberen Bereich nahe 100 % befinden und die Daten, wie an den geringen Interquartilsabständen und Spannweiten zu erkennen, wenig streuen.

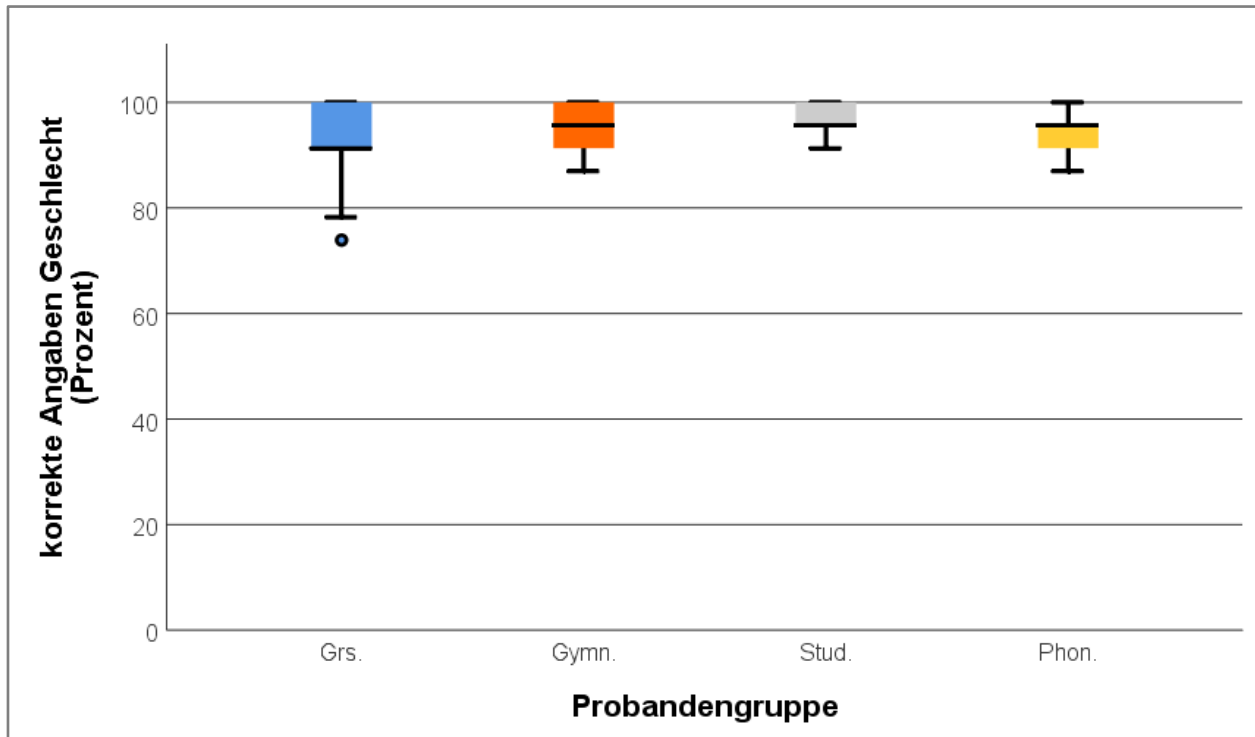


Abb. 12: Boxplot der korrekten Angaben des Geschlechts (Prozent) sortiert nach den Probandengruppen

In der Welch-ANOVA unterschieden sich die genannten Ergebnisse der 4 Gruppen signifikant (Welch-Test $F(3, 2117,74) = 4,81$, $p = 0,002$), nach Games-Howell Test zwischen den Grundschulern und Studierenden.

Im Folgenden sollen die korrekten Identifikationen bei den Sprechern im Kindes- und Erwachsenenalter verglichen werden. Während das Geschlecht der erwachsenen Sprecher in 98,30 % richtig identifiziert wurde, war dies bei den Kindern in 72,42 % der Fall. In Tab. 10 ist die Ursache dieser Differenz dargestellt: Die Jungen wurden von den Probanden insgesamt in lediglich 59,90 % richtig als männlich identifiziert, d.h. sie wurden in ca. 40 % der Fälle für weiblich gehalten. Das Mädchen wurde hingegen mit 97,46 % korrekten Angaben sicher identifiziert. Dieser Unterschied ist bei nicht gegebener Normalverteilung im Kolmogorov-Smirnov-Test ($p < 0,001$) gemäß des Mann-Whitney-U-

Tests signifikant, $U = 24231,00$, $p < 0,001$. Bei den erwachsenen Sprechern wurden die Männer geringgradig, aber statistisch signifikant häufiger korrekt identifiziert als die Frauen (Kolmogorov-Smirnov-Test: $p < 0,001$; $U = 1845496,00$, $p = 0,027$).

Tab. 10: Korrekte Angaben des Geschlechts (Prozent) sortiert nach Alter (Jahre) und Geschlecht der Sprecher sowie den Probandengruppen

Alter Sprecher (Jahre)	Geschlecht Sprecher	Probandengruppe				
		Insgesamt	Grs.	Gymn.	Stud.	Phon.
Kind (5-14)	männlich	59,90	57,29	63,04	65,00	48,33
	weiblich	97,46	93,75	98,55	100,00	96,67
Erwachsener (15-94)	männlich	98,86	97,40	98,91	99,75	99,58
	weiblich	97,93	96,70	97,71	99,50	97,78

3.4.5 Akzent/Dialekt/Hochdeutsch

Frage 3 des Fragebogens umfasste die Antwortmöglichkeiten „Hochdeutsch“, „ausländischer Akzent“ und „Dialekt“ (zu den Sprechern mit ausländischem Akzent s. 2.3.1). Es sollte bei Vermutung eines Akzents oder Dialekts zusätzlich angekreuzt werden, *welcher* Akzent oder Dialekt genau angenommen wurde. Es gab keinen Sprecher, der sowohl mit Akzent als auch Dialekt sprach, somit war auch bei dieser Frage immer nur eine Antwortmöglichkeit richtig.

Insgesamt antworteten die Probanden in 2952 von 4531 Fällen (65,15 %) richtig, d.h. gaben den jeweiligen Akzent, Dialekt oder das Hochdeutsch der Sprecher korrekt an. Die Grundschüler wiesen in 42,03 % korrekte Angaben auf, die Gymnasiasten in 66,48 % und die Studierenden in 76,17 %. Die Phoniater erzielten mit 80,72 % die meisten richtigen Zuordnungen. Im Boxplot (Abb. 13) sind anhand des Interquartilsabstandes und der Spannweite eine höhere Streuung der Werte bei den Grundschülern sowie eine niedrigere Streuung bei den Phoniatern zu erkennen.

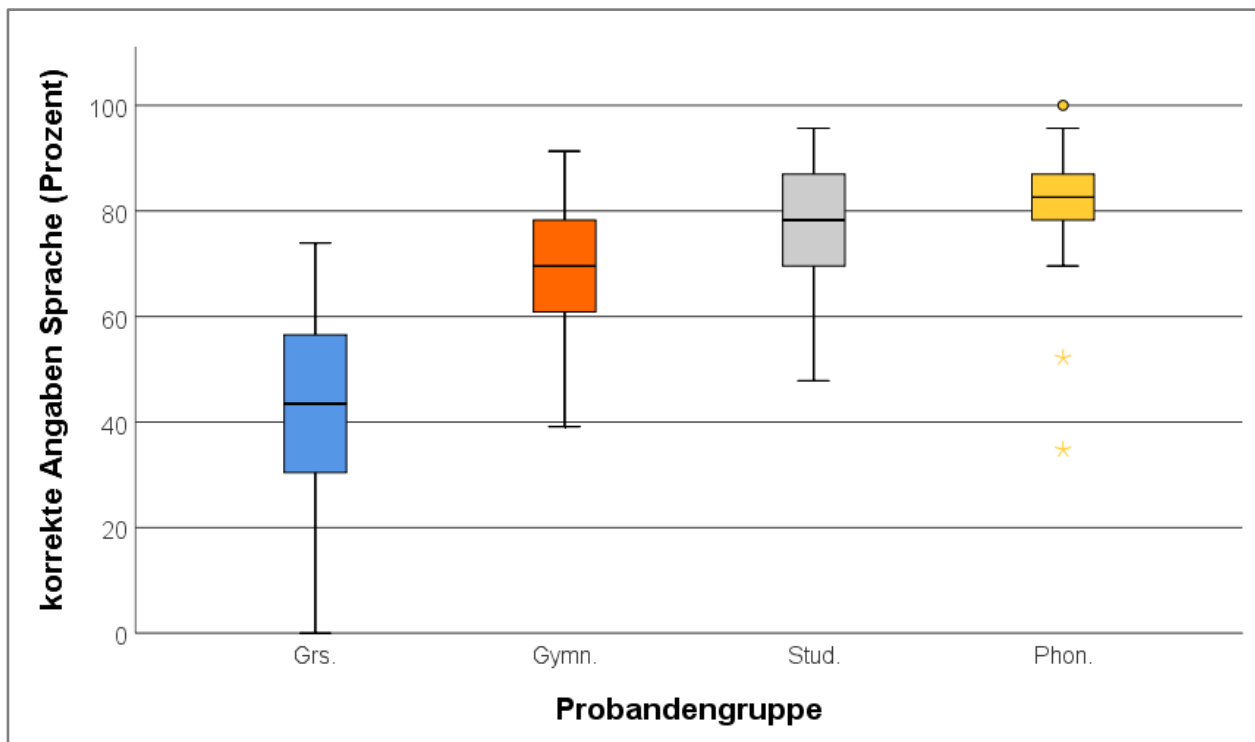


Abb. 13: Boxplot der korrekten Angaben von Hochdeutsch bzw. dem jeweiligen Akzent/Dialekt (Prozent) sortiert nach den Probandengruppen

In der Welch-ANOVA ergab sich ein signifikanter Unterschied (Welch-Test $F(3, 2219,22) = 139,63$, $p < 0,001$), nach Games-Howell Test zwischen allen Zuhörergruppen außer zwischen Studierenden und Phoniatern.

Weiterhin sollen die Ergebnisse in den einzelnen Antwortkategorien aufgezeigt werden. Sprecher mit einem Akzent wurden von den Grundschulern in 63,19 % als solche erkannt und von den restlichen Zuhörergruppen in über 95 % der Fälle (Gymnasiasten: 98,07 %, Studierende: 98,00 %, Phoniater: 96,67 %). Bei Sprechern mit Dialekt wurde von den Grundschulern in 43,52 % richtigerweise ein Dialekt angekreuzt, von den restlichen Hörergruppen in über 80 % der Fälle (Gymnasiasten: 83,25 %, Studierende: 84,44 %, Phoniater: 93,33 %). Die korrekten Angaben der Zuhörergruppen in den Kategorien Hochdeutsch, Akzent und Dialekt sind in Tab. 11 und Abb. 14 dargestellt. Bei den Sprechern mit Dialekt im Vergleich zu denjenigen mit Akzent oder Hochdeutsch differieren die Ergebnisse zwischen den Gruppen stärker und fallen insgesamt schlechter aus (46,93 % korrekte Angaben).

Tab. 11: Korrekte Angaben von Hochdeutsch bzw. dem jeweiligen Akzent/Dialekt (Prozent) sortiert nach den Probandengruppen

Stimmaufnahmen	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
Hochdeutsch	76,14	65,53	79,18	83,09	74,55
Akzent	79,53	34,72	94,20	93,33	94,44
Dialekt	46,93	15,74	41,71	62,00	83,70

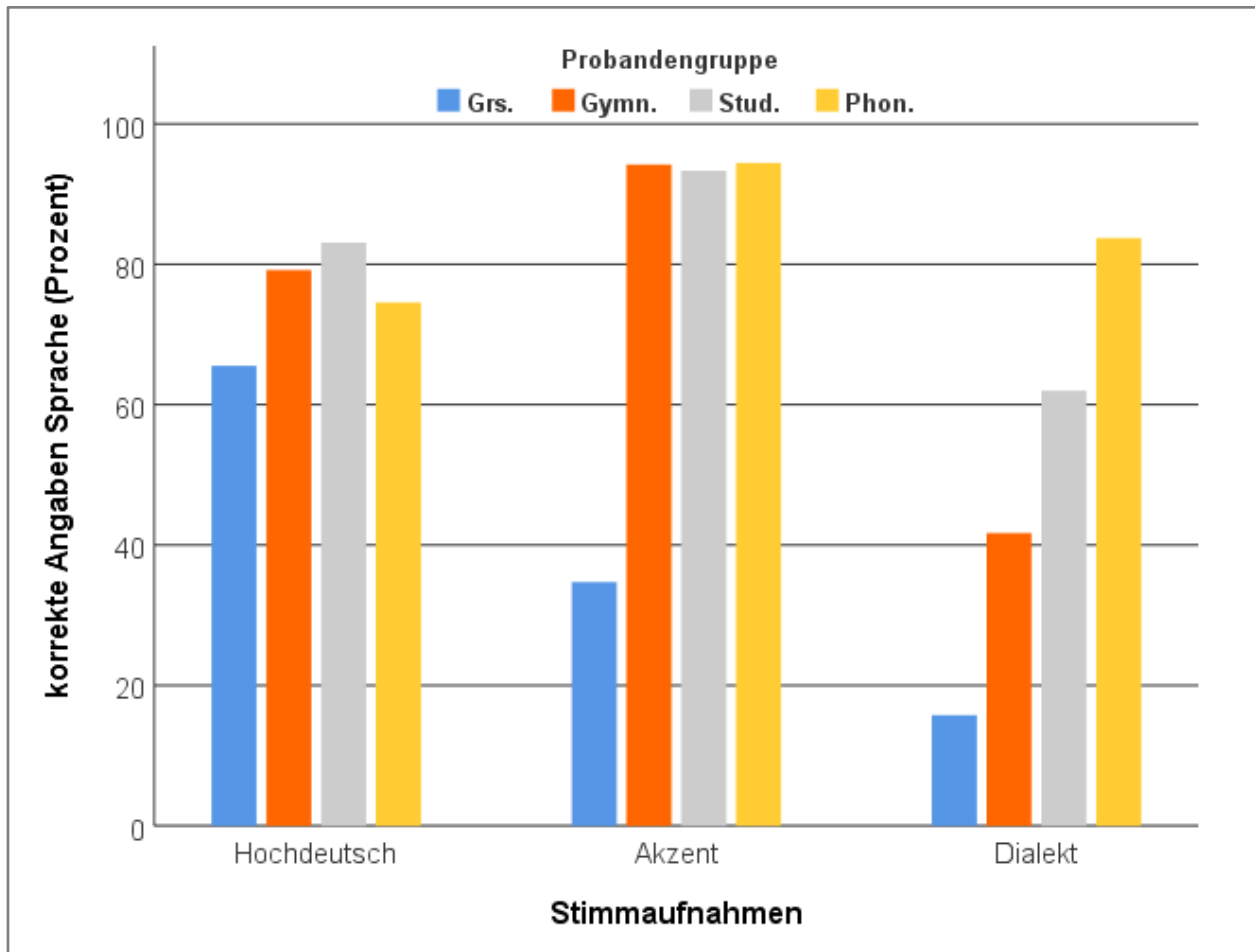


Abb. 14: Korrekte Angaben von Hochdeutsch bzw. dem jeweiligen Akzent/Dialekt (Prozent) sortiert nach den Probandengruppen

Tab. 12 zeigt die korrekten Zuordnungen, die bei den einzelnen Akzenten und Dialekten erzielt wurden. Während bei den 3 ausländischen Akzenten die Ergebnisse ähnlich gut ausfielen, waren die Dialekte offenbar unterschiedlich leicht bzw. schwierig zuzuordnen. Der schwäbische Dialekt konnte insgesamt nur in ca. einem Drittel der Fälle korrekt identifiziert werden, Bayerisch, Sächsisch und Kölsch dagegen in mehr als der Hälfte der Fälle. Im Vergleich der Probandengruppen hatten v.a. die Grundschüler Schwierig-

keiten, die sächsischen, schwäbischen und norddeutschen Stimmufnahmen richtig zuzuordnen, dies gelang ihnen in jeweils weniger als 10 % der Fälle.

Tab. 12: Korrekte Angaben von Akzenten/Dialekten (Prozent) sortiert nach den Probandengruppen

Stimmufnahmen	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
Hochdeutsch	76,14	65,53	79,18	83,09	74,55
Französisch	77,66	33,33	88,41	92,00	100,00
Englisch/Amerikanisch	79,70	31,25	97,10	96,00	90,00
Türkisch/Arabisch	81,22	39,58	97,10	92,00	93,33
Bayerisch	51,27	25,00	44,93	63,00	88,33
Sächsisch	54,82	8,33	55,07	74,00	96,67
Schwäbisch	33,76	5,21	22,46	50,00	78,33
Kölsch	50,76	22,22	46,38	66,00	81,11
Norddeutsch	45,18	6,25	46,38	60,00	80,00

Zwecks Darstellung der fehlenden und ungültigen Werte ist in Tab. 13 aufgeführt, wie sich die Antworten der Zuhörergruppen zusammensetzen. Die Antworten der Grundschüler bei der Frage nach Akzent/Dialekt/Hochdeutsch waren in 132 von 1104 Fällen und somit in rund 12 % als ungültig zu werten; meist aufgrund des Ankreuzens von mehr als einer Antwortmöglichkeit (z.B. wurde sowohl die Kategorie Akzent als auch einer der aufgeführten Dialekte angekreuzt). Die 3 Grundschüler, die keinen der 23 Sprecher korrekt zuordneten, wiesen in mindestens 12 von 23 Fällen und somit in mehr als der Hälfte der Fälle ungültige Antworten auf. Fehlende Werte lagen bei allen 4 Zuhörergruppen in weniger als 1 % der Fälle vor.

Tab. 13: Zusammensetzung der Angaben zu Akzent/Dialekt/Hochdeutsch (Prozent), sortiert nach den Probandengruppen

Angabe	Probandengruppe				
	Insgesamt	Grs.	Gymn.	Stud.	Phon.
korrekt	65,15	42,03	66,48	76,17	80,72
nicht korrekt	30,59	45,47	31,51	22,87	17,54
ungültig	3,75	11,96	1,39	0,61	1,30
fehlend	0,51	0,54	0,63	0,35	0,43
Gesamt	100,00	100,00	100,00	100,00	100,00

3.4.6 Rauchverhalten

Alle kindlichen Sprecher wurden von den Probanden in über 95 % der Fälle richtigerweise als Nichtraucher identifiziert. Im Folgenden werden somit die Einschätzungen, die die Probanden bei den erwachsenen Sprechern trafen, dargestellt. Sprecher, die zum Zeitpunkt der Durchführung der Studie angaben, 10 oder mehr Zigaretten am Tag, entsprechend mindestens einer halben Zigarettenpackung täglich, zu rauchen, wurden als Raucher eingestuft. Zusätzlich wurde eine Frau, die zwar weniger als 10 Zigaretten am Tag rauchte, dies jedoch schon seit ca. 45 Jahren, ebenfalls zu den Rauchern gezählt. Die restlichen Probanden waren Nichtraucher. Insg. sind somit 5 der 20 erwachsenen Sprecher, darunter 3 Frauen und 2 Männer, Raucher, ansonsten handelt es sich um Nichtraucher.

Bezüglich ihrer Antwortauswahl kreuzten die Probanden in fast drei Viertel der Fälle (71,73 %, 2826/3940 Fälle) und somit am weitaus häufigsten die Antwort „Nichtraucher“ an. In nur ca. einem Viertel der Fälle (26,19 %, 1032/3940 Fälle) wurde die Antwortmöglichkeit „Raucher“ ausgewählt. Dies ist auch ersichtlich, wenn man die Angaben der Probanden für die nichtrauchenden und rauchenden Sprecher vergleicht (Tab. 14). Die tatsächlichen Raucher wurden ebenso wie die Nichtraucher in je über 70 % der Fälle für Nichtraucher gehalten (s. grün hinterlegte Zeilen); entsprechend wurden sie, unabhängig von ihrem tatsächlichen Rauchverhalten, in ca. 25 % der Fälle für Raucher gehalten (s. gelb hinterlegte Zeilen). Die Grundschüler identifizierten mit 34,58 % mehr Raucher korrekt als die anderen Zuhörergruppen, allerdings hielten sie ebenso die Nichtraucher häufiger für Raucher (32,22 %). Der Chi-Quadrat-Test ergab keinen signifikanten Zusammenhang zwischen dem tatsächlichen und vermuteten Rauchverhalten ($\chi^2(1) = 0,93$, $p = 0,334$). Es lässt sich konstatieren, dass es den Probanden offenbar nicht gelang, das Rauchverhalten der Sprecher korrekt zu identifizieren.

Tab. 14: Angaben der Probanden über das Rauchverhalten der Sprecher (Prozent) in Abhängigkeit von ihrem tatsächlichen Rauchverhalten, sortiert nach den Probandengruppen

Rauchverhalten Sprecher	Angabe der Probanden	Probandengruppe				
		Insgesamt	Grs.	Gymn.	Stud.	Phon.
Nichtraucher	Nichtraucher	71,07	62,78	74,30	73,07	73,56
	Raucher	26,53	32,22	24,06	25,07	25,56
Raucher	Nichtraucher	73,60	62,92	82,32	73,20	71,33
	Raucher	25,28	34,58	17,10	26,00	28,00

Es sollen auch die Antworten der Probanden in Abhängigkeit vom Geschlecht der Sprecher wiedergegeben werden. Die Männer wurden in 34,77 % (548/1576 Fälle) für Raucher gehalten, die Frauen in 20,47 % (484/2364 Fälle). Dieser Unterschied konnte mittels des Chi-Quadrat-Tests bestätigt werden, d.h. es besteht ein signifikanter Zusammenhang zwischen Geschlecht der Sprecher und Angaben der Probanden über deren Rauchverhalten ($\chi^2(1) = 102,53$, $p < 0,001$). In Abb. 15 sind die prozentualen Häufigkeiten der Angabe Raucher nach Geschlecht der Sprecher veranschaulicht. Diese unterscheiden sich stärker bei den Grundschulern und Gymnasiasten; die Gymnasiasten schätzten die Männer im Vergleich zu den Frauen in mehr als doppelt so vielen Fällen als rauchend ein.

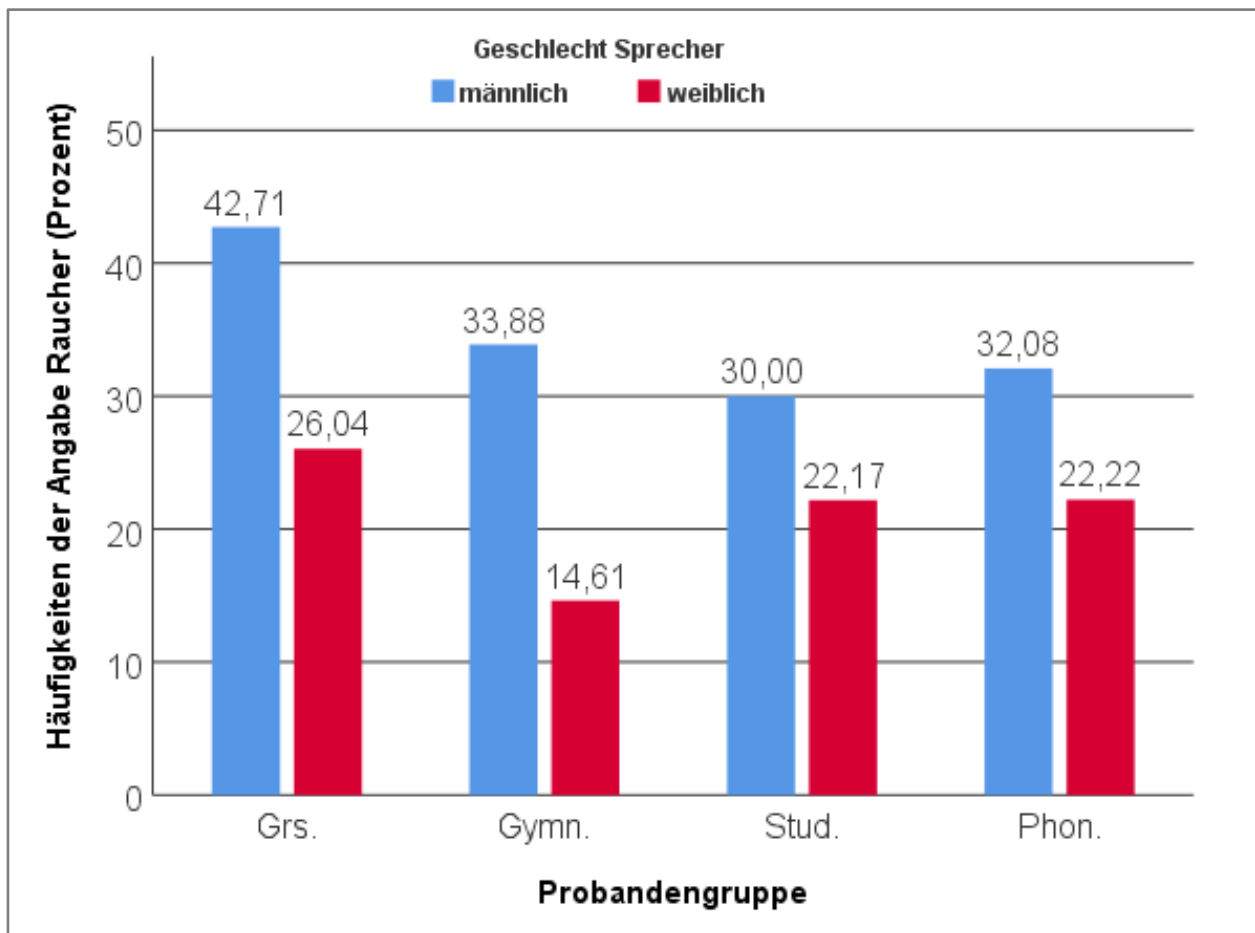


Abb. 15: Häufigkeiten der Angabe Raucher (Prozent) nach Geschlecht der Sprecher, sortiert nach den Probandengruppen

3.4.7 Schulabschluss

Bei der Frage nach dem Schulabschluss der Sprecher (s. Frage 7 des Fragebogens) wählten die Probanden bei den kindlichen Sprechern, die noch die Schule besuchen, in über 80 % der Fälle die richtige Antwort „noch SchülerIn“ aus (in 81,22 %, 480/591 Fälle). Im Einzelnen gelang dies den Grundschulern in über 50 % der Fälle (57,64 %) sowie den restlichen Zuhörergruppen in mindestens 75 % der Fälle (Gymnasiasten: 95,17 %, Studierende: 88,00 %, Phoniater: 75,56 %). Im Folgenden werden die Ergebnisse, die die Probanden bei den erwachsenen Sprechern erzielten, dargestellt.

Die Probanden insgesamt gaben in 1842 von 3940 Fällen (46,75 %) den Schulabschluss der Sprecher korrekt an. Die Grundschüler erreichten in 33,96 % der Fälle, die Gymnasiasten in 47,97 % der Fälle korrekte Zuordnungen; Studierende und Phoniater erzielten mit 53,50 % und 53,17 % die meisten richtigen Zuordnungen. Im Boxplot (Abb.

16) zeigt sich wiederum die höhere Streuung der Werte bei den Grundschulern (ein Schüler gab bei keinem erwachsenen Sprecher den korrekten Schulabschluss an).

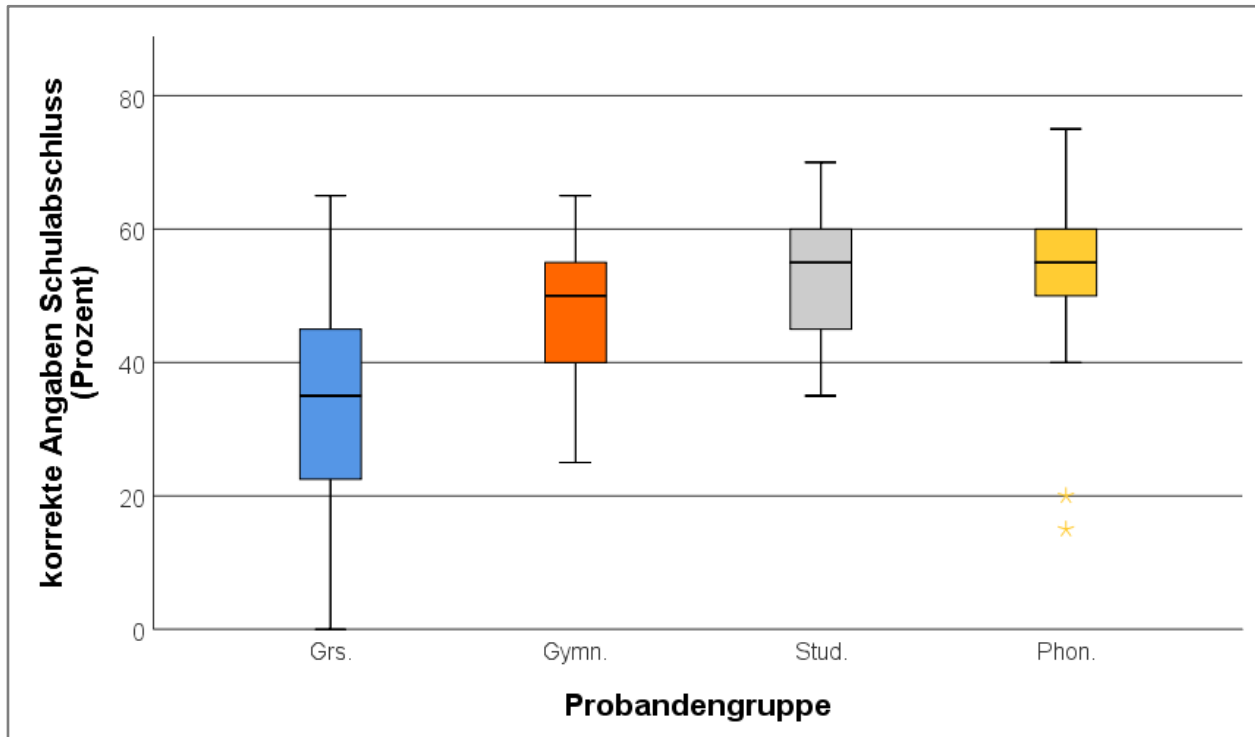


Abb. 16: Boxplot der korrekten Angaben des Schulabschlusses (Prozent) sortiert nach den Probandengruppen

In der Welch-ANOVA unterschieden sich die erreichten Ergebnisse signifikant (Welch-Test $F(3, 1880,68) = 32,92, p < 0,001$), nach Games-Howell Test zwischen den Grundschulern und jeweils den 3 weiteren Zuhörergruppen ($p < 0,001$) sowie zwischen den Gymnasiasten und den Studierenden ($p = 0,039$).

Weiterhin sollen die Angaben der Probanden zum Schulabschluss der Sprecher unter Berücksichtigung deren tatsächlichen Schulabschlusses betrachtet werden (Tab. 15). Aus Übersichtlichkeitsgründen sind die Häufigkeiten der Probandenangaben zu den Antwortmöglichkeiten „kein Schulabschluss“ und „noch SchülerIn“ nicht in der Tabelle aufgeführt; die Zuhörer wählten die beiden genannten Antworten in weniger als 5 % bzw. 2 % aller Fälle aus. Es fällt auf (s. grün hinterlegte Zeilen), dass die Probanden bei den Sprechern mit Gymnasial-/Gesamtschulabschluss diese Antwort häufiger ankreuzten (in 68,02 % der Fälle) als bei den Sprechern mit Real- bzw. Hauptschulabschluss (in 39,97 % bzw. 33,16 % der Fälle). Ebenso wurden die Sprecher mit Realschulabschluss

häufiger als solche identifiziert als die Sprecher mit Gymnasial-/Gesamtschulabschluss (s. gelb hinterlegte Zeilen). Auch bei den Sprechern mit Hauptschulabschluss zeigt sich, dass diese geringgradig häufiger als solche erkannt wurden als die Sprecher mit den beiden anderen Abschlüssen (s. blau hinterlegte Zeilen). Laut Chi-Quadrat-Test liegt ein signifikanter Zusammenhang zwischen tatsächlichem Schulabschluss der Sprecher und den Angaben der Probanden darüber vor ($\chi^2(8) = 399,99$, $p < 0,001$). Dies lässt darauf schließen, dass die Zuhörer, insbesondere Gymnasiasten, Studierende und Phoniater, zumindest teilweise den Schulabschluss der Sprecher korrekt identifizieren konnten.

Tab. 15: Angaben der Probanden zum Schulabschluss der Sprecher (Prozent) in Abhängigkeit von ihrem tatsächlichen Schulabschluss, sortiert nach den Probandengruppen

Schulabschluss Sprecher	Angabe der Probanden	Probandengruppe				
		Insgesamt	Grs.	Gymn.	Stud.	Phon.
Gymnasium/Gesamtschule	Gymnasium/Gesamtschule	68,02	53,01	73,75	74,00	68,89
	Realschule	18,84	14,58	18,84	21,11	21,85
	Hauptschule	8,52	18,06	4,99	3,78	9,26
Realschule	Gymnasium/Gesamtschule	39,97	50,26	43,48	33,50	26,25
	Realschule	34,96	19,01	33,15	44,75	48,33
	Hauptschule	14,53	14,06	12,14	14,75	20,42
Hauptschule	Gymnasium/Gesamtschule	33,16	47,22	42,03	15,33	20,00
	Realschule	46,19	19,44	44,93	66,67	57,78
	Hauptschule	14,89	18,06	10,63	15,33	18,89

3.4.8 Berufsausbildung

Analog zur Darstellung bei der Variable Schulabschluss (s. 3.4.7) sollen die Ergebnisse hinsichtlich der Frage nach der Berufsausbildung der Sprecher (s. Frage 8 des Fragebogens) aufgezeigt werden. Die Zuhörer gaben bei den Kindern unter den Sprechern in 522 von 591 Fällen (88,32 %) richtigerweise die Antwort „noch nicht berufstätig“ (Grundschüler: 75,69 %, Gymnasiasten: 98,07 %, Studierende: 92,00 %, Phoniater: 80,00 % korrekte Angaben).

Die korrekte Berufsausbildung der erwachsenen Sprecher vermuteten die Probanden insgesamt in 2327 von 3940 Fällen (59,06 %); die Grundschüler in 40,73 %, die Gymnasiasten in 59,35 %, die Studierenden in 70,20 % sowie die Phoniater in 69,17 % der Fäl-

le. Gemäß der Welch-ANOVA ergab sich ein signifikanter Unterschied (Welch-Test $F(3, 1895,70) = 73,57, p < 0,001$), nach Games-Howell Test zwischen allen Zuhörergruppen außer Studierenden und Phoniatern.

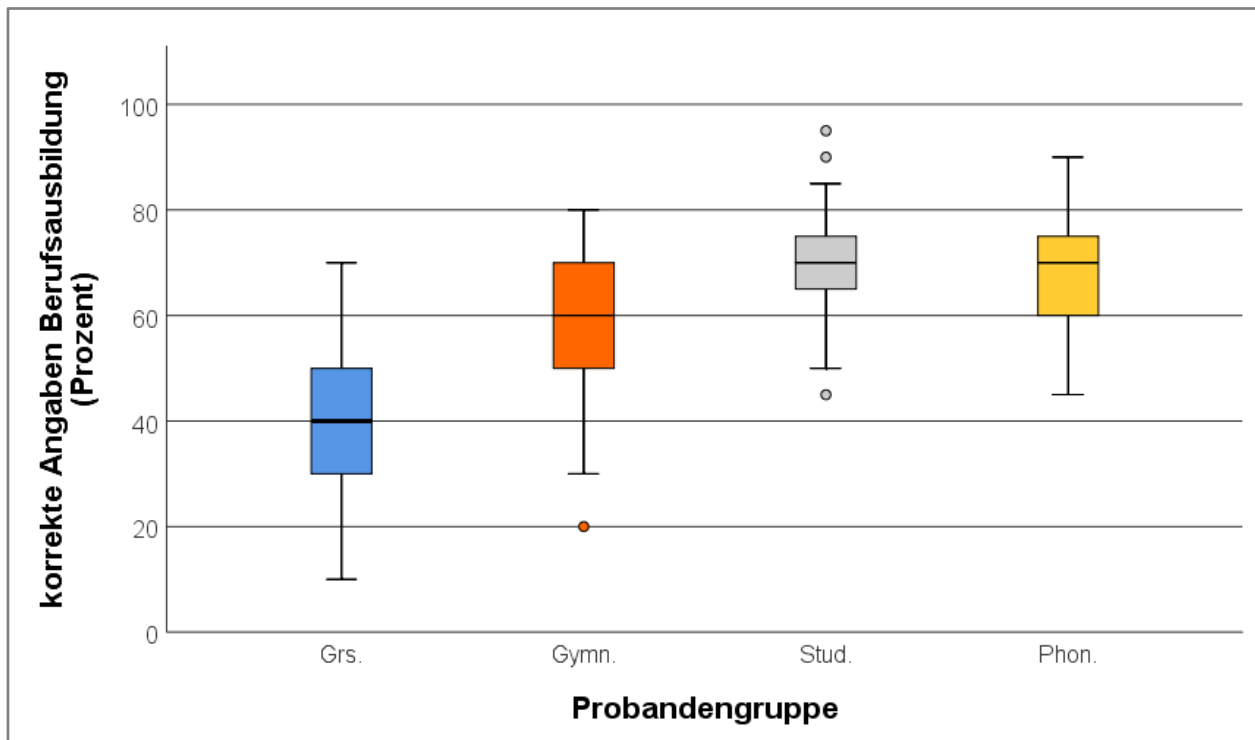


Abb. 17: Boxplot der korrekten Angaben der Berufsausbildung (Prozent) sortiert nach den Probandengruppen

Hinsichtlich der Antworten der Probanden in Abhängigkeit von der tatsächlichen Ausbildung der Sprecher (s. Tab. 16) ergeben sich ähnliche Ergebnisse wie bei der Variable Schulabschluss, die darauf hinweisen, dass die Zuhörer häufig zwischen Sprechern mit Studium und Ausbildung unterscheiden konnten: Bei den Sprechern mit einem Studium wurde fast doppelt so häufig die Antwort „Studium“ gewählt (in 67,22 %) wie bei den Sprechern mit einer Ausbildung (in 34,48 % der Fälle), s. grün hinterlegte Zeilen. Selbiges zeigte sich bei den Sprechern mit einer Ausbildung, die häufiger als solche eingeschätzt wurden als die Sprecher mit Studium (s. gelb hinterlegte Zeilen). Gemäß Chi-Quadrat-Test besteht ein signifikanter Zusammenhang zwischen Berufsausbildung der Sprecher und den Zuordnungen der Probanden ($\chi^2(2) = 388,25, p < 0,001$).

Tab. 16: Angaben der Probanden zur Berufsausbildung der Sprecher (Prozent) in Abhängigkeit von ihrer tatsächlichen Ausbildung, sortiert nach den Probandengruppen

Berufsausbildung Sprecher	Angabe der Probanden	Probandengruppe				
		Insgesamt	Grs.	Gymn.	Stud.	Phon.
Studium	Studium	67,22	50,60	72,88	72,57	71,90
	Ausbildung/Lehre	27,77	34,23	24,43	26,57	27,14
	keine Berufsausbildung	3,99	12,80	1,66	0,86	0,48
Ausbildung/Lehre	Studium	34,48	46,79	37,68	24,92	23,33
	Ausbildung/Lehre	54,71	35,74	51,95	68,92	67,69
	keine Berufsausbildung	9,49	15,87	9,25	5,38	6,67

3.4.9 Beruf

Die Erläuterung der Ergebnisse zum Beruf der Sprecher (s. Frage 9 des Fragebogens) erfolgt entsprechend der Darstellung bei Schulabschluss und Berufsausbildung (s. 3.4.7, 3.4.8). Die kindlichen Sprecher wurden von den Probanden in über 90 % der Fälle (547/591 Fälle) korrekt als „noch nicht berufstätig“ eingeordnet (Grundschüler: 78,47 %, Gymnasiasten: 100 %, Studierende: 96,00 %, Phoniater: 92,22 % korrekte Angaben). Bei den erwachsenen Sprechern gaben die Zuhörer insgesamt in 1459 von 3940 Fällen (37,03 %) die richtige Antwort an; die Grundschüler in 22,81 %, die Gymnasiasten in 37,61 %, die Studierenden in 44,00 % sowie die Phoniater in 46,83 % der Fälle. In der Welch-ANOVA zeigte sich ein signifikanter Unterschied (Welch-Test $F(3, 1871,37) = 50,08$, $p < 0,001$), nach Games-Howell Test zwischen allen Zuhörergruppen außer Studierenden und Phoniatern. Im Vergleich fällt auf, dass die Prozentzahlen der korrekten Antworten niedriger ausfallen als bei den Variablen Schulabschluss und Berufsausbildung, was möglicherweise den Schwierigkeitsgrad der Frage nach dem Beruf widerspiegelt.

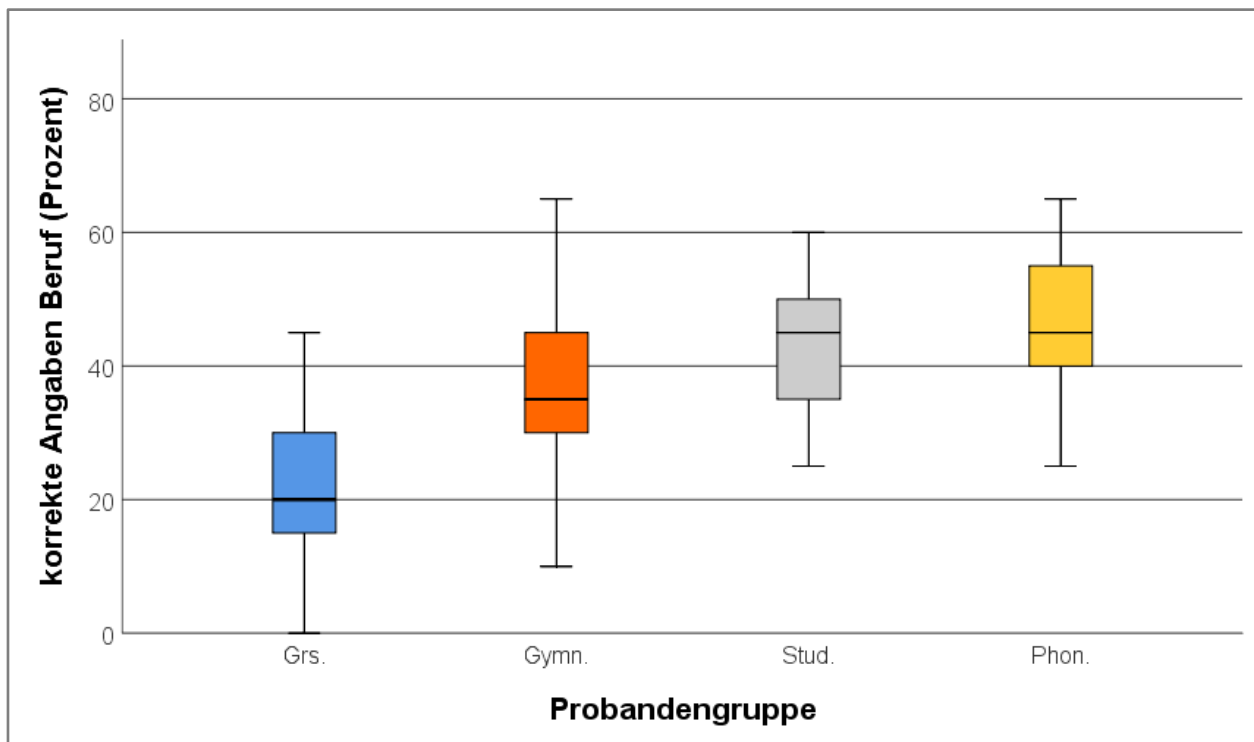


Abb. 18: Boxplot der korrekten Angaben des Berufes (Prozent) sortiert nach den Probandengruppen

In Bezug auf die Angaben der Zuhörer sortiert nach tatsächlichem Beruf der Sprecher (Tab. 17) ist festzuhalten, dass die Sprecher, die Akademiker sind, auch häufiger als solche eingeschätzt wurden (in 50,76 %) als die Sprecher der anderen beiden Berufsgruppen (in 26,14 % bzw. 15,95 % der Fälle), s. grün hinterlegte Zeilen. Dieselbe Feststellung konnte jeweils auch bei den Sprechern mit Bürotätigkeit (s. gelb hinterlegte Zeilen) und Handwerkstätigkeit (s. blau hinterlegte Zeilen) getroffen werden. Der Chi-Quadrat-Test ergab einen signifikanten Zusammenhang zwischen tatsächlichem Beruf und Angaben der Zuhörer ($\chi^2(8) = 618,03$, $p < 0,001$). In der Tabelle sind aus Übersichtlichkeitsgründen die Probandenangaben zu der Antwortmöglichkeit „noch nicht berufstätig“ nicht aufgeführt, da letztere insgesamt in lediglich 7,14 % aller Fälle von den Probanden ausgewählt wurde.

Tab. 17: Angaben der Probanden zum Beruf der Sprecher (Prozent) in Abhängigkeit von ihrem tatsächlichen Beruf, sortiert nach den Probandengruppen

Beruf Sprecher	Angabe der Probanden	Probandengruppe				
		Insgesamt	Grs.	Gymn.	Stud.	Phon.
Akademiker (z.B. Lehrer)	Akademiker	50,76	22,62	53,83	63,71	67,14
	Bürotätigkeit	20,01	28,57	21,12	13,43	14,76
	Handwerker	15,01	16,67	14,70	13,71	15,24
	Arbeiter (z.B. Fabrikarbeiter)	9,72	22,92	8,07	4,00	1,90
Bürotätigkeit (z.B. Sekretär)	Akademiker	26,14	22,22	26,33	29,33	26,67
	Bürotätigkeit	41,79	28,82	39,61	50,67	52,78
	Handwerker	7,61	14,24	7,73	3,67	3,33
	Arbeiter	10,41	17,01	11,35	4,33	7,78
Handwerker (z.B. Tischler)	Akademiker	15,95	21,13	16,56	10,86	14,76
	Bürotätigkeit	37,93	29,17	37,06	48,57	36,19
	Handwerker	19,07	17,26	19,67	18,57	21,43
	Arbeiter	21,32	20,83	22,36	18,57	24,29

3.4.10 Persönlichkeit

Die Antworten auf die Frage nach der Persönlichkeit der Sprecher bzw. ihrer Intro-/Extraversion, für die jüngeren Zuhörer vereinfacht als „nach innen“ bzw. „nach außen gekehrt“ formuliert (s. Frage 10 des Fragebogens), beruhen auf einer Selbsteinschätzung der Sprecher, wie in 2.5 dargelegt; daher werden sie nicht als „richtig“ oder „falsch“ gewertet. Im Folgenden werden die Übereinstimmungen zwischen den Antworten der Probanden und den tatsächlichen Angaben der Sprecher über sich selbst untersucht.

Die Antworten aller Zuhörer stimmten in 2864 von 4531 Fällen (63,21 %) mit denjenigen der Sprecher überein; den Grundschulern gelang dies in 56,16 %, den Gymnasiasten in 65,28 % sowie den Studierenden und Phoniatern in 65,22 % bzw. 66,38 % der Fälle. Gemäß der Welch-ANOVA und Games-Howell Test ergab sich ein signifikanter Unterschied zwischen dem Ergebnis der Grundschüler und jeweils dem der 3 weiteren Probandengruppen (Welch-Test $F(3, 2163,55) = 10,12, p < 0,001$).

Wie in Tab. 18 ersichtlich, stimmten Selbsteinschätzung und Eindruck der 4 Zuhörerguppen v.a. bei den extrovertierten Sprechern in mindestens zwei Drittel der Fälle überein. Die extrovertierten Sprecher wurden häufiger für extrovertiert gehalten (in 70,13 %) als die introvertierten Sprecher (in 45,48 % der Fälle), s. gelb hinterlegte Zeilen,

ebenso wurden die introvertierten Sprecher häufiger als introvertiert eingeschätzt (s. grün hinterlegte Zeilen). Im Chi-Quadrat-Test war der Zusammenhang zwischen Selbsteinschätzungen der Sprecher und Vermutungen der Probanden signifikant ($\chi^2(1) = 283,74$, $p < 0,001$).

Tab. 18: Angaben der Probanden zur Persönlichkeit der Sprecher (Prozent) in Abhängigkeit von ihrer Selbsteinschätzung, sortiert nach den Probandengruppen

Persönlichkeit Sprecher	Angabe der Probanden	Probandengruppe				
		Insgesamt	Grs.	Gymn.	Stud.	Phon.
introvertiert	introvertiert	54,16	42,71	55,07	62,80	56,00
	extrovertiert	45,48	56,88	44,64	36,60	44,00
extrovertiert	introvertiert	29,40	32,85	26,64	32,46	25,13
	extrovertiert	70,13	66,35	73,13	67,08	74,36

3.4.11 Äußere Erscheinung

Auch bei Frage 11 des Fragebogens sollen die Antworten der Zuhörer mit den Selbsteinschätzungen der Sprecher verglichen werden. Die äußere Erscheinung der Sprecher umfasst v.a. deren Kleidungsstil bzw. Auftreten.

Die Zuhörer kreuzten in 3660 von 4531 Fällen und somit in 80,78 % der Fälle an, dass die äußere Erscheinung der Sprecher „alltäglich/normal“ sei. In lediglich 18,80 % (850/4531 Fälle) wurde die Antwortmöglichkeit „vornehm/elegant“ ausgewählt.

Dies lässt sich auch in Tab. 19 erkennen: Sowohl bei den Sprechern mit alltäglicher Erscheinung als auch bei denjenigen mit vornehmer Erscheinung wurde von den Probanden in ca. 80 % der Fälle eine alltägliche äußere Erscheinung vermutet (s. grün hinterlegte Zeilen); entsprechend wurde grob bei ca. 20 % der Sprecher ein vornehmes Äußeres angenommen (s. gelb hinterlegte Zeilen). Im Chi-Quadrat-Test ergab sich kein signifikanter Zusammenhang zwischen den Selbsteinschätzungen und den Angaben der Zuhörer ($\chi^2(1) = 3,00$, $p = 0,083$). Somit konnten die Zuhörer offenbar nicht zwischen Sprechern, die ihr Äußeres als alltäglich bzw. als elegant einschätzten, anhand des Hörens ihrer Stimmaufnahmen unterscheiden.

Tab. 19: Angaben der Probanden zur äußeren Erscheinung der Sprecher (Prozent) in Abhängigkeit von ihrer Selbsteinschätzung, sortiert nach den Probandengruppen

Äußere Erscheinung Sprecher	Angabe der Probanden	Probandengruppe				
		Insgesamt	Grs.	Gymn.	Stud.	Phon.
alltäglich/normal	alltäglich/normal	81,24	79,93	82,76	79,37	82,98
	vornehm/elegant	18,30	18,75	16,93	20,63	16,84
vornehm/elegant	alltäglich/normal	78,55	83,33	79,35	74,00	76,67
	vornehm/elegant	20,94	15,10	20,29	26,00	23,33

3.4.12 Höreindruck der Stimme

Die Zuhörer sollten ihren subjektiven Höreindruck der Stimmen der Sprecher angeben (s. Frage 12 des Fragebogens).

Die Zuhörer wählten in rund der Hälfte der Fälle (50,54 %, 2290/4531 Fälle) die Antwort „neutral“ aus. In 34,03 % (1542/4531 Fälle) empfanden sie die Stimmen der Sprecher als angenehm und in 15,27 % der Fälle (692/4531 Fälle) als nicht angenehm.

Im Folgenden sollen die Angaben der Zuhörer in Abhängigkeit von Geschlecht und Alter der Sprecher dargestellt werden. Wie in Tab. 20 ersichtlich, wurden die männlichen Stimmen geringgradig häufiger als angenehm empfunden (in 36,14 %) als die weiblichen Stimmen (in 32,41 %). Der Chi-Quadrat-Test ergab einen schwachen, aber signifikanten Zusammenhang zwischen dem Geschlecht der Sprecher und dem Höreindruck der Zuhörer ($\chi^2(2) = 9,13$, $p = 0,010$).

Tab. 20: Angaben der Probanden zu den Stimmen der Sprecher (Prozent) in Abhängigkeit vom Geschlecht der Sprecher

Geschlecht Sprecher	Höreindruck Probanden		
	angenehm	neutral	nicht angenehm
männlich	36,14	49,85	13,91
weiblich	32,41	51,07	16,32

Hinsichtlich des Alters der Sprecher wurden diese in 3 breitere Altersgruppen gegliedert, wie es bereits bei der Variable Alter durchgeführt wurde (s. 3.4.1). Es ist erkennbar (s. Tab. 21), dass mit steigendem Alter der Sprecher deren Stimmen tendenziell als angenehmer empfunden wurden. Im Chi-Quadrat-Test zeigte sich ebenfalls ein schwacher,

jedoch signifikanter Zusammenhang zwischen dem Alter der Sprecher und dem Höreindruck der Probanden ($\chi^2(4) = 49,70$, $p < 0,001$).

Tab. 21: Angaben der Probanden zu den Stimmen der Sprecher (Prozent) in Abhängigkeit vom Alter der Sprecher (Jahre)

Altersgruppe Sprecher (Jahre)	Höreindruck Probanden		
	angenehm	neutral	nicht angenehm
5-34	26,76	55,84	17,19
35-64	36,27	48,68	14,91
65-94	39,29	47,21	13,40

4. Diskussion

4.1 Erörterung der Ergebnisse in den einzelnen Variablen

Bei der Betrachtung der Ergebnisse der 4 Probandengruppen fällt insgesamt auf, dass in keiner der untersuchten Kategorien ein signifikanter Unterschied in der Anzahl der korrekten Antworten zwischen Phoniatern und Studierenden besteht. Somit fallen die Ergebnisse der beiden Gruppen ähnlich aus, wobei in der Mehrzahl der Fälle die Studierenden sogar eine etwas höhere Anzahl an richtigen Antworten erzielten. Somit erreichten letztere im Vergleich zu den Phoniatern auch häufiger signifikant bessere Ergebnisse als die Gymnasiasten. Die Grundschüler erzielten in fast allen Kategorien signifikant weniger korrekte Zuordnungen als die anderen Hörergruppen. Diese Beobachtungen werden in 4.2 näher behandelt. Im Folgenden sollen die Ergebnisse in den einzelnen Kategorien diskutiert werden.

4.1.1 Alter

Es zeigte sich ein deutlicher Zusammenhang zwischen der durch die Zuhörer geschätzten und der tatsächlichen Altersgruppe der Sprecher (Spearman-Korrelationskoeffizient: $r = 0,772$, $p < 0,001$). Dies bestätigt die Ergebnisse früherer Studien, die ebenfalls diese Korrelation untersuchten; allerdings sollten die Zuhörer in der Mehrzahl der Studien direkte Schätzungen des Sprecheralters abgeben, somit wurde anstelle des Spearman-Korrelationskoeffizienten für ordinalskalierte Variablen der Korrelationskoeffizient nach Pearson berechnet. Es wurden stark positive Korrelationen beschrieben, u.a. variierend von 0,68 bei Braun (1996) bis zu 0,88 bei männlichen Sprechern und 0,93 bei weiblichen Sprechern in der Studie von Ryan und Capadano (1978).

Die Zuhörer erzielten in der vorliegenden Arbeit in insg. 35,51 % der Fälle korrekte Alterszuordnungen. Dies mag gering erscheinen, jedoch ist zugleich zu bedenken, dass es bei dieser Frage 9 Antwortmöglichkeiten (Altersdekaden im Bereich von 5-94 Jahren entsprechend) gibt. Somit beträgt die Wahrscheinlichkeit, zufällig die richtige Antwort auszuwählen, ca. 11 % - mehr als 3x so häufig antworteten die Probanden tatsächlich richtig. Cerrato et al. (2000) berichteten über ähnliche Resultate: Der Prozentsatz der richtigen Antworten der Zuhörer betrug im Durchschnitt lediglich 27 %. Dies begründe-

ten sie jedoch mit der geringen Klassenbreite der Antwortmöglichkeiten, welche in ihrer Studie 7 Jahre betrug. Diese Vermutung bestätigend, fiel auch in der vorliegenden Studie die Anzahl der richtigen Antworten mit 74,18 % deutlich höher aus, wenn ein Fehler von ± 1 Altersgruppe gewährt wurde, die Klassen also statt mit 10 Jahren mit 30 Jahren breiter gewählt wurden. Somit konnte gezeigt werden, dass Zuhörer in der Lage sind, relativ präzise Altersschätzungen anhand des Hörens von Sprecherstimmen vorzunehmen; so auch in der Studie von Neiman und Applegate (1990), in der die Probanden Stimmaufnahmen 7 Alterskategorien von 20 bis 75 Jahren zuordnen sollten und der Prozentsatz der Übereinstimmungen, gebildet durch den Quotienten aus chronologischem und vermutetem Alter, bei 80,25 % lag.

Wie mittels des Streudiagramms (Abb. 4) sowie der berechneten Abweichungen (Tab. 2, Abb. 5) gezeigt wurde, überschätzten die Probanden das Alter der jüngeren Sprecher (5-34 Jahre) und unterschätzten das Alter der älteren Sprecher (65-94 Jahre). Diese Beobachtung wurde auch in einigen vorherigen Studien beschrieben (bspw. Cerrato et al., 2000; Schötz, 2005). Bei den Grundschulern fielen die entsprechenden Abweichungen am stärksten aus. Huntley et al. (1987) vermuteten als mögliche Ursache das sich noch in Entwicklung befindliche ‚Alterskonzept‘ von jüngeren Zuhörern: Diese würden aufgrund ihres jungen Alters andere Personen tendenziell als älter als sich selbst wahrnehmen, was zu einer Überschätzung des Alters führen kann. Zugleich fällt ihnen aufgrund ihres Mangels an Erfahrung eine Altersdifferenzierung innerhalb einer Gruppe von älteren Sprechern schwer, weshalb sie ältere Sprecher möglicherweise jünger schätzen. Eppley und Mueller (2001) dagegen stellten fest, dass jüngere Zuhörer das Alter älterer Sprecher nicht signifikant stärker unterschätzten als ältere Zuhörer, weshalb sie vermuteten, dass wahrnehmbare stimmliche Merkmale des Alterns sich langsamer als der chronologische Alterungsprozess entwickeln und daher ältere Sprecher generell häufig für jünger gehalten werden.

Weiterhin lässt sich konstatieren, dass den Zuhörern mit zunehmendem Alter der Sprecher signifikant weniger korrekte Altersschätzungen gelangen. Überraschend erscheint, dass die Phoniater bei den Sprechern im Alter von 65-94 Jahren weniger richtige Alterszuordnungen erzielten als alle anderen Zuhörergruppen einschließlich der Grundschüler (s. Tab. 4). Moyse (2014) dokumentierte ebenfalls eine Abnahme der Genauigkeit der Altersschätzungen, wenn die Sprecher älter waren. Möglicherweise ist diese statistisch

dadurch zu erklären, dass bei jüngeren Sprechern, bspw. im Alter von 15-24 Jahren, die möglichen Abweichungen der Schätzungen nach unten begrenzt sind, während bei älteren Sprechern Abweichungen ober- oder unterhalb des tatsächlichen Alters möglich sind (Kreiman und Sidtis, 2011). Zudem schlägt Moyse (2014) vor, dass der Alterungsprozess eines Sprechers individuell auch von extrinsischen Faktoren wie z.B. Rauchverhalten und Arbeitsbelastung beeinflusst wird, wodurch die Variabilität der Altersschätzungen bei älteren Sprechern höher sein kann. Die geringe Anzahl der richtigen Altersschätzungen, die die Phoniater bei den älteren Sprechern erzielten (s. oben), lässt sich durch die Vermutung von Huntley et al. (1987), dass ältere Zuhörer Sprecher tendenziell als jünger als sich selbst einschätzen und somit ältere Sprecher häufig für jünger halten, nur bedingt erklären - schließlich sind in der vorliegenden Studie viele Sprecher deutlich älter als die Phoniater, deren Altersdurchschnitt bei ca. 49 Jahren liegt (s. 3.2). Rückblickend wäre es sicher sinnvoll gewesen, noch eine weitere Zuhörergruppe mit Erwachsenen fortgeschritteneren Alters in die Studie miteinzubeziehen, um den Zusammenhang von Zuhöreralter und Genauigkeit der Altersschätzungen noch genauer zu untersuchen.

Bei den jüngeren Sprechern (5-34 Jahren), die auch Kinder und Jugendliche beinhalten, gelangen jedenfalls in mehr als der Hälfte der Fälle (in 54,24 %) richtige Altersschätzungen. Als vergleichbare Studie fand sich nur eine Studie, die ebenfalls untersuchte, wie genau die Altersschätzungen für kindliche Sprecher ausfallen: Amir et al. (2012) stellten ebenso fest, dass Zuhörer das Alter von Kindern deutlich über dem Zufallsniveau identifizieren konnten.

4.1.2 Körperbau

Wie bereits in der Einleitung (s. 1.3) ausgeführt, sollten die Zuhörer in den vorhandenen Studien Körpergröße und -gewicht der Sprecher einschätzen. Zunächst soll angemerkt werden, dass die vorliegende Studie von dieser Vorgehensweise aufgrund der Zusammensetzung ihrer Probanden bewusst abwich: Statt das Körpergewicht zu schätzen, bestand die Aufgabe darin, die Sprecher vereinfachten, schematischen Zeichnungen der Körperformen zuzuordnen. Die Körperformen entsprechen den unterschiedlichen BMI-Kategorien (s. 2.2.1). Dieses Vorgehen wurde gewählt, um auch den jüngeren Zuhörern

die Aufgabenstellung so verständlich wie möglich zu machen und so allen Zuhörern zumindest ähnliche Voraussetzungen für die Beantwortung dieser Frage zu ermöglichen. Insbesondere für die Zuhörer im Grundschulalter dürfte die Differenzierung zwischen verschiedenen Körpergrößen und v.a. Gewichtsangaben mitsamt der verwendeten Einheiten (cm, kg) eine aus ihrer Sicht abstrakte Aufgabe darstellen.

Studien zur direkten Schätzung von Körpergröße und -gewicht von Sprechern zeigten inkonsistente Ergebnisse: So führten etwa Bruckert et al. (2006) eine Studie mit weiblichen Zuhörern durch und berichteten, dass diese das Gewicht von männlichen Sprechern einschätzen konnten (Korrelationskoeffizient nach Pearson: $r = 0,4$, $p < 0,05$), jedoch keine signifikante Korrelation zwischen der geschätzten und tatsächlichen Körpergröße der Sprecher bestand. Van Dommelen und Moxness (1995) stellten fest, dass nur für männliche Sprecher signifikante Korrelationen zwischen geschätzter und tatsächlicher Körpergröße/-gewicht vorlagen, wobei die korrekten Schätzungen meist von männlichen Zuhörern stammten. Dagegen konnten die entsprechenden Körpermaße von Frauen weder von männlichen noch weiblichen Zuhörern korrekt geschätzt werden. Gonzalez (2003) wiederum ermittelte in einer Reanalyse älterer Studien (Lass und Kollegen), dass nur 14 % der Schätzungen von Größe und Gewicht signifikant mit den tatsächlichen Werten korrelierten, und folgerte, dass Probanden Größe und Gewicht von Personen nicht genau identifizieren können, wenn sie lediglich ihre Stimmen hören. In der vorliegenden Studie ergab sich eine deutliche Korrelation zwischen der vermuteten und tatsächlichen Größenkategorie (Spearman-Korrelationskoeffizient: $r = 0,617$, $p < 0,001$) und eine schwach positive Korrelation zwischen der geschätzten und tatsächlichen BMI-Klasse ($r = 0,202$, $p < 0,001$). Den Probanden gelangen insgesamt in 39,17 % der Fälle korrekte Größenzuordnungen und in 27,70 % der Fälle korrekte Angaben der BMI-Klassen der Sprecher. Dass im Unterschied zu den genannten Studien in beiden Kategorien signifikante Korrelationen vorlagen, könnte u.a. damit zusammenhängen, dass eine Kategorisierungsaufgabe mit vorgegebenen Antwortmöglichkeiten einen geringeren Schwierigkeitsgrad für die Probanden aufweist als eine direkte Schätzungsaufgabe, bei der ihnen bspw. nur der mögliche Wertebereich angegeben wird (Bruckert et al., 2006). Rendall et al. (2007) führten eine Kategorisierungsaufgabe durch, bei der die Zuhörer die Stimmen von jeweils 2 Sprechern, deren Größenunterschied mindestens 14

cm betrug, hörten und anhand dessen den größeren der beiden Sprecher benennen sollten - dies gelang ihnen in rund 80 % der Fälle.

Es fällt außerdem auf, dass die Zuhörer bei der Einschätzung der Größe bessere Ergebnisse erzielten als beim BMI. Ein möglicher Grund könnte sein, dass die verschiedenen Alters- und Geschlechtergruppen stärker auf die Körpergröße schließen lassen als auf den BMI. So ist bspw. ein Kind im Normalfall nicht so groß wie ein Erwachsener und Frauen sind im Durchschnitt kleiner als Männer; jedoch können sowohl Kinder als auch Erwachsene untergewichtig bis adipös sein. Es ergaben sich Hinweise darauf, dass die Probanden, auch die Zuhörer im Kindesalter, eine grobe Vorstellung davon haben, wie groß Kinder im Vergleich zu Erwachsenen sind (s. Tab. 6). Diese Vermutung lässt sich auch mit dem akustischen Korrelat stützen: Wie in 1.2.3 beschrieben wurde, unterscheiden sich die Grundfrequenzen der Stimmen von Kindern, Männern und Frauen.

Allerdings muss konstatiert werden, dass es den Zuhörern nicht gelang, innerhalb der Gruppe der erwachsenen Sprecher zwischen Personen mit verschiedenen Ausprägungen von Körpergröße und BMI zu differenzieren. Sie wiesen vielmehr eine Tendenz zur Mitte bei der Auswahl ihrer Antworten auf: Somit erzielten sie hinsichtlich der Größe bei den mittleren Größenkategorien die meisten korrekten Zuordnungen, während bspw. Frauen mit geringer und v.a. Männer mit hoher Körpergröße seltener richtig zugeordnet wurden (s. Abb. 9). Bei der Einschätzung des BMI waren sie entsprechend bei den Kategorien Normal- und Übergewicht am erfolgreichsten, wohingegen untergewichtige und adipöse Sprecher selten korrekt identifiziert wurden (s. Abb. 11). In Entsprechung dazu wurde festgestellt (s. 1.2.3), dass innerhalb der Geschlechter- und Altersgruppen keine signifikante Korrelation zwischen Grundfrequenz und Körpergröße/-gewicht vorliegt, also bspw. die Stimme eines größeren Mannes nicht notwendigerweise tiefer ist als die eines kleineren. Problematisch ist, dass in der Studie Sprecher mit mittleren Größen- und BMI-Kategorien tatsächlich in der Mehrzahl der Fälle vertreten waren, wodurch die Ergebnisse der Zuhörer möglicherweise fälschlich gut ausfallen. Auch Gunter und Manning (1982) schilderten die Problematik, dass die Genauigkeit der Größen- und Gewichtsschätzungen der scheinbar ‚besseren‘ Zuhörer hauptsächlich darauf zurückzuführen war, dass diese wiederholt Werte nahe dem Durchschnitt bzw. Populationsmittelwert schätzten. Somit ist die Probandenauswahl in vorliegender Studie nicht ideal, da nicht

gleich viele Sprecher in jeder Größen-/BMI-Kategorie vorliegen. Dies war jedoch aufgrund der Vielzahl der untersuchten Kategorien nicht umsetzbar.

4.1.3 Geschlecht

Die Zuhörer konnten das Geschlecht der Sprecher anhand des Hörens ihrer Stimmnahmen in insg. 94,92 % der Fälle korrekt und somit zuverlässig identifizieren. Auch in früheren Studien wurde gezeigt, dass Zuhörer in der Lage sind, selbst bei kurzer Dauer der Stimmaufnahme das Geschlecht sicher zu erkennen. Anhand von stimmhaften Vokalen etwa konnten sie das Geschlecht nach Lass et al. (1976) in 96 % der Fälle identifizieren, bei Darbietung kurzer Vokalsegmente von 50-100 ms Dauer nach Whiteside (1998) in 98,9 %. Lass et al. (1979) spielten den Zuhörern verschiedene Hörreize von Sprechern vor (einzelne Vokale, ein- und zweisilbige Wörter, Sätze) und stellten fest, dass kein regelmäßiger Trend in der Hörergenaugigkeit von einfachen zu komplexen Reizen gegeben war. Stattdessen waren die Ergebnisse bei allen 4 Stufen der phonetischen Komplexität ähnlich gut und es ergab sich insg. ein Prozentsatz der korrekten Geschlechtsidentifikationen von 97,6 %. Bei Cerrato et al. (2000) wiederum lasen die Sprecher einen Text von ca. 40 s Dauer vor und die Zuhörer konnten anhand dessen das Geschlecht fehlerfrei zuordnen. Diese Resultate decken sich damit, dass in Sprachsignalen Informationen über das Geschlecht eines Sprechers redundant dargeboten werden, s. 1.2.2 (Kreiman und Sidtis, 2011); Grund- und Formantfrequenzen stellen die wichtigsten akustische Hinweise dar und erlauben eine Identifikation des Geschlechts mit nur geringen Fehlerraten (Bachorowski und Owren, 1999).

Der Prozentsatz der richtigen Geschlechtsidentifikationen wurde in der vorliegenden Studie auch davon beeinflusst, dass Sprecher im Kindesalter miteinbezogen wurden. Wie in Tab. 10 gezeigt, waren die Zuhörer sich über das Geschlecht der Kinder, genauer gesagt der beiden 9-jährigen Jungen häufig unsicher und hielten diese in rund 40 % der Fälle für weiblich. Diese Fehleinschätzung ist insofern erklärbar, als sich bei Mädchen und Jungen vor der Pubertät die Grundfrequenz respektive Tonhöhe als für die Zuhörer wichtiger akustischer Parameter, der grundsätzlich Rückschlüsse auf Geschlecht und Alter erlaubt, nicht unterscheidet (s. 1.2.2). Da sich somit erst im Verlauf der Pubertät die physiologischen Unterschiede des Stimmapparates von Mädchen und

Jungen mit den entsprechenden akustischen Korrelaten ausbilden, fungieren in den meisten Studien Erwachsene als Sprecher und nur wenige Studien haben untersucht, wie sicher auch das Geschlecht von kindlichen Sprechern erkannt wird. Nach Amir et al. (2012) wurde das Geschlecht von Kindern und Jugendlichen im Alter von 8, 10, 12, 14, 16 und 18 Jahren in insg. 81,81 % der Fälle richtig erkannt. Sie beobachteten ebenfalls, dass in den jüngeren Altersgruppen bis 12 Jahre Mädchen häufiger korrekt identifiziert wurden, in den Altersgruppen ab 14 Jahren hingegen Jungen, was sie mit dem Stimmwechsel zur Männerstimme mit einer deutlich niedrigeren Grundfrequenz begründeten. Dies passt zu den Ergebnissen der vorliegenden Studie, in der die beiden Jungen sich mit 9 Jahren noch vor Beginn der Pubertät befinden. Perry et al. (2001) stellten ebenso fest, dass Zuhörer das Geschlecht von Kindern anhand der Stimme unterscheiden konnten, bereits bei Kindern im Alter von 4 Jahren gelang ihnen dies in 64,5 % der Fälle. Bei Traunmüller und van Bezooijen (1994) allerdings lag die Rate der korrekten Zuordnungen für Kinder im Alter von 5, 7, 9 und 11 Jahren mit 8-40 % deutlich niedriger. Auch hier wurde die Tendenz beschrieben, dass jüngere Kinder öfter als Mädchen und ältere Kinder öfter als Jungen klassifiziert wurden. Insgesamt ist festzuhalten, dass in der vorliegenden Studie die Zuhörer das Geschlecht der kindlichen Sprecher zwar weniger häufig korrekt erkannten als das der Erwachsenen, ihnen dies aber dennoch in der Mehrzahl der Fälle (in 72,42 %) - wie in den Studien von Amir et al. (2012) und Perry et al. (2001) - gelang. Daher ist anzunehmen, dass die Einschätzungen der Zuhörer in Bezug auf Kinder auf den Formantfrequenzen beruhen, die sich bereits zwischen Mädchen und Jungen im Alter von 4 Jahren unterscheiden (Perry et al., 2001).

Im Hinblick auf die erwachsenen Sprecher wurde festgestellt, dass Männer zwar geringgradig, aber signifikant häufiger korrekt erkannt wurden als Frauen. Lass et al. (1976) und Owren et al. (2007) machten ebenso die Beobachtung, dass Zuhörer bei Darbietung von Vokalen bzw. kurzen Vokalsegmenten die männlichen Sprecher häufiger korrekt identifizierten. Owren et al. (2007) schlagen als Erklärung vor, dass die Stimmen postpubertärer Männer gewissermaßen durch abgesenkte Grund- und Formantfrequenzen evolutionsbiologisch markiert sind. Somit handelt es sich beim Vorliegen dieser Merkmale mit großer Sicherheit um die Stimme eines erwachsenen Mannes, was auch die Zuhörer erkennen. Beim Fehlen der für einen Mann charakteristischen Grund- und Formantfrequenzen kann es sich um eine erwachsene Frau handeln, jedoch möglicher-

weise auch um einen präpubertären Mann oder um einen erwachsenen Mann mit geringerem oder unvollständigem Stimmwechsel.

Schließlich ist anzumerken, dass alle 4 Probandengruppen das Geschlecht der Sprecher sicher erkennen konnten und ähnlich gute Ergebnisse erzielten, die sich nur zwischen Grundschulern und Studierenden signifikant unterschieden. Somit gelangen auch den Probanden im Kindesalter mit 93,39 % ähnliche viele korrekte Zuordnungen wie den Erwachsenen. Dies ist konsistent mit den Ergebnissen der Studie von Bennett und Montero-Diaz (1982), in der Kinder im Alter von 6 bis 9 Jahren u.a. das Geschlecht von erwachsenen Sprechern zuordnen sollten und ähnlich gute Ergebnisse wie erwachsene Zuhörer erzielten (97-100 % korrekte Zuordnungen für stimmhafte Vokale).

4.1.4 Akzent/Dialekt/Hochdeutsch

Während die Zuhörer bei den Sprechern, die Hochdeutsch bzw. mit ausländischem Akzent sprachen, in über drei Viertel der Fälle korrekte Zuordnungen erzielten, schien ihnen die Identifizierung des Dialekts schwerer zu fallen. Im Folgenden sollen die erzielten Ergebnisse hinsichtlich Akzent und Dialekt erörtert werden.

Sprecher ohne Akzent oder Dialekt wurden in 76,14 % der Fälle auch als solche identifiziert. Bei den Sprechern mit ausländischem Akzent erkannten die Zuhörer sogar in insg. 89,34 % der Fälle, dass ein Akzent vorliegt, und identifizierten den jeweilig vorliegenden Akzent in mehr als drei Viertel der Fälle korrekt (in 79,53 %). In vergleichbaren Studien wurde meist untersucht, inwiefern Zuhörer überhaupt zwischen Sprechern mit und ohne ausländischen Akzent unterscheiden können. In der Studie von Cunningham-Andersson und Engstrand (1989) hörten schwedische Zuhörer Stimmaufnahmen von Sprechern mit verschiedenen ausländischen Muttersprachen und somit ausländischen Akzenten, die einen schwedischsprachigen Text vorlasen. Die Zuhörer sollten auf einer Skala von 1 (kein ausländischer Akzent) bis 5 (starker ausländischer Akzent) die Stärke des Akzents auswählen. Es war ersichtlich, dass dies den Laienhörern gelang, da ihre Einschätzungen signifikant mit den von Experten beschriebenen Abweichungen von der Muttersprache korrelierten. Major (2007) stellten fest, dass Zuhörer sogar in dem Fall zwischen Sprechern ohne Akzent und Sprechern mit ausländischem Akzent unterscheiden konnten, wenn es sich bei dem Gehörten nicht um ihre eigene Muttersprache handelt. Dass

auch Kinder einen ausländischen Akzent schon wahrnehmen können, zeigten Girard et al. (2008): 5- bis 6-jährigen französischsprachigen Kindern gelang es, zwischen Stimm-aufnahmen zu unterscheiden, die zum einen in ihrem Heimatdialekt und zum anderen von englischen Muttersprachlern mit starkem fremdsprachigen Akzent erstellt wurden. Auch in der vorliegenden Studie gelang es den 8- bis 11-jährigen Grundschulern in 63,19 % der Fälle, die Sprecher mit Akzent richtigerweise als solche einzuordnen.

Darüber hinaus sollten die Zuhörer in dieser Studie auch auswählen, welchen von 3 ausländischen Akzenten sie vermuteten. Gymnasiasten, Studierende und Phoniater entschieden sich in über 90 % der Fälle für den korrekten Akzent. Durchaus überraschend erscheint, dass bereits die Gymnasiasten mit so hohem Erfolg die vorhandenen Akzente identifizieren konnten. Den Zuhörern im Kindesalter gelang dies deutlicher seltener (in 31-39 % der Fälle).

Im Hinblick auf die Sprecher mit Dialekt kreuzten die Zuhörer insg. in 75,41 % richtigerweise einen Dialekt an und identifizierten den jeweilig vorliegenden Dialekt insg. in 46,93 % der Fälle korrekt. Während in der vorliegenden Studie die Zuhörer einen von mehreren Dialekten auswählen sollten, sollten sie in den bisherigen Studien die Sprecher vorgegebenen geografischen Regionen zuordnen; dabei wurde ihnen zur Hilfe eine Landkarte mit den entsprechenden Regionen zur Verfügung gestellt. Bei van Bezooijen und Gooskens (1999) sollten die Zuhörer niederländische Stimm-aufnahmen Regionen in den Niederlanden und Belgien sowie englische Stimm-aufnahmen Regionen im Vereinigten Königreich zuordnen. Es wurde festgestellt, dass die Zuhörer insg. in über der Hälfte der Fälle richtig angeben konnten, aus welcher Region im jeweiligen Land die Sprecher stammten (z.B. ob aus Nord-, Zentral- oder Südengland). Auch Clopper und Pisoni (2004) stellten fest, dass es Zuhörern gelang, Sprecher des amerikanischen Englisch 3 breiteren Dialektclustern zuzuordnen, sie aber Schwierigkeiten bei der Zuordnung zu 6 kleineren Regionen hatten.

In der vorliegenden Studie zeigte sich bei den einzelnen Dialekten ebenfalls eine Variabilität der Ergebnisse: Einige Dialekte, wie in 3.4.5 beschrieben Bayerisch, Sächsisch und Kölsch, waren offenbar leichter zu erkennen, während der schwäbische und der norddeutsche Dialekt weniger häufig identifiziert wurden. Hundt et al. (2017) führten eine vergleichbare Studie im deutschsprachigen Raum durch und erhielten ähnliche Resultate: Die Rate der korrekten Zuordnungen differierte für die einzelnen Dialekte zwischen

39 und 91 %. Außerdem fiel in der Studie dieser Doktorarbeit auf, dass die 8- bis 11-jährigen Grundschüler die Dialekte deutlich seltener korrekt identifizierten, mit erheblichen Schwankungen für die jeweiligen Dialekte: Kölsch und Bayerisch erkannten sie immerhin in mehr als 20 % der Fälle. Für Kölsch rührt dieses Ergebnis möglicherweise daher, dass ihnen dieser Dialekt aufgrund ihrer Herkunft oder ihres Umfeldes vertrauter oder bekannt ist - schließlich wurde die Befragung in Schulen im Bonner und Kölner Raum durchgeführt. Die restlichen Dialekte schienen ihnen eher fremd zu sein, da sie diese in weniger als 10 % der Fälle erkannten. Girard et al. (2008) stellten fest, dass 5- bis 6-jährige französischsprachige Kinder anhand von Stimmaufnahmen nicht zwischen ihrem Heimatdialekt aus dem Nordosten Frankreichs und einem Dialekt aus Südfrankreich differenzieren konnten. Wagner et al. (2014) kamen zu demselben Ergebnis bei englischsprachigen Kindern bzw. Sprechern.

Im Folgenden wird ein zusammenfassender Vergleich der Ergebnisse der 4 Probandengruppen angestellt, da sie sich bei dieser Kategorie relevant unterscheiden. Hinsichtlich des Akzents waren diese Unterschiede noch moderater ausgeprägt. So erzielten, wie beschrieben, Gymnasiasten, Studierende und Phoniater einheitlich gute Ergebnisse, die Grundschüler erkannten den jeweiligen Akzent weniger häufig, aber immerhin - im Vergleich zu ihren Resultaten bei den Sprechern mit Dialekt - in ca. einem Drittel der Fälle korrekt. Girard et al. (2008) vermuten, dass Zuhörer, auch Kinder, einen Akzent leichter als einen Dialekt erkennen, da ein ausländischer Akzent sich durch gröbere Abweichungen wie bspw. eine nicht flüssige oder stockende Aussprache auszeichnet, während feinere Unterschiede die verschiedenen regionalen Dialekte charakterisieren. Zudem halten sie es für wahrscheinlich, dass Kenntnisse fremder Länder und anderer Sprachen verbreiteter sind als ein detailliertes Wissen über regionale Dialekte innerhalb eines Landes, welches genauere geografische Kenntnisse erfordert.

Hinsichtlich der Identifizierung des Dialekts weichen die Ergebnisse der Hörergruppen stärker voneinander ab, von ca. 15 % korrekten Zuordnungen der Grundschüler bis zu 83 % bei den Phoniatern. Je älter die Zuhörer sind bzw. je höher ihr Ausbildungsstand, desto häufiger erkannten sie den vorliegenden Dialekt. Somit sind die Phoniater bei dieser Kategorie auch deutlich besser als die Studierenden. Dies lässt den Schluss zu, dass es sich bei der Wahrnehmung von v.a. regionalen Sprachvarietäten um einen Lernprozess handelt, der u.a. sicherlich von Lebensalter und Erfahrung abhängig ist.

Hundt et al. (2017) stellten ebenfalls fest, dass ältere Zuhörer die Sprachbeispiele besser verorten konnten. Auch Baker et al. (2009) postulieren, dass Herkunft und Erfahrung wichtige Faktoren für die Erkennung von Dialekten darstellen. So konnten Zuhörer, die in einer Region aufgewachsen waren oder später dort lebten, besser zwischen dem dort vorherrschenden regionalen Dialekt und anderen Dialekten differenzieren. Jedoch könnten zu den wenigen richtigen Zuordnungen der Grundschüler auch beigetragen haben, dass einige möglicherweise die Aufgabenstellung nicht verstanden haben. Schließlich waren ihre Antworten deutlich häufiger ungültig als die der anderen Zuhörer (s. 3.4.5), da sie bspw. die Kategorie Akzent ankreuzten, aber auch einen der aufgeführten Dialekte.

Daher ist anzumerken, dass die Aufgabenstellung dieser Kategorie sich nicht als optimal herausstellte, da sie einige Grundschüler offenbar überforderte. Eine Zuordnung zu Regionen wäre aber wahrscheinlich auch nicht zielführender gewesen, da davon auszugehen ist, dass diese das geografische Verständnis von 8- bis 11-jährigen Kindern ebenso übersteigt. Allenfalls wäre eine freie Klassifizierungsaufgabe der Dialekte sinnvoll gewesen, um zu explorieren, ob Kinder grundsätzlich Unterschiede zwischen Sprachvarietäten wahrnehmen. Clopper und Pisoni (2007) führten eine derartige Aufgabe durch, bei der keine Dialektkategorien vorgegeben, sondern die Zuhörer gebeten wurden, die Sprecher hinsichtlich des Dialekts in beliebig viele und beliebig große Gruppen einzuteilen. Es stellte sich heraus, dass diese mehr Gruppen als erwartet bildeten, was nahelegte, dass sie durchaus in der Lage waren, feine Differenzierungen zwischen Dialekten vorzunehmen.

4.1.5 Rauchverhalten

Es gibt bislang wenige Studien, die sich mit den Auswirkungen des Rauchens auf akustische Eigenschaften der Stimme - welche möglicherweise auch Zuhörern Hinweise auf das Rauchverhalten liefern könnten - beschäftigen. Wie in der Einleitung (s. 1.3) erläutert, wurden in diesen Studien meist die Grundfrequenzen der Stimmen von Rauchern und Nichtrauchern untersucht und festgestellt, dass Raucher niedrigere Grundfrequenzen aufweisen. Die Grundfrequenz hat nicht zuletzt einen starken Einfluss auf das wahrgenommene Alter von Sprechern (s. 1.2.1). So überprüften Braun und Rietveld

(1995), inwieweit das Rauchverhalten eines Sprechers sein geschätztes Alter beeinflusst. Sie stellten fest, dass Raucher signifikant älter geschätzt wurden als gleichaltrige Nichtraucher. Bei Winkler (2010) war dieser Effekt schwächer ausgeprägt und erreichte keine statistische Signifikanz. Somit konstatierte er, dass die Altersschätzungen der Zuhörer nicht signifikant von den Rauchgewohnheiten der Sprecher beeinflusst wurden. Als eine mögliche Begründung führte er an, dass in seine Studie entweder junge (Alter bis 28 Jahre) oder ältere Sprecher (ab 55 Jahren) einbezogen wurden, während die Sprecher bei Braun und Rietveld (1995) zwischen 25 und 59 Jahre alt waren; bei jungen und älteren Sprechern seien die Altersschätzungen stärker durch Über- und Unterschätzung gekennzeichnet - wie es auch in vorliegender Studie beobachtet werden konnte (s. 3.4.1). Somit seien die Altersschätzungen in seiner Studie stärker von diesem Zentralisierungseffekt beeinflusst als vom Rauchverhalten der Sprecher. In vorliegender Studie wurde der Einfluss des Rauchverhaltens auf das wahrgenommene Alter nicht untersucht, da aufgrund der Vielzahl der untersuchten Merkmale der Sprecher die Voraussetzungen dafür nicht gegeben waren. So lagen u.a. aufgrund der großen Altersspanne der Sprecher nicht genügend Raucher und Nichtraucher desselben Alters und Geschlechts vor, deren Altersschätzungen man sinnvoll hätte vergleichen können.

In der Studie dieser Doktorarbeit wurde überprüft, inwieweit die Probanden das Rauchverhalten der Sprecher anhand des Hörens ihrer Stimmufnahmen angeben konnten. Es wurde gezeigt, dass sie dieses nicht korrekt identifizieren konnten, da kein signifikanter Zusammenhang zwischen dem tatsächlichen und vermuteten Rauchverhalten vorlag. Vielmehr hielten sämtliche Probandengruppen sowohl die rauchenden als auch die nichtrauchenden Sprecher deutlich häufiger für Nichtraucher (in insg. über 70 % der Fälle) als für Raucher (in ca. 25 % der Fälle). Zudem wurde beobachtet, dass sie die Männer signifikant häufiger als rauchend einschätzten als die Frauen. Dies lässt die Vermutung zu, dass die Zuhörer sich in ihren Urteilen eher auf die angenommene Verteilung des Merkmals in der Gesellschaft verließen. Sie mögen gewusst haben, dass es, zumindest in Deutschland, gegenwärtig deutlich mehr Nichtraucher als Raucher gibt und möglicherweise auch, dass unter den Rauchern die Männer tendenziell überwiegen. Dies bestätigt der Mikrozensus des statistischen Bundesamtes (2017), wonach rund 78 % der Deutschen, einschließlich der früheren Raucher, Nichtraucher sind und ca. 22 % aktive Raucher; ca. 24 % der Männer und ca. 19 % der Frauen sind Raucher.

Mit Winkler (2010) findet sich nur eine einzige Studie, die ebenfalls die Wahrnehmung des Rauchverhaltens durch Zuhörer untersuchte. Dass so wenige Studien zu dieser Thematik vorliegen, liegt möglicherweise daran, dass die Rauchgewohnheiten von Personen schwierig zu quantifizieren sind. Es ist nicht einheitlich festgelegt, ab wann eine Person als Raucher oder starker Raucher gilt, wie Gelegenheitsraucher oder Personen mit geringem Zigarettenkonsum einzuordnen sind. Zudem lassen sich Raucher verschiedenen Alters nur eingeschränkt miteinander vergleichen, da ein älterer Raucher natürlich mehr Raucherjahre aufweist als ein jüngerer Raucher mit ähnlichem Zigarettenkonsum (Winkler, 2010). Viele Faktoren dürften die Einschätzung, ob ein Sprecher raucht, mit beeinflussen; beispielhaft zu nennen dessen Alter und Vorerkrankungen.

Bei Winkler (2010) gelang es den Zuhörern im Unterschied zu den Ergebnissen vorliegender Studie durchaus, den Tabakkonsum der Sprecher einzuschätzen. Allerdings wurden ihre Einschätzungen wie vermutet vom Sprecheralter beeinflusst, da sie ältere Sprecher grundsätzlich häufiger für Raucher hielten als jüngere Sprecher.

4.1.6 Bildungsstand

Aus forensischer Perspektive sind Hinweise über das Bildungsniveau bzw. den sogenannten Soziolekt (s. 1.3) einer Person in der Tat von Relevanz. So kann ebenso wie etwa die geografische Herkunft auch die soziale Herkunft einen wichtigen Anhaltspunkt liefern, der letztendlich zur Identifikation eines Straftäters beitragen kann. Nach Jessen (2007) können Äußerungen von Sprechern bspw. aufgrund der verwendeten Wortwahl und Syntax durchaus Rückschlüsse auf deren Bildungsstand erlauben. Allerdings haben sich bislang Studien nicht damit befasst, inwieweit Zuhörer anhand von Stimmaufnahmen von Sprechern deren Bildungsstand beurteilen können. Möglicherweise ist dies damit begründbar, dass es sich um ein Merkmal handelt, welches einem Menschen nicht primär vorgegeben ist - anders als bspw. Körpergröße und Geschlecht - und aus dem man daher nur eingeschränkt Rückschlüsse ziehen kann. Es kann sich z.B. eine Person mit Abitur, die zum Erwerb eines Hochschulabschlusses befähigt wäre, ebenso gegen eine Berufsausbildung entscheiden. Dann erlaubt der Bildungsabschluss wenige Rückschlüsse etwa auf die Intelligenz der Person.

In der vorliegenden Studie hörten die Zuhörer die Stimmufnahmen der Sprecher, welche einen phonetischen Mustertext (s. 2.2.2) vorlasen, und sollten anhand dessen ihren Schulabschluss, ihre Berufsausbildung und ihren Beruf einschätzen. Es ergaben sich Hinweise darauf, dass die Zuhörer insg. durchaus in der Lage waren, zwischen Sprechern mit unterschiedlichem Bildungsstand zu unterscheiden. So ergab sich jeweils ein signifikanter Zusammenhang zwischen dem tatsächlichen und vermuteten Schul-, beruflichen Abschluss und Beruf. Dies erscheint überraschend v.a. vor dem Hintergrund, dass den Zuhörern die möglichen Hinweise, die sich bei freien Äußerungen der Sprecher bspw. anhand von Wortwahl und Komplexität des Satzbaus ergeben hätten, nicht zur Verfügung standen, da die Sprecher alle denselben Text vorlasen. Im Hinblick auf die einzelnen Probandengruppen muss jedoch einschränkend festgehalten werden, dass den Grundschulern die Unterscheidung zwischen Sprechern mit unterschiedlichen Bildungsabschlüssen offenbar nicht gelang, wie in den Tab. 15, 16 und 17 ersichtlich ist; den anderen Zuhörergruppen aber schon. Rammstedt (2013) untersuchte im Rahmen einer internationalen Studie zur Erfassung von grundlegenden Kompetenzen Erwachsener, inwiefern sich die Lesekompetenzen von Personen mit unterschiedlichem Bildungsabschluss unterscheiden. Die Lesekompetenz umfasst die Fähigkeit des flüssigen Lesens von Textpassagen und des Erschließens ihrer Inhalte, wobei im Falle der vorliegenden Studie die Sprecher lediglich den Text vorlesen sollten. Es wurde festgestellt, dass die Lesekompetenz von Personen mit niedrigerem Bildungsabschluss geringer ausfällt als von Personen mit höherem Abschluss. Offenbar bestimmte v.a. die Höhe des *Schulabschlusses* die Lesekompetenz nachhaltig.

Die Zuhörer in der vorliegenden Studie registrierten offenbar Unterschiede im Vorlesen und konnten so die Sprecher teilweise korrekt zuordnen. Mögliche Kriterien könnten sein, wie flüssig der Text vorgelesen wurde, ob ein sicheres und fehlerfreies Lesen gelang, ob die Betonungen im Text sinnvoll gesetzt wurden etc. Es wäre sicher hilfreich gewesen, die Zuhörer nach ihren subjektiven Beurteilungskriterien zu befragen. Dies hätte jedoch aufgrund der von den restlichen Aufgaben (Multiple Choice) abweichenden Aufgabenstellung die Grundschüler wahrscheinlich methodisch überfordert.

4.1.7 Persönlichkeit

Bei Studien zur Beurteilung der Persönlichkeit ergibt sich das Problem, dass diese aufgrund ihrer Komplexität schwierig zu definieren ist. Die individuellen Charakterzüge einer Person lassen sich nicht eindeutig klassifizieren und eingrenzen, im Unterschied bspw. zu physischen Parametern wie Körpergröße und Alter. Daraus resultiert, dass die Einschätzungen von Zuhörern über die Persönlichkeit von Sprechern einer objektiven Bewertungsgrundlage entbehren. In den wenigen und älteren Studien wurden unterschiedliche Ansätze zur Persönlichkeitserfassung angewendet; meist beruhten diese auf einer Selbsteinschätzung der Sprecher, welche etwa je eines von zwei gegensätzlichen Adjektiven sich selbst zuordnen sollten (z.B. ruhig - gesprächig, Hunt und Lin, 1967). Gelegentlich wurde ein standardisiertes Persönlichkeitsinventar verwendet (Scherer, 1972). Zudem liegt nahe, dass vorhandene Studien sich kaum vergleichen lassen oder Schlussfolgerungen gezogen werden können (Kreiman und Sidtis, 2011).

In der vorliegenden Arbeit, bei der die Angaben der Zuhörer ebenfalls mit den Selbsteinschätzungen der Sprecher verglichen wurden, gelangen allen 4 Probandengruppen korrekte Zuordnungen über dem Zufallsniveau (in über 50 % der Fälle) und es ergab sich ein signifikanter Zusammenhang zwischen den Angaben von Sprechern und Zuhörern. Offenbar boten sich den Zuhörern Hinweise auf die Intro-/Extraversion der Sprecher (z.B. Lebhaftigkeit, Lautstärke, Sicherheit des Vorlesens), die man, wie beim Bildungsstand (s. 4.1.6), genauer hätte explorieren können.

Bei Hunt und Lin (1967) lag die Richtigkeit der Einschätzungen der Zuhörer insg. ebenfalls über dem Zufallsniveau, jedoch variierten die Fehlerraten zwischen den bewerteten Eigenschaften der Sprecher stark.

4.1.8 Äußere Erscheinung

Wie erwartet zeigte sich kein signifikanter Zusammenhang zwischen den Selbsteinschätzungen der Sprecher über ihr Erscheinungsbild und den Vermutungen der Zuhörer. Die Zuhörer nahmen in den meisten Fällen (in über 80 %) eine alltägliche äußere Erscheinung an. Es ist naheliegend, dass Zuhörer nicht von der Stimme einer Person auf äußere Merkmale wie Gesicht, Haare oder Kleidungsstil schließen können, da diese nicht mit physiologischen und akustischen Stimmeeigenschaften in Zusammenhang ste-

hen (Kreiman und Sidtis, 2011). Bei McGlone und Hicks (1979) lasen Frauen im jungen Erwachsenenalter, die mit ähnlichem Dialekt sprachen, zwei vorgegebene Sätze vor. Den Probanden, die die Stimmaufnahmen der Frauen hörten, lagen je Durchgang zwei Fotografien der Frauen von der Taille aufwärts vor und sie sollten versuchen, die Stimmaufnahme der richtigen Fotografie zuzuordnen. Korrekte Zuordnungen gelangen ihnen nicht über dem Zufallsniveau.

4.1.9 Höreindruck der Stimme

Die Antworten der Probanden bei der Frage nach dem Höreindruck waren von der zuvor schon beschriebenen Tendenz zur Mitte geprägt, so kreuzten sie am häufigsten (in rund 50 % der Fälle) einen neutralen Höreindruck an und legten sich seltener auf „angenehm“ oder „nicht angenehm“ fest. Obgleich es sich um einen schwachen Zusammenhang handelt, wurden die männlichen Stimmen etwas häufiger als angenehm empfunden als die weiblichen Stimmen. Deal und Oyer (1991) berichteten ebenfalls, dass männliche Sprecher als angenehmer bewertet wurden, die Gründe dafür blieben jedoch ebenso unklar. Sie schlugen vor, dass Unterschiede im Alterungsprozess von Männern und Frauen, der auch die Stimme betrifft, eine Rolle spielen könnten. Weitere Studien untersuchten den Zusammenhang von Geschlecht und Höreindruck geschlechterspezifisch: Frauen empfanden männliche Stimmen mit niedrigeren Grundfrequenzen im Vergleich zu höheren Grundfrequenzen als angenehmer (Bruckert et al., 2006; Feinberg et al., 2005), während Männer Frauenstimmen mit höheren Grundfrequenzen bevorzugten (Collins und Missing, 2003).

Zudem wurde festgestellt, dass eine Tendenz der Zuhörer bestand, die Stimmen älterer Sprecher als angenehmer einzustufen. In anderen Studien liegen inkonsistente Ergebnisse vor: Während Pittum (1984) ebenfalls positive Einschätzungen für ältere Sprecher dokumentierte, berichteten u.a Deal und Oyer (1991), dass die Stimmen von Sprechern mit zunehmendem Alter als weniger angenehm bewertet wurden. Ob dies auf bestimmten Stimmmerkmalen der älteren Sprecher beruhte oder die Zuhörer jüngere Sprecher aufgrund gesellschaftlicher Vorurteile gegen ältere Personen grundsätzlich bevorzugten, konnte nicht exploriert werden.

In der vorliegenden Studie hätte eine Analyse der akustischen Parameter der Sprecherstimmen Hinweise auf die Merkmale, die die Zuhörer möglicherweise zu ihren Einschätzungen führten, liefern können.

4.2 Erörterung der Ergebnisse im Hinblick auf Alter und Expertise

Es lässt sich, wie auch bereits bei der Darstellung in den einzelnen Variablen ersichtlich, ein deutlicher Unterschied zwischen den Ergebnissen der Grundschüler und der anderen Zuhörergruppen konstatieren. Die Grundschüler erzielten in allen bewerteten Kategorien signifikant weniger korrekte Zuordnungen als jeweils die drei anderen Hörergruppen, mit Ausnahme der Kategorien BMI und Geschlecht, somit in 7 von 9 bewerteten Kategorien. Dies legt nahe, dass die Fähigkeit, bestimmte Merkmale einer Person anhand des Hörens ihrer Stimme zu erkennen, von Erfahrung und damit u.a. auch vom Lebensalter abhängt. So fällt Erwachsenen bspw. die Altersdifferenzierung von Sprechern leichter als Kindern, da sie selbst schon mehr Altersstufen durchlaufen haben.

Die Ergebnisse der anderen 3 Probandengruppen liegen weniger weit auseinander. Den Gymnasiasten gelangen mehr korrekte Zuordnungen als im Vorhinein erwartet, sodass sie häufig ähnlich gute Ergebnisse wie die erwachsenen Zuhörer erzielten, was sich auch an den Gesamtpunktzahlen zeigt (s. 3.3). Weiterhin ist festzuhalten, dass die Phoniater offenbar keinen eindeutigen Vorteil durch ihr höheres Alter oder ihre Expertise im Vergleich zu den Studierenden und teilweise auch Gymnasiasten hatten. Stattdessen unterschieden sich die von Phoniatern und Studierenden erzielten Ergebnisse in keiner Kategorie signifikant und die Phoniater schnitten in lediglich 3 Kategorien signifikant besser ab als die Gymnasiasten. Es zeigte sich sogar, dass die Studierenden in der Mehrzahl der Kategorien eine etwas höhere, wenn auch nicht signifikant höhere, Anzahl an korrekten Antworten erzielten. Ein möglicher Erklärungsansatz für diese durchaus überraschende Beobachtung könnte sein, dass die Befragung einiger Phoniater nicht in einem so kontrollierten Setting stattfinden konnte wie im Falle der Studierenden und Schüler. Zur Gewinnung einer höheren Anzahl an teilnehmenden Phoniatern wurden Fragebögen und Stimmaufnahmen an einen Teil von ihnen postalisch verschickt und sie wurden gebeten, in Eigeninitiative die Aufnahmen anzuhören und den Fragebogen auszufüllen. Daher fehlte eine ‚Überprüfung‘ des Ausfüllens und tatsächlich zeigten sich

etwas höhere Raten der ungültigen und fehlenden Werte im Vergleich zu den Studierenden und Gymnasiasten (s. Abb. 3). Man beachte in diesem Zusammenhang, dass 3 Phoniater wegen einer zu hohen Anzahl fehlender und/oder ungültiger Werte aus der Studie ausgeschlossen werden mussten (s. 3.1). Bei der Kategorie Akzent/Dialekt allerdings zeigten sich wie beschrieben stärkere Unterschiede, insbesondere bei den Dialekten der Sprecher zwischen *allen* Zuhörergruppen, sodass die Anzahl der korrekten Zuordnungen von Gymnasiasten zu Studierenden und auch von Studierenden von Phoniatern deutlich zunahm. In diesem Fall gereichten den Phoniatern ihre Lebenserfahrung und auch Expertise offensichtlich zum Vorteil.

Bislang wurde in anderen Studien kein umfassender Vergleich phonetisch geschulter Hörer und Laienhörer hinsichtlich ihrer Wahrnehmung von Stimmmerkmalen durchgeführt. Es finden sich nur wenige Studien, die bei der Untersuchung eines Stimmmerkmals auch geschulte Hörer miteinbezogen. U.a. Braun und Rietveld (1995) stellten fest, dass Phonetiker mit Erfahrung in der forensischen Stimmenanalyse das Alter von Sprechern nicht signifikant besser schätzen konnten als Laienhörer.

4.3 Evaluation des Studiendesigns

Die vorliegende Arbeit sollte einen Überblick darüber ermöglichen, inwiefern verschiedene Zuhörergruppen diverse Merkmale von Sprechern beurteilen können. Um die Ergebnisse der Gruppen miteinander in Vergleich setzen zu können, mussten der Studienablauf und die Fragebögen für alle Probanden einheitlich sein. Zwar wurden im Hinblick auf die jüngeren Zuhörer bspw. die Formulierungen im Fragebogen vereinfacht gehalten; möglicherweise war dies aber nicht kindgerecht genug und überforderte die Grundschüler teilweise. So war festzustellen, dass die Antworten der Grundschüler im Vergleich zu den restlichen Hörergruppen im Durchschnitt häufiger fehlend oder ungültig waren (s. 3.3). Dass sie bei einigen Fragen keine oder mehr als eine Antwortmöglichkeit auswählten, kann etwa auf mangelnde Konzentration zurückzuführen sein, aber auch darauf, dass sie die Aufgabenstellung nicht verstanden haben oder ihnen Begriffe unbekannt waren. Verständnisprobleme bestanden offenbar insbesondere bei der Frage nach dem Akzent oder Dialekt der Sprecher, wie in 4.1.4 dargelegt wurde. Hinzu kommt, dass durch die Abfrage von 12 Merkmalen für insg. 23 Sprecher allein die Dauer der

Befragung (bei den Grundschulern 1-2 Schulstunden) einen kritischen Faktor für die Konzentration der jüngeren Zuhörer darstellte. Dem wurde mit Pausen und motivierendem Zuspruch versucht entgegenzuwirken. Zusammenfassend kann nicht ausgeschlossen werden, dass bei einer kindgerechteren Durchführung die Grundschüler möglicherweise noch besser hätten motiviert werden können und bessere Ergebnisse erzielt hätten.

Aufgrund der Vielzahl der untersuchten Merkmale war keine ideale respektive gleichmäßige Verteilung der Sprecher in den einzelnen Kategorien gegeben, wie in Tab. 1 ersichtlich ist; so fanden sich z.B. unter den 23 Sprechern nur 5 Raucher, ansonsten Nichtraucher.

Weiterhin muss angemerkt werden, dass die Perzeption der Zuhörer untersucht wurde und keine akustische Analyse der Sprecherstimmen vorgenommen wurde. Dies limitiert die Aussagekraft der Ergebnisse insofern, als die akustischen Merkmale der Stimmen, auf die die Zuhörer sich in ihren Beurteilungen tatsächlich stützten, nicht bekannt sind. Dieser Aspekt sollte in weiteren Untersuchungen auf jeden Fall Berücksichtigung finden, um zu explorieren, aus welchen Gründen bzw. anhand welcher Hinweise Zuhörer bestimmte Merkmale identifizieren können. Dies gilt in besonderem Maße für die bisher noch wenig untersuchten Variablen wie z.B. Rauchverhalten und Bildungsstand. Wie in 2.3.2 ausgeführt, war ursprünglich die Teilnahme von forensischen Phonetikern des Bundeskriminalamtes als weitere Probandengruppe geplant. Diese hätten in ihrer Analyse auch akustische Parameter gemessen, welche zu den Einschätzungen der Zuhörer in Beziehung hätten gesetzt werden können. Es wäre auch aufschlussreich gewesen zu erfahren, wie genau die Urteile der Experten für Sprechererkennung unter Zuhilfenahme akustischer Methoden und bestehender Forschungsliteratur ausfallen.

4.4 Einordnung im Hinblick auf die Fragestellung

Zusammenfassend lassen sich die eingangs gestellten Fragen (s. 1.3) wie folgt beantworten: Die Fähigkeiten der Stimmerkennung von Kindern unterscheiden sich insofern von denen Jugendlicher und Erwachsener, als die Zuhörer im Grundschulalter in allen

bis auf 2 Kategorien signifikant weniger korrekte Antworten als die Gymnasiasten, Studierenden und Phoniater erzielten. Somit konnte eine Altersabhängigkeit der Stimmerkennung bestätigt werden. Dennoch sollte nicht vernachlässigt werden, dass auch die kindlichen Zuhörer bereits in vielen Kategorien richtige Zuordnungen über dem Zufallsniveau erreichten, so konnten sie etwa das Geschlecht von Sprechern zuverlässig angeben und auch ihr Alter grob einordnen. Die Fachärzte für Phoniatrie und Pädaudiologie erzielten keine signifikant besseren Ergebnisse als die Studierenden respektive die erwachsenen Laienhörer.

Für forensische Fragestellungen (z.B. Zeugenaussagen) könnte es durchaus relevant sein, dass bereits Kinder und Jugendliche in der Lage sind, anhand des Hörens von Stimmufnahmen Rückschlüsse v.a. auf physische Parameter wie Geschlecht, Alter und Körpergröße der Sprecher zu ziehen.

5. Zusammenfassung

Die vorliegende Studie beschäftigt sich mit der Wahrnehmung unterschiedlicher Sprechermerkmale durch Zuhörer anhand des Hörens ihrer Stimmaufnahmen. In bisherigen Studien wurden v.a. die Fähigkeiten von Zuhörern zur Identifikation physischer Merkmale, d.h. Alter, Geschlecht, Körpergröße und -gewicht, überprüft. Meist handelte es sich bei den Zuhörern um Erwachsene ohne Erfahrung im Bereich der Stimmerkennung. In dieser Arbeit wurden zusätzlich zu den genannten Kategorien auch bisher weniger berücksichtigte Merkmale untersucht (Akzent/Dialekt, Rauchverhalten, Bildungsstand, Persönlichkeit, äußere Erscheinung, Höreindruck der Sprecher). Es wurden mehrere Zuhörergruppen unterschiedlicher Altersstufen und Expertise einbezogen. Das Ziel der Arbeit war somit der Vergleich der Stimmerkennung von Kindern/Jugendlichen und Erwachsenen sowie von Laien und geschulten Zuhörern.

Insg. 197 Probanden, gegliedert in 4 Gruppen (Grundschüler, Gymnasiasten, Studierende, Fachärzte für Phoniatrie und Pädaudiologie), wurden Stimmaufnahmen von 23 ihnen unbekanntem Sprechern vorgespielt. Anhand dessen sollten sie einen Fragebogen, der die genannten 12 Merkmale der Sprecher abfragte, ausfüllen.

Zwischen vermuteter und tatsächlicher Altersgruppe bestand eine hohe Korrelation und allen Probandengruppen gelang eine grobe Einordnung des Sprecheralters, jedoch überschätzten sie häufig das Alter jüngerer Sprecher und unterschätzten das Alter älterer Sprecher in Entsprechung zu vorherigen Studien. Sie konnten innerhalb der Gruppen der Frauen und Männer nicht zwischen verschiedenen Ausprägungen von Körpergröße und BMI differenzieren. Das Geschlecht der Sprecher konnten sie mit hoher Genauigkeit identifizieren. Die Akzente wurden von den Grundschulern in ca. einem Drittel der Fälle korrekt identifiziert und von den restlichen Hörergruppen mit einheitlich hoher Genauigkeit den Sprechern zugeordnet, während bei den Dialekten zunehmendes Alter und höherer Ausbildungsstand seitens der Zuhörer zu stetig besseren Ergebnissen führten; somit gelang den Phoniatern die Identifizierung der Dialekte am häufigsten. Die Probanden konnten nicht zwischen Rauchern und Nichtrauchern unterscheiden, es zeigte sich somit kein signifikanter Zusammenhang zwischen tatsächlichem und angenommenem Rauchverhalten. Insbesondere Gymnasiasten, Studierende und Phoniater waren durchaus zu einer Differenzierung zwischen Sprechern mit unterschiedlichem Bildungsstand

(Schulabschluss, Berufsausbildung, Beruf) in der Lage. Es ergab sich ein signifikanter Zusammenhang zwischen den Angaben der Zuhörer und den Selbsteinschätzungen der Sprecher über ihre Persönlichkeit (Intro-/Extraversion); jedoch konnten die Zuhörer nicht von der Stimme der Sprecher auf ihre äußere Erscheinung (alltäglich oder vornehm) schließen. Die Probanden empfanden eher männliche Stimmen als angenehmer als weibliche und stufen die Stimmen älterer Sprecher im Vergleich zu jüngeren Sprechern als angenehmer ein.

Im Hinblick auf das Alter der Zuhörer lässt sich zusammenfassen, dass eine Altersabhängigkeit der Stimmerkennung bestätigt werden konnte: Die Zuhörer im Grundschulalter erzielten in allen außer 2 Kategorien signifikant weniger korrekte Antworten als die Gymnasiasten, Studierenden und Phoniater. Dennoch ist zu konstatieren, dass die kindlichen Zuhörer teilweise bereits ähnliche Informationen aus den Stimmen der Sprecher gewinnen können wie die Erwachsenen, etwa im Hinblick auf die physischen Merkmale. Diese Beobachtung könnte für forensische Untersuchungen durchaus von Relevanz sein. Bspw. können auch die Aussagen, die Kinder und Jugendliche im Rahmen von Zeugenaussagen oder auditiven Gegenüberstellungen treffen, durchaus valide und wichtig für die Identifizierung von Straftätern sein.

Die Phoniater und Pädaudiologen hatten offenbar, außer für die Angabe der Dialekte, keinen eindeutigen Vorteil durch ihre Expertise oder ihre größere (Lebens-)Erfahrung. Ihre Ergebnisse waren nicht signifikant besser als die der Studierenden, d.h. der erwachsenen Laienhörer.

In zukünftigen Untersuchungen sollten durch eine akustische Analyse der Sprecherstimmen die akustischen Merkmale exploriert werden, die von Bedeutung für die Urteile der Zuhörer sind. So könnte noch besser nachvollzogen werden, aus welchen Gründen den Zuhörern die Identifikation bestimmter, v.a. bisher noch wenig untersuchter Sprechermerkmale gelingt.

6. Anhang

Universitätsklinikum Bonn
 Prof. Dr. med. Götz Schade
 Leitung der Abteilung für Phoniatrie und Pädaudiologie
 Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde/ Chirurgie
 Doktorandin Johanna Thiele

Fragebogen zur Stimmerkennung

Alter: _____ Jahre Geschlecht: männlich () weiblich ()

Im Folgenden hörst Du eine Geschichte. Diese Geschichte wird von verschiedenen Personen vorgelesen.

Konzentriere Dich dabei bitte auf die Stimme der Sprecherin/ des Sprechers und versuche die Fragen über diese zu beantworten. Kreuze pro Frage bitte immer nur ein Kästchen an. Danke für deine Mitarbeit! 😊

Person 1

1. *Geschlecht der Sprecherin/ des Sprechers*

männlich weiblich

2. *Alter der Sprecherin/ des Sprechers in Jahren*

5-14 45-54 85-94
 15-24 55-64
 25-34 65-74
 35-44 75-84

3. *Liest die Sprecherin/ der Sprecher den Text...*

ohne Akzent/Dialekt (Hochdeutsch)
 mit ausländischem Akzent, und zwar
 Französisch
 Englisch/Amerikanisch
 Türkisch/Arabisch

- mit Dialekt, und zwar
 - Bayerisch
 - Sächsisch
 - Schwäbisch
 - Kölsch
 - Norddeutsch

4. Körperbau der Sprecherin/ des Sprechers

- Abbildung 1
- Abbildung 2
- Abbildung 3
- Abbildung 4
- Abbildung 5



Abb. 1



Abb. 2



Abb. 3



Abb. 4



Abb. 5

5. Körperlänge der Sprecherin/ des Sprechers in cm

- 131-140
- 141-150
- 151-160
- 161-170
- 171-180
- 181-190
- 191-200

6. Rauchverhalten der Sprecherin/ des Sprechers

- RaucherIn
- kein/e RaucherIn

7. Schulabschluss der Sprecherin/ des Sprechers

- Gymnasium/Gesamtschule
- Hauptschule
- noch SchülerIn
- Realschule
- kein Schulabschluss

8. Beruflicher Abschluss der Sprecherin/ des Sprechers

- Studium, z. B. an einer Universität Ausbildung/Lehre, z. B. in einer Firma
 keine Berufsausbildung

9. Beruf der Sprecherin/ des Sprechers

- ArbeiterIn, z. B. BauarbeiterIn, FabrikarbeiterIn
 Bürotätigkeit, z. B. SekretärIn, SachbearbeiterIn
 HandwerkerIn, z. B. TischlerIn, MechanikerIn
 AkademikerIn, z. B. LehrerIn, Rechtsanwalt/-anwältin
 noch nicht berufstätig

10. Persönlichkeit der Sprecherin/ des Sprechers

- nach innen gekehrt (zurückhaltend)
 nach außen gekehrt (kontaktfreudig)

11. Äußere Erscheinung der Sprecherin/ des Sprechers

- alltäglich/normal vornehm/elegant

12. Höreindruck der Stimme der Sprecherin/ des Sprechers

- angenehm neutral nicht angenehm
-

Universitätsklinikum Bonn
Prof. Dr. med. Götz Schade
Leitung der Abteilung für Phoniatrie und Pädaudiologie
Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde/ Chirurgie
Doktorandin Johanna Thiele | thiele.hennef@hotmail.de

Betreff: Probandeninformation und Einverständniserklärung zur Durchführung der Studie zur altersspezifischen forensischen Stimmerkennung

Sehr geehrte Erziehungsberechtigte,

wir möchten Sie um Ihr Einverständnis zur Teilnahme Ihrer Tochter/ Ihres Sohn an einer Studie zur **altersspezifischen forensischen Stimmerkennung** bitten. Die Studie wird von der Abteilung für Phoniatrie & Pädaudiologie der Universitäts-HNO-Klinik Bonn durchgeführt und wurde von der Ethik-Kommission der Medizinischen Fakultät der Universität Bonn genehmigt.

Im Folgenden möchten wir Ihnen und Ihrer Tochter/ Ihrem Sohn nähere Informationen zu der Studie geben, damit Sie Ihre Entscheidung über die Teilnahme treffen können. Bitte lesen sowohl Sie als auch Ihre Tochter/ Ihr Sohn diese Informationsschrift sorgfältig durch. Bei evtl. verbleibenden Fragen stehen wir Ihnen selbstverständlich für weitere Informationen gern zur Verfügung (siehe oben stehende E-Mail-Adresse).

Ihre Einwilligung erfolgt freiwillig und schriftlich. Sie können Ihre Einwilligung jederzeit ohne Angabe von Gründen widerrufen, ohne dass Ihnen daraus Nachteile entstehen.

Vielen Dank im Voraus,
mit herzlichen Grüßen,



Johanna Thiele | Durchführung der Studie

1. Probandeninformation

Das Ziel der Studie ist es herauszufinden, ob Kinder und Jugendliche bereits ähnliche Informationen aus der Stimme eines Probanden gewinnen können wie junge Erwachsene und 'Stimmprofis' (Fachärzte für Phoniatrie & Pädaudiologie).

Zu diesem Zweck soll eine erstmalig realisierte Studie an Probanden unterschiedlicher Altersgruppen durchgeführt werden. Hierbei sollen die Schüler je zweier 3. Grundschulklassen der Waldschule Bonn-Venusberg und der Gemeinschaftsgrundschule Drabenderhöhe, zwei 9. Klassen der gymnasialen Mittelstufe des Beethoven-Gymnasiums Bonn, ca. 40 Student*innen sowie eine Gruppe von Fachärzten für Phoniatrie & Pädaudiologie aus verschiedenen bundesdeutschen Kliniken einbezogen werden.

Konkret wird die Studie folgendermaßen durchgeführt: Ihrer Tochter/ Ihrem Sohn wird ein Mustertext ('Der Nordwind und die Sonne') vorgespielt. Dieser Text von etwa einer Minute Dauer wurde zuvor von mehreren verschiedenen Personen vorgelesen und von uns digitalisiert. Die Tonaufnahmen werden nacheinander abgespielt. Ihre Tochter/ Ihr Sohn soll währenddessen versuchen, verschiedene Merkmale der jeweiligen vorlesenden Personen „herauszuhören“ bzw. zu identifizieren und nach jedem Stimmbeispiel einen entsprechenden von uns entwickelten Fragebogen ausfüllen. Zu identifizierende Merkmale sind beispielweise ungefähres Alter, Dialekt, Ernährungszustand (sehr dünn bis sehr dick), Bildungsstand, etc. Die Durchführung der Befragung wird insgesamt maximal eine Schulstunde, also 45 Minuten, in Anspruch nehmen.

2. Einwilligungserklärung

Hiermit erkläre ich meine Einwilligung zur Teilnahme meiner Tochter/ meines Sohnes an der genannten Studie gemäß der vorstehenden Information. Durch meine Unterschrift bestätige ich, dass ich den Inhalt der Information gelesen und verstanden habe. Ich hatte Gelegenheit, Fragen zu stellen, die mir zufriedenstellend beantwortet wurden. Ich weiß, dass die Teilnahme freiwillig ist und ich meine Einwilligung jederzeit ohne Angabe von Gründen widerrufen kann, ohne dass mir daraus irgendwelche Nachteile entstehen.

Muttersprache der Studienteilnehmerin/ des Studienteilnehmers

Name Studienteilnehmer*in in Druckbuchstaben

Geburtsdatum Studienteilnehmer*in

Ort, Datum

Unterschrift Eltern (beider Elternteile) / Erziehungsberechtigte

Universitätsklinikum Bonn
Prof. Dr. med. Götz Schade
Leitung der Abteilung für Phoniatrie und Pädaudiologie
Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde/ Chirurgie
Doktorandin Johanna Thiele | thiele.hennef@hotmail.de

Betreff: Information der Schülerinnen und Schüler zur geplanten Studie (Studie zur altersspezifischen forensischen Stimmerkennung)

Liebe Schülerin, lieber Schüler,

wir möchten Dich fragen, ob Du einverstanden bist, an einer Befragung zur Erkennung von Stimmen teilzunehmen. Deine Teilnahme ist freiwillig und Du kannst sie jederzeit abbrechen, wenn Du nicht mehr an der Befragung teilnehmen möchtest. Gib dazu einfach uns oder Deinen Eltern/ Erziehungsberechtigten Bescheid.

Wir möchten Dir im Folgenden Informationen zu der Befragung geben.

Das Ziel der Befragung ist herauszufinden, ob es möglich ist, Menschen nur anhand ihrer Stimme zu beschreiben. Kann man beispielsweise anhand der Stimme „heraus hören“, wie alt eine Person ist, woher sie kommt etc.? Können Kinder/Jugendliche und Erwachsene dies vielleicht unterschiedlich gut „heraus hören“?

Zu diesem Zweck wäre es hilfreich, wenn Du an der Befragung teilnimmst.

Es werden Dir verschiedene Klangbeispiele von Stimmen vorgespielt. Alle tragen dieselbe kurze Geschichte vor. Du sollst dabei versuchen, Eigenschaften der Erzählerin/ des Erzählers am Klang der Stimme herauszufinden und einen entsprechenden Fragebogen ausfüllen.

Zur Unterstützung der Studie würde ich mich über Deine Hilfe freuen!



Johanna Thiele | Durchführung der Studie

Universitätsklinikum Bonn
 Prof. Dr. med. Götz Schade
 Leitung der Abteilung für Phoniatrie und Pädaudiologie
 Klinik und Poliklinik für Hals-Nasen-Ohrenheilkunde/ Chirurgie
 Doktorandin Johanna Thiele | thiele.hennef@hotmail.de

Datenerfassung: Vorlesende Personen

Name, Vorname _____

Straße _____

PLZ _____ Ort _____

Geburtsdatum _____

Geburtsland _____

Regionale Herkunft/sprachliche Besonderheiten (z. B. Dialekt, Akzent): _____

Körperlänge (in m) _____ Körpergewicht (in kg) _____ BMI* _____

Beruf _____

Schulischer Bildungsabschluss

Gymnasium/ Gesamtschule Realschule Hauptschule

Beruflicher absolvierter oder zu erwartender Bildungsabschluss

Hochschulabschluss (z.B. Universität, Fachschule) Berufsausbildung

Raucher/in

Ja Nein _____

Persönlichkeit (Selbsteinschätzung)

sehr extrovertiert eher extrovertiert eher introvertiert sehr introvertiert

Äußere Erscheinung (Selbsteinschätzung)

einfach/schlicht alltäglich/normal vornehm/elegant

Ort, Datum

Unterschrift vorlesende Person

* Dieses Feld ist nicht von der vorlesenden Person auszufüllen.

7. Abbildungsverzeichnis

Abb. 1: Darstellung von Kehlkopf und inneren Kehlkopfmuskeln (nach Schüssler et al., 2017).....	8
Abb. 2: Boxplot der Punktzahl im Fragebogen (Prozent) sortiert nach den Probandengruppen (Grs. = Grundschüler; Gymn. = Gymnasiasten; Stud. = Studierende; Phon. = Phoniater).....	30
Abb. 3: Mittelwerte der ungültigen und fehlenden Werte sortiert nach den Probandengruppen.....	32
Abb. 4: Streudiagramm der Mittelwerte der geschätzten und tatsächlichen Altersgruppen der Sprecher. Die durchgezogene Linie stellt dar, wo geschätztes und korrektes Alter übereinstimmen	33
Abb. 5: Mittelwerte der Differenzen geschätzte Altersgruppe - tatsächliche Altersgruppe, sortiert nach Altersgruppen der Sprecher (Jahre) und Probandengruppen.....	34
Abb. 6: Boxplot der korrekten Alterszuordnungen (Prozent) sortiert nach den Probandengruppen.....	35
Abb. 7: Liniendiagramm der korrekten Alterszuordnungen (Prozent) sortiert nach Altersgruppen der Sprecher (Jahre) und Probandengruppen.....	36
Abb. 8: Boxplot der korrekten Größenzuordnungen (Prozent) sortiert nach den Probandengruppen.....	38
Abb. 9: korrekte Größenzuordnungen (Prozent) sortiert nach Größe (cm) und Geschlecht der Sprecher.....	39
Abb. 10: Boxplot der korrekten BMI-Zuordnungen (Prozent) sortiert nach den Probandengruppen.....	41
Abb. 11: Korrekte BMI-Zuordnungen (Prozent) sortiert nach BMI-Klasse der Sprecher.....	42

Abb. 12: Boxplot der korrekten Angaben des Geschlechts (Prozent) sortiert nach den Probandengruppen.....	43
Abb. 13: Boxplot der korrekten Angaben von Hochdeutsch bzw. dem jeweiligen Akzent/Dialekt (Prozent) sortiert nach den Probandengruppen.....	45
Abb. 14: Korrekte Angaben von Hochdeutsch bzw. dem jeweiligen Akzent/Dialekt (Prozent) sortiert nach den Probandengruppen.....	46
Abb. 15: Häufigkeiten der Angabe Raucher (Prozent) nach Geschlecht der Sprecher, sortiert nach den Probandengruppen.....	50
Abb. 16: Boxplot der korrekten Angaben des Schulabschlusses (Prozent) sortiert nach den Probandengruppen.....	51
Abb. 17: Boxplot der korrekten Angaben der Berufsausbildung (Prozent) sortiert nach den Probandengruppen.....	53
Abb. 18: Boxplot der korrekten Angaben des Berufes (Prozent) sortiert nach den Probandengruppen.....	55

8. Tabellenverzeichnis

Tab. 1: Überblick über die Merkmalsverteilung der vorlesenden Personen.....	22–23
Tab. 2: Mittelwerte \pm Standardabweichungen der Differenzen geschätztes Alter - tatsächliches Alter sortiert nach Altersgruppen der Sprecher (Jahre) und Probanden- gruppen (Grs. = Grundschüler; Gymn. = Gymnasiasten; Stud. = Studierende; Phon. = Phoniater.....	34
Tab. 3: Korrekte Alterszuordnungen sowie Alterszuordnungen bei einem Fehler von ± 1 und ± 2 Altersgruppe(n) (Prozent) sortiert nach den Probandengruppen.....	35
Tab. 4: Korrekte Alterszuordnungen (Prozent) sortiert nach Altersgruppen der Sprecher (Jahre) und Probandengruppen.....	36
Tab. 5: Korrekte Größenzuordnungen sowie Größenzuordnungen bei einem Fehler von ± 1 und ± 2 Größenkategorie(n) (Prozent) sortiert nach den Probandengruppen.....	37
Tab. 6: Korrekte Größenzuordnungen (Prozent) sortiert nach dem Alter der Sprecher (Jahre) und den Probandengruppen.....	38
Tab. 7: Mittelwerte \pm Standardabweichungen der Differenzen geschätzte Größe - tatsächliche Größe sortiert nach Größenklassen der Sprecher (cm) und Probandengruppen.....	40
Tab. 8: Korrekte BMI-Zuordnungen sowie BMI-Zuordnungen bei einem Fehler von ± 1 und ± 2 BMI-Klasse(n) (Prozent) sortiert nach den Probandengruppen.....	41
Tab. 9: Mittelwerte \pm Standardabweichungen der Differenzen geschätzter BMI - tatsächlicher BMI sortiert nach BMI-Klassen der Sprecher und Probandengruppen.....	42
Tab. 10: Korrekte Angaben des Geschlechts (Prozent) sortiert nach Alter (Jahre) und Geschlecht der Sprecher sowie den Probandengruppen.....	44

Tab. 11: Korrekte Angaben von Hochdeutsch bzw. dem jeweiligen Akzent/Dialekt (Prozent) sortiert nach den Probandengruppen.....	46
Tab. 12: Korrekte Angaben von Akzenten/Dialekten (Prozent) sortiert nach den Probandengruppen.....	47
Tab. 13: Zusammensetzung der Angaben zu Akzent/Dialekt/Hochdeutsch (Prozent), sortiert nach den Probandengruppen.....	47
Tab. 14: Angaben der Probanden über das Rauchverhalten der Sprecher (Prozent) in Abhängigkeit von ihrem tatsächlichen Rauchverhalten, sortiert nach den Probandengruppen.....	49
Tab. 15: Angaben der Probanden zum Schulabschluss der Sprecher (Prozent) in Abhängigkeit von ihrem tatsächlichen Schulabschluss, sortiert nach den Probandengruppen.....	52
Tab. 16: Angaben der Probanden zur Berufsausbildung der Sprecher (Prozent) in Abhängigkeit von ihrer tatsächlichen Ausbildung, sortiert nach den Probandengruppen.....	54
Tab. 17: Angaben der Probanden zum Beruf der Sprecher (Prozent) in Abhängigkeit von ihrem tatsächlichen Beruf, sortiert nach den Probandengruppen.....	56
Tab. 18: Angaben der Probanden zur Persönlichkeit der Sprecher (Prozent) in Abhängigkeit von ihrer Selbsteinschätzung, sortiert nach den Probandengruppen.....	57
Tab. 19: Angaben der Probanden zur äußeren Erscheinung der Sprecher (Prozent) in Abhängigkeit von ihrer Selbsteinschätzung, sortiert nach den Probandengruppen.....	58
Tab. 20: Angaben der Probanden zu den Stimmen der Sprecher (Prozent) in Abhängigkeit vom Geschlecht der Sprecher.....	58
Tab. 21: Angaben der Probanden zu den Stimmen der Sprecher (Prozent) in Abhängigkeit vom Alter der Sprecher (Jahre).....	59

9. Literaturverzeichnis

Amir O, Engel M, Shabtai E, Amir N. Identification of children's gender and age by listeners. *J Voice* 2012; 26: 313–321

Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter, 2015: Konsensbasierte (S2) Leitlinie zur Diagnostik, Therapie und Prävention von Übergewicht und Adipositas im Kindes- und Jugendalter. https://www.adipositas-gesellschaft.de/fileadmin/PDF/Leitlinien/AGA_S2_Leitlinie.pdf (Zugriffsdatum: 02.02.2020)

Arbeitsgemeinschaft Adipositas im Kindes- und Jugendalter, 2020: AGA - my BMI 4 Kids. <https://aga.adipositas-gesellschaft.de/mybmi4kids/> (Zugriffsdatum: 02.02.2020)

Aumüller G, Wennemuth G. Halsorgane. In: Aumüller G, Aust G, Conrad A, Engele J, Kirsch J, Maio G, Mayerhofer A, Mense S, Reißig D, Salvetter J, Schmidt W, Schmitz F, Schulte E, Spanel-Borowski K, Wennemuth G, Wolff W, Wurzinger LJ, Zilch H-G, Hrsg. *Duale Reihe Anatomie*. Stuttgart: Georg Thieme Verlag, 2020: 914–937

Bachorowski JA, Owren MJ. Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *J Acoust Soc Am* 1999; 106: 1054–1063

Baker S, Weinrich B, Bevington M, Schroth K, Schroeder E. The effect of task type on fundamental frequency in children. *Int J Pediatr Otorhinolaryngol* 2008; 72: 885–889

Baker W, Eddington D, Nay L. Dialect identification: The effects of region of origin and amount of experience. *Am Speech* 2009; 84: 48–71

Bennett S. Vowel formant frequency characteristics of preadolescent males and females. *J Acoust Soc Am* 1981; 69: 231–238

Bennett S, Montero-Diaz L. Children's perception of speaker sex. *J Phon* 1982; 10: 113–121

Biever DM, Bless DM. Vibratory characteristics of the vocal folds in young adult and geriatric women. *J Voice* 1989; 3: 120–131

Boliek CA, Hixon TJ, Watson PJ, Morgan WJ. Vocalization and breathing during the first year of life. *J Voice* 1996; 10: 1–22

Braun A. Age estimation by different listener groups. *Int J Speech Lang La* 1996; 3: 65–73

Braun A, Rietveld T. The influence of smoking habits on perceived age. In: Elenius K, Branderud P, Hrsg. *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Stockholm: KTH, 1995: 294–297

Broeders APA, van Amelsvoort AG. Lineup construction for forensic earwitness identification: A practical approach. In: Ohala JJ, Hrsg. *Proceedings of the XIVth International Congress of Phonetic Sciences*. Berkeley, Calif.: Univ. of California Department of Linguistics, 1999: 1373–1376

Bruckert L, Liénard J-S, Lacroix A, Kreutzer M, Leboucher G. Women use voice parameters to assess men's characteristics. *Proc Biol Sci* 2006; 273: 83–89

Cerrato L, Falcone M, Paoloni A. Subjective age estimation of telephonic voices. *Speech Commun* 2000; 31: 107–112

Clopper CG, Pisoni DB. Some acoustic cues for the perceptual categorization of American English regional dialects. *J Phon* 2004; 32: 111–140

Clopper CG, Pisoni DB. Free classification of regional dialects of American English. *J Phon* 2007; 35: 421–438

Collins SA, Missing C. Vocal and visual attractiveness are related in women. *Anim Behav* 2003; 65: 997–1004

Cunningham-Andersson U, Engstrand O. Perceived strength and identity of foreign accent in Swedish. *Phonetica* 1989; 46: 138–154

Daly N, Warren P. Pitching it differently in New Zealand English: Speaker sex and intonation patterns. *J Socioling* 2001; 5: 85–96

Deal LV, Oyer HJ. Ratings of vocal pleasantness and the aging process. *Folia Phoniatri (Basel)* 1991; 43: 44–48

Deutsche Adipositas-Gesellschaft, 2014: Interdisziplinäre Leitlinie der Qualität S3 zur „Prävention und Therapie der Adipositas“. https://www.adipositas-gesellschaft.de/fileadmin/PDF/Leitlinien/S3_Adipositas_Praevention_Therapie_2014.pdf
(Zugriffsdatum: 02.02.2020)

Eichhorn JT, Kent RD, Austin D, Vorperian HK. Effects of Aging on Vocal Fundamental Frequency and Vowel Formants in Men and Women. *J Voice*; 32: 644.e1–644.e9

Eppley BD, Mueller PB. Chronological Age Judgments of Elderly Speakers: The Effects of Listeners' Age. *Contemp Issues Commun Sci Disord* 2001; 28: 5–8

Feinberg DR, Jones BC, Little AC, Burt DM, Perrett DI. Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Anim Behav* 2005; 69: 561–568

Fitch WT, Giedd J. Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *J Acoust Soc Am* 1999; 106: 1511–1522

Fried MP, Kelly JH, Strome M. Comparison of the adult and infant larynx. *J Fam Pract* 1982; 15: 557–558, 561

Geiger J. Hören und Sprechen: Kommunikation des Menschen. In: Pape H-C, Kurtz A, Silbernagl S, Hrsg. Physiologie. Stuttgart: Georg Thieme Verlag, 2019: 740–759

Girard F, Floccia C, Goslin J. Perception and awareness of accents in young children. *Br J Dev Psychol* 2008; 26: 409–433

Gonzalez J. Estimation of speakers' weight and height from speech: a re-analysis of data from multiple studies by Lass and colleagues. *Percept Mot Skills* 2003; 96: 297–304

González J. Formant frequencies and body size of speaker: a weak relationship in adult humans. *J Phon* 2004; 32: 277–287

Gunter CD, Manning WH. Listener estimations of speaker height and weight in unfiltered and filtered conditions. *J Phon* 1982; 10: 251–257

Haan J, van Heuven VJ. Male vs. female pitch range in dutch questions. In: Ohala JJ, Hrsg. Proceedings of the XIVth International Congress of Phonetic Sciences. Berkeley, Calif.: Univ. of California Department of Linguistics, 1999: 1581–1584

Hatley W, Samuel E, Evison G. The pattern of ossification in the laryngeal cartilages: A radiological study. *Br J Radiol* 1965; 38: 585–591

Hazan V, Markham D. Acoustic-phonetic correlates of talker intelligibility for adults and children. *J Acoust Soc Am* 2004; 116: 3108–3118

Hillenbrand JM, Clark MJ. The role of f_0 and formant frequencies in distinguishing the voices of men and women. *Atten Percept Psychophys* 2009; 71: 1150–1166

Hirano M, Kurita S, Sakaguchi S. Ageing of the vibratory tissue of human vocal folds. *Acta Otolaryngol* 1989; 107: 428–433

Hirano M, Kurita S, Toh Y. Growth, development and aging of the vocal fold. *Pract Otol* 1981; 74: 1791–1802

Honjo I, Isshiki N. Laryngoscopic and voice characteristics of aged persons. *Arch Otolaryngol* 1980; 106: 149–150

Hudson AI, Holbrook A. Fundamental frequency characteristics of young Black adults: spontaneous speaking and oral reading. *J Speech Hear Res* 1982; 25: 25–28

Hundt M, Palliwoda N, Schröder S. *Der deutsche Sprachraum aus der Sicht linguistischer Laien. Ergebnisse des Kieler DFG-Projektes.* Berlin, Boston: De Gruyter, 2017

Hunt RG, Lin TK. Accuracy of judgments of personal attributes from speech. *J Pers Soc Psychol* 1967; 6: 450–453

Huntley R, Hollien H, Shipp T. Influences of listener characteristics on perceived age estimations. *J Voice* 1987; 1: 49–52

Imhof M. Listening to Voices and Judging People. *Int J List* 2010; 24: 19–33

International Phonetic Association. *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet.* Cambridge: Univ. Press, 1999

Jessen M. Speaker classification in forensic phonetics and acoustics. In: Müller C, Hrsg. *Speaker classification I. Fundamentals, features, and methods.* Berlin, Heidelberg: Springer, 2007: 180–204

Jessen M. Forensic Phonetics. *Lang Linguist Compass* 2008; 2: 671–711

Jiang J, Lin E, Hanson DG. Vocal fold physiology. *Otolaryngol Clin North Am* 2000; 33: 699–718

Kahane JC. Growth of the human prepubertal and pubertal larynx. *J Speech Hear Res* 1982; 25: 446–455

Kahane JC. Connective tissue changes in the larynx and their effects on voice. *J Voice* 1987; 1: 27–30

Klatt DH, Klatt LC. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *J Acoust Soc Am* 1990; 87: 820–857

Kreiman J, Sidtis D. *Foundations of Voice Studies*. Oxford, UK: Wiley-Blackwell, 2011

Kromeyer-Hauschild K, Wabitsch M, Kunze D, Geller F, Geiß HC, Hesse V, Hippel A von, Jaeger U, Johnsen D, Korte W, Menner K, Müller G, Müller JM, Niemann-Pilatus A, Remer T, Schaefer F, Wittchen H-U, Zabransky S, Zellner K, Ziegler A, Hebebrand J. Perzentile für den Body-mass-Index für das Kindes- und Jugendalter unter Heranziehung verschiedener deutscher Stichproben. *Monatsschr Kinderheilkd* 2001; 149: 807–818

Künzel HJ. How well does average fundamental frequency correlate with speaker height and weight? *Phonetica* 1989; 46: 117–125

Kurita S. Growth, development and aging of the vocal fold. *Japan J Logop Phoniatr* 1988; 29: 185–193

Lass NJ, Hughes KR, Bowyer MD, Waters LT, Bourne VT. Speaker sex identification from voiced, whispered, and filtered isolated vowels. *J Acoust Soc Am* 1976; 59: 675–678

Lass NJ, Tecca JE, Mancuso RA, Black WI. The effect of phonetic complexity on speaker race and sex identifications. *J Phon* 1979; 7: 105–118

Linville SE. *Vocal aging*. San Diego: Singular Thomson Learning, 2001

Linville SE, Fisher HB. Acoustic characteristics of perceived versus actual vocal age in controlled phonation by adult females. *J Acoust Soc Am* 1985; 78: 40–48

Major RC. Identifying a foreign accent in an unfamiliar language. *Stud Second Lang Acquis* 2007; 29: 283

Martins RHG, Gonçalves TM, Pessin ABB, Branco A. Aging voice: Presbyphonia. *Aging Clin Exp Res* 2014; 26: 1–5

McGlone RE, Hicks JW. Speaker identification from photographs. *J Acoust Soc Am* 1979; 65: 116

Mead J, Bouhuys A, Proctor DF. Mechanisms generating subglottic pressure. *Ann N Y Acad Sci* 1968; 155: 177–182

Moyse E. Age Estimation from Faces and Voices: A Review. *Psychol Belg* 2014; 54: 255–265

Murphy CH, Doyle PC. The effects of cigarette smoking on voice-fundamental frequency. *Otolaryngol Head Neck Surg* 1987; 97: 376–380

Neiman GS, Applegate JA. Accuracy of listener judgments of perceived age relative to chronological age in adults. *Folia Phoniatr (Basel)* 1990; 42: 327–330

Owren MJ, Berkowitz M, Bachorowski J-A. Listeners judge talker sex more efficiently from male than from female vowels. *Percept Psychophys* 2007; 69: 930–941

Perry TL, Ohde RN, Ashmead DH. The acoustic bases for gender identification from children's voices. *J Acoust Soc Am* 2001; 109: 2988–2998

Pisanski K, Fraccaro PJ, Tigue CC, O'Connor JJM, Röder S, Andrews PW, Fink B, DeBruine LM, Jones BC, Feinberg DR. Vocal indicators of body size in men and women: a meta-analysis. *Anim Behav* 2014; 95: 89–99

Pisanski K, Rendall D. The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *J Acoust Soc Am* 2011; 129: 2201–2212

Prakash M, Johnny JC. Whats special in a child's larynx? *J Pharm Bioallied Sci* 2015; 7: 55–58

Preston DR. Perceptual Dialectology. In: Boberg C, Nerbonne J, Watt D, Hrsg. *The Handbook of Dialectology*: Wiley-Blackwell, 2017: 177–203

Rammstedt B. Grundlegende Kompetenzen Erwachsener im internationalen Vergleich. Ergebnisse von PIAAC 2012. Münster: Waxmann Verlag, 2013

Rendall D, Vokey JR, Nemeth C. Lifting the curtain on the Wizard of Oz: biased voice-based impressions of speaker size. *J Exp Psychol Hum Percept Perform* 2007; 33: 1208–1219

Ryan EB, Capadano HL. Age perceptions and evaluative reactions toward adult speakers. *J Gerontol* 1978; 33: 98–102

Sapienza CM, Ruddy BH, Baker S. Laryngeal Structure and Function in the Pediatric Larynx. *Lang Speech Hear Serv Sch* 2004; 35: 299–307

Sato T, Tauchi H. Age changes in human vocal muscle. *Mech Ageing Dev* 1982; 18: 67–74

Schade G, Hess M, Bubenheim M, Berger J. Gibt es ein einheitliches Vorgehen bei der Einteilung der Stimmlippen in Drittel? *Laryngorhinootologie* 2004a; 83: 236–242

Schade G, Leuwer R, Kraas M, Rassow B, Hess MM. Laryngeal morphometry with a new laser 'clip on' device. *Lasers Surg Med* 2004b; 34: 363–367

Scherer KR. Judging personality from voice: a cross-cultural approach to an old issue in interpersonal perception. *J Pers* 1972; 40: 191–210

Schötz S. Effects of stimulus duration and type on perception of female and male speaker age. *Proceedings of Fonetik* 2005; 87–90

Schünke M, Schulte E, Schumacher U, Voll M, Wesker KH. Atmungssystem. In: Schünke M, Schulte E, Schumacher U, Voll M, Wesker KH, Hrsg. *Prometheus Innere Organe*. Stuttgart: Georg Thieme Verlag, 2018: 32–37

Schüssler J, Bonkowsky V, Strutz J. Untersuchung des Larynx und der Trachea. In: Strutz J, Mann WJ, Hrsg. *Praxis der HNO-Heilkunde, Kopf- und Halschirurgie*. Stuttgart: Georg Thieme Verlag, 2017: 104

Schutte HK, Seidner W. Physiologische Grundlagen. In: Wendler J, Seidner W, Eysholdt U. *Lehrbuch der Phoniatrie und Pädaudiologie*. Stuttgart: Georg Thieme Verlag, 2015: 71–90

Sharma G, Goodwin J. Effect of aging on respiratory system physiology and immunology. *Clin Interv Aging* 2006; 1: 253–260

Shipp T, Hollien H. Perception of the aging male voice. *J Speech Hear Res* 1969; 12: 703–710

Shipp T, Qi Y, Huntley R, Hollien H. Acoustic and temporal correlates of perceived age. *J Voice* 1992; 6: 211–216

Siebel VS. Physiologie. In: Remmert S, Hrsg. *Funktionelle Wiederherstellung der oberen Luft- und Speisewege*. Stuttgart, New York: Georg Thieme Verlag, 2018: 39–46

Smith DRR, Walters TC, Patterson RD. Discrimination of speaker sex and size when glottal-pulse rate and vocal-tract length are controlled. *J Acoust Soc Am* 2007; 122: 3628–3639

Södersten M, Hertegård S, Hammarberg B. Glottal closure, transglottal airflow, and voice quality in healthy middle-aged women. *J Voice* 1995; 9: 182–197

Sorenson DN. A fundamental frequency investigation of children ages 6–10 years old. *J Commun Disord* 1989; 22: 115–123

Statistisches Bundesamt, 2017: Rauchgewohnheiten nach Altersgruppen und Geschlecht. <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Gesundheitszustand-Relevantes-Verhalten/Tabellen/liste-rauchverhalten.html> (Zugriffsdatum: 11.06.2020)

Titze IR. On the relation between subglottal pressure and fundamental frequency in phonation. *J Acoust Soc Am* 1989; 85: 901–906

Traunmüller H, van Bezooijen R. The auditory perception of children's age and sex. *Third International Conference on Spoken Language Processing* 1994; 1171–1174

van Bezooijen R, Gooskens C. Identification of Language Varieties. *J Lang Soc Psychol* 1999; 18: 31–48

van den Berg J. Myoelastic-aerodynamic theory of voice production. *J Speech Hear Res* 1958; 1: 227–244

van Dommelen WA, Moxness BH. Acoustic parameters in speaker height and weight identification: Sex-specific behaviour. *Lang Speech* 1995; 38: 267–287

Vorperian HK, Kent RD. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *J Speech Lang Hear Res* 2007; 50: 1510–1545

Vorperian HK, Kent RD, Lindstrom MJ, Kalina CM, Gentry LR, Yandell BS. Development of vocal tract length during early childhood: a magnetic resonance imaging study. *J Acoust Soc Am* 2005; 117: 338–350

Wagner L, Clopper CG, Pate JK. Children's perception of dialect variation. *J Child Lang* 2014; 41: 1062–1084

Wallner LJ. Smoker's larynx. *Laryngoscope* 1954; 64: 259–270

Whiteside SP. Temporal-based acoustic-phonetic patterns in read speech: some evidence for speaker sex differences. *J Int Phon Assoc* 1996; 26: 23–40

Whiteside SP. Identification of a speaker's sex: a study of vowels. *Percept Mot Skills* 1998; 86: 579–584

Winkler R. Influence of smoking on age estimation from voice. In: Fuchs S, Hrsg. *Between the regular and the particular in speech and language*. Frankfurt am Main: Lang, 2010: 77–96

World Health Organization. *Obesity - Preventing and Managing the Global Epidemic*. Geneva: World Health Organization, 2000