



universität
wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

Zur Perzeption emotionaler Sprechweisen durch
prosodische Merkmale im Deutschen und Italienischen

verfasst von / submitted by

Mag. Barbara Ladurner, BA MA

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of
Doktorin der Philosophie (Dr. phil.)

Wien, 2020 / Vienna 2020

Studienkennzahl lt. Studienblatt /
degree programme code as it appears on the stu-
dent record sheet:

A 792 332

Dissertationsgebiet lt. Studienblatt /
field of study as it appears on the student record
sheet:

Deutsche Philologie

Betreut von / Supervisor:

Univ.-Prof. Mag. Dr. Peter Ernst

Vorwort

Die mündliche Kommunikation insgesamt, die Wahrnehmung von Emotionen im sprachlichen Kontext und die besondere Rolle der Prosodie im Perzeptionsprozess gesprochener Sprache sind in den letzten Jahren aufgrund meiner eigenen interdisziplinären Ausbildung und Tätigkeit immer stärker in den Fokus meiner Aufmerksamkeit gerückt. Bald war mir klar, dass ich die Hörer*innen selbst und ihren subjektiven Höreindruck ins Zentrum meiner Forschung stellen wollte – nicht nur, weil sie in bisherigen Untersuchungen zur Mündlichkeit oft eher stiefmütterlich behandelt worden waren, sondern vielmehr deswegen, weil in der Alltagskommunikation in den meisten Fällen doch die ganz „normale“ Hörer*in emotionale Äußerungen wahrnimmt und interpretiert. Dass in diesem Prozess allerhand Missverständnisse auftreten können, war mir als Südtirolerin und Sprachlehrerin für Deutsch und Italienisch aus eigener Erfahrung klar. Und als „Mensch“ hatte ich schon oft die Erfahrung gemacht, dass es häufig nicht die Worte sind, die den Ausschlag geben, sondern dass „der Ton die Musik“ macht. Diesen Aspekt finde ich als Sängerin und Alumna der Sprechwissenschaft besonders spannend. Daher kristallisierte sich nach eingehender Literaturrecherche und einigem Kopfzerbrechen bald „mein“ Thema heraus – ich wollte die Perzeption natürlicher emotionaler Sprechweisen (in den Ausprägungen Freude und Ärger) anhand prosodischer Merkmale untersuchen, um der Bedeutung dieser „Zauberelemente“ näher auf die Spur zu kommen. Dabei interessierte mich zudem die Frage, ob es wohl einen Unterschied machen würde, ob deutsch- oder italienischsprachige Personen, Männer oder Frauen sowie Laien oder Stimmexpert*innen die emotionalen Äußerungen wahrnehmen und beschreiben.

Dank der großartigen Unterstützung vieler Personen durfte ich diesen Fragen in einem mehrjährigen Projekt auf den Grund gehen. Ich freue mich über diese Möglichkeit, die ich sehr zu schätzen weiß, und möchte mich von Herzen bei all jenen bedanken, die mich in diesem Prozess begleitet haben.

Zunächst gilt mein herzlicher Dank „meinem“ Professor, Univ.-Prof. Mag. Dr. Peter Ernst der Universität Wien, der mich als Doktorandin stets mit größter Professionalität und Herzlichkeit betreut hat. Sein Einsatz, seine raschen und immer freundlichen Antworten sind alles andere als selbstverständlich und seine Unterstützung weiß ich entsprechend zu schätzen.

Ein großer Dank gebührt auch Prof. Dr. phil. habil. Dr. h.c Ursula Hirschfeld der Universität Halle-Wittenberg für die wertvollen fachlichen Hilfestellungen und die Möglichkeit der Teilnahme an Doktoranden-Kolloquien zum fachlichen Austausch und zur Vernetzung.

Weiters möchte ich mich bei Professor Paolo Mastandrea und bei Dr. Antonio Pistellato des Dipartimento di Studi Umanistici der Universität Ca`Foscari in Venedig sowie bei Stefano Gibellato, Dozent am Konservatorium Venedig, für die äußerst freundliche und tatkräftige Unterstützung bei der Durchführung der empirischen Studien mit den italienischsprachigen Proband*innen bedanken.

Dr. Brigitte Teuchert, Studiengangsleiterin an der Universität Regensburg, sei ein größtes Dankeschön für viele fachliche Anregungen und für die Unterstützung bei der Durchführung der empirischen Untersuchung ausgesprochen. Auch bei PD Dr. Wieland Kranich der Universität Regensburg möchte ich mich aufrichtig für die fachlichen Hinweise bedanken.

Des Weiteren gilt meiner Schwester Maria Ladurner, MA und Simon Wieser mein größter Dank für die äußerst wertvolle Hilfe bei der statistischen Auswertung der Daten.

Mein herzlicher Dank geht auch an Christian Cermak für die technische Hilfestellung bei der Aufbereitung des Audiomaterials.

All meinen Studienkolleg*innen, die mich inspiriert und motiviert haben, sei an dieser Stelle auch gedankt.

Bei meiner Familie, allen voran bei meinem Partner, möchte ich mich für die Geduld und die Unterstützung im Entstehungsprozess der Arbeit bedanken.

Vielen Dank auch der Universität Wien und dem OeAD für die Gewährung der Auslandsstipendien sowie der Provinz Bozen, Abteilung Bildungsförderung, für die finanzielle Unterstützung zur Durchführung dieses Forschungsvorhabens.

Nicht zuletzt gilt mein herzliches Dankeschön allen Probandinnen und Probanden, die sich für die aufwändigen und langwierigen Befragungen und Analysen zur Verfügung gestellt haben, ohne sie hätte die Durchführung der Arbeit nicht gelingen können.

Inhalt

1 Einleitung	8
1.1 Thema und Forschungsfrage.....	8
1.2 Methodik.....	11
1.3 Aufbau der Arbeit.....	15
2 Theoretische Grundlagen und Forschungssituation	16
2.1 Prosodie und Sprechausdruck	16
2.1.1 Definition Prosodie	16
2.1.2 Prosodische Merkmale	23
2.1.2.1 Akzent.....	23
2.1.2.2 Rhythmus.....	25
2.1.2.3 Tonhöhe und Melodieführung	27
2.1.2.4 Lautstärke	29
2.1.2.5 Sprechgeschwindigkeit.....	30
2.1.2.6 Dauer	31
2.1.2.7 Pausen.....	32
2.1.2.8 Stimmfärbung (Timbre).....	33
2.1.3 Sprechausdruck.....	36
2.2 Prosodie und Emotion	38
2.2.1 Begriff Emotion	38
2.2.2 Emotionstheorien: universell vs. kulturell	40
2.2.3 Theoretische Konzepte und Modelle	41
2.2.3.1 Scherer: Integrative Theorie	42
2.2.3.2 Linguistische Ansätze.....	43
2.2.3.3 Klassifikation von Emotionen.....	44
2.2.3.4 Das Linsenmodell von Brunswik in einer Adaption von Scherer.....	45
2.2.4 Forschungsüberblick Prosodie, Emotion, Perzeption	48
2.3 Prosodie und Sprachperzeption.....	53
2.3.1 Perzeptive Phonetik	53
2.3.2 Theorien zur Sprachperzeption.....	55

2.3.2.1 Bottom-up-Ansatz der Verstehenstheorie	55
2.3.2.2 Top-down-Ansatz der Verstehenstheorie	56
2.3.2.3 Sprachperzeption als aktiver Prozess	58
2.3.3 Modelle zur Sprachperzeption.....	59
2.3.3.1 Die Motor-Theorie der Sprachperzeption.....	59
2.3.3.2 Auditive Sprachperzeptionstheorien	62
2.4 Zur Perzeption prosodischer Merkmale	63
2.4.1 Wahrnehmung von Rhythmus	63
2.4.2 Wahrnehmung von Akzenten	64
2.4.3 Wahrnehmung von Tonhöhe/Melodieführung	65
2.4.4 Wahrnehmung von Lautstärke.....	66
2.4.5 Wahrnehmung von Sprechgeschwindigkeit und Pausen	67
2.4.6 Wahrnehmung von Stimmfärbung.....	68
2.4.7 Allgemeine Probleme bei Untersuchungen der Prosodieperzeption	69
2.5 Einflussfaktoren auf die Prosodieperzeption	71
2.5.1 Einflussfaktor Kultur.....	71
2.5.1.1 Kulturbegriff	72
2.5.1.2 Interkulturelle Kommunikation.....	73
2.5.1.3 IK und Prosodie	74
2.5.2 Einflussfaktor Sprache	75
2.5.3 Einflussfaktor Geschlecht.....	79
2.5.4 Einflussfaktor Beruf.....	80
3 Empirische Untersuchung.....	82
3.1 Forschungsdesign.....	82
3.1.1 Fragestellung und Zielsetzung.....	82
3.1.2 Methodenkritik	83
3.1.3 Hypothesen	87
3.2 Untersuchungsaufbau und –durchführung.....	89
3.2.1 Aufbereitung und Selektion des Audiomaterials.....	89
3.2.2 Fragebogenerhebung.....	96

3.3 Ergebnisse	98
3.3.1 Darstellung der für die Emotionserkennung zentralen prosodischen Merkmale	98
3.3.1.1 Gesamtüberblick aller Antworten	99
3.3.1.2 Deutsch- und italienischsprachige Proband*innen	100
3.3.1.3 Laien und Stimmexpert*innen	102
3.3.1.4 Deutsch- und italienischsprachige Laien und Stimmexpert*innen	103
3.3.1.5 Männer und Frauen	105
3.3.1.6 Deutsch- und italienischsprachige Männer und Frauen	107
3.3.1.7 Ergebnisse in Bezug auf die freudigen Sprechweisen	109
3.3.1.8 Ergebnisse in Bezug auf die Sprechweisen zu Ärger	113
3.3.1.9 Gegenüberstellung Freude – Ärger	117
3.3.2 Relevanz der prosodischen Merkmale in der Perception der emotionalen Sprechweisen	118
3.3.2.1 Gesamtüberblick aller Antworten	119
3.3.2.2 Deutschsprachige Proband*innen	120
3.3.2.3 Italienischsprachige Proband*innen	122
3.3.2.4. Alle Laien	124
3.3.2.5 Alle Stimmexpert*innen	125
3.3.2.6 Deutsch- und italienischsprachige Laien und Stimmexpert*innen	127
3.3.2.7 Männer	128
3.3.2.8 Frauen	130
3.3.2.9 Deutsch- und italienischsprachige Männer und Frauen	132
3.3.2.10 Ergebnisse in Bezug auf die freudigen Sprechweisen	134
3.3.2.11 Ergebnisse in Bezug auf die Sprechweisen zu Ärger	137
3.3.2.12 Gegenüberstellung Freude – Ärger	140
3.3.3 Ausprägungsgrad der jeweiligen Emotion	141
3.3.4 Perception und Analyse der prosodischen Merkmale	146
3.3.4.1 Perception und Analyse der prosodischen Merkmale – Freude	148
3.3.4.1.1 Akzent/Rhythmus	149
3.3.4.1.2 Tonhöhe/Melodieführung	153
3.3.4.1.3 Lautstärke	157
3.3.4.1.4 Sprechgeschwindigkeit	161
3.3.4.1.5 Pausen	165
3.3.4.1.6 Stimmfärbung (Timbre)	168
3.3.4.1.7 Gesamtergebnis zur Perception und Analyse der prosodischen Merkmale – Freude	175
3.3.4.2 Perception und Analyse der prosodischen Merkmale – Ärger	177

3.3.4.2.1 Akzent/Rhythmus.....	177
3.3.4.2.2 Tonhöhe/Melodieführung	180
3.3.4.2.3 Lautstärke.....	184
3.3.4.2.4 Sprechgeschwindigkeit	187
3.3.4.2.5 Pausen.....	191
3.3.4.2.6 Stimmfärbung (Timbre).....	193
3.3.4.2.7 Gesamtergebnis zur Perzeption und Analyse der prosodischen Merkmale – Ärger.....	198
3.3.4.3 Gegenüberstellung Freude – Ärger	200
3.3.4.4 Anmerkungen (optional).....	208
4 Ergebnisinterpretation und Hypothesenbeantwortung	210
4.1 Forschungsfrage 1	210
4.1.2 Akzent/Rhythmus	212
4.2 Forschungsfrage 2	231
4.3 Forschungsfrage 3	250
4.4 Forschungsfrage 4	251
4.5 Hypothesenbeantwortung	254
4.6 Zusammenfassung	259
5 Literaturverzeichnis	261
Anhang.....	284
Vorstudie: Hörbeispiele.....	284
Tabelle 1: Auswertung Hörprobe 1, Vorstudie, deutsche Proband*innen	294
Tabelle 2: Auswertung Hörprobe 2, Vorstudie, italienische Proband*innen	296
Fragebogen 1: Vorstudie, deutsche Proband*innen.....	298
Fragebogen 2: Vorstudie, italienische Proband*innen.....	302
Fragebogen 3: Untersuchung zur Perzeption emotionaler Sprechweisen durch Prosodie – Deutsch.....	306
Fragebogen 4: Untersuchung zur Perzeption emotionaler Sprechweisen durch Prosodie – Italienisch.....	314
Abstract.....	322

1 Einleitung

In den folgenden Kapiteln soll in das Thema eingeführt werden, wobei die zentralen Forschungsfragen und wichtigsten Hypothesen überblicksmäßig dargestellt werden. Auch die methodische Vorgehensweise wird kurz erläutert. Eine ausführlichere Beschreibung des Forschungsdesigns sowie der Fragestellungen erfolgt zu Beginn des zweiten Teils der Arbeit.

1.1 Thema und Forschungsfrage

Die menschliche Kommunikation umfasst die Fähigkeit des Sprechens sowie jene der Dekodierung sprachlicher Information auf der Hörer*innenseite. Dabei stehen prosodische Eigenschaften im Mittelpunkt – in der mündlichen Kommunikation wird sprachliche Information mittels der prosodischen Parameter Tonhöhenverlauf, Lautheit, Dauer, Akzent und Rhythmus, Sprechgeschwindigkeit und Pausen übertragen (Alter 2003: 339). Die Klangfarbe der Stimme (Timbre) zeigt Gefühle und körperliche Zustände des Sprechers an und erfüllt zusammen mit den anderen prosodischen Parametern relevante rhetorische (Eindringlichkeit, Emphase), beziehungsregulierende und emotionsexpressive Funktionen. Wie etwas gesagt wird, ist oft entscheidender als die inhaltliche Mitteilung (Schwitalla 2012: 82).

Aus diesem Grund soll der Zusammenhang von Emotion und Prosodie untersucht werden, wobei der Fokus auf die Perzeption emotionaler Sprechweisen durch prosodische Merkmale, also auf die Hörer*innenseite, gelegt wird. Schließlich ist es im Normalfall die phonetisch ungeübte Hörer*in, die in der Alltagskommunikation emotionale Sprechweisen wahrnimmt, decodiert und versteht (vgl. Wendt 2007). Anhand einer empirischen Studie mit je deutschen und italienischen Hörbeispielen und Fragebögen, die neben der theoretischen Basis den Kern dieser Forschungsarbeit darstellen wird, soll herausgefunden werden, welche Relevanz prosodische Merkmale für die Perzeption emotionaler Äußerungen haben und wie sie von den einzelnen Hörer*innen rezipiert werden. Dabei soll ergründet werden, inwieweit sich sprach-/kulturspezifische, geschlechterspezifische und in Relation zur individuellen Fähigkeit des funktionellen Hörens (durch eine bestimmte Ausbildung bzw. berufliche Tätigkeit) stehende Differenzen auf die Perzeption der prosodischen Merkmale auswirken. Gearbeitet werden soll mit zwei Basemotionen, nämlich Ärger und Freude, und je zwei männlichen und zwei weiblichen Sprecher*innen, je Italienisch und Deutsch. Es handelt sich bei den Hörbeispielen um kurze Äußerungen, da sich emotionale Sprechweisen und emotionale Ausprägungen schnell ändern. Ziel

der Arbeit ist es, Aussagen über die Perzeption natürlicher Sprechweisen machen zu können, daher wird mit entsprechenden Aufnahmen aus TV und Fernsehen (z. B. die Sendung „Verstehen Sie Spaß“, in der mit versteckter Kamera gedreht wird) gearbeitet, wobei aus Gründen der Vergleichbarkeit möglichst abgeschlossene Turns und syntaktisch ähnliche Sätze mit einer Länge von circa sieben bis zehn Wörtern verwendet werden.

In einem ersten Durchgang sollen im Rahmen der Vorstudie verschiedene Hörbeispiele von den Proband*innen den Emotionen Ärger und Freude zugeordnet werden. Dadurch werden die besten Aufnahmen selektiert und sichergestellt, dass bei der Hauptuntersuchung nur mit Sprachmaterial gearbeitet wird, welches eindeutig den entsprechenden Emotionen und den qualitativen Anforderungen an die Studie entspricht. Im zweiten Durchgang erfolgt dann die eigentliche Untersuchung der ausgewählten Hörbeispiele durch die Hörer*innen anhand eines Fragebogens mit dem Ziel, Aussagen über die individuelle Perzeption und Analyse der suprasegmentalen Elemente zu treffen. Folgende prosodische Merkmale sollen untersucht werden: Akzent/Rhythmus, Tonhöhe/Melodieführung, Lautstärke, Sprechgeschwindigkeit, Pausen und Stimmfärbung (Timbre). Anhand vier verschiedener Teilnehmergruppen sollen bei der Analyse geschlechter-, sprach-/kultur- und berufsspezifische Aspekte berücksichtigt werden. Die Proband*innengruppen setzen sich aus deutschen Muttersprachler*innen ohne und mit Stimmbetrieb (z. B. Sprecherzieher*in) sowie italienischen Muttersprachler*innen ohne und mit Stimmbetrieb zusammen. In den jeweiligen Gruppen wird zwischen Männern und Frauen differenziert, außerdem wird darauf geachtet, dass die Personen über keine Kenntnisse der jeweils anderen Sprache verfügen. Diese Differenzierung erlaubt es, oben genannte Punkte als Einflussfaktoren auf das Perzeptionsvermögen zu untersuchen und so zu einer umfassenden Darstellung der Emotionsperzeption durch Prosodie zu gelangen. Gearbeitet wird nur mit authentischem Sprachmaterial, also mit natürlichen Äußerungen, was eine relevante Neuerung im Vergleich zu früheren Forschungen darstellt (vgl. Schwitalla 2014).

Es wird davon ausgegangen, dass es dadurch sowie aufgrund des interdisziplinären Zugangs möglich ist, Muster der Emotionsperzeption mittels prosodischer Merkmale zu erkennen und zu beschreiben. Diesem Forschungsdesign liegt die Annahme zugrunde, dass die Perzeptionsfähigkeit von Emotionen weniger biologisch als vielmehr sozial bedingt ist (vgl. Féry 1993), von der die jeweilige Person umgebende Sprache und Kultur beeinflusst ist (vgl. Rabanus 2001) und auch von einer mehr oder weniger ausgeprägten Stimmaffinität der jeweiligen Person abhängig ist, wobei von den Stimmexpert*innen eine sensiblere Wahrnehmung der prosodischen

Parameter erwartet wird. Was die potentielle Funktion der prosodischen Variablen im Indizieren unterschiedlicher Emotionen betrifft, so wird grundsätzlich eine Zusammenwirkung aller prosodischen Merkmale mit verschiedenartiger Dominanz angenommen, d. h. die Versuchsteilnehmer*innen nehmen wahrscheinlich die einzelnen Emotionen ganzheitlich als komplexe Phänomene wahr (vgl. Kranich 2003, Kehrein 2002: 321). Besonders den Tonhöhe- und Timbreigenschaften wird jedoch eine besondere Bedeutung zugesprochen (vgl. Kranich 2003: 223). Im Folgenden sollen die Hypothesen zu den einzelnen prosodischen Merkmalen kurz genannt werden.

- Rhythmus/Akzente: eher weniger relevant aufgrund der Schwierigkeit diese Parameter an emotionalen Merkmalen festzumachen;
- Tonhöhe/Melodieführung: sehr relevant, v. a. bei Freude (hohe Sprechstimmlage, großer Tonumfang erwartet) (Kranich 2003: 198f.);
- Lautstärke: insgesamt etwas weniger relevant, wichtig jedoch evtl. bei Ärger (hohe Lautstärke);
- Sprechgeschwindigkeit: in Anbetracht aller Emotionen etwas weniger relevant, wichtig evtl. bei Ärger (hohe Sprechgeschwindigkeit) und Freude (hohe Sprechgeschwindigkeit);
- Pausen: in Bezug auf Ärger und Freude weniger relevant, evtl. wichtig bei Traurigkeit („geräuschvolle“ Pausen mit Seufzen, Atmen), Überraschung (geräuschvolles Einatmen);
- Stimmfärbung (Timbre): sehr relevant für alle Emotionen; Ärger (hell, scharf, klangarm, behaucht); Freude (hell, weich, warm, weit); (Kranich 2003: 198).

Natürlich tragen die prosodischen Merkmale nicht nur emotionale Information, sondern erfüllen auch sprachlich konventionalisierte Funktionen. Eine grundlegende Unterscheidung, die im Beschreibungsmodell von Prosodie erfasst werden soll, ist also die zwischen der phonetischen und der funktionalen, phonologischen Ebene: Auf der phonetischen Ebene werden die prosodischen Merkmale, auf der funktionalen Ebene prosodische Einheiten beschrieben (Kehrein 2002: 79). Es gilt in der Arbeit also auch die diversen Funktionen (z. B. grammatisch-syntaktische und gesprächsorganisierende) von Prosodie zu berücksichtigen und voneinander zu unterscheiden (vgl. Kehrein 2002: 226).

1.2 Methodik

Im zweiten Kapitel der Dissertation sollen zunächst das Thema theoretisch begründet und der Forschungsstand zusammengefasst und abgebildet werden. Der Hauptteil der Arbeit bildet eine empirische Untersuchung zu den oben dargelegten Fragestellungen. Die Ergebnisse sollen dabei anhand von Hörbeispielen und Fragebögen gewonnen werden. Eine detailliertere Beschreibung des methodischen Vorgehens findet sich im dritten Teil, Kapitel 3.1 „Forschungsdesign“.

Hörbeispiel

In dieser Forschungsarbeit soll mit natürlichen emotionalen Sprechweisen gearbeitet werden, weil sich wissenschaftliche Untersuchungen zur Kommunikation mit möglichst alltagsnahen und nicht-konstruierten Situationen beschäftigen sollten (vgl. Hielscher 2003; Schwitalla 2012;). Außerdem gewährt die Arbeit mit natürlichen Äußerungen eine authentische und erfolgreiche Perzeption der jeweiligen Emotion, wie auch eine Studie Kehreins zeigt, in der Aussagen des Lego-Korpus von drei Personen auf Emotionalität hin angehört bzw. analysiert wurden. Dabei wurden in keinem Fall einem Redebeitrag zwei unvereinbare emotionale Qualitäten zugeordnet. „Stellt man diesem Befund die niedrigen Erkennungsraten und häufigen Fehlinterpretationen bei der Einschätzung kontextfreier Stimuli in anderen Untersuchungen gegenüber, wird deutlich, dass es in wissenschaftlichen Untersuchungen keine Alternative zur Wahrnehmung von Emotionen *im Kontext* gibt.“ (Kehrein 2002: 227) Es handelt sich bei den Hörbeispielen also um kurze emotionale Sprechweisen, die den authentischen, alltäglichen Emotionen Ausdruck verleihen. Diese sind oft dadurch gekennzeichnet, dass sie wenig intensiv sind (kein „In-die-Luft-gehen“), aber im täglichen Leben häufig vorkommen. Die Wahrnehmung von Emotionen im Kontext stellt einen grundlegenden Fortschritt in der Emotionsausdrucksforschung dar (Kehrein 2002: 309). Die alltagssprachliche Emotionsattribution beruht auf einer ganzheitlichen Wahrnehmung des emotionalen Mitmenschen in einem sprachlichen und situativen Kontext, d. h. die Zeichen aller Signalisierungssysteme tragen zu der Wahrnehmung einer bestimmten Emotion bei (Kehrein 2002: 146). Das Konzept „Ganzheitliche Wahrnehmung“ geht zurück auf Fiehler (1990b: 54-56, 104-106), der von der ganzheitlichen Selbstwahrnehmung von Emotionen einerseits und von der Relevanz des Gesamtausdrucks bei der Zuschreibung von Emotionen andererseits ausgeht. Auch Watzlawick geht von der zentralen Annahme aus, dass das Material jeglicher Kommunikation keineswegs nur Worte, sondern auch alle paralinguistischen Phänomene (z. B. Tonfall, Lachen, Seufzen usw.), Körperhaltungen, Körpersprache usw. innerhalb eines bestimmten Kontextes umfasst (Watzlawick 1996(9): 51).

Um also mit natürlichen Äußerungen arbeiten zu können, ist möglichst authentisches Sprachmaterial vonnöten. Da dieses trotzdem von guter akustischer Qualität sein soll, wird auf TV-Aufnahmen aus Talkshows, Interviews und anderen Programmen (z. B. Sendung „Verstehen Sie Spaß“, in der mit versteckter Kamera gearbeitet wird) zurückgegriffen, in denen emotionale Äußerungen in guter akustischer Qualität vorliegen. Der Vorteil liegt darin, dass es sich trotz der öffentlichen Situation mit Publikum und/oder Kamera um natürliche Aussagen mit dem vorrangigen Handlungsziel der Mitteilungsäußerung und nicht um von Schauspieler*innen kreierte, vom Kontext losgerissene emotionale Äußerungen handelt. Damit kann der Nachteil, dass sich andere, zusätzliche Informationen (etwa durch die Lexik) nicht ganz ausschalten lassen, in Kauf genommen werden. Es ist darauf zu achten, dass die Videofunktion abgeschaltet und die lexikalische Information der ausgewählten Episode allein nicht aussagekräftig zur Bestimmung der Emotion ist, sodass die prosodischen Merkmale in den Vordergrund treten. Die selektierten Audioaufnahmen werden in einer Vorstudie einer Gruppe von Proband*innen vorgespielt, um deren Höreindrücke zur Emotionalität der Äußerungen zu sammeln. Erst in einem zweiten Schritt werden anhand der empirischen Studie detailliertere Informationen zur Wahrnehmung der Emotionalität durch Prosodie erfragt werden. Dieser Vorgang hat den Vorteil, dass eine subjektive Bewertung der einzelnen Aufnahmen bei der Fragebogen-Erstellung vermieden werden und induktiv ein Emotionalitätsraster erfasst werden kann. Um die Faktoren Geschlecht und Sprache/Kultur auch sprecherseitig zu berücksichtigen und entsprechend auszugleichen, wird folgende Konstellation der Sprecher*innen für sinnvoll erachtet:

	Emotion Freude	Emotion Ärger
Italienischsprachige Sprecher*innen	Mann / Frau	Mann / Frau
Deutschsprachige Sprecher*innen	Mann / Frau	Mann / Frau

Beide emotionale Sprechweisen zu Freude und Ärger werden von italienisch- und deutschsprachigen Sprecher*innen, je männlich und weiblich, realisiert. Die Emotionen Freude und Ärger wurden ausgewählt, da es sich hierbei um übernational gültige Basisemotionen handelt (z. B. Ekman 1992: 174; Johnson-Laird [u. a.] 1989: 105; Schmidt-Atzert 1987: 140; Scherer 1986b: 174).

Fragebogen

Die Fragebögen sollen so konzipiert sein, dass die Proband*innen zuerst einige relevante persönliche Daten ausfüllen (Alter, Geschlecht, Nationalität, Muttersprache, Dialektsprecher*in/Standardsprache, Beruf) und dann das Hörbeispiel abhören, das sie so oft anhören können, wie sie möchten. Anhand der Fragebogenerhebung sollen dann zum jeweiligen Hörbeispiel folgende Informationen erfragt werden:

- Welche Parameter werden zur Emotionserkennung verwendet?
- Wie wird die Relevanz der einzelnen prosodischen Merkmale für den Perzeptionsprozess eingeschätzt?
- Wie stark ausgeprägt ist die jeweilige Emotion für die Proband*innen?
- Wie werden die prosodischen Merkmale der freudigen bzw. ärgerlichen Sprechweisen perzipiert und beschrieben?
- Gibt es bei oben genannten Fragen sprach-/kulturbedingte, berufsbezogene oder geschlechtsspezifische Differenzen?

Dabei werden den Hörer*innen Fragen zu den genannten prosodischen Parametern gestellt, welche einerseits eine möglichst genaue Beschreibung des prosodischen Höreindrucks erlauben und andererseits die nicht auf mündliche Kommunikation spezialisierten Laien nicht überfordern. Daher wird auf eine adäquate Sprachverwendung geachtet. Außerdem gibt es die Möglichkeit, die Fragen auch mit „Weiß ich nicht“ zu beantworten, um falsche Ergebnisse und eine Überforderung der Teilnehmer*innen zu vermeiden. Von den phonetisch geschulten Expert*innen hingegen sind differenziertere Aussagen zu den Aufnahmen zu erwarten. Bei der Konzipierung dieser Fragestellungen zur Analyse der einzelnen prosodischen Merkmale im dritten Teil des Fragebogens wird der Katalog zur auditiven Merkmalsbeschreibung von Sprechausdrucksmustern von Ines Bose (Bose 2001) zur Orientierung herangezogen. In Anlehnung an den von Bose sehr detailliert beschriebenen möglichen Ausprägungen jedes prosodischen Parameters werden die wichtigsten Beschreibungskategorien derselben in stark komprimierender Weise neu zusammengesetzt. Damit erschöpfen sich natürlich keinesfalls die Beschreibungsmöglichkeiten der prosodischen Merkmale, die Reduktion ist vielmehr als notwendige Voraussetzung für das Gelingen der Analyse durch alle, v. a. auch in diesem Bereich nicht vorgebildeten Proband*innen anzusehen.

Auswertung

Die Fragebögen werden zunächst auf der Plattform UmfrageOnline und dann mithilfe des Statistikprogramms IBM SPSS ausgewertet und die daraus gewonnenen Daten schließlich anhand der dargestellten verschiedenen Gesichtspunkte miteinander verglichen. Bei der Auswahl der Proband*innen ist auch darauf zu achten, dass es sich um möglichst standardaffine Sprecher*innen handelt bzw. dass Dialektsprecher*innen als solche markiert werden. Dieser Aspekt sowie auch andere, etwa das Alter, spielen bei der Perzeption natürlich eine wichtige Rolle, können aber im Einzelnen in dieser Forschungsarbeit nicht berücksichtigt werden. Um größere altersbedingte Divergenzen auszuschließen, wird v. a. mit jüngeren Personen (Student*innen) gearbeitet, in der Gruppe der Stimmexpert*innen ist das Durchschnittsalter etwas höher.

Da das Augenmerk in dieser Studie auf die individuelle bzw. nach Gruppen differenzierte Wahrnehmung der emotionalen Sprechäußerungen liegt, wird auf eine phonetische Analyse der Hörbeispiele und eine Transkription, etwa nach den Notationskonventionen des „Gesprächsanalytischen Transkriptionssystems“ (GAT) (vgl. Selting et. al. 1998), verzichtet. Im Zentrum stehen die subjektiven Höreindrücke der Proband*innen in den diversen Referenzgruppen. Hier ist auch ein Vergleich mit den objektiven Daten zu den phonetischen und prosodischen Charakteristika von Emotionen interessant und durchaus aufschlussreich für das Verstehen der Perzeption emotionaler Prosodie. Außerdem wird mit einer erfahrenen Sprechwissenschaftlerin gearbeitet, welche die Sprechweisen als Referenzhörer*in auditiv analysiert.

Die für die Studie relevanten Vergleichsgruppen sind folgende:

- Deutsche Muttersprachler*innen
- Deutsche Muttersprachler*innen mit einem stimmaffinen/sprecherischen Beruf (z. B. Sprecherzieher*in)
- Italienische Muttersprachler*innen ohne Deutschkenntnisse
- Italienische Muttersprachler*innen mit einem stimmaffinen/sprecherischen Beruf (z. B. Sänger*in)

Oben genannte Vergleichsgruppen wurden u. a. deswegen ausgewählt, um interlinguale/interkulturelle Sprecher-Beobachter-Differenzen herauszuarbeiten. Dass systematische Untersuchungen dazu nötig sind, zeigt auch eine exemplarische Studie von Missaglia (2001). Hier wird – anhand simulierter Emotionen – durch einen Vergleich deutscher und italienischer Muttersprachler*innen „die Hypothese divergierender Verhaltensweisen bei der Dekodierung emo-

tionsgeladener Sprechakte in Erst- und Zweitsprache“ bestätigt. Bereits Gumperz (1982) beobachtet, dass bestimmte prosodische Muster, die im Indian bzw. Pakistani English emotional nicht markiert sind, von Sprecher*innen des British English als emotional fehlinterpretiert werden (Kehrein 2002: 325). Auch aufgrund neuerer Forschungen kann davon ausgegangen werden, dass die Emotionsproduktion und -perzeption sozio-kulturell normiert und daher durchaus Unterschiede in den deutsch/italienischen Vergleichsgruppen zu erwarten sind.

1.3 Aufbau der Arbeit

Im zweiten Kapitel der Arbeit sollen die theoretischen Grundlagen für die empirische Untersuchung erarbeitet werden. Dabei wird besonders auf eine genaue Definition der Begrifflichkeiten, eine klare Verortung des Themas und eine kritische Erörterung der Möglichkeiten und Schwierigkeiten von Untersuchungen zur Prosodieperzeption Wert gelegt.

Der dritte Teil ist der Beschreibung und Auswertung der empirischen Untersuchung gewidmet. Hier werden die Forschungsfragen, Ziele, Hypothesen und Methoden nochmals näher erläutert, Untersuchungsaufbau und -durchführung beschrieben und die Ergebnisse vorgestellt.

Die Diskussion der Forschungsergebnisse, die Beantwortung der Forschungsfragen und die Hypothesenbeantwortung erfolgen schließlich im vierten Kapitel der Arbeit.

2 Theoretische Grundlagen und Forschungssituation

2.1 Prosodie und Sprechausdruck

2.1.1 Definition Prosodie

In dieser Arbeit wird u. a. der Frage nachgegangen, auf welche Art und Weise prosodische Merkmale rezipiert werden. Daher soll zunächst der Prosodiebegriff näher umrissen werden. In dieser Arbeit liegt der Fokus auf der Perzeption prosodischer Parameter, was notwendig erscheint, da die Prosodie in der wissenschaftlichen Literatur bisher vor allem unter sprecher- bzw. signalseitiger Perspektive betrachtet wurde. Da die Prosodie Untersuchungsgegenstand unterschiedlicher Teildisziplinen ist, welche je nach Forschungsinteresse verschiedene Aspekte dieses komplexen Phänomens thematisieren, existieren eine Fülle an Definitionen und Funktionszuschreibungen nebeneinander (Kranich 2016: 17). In dieser Arbeit liegt der Schwerpunkt jedoch auf der perceptiven Seite, weshalb hier keine fachübergreifende Gesamtdarstellung zur Prosodie angestrebt wird, sondern nur ein Abriss über die diversen Sichtweisen dargestellt wird. Dabei werden primär phonetische, phonologische und sprechwissenschaftliche Betrachtungsweisen berücksichtigt.

Prosodie ist ein Spezifikum der gesprochenen Sprache, d. h. wie alle Aspekte derselben handelt es sich hierbei um ein „flüchtiges“ Phänomen, dessen systematische Analyse nur in jüngerer Zeit möglich geworden ist. In der Duden-Grammatik 2009 wird Intonation¹ als melodische Gestalt einer Äußerung bezeichnet, welche sich aus der Wahrnehmung von Tonhöhen-eigenschaften durch das Gehör ergibt. Somit ist die Intonation ein auditives Phänomen, das nur anhand von Tonaufzeichnungen wissenschaftlich untersucht werden kann. Aufgrund dieser Technikabhängigkeit ist die Prosodieforschung auch ein vergleichsweise junger Zweig linguistischer Forschung, wenngleich das Bewusstsein für die sprachlich-kommunikative Bedeutung von Prosodie schon seit der Antike besteht. Vor allem unter Begriffen wie „vox“ und „pronuntiatio“ bemüht man sich in der antiken Rhetorik um Phänomene des Wohlklangs, der rhythmischen Gliederung und um Ausdruck und Auslösung von Affekten (Kehrein 2002: 6). Wunderli [u. a.] kritisieren zwar den geringen Erkenntnisfortschritt seit jener Zeit bis zur Mitte des 20. Jahrhunderts. So sei die Intonationsforschung „vor allem durch ein kaum mehr zu überbietendes Maß

¹ In den diversen Forschungsbereichen, welche im Folgenden ansatzweise umrissen werden, finden sich als Synonyme zu Prosodie auch die Termini Intonation und Suprasegmentalia.

an Dilettantismus“ gekennzeichnet (Wunderli u. a. 1978: 9). Allerdings gilt es durchaus relevante Forschungsbemühungen festzuhalten, so beschreibt Sievers z. B. in seinen „Grundzügen der Lautphysiologie“ von 1876 und vor allem in der Auflage von 1901 sehr ausführlich prosodische Phänomene, wenngleich nur als stilisierte Verläufe.

Die Erforschung prosodischer Eigenschaften der Rede geht in Deutschland aber hauptsächlich auf die 1950er Jahre zurück und legte ihren Schwerpunkt vor allem auf den Fokusakzent und den Zusammenhang zwischen Intonations- und Satztypen (Schwitalla 2012: 56). Als wichtige Schrift in diesem Zusammenhang kann Otto von Essens „Gründzüge der hochdeutschen Satzintonation“ von 1956 genannt werden. Seit den 1980er Jahren wurden vermehrt auch interaktive Phänomene berücksichtigt. Unter dem Einfluss der Habilitationsschrift Margret Seltings gewinnen seit Anfang der 1990er Jahre denn auch konversationsanalytische Methoden für die bis dahin hauptsächlich satzbezogene Prosodieforschung an Bedeutung (Kehrein 2002).

Im komplexen Symbolsystem „Prosodie“ können theoretisch folgende Phänomenbereiche als die wichtigsten unterschieden werden: Akzent und Rhythmus, Tonhöhe (Tonhöhenrichtung, Tonhöhenverlauf, -bandbreite und -register), Lautstärke, Sprechgeschwindigkeit, Pausen und Stimmfärbung (Timbre). In der sprachlichen Realität kommen alle Teilbereiche immer gemeinsam vor. Die meisten der genannten Kategorien haben zeitlich gesehen eine lokale und eine globale Ausdehnung, d. h. Tonhöhe, Lautstärke und Tempo etwa können sowohl bei ein paar Silben als auch bei längeren Äußerungen untersucht werden (Schwitalla 2012: 56). In der Forschungsliteratur werden die Termini Prosodie, Suprasegmentalia und Intonation oft für die Beschreibung gleicher oder ähnlicher akustisch-auditiver Phänomene verwendet. Je nach wissenschaftlicher Schule können sie jedoch unterschiedlich geprägt sein (vgl. Neuber 2002; Wendt 2006).

Prosodie kann als Oberbegriff für die suprasegmentalen Aspekte einer Äußerung verstanden werden bzw. wurde gerade in den 70er und 80er Jahren des 20. Jahrhunderts häufig synonym zu Suprasegmentalia verwendet (Rabanus 2001). Dieser Begriff meint diejenigen sprachlichen Elemente, deren Geltungsbereich über ein Einzelsegment hinausreicht, wobei die Suprasegmentalia für ihre Realisierung jedoch die segmentale Basis benötigen (Pétursson/Neppert 1991). Lindner verwendet den Begriff Suprasegmentalia für alle akustischen Eigenschaften, die Aussprüche unabhängig von den in ihnen enthaltenen Segmenten gliedern und typisieren (Lindner 1977), wobei als messbares Korrelat Intensitätsveränderung, zeitliche Strukturierung und Grundfrequenzverlauf genannt werden. Letzteres sei besonders für den Ausdruck des emotionalen Anteils einer Äußerung geeignet. Stock/Suttner definieren die Suprasegmentalia als stets

verschiedene Kombination aus den Elementen Grundfrequenz, Lautheit, Sprechgeschwindigkeit und Pausierung (Stock/Suttner 1991). Suprasegmentalia erfüllen nach Kranich eine Doppelfunktion, da sie als linguales Merkmal zum einen der Textstrukturierung und Signalisierung der Satzmodalität dienen und als paralinguales Merkmal zum anderen dem Text seine konkrete prosodische Struktur verleihen, wobei die Eigenschaften Tonhöhe, Lautheit, Temporalität und stimmklangliche Charakteristika zum Tragen kommen (Kranich 2003). Die suprasegmentalen Merkmale treten immer kombiniert auf und werden situationsbedingt bzw. individuell unterschiedlich verwendet, allerdings stets im Rahmen eines bestimmten Regelbereiches bzw. der Erwartungsnormen. Unterschieden werden muss dabei zwischen frei Gesprochenem, z. B. beim spontanen Gespräch, und Reproduziertem, also Vorgelesenem oder Memoriertem. Je nach Situation ergeben sich so unterschiedlich strukturierte, rhythmisch-melodische Strukturen, welche charakterisiert sind durch die Akzentgruppe, deren Abgrenzung mittels Pausen und anderer Grenzsignale und durch die Melodisierung der rhythmischen Gruppen. Neben der kommunikativen, strukturierenden, phonologischen und gesprächsorganisierenden Funktion, welche die Suprasegmentalia nach Bose/Hirschfeld/Neuber/Stock erfüllen (2013: 40f.), spielt für diese Arbeit besonders die Realisierung der expressiven bzw. affektiven Funktion eine große Rolle, sprich der Ausdruck von Emotionen durch suprasegmentale Mittel.

Mit der Bestimmung suprasegmentaler Eigenschaften gehen in der Literatur die Begriffe Prosodie und Intonation einher, wobei insbesondere letzterer unterschiedlich ausgelegt wird. Ausgehend von der Linguistik wurde der Begriff Intonation in der Vergangenheit am häufigsten verwendet. Bereits die Prager Strukturalisten bezeichneten damit in Anlehnung an die Musik den Tonhöhenverlauf einer Äußerung zur Kennzeichnung des Satztypus (Trubetzkoy 1971: 201f.). In der Literatur kennt man einen engen und einen weiten Intonationsbegriff (Wendt 2006: 44f.) bzw. zwei international bedeutende Ansätze zur Beschreibung von Prosodie, welche unter den Namen „britische Tradition“ und „amerikanische Tradition“, die heute zur sog. „autosegmentalen-metrischen Phonologie“ weiterentwickelt wurde, bekannt sind (Kehrein 2002: 8ff.). Die britische Tradition modelliert Tonhöhenverläufe als Tonhöhenbewegungen (Rabanus 2001) und es wird versucht, prosodische Einheiten unterschiedlicher Größe (prosodic domains) zu eruieren, um diese in Abhängigkeit vom stärksten Akzent einer Äußerung als Tonhöhenverläufe (nuclear tones) zu beschreiben. Dabei gestaltet sich die Abgrenzung der Domänen jedoch als problematisch. Die autosegmentale-metrische Phonologie entzieht sich dieses Problems, da sie von der Silbe als zentraler Beschreibungseinheit einer linearen Abfolge

phonologischer Ereignisse (phonological events) ausgeht (Kehrein 2002) und Tonhöhenbewegungen als Folge von Tonhöhenstufen modelliert (Rabanus 2001).

Für diese Arbeit besonders relevant ist der Ansatz der neueren sprechwissenschaftlichen Forschung, die zeigt, dass die der suprasegmentalen Form entsprechenden Wahrnehmungen nie nur auf ein Signalmerkmal zurückzuführen sind (vgl. Bose 2003; Neuber 2002). So definiert Stock z. B. Intonation in weitem Sinne als Kombination von Tonhöhen-, Lautheits- und Tempoveränderungen, die gesprochene Sprache als Ganzheit kennzeichnet und ihr durch Akzente und Pausen eine Struktur, eine „innere Gestalt“ verleiht (Stock 1996: 216). Stock verwendet diesen Begriff in Abgrenzung zu dem der Intonation im engeren Sinn, welcher die bis dahin mehrheitlich verwendete Umschreibung der Tonhöhenverlaufskontur bezeichnet. Intonation im weiten Sinn hingegen umfasst auch zusätzliche Eigenschaften, die heute als prosodische Merkmale subsummiert werden. So scheint sich in den letzten Jahren tendenziell der Begriff Prosodie als eigenständiges, von anderen linguistischen Teilsystemen unabhängiges Signalsystem zu etablieren, während damit in der Vergangenheit unter der Bezeichnung Intonation hauptsächlich die enge Beziehung zur grammatikalischen Funktion betrachtet wurde (Kranich 2016: 18). So sah Pheby beispielsweise die Funktion der Intonation in der Zuschreibung des Satztypus, welche sich durch unterschiedliche Tonmuster realisieren und innerhalb verschiedener Satzarten zwischen markierten und unmarkierten Satztypen unterscheiden (Pheby 1981: 874ff.). Eine vergleichbare Position vertrat auch die Münchner Forschungsgruppe um Altmann, nur dass statt verschiedener Satzarten fünf verschiedene Satzmodi (Assertion, Interrogation, Wunsch, Aufforderung, Exklamation) durch intonatorische Merkmale gekennzeichnet werden (Altmann 1989). Eine erweiterte Sichtweise findet sich bei Bolinger, welcher die Intonation als ein von der Grammatik unabhängiges, autonomes Signalisierungssystem betrachtet. Vielmehr sei die Intonation Teilbereich der Gestik, denn beide Systeme weisen auf etwas hin und seien also unabhängig von der Grammatik. Die Funktionen der Intonation sind laut Bolinger vielfältig, so können z. B. Grenzsingale gesetzt, verschiedene Äußerungstypen markiert oder durch Akzentuierungen Interpretationshilfen nicht grammatikalischer Natur gegeben werden (Bolinger 1983). Derartige Begriffsbestimmungen sind für empirische Forschungen sinnvoll und notwendig, für eine umfassende Definition greifen sie jedoch zu kurz.

Einen umfassenderen Begriff verwendet Selting, die Prosodie definiert als „Oberbegriff für diejenigen suprasegmentalen Aspekte der Rede, die sich aus dem Zusammenspiel der akusti-

schen Parameter Grundfrequenz (F_0), Intensität und Dauer in silbengroßen und größeren Domänen ergeben“ (Selting 1995: 1). Dabei werden auditive Phänomene wie die Intonation im engeren Sinn, Lautheit, Länge, Pause, Sprechgeschwindigkeit und Rhythmus ausdrücklich eingeschlossen. Mit den vielfältigen Funktionen der Prosodie beschäftigen sich in neueren linguistischen Forschungen z. B. Fox (2002: 271) und Tatham/Morton (2007: 212), welche Prosodie als eine Dimension jenseits des individuellen Stimmklangs verstehen, die doch mit diesem verbunden sein muss. Nach Dogil ermöglicht die Prosodie die grundlegende Organisation des Sprechens durch die Integration semantischer Informationen (Bedeutung), der syntaktischen (Phrasierung) und morphologischen Struktur (metrisches Gerüst) und die segmentale Sequenzierung. Auf weitere nicht-linguistische Funktionen verweist der Autor nicht (Dogil 2003: 546). Nach Hielscher hingegen bildet sich der Funktionsaspekt der Prosodie auf einem Kontinuum von nicht-linguistischen (z. B. Stimmqualität und Spannungsverhältnisse) zu linguistischen Funktionen (Betonung und Frequenzmuster) ab. Im mittleren Bereich lassen sich die paralinguistischen Merkmale, so z. B. die emotionale Prosodie, einordnen. (Hielscher 2003: 680) Hier zeigt sich die Besonderheit der Prosodie im Vergleich zu anderen Teilen des Sprachsystems – lässt sich bei einem Phonem eine relativ eindeutige Zuordnung zwischen Inhalt und Form vornehmen, sind prosodische Merkmale jedoch polyfunktional (Neuber 2001: 103). Dies zeigt sich sehr deutlich etwa für die Klangqualität, welche indexikalische, biophonetische und linguistische Funktionen erfüllen kann.

Eine Unterscheidung der Begriffe suprasegmental und prosodisch nimmt Grassegger vor, er bezeichnet Prosodie als Oberbegriff aller sprachlichen Funktionen mit Suprasegmentalia. Die aus phonetischer Sicht den Einzellaut übergreifenden Erscheinungen seien dabei als suprasegmental zu bezeichnen, jene unter phonologischer Perspektive als prosodisch (Grassegger 2010: 63). Kohler (1995: 110) und Kehrein (2002: 149) hingegen schlagen eine Brücke zwischen phonetischen und phonologischen Aspekten und differenzieren zwischen phonetisch wahrnehmbaren bzw. messbaren Größen und phonologischen Größen mit linguistischer Bedeutung. Eine ausführliche Übersicht zum Stand der Terminologie-Diskussion findet sich bei Heilmann. Sie stellt fest, dass in der phonetisch orientierten Literatur hauptsächlich der Terminus Prosodie als Oberbegriff für die suprasegmentalen Aspekte des Sprechens Verwendung findet. Hier wird außerdem unterschieden zwischen akustischen Parametern, die messtechnisch aus dem Schallsignal abgeleitet werden (Grundfrequenz, Dauer, Intensität u. Ä.) und weiteren Merkmalen (auch paralinguistische Parameter oder Begleitphänomene genannt). Dazu zählen z. B. Sprechtempo, Rhythmus, Artikulations- und Sprechspannung, Pausen und Stimmqualität. Während

die messbaren Merkmale sowohl segmental als auch suprasegmental wirken können, haben die weiteren Merkmale ausschließlich suprasegmentalen Charakter (Heilmann 2002: 20).

Einen guten Überblick zum Stand der Begriffs-Diskussion bietet auch Kranich (Kranich 2016: 17ff.).

Prosodie wird im Allgemeinen zum einen auf der phonetischen Ebene und zum anderen auf der phonologischen Ebene beschrieben. Zentrale Ziele und Schwierigkeiten sind dabei zunächst die Segmentierung prosodischer Einheiten, deren Form- und Funktionsbeschreibung und das Verhältnis von Prosodie zu anderen sprachlichen Signalisierungssystemen. Prosodische Merkmale sind Phänomene, die sich in der Zeit kontinuierlich verändern (v. a. Tonhöhe und Lautstärke) und sich normalerweise über mehr als ein Segment erstrecken. Dabei gilt, dass Prosodie ohne segmentelle Basis ebenso wenig existieren kann wie gesprochene Sprache ohne prosodische Eigenschaften (Kehrein 2002: 6f.). Daraus resultiert auch die Notwendigkeit von Forschungen mit natürlicher Sprache, wie hier angesetzt werden soll. Weiters gilt es die phonologische/funktionale und phonetische Ebene der prosodischen Parameter klar zu unterscheiden. Dabei sollte sich die Analyse zunächst auf die Ermittlung der funktionalen Einheiten und erst dann auf die Formseite richten. Diesen Ansatz vertreten auch Kehrein (2002: 7) und Hammarström, welcher dafür plädiert, zunächst die „Prosodkomponenten als Lauteindrücke des Hörers“ zu eruieren. „Erst danach kann festgestellt werden, wie sie vom Sprecher erzeugt werden und welche akustische [sic!] Stimuli sie hervorrufen. Wie die Komponenten hier verstanden werden, können sie nicht durch akustische Messungen ermittelt werden, denn ein ‚reines‘, nicht durch schon bekannte Höreindrücke gesteuertes, akustisches Studium könnte teilweise Komponenten [...] ermitteln, die nichts mit Höreindrücken zu tun haben.“ (Hammarström 1963b: 198)

Die Relevanz der Differenzierung zwischen einer parameterorientierten akustischen und einer funktionalen Definition unterstreicht auch Neuber als Voraussetzung zur Begriffsbestimmung von Prosodie. Diese Unterscheidung resultiert aus der Tatsache, dass prosodische Merkmale eine materielle und eine immaterielle Dimension aufweisen, die jeweils unterschiedlichen Kategorien zuzuordnen sind (Neuber 2002: 51). So ist Prosodie laut Hirschfeld/Neuber zu verstehen als „multiparametrischer Merkmalskomplex“, in dem die einzelnen Merkmale „einzeln oder in Kombination (als Akzentuierung, Gliederungssignale, rhythmische Muster) bestimmte Funktionen in gesprochenen Äußerungen [erfüllen], wie zum Beispiel hervorzuheben oder zu strukturieren. Merkmale und Funktionen der Prosodie sind sowohl Gegenstand der Phonologie (als Teil des Sprachsystems) als auch der Phonetik (als konkrete Realisationsformen) (Hirschfeld/Neuber 2010: 10f.).“ Wesentlich bei Neuber (2002) ist die Hervorhebung zweier Aspekte

der Prosodie – zum einen haben prosodische Merkmale eine zeitlich-akustische, dem Signal zuzuordnende Korrelate, zum anderen müssen diese signalseitigen, materiellen Gegebenheiten jedoch auch auditiv wahrgenommen werden. Erst dieser Aspekt ermöglicht, dass die Prosodie ihre vielfältigen linguistischen (z. B. als Distinktionsfunktion durch Akzentverschiebungen oder die Markierung der Satzart) und nicht-linguistischen (z. B. indexikalische, soziale, biophonetische, gesprächsstrukturierende) Funktionen ausüben kann (Kranich 2016: 21f.). Die Prosodie wirkt auf der Beziehungsebene, der Sachebene und der Stilisierung des Sprechens und spielt auch im Rahmen der interkulturellen Kommunikation eine zentrale Rolle (Hirschfeld/Stock 2010: 24). Prosodische Merkmale stehen außerdem in Verbindung zu anderen sprachlichen Signalisierungssystemen, wie z. B. Verbalsprache, Mimik, Gestik, Körpersprache etc. (Kehrein 2002: 8).

Aus den referierten Forschungsrichtungen werden zum einen die funktionale Vielfalt prosodischer Merkmale und zum anderen die Verwendung verschiedener Begrifflichkeiten für den Merkmalskomplex Prosodie ersichtlich. Da der Terminus Intonation sowohl in einem engeren Sinn als Tonhöhe bzw. Grundfrequenz als auch in weiterem Sinn als der dieser Arbeit zugrunde liegende Prosodiebegriff verstanden werden kann, scheint eine eindeutige Zuordnung schwierig. Auch der Begriff Suprasegmentalia ist durch die Einführung des Begriffs ‚autosegmental‘ potenziell ambig geworden und könnte auf eine nicht intendierte Nähe zum theoretischen Umfeld der autosegmentalen-metrischen Phonologie hinweisen (Kehrein 2002: 5). Daher wird in dieser Arbeit der Terminus Prosodie verwendet. Dieser scheint auch deshalb geeignet, da er in den letzten Jahren in entsprechenden Forschungsansätzen vermehrt und in weiterem Sinn Anwendung findet (vgl. Neuber 2002, Kehrein 2002, Kranich 2003 u. a.) und die wesentlichen prosodischen Elemente umfasst, welche für die Emotionsperzeption bzw. den hier vorliegenden Forschungsansatz relevant sind.

In Anlehnung an Neuber (2002) wird Prosodie hier also als Zusammenfassung der „auditiv wahrnehmbaren Merkmale Melodieführung, Lautheit und Lautheitsdifferenz, Akzent, Sprechgeschwindigkeit (Tempo) und Tempowechsel, Sprechrhythmus, Pausen, (indexikalisch bedingte) Stimmqualität, Stimmausdruck (Timbre) und Gesamtdauer der prosodischen Erscheinung“ (Neuber 2002: 45ff.) definiert.

2.1.2 Prosodische Merkmale

Wie schon bei der allgemeinen Definition von Prosodie zeigen sich auch bei der Bezeichnung der prosodischen Merkmale unterschiedliche Begriffsfestlegungen, was hauptsächlich in der potentiellen Mehrdimensionalität der Prosodie bezüglich möglicher Repräsentationsebenen begründet ist (Günther 1999). In der Phonologie z. B. kommt dem Ton (tone) meist eine größere Rolle als dem Tonhöhenverlauf, der Betonung und dem Akzent zu (Fox 2002: 1), während andere phonetisch interessante Eigenschaften, wie z. B. Tempo oder Klangqualität, oft keine Berücksichtigung finden (Kranich 2016: 22). Prosodische Merkmale lassen sich also als abstrakte Größen unter phonologischem Gesichtspunkt definieren oder als artikulatorische, akustische und auditive Parameter beschreiben, wobei hier eine klare Abgrenzung angestrebt wird. Die einzelnen prosodischen Merkmale sollen in Bezug zu den im vorigen Abschnitt dargestellten Ansätzen und der vorgenommenen Positionierung ohne Anspruch auf Vollständigkeit so definiert werden, wie sie im Rahmen dieser Arbeit relevant sind. Es gilt natürlich zu beachten, dass prosodische Merkmale gesprochener Sprache niemals isoliert auftreten und sich wechselseitig beeinflussen (Kranich 2016: 22).

2.1.2.1 Akzent

Der Begriff Akzent bezeichnet die Prominenz von Äußerungsteilen im Syntagma. Es handelt sich dabei um ein relationales Merkmal, da ein Teil einer Äußerung nur in Bezug auf einen anderen Teil derselben hervorgehoben wird (Günther 1999: 48). Andere Begriffe, die zum Teil synonym Verwendung finden, sind Betonung, Hervorhebung, Fokus, Iktus und Prominenz (Kranich 2016: 27). Der Akzent ist nach Fox das am meisten kontrovers diskutierte prosodische Merkmal. Über seine phonetische Natur und phonologische Rolle besteht bisher noch kein Konsens (Fox 2002: 14). Die Akzentuierung ist ein komplexes Phänomen, welches sich isoliert als Tonhöhen-, Dehnungs-, Intensitäts- und Klangqualitätsakzent manifestieren kann, praktisch jedoch stets als Kombination mehrerer dieser Eigenschaften auftritt (Kranich 2016: 35). So können z. B. eine größere Lautstärke und eine höhere, teilweise auch tiefere Tonhöhe für starke Akzentuierungen gebraucht werden (Schwitalla 2012: 57). Aus linguistischer Funktion werden „stress“ und „accent“ differenziert. Zum einen bezeichnet Akzent als abstraktes Merkmal (engl. stress) die potentiell gegebene Möglichkeit einer Hervorhebung, die durch grammatische Strukturgegebenheiten bedingt ist. Im Allgemeinen ist damit der Wortakzent gemeint, d. h. die durch

lexikalisch-phonologische Prinzipien zurückführbare Betonung der Silben. Des Weiteren werden auch auf Phrasen- und Satzebene abstrakte Akzente (engl. phrasal stress, sentence stress) entsprechend semantisch-syntaktischer und allgemeiner metrischer Regeln auf Worte und Konstituenten im Syntagma gesetzt. Zum anderen ist mit Akzent auch die real vollzogene und wahrnehmbare Hervorhebung (engl. accent) in der Äußerung gemeint (Günther 1999: 48f.). Die Funktionen von Akzentuierungen sind vielfältig, so dienen sie beispielsweise dazu, einen Redegegenstand zu thematisieren, inhaltlich Neues oder kommunikativ Relevantes hervorzuheben, etwas herauszustellen oder in Kontrast zu etwas anderem zu stellen usw. Die wichtigste Funktion des primären Akzentes ist es also, den rhematischen Bereich einer Äußerung zu markieren (Stock 1996: 51). Eine weitere wichtige Funktion bekommt der Akzent zum Ausdruck von Emotion, Wertung, Eindringlichkeit und Emphase, welche v. a. durch Dehnung und eine weite Tonhöhenbandbreite realisiert werden. Dabei gibt es eine direkte Entsprechung zwischen der Tonhöhenbandbreite und dem Grad der Expressivität. So kennzeichnen Sprecher*innen durch starke Akzentuierungen auf hoher, z. T. auch auf tiefer Tonebene Äußerungseinheiten als etwas, das nicht ihren Erwartungen entspricht (Schwitalla 2012: 60f.). Ebenso verhält es sich mit den Wörtchen „ja“ und „nein“ – tief und gedehnt gesprochen drückt letzteres Erstaunen bzw. Fassungslosigkeit aus. Ein laut und gedehnt realisiertes „ja“ hingegen signalisiert eine feste eigene Überzeugung im Gegensatz zu einem prosodisch flach gesprochenen Antwort-Ja (Rabanus 2001: 219).

Akzente wirken als Träger und Wendepunkte für den Melodieverlauf. Bei den Tonhöhenbewegungen auf der akzentuierten Silbe und danach fortgesetzt auf nicht-akzentuierte Silben lassen sich steigende (L*H), hohe (H*), fallende (H*L) und tiefe Akzente (T*) unterscheiden, wobei H ‚hoch‘ (high) und L ‚tief‘ (low) meint (Selting 2009: 378ff.). Die Tonhöhenbewegung nach dem nuklearen Akzent, d. h. der letzten akzentuierten Silbe, bis zum letzten Ton der Tonhöhenbewegung ist entscheidend für deren Gesamteindruck. Transkribiert werden sollte nach dem auditiven Eindruck, wobei zwischen stark akzentuierten Silben (!BUCH!), primären oder Hauptakzenten ((BUCH), sekundären oder Nebenakzenten (bUch) und nichtbetonten Silben unterschieden werden kann (Schwitalla 2012: 57).

Wie die Akzente phonetisch genau konfiguriert werden ist abhängig von der jeweiligen Sprache. Im Deutschen ist die Zusammensetzung aus Tonhöhen- (Maxima bzw. Minima) und Intensitätsmerkmalen am häufigsten zu beobachten. Auch temporale Merkmale, z. B. die Verlängerung eines Silbensegments, und eine deutlichere Artikulationspräzision treten oft in Erschei-

nung. Meist werden Akzente durch das Zusammenwirken verschiedener prosodischer Merkmale konfiguriert. In emotionalen Sprechweisen manifestieren sich an Akzentpositionen auch Änderungen der Klangqualität (Kranich 2016: 27f.).

2.1.2.2 Rhythmus

Ein weiteres komplexes Phänomen ist auch der Rhythmus, der generell im Rahmen menschlicher Aktivitäten und natürlich auch in der Sprache eine große Rolle spielt und zu welchem eine Fülle an nicht zur wissenschaftlichen Publikationen vorliegt. Tatsächlich sind systematische wissenschaftliche Untersuchungen zu diesem prosodischen Phänomen eher rar und es besteht keine Einigkeit über die linguistische Funktion desselben (Fox 2002: 86). Eine allgemeine Definition des Begriffs Rhythmus kann dem Lehrbuch Otto von Essens entnommen werden: „Rhythmus ist die periodische Wiederkehr gleicher und in gleicher Ordnung abgestufter, durch eine dynamische Gipfelbildung zur Einheit gebundener Vorgänge. Es ist nicht notwendig, dass die mit der gleichen Ereignisfolge erfüllten *Zeitabschnitte* gleich sind; die Perioden können lang und kurz, die „Agogik“ kann lebhaft sein; der Begriff des Rhythmus bezieht sich allein auf den Bau, nicht auf die *Dauer* der Periode.“ (Essen 1957: 128) Fährmann bezeichnet Rhythmus als „etwas sich immer wieder aus sich selbst heraus Erneuerndes“ und betont die Abgrenzung dieses Begriffs vom Takt, welcher nach ihm „ein unlebendiges, starres, schemahaftes Metrum“ darstelle (Fährmann 1982: 149).

In der Tat gilt es die allgemein rhythmische Gliederung lautsprachlicher Äußerungen in sog. „Takte“ von den mit der phrasalen Gliederung von Äußerungen verbundenen lautbezogenen Dauerphänomenen zu differenzieren.

Als Takt oder Fuß wird die gegenüber der Silbe zeitlich größere und prosodisch komplexer strukturierte phonetische Einheit bezeichnet. Eine bzw. mehrere zusammengesetzte Takte ergeben schließlich die prosodisch kohärente intonatorische Phrase, wobei die einzelnen Sprachen durchaus ähnliche Präferenzen im Taktrhythmus aufweisen (Pompino-Marschall 2009: 249). So wird für die verschiedenen Sprachen bezüglich ihrer rhythmischen Eigenschaften eine typologische Unterteilung in betonungszählende (engl. stress timed), silbenzählende (engl. syllable timed) und morenzählende (engl. mora timed) Sprachen vorgeschlagen. Diese seien dadurch gekennzeichnet, dass die jeweils ausgezeichneten Einheiten, also die betonten Silben, die einzelnen Silben oder die einzelnen Moren, in gleichmäßigem zeitlichen Abstand aufeinander folgen. Diese sog. Isochroniehypothese lässt sich zwar in ihrer strengen Form messphonetisch am akustischen Signal nicht nachvollziehen, allerdings lassen sich durchaus Tendenzen

feststellen. So zeigen betonungszählende Sprachen wie das Deutsche oder Englische beispielsweise eine starke Tendenz zur zeitlichen Verkürzung unbetonter Silben, vor allem je mehr unbetonte Silben zwischen bei betonten Silben vorkommen. Hauptsächlich die vokalischen Silbenkerne, v. a. der Schwa-Laut als Reduktionsvokal, sind von dieser Verkürzung bzw. Kompression betroffen. In den morenzählenden Sprachen, z. B. im Japanischen, zeigen sich hingegen keine derartigen betonungsbedingten Variationen, vielmehr lässt sich ein recht stabiler Dauerunterschied zwischen zweimorigen (schweren) und einmorigen (leichten) Silben erkennen. Zwischen diesen beiden Polen stehen die silbenzählenden Sprachen, zu denen z. B. auch das Italienische und alle weiteren romanischen Sprachen gehören. Hier hält sich die Variabilität der Silbendauer in Grenzen, auch zeigen sich hier im Gegensatz zu den betonungszählenden Sprachen keine Kompressionseffekte (Pompino-Marschall 2009: 248f., vgl. Fox 2002: 87ff.). Von diesem Taktrhythmus zu trennen sind einzelne Phrasierungseinheiten, die durch wahrnehmbare Grenzsignale (sog. Junktoren) dem Gehör als prosodisch abgrenzbare Einheit zugänglich gemacht werden (Pompino-Marschall 2009: 249). In der gesprochenen deutschen Alltagssprache ist ein hörbarer Rhythmus eher selten, er erfüllt dann rhetorische Funktionen (Uhmann 1996: 326) oder zählt zur Ressource bestimmter Gruppen, z. B. das „stampfende“ Sprechen von Türkinnen der zweiten Generation (Keim 2001: 281ff.). Erickson (1992) z. B. konnte im Rahmen einer amerikanischen Unterhaltung beim Essen nachweisen, wie die Beteiligten durch Sprechen, Gestik und aktionale Handlungen einen gemeinsamen Rhythmus entwickeln. Isochronie wird in einem Gespräch, häufig auch bei Verabschiedungsformeln, zur Demonstration von Gemeinsamkeit hergestellt (Schwitalla 2012: 64). Außerdem werden bestimmte Sprechrhythmen als emphatische Verstärkung von Meinungen und Wertungen produziert. So zeigt Uhmann (1996: 329ff.) z. B., dass bei starken Wertungen im Deutschen bis zu fünf akzentuierte Silben nacheinander ohne Pause realisiert werden, was ja dem Prinzip der Vermeidung direkt aufeinander folgender Akzentuierung eigentlich widerspricht. Es handelt sich also um eine sehr auffällige prosodische Form und sie wirkt ansteckend auf ihre Umgebung (Uhmann 1996: 327ff.). Weiters können durch das prosodische Merkmal des Rhythmus die mit den Äußerungen verbundenen Sprechakte intensiviert und die Verständnissicherung unterstützt werden. Rhythmische Redepassagen heben sich von ihrer Umgebung ab und erhalten die kommunikative Bedeutung von Eindringlichkeit, außerdem werden sie von den Hörern in ähnlicher Weise beantwortet, wodurch auch längere Passagen gemeinsam rhythmisch gesprochen werden können (Schwitalla 2012: 61ff.).

Unter Sprachrhythmus kann also die bestimmten Regularitäten und kommunikativen Absichten folgende zeitliche Abfolge rezipierter phonetischer Ereignisse verstanden werden, wobei in dieser Arbeit v. a. der emotional markierte Sprechrhythmus unter dem Aspekt der Mitteilungsäußerung bzw. Perzeption von Interesse ist. Dabei gilt zu beachten, dass der Rhythmus je nach Realisationsform (z. B. rhythmusgebundene Dichtung) mehr oder weniger bewusst produziert und perzipiert werden kann und häufig auch mit Tonhöhen- und Intensitätsveränderungen kombiniert sein kann (vgl. Kranich 2016: 35).

2.1.2.3 Tonhöhe und Melodieführung

Auch in Bezug auf die prosodischen Merkmale der Tonhöhe und Melodieführung gilt es kurz die Begrifflichkeiten zu klären.

Von der Analyse und Beschreibung intonatorischer Phänomene ausgenommen sei an dieser Stelle der für das Deutsche und Italienische nicht relevante Bereich des lexikalisch distinktiven Tons, dem sog. Tonem, in Tonakzent- und Tonsprachen (engl. pitch accent languages). In Tonakzentsprachen, wie z. B. dem Schwedischen, Norwegischen oder Japanischen, wird lexikalisch den Worten ein Tonem, d. h. ein Wortakzent, zugewiesen, der über dem gesamten Wort realisiert wird. In Tonsprachen wie dem Chinesischen, den mittelamerikanischen Oto-Mangue-Sprachen oder dem Yoruba wird hingegen fast jeder Silbe lexikalisch ein Tonem zugewiesen. Dieses trägt in diesen Sprachen dieselbe distinktive Funktionalität wie Vokal- oder Konsonantenphoneme und manifestiert sich in verschiedenen Tonregistern (Registertonsprachen) oder Tonkonturen (Konturtonsprachen) (Günther 1999: 62).

Auf die divergierenden Darstellungen des Begriffs der Intonation wurde im Abschnitt 2.1.1 bereits hingewiesen, wobei hier auf die melodischen Merkmale (Melodieverlauf über die Äußerung hinweg, engl. tune) von Äußerungen rekurriert wird. Die Begriffe Melodieführung, Sprechmelodie und Tonhöhenverlauf, als auditiv wahrnehmbare Größe, werden dabei synonym verwendet, gemeint ist damit die distinktive Variation der Grundfrequenz in Äußerungen (Günther 1999: 62). Die Grundfrequenz (f_0) der stimmhaften Abschnitte des akustischen Sprachsignals ist das physikalische Korrelat der Tonhöhe (engl. pitch) und spiegelt den Zeitverlauf des periodischen Schließens der schwingenden Stimmlippen wider. Dabei ist für die lautsprachliche Kommunikation weniger die absolute Tonhöhe von Bedeutung als vielmehr die Form des Tonhöhenverlaufs in Relation zur Frequenzlage des jeweiligen Sprechers (Pompino-Marschall 2009: 246). Die rezipierte Tonhöhe entspricht direkt der Periodenfrequenz des gesamten kom-

plexen Schwingungsverlaufs und somit der Grundfrequenz des Stimmklangs, wobei die Tonhöhenwahrnehmung jedoch auch von der Lautlänge und Intensität beeinflusst sein kann (Günther 1999: 63). Globale Melodieverläufe können auf insgesamt tiefem, mittleren oder hohem Tonhöhenregister gesprochen werden, wobei z. B. untergeordnete und „herabspielende“ Sprechhandlungen oft im tiefen Register, „drängende“ und für relevant erachtete Aktivitäten im hohen Register vollzogen werden (Selting 1995: 159ff.). Die Sprechmelodie, akustisch Grundfrequenz, ist der beim Sprechen erzeugte gestalthafte Tonhöhenverlauf, wobei die Lage innerhalb des Frequenzbereichs der menschlichen Stimme physiologisch und kulturell bedingt sowie individuell verschieden ist (Bose/Hirschfeld/Neuber/Stock 2016: 39).

Es gibt eine Reihe von Untersuchungen, die sich mit Typen von Tonhöhenverläufen und deren Bedeutung auseinandersetzen. Féry z. B. nennt folgende Funktionen von globalen Tonhöhenverläufen: Aussage-, W-Fragesätze und Antworten sind durch eine fallende Intonationskontur gekennzeichnet, Fragen, Rückversicherungssignale, unvollständige Äußerungen, Warnungen oder Aufforderungen in Deklarativsätzen, Antworten auf Fragen nach etwas Offensichtlichem, Unsicherheit und Entrüstung hingegen durch eine steigende. Markierte Fragen und Drohungen werden in fallend-steigenden Melodiekurven ausgedrückt und steigend-fallende Intonationsbögen (selten) werden angewendet, um Aufmerksamkeit zu erregen, zu drohen oder Selbstverständliches auszudrücken (Féry 1993: 81ff.). In linguistischen Forschungen versuchte man bestimmte Intonationskonturen mit Satztypen in Verbindung zu bringen (vgl. Wunderlich 1988: 6; Bannert 1985a; Altmann [u. a.] 1989; Uhmman 1991: 128ff.; Pheby 1981: 857). Zwar sind Muster von Tonhöhenverläufen unabhängig von syntaktischen Strukturen, sie fallen aber dennoch oft mit Satzgrenzen zusammen. Hier sei kurz angemerkt, dass die Korrelationsstudien von Satz- und Intonationstypen von isolierten Äußerungen ausgehen und daher deren Ergebnisse nicht problemlos auf die natürliche Rede übertragen werden können (Schwitalla 2012: 67f.). Auch Vorgelesenes muss von frei Gesprochenem unterschieden werden (Stock 1996: 224f.). Feste syntaktische Funktionen der Tonhöhenbewegung für natürliche Äußerungen werden daher bezweifelt (Schönherr 1997: 99; Schmidt 2001: 16).

Dennoch erfüllen die grundlegenden prosodischen Merkmale Tonhöhe und Melodieführung natürlich vielfältige Funktionen, so fungieren sie praktisch als Transportmedium für Silben und segmentale Elemente und in der Art ihrer Realisation zeigen sich indexikalische und biophonetische Eigenschaften, welchen als inhärenten Kommunikationssignalen große Bedeutung zukommt (Kranich 2016: 34). Anhand von Intonationsbögen ist es z. B. möglich, Sprecheinheiten zu gliedern, Redezeiten zu organisieren, inhaltlich oder kommunikativ Zusammengehörendes zu markieren und auch regionalspezifische Unterschiede festzustellen. Besonders relevant ist

im Rahmen dieser Arbeit der Zusammenhang von Tonhöhenverlauf und expressivem Sprechen. So können mit plötzlichen Tonsprüngen, häufig verbunden mit Änderungen der Lautstärke und Sprechgeschwindigkeit, auch psychische Erregung, Emphase, Überraschung u. Ä. ausgedrückt werden (Schwitalla 2012: 69ff.).

2.1.2.4 Lautstärke

Die auf der auditiven Ebene verankerten Begriffe Lautstärke bzw. Lautheit (akustisch: Dynamik, Schalldruck, -intensität) sollen hier synonym verwendet werden und bezeichnen die subjektiv empfundene Lautstärke des Gesprochenen. Die Lautstärke ist vom Atemdruck und der damit verbundenen Sprechspannung bedingt (Bose/Hirschfeld/Neuber/Stock 2016: 39). Dabei ist die kommunikative Angemessenheit der Lautstärke entscheidend, um das Signal wahrzunehmen und das Gesagte verstehen zu können. Absolute Werte sind als Maßstab diesbezüglich nicht aussagekräftig, vielmehr sind das Sprechen und die Rezeption im Sinne einer mittleren Lautheit grundlegend. Auch gilt zu beachten, dass signalseitig eine Abhängigkeit vom segmentalen Material besteht. So transportieren sonorantische Anteile den Sprechschall besser als etwa die hochfrequenten Geräuschbänder der Obstruenten. Auch durch Akzentuierungen können sich natürliche Schwankungen ergeben. Da das menschliche Perzeptionsvermögen auf derartige Schwankungen eingestellt ist, wird im Normalfall die Sprechweise eines Sprechers bzw. einer Sprecherin als ungefähr gleich laut empfunden (Kranich 2016: 34). Modifikationen der Lautstärke erfüllen jedoch auch wichtige Funktionen, sie dienen z. B. der Hervorhebung (Akzentuierung) in Wort, Wortgruppe und Äußerung und beeinflussen die Artikulationspräzision und damit den Umfang der Assimilationen (Lautangleichungen) und die Zahl der Elisionen (Lautausfälle) (Bose/Hirschfeld/Neuber/Stock 2016: 39). Die Lautstärke trägt auch paraverbale Funktionen, so kann eine erhöhte Lautheit z. B. auf entsprechende emotionale Zustände hinweisen (Gadler 2006: 53). Dies ist im Rahmen dieser Arbeit von besonderem Interesse. Leises Sprechen oder Flüstern beispielsweise kann die Vertraulichkeit einer Nachricht anzeigen oder auf eine untergeordnete Redeeinheit hinweisen. Auch können damit die Geltung des Gesagten oder dessen Adressierung an andere Personen vermindert werden sowie kritische Bemerkungen als nicht-öffentliche Äußerungen kontextualisiert und damit die eigene Verantwortung gemindert werden. Durch leises Sprechen kann auch Tabuthemen etwas von deren Brisanz genommen werden (Schwitalla 2012: 74). Des Weiteren werden nicht-kompetitive Interjektionen und Nebenbemerkungen prosodisch oft durch eine geringe Lautstärke gekennzeichnet und nicht als turnkompetitive Unterbrechungen behandelt (Selting 1995: 209). Erhöhte Lautstärke, oft auch

in Kombination mit höherer Tonhöhe, signalisiert hingegen den 'Kampf' um die Turnfortsetzung (Selting 1995: 215ff.). Auch bei einem Themenwechsel wird das neu einsetzende Thema oft lauter als das auslaufende Thema realisiert, z. T. wird auch schneller und höher gesprochen (Hausendorf/Quasthoff 1995: 236). Die Parallele zwischen lautem Sprechen und größerer Relevanz, Dringlichkeit oder Aufmerksamkeitsweckung gilt überdies nicht immer, entscheidende, empörende oder überraschende Mitteilungen z. B. können auch sehr leise gesprochen werden, um den rhematischen Teil der Äußerung zu unterstreichen (Schwitalla 2012: 75).

2.1.2.5 Sprechgeschwindigkeit

Der Sprechgeschwindigkeit entspricht akustisch die Häufigkeit von Signalabschnitten pro Zeiteinheit und sie wird als Silben- und Phonrate perzipiert. Unter Phon werden Allophone und Sprosslaute subsumiert. Die Sprechgeschwindigkeit ist vor allem bei Akzelerationen und Retardationen auffällig. Eine erhöhte Sprechgeschwindigkeit verstärkt außerdem die Tendenz zu Elisionen, d. h. Lautausfällen, und Assimilationen (Bose/Hirschfeld/Neuber/Stock 2016: 39). Die Artikulationsrate und deren Fluktuationen sind durch individuelle und auch situative Faktoren determiniert. So beschreibt Harf die Abhängigkeit der Artikulationsrate von bestimmten Epochen anhand einer Untersuchung, die die Zunahme der Artikulationsrate in den Industrienationen zwischen 1945 und 1995 um 50% aufzeigt (Harf 2011: 122). Als Maß für die Artikulationsrate als praktisch objektives Maß der Sprechgeschwindigkeit hat sich das Maß Silben pro Sekunde durchgesetzt (Kranich 2016: 24). Bei einer normalen Sprechgeschwindigkeit liegt eine Rate von ungefähr fünf bis acht Silben pro Sekunde vor, bei sehr schnellem Sprechen können es bis zu 400 Wörter pro Minute sein.

Signalphonetisch geschehen bei einer Veränderung der Sprechgeschwindigkeit komplexe Umstrukturierungen. Vor allem die vokalischen Silbenkerne werden bei erhöhter Sprechgeschwindigkeit zeitlich verkürzt und die Artikulationsbewegung u. U. nicht so ausgeprägt vorgenommen, wodurch die Zunge ihre Zielkonfiguration (engl. target) nicht vollständig erreicht (engl. undershoot). Die konsonantischen Artikulationsbewegungen hingegen werden zwar schneller ausgeführt, weisen aber einen wesentlich geringeren zeitlichen Kompressionseffekt auf (Pompino-Marschall 2009: 250).

Die Sprechgeschwindigkeit ist schwer zu operationalisieren und bereitet in Untersuchungen häufig Probleme, weil viele Phänomene berücksichtigt werden müssen. So können bei Silben Verschleifungen, bei Wörtern ungleiche Längen, in Takten Tempoveränderungen und außerdem Pausen und Dehnungen vorkommen. Zudem gibt es Schnell- und Langsamsprecher*innen. Ob eine Äußerung also als schnell oder langsam wahrgenommen wird, hängt weniger von der

absoluten Zahl der Silben pro Sekunde als vielmehr von der Relation zur Durchschnittsgeschwindigkeit eines Sprechers ab (Schwitalla 2012: 72).

Das Sprechtempo bzw. die Sprechgeschwindigkeit, die Begriffe sollen hier synonym verwendet werden, erfüllt wichtige Funktionen in Bezug auf das Rederecht und der Konstruktion mehrteiliger Redeteile. So kann z. B. schnelles Reden zu Beginn oder am Ende einer Redeeinheit dazu dienen, das Rederecht zu beanspruchen. Auch neue Themen werden oft durch eine erhöhte Sprechgeschwindigkeit eingeführt, um die Sprecherrolle zu sichern (Schwitalla 2012: 72). Schneller gesprochene Passagen, z. T. auch mit leiserer Stimme, weisen nach Uhmann (1992) außerdem auf einen untergeordneten Status der Äußerungen in Bezug auf deren Relevanz hin. Umgekehrt werden die für die Sprecher*innen zentralen Inhalte eher langsamer gesprochen (Paul 2000: 125). Überdies kann ein schnelles bzw. langsames Sprechtempo auch eine ikonisch abbildende Funktion erfüllen (Schwitalla 2012: 74). Verlangsamte Passagen werden spannungsfördernd eingesetzt, um das Gesagte zu unterstreichen (Redder 1994: 248), oder bereiten in Erzählungen und Witzen durch gedehntes Sprechen, Pausen etc. die Pointe vor (Kallmeyer 1981: 414; Kotthoff 1997: 129).

2.1.2.6 Dauer

Die Dauer als temporales Merkmal der Prosodie spielt in dieser Arbeit nur indirekt eine Rolle, liegt aber natürlich jeder Äußerung zugrunde und soll daher kurz genannt werden. Akustische Phänomene manifestieren sich schließlich immer in der Zeit, weshalb diese Dimension die eigentliche Grundvoraussetzung jedes Sprech- und Hörvorgangs ist (Kranich 2016: 23f.). Die Dauer muss also aus zwei Perspektiven betrachtet werden. Zum einen realisiert sich das Sprechen wie jede akustische Erscheinung immer in der Zeit und stellt so den Zeitraum dar, in dem Sprache auditorisch wahrnehmbar ist. Zum anderen lassen sich im Sprechen selbst signalhafte und signalfreie Phasen unterscheiden, welche sich als Artikulationsrate und deren Fluktuationen oder als Pausen manifestieren (Kranich 2016: 34). Akustisch wird unter Dauer die gemessene zeitliche Länge von Signalabschnitten verstanden, auditiv die subjektiv bewertete Sprechzeit, also die Zeit, die verwendet wird, um das Gesagte zu erzeugen. Dabei hängt die Dauer von der Sprechgeschwindigkeit und auch von der Wortlänge ab. Sie wird im Deutschen als distinktives Merkmal der Vokale gebraucht (lange vs. Kurze Vokale, z. B. Staat – Stadt) und dient lautübergreifend auch der Hervorhebung in Wort, Wortgruppe und Äußerung (Bose/Hirschfeld/Neuber/Stock 2016: 39).

2.1.2.7 Pausen

Als Pause wird allgemein die Zeit bezeichnet, die schallphysikalisch als akustische Nullphase in Erscheinung tritt (Kranich 2016: 25). Sie spielen für das Verständnis und die Verarbeitung des Gesprochenen eine wichtige Rolle und sind ein wichtiges Gliederungs- und Strukturierungssignal, welches semantische und gleichzeitig rhythmisch-melodische Einheiten voneinander abgrenzt. Pausen können gefüllt sein (z. B. mit Atmung, Gestik oder Mimik oder Füllwörtern wie „hm“ und „äh“) und werden im Normalfall durch andere Merkmale begleitet, die auch ohne Schallabwesenheit als Grenzsinal fungieren können (Bose/Hirschfeld/Neuber/Stock 2016: 39). Durchschnittlich alle 11 bis 13 Silben wird eine gefüllte oder stille Pause gemacht (Bose 1994: 121). Diese Pausentypen sind funktional nicht gleichwertig – gefüllte Pausen zählen zu den Gesprächspartikeln und signalisieren, dass der Sprecher bzw. die Sprecherin die Sprecherrolle ergreifen oder behalten möchte und dauern zusammen mit anderen Verzögerungen umso länger, je größer die Komplexität der beabsichtigten Mitteilung ist. Beide Pausentypen kommen oft auch kurz nach dem Redeeinsatz vor (Schwitalla 2012: 75). Aus Sprecher*innenperspektive ergeben sich vielfältige Gründe für Pausen, so z. B. physiologische Gegebenheiten (Atmung, Husten, Gähnen usw.), psychische Gründe oder sprechplanerische Gesetzmäßigkeiten (Kranich 2016: 25). Auch Planungslücken, Denktätigkeit oder die Suche nach Argumenten können im Rahmen komplexer Sprechprozesse zu Pausen führen (Meinhold 1995: 21). Natürlich können Pausen auch rhetorisch genutzt werden, um die Aufmerksamkeit der Zuhörer*innen zu sichern, um Spannung vor einer wichtigen Information aufzubauen oder um unwilliges bzw. vorsichtiges Sprechen anzuzeigen (Schwitalla 2012: 76). Zu unterscheiden sind Pausen („Haltepause“), die eine Äußerungseinheit unterbrechen und Pausen („Auslaufpause“), die nach einer Äußerungseinheit eine segmentierende Funktion ausüben, wobei erstere länger sind als letztere (Bose 1994: 108). Beachtet werden müssen hier auch kulturelle Unterschiede, denn Pausen nach einem Sprecherbeitrag können zu Problemen führen mit Personen, die wesentlich kürzere oder längere Pausen gewöhnt sind. Schweigen kann auch ein Zeichen für eine Verstehensschwierigkeit oder für Nicht-Zustimmen sein, sodass der Sprecher bzw. die Sprecherin eine Reparatur beginnt und zusätzliche Informationen einräumt (Uhmann 1997a: 78-85). Pausen nehmen zusammen mit anderen prosodischen Merkmalen (z. B. leises, tiefes Sprechen) am Ende von thematischen Einheiten außerdem zu und zeigen an, dass das alte Thema ausdiskutiert ist und ein neues Thema begonnen werden kann (Meise 1996: 39ff.). Unter bestimmten Bedingungen erhält eine Pause auch die interaktive Bedeutung einer Verweigerung (Baldauf 2002: 167ff.) und wird nach einer Äußerung mit aufforderndem Charakter als „Schweigezug“ interpretiert (Meise 1996: 57ff.).

2.1.2.8 Stimmfärbung (Timbre)

Für dieses komplexe prosodische Merkmal werden unterschiedliche Begrifflichkeiten verwendet, die hier synonym Gebrauch finden. Seitens der Akustik wird meist der Terminus Klangfarbe verwendet, wobei die Referenz auf die visuelle Modalität aus der Farblehre entlehnt ist und die Klangfarbe also akustisch eine Zusammensetzung verschiedener Harmonischer ist, die mit einer bestimmten Frequenz über eine bestimmte Dauer auftreten. Damit handelt es sich also um einen spektralen Raum (Kranich 2016: 26). Die auditorische Größe hingegen wird oft als Klangqualität bezeichnet, wobei für dieses Merkmal sprecherische (Larynx, Ansatzrohr) und sprachliche (Formanten) Determinanten konstituierend sind (Kranich 2014a: 37).

Bose [u. a.] sprechen von Stimmqualität bzw. Stimmausdruck (akustisch Teiltonstruktur) und weisen darauf hin, dass sie stark individuell geprägt ist (Bose/Hirschfeld/Neuber/Stock 2016: 39). Kranich bezeichnet das komplexe Phänomen als Klangqualität und definiert es akustisch als die Konfiguration bestimmter Teiltöne mit jeweils verschiedenen Intensitäten als Funktion der Zeit. Unter funktionalem Aspekt zeige sich das in der Wahrnehmung spezifischer Sprachlaute sowie indexikalischen und biophonetischen Merkmalen. Als weitere Kennzeichen nennt er das Verhältnis tonaler vs. geräuschhafter Anteile, spektrale und zeitliche Hüllkurven, Formanten und Ein- und Ausschwingvorgänge (Kranich 2016: 35). Scherer spricht von Klangfarbe bzw. Timbre und nennt es das subtilste und zugleich „geheimnisvollste“ Merkmal der menschlichen Stimme und Sprache, welches auf das Engste mit den Tiefenschichten der Persönlichkeit verbunden ist. In ihm kämen selbst die differenziertesten und sublimsten seelischen Regungen zum Ausdruck. So läge in der Klangfarbe am meisten „Seele“, sie sei am stärksten vom materiellen Untergrund gelöst und sei „Ausdruck des Freien und Spirituellen – und damit wahrhaft „Göttlichen“ in uns (Scherer 1982: 145)“. Unter Klangfarbe kann nach Wolff akustisch die Art der Zusammensetzung eines Sprechtones aus dem Grundton und seinen verschiedenen Obertönen (Partialtönen) verstanden werden. Diese Mischung, welche die Klangganzheit in der ihr eigentümlichen Farbqualität ergibt, kann bei den einzelnen Vokalen und Konsonanten variieren. Außerdem verändert sich der Klang der Stimme auch in Hinblick auf die Verwendung des resonanzreicheren, kräftigeren Brustregisters oder der resonanzärmeren, klangdünneren Regionen (Wolff 1942). Nach Helmholtz (1877) ist das Timbre hingegen nur von der Anzahl, Lage und Stärke der Partialtöne abhängig – so klängen obertonarme Zusammensetzungen matt und leer, obertonreiche Mischungen dagegen farbig. Scherer weist auf die physiologischen Variationsmöglichkeiten zwischen den einzelnen Individuen hin. Das Timbre ist schließlich auch von der Beschaffenheit der Stimmbänder, des Kehlkopfs und v. a. des Ansatzrohres bedingt. Träger der Klangfarbe sind nach Scherer primär die Vokale, sekundär dann die stimmhaften Reibelaute

und Resonanten (Scherer 1982: 144f.). Pompino-Marschall beschreibt, dass die wahrgenommene Stimmqualität auf der Form des glottalen Schwingungsverhaltens beruht. Bei ‚behauchter Stimme‘ (engl. *breathy voice*) können die schwingenden Stimmlippen wegen ihrer geometrischen Einstellung nicht ganz schließen bzw. zeigen aufgrund ihrer zu geringen Steife nur mäßigen Widerstand gegenüber dem subglottalen Druck. So entsteht ein geräuschbehaftetes, spektral flaches und intensitätsschwaches Rohschallsignal. Eine ‚raue Stimme‘ (engl. *harsh voice*) hingegen entsteht durch Irregularitäten in der Form bzw. der zeitlichen Abfolge der Stimmlippenschwingungen aufgrund der damit verbundenen akustischen Auswirkungen eines variablen Grundfrequenz- (engl. *jitter*) bzw. Intensitätsverlaufs (engl. *shimmer*). Das gemeinsame Auftreten der Merkmale der behauchten und der rauhen Stimme kennzeichnet schließlich die ‚heisere Stimme‘ (engl. *hoarse voice*) (Pompino-Marschall 2009: 40f.).

Andere Arten der Klangfarbe sind sehr schwer zu beschreiben, dabei ist diese oft wichtiger als die anderen prosodischen Merkmale. Das Problem ist zum einen die große Differenzierung psychischer Zustände und zum anderen die genaue Bestimmung der prosodischen und artikulatorischen Faktoren im Einzelnen, die beim Hörer bzw. der Hörerin einen manchmal vagen, manchmal eindeutigen Eindruck eines Gefühls entstehen lassen. Schon beim Telefonieren kann rasch festgestellt werden, in welcher Stimmung sich der Sprecher bzw. die Sprecherin befindet. Belustigt, verärgert, fröhlich, genervt, erstaunt und viele weitere sind Alltagsbezeichnungen für solche gefühls- und einstellungsbedingte Sprechweisen. Manche sind auch an Rollenstereotypen festgemacht, z. B. pastoral, lehrerhaft, militärisch etc. oder an sozialen Merkmalen, wie vornehm, rockerartig, preziös usw. Es liegen aber erst wenige linguistische Arbeiten über Gefühlsexpressionen in natürlichen Interaktionen vor, die intensiv auf prosodische Phänomene eingehen (Schwitalla 2012: 79). Auch deshalb wird hier versucht, näher auf diese komplexe Thematik einzugehen.

Kranich (2016: 26) bezeichnet die Klangqualität als eine der interessantesten prosodischen Merkmale, da ihr zahlreiche Funktionen zukommen, sie sich aber akustisch nicht auf eine Größe reduzieren lässt. Trotz ihrer unbestrittenen Bedeutung liegen nur sehr wenige systematische Untersuchungen zu diesem Merkmal vor. Zu dessen Bestimmung schlägt er einen minimalen Merkmalskatalog mit möglichst großer Trennschärfe vor und unterteilt folgende Kategorien: Gespanntheit (unterspannt, neutral, gespannt), Klangfülle (klangarm, neutral, klangvoll), Helligkeit (dunkel, neutral, hell), Knarren (kein Knarren, Knarren) (Kranich 2014b: 34). Bose hingegen schlägt folgenden Katalog vor: Klangfülle, Klangfarbe, faukale Distanz, Stimmklang (Ein-, Ausschwingphasen), Geräuschhaftigkeit, Stimmansatz, Stimmklangmodifikation

(zeitlich, umfänglich, formal) (Bose 2003: 413). Scherer (1982: 145) hat die wichtigsten unterscheidenden Merkmale der Klangfarbe als Gegensatzpaare bzw. –gruppen zusammengefasst und differenziert zwischen folgenden Kategorien: a) hell – dunkel, wobei der helle Klang Durcharakter und der dunkle Mollcharakter habe; b) farbig-sonor, glänzend, klar – stumpf, rauh, bedeckt, verschleiert. Hierbei geht es um die faukale Distanz, die Trojan beschreibt. Der farbig-sonore, glänzende, klare Klang resultiere aus der (aus dem Gefühl der Lust heraus) resultierenden Rachenweite, der sog. „faukale Weite“. Das stumpfe, rauhe, abgedeckte Klangbild hingegen entstehe, verbunden mit einem Gefühl der Unlust und inneren Spannung, durch das Zusammenziehen des Ansatzrohres und v. a. der hinteren Gaumenbögen („faukale Enge“) (vgl. Trojan 1975: 69ff.). Im letztgenannten Klangbild können außerdem der gutturale (kehlige) Stimmklang, das sog. Näseln und die heisere, belegte Stimme (als Erregungsheiserkeit, z. B. bei Zorn, oder habituell als Gewöhnungsheiserkeit) unterschieden werden. Des Weiteren nimmt Scherer die Trennung von c) warmem, weichem, melodischem vs. kaltem, hartem, scharfem, schrillum, schneidendem Stimmklang vor. Als Varianten des kalten Stimmklangs werden außerdem der gepresste Stimmklang (bedingt durch bestimmte Atmung, den hohen Spannungsgrad und die „faukale Enge“) und das schnarrende, rasselnde Sprechen (z. B. zum Karikieren von Standespersonen) genannt. Zuletzt werden noch die Pole d) zart, aufgelockert – leer, dürr, drahtig, spitz angeführt (Scherer 1982: 145f.).

In dieser Untersuchung wird eine eigene, der Konstellation der Studie angepasste Klassifizierung, v. a. auf Grundlage des Merkmalkatalogs von Bose, vorgenommen.

Die Stimmqualität bzw. Klangfarbe erfüllt viele wichtige Funktionen, sie wird vom Hörer bzw. der Hörerin rezipiert als Signal bzw. Teilsignal für den Ausdruck modaler und emotionaler Befindlichkeiten (pathognomischer Ausdruck) und Persönlichkeitseigenschaften (physiognomischer Ausdruck, z. B. vermutetes Alter, psychophysischer Zustand etc.) (Bose/Hirschfeld/Neuber/Stock 2016: 39). Außerdem dient die vokalische Klangfarbe in einigen Sprachen auch als wichtiges Element der Akzentuierung. So haben z. B. das Englische und Russische besondere Vokalinventare für betonte Silben und andere für unbetonte Silben. Auch im Deutschen spielt die Klangfarbe eine gewisse Rolle bei der Akzentuierung, so kommen die Vokale [ɐ] und [ə] nur in unbetonten Silben vor. In anderen Sprachen, z. B. auch dem Italienischen, spielt die Klangfarbe bei der Akzentuierung jedoch kaum eine Rolle (Pétursson/Neppert 2002: 157f.).

Diese Beschreibung der prosodischen Merkmale zeigt die unterschiedlichen definitiven Ansätze sowie deren Relevanz für die vorliegende Arbeit auf und lässt die Bedeutung der prosodischen Parameter in allen Aspekten der gesprochenen Sprache sowie deren vielfältigen Funktionen linguistischer und nicht-linguistischer Art erkennen. Dabei wurde insbesondere auf die emotionssignalisierende Funktion der prosodischen Mittel hingewiesen, die im Zentrum dieser Arbeit steht.

2.1.3 Sprechausdruck

Um die Frage nach der etwaigen Musterhaftigkeit von prosodischer Perzeption emotionaler Sprechweisen adäquat bearbeiten zu können, bedarf es auch hier zunächst einer begrifflichen Erläuterung. Dabei soll an der Tradition der sprechwissenschaftlichen Gesprächsforschung angeknüpft werden, da die Methoden und Forschungserfahrungen in diesem Bereich auch linguistische und gesprächsanalytische Arbeiten bereichern. Aufgreifen möchte ich hier den im Vergleich zum Begriff Intonation weiter gefassten Begriff des Sprechausdrucks. Schließlich stehen in dieser Arbeit die auditive Deskription von Sprechausdrucksmerkmalen durch externe Hörer*innen und die Rekonstruktion der interaktiven Relevanz der Sprechausdrucksgestaltung in Bezug auf die Emotionsperzeption im Mittelpunkt. Außerdem wird angenommen, dass die Sprechmuster in engstem Zusammenhang mit den Hörmustern der Kommunizierenden stehen, weshalb ein kurzer Blick auf dieses Phänomen vonnöten scheint.

Der Terminus Sprechausdruck ist ein Oberbegriff für konventionelle Gestaltungsweisen und -muster im Sprechschall, die durch Merkmale wie Lautheit, Sprechstimmhöhe, Stimmklang, Artikulation (und ihre jeweiligen Veränderungen), Sprechgeschwindigkeit und Merkmalskomplexe wie Gliederung, Akzentuierung, Sprechspannung und Sprechrhythmus und andere zustande kommen (Bose 2001: 262). An sprachliche Formulierungen gebundene bzw. sie komplettierende sprecherische Ausdrucksformen, die zwar individuell ausgeführt, aber stimmungs- und situationsadäquat realisiert sind, werden mit dem Begriff Sprechausdruck bezeichnet. Wie der in der Gesprächslinguistik verwendete Terminus Sprechstil (vgl. Sandig/Selting 1997) verweist der Begriff Sprechausdruck zwar auch auf die Interdependenz von Sprachlichem und Sprecherischem, aber hier stehen die im Bereich des Sprechschalls möglichen Gestaltungsweisen und -muster im Mittelpunkt (Bose 2001: 265f.).

In konkreten Kommunikationsereignissen werden sog. Sprechausdrucksweisen bzw. Formulierungsweisen produziert und rezipiert. Wenn sich bestimmte Sprechausdrucksweisen regelmäßig wiederholen und an ähnliche Bedingungen geknüpft sind, sie also geordnet und regelhaft erscheinen, dann liegen vermutlich Sprechausdrucksmuster in den Köpfen der Kommunizierenden vor. Gutenberg spricht von Formulierungsmustern, die er als „Strukturmuster, nach denen der Schall der konkreten Sprechereignisse gebildet wird“, bezeichnet (Gutenberg 1998: 25f.). Dazu zählen unter anderem rollen- und gruppenspezifische sowie habituelle oder medienbedingte Muster (Gutenberg 2001: 126).

Im Sprechen selbst lassen sich der pathognomische, d. h. situativ bedingte Sprechausdruck, vom physiognomischen, d. h. angeborenen und gewohnheitsmäßigen Sprechausdruck, nicht so einfach differenzieren (Bose 2001: 266). Schließlich wird Sprachliches und Nichtsprachliches gleichzeitig wahrgenommen und sowohl als pathognomischer und physiognomischer Ausdruck bewertet, wobei nach Stock die Bewertung des pathognomischen Ausdrucks von der des physiognomischen abhängt. Das bedeutet, dass eine konkrete Sprechweise in einer bestimmten Situation vor dem Hintergrund der individuellen-spezifischen Sprechweise „als eine Variante erfasst [werde], die erst durch die Bedingungen oder Erfordernisse der Situation ausgelöst wurde und die dadurch als situationsbezogener Ausdruck verstanden werden kann (Stock 1987: 57f.)“. Auch innerhalb eines Kulturkreises, einer Sprachgemeinschaft oder einer sozialen Gruppe bilden sich bestimmte Erwartungsvorstellungen über das mögliche bzw. zulässige Ausdrucksverhalten heraus (Stock 1987: 58). Diese Überlegungen liegen u. a. auch den zahlreichen Studien zugrunde, die emotionale Sprechmuster beschreiben und ggwf. interkulturell vergleichen, wie im Abschnitt zum Einflussfaktor Kultur näher gezeigt wird. Außerdem bedingen die Sprechmuster wohl auch die Hörmuster der Kommunizierenden in wechselseitigem Einfluss. Als Hörmuster können in diesem Zusammenhang allgemeine Gesetzmäßigkeiten bezeichnet werden, also mental gespeicherte, internalisierte Abstraktionen charakteristischer prosodischer Konfigurationen, welche eine überindividuell-soziale Prägung aufweisen (Bose 2003: 19; Kranich 2016: 242). Hörmuster sind Präferenzen für bestimmte Laute und Silben, die bereits im Kindesalter beim Erlernen der Muttersprache im Gehirn gebildet werden (Lipkowski/Schüller 2017: 34).

So wird angenommen, dass der Sprechausdruck in der Ontogenese auf der Grundlage von angeborenen Grundmustern stimmlichen Ausdrucks durch Anregung und kommunikative Vorbilder gelernt wird. Während der vorsprachlichen Phase bilden die stimmlichen Ausdrucksfähigkeiten der Bezugspersonen und deren kommunikationsbezogenen Interpretationen der kind-

lichen Laute die Basis in der Verständigung mit dem Säugling. Über die Deutung und Produktion dieser Ausdrucksformen kommen die Kinder also zur Sprache. Bereits kleine Kinder verfügen über eine gewisse Sprechausdruckskompetenz, die ihre sprachliche Kompetenz übersteigt und eine wichtige kommunikative Rolle spielt. So verfügen die Kinder über v. a. implizite Kenntnisse über und Fähigkeiten zur situations- und stimmungsabhängigen Variation von Sprechausdrucksweisen, die als Regulationsbasis für kommunikatives Handeln dienen und in Bezug zur jeweiligen Kommunikations- und Kulturgemeinschaft stehen (Bose 2001: 266f.).

2.2 Prosodie und Emotion

2.2.1 Begriff Emotion

Eine erste Schwierigkeit wissenschaftlicher Forschungsbemühungen zum Themenkomplex Emotion und Prosodie liegt darin, das Konstrukt wissenschaftlich exakt zu beschreiben. Ein Grund dafür ist wohl die Tatsache, dass der Forschungsgegenstand in einem Überschneidungsbereich der zwei wissenschaftlichen Disziplinen Sprachwissenschaft und Psychologie liegt. Problematisch dabei ist, dass in beiden Einzeldisziplinen zu den Begriffen Prosodie bzw. Emotion bisher kaum ein breiter Konsens erreicht worden ist, vielmehr konkurrieren zahlreiche Theorien und Definitionen (vgl. Kehrein 2002). Der Terminus Prosodie wurde in den vorigen Kapiteln schon näher beleuchtet, nun soll der Begriff Emotion für den Rahmen dieser Arbeit kurz umrissen werden.

Begriffsdefinition

Während nun eine wissenschaftliche Untersuchung nicht durchgeführt werden kann, ohne das Forschungsgebiet abzugrenzen, so ist doch gleichzeitig eine exakte Definition des untersuchten Phänomens nicht zwingend nötig bzw. auch gar nicht möglich. Tatsächlich gibt es bis heute keine allgemein anerkannte Antwort auf die Frage, was eine Emotion überhaupt sei. Um im Rahmen dieser Untersuchung Emotionen (bzw. emotionale Sprechweisen) betrachten zu können, wird also ungefähre Charakterisierung in Form einer Arbeitsdefinition angestrebt, um den Untersuchungsgegenstand grob abzugrenzen und die Leser*innen darüber zu verständigen, welche Phänomene untersucht werden.

Zunächst muss zwischen verschiedenen Bezeichnungen für das Phänomen Emotion unterschieden werden. Stimmungen und Gefühlszustände werden oft als kleinere, alltägliche Emotionen

verstanden, (vgl. Isen 1984) wobei Stimmungen in der Literatur häufig hinsichtlich ihrer Intensität und Objektbezogenheit von geringerer und bezüglich ihrer Dauer von größerer Ausprägung als Emotionen angesehen werden (vgl. Scherer 1990). Meyer et al. (1993) halten diese Unterscheidungskriterien jedoch für fragwürdig und definieren die Abgrenzung von Stimmung und Emotion als problematisch. Weiters abzugrenzen sind die Begriffe Affekt (vgl. Schwarz-Friesel 2013), welcher in der englischsprachigen Literatur oft als Oberbegriff für Emotionen verwendet wird (Meyer et al. 1993), sowie Gefühl. Während schon McDougall (1928/1969) eine Differenzierung der Begriffe Gefühl und Emotion vorschlug, haben andere Autoren diese wiederum synonym gebraucht (vgl. Scherer 1989). In dieser Forschungsarbeit werden die Begriffe Emotion, Gefühl, Affekt und ähnliche synonym verwendet.

Im Rahmen seiner Komponenten-Prozess-Theorie schlägt Scherer (1993: 4) eine zeitgemäße Arbeitsdefinition von Emotion vor, die ihre Bedeutung durch den großen Einfluss auf die Forschung erlangt hat und an der ich mich auch hier orientieren möchte. In Anlehnung an Scherers Arbeitsdefinition von Emotion wird Emotion also in dieser Arbeit verstanden als eine „Episode zeitlicher Synchronisation aller bedeutender Subsysteme des Organismus, die fünf Komponenten bilden (Kognition, physiologische Regulation, Motivation, motorischer Ausdruck [motor expression] und Monitoring/Gefühl), und die eine Antwort auf die Bewertung eines externalen oder internalen Reizereignisses als bedeutsam für die zentralen Bedürfnisse und Ziele des Organismus darstellt (Scherer 1990: 6; Übersetzung vgl. Otto [u. a.] 2000: 5).“ Scherer fasst Emotionen hauptsächlich als eine Schnittstelle zwischen der Umwelt und dem Organismus auf, d. h. Emotionen vermitteln zwischen ständig wechselnden Umweltsituationen und –ereignissen und dem Individuum. In der vorgestellten Definition ist die kognitive Bewertung (appraisal) ein zentraler Prozess, der für die Entstehung und auch für die kontinuierliche und rekursive Informationsverarbeitung (reappraisal) einer Emotion als wesentlich betrachtet wird (Otto [u. a.] 2000: 4f.). Dieser Aspekt ist von größter Relevanz im Zusammenhang mit der Frage nach der kulturell geprägten Emotionsperzeption. Ein weiteres wichtiges Merkmal von Emotionen ist deren Gerichtetheit auf reale oder imaginierte Objekte und deren Episodenhaftigkeit, wobei diese aktuellen Episoden meist von relativ kurzer Dauer sind (Meyer [u. a.] 1993: 26f.). Daher sind die ausgewählten emotionalen Sprechweisen auch eher kurz.

2.2.2 Emotionstheorien: universell vs. kulturell

Tatsächlich dachte man lange, der Emotionsausdruck sei universell. Charles Darwin vertrat als Erster explizit diese Position, er schreibt „the different races of man express their emotions and sensations with remarkable uniformity throughout the world (Darwin 1872/1998: 130-1)“. Viele Wissenschaftler*innen griffen diesen Standpunkt auf, Emotionen seien biologisch im menschlichen Körper verankert und damit universell. William James, der sich um eine globale Emotionstheorie bemühte, ging davon aus, dass Emotion die subjektive Erfahrung der physiologischen Veränderungen im Körper sei (James 1884). Wenn also Emotionen aus den grundlegenden biologischen Prozessen im Körper hervorgehen, sollten sie für alle Menschen gleich sein und selbiges gilt natürlich auch für den Emotionsausdruck. Anfang des zwanzigsten Jahrhunderts war die Idee der biologisch verankerten und universell geltenden Emotionen allgemein anerkannt. Paul Ekman und Carroll Izard z. B. konnten empirisch nachweisen, dass Emotionen in unterschiedlichen Kulturkreisen sehr gut erkannt werden (Ekman/Sorenson/Friesen 1969; Izard 1971). Wenngleich die Ergebnisse heute umstritten sind, haben sie die nachfolgenden Forscher*innen stark beeinflusst. Silvan Tomkins schreibt in seiner neo-darwinschen Emotionstheorie, dass menschliche Emotionen mit jenen im Tierreich vergleichbar und interkulturell gültig seien (Tomkins 1962). In vielen Tiergesellschaften erfüllen vokale Signale eine für die soziale Organisation äußerst wichtige Kommunikationsfunktion und aufgrund der angenommenen phylogenetischen Kontinuität emotionalen Ausdrucks finden sich auch beim Menschen diskrete Affektvokalisationen, die an die Rufsysteme der Tiere erinnern. Allerdings treten diese oft nur in Situationen extremer emotionaler Erregung auf, wenn die Intensität der neuronalen Erregung in subkortikalen Gehirnregionen die hemmende und kontrollierende Einwirkung des Großhirns außer Kraft setzt (Scherer 1982: 282).

Neben der Annahme eines universell gültigen Emotionsausdrucks kam es nun zu einem Wandel in der Emotionspsychologie. Einige Wissenschaftler*innen legten die Bedeutung der kognitiven Komponente in der Emotionsbewertung nahe (appraisal) (Arnold 1960; Lazarus 1968; Lazarus/Coyne/Folkman 1984, Scherer 1982). Schon Wundt wies darauf hin, dass es kulturell verbindliche Ausdrucksregeln gibt (Wundt 1903: 285). Die kognitive Fähigkeit ist weitaus komplexer und entwickelter als die basalen physiologischen Prozesse und während sich zwar nicht die kognitive Fähigkeit an sich in den verschiedenen Kulturen unterscheidet, so doch die Denkweise der Personen. Dies suggeriert, dass Emotionen doch kulturell geprägt sind. Während Biolog*innen und Psycholog*innen eher Gemeinsamkeiten suchen, konzentrieren sich

Anthropolog*innen und Ethnolog*innen eher auf Unterschiede in Bezug auf Emotionen in diversen Kulturen (Scherer [u. a.] 2011: 404). Untersuchungen zu kulturellen Differenzen in Denkweise, Sprache und Bräuchen zeigten auch potentielle Unterschiede im kulturell geprägten Emotionsausdruck auf. Ray Birdwhistell und Margaret Mead z. B. argumentierten, emotionaler Ausdruck müsse wie Sprache als Komplex von Zeichen verstanden werden, die gelernt und kulturell geprägt sind (Barfield 1997). Aus diesem Grund können in der interkulturellen Kommunikation auch Missverständnisse oder Konflikte entstehen. Besonders Emotionen wie Ärger, Frustration, Nervosität oder Traurigkeit werden interkulturell unterschiedlich gehandhabt und bilden Konfliktpotenzial. Um zu einem besseren gegenseitigen Verständnis zu gelangen, ist das Wissen um den potentiell diversen Umgang mit Emotionen hilfreich (Matsumoto [u. a.] 2007: 80f.). Matsumoto [u. a.] plädieren daher auf die Ausbildung bestimmter Kompetenzen, um interkulturelle Konflikte lösen bzw. adäquat miteinander kommunizieren zu können (Matsumoto [u. a.] 2007: 77). Scherer legt in seiner Appraisal-Theorie einen Schwerpunkt auf die kognitive Komponente der Situationsbewertung. Ekman/Friesen prägten schließlich den Begriff „display rules“, der eben auf die sozio-kulturelle Fundierung der Emotionen verweist. Die Autoren gehen davon aus, dass der Emotionsausdruck kontrollierbar sei und nennen hierfür vier Gründe: „cultural display rules“, „personal display rules“, „vocational requirements“ (z. B. bei Diplomaten oder Schauspielern) und „need of the moment“ (z. B. vor Gericht) (Ekman/Friesen 1975: 138f.). Der Sprach- und Kommunikationswissenschaftler Fiehler nimmt nicht nur Manifestationsregeln an, sondern postuliert darüber hinaus auch Emotionsregeln (kodifizieren, welches Gefühl in bestimmten Situationen sozial erwartbar und angemessen ist), Kodierungsregeln (Konventionen bestimmen die Verhaltensweise zur Emotionsmanifestation) und Korrespondenzregeln (kodifizieren, welche Emotionsmanifestationen bei mir in Bezug auf die Emotionalität meines Gegenübers angemessen oder erwartbar sind) (Fiehler: 1990b: 46;80). In dieser Arbeit werden also, an den Überlegungen von Scherer u. a. anschließend, Unterschiede zwischen den deutschsprachigen und italienischen Proband*innengruppen sowie zwischen den Expert*innen- und Laienhörer*innen in der Emotionsperzeption für wahrscheinlich gehalten.

2.2.3 Theoretische Konzepte und Modelle

Folgend sollen einige ausgewählte, für dieses Thema relevante theoretische Konzepte und Modelle vorgestellt werden.

2.2.3.1 Scherer: Integrative Theorie

Trotz des theoretischen Umschwungs waren empirische Arbeiten zum Thema bis vor Kurzem noch hauptsächlich in der universellen Tradition verhaftet und ließen kulturelle Differenzen eher außen vor. Erst in jüngster Zeit ist ein Wandel dahingehend zu beobachten, beide Annahmen zu beleuchten und empirisch aufzugreifen (Scherer [u. a.] 2011: 404f.). Es können also nach Scherer zwei Faktoren genannt werden, die emotionale Episoden beeinflussen – zum einen sind dies die sog. push-Effekte, d. h. die physiologischen Veränderungen, die die Emotion begleiten, und zum anderen die sog. Pull-Effekte. Darunter werden alle Merkmale verstanden, die als soziale Faktoren Einfluss auf die Persönlichkeitsentwicklung des Individuums hatten (Scherer 1989b: 166). Die kognitive Komponente spielt in Scherers Überlegungen also eine zentrale Rolle. Er versucht die Bewertung von Reizen (appraisal), die jeweils spezifische emotionale Reaktionen auslösen, zu systematisieren. Dieser Ansatz soll hier kurz vorgestellt werden, da der Situations-/Kontextbezug im Rahmen dieser sowie anderer sprach- und kommunikationswissenschaftlicher Untersuchungen von Emotionen relevant ist, da die Situationsbewertung eine zentrale Rolle spielt. Das folgende Schema bietet eine Übersicht zu den Funktionen, Subsystemen und Komponenten von Emotionsprozessen, wobei in Anlehnung an Kehrein gegenüber dem Original der Verständlichkeit wegen von oben nach unten Auslöser, Reaktions-trias und Folgen angegeben sind (Kehrein 2002: 106f.).

Funktionen	Subsysteme	Komponenten
Reizbewertung	Informationsverarbeitungssystem	Kognitive Komponente
Systemregulation	Versorgungssystem	Neurophysiologische Komponente
Reflexion und Kontrolle	Monitorsystem	Gefühlskomponente
Kommunikation von Reaktion und Intention	Aktionssystem	Ausdruckskomponente
Handlungsvorbereitung	Steuerungssystem	Motivationale Komponente

Abb. 1: „Organismische Subsysteme in Beziehung zu Funktionen und Komponenten von Emotionen (vgl. Scherer 1990: 4; Kehrein 2002: 107)“

Der Systematisierungsversuch der Reizbewertung ist als Komponenten-Prozess-Modell bekannt. Die Grundidee besagt, dass dynamisch aktive Prozesse als Muster von Zustandsänderungen auftreten, die so in den relevanten Subsystemen des Organismus konzeptualisiert werden können. Der Prozess des Affektes wird als Sequenz eines Komplexes betrachtet. Die Bewertung von Ereignissen bzw. Situationsveränderungen verlaufe in Form von sequenziell angeordneten Reizverarbeitungsschritten (engl. Stimulus Evaluation Checks – SEC). Scherer postuliert fünf Gruppen von SECs, die z. T. in weitere subchecks gegliedert werden.

1. Neuigkeitsprüfung des Inputs
2. Behaglichkeitsprüfung des Inputs für das Individuum
3. Ziel-/Nutzenanalyse für das Individuum
4. Überprüfen der Bewältigungsmöglichkeiten
5. Überprüfung hinsichtlich sozialer Normen und des Selbstkonzepts

Die multisystemen Mikrozustände ändern sich ständig und dadurch auch die Konfiguration in allen Subsystemen, so können Emotionsmischungen entstehen und verschiedene expressive Ausprägungen vorausgesagt werden. Vorhersagen in der Art eines umfassenden Reaktionskomplexes gestalten sich jedoch lückenhaft (Kranich 2003: 45f.; vgl. Scherer 1993).

2.2.3.2 Linguistische Ansätze

Wie auch viele Primaten die Möglichkeit einer sehr feinen Abstufung der akustischen Charakteristika der Rufe mit Übergängen von einem Ruftyp zum anderen in Bezug auf den Emotionsausdruck haben, wodurch komplexe Mischungen verschiedener emotionaler Inhalte und unterschiedliche Ausprägungen übermittelt werden, so spielt auch beim Menschen die Einwirkung der emotionalen Prozesse auf die Lautproduktion beim Sprechen eine zentrale Rolle. Anstelle der kontinuierlichen Variation akustischer Merkmale bei diskreten Rufen finden sich beim Menschen kontinuierliche Veränderungen der akustischen Merkmale der Sprachlaute, oft als ‚paralinguistische Phänomene‘ bezeichnet. Diese erfüllen neben der emotionalen Ausdruckswirkung auch semantische oder syntaktische Funktionen. Die emotionalen Zustände wirken sich auf die Sprachproduktion, die akustischen Charakteristika und damit auch auf die Perception emotionaler Sprechweisen aus (Scherer 1982: 282f.). Drach stellt sogar die Behauptung auf, alles Sprechen sei ursprünglich den Affektausdruck entwachsen. Die Schwächung dieser Ausdrucksmerkmale sei auf die Schulbildung, den Sozialisationsprozess und die Erziehung zurückzuführen (Drach 1926: 14). Ausgehend von der Überzeugung der Relevanz emotionaler

Faktoren für die Sprachproduktion und -rezeption hat sich in der Vergangenheit auch die Linguistik mit diesem Thema auseinandergesetzt, wobei deren Forschungsansatz natürlich primär auf Erkenntnissen der Emotionspsychologie basiert. Teilweise wurden jedoch auch eigene Ansätze zum Emotionskonzept entwickelt (Kranich 2003: 42f.).

Nach Janney z. B. sind Emotionen in zwei grundlegende Prozesse verwurzelt. Der individuelle Prozess subsummiert alle physiologischen, psychischen und neuronalen Prozesse, die die Emotionen auslösen und steuern. Er ermöglicht, dass diese durch diskrete Muster innerhalb linguistischer Komponenten ausgewählt und ausgedrückt werden (Janney 1990: 40ff.). Die Gruppenprozesse hingegen üben einen interpersonalen, sozialen und kulturellen Einfluss aus. Sie sind erlernt worden und bestimmen, was als normal gilt. Der emotive Ausdruck sei ein soziales und strategisches Muster menschlicher Interaktion. Dabei besteht zwischen der Wahrnehmung der Kommunikationssituation in ihrer Komplexität und den individuellen inneren Prozessen eine Wechselbeziehung, weshalb die Kommunikation von Emotionen weder als rein instinktiv noch als ausschließlich konventionell betrachtet werden kann (Janney 1990: 10). Ausgehend von Janneys Aussagen besteht die Möglichkeit, dass auch geschlechtermarkierte Unterschiede in der Emotionsperzeption möglich sind, da Geschlechterrollen als soziale Einflussfaktoren auf das Individuum eine große Rolle spielen.

2.2.3.3 Klassifikation von Emotionen

Was die Klassifikation von Emotionen betrifft, so gibt es keinen einheitlichen Ansatz (Kehrein 2002: 112). Neben dem Komponenten-Prozess-Modell von Scherer werden hauptsächlich dimensionale und diskrete Modelle diskutiert.

Die dimensionalen Erklärungsmodelle gehen auf Wundt zurück, der selbst einen Ansatz Spencers aufgriff und erweiterte und schon 1874 die Emotionen in ein dreidimensionales System einordnete (Kranich 2003: 45). Er benannte die Dimensionen mit Lust/Unlust, Erregung/Ruhe und Spannung/Lösung (Wundt 1910: 294-301). Bis auf die letztgenannte sind diese Dimensionen auch heute noch unter anderen Begrifflichkeiten gebräuchlich. Lust/Unlust entspricht dem Terminus Valenz, wobei die Valenzskala die Polarität positive – negative Befindlichkeit umfasst, und Erregung/Ruhe kann unter der Bezeichnung Aktivierung/Erregung subsummiert werden (Kehrein 2002: 111). Auch Schlossberg bedient sich in seinem Dimensionsmodell dieser Skalen, nämlich der Stärke-Skala (Achtung-Abwehr-Dimension), der Valenzskala (Polarität positiv/angenehm – negativ/unangenehm) und der Aktivitätsdimension (Schlaf-

Spannung). Der Schwachpunkt der Dimensionsmodelle liegt in der vorausgesetzten Orthogonalität der Achsen, tatsächlich können diese nicht unabhängig voneinander betrachtet werden (zit. nach Kranich 2003: 45).

Die Vertreter der diskreten Emotionsmodelle hingegen versuchen Emotionen kategorial einzuteilen. Dabei postulieren sie ein Set von Basisemotionen, deren Zahl zwischen zwei und 18 schwankt. Diesen Prototypen werden andere Emotionen oder Mischformen zugeordnet. Die Basisemotionen, die am häufigsten genannt werden, sind Ärger bzw. Wut, Freude, Angst, Traurigkeit und Ekel. Manche Untersuchungen unterscheiden zwischen Basisemotionen und Zivilisatorischen Emotionen, die nur bestimmten Zivilisationen bekannt sind, wobei Ekel oft zu letzteren gezählt wird (Kehrein 2002: 109). Scherer z. B. konnte in einer Studie mittels Fragebogen die Verbreitung von vier Grundemotionen in acht europäischen Ländern beweisen. In abnehmender Häufigkeit wurden dabei Ärger, Freude, Traurigkeit und Wut ermittelt (Scherer 1986b: 174). Trotz kultureller Unterschiede in Bezug auf die Realisation der Emotion ließen sich diverse Stereotype erkennen, die universal verbreitet sind.

Auch in diesem Ansatz der Basisemotionen gibt es keine einheitliche Klassifikation. Allerdings zählen zahlreiche Autoren (z. B. Ekman 1992: 174; Johnson-Laird [u. a.] 1989: 105; Schmidt-Atzert 1987: 140; Scherer 1986b: 174) die Emotionen Ärger und Freude zu den Basisemotionen und daher habe ich mich in dieser Studie zur Verwendung von ärgerlichen und freudigen Sprechweisen entschieden.

2.2.3.4 Das Linsenmodell von Brunswik in einer Adaption von Scherer

Relevant für diese Arbeit und alle Untersuchungen, die an der Schnittstelle von Emotion und Prosodie interessiert sind, ist das brunswiksche Linsenmodell in einer Version von Scherer. Zur Beschreibung prosodischer Ausdrucksphänomene finden sich in der Literatur zwei methodisch diverse Herangehensweisen. Zum einen wird aus der Sprecherperspektive das Ziel verfolgt, anhand der Beschreibung konkreter akustischer Parameter objektive Aussagen über prosodische Veränderungen in der Stimme und Sprechweise nachzuweisen. Zum anderen wird aus der Hörerperspektive durch Perzeptionstests versucht, Aussagen über die Wahrnehmung emotionaler Sprechweisen zu treffen. Das Linsenmodell von Brunswik (1956) wurde von Scherer (1978) für die vokale Charakterisierung des Emotionsausdrucks einerseits und für die Wahrnehmung von Emotionen andererseits adaptiert und ermöglicht dadurch eine Synthese beider Herangehensweisen (Redecker 2006: 35).

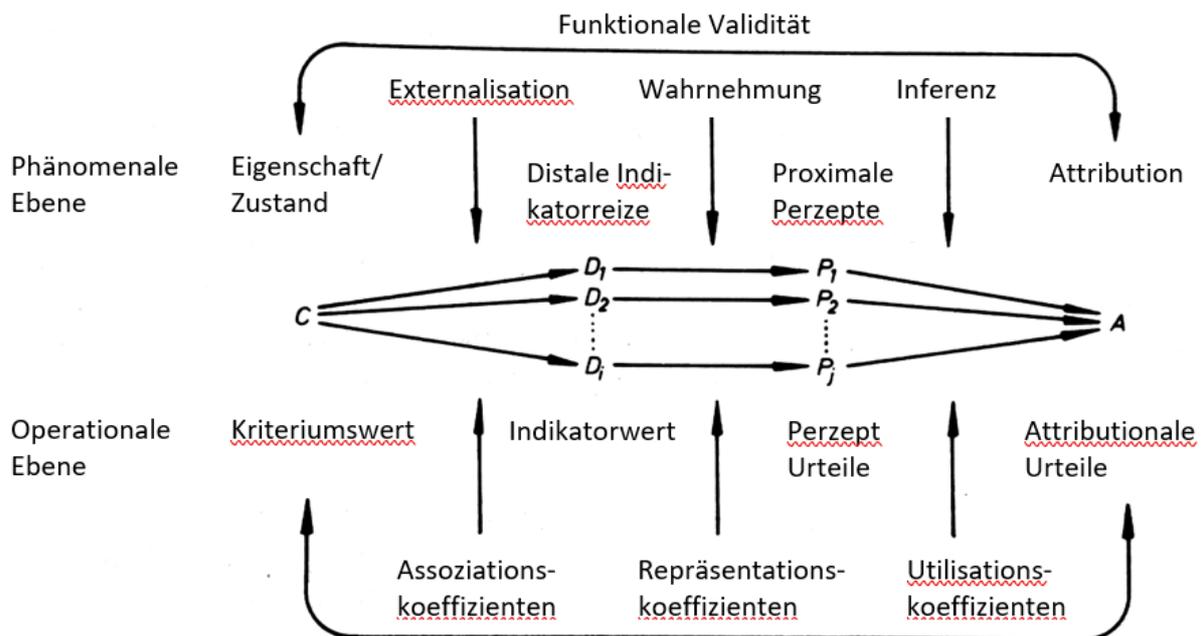


Abb. 2: Modifizierte Version des brunswikschen Linsenmodells nach Scherer (Scherer 1982: 189)

Scherer gliedert den Inferenzprozess auf Phänomenaler Ebene in die Sprecherbefindlichkeit („Eigenschaft/Zustand“), die objektiven Parameter, die durch die Sprecherbefindlichkeit beeinflusst werden („Distale Indikatorreize“), die subjektiv wahrgenommenen Korrelate dieser Parameter („Proximale Perzepte“) sowie die zugeschriebene Sprecherbefindlichkeit („Attribution“). Die sich aus den einzelnen Punkten als Bindeglieder ergebenden Phasen der „Externalisation“, „Wahrnehmung“ und „Inferenz“ verdeutlichen die Prozesshaftigkeit des Modells. Scherer nimmt an, dass sich stabile Eigenschaften und vorübergehende Zustände in der Form von distalen Hinweisreizen im Verhalten und in der Erscheinung eines Individuums „externalisieren“. Darunter versteht der Autor sowohl die absichtliche Kommunikation von Zuständen oder Eigenschaften durch verbales oder nonverbales Verhalten als auch „unabsichtliche“ physiologische oder verhaltensmäßige beobachtbare Reaktionen.

Das Modell (Abb.2) schematisiert außerdem den Expressions- und Perzeptionsprozess zweier Personen. Eine Person (A) perzipiert die emotionalen Attribute einer anderen Person (C) anhand externer Sprechermerkmale, wobei der Sprecher/die Sprecherin (C) durch habituelle und aktuelle Merkmale näher spezifiziert wird. Diese Merkmale können durch die Stimme und Sprechweise ausgedrückt werden. Die habituellen Merkmale betreffen die Persönlichkeit oder

Einstellung des Sprechers/der Sprecherin und werden als sehr stabil eingeschätzt, bei den aktuellen Merkmalen hingegen handelt es sich um situationsabhängige und variable Merkmale, wie sie besonders beim Ausdruck von Emotionen verwendet werden (Redecker 2006: 36).

Auf der Operationalen Ebene sind Eigenschaften und Zustände als „Kriteriumswerte“ und distale Hinweisreize als „Indikatorwerte“ bezeichnet. Die empirische Kovariation zwischen Kriterien und Indikatoren stellt Scherer durch den „Assoziationskoeffizienten“ dar. Die distalen Hinweisreize sind Perzepte, weil sie aus dem Wahrnehmungsprozess resultieren. Sie können operational z. B. durch Beurteilungen gemessen werden, wobei die Korrelationen zwischen Indikatorwerten und Perzepturteilen als „Repräsentationskoeffizienten“ bezeichnet werden. Sie zeigen an, mit welcher Genauigkeit distale Hinweisreize perzipiert werden. Auf die Perzepte distaler Hinweisreize folgen Inferenzprozesse, woraus Attributionen von Eigenschaften oder Zuständen resultieren. Diese Attributionen werden z. B. durch Beurteilungen von Beobachtern mittels psychologischer Dimensionen ermittelt. In Anlehnung an Brunswik (1956) werden die Korrelationen zwischen perzeptuellen und attributionalen Einschätzungen als „Utilisationskoeffizienten“ genannt. Sie zeigen, inwieweit ein Hinweisreiz bei der Inferenz benutzt bzw. wie er gewichtet wird.

Die Korrelation zwischen Kriteriumswert und attributionalem Urteil ist die Genauigkeit, mit der Attributionen die realen Eigenschaften bzw. Zustände des beobachteten Individuums repräsentieren. Diese Korrelationen bezeichnet Scherer als „Genauigkeitskoeffizienten“.

Methodisch empfiehlt Scherer eine Synthese aus objektiven und subjektiven Untersuchungsmethoden (akustische Analyse und Perzeptionstest), um den Fragestellungen zur Korrelation von der Stimmqualität und der Persönlichkeit eines Sprechers nachzugehen (Scherer 1982: 189f.).

Dieser Ansatz und der im Rahmen des modifizierten brunswikschen Linsenmodells aufgezeigte methodische Weg Scheres können auch im Rahmen dieser Arbeit aufgegriffen werden.

Scherer entwickelt das vorgestellte Modell, um methodisch aufzuzeigen, wie Schlussfolgerungen über den Emotionszustand einer Person anhand stimmlicher und sprecherischer Merkmale gezogen werden können. In seiner Untersuchung (1978: 467ff.) verwendet der Autor das Modell als methodische Grundlage, um Persönlichkeitseigenschaften aufgrund der Stimmqualität einzuschätzen. Dabei stellt er fest, dass sowohl amerikanische als auch deutsche Beurteiler*innen bei deutschen Sprecher*innen Dominanz und bei amerikanischen Sprecher*innen Extraversion mit überzufällig großer Wahrscheinlichkeit vorhersagen konnten. Auch in späteren Forschungsarbeiten greift Scherer das Modell auf, z. B. um eine Systematik für die Einordnung der

Forschungsliteratur zu entwickeln. Diese kann untergliedert werden in drei verschiedene wissenschaftliche Ansätze, wobei meistens Elemente aus verschiedenen Methoden zusammengeführt werden. Erstens unterscheidet Scherer also Arbeiten, die den Enkodierungsvorgang fokussieren, zweitens Arbeiten im Schnittpunkt von Enkodierung und Dekodierung, die den Übergang von prosodischen Parametern zu Perzeptionskonzepten untersuchen und drittens Arbeiten, die das Dekodieren in den Mittelpunkt des Interesses rücken (Redecker 2006: 38). Auch Schröder nimmt diese Einteilung in einer Vereinfachung vor, wenn er sprecherzentrierte Studien (Enkodierung) und hörerzentrierte Studien (Dekodierung) differiert (Schröder 2004: 37). Forschungen im Schnittbereich zwischen Stimme, Sprechweise und Emotion greifen jedoch im Normalfall immer sowohl Enkodierungs- als auch Dekodierungsaspekte auf. So werden akustische Korrelate gemessen und auch durch Perzeptionstests überprüft, um objektiv und subjektiv gewonnene Daten im Ergebnis zu synthetisieren. Meistens wird in den Untersuchungen jedoch ein Aspekt näher fokussiert (Redecker 2006: 38).

2.2.4 Forschungsüberblick Prosodie, Emotion, Perzeption

Im folgenden Abschnitt sollen einige wichtige Forschungen zum Themenkomplex Emotion und Prosodie genannt werden, dabei werden Enkodierungs- als auch Dekodierungsuntersuchungen berücksichtigt. So soll der Stand der Forschung kurz umrissen werden. Es sollen außerdem exemplarisch einige Autoren und Autorinnen genannt werden, die für die vorliegende Arbeit unter diversen Gesichtspunkten besonders relevant sind und oder die mit ihren Forschungen die inhaltlichen und methodischen Grundsteine im relevanten Themenkomplex gelegt haben bzw. den entsprechenden Forschungsbereich vorantreiben. Die Auswahl hier ist natürlich als überaus lückenhaft und auf die vorliegende Untersuchung bezogen anzusehen.

In linguistischen wie in sprechwissenschaftlichen Publikationen zeigt sich, dass dem Sprecher bzw. der Sprecherin und der Botschaft bis heute viel mehr Aufmerksamkeit zuteil wurde als den Rezipient*innen (Kranich 2015: 107). Lediglich in der Sprechwirkungsforschung wird mit den Folgen der Botschaft bei den Hörer*innen ein relevanter Aspekt der mündlichen Kommunikation thematisiert (vgl. Stock 1991: 9ff.). Der Fokus zukünftiger Arbeiten sollte sich also auch vermehrt auf das „Wie“ des Perzeptionsvorgangs richten.

In den letzten Jahrzehnten wurde im Rahmen der Bemühungen um eine automatische Spracherkennung der Forschung zur Sprachwahrnehmung viel Aufmerksamkeit geschenkt. Trotz vie-

ler Fortschritte und wertvoller Erkenntnisse aus experimentellen Designs, z. B. mit synthetischem Material, ist die maschinelle Erkennung der menschlichen Rezeptionsfähigkeit noch weit unterlegen (Wendemuth 2004: 193f.). Das liegt unter anderem daran, dass diverse Menschen dieselbe Emotion auf unterschiedliche Art und Weise ‚lesen‘. Diese Differenzen gilt es also zu untersuchen und zu verstehen (Cowie/Douglas-Cowie 2009: 217). Zudem wird gerade in jüngerer Zeit betont, dass die Prosodie in diesen Studien eine zentrale Rolle spiele, weil sie gerade für die Erkennung emotionaler Sprechweisen grundlegend ist. Schließlich ist die paraverbale Ebene das „Vehikel“, mittels dessen Sprache „transportiert“ wird (Kranich 2015: 107). Im Rahmen einer Studie zur automatischen Spracherkennung wurde von Schuller/Wöllmer/Eyben/Rigoll (2009:302f.) der Versuch gestartet, diejenigen prosodischen Merkmale zu eruieren, die besonders relevant für die Emotionserkennung sind. Das Resultat zeigte, dass v. a. die spektralen Eigenschaften der gesprochenen Sprache ausschlaggebend sind. Im Rahmen einer anderen Untersuchung zur automatischen Spracherkennung wurde die Tonhöhe bzw. der Tonhöhenverlauf als relevantestes Merkmal zur Erkennung von natürlichen Emotionen genannt, allerdings in Kombination mit anderen prosodischen Merkmalen (Busso/Bulut/Lee/Narayanan 2009: 328f.).

Viele Wissenschaftler*innen weisen gerade der Prosodie einen besonderen Status für den Ausdruck und die Kommunikation von Emotionen zu (vgl. Selting 1994). Nicht vergebens finden sich in Wörterbüchern für deutsche Redewendungen etwa Einträge wie „der Ton macht die Musik“, „in den höchsten Tönen“, „was ist das für ein Ton“, „das ist wie Musik in meinen Ohren“ und viele weitere (Kehrein 2002: 1). Die Funktion der Prosodie ist sowohl aus linguistischer als auch aus psychologischer Sicht, wie bereits im vorherigen Abschnitt dargestellt, zum Teil sogar evolutionär begründet worden (vgl. Meyer et al. 1997). Demnach können vokale und mimische Muster als Universalien des affektiven Ausdrucksverhaltens einer Reihe von Basiseemotionen aufgefasst werden. Entsprechend sollte beim Menschen die Fähigkeit zur Identifikation und Klassifikation emotionaler Prosodie gut ausgeprägt sein, was in diversen wissenschaftlichen Arbeiten belegt werden konnte (z. B. Scherer 1982). Allerdings handelt es sich bei den meisten bisherigen Ergebnissen um Untersuchungen mit experimentellen Situationen, in denen professionelle Sprecher*innen, z. B. Schauspieler*innen, kontextfreie Phrasen mit meist gleichem Wortlaut vorlesen und bestimmte Emotionen vorführen. Die Generalisierbarkeit auf freie Kommunikationssituationen erscheint so äußerst fragwürdig, die Ergebnisse müssten an natürlichen Gesprächen überprüft werden (Hielscher 2003: 683).

Daher möchte ich meine empirische Untersuchung zur Perzeption emotionaler Sprechweisen durch prosodische Parameter auf natürliche Äußerungen basieren.

In der Psychologie, Sprechwissenschaft und experimentellen Phonetik gibt es sehr viele Untersuchungen zu Korrelationen von Gefühlszuständen und prosodischen Eigenschaften. Neuber z. B. lässt das Kunstwort [kataˈkalama] in 33 Varianten sprechen, um drei Prototypen festzustellen: freundlich-fragend, wütend/ärgerlich/aggressiv sowie wütend-schimpfend (Neuber 2002: 58ff.). Drescher analysiert französische Gespräche und findet heraus, dass die prosodischen Formen für Gefühlstypen sehr unspezifisch sind und eher die höhere Relevanz der emotionalen Erregung anzeigen (Drescher 2001: 190). Kehrein untersucht 2002 den Zusammenhang von Prosodie und Emotionen unter den Fragestellungen, weshalb die Hörer*innen bestimmte Äußerungen im Kontext als emotional wahrnehmen und ob es prosodische Einheiten mit genuin emotionalen Bedeutungen gibt (Kehrein 2002: 2). Dazu erzeugt er im Rahmen des Lego-Korpus möglichst natürliche Emotionen und stellt fest, dass mehrere Signalisierungssysteme an der Konstitution von Emotionswahrnehmungen beteiligt sind und die Hörer*innen einzelne Emotionen ganzheitlich als Kompositions-/Kontextbedeutungen wahrnehmen. Außerdem eruiert er folgende prosodische Einheiten mit emotionaler Bedeutung: lokales Intonationsmuster zum Ausdruck positiver Befindlichkeit, Sprechtempo, Prominenz und erhöhtes F₀-Maximum. So korrelieren z. B. beim Eindruck ‚freudig‘ hohe Intonation und lange Dehnung bei unverminderter Lautstärke der akzentuierten Silbe (Kehrein 2002: 321ff.). Paeschke geht es im Rahmen ihrer empirischen Untersuchung mit über 500 Äußerungen von Schauspieler*innen, die die Emotionen Freude, Ärger, Angst, Trauer und Ekel simulieren, darum, spezifische Merkmale der Sprechweisen zu identifizieren, welche bei den Hörer*innen den Eindruck bestimmter Emotionen hinterlassen (vgl. Paeschke 2003). Kranich eruiert diverse prototypische prosodische Muster emotionaler Sprechweisen auf der Grundlage eines kombiniert auditiv-instrumentalphonetischen Verfahrens (Kranich 2003: 10f.). Bisherige Ergebnisse dazu fasst er übersichtlich ab S. 50 zusammen. In dieser Arbeit stütze ich mich auf die von ihm eruierten phonetischen Charakteristika von Ärger (Temporalität: hohe Sprechgeschwindigkeit; Tonhöhereigenschaften: großer Tonraum, mehr Abwärtsschritte; Intensitätseigenschaften: geringe Dynamikbreite; Stimmklang: hell, scharf, klangarm, Behauchung) und Freude (Temporalität: hohe Sprechgeschwindigkeit; Tonhöhereigenschaften: viele Tonschritte, große Tonschritte, hohe Sprechstimmlage; Intensitätseigenschaften: kein Ergebnis; Stimmklang: hell, weich, warm, faukale Weite) (Kranich 2003: 198). Allerdings gilt zu beachten, dass in Kranichs Studie simulierte expressive Sprechweisen von Schauspieler*innen, Sänger*innen und Lehrer*innen untersucht wurden

(Kranich 2003: 79). Die Gültigkeit der eruierten phonetischen Charakteristika der emotionalen Sprechweisen kann also nicht problemlos auch für natürliche emotionale Sprechweisen angenommen werden.

Es gibt auch einige Untersuchungen, in denen akustische Stimmanalysen in realistischen Situationen durchgeführt wurden. Diese beschäftigen sich hauptsächlich mit der Erfassung bzw. Kontrolle des Emotionszustandes von Pilot*innen oder Raumfahrer*innen während des Fluges. Williams und Stevens z. B. analysierten eine Unterhaltung zwischen einem Piloten und einem Fluglotsen im Kontrollturm, dabei lagen ernsthafte Schäden am Flugzeug vor und die Maschine stürzte ab. Eine weitere Aufnahme (1969), die analysiert wurde, ist die des Radiosprechers, der den Brand und Absturz des Zeppelins Hindenburg beschreibt (Williams/Stevens 1982: 321ff.). Wendt entwickelte ein Sprachkorpus zur Erforschung der Verarbeitungsmodalitäten gesprochener Sprache, wobei das Stimulusmaterial den speziellen Anforderungen nicht-invasiver Messmethoden der kognitiven Neurowissenschaft entsprechen sollte und in Form einer Sprachperzeptionsstudie mit der funktionellen Kernspintomographie angewendet werden sollte (Wendt 2006: 11). Redecker nimmt eine Untersuchung vor, die der Frage nach der Perzeption prosodischer Merkmale im Persuasionsprozess einer Parfumwerbung nachgeht und findet heraus, dass die Veränderung stimmlicher und sprecherischer Merkmale verschiedener Sprechweisen eines Werbeslogans zu dessen veränderten Wahrnehmung und Wirkung führt (Redecker 2006: 135). Ein Team von Wissenschaftler*innen um den Psychologen Ralf Rummer und die Phonetikerin Martine Grice hat Experimente zum Zusammenhang von Sprache und Emotionen durchgeführt, um herauszufinden, inwieweit die Bedeutung von Wörtern mit deren Klang zusammenhängt. Dabei wurden Versuchspersonen in negative oder positive Stimmung versetzt und gebeten, Kunstwörter zu bilden und auszusprechen. Diese enthielten mehr /i:/s als /o:/s, wenn die Proband*innen positiv gestimmt waren und umgekehrt mehr /o:/s, wenn diese negativ gestimmt waren. Die Sprachverwender lernen also, dass die Artikulation bestimmter Vokale mit gewissen Gefühlszuständen korreliert und tendieren also zur entsprechenden Sprachverwendung (Rummer/Schwepe/Schlegelmilch/Grice 2014: 246). In einer Studie von Chen/Gussenhoven/Rietveld wurde die Perzeption von Intonationskurven diverser Stimuli in Englisch (britisch) und Niederländisch von jeweils muttersprachlichen Hörer*innen untersucht. Diese sollten die Stimuli in den Kategorien selbstsicher-nicht selbstsicher, freundlich-unfreundlich, überrascht-nicht überrascht und emphatisch-nicht emphatisch einordnen. Es konnten relevante Unterschiede in der Perzeption zwischen Engländer*innen und Niederländer*innen festgestellt werden (Chen/Gussenhoven/Rietveld 2004: 311). Neben Kehrein, Kranich, Grice, Gussenhoven und den weiteren in diesem Abschnitt bereits angeführten Wissenschaftler*innen und

Wissenschaftlern müssen noch folgende wichtige Autor*innen (Auswahl) genannt werden, die im Themenkomplex Sprachperzeption-Emotion-Prosodie forschend tätig sind bzw. waren: Julia Hirschfeld, sie untersuchte z. B. die Perzeption von Charisma anhand bestimmter akustischer, prosodischer und lexikalisch-syntaktischer Charakteristika (Hirschfeld/Rosenberg 2009: 640); Janet B. Pierrehumbert, sie konnte z. B. in einer Studie herausfinden, wie Erwartungen an die prosodische Realisierung von Äußerungen die Perzeption beeinflussen (Zheng/Pierrehumbert 2010: 851) und Robert Ladd, der sich u. a. mit der Phonologie der Intonationssysteme und der komparativen Phonetik beschäftigte (vgl. Ladd 2008).

Zahlreiche Autor*innen betrachten also den Zusammenhang von psychischen Gestimmtheiten und Prosodie. Dabei liegt das Augenmerk der Forscher*innen aber meist entweder auf der phonetischen Beschreibung und Klassifizierung der unterschiedlichen Gefühlsexpressionen oder auf der Fähigkeit des perzeptorischen Zuordnens der gehörten Phänomene zu bestimmten Emotionen, ohne jedoch näher auf diesen Prozess einzugehen. Es erscheint jedoch notwendig, die Thematik unter neuen Gesichtspunkten aufzugreifen und zu ergründen, schließlich hat sie nichts an ihrer Relevanz für die Linguistik, Psychologie und Sprechwissenschaft und auch andere Disziplinen verloren. Sie gewinnt im Gegenteil sogar immer mehr auch in angrenzenden Disziplinen, wie z. B. der Neurowissenschaft und Sprachdidaktik an Bedeutung. So wird in den Neurowissenschaften intensiv der Zusammenhang von Emotion und Sprache erforscht, wobei etwa die enge Verbindung von Sprache und Kognition nachgewiesen werden konnte (vgl. u. a. Borod 2000, Panksepp 2003, Damasio 2005). Ausgehend von der Entdeckung der Spiegelneuronen (Rizzolatti et al. 1996) wird die Verbindung von Emotion und Sprache in der „Innate Intersubjectivity Theorie“ dargelegt und durch die Gruppe um Trevarthen (vgl. Trevarthen 2005) weiter erforscht. So konnte etwa herausgefunden werden, dass bereits Neugeborene durch ihre neurobiologische Bereitschaft der IMF (Intrinsic Motive Formation) und des EMS (Emotional Motor System) relationale Emotionen spiegeln und intentional kommunizieren können. Die Fähigkeit der Emotionsperzeption ist also bereits pränatal psychophysiologisch angelegt (vgl. Trevarthen 2005). In der kognitiven Neurowissenschaft, die sich als Schnittstelle zwischen Psychologie und Hirnforschung versteht und sich mit Bereichen wie Gedächtnis- und Lernfunktionen, Sprachfunktionen, Emotion und Emotionsregulation und kindliche Entwicklung beschäftigt, ist die Prosodie ebenfalls als Forschungsgegenstand relevant (Wendt 2006: 10).

Auch die bilingual bzw. kontrastiv ausgerichtete Emotionsforschung im Hinblick auf sprachdidaktische Implikationen ist aktuell äußerst relevant, für Lehr-/Lernmethoden ebenso wie für die

Aufbereitung von Arbeits-, insbesondere Audiomaterial zu Unterrichtszwecken (Zur Problematik im Zweit- und Fremdsprachenerwerb vgl. Missaglia 2001; zum Zusammenhang von Sprache, Emotion und Didaktik vgl. Lüdtker 2006). Das Modell der „Relationalen Didaktik“ (Lüdtker 2002a, 2004, 2004b, 2005a, 2006a-c) etwa legt den Fokus auf die intersubjektive, bedeutungskonstruierende Rolle der Emotionen und betont im sprachlichen Lehr- Lernprozess entsprechend die Unterstützung durch einen professionellen emotionalen Kontext. Erwähnenswert sind hier nicht nur die Implikationen für den (Fremd-) Sprachenerwerb, sondern auch für Möglichkeiten der Sprachtherapie. Auch im Rahmen der KI (Künstlichen Intelligenz) und für die Entwicklung von Spracherkennungsprogrammen könnten die Ergebnisse dieser Untersuchung eine Rolle spielen. Zukünftig wird es immer mehr Schnittstellen zwischen Computern bzw. digitalen Systemen und dem Menschen geben, wobei die gegenseitige sprachliche Kommunikation an enormer Relevanz gewinnt. Für eine Kommunikationssoftware könnte auch die Interpretation prosodischer Merkmale von Bedeutung sein, z. B. bei robotischen Helfern im Pflegedienst oder in anderen sozial sensiblen Bereichen.

2.3 Prosodie und Sprachperzeption

Die Prozesse der Perzeption gesprochener Sprache werden von verschiedenen Wissenschaftsdisziplinen unterschiedlich beleuchtet. Es gibt verschiedene Ansätze, sich diesem komplexen Phänomen zu nähern, es gibt jedoch keine einheitliche, allgemein gültige Gesamtdarstellung desselben (vgl. Miller/Johnson-Laird 1976: 33).

2.3.1 Perzeptive Phonetik

Die auditive Sprachwahrnehmungsforschung bzw. die perzeptive Phonetik entstanden durch die Entwicklung der entsprechenden elektroakustischen Geräte und wurde durch deren technische Probleme und Anforderungen erst stimuliert. Abgesehen von frühen Experimenten im Bereich der akustischen Phonetik mit dem Ziel, mit dem Ohr akustische Eigenschaften zu messen (z. B. Wolf 1871), waren es bis zur Mitte des letzten Jahrhunderts hauptsächlich die Telefongesellschaften, v. a. die Bell-Laboratorien in den USA, welche die Sprachwahrnehmung untersuchten. Dabei ging man vor allem der Frage nach, inwieweit und wie das Sprachsignal zur bestmöglichen Nutzung der Kapazität der Übertragungsmedien reduziert werden kann. Auf der

Grundlage der damals durchgeführten Verständlichkeitsforschungen wird bis heute beim Telefon nur das Frequenzband zwischen 300 Hz und 3 kHz übertragen (Pompino-Marschall 2009: 158).

Auch die eigentliche perzeptive Phonetik entstand im Kontext der technischen Entwicklung, genauer gesagt aufgrund eines Fehlschlags. Die Haskins-Laboratorien in New York verfolgten in den 50er Jahren das Ziel, ein Lesegerät für Blinde zu entwickeln und verfolgten zu diesem Zweck die Absicht, die Buchstaben des Textes durch spezielle akustische Signale zu chiffrieren (ähnlich wie beim Morsen). Dadurch konnten die Übertragungsraten der alltäglichen lautsprachlichen Kommunikation jedoch in keiner Weise erreicht werden. Die auditorische Sprachwahrnehmung muss also auf andere Art und Weise erfolgen. Die perzeptive Phonetik wurde zu einer neuen Wissenschaftsdisziplin innerhalb der Phonetik (Pompino-Marschall 2009: 159). Sie ist auch ein Teilgebiet der Wahrnehmungspsychologie und wird auch Psychophonetik genannt. Methodisch wird so vorgegangen, dass ein bestimmter Redeabschnitt sowohl instrumentell analysiert als auch von Proband*innen beschrieben wird. Dann wird das Abhörresultat mit der akustischen Substanz verglichen und das Verhältnis zwischen auditivem Erfassen und physikalischer Wirklichkeit ausgewertet. Mit der Entwicklung der Computer stehen natürlich viel differenziertere Forschungsmittel zur Verfügung (Wendt 2007: 31). Durch die technische Möglichkeit der Erzeugung synthetischer Sprache konnten die Perzeptionsindizien der Sprache so erstmals systematisch untersucht werden. Der Vorteil der vereinfachten und von zufälligen Abweichungen freie akustische Struktur des synthetischen Sprachmaterials liegt darin, dass einzelne Signalelemente isoliert oder in definierter Kombination systematisch variiert und auf ihre sprachperzeptuelle Relevanz hin erforscht werden können. So konnte festgestellt werden, dass in Zuordnung zu sprachlichen Einheiten im akustischen Signal die bis dahin angenommene Invarianz nicht vorhanden ist. Das gilt auch dann, wenn sprachperzeptiv jeweils die gleiche invariante sprachliche Einheit erkannt wird. Demonstriert werden kann das in einem bekannten Beispiel aus der Locus-Forschung, das sich um die Transitionen der Silben [di] und [du] handelt. Während in der Silbe [di] die Mittenfrequenz des F2 von 2000 auf 2400 Hz steigt, fällt sie in der Silbe [du] von 1200 auf 800 Hz. Hier liegt also eine klare akustische Diskontinuität vor, die aber sprachperzeptiv jeweils eindeutig als [d] wahrgenommen wird (Pétursson/Neppert 2002: 191). Die Locus-Theorie führt diese Tatsache darauf zurück, dass im Gehirn des Hörers der Locus des /d/, also ein zwischen 1200 Hz und 2000 Hz liegender /d/-spezifischer Wert ermittelt wird (Jussen/Kloster-Jensen/Wisotzki 1994: 72ff.). Andererseits gibt es auch Fälle, in denen bei der Schallerzeugung für bestimmte Signalelemente akustische Invarianz gegeben war, die perzeptiv aber als verschiedene sprachliche Einheiten wahrgenommen wurden. Die

Deutung der Wissenschaftler*innen hierzu ist unterschiedlich, als wesentlich kann festgestellt werden, dass das akustisch gleiche Sprachsignal unterschiedlich wahrgenommen und verschiedene akustische Ereignisse als ein und dieselbe sprachliche Einheit kategorisiert werden können (Pétursson/Neppert 2002: 191f.). Wie die Decodierung genau vor sich geht, entzieht sich bis jetzt unserer Kenntnis (Wendt 2007: 32).

2.3.2 Theorien zur Sprachperzeption

Sprachverstehen scheint also viel mehr zu sein als eine bloße Umkehr des Sprachproduktionsprozesses. Hierzu gibt es aber unterschiedliche Auffassungen – teilweise werden die Vorgänge der Sprachperzeption als eine Umkehrung der produktiven Vorgänge betrachtet (z. B. Schade 2003: 84), andere Vertreter*innen sprechen sich gegen diese komplementäre Betrachtungsweise aus (vgl. Imhof 2003: 115; Dietrich 2007: 212; Prestin 2003: 491). Selbst möchte ich mich der letztgenannten Auffassung anschließen. Derselbe Satz kann von unterschiedlichen Rezipient*innen verschieden verstanden und in unterschiedlichen Situationen anders interpretiert werden. Verstehen ist demnach ein konstruktiver und kontextabhängiger Prozess. Welche kognitiven Prozesse mit dem Verstehen verbunden sind, erforscht die psycholinguistische Sprachrezeptionsforschung. Sie unterscheidet die Ebenen des perzeptuellen Verstehens, des morpho-syntaktischen Verstehens, des semantischen und des pragmatischen Sinnverstehens (vgl. Rickheit/Strohner 1993). Dabei herrschen unterschiedliche Auffassungen darüber vor, ob und inwieweit diese Ebenen miteinander interagieren.

2.3.2.1 Bottom-up-Ansatz der Verstehenstheorie

Vertreter*innen der autonomen Prozessmodelle gehen davon aus, dass die Verarbeitung auf den einzelnen Ebenen isoliert bzw. autonom abläuft und die vollständigen Ergebnisse einer Ebene an die nächste weitergegeben wird. Nach Abschluss der rein sprachlichen Verarbeitung würde das Weltwissen der Hörer hinzugezogen, um die Information einzuordnen und zu gewichten (Prestin 2003: 491). Die Sprachrezeption wird also als das serielle (nacheinander erfolgende) Erkennen bzw. Verstehen von Phonemen, Morphemen, Wörtern, Sätzen und Texten beschrieben, wobei die Bedeutung schrittweise extrahiert wird. Diese modular-serielle Annahme wird auch als Bottom-up-Ansatz der Verstehenstheorie bezeichnet und wird gewöhnlich in der Psycholinguistik vertreten (Bose/Hirschfeld/Neuber/Stock 2016: 7f.). Sie wird nach Sa-

downik auch durch neurowissenschaftliche Befunde stark validiert. Demnach unterstützen bestimmte Gehirnareale die Sprachverarbeitungsprozesse und syntaktische, semantische und prosodische Prozesse können einer festgelegten Hirnstruktur zugeordnet werden (Sadownik 2014: 63). Bekannt geworden ist inzwischen das neurokognitive Modell des Sprachverstehens von Friederici (1995). Sie nimmt an, dass der kognitive Prozess auf den gleichen Ebenen wie bei der Sprachproduktion verläuft, allerdings in umgekehrter Richtung. Probleme treten dabei auf bei der Frage, wie Phoneme identifiziert werden, schließlich variieren deren Realisationen (Allophone) oft stark akustisch und artikulatorisch, sind oft nicht segmentierbar und müssen z. T. phonologisch zwei oder mehr Phonemen zugeordnet werden. Dadurch ist eine autonome Identifizierung nicht möglich. Es scheinen also keine eindeutigen Beziehungen zwischen Signalmerkmalen und Wahrnehmungskategorisierung zu bestehen. Nur interaktive Abläufe ermöglichen, dass auch bei Mehrfachzuordnungen der Allophone Phoneme ohne Schwierigkeit erkannt werden. Auch Ergebnisse aus Experimenten der Mitte des letzten Jahrhunderts sprechen gegen eine strenge seriell-autonome Verarbeitung. Bei der Wahrnehmung von Wörtern zeigte sich nämlich, dass die Phonemidentifikation abgebrochen wird, wenn das Wort an seiner Gestalt, seiner prosodischen Form oder aus dem Zusammenhang heraus erraten werden konnte. Dies erfolgt also bereits am Anfang des Verarbeitungsprozesses und nicht, wie von Friederici angenommen, in späteren Phasen (Bose/Hirschfeld/Neuber/Stock 2016: 7f.). Friederici schränkt selbst die Gültigkeit des Bottom-up-Modells ein und erklärt, dass nach den bisherigen Verhaltensdaten nicht gesagt werden kann, ob seriell oder interaktive Prozesse ablaufen. Neuere physiologische Daten weisen darauf hin, dass einzelne Komponenten in der ersten Verarbeitungsphase zunächst autonom arbeiten und dann in einer späteren Phase miteinander interagieren (Friederici 2006: 346ff.).

2.3.2.2 Top-down-Ansatz der Verstehenstheorie

Verfechter*innen der interaktiven Prozessmodelle hingegen nehmen einen ständigen Austausch der Ebenen während der laufenden Verarbeitung an. Auch das Weltwissen sei von Anfang an zentral und beeinflusse direkt die phonologische, morpho-syntaktische, semantische und pragmatische Analyse des Sprachsignals (Prestin 2003: 491). Diese gegenläufige Theorie wird auch als Top-down-Verarbeitung bezeichnet. Hier werden interaktive Abläufe angenommen, d. h. zentral gesteuerte Antizipationen beeinflussen die Reizanalyse, welche bei Übereinstimmung abgebrochen wird (Rickheit [u. a.] 2007: 108ff.). Theorien mit Top-down-Ansatz beziehen möglichst früh Weltwissen in die Verarbeitung mit ein, also z. B. Schema- oder Script-Wissen, mentale Modelle usw., und gehen davon aus, dass Rezipierende auf diesem Weg

durch den Aufbau von textbezogenen mentalen Szenarien mit Rollen und Details am ehesten das Gesagte identifizieren und erfassen können, was der Sprecher bzw. die Sprecherin meint (Bose/Hirschfeld/Neuber/Stock 2016: 8). Ein relevanter Vertreter des Top-Down-Ansatzes ist z. B. Herrmann, der im Sprachrezeptionsprozess anders als die Psycholinguistik nicht die Tätigkeit eines spezialisierten, abgekapselten, immer in Höchstform arbeitenden Apparats sieht. Die Sprachrezeption ist für ihn im Rahmen des Gesamtzusammenhangs des menschlichen Handelns und Erlebens vielmehr „eine situationsspezifisch variierende Modifikation des mentalen Gesamtsystems“, welche sporadisch auftritt und zu anderen Modifikationen dazukommt (Herrmann 2005: 251). Er erläutert hierzu, dass unser Bewusstsein im Wachzustand durch Wahrnehmungen aus der Innen- und Außenwelt laufend durch wechselnde Kognitionen, Emotionen, Volitionen, Erinnerungen, aktivierte Wissensschemata usw. durchflutet und modifiziert wird. Dies geschehe auch durch gesprochene Sprache, die im Laufe des Tages immer wieder an unser Ohr gelangt, aber unterschiedlich wichtig sei und daher verschieden genau aufgegriffen werde. Herrmann erfasst diese unterschiedliche Bedeutsamkeit mit sieben Modifikationsstufen, auf welchen sich die Rezeptionsintensität schrittweise steigere, wobei der Rezipient bzw. die Rezipientin auf Stufe eins nur wahrnehme, dass irgendwo gesprochen wird, während er auf Stufe sieben das Verhältnis von Gesagtem und Gemeintem rekonstruiere (Herrmann 2005: 251ff.). Dieser Ansatz erscheint mir auch für diese Arbeit relevant, die auf das Erkennen und Analysieren von prosodischen Merkmalen fokussiert. Auch in anderen Bereichen der mündlichen Kommunikation werden z. T. nur einzelne Aspekte der Äußerung erfasst, z. B. in der Sprechkunst oder in der Phonetik bei der Nachahmung von Lauten. Nicht bei jeder Rezeption läuft also das komplette Verarbeitungsprogramm ab. Anders als bei der Sprachproduktion sind bei der Rezeption einzelne Teilprozesse verzichtbar oder kompensierbar (vgl. Bose/Hirschfeld/Neuber/Stock 2016: 9).

Seit einiger Zeit gibt es auch situierte Modelle der Sprachverarbeitung, in denen der situativen Einbettung der Kommunikation ein großer Einfluss zugeschrieben wird (Prestin 2003: 491).

Fazit

Zusammenfassend lässt sich also konstatieren, dass die Diskussion um modulare versus holistische oder andere Ansätze zur Erklärung der neurokognitiven Organisation des Sprachverarbeitungssystems keineswegs abgeschlossen ist, sondern im Gegenteil vor vielen Herausforderungen steht. Seit der Einführung nicht-invasiver bildgebender Verfahren konnten auch die bei

der Sprachverarbeitung beteiligten neuroanatomischen Strukturen mit hoher räumlicher Auflösung besser identifiziert werden. So hat sich gezeigt, dass Sprachproduktion und Sprachverstehen nicht nur innerhalb eines Zwei-Areale-Modells stattfinden, sondern weitere Areale involviert sind. Zusätzlich zum Broca- und Wernicke-Areale wurden also neue Hirnbereiche erkannt, die für die Sprachverarbeitung ebenfalls eine zentrale Rolle spielen. Zudem belegen zahlreiche Daten zur funktionellen Neuroanatomie der Sprachproduktion und des Sprachverständnisses, dass die Funktionen dieser Sprachzentren weit komplexer sind als bisher vermutet (Sadownik 2014: 64). Die Ergebnisse aus der kognitiven Neurowissenschaft, anderen wissenschaftlichen Disziplinen und aus der Klinik helfen, die Prozesse der Sprachrezeption besser zu verstehen, und sind u. a. auch die Grundlage für die Rehabilitation und Therapie bei Sprachstörungen (Wendt 2007: 18).

2.3.2.3 Sprachperzeption als aktiver Prozess

In der Annahme, dass es sich bei der Verarbeitung und dem Verstehen von Sprache um einen konstruktiven Prozess handelt, besteht weitgehende Einigkeit (z. B. Kelter 2003: 510). Dem Verstehensprozess wird dabei mehrheitlich eine aktive Rolle zugeschrieben, einige Vertreter*innen der reinen Bottom-up-Verarbeitung betrachten ihn jedoch als passiv. Hier wird der Sprachperzeption eine aktive Rolle zugeordnet. Es gilt in diesem Zusammenhang weiters zu beachten, dass die Begriffe Hören und Perzeption nach Tatham/Morton unterschieden werden in einen passiven physikalischen Prozess in unserem Ohr, den wir kaum durch unser Denken beeinflussen (Hören) und in einen hauptsächlich aktiven kognitiven Prozess in unserem Gehirn, der Perzeption. Diese Begriffe werden oft synonym verwendet, die Unterscheidung des kognitiven und physikalischen Prozesses ist jedoch relevant (Tatham/Morton 2011: 126). Wird in dieser Arbeit vom ‚Hören‘ der emotionalen Sprechweisen (empirischer Teil, Analyse der Hörbeispiele), vom ‚Hörer‘ bzw. der ‚Hörerin‘ (bzw. Rezipient*in/Proband*in im Rahmen der durchgeführten Studie) o.Ä. gesprochen, referiere ich im Regelfall auf den komplexen Gesamtprozess der Sprachperzeption, der das physikalische Hören und das kognitive Perzipieren im Sinne Tathams/Mortons umfasst.

An dieser Stelle gilt es außerdem, das Zuhören als einen Prozess in der Zeit vom Verstehen als eher punktuellere Ereignis zu differenzieren (Kranich 2016: 84ff.). Tatsächlich kann es beim Hörer bzw. der Hörerin zu Ermüdungserscheinungen führen, wenn dieser die Information schon vor Beendigung der Äußerung kennt (Hagen 2003: 81). Daher spricht Mangold auch vom Verstehen als den „Feind des Hörers“ (Mangold 1984: 155). Die vom Weltwissen getragenen Top-down-Prozesse der Sprachperzeption können also bewirken, dass weniger genau

zugehört wird und Einzelheiten nicht realisiert werden. Die Verarbeitung allgemeiner akustischer Eigenschaften und das für den Verständnisprozess notwendige Weltwissen müssen in den Modellen zur Sprachperzeption also gleichermaßen berücksichtigt werden (Kranich 2016: 86).

2.3.3 Modelle zur Sprachperzeption

Zahlreiche Modelle und noch mehr Modellergänzungen versuchen die Wahrnehmung und Verarbeitung gesprochener Sprache zu erklären und zu beschreiben. Auch existiert eine Vielzahl an Klassifikationsversuchen der Modelltypen (z. B. in Bottom-up und Top-down-Ansätze: diese Theorien müssen aber eher als grundlegende Annahmen der Wahrnehmung betrachtet werden), wobei auch hier kein Konsens auszumachen ist (Kranich 2016: 95). An dieser Stelle sollen also nur zwei der bekanntesten Modelle kurz vorgestellt werden, um einen Eindruck über die möglichen Sprachverarbeitungsprozesse zu vermitteln. Dabei soll auf die jeweils eher aktive bzw. passive Rolle des Hörers bzw. der Hörerin hingewiesen werden.

2.3.3.1 Die Motor-Theorie der Sprachperzeption

Die Motor-Theorie der Sprachperzeption ist eine der am häufigsten zitierten aktiven Sprachtheorien, die aus der Locus-Forschung entstanden ist (Liberman [u. a.] 1963) In dieser Theorie wird eine eindeutige Beziehung zwischen dem akustischen Signal und dem Phonem hergestellt (Pétursson/Neppert 2002: 193).

In der ursprünglichen Version von 1967 betrachtete Liberman die kategoriale Wahrnehmung als Ergebnis einer gelernten Verbindung zwischen initialen auditiven Repräsentationen und deren sensorischen Konsequenzen. Eine beim Sprechen geformte artikulatorische Geste, die als typische Einstellung des Vokaltrakts quasi als Schablone der Lautbildung verantwortlich ist, bilde die Grundlage hierfür. Wird dieser Prozess nämlich umgekehrt, so wird das Hörmuster mit dem gespeicherten artikulatorischen Muster verglichen und es wird innerlich ein Lautmuster erzeugt (Kranich 2016: 102). Die Dekodierung des kodierten Sprechsignals funktioniert nach der Motor-Theorie also so, dass die Hörer*innen das empfangene Signal aufgrund ihrer artikulatorischen Fähigkeit bewerten und durch den Vergleich mit den eigenen artikulatorischen Mustern sprachlich kategorisieren. Erst danach wird das Sprachsignal zur weiteren Bearbeitung in die Spracherkennungsregionen des Gehirns geleitet. Streng genommen wäre die Artikulationsfähigkeit also eine notwendige Voraussetzung für die Sprachperzeption (Pétursson/Neppert 2002: 193f.). Außerdem wurde eine parallele Kodierung von Konsonanten und

Vokalen, die artikulatorische Geste, angenommen. Die Mechanismen des sprechmotorischen Apparates würden dabei transformiert und als sequenzielle phonologische Information behandelt werden. Als problematisch zu betrachten ist hierbei, dass das sprachliche Wissen kaum berücksichtigt wird (Kranich 2016: 103).

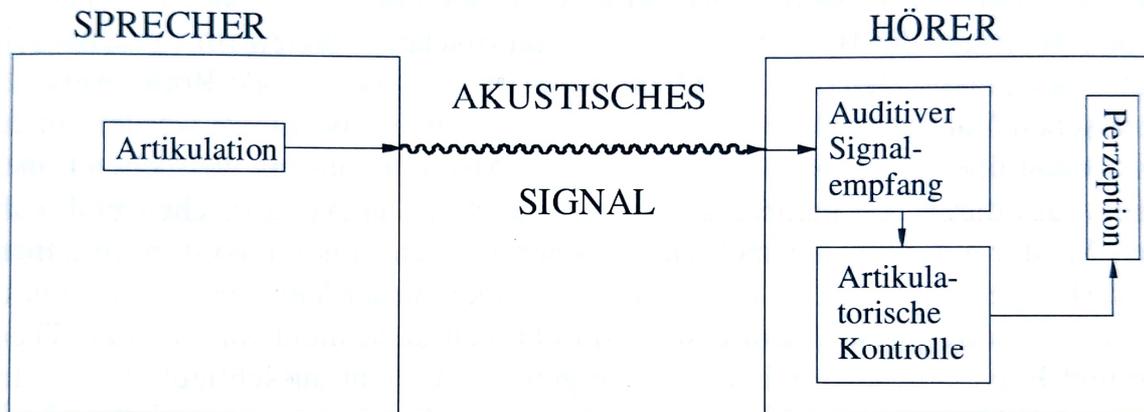


Abb. 3: Eine schematische Darstellung der Stufen in der Sprechschallperzeption nach der Motor-Theorie. Das vom Sprecher gesendete akustische Signal wird vom Hörer empfangen, als Hörmuster erkannt und mit gespeicherten artikulatorischen Mustern des Hörers verglichen, bevor es im Gehirn weiter verarbeitet wird (Pétursson/Neppert 2002: 195). Das Modell beruht auf dem von Shannon/Weaver (1949).

Diese Auffassung wurde anhand neuerer Forschungsergebnisse vom Autor revidiert, wobei die akustische Information jedoch auch in der neueren Version eine zentrale Rolle spielt. Liberman und Mattingly (1985) vertraten nun die Annahme, dass es ein auf die Erkennung der vom Sprecher bzw. der Sprecherin beabsichtigten artikulatorischen Bewegungen spezialisiertes Modul im Gehirn gebe (Pétursson/Neppert 2002: 194). Es wird also eine direkte Verbindung zwischen Sprachproduktion und –perzeption hergestellt, indem der Lautwahrnehmungsprozess einer phonologischen Kategorie zugeschrieben werden kann, da der Sprecher bzw. die Sprecherin über ein Wissen um akustische Korrelate der Lautstruktur verfügt (Kranich 2016: 103). Das sog. Phonetikmodul ermöglicht also die Wahrnehmung der phonetischen Struktur, ohne dass zuvor gesammelte auditive Erfahrungen berücksichtigt werden müssen.

Diese Theorie erklärt die bekannten Phänomene, dass der Mensch beim Hören von Sprache häufig entsprechend den gehörten Sprechabläufen mit atmet und auch der Ablauf der Artikulationsvorgänge mehr oder weniger intensiv mitartikuliert wird. Auch kann eine Fremdsprache besser wahrgenommen und kategorisiert werden, wenn sie artikuliert werden kann. Vermutlich

sind die artikulatorischen Referenzmuster erst dann vorhanden, wenn die fremden Laute erzeugt werden können oder aber das gleiche Gehirn steuert die eigene Sprechlautproduktion und dessen Rezeption bzw. Perzeption über die innerkörperlichen kinästhetischen/propriozeptiven Rückmeldekreise und zusätzlich auch über den äußeren akustisch-auditiven Rückmeldekreis. Produktion und Perzeption des mit dem Sprechapparat erzeugten Schalls und damit auch der Sprechlaute wären also wechselseitig aufeinander bezogen (Pétursson/Neppert 2002: 194f.). Die Motor-Theorie legt also großen Wert auf die Bedeutung relativer akustischer Indizien im Sprechschallkontinuum. Der lautliche Kontext spielt eine zentrale Rolle, weil die akustischen Varianten vor allem durch den Kontext bedingt sind. Die Auffassung, dass sich die gesprochene Sprache von allen anderen akustischen Signalen unterscheidet, wird jedoch vor allem von Psycholog*innen stark angegriffen (vgl. Studdert-Kennedy [u. a.] 1970) und steht in Kontrast zu Theorien zur vorsprachlichen Schallerzeugungs- und Schallimitationsfähigkeit.

Gegen die Motor-Theorie liegen verschiedene Einwände vor. So zeigt sich z. B. auch bei Menschen mit motorischen Aphasien oder neurogenen Stammererscheinungen eine beinahe normale Fähigkeit der Sprachperzeption (Pétursson/Neppert 2002: 196f.). Auch Bauchredner*innen und Papageien können verstanden werden, obgleich sie bei der Spracherzeugung nicht die erforderlichen artikulatorischen Gesten ausführen. Ein Verständnis ist auch von Menschen möglich, die während des Sprechens essen oder Anomalien des Artikulationsapparates aufweisen (Dietrich 2007: 221). Diese Tatsachen müssten in den neueren Varianten der Motor-Theorie berücksichtigt werden.

Die Faszination der Motor-Theorie liegt u. a. sicherlich auch darin begründet, dass sie die schwierig zu erfassenden Prozesse der Sprachperzeption in einen Bereich verlagert, der empirisch leichter überprüfbar ist. Außerdem lassen sich durch dieses Modell zahlreiche Befunde erklären, die in anderen Modellen und Theorien nicht integrierbar sind (Kranich 2016: 104). Beispielsweise steht die Motor-Theorie nicht zu Forschungen im Widerspruch, die die Silbe als Grundeinheit der Sprachperzeption postulieren.

Trotz der vielfach vorgebrachten Kritik an der Motor-Theorie ist es zweifelsohne so, dass eine gute Artikulationsfähigkeit in einer bestimmten Sprache auch deren Perzeption erleichtert. Als Minimalkonsens kann weitestgehend anerkannt werden, dass durch den natürlichen Sprecherwerb die artikulatorischen und auditiven Muster auf irgendeiner Ebene im Zentralnervensystem miteinander verbunden sind. Art, Grad und Auswirkung dieser Verkopplung sowie weitere relevante Faktoren im Perzeptionsprozess müssen noch erforscht werden (Pétursson/Neppert 2002: 197).

2.3.3.2 Auditive Sprachperzeptionstheorien

Die auditiven oder akustischen Sprachperzeptionstheorien behaupten, dass die Sprachperzeption nur durch den akustischen Stimulus ausgelöst wird. Diesen passiven Theorien ist gemeinsam, dass sie die Hörer*innen, in Gegenüberstellung zur Motor-Theorie, als passive Empfänger*innen des Sprachsignals betrachten. Sie verarbeiten das Signal, gestalten es jedoch selber bei der Erkennung der auditiven Muster und der sprachlichen Formen nicht mit. Bei diesen Theorien spielt die Artikulation bei der Sprachperzeption also keine Rolle (Wendt 2007: 34). Auslöser für den Perzeptionsprozess sei der akustische Signalempfang.

Zahlreiche auditive Sprachperzeptionstheorien nehmen an, dass der Hörer bzw. die Hörerin aus dem Signal Merkmale extrahiert. Um diese zu identifizieren, würden die Merkmale oder Merkmalskombinationen mit vorhandenen gespeicherten Referenzmustern verglichen, indem die Merkmale zu Phonemen, die Phoneme zu Morphemen und die Wörter schließlich zu Sätzen kombiniert werden. So würden am Ende dieses Prozesses die Informationen, also die semantischen Inhalte aus dem Trägersignal herausfiltriert.

Solche Merkmalsextraktionstheorien gehen also implizit davon aus, dass der Mensch über eine natürliche, angeborene Sprachkapazität (LAD = Language Acquisition Device) verfügt, welcher im Gehirn bestimmte Zellen entsprechen, die auf gewisse Merkmale spezialisiert sind. Die im Signal gespeicherten Informationen lägen beim Hörer bzw. der Hörerin vor, sobald die Merkmale durch diese Zellen aus dem akustischen Signal extrahiert worden sind.

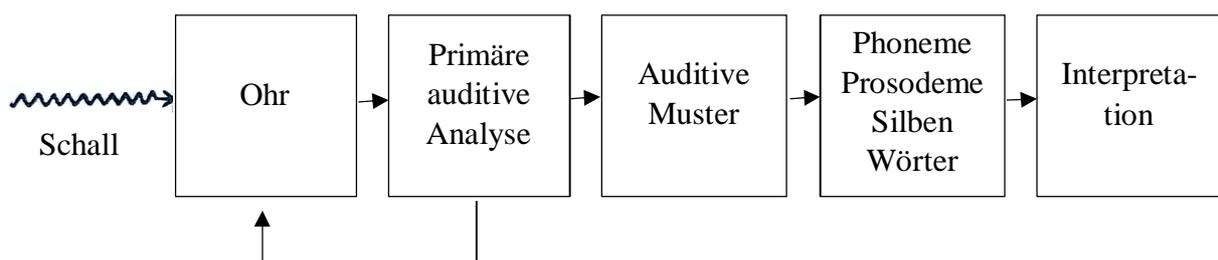


Abb. 4: Eine schematische Darstellung des Signalempfangs und der Interpretation nach der Auffassung vieler auditiver Theorien der Sprachperzeption (Pétursson/Neppert 2002: 198)

Vertreter*innen einer anderen Modellvorstellung gehen davon aus, dass das vom akustischen Signal erzeugte auditive Muster über das Kurzzeitgedächtnis zu einem besonderen Bearbeitungsmechanismus geleitet wird, in welchem ein komplexes Sortierungs- und Verarbeitungsverfahren stattfindet, bis die Information aus dem Signal herausgefiltert werden kann. Es wird

davon ausgegangen, dass im Langzeitgedächtnis die Sprache als System gespeichert ist, während im Kurzzeitgedächtnis nur die ankommenden auditiven Muster aufgenommen werden. In dieser Theorie findet sich kein Hinweis auf die Rolle der Artikulation bei der Sprachperzeption, denn diese wird als Sinneswahrnehmung mit nachgeschalteter Erkennungsprozedur geschaltet. Modelle dieser Art scheinen spekulativ, solange die Beziehung zwischen dem Schallereignis und dem Phonem nicht eindeutig geklärt ist (Pétursson/Neppert 2002: 199f.).

Bei diesen auditiven Theorien werden die neurophysiologisch und psychologisch beschreibbaren Aktivitäten außer Acht gelassen, die jedoch auch beim Prozess der Wahrnehmung ausgelöst werden. Insgesamt kann festgehalten werden, dass noch relativ wenig gesicherte Daten über den menschlichen Sprachperzeptionsprozess vorliegen (Wendt 2007: 35f.).

2.4 Zur Perzeption prosodischer Merkmale

2.4.1 Wahrnehmung von Rhythmus

Der Rhythmus in der gesprochenen Sprache ist ein recht intensiv erforschtes Phänomen, finden sich rhythmische Phänomene doch in vielen Lebensbereichen, z. B. der Biologie und Musik. Versucht man, einen Konsens aus einer größeren Zahl an Modellvorstellungen zur Rhythmusproduktion und -perzeption zu ziehen, so kann Rhythmus als ein eigenständiges, zeitlich-ordnendes Gestaltungsprinzip verstanden werden, bei dem Gleichförmigkeit und Abwechslung eine Rolle spielen (Kranich 2016: 175). Dabei sind die Phänomene Spannung und Entspannung für die rhythmische Konfiguration entscheidend, wobei die Spannungsmomente als Hervorhebungen bzw. Akzente erlebt werden (Fischinger 2009: 22ff.). Somit handelt es sich beim Rhythmus primär um ein psychologisches Konstrukt, welches sich auf mehreren Ebenen manifestiert und einen unterschiedlichen Bezug zum Bewusstsein aufweist (Fischinger 2009: 41f.).

Eine besondere Form rhythmisierter Sprache ist die rhythmusgebundene Dichtung. In Vergangenheit wurden teilweise die Ordnungsprinzipien der rhythmisch gebundenen Dichtung auch in der Alltagssprache gesucht. Weiters bemühte man sich um die Unterscheidung zwischen silben- und akzentzählenden Sprachen (Isochronie-Hypothese, vgl. Abercrombie 1967: 35ff.). Den gesprochenen Sprachen wohne demnach ein isochroner Rhythmus inne, der auf der Isochronie von Takten oder Silben beruhe. So müsste das Deutsche als akzentzählende Sprache einen zeitlich fast gleichen Abstand akzentuierter Silben aufweisen. Die Gültigkeit dieser Hypothese konnte in ihrer reinen Ausprägung nicht verifiziert werden, weil die Probanden eher

geneigt waren, mehr rhythmische Struktur wahrzunehmen, als tatsächlich im Signal vorhanden ist (Kranich 2016: 180). Das bedeutet, dass die Fortsetzungserwartung des Rhythmus die entscheidende Komponente in der rhythmischen Wahrnehmung ist (Stock/Velickova 2002: 16). Außerdem konnten Stock und Velickova (2002: 247) nachweisen, dass die Grobstruktur der rhythmischen Verläufe über den Silbenumfang, die Akzentzahl und die Position der Akzentsilben abgebildet werden können. Darüber hinaus zielt der rhythmische Kontext nach Noel perzeptionsseitig nicht nur auf die Position primärer Akzente, sondern sei auch stark von den Nebenakzenten beeinflusst (Noel 2003: 141f.).

In der Spontansprache wird die Rhythmizität des Sprechens kaum bewusst, weil aufgrund der sprachspezifisch geprägten Hörmuster nur Abweichungen vom Erwarteten auffällig erscheinen. Außerdem kommen stets auch Merkmale wie Melodieführung, Klangqualität, temporale Gestaltung und andere prosodische Parameter bei der Eindrucksbildung hinzu. Temporale und rhythmische Aspekte sind also Bestandteil des Gesamterlebens und werden nur bei Fokussierung auf die entsprechenden Parameter bewusst perzipiert (Kranich 2016: 181f.). Auch der Einfluss der Erstsprache auf die Rhythmuswahrnehmung ist experimentell bestätigt worden (Iversen et al. 2008: 2263ff.).

2.4.2 Wahrnehmung von Akzenten

Was die Akzentuierung unter rhythmischem Aspekt betrifft, so sind vor allem Fragestellungen im Kontext der Isochronie-Hypothese relevant, wobei in derartigen Untersuchungen hauptsächlich signal- und sprecherseitige Akzentuierungsphänomene im Zentrum stehen. Zur Perzeption von Akzenten, also hinsichtlich der Frage, wie es zu einer Akzentwahrnehmung kommt, liegen recht wenige Forschungen vor. Das liegt möglicherweise auch daran, dass Akzente im Normalfall als eine Kombination verschiedener prosodischer Eigenschaften auftreten, sodass kein einheitlicher prototypischer Akzenttypus definiert werden kann. Es kann also davon ausgegangen werden, dass Akzente ebenso wie Stimmfärbung bzw. Klangqualität konstruktive Phänomene der Perzipient*innen sind. Dabei spielen bei der Akzentwahrnehmung die Erwartungshaltung der Hörer*innen sowie das Wissen über melodische Phrasierung und Rhythmus eine relevante Rolle (Kranich 2016: 190ff.).

Fry führte für das Englische Untersuchungen zur Akzentperzeption durch und stellte fest, dass hierfür vor allem Dauer und Intensität entscheidend sind (Fry 1958: 126ff.). O'Shaughnessy (1996: 164) hingegen sieht für das Englische die Grundfrequenzveränderung als wichtigstes Merkmal an, gefolgt von Dauer und Intensität. Diese Reihenfolge wird auch für das Deutsche

angenommen (vgl. z. B. Pétursson/Neppert 2002: 145). Auch Veränderungen der Klangqualität können Akzente bewirken, v. a. zur Darstellung von Expressivität. Des Weiteren ist meist eine Zunahme der Artikulationsspannung festzustellen. Für das Deutsche ergeben sich also folgende Mittel der Akzentrealisation: Die Akzentsilben werden langsamer, aber präziser als die umgebenden Silben artikuliert, sind gedehnt und zeichnen sich auch durch Tonhöhenkontraste aus. In der Perzeption ist ein Wechsel von akzentuierten und akzentlosen Silben in unregelmäßiger Folge festzustellen, wobei die akzentlosen Silben meist schneller und leiser gesprochen werden (Hirschfeld/Stock 2013a: 41).

2.4.3 Wahrnehmung von Tonhöhe/Melodieführung

Bei der Analyse der Sprechweise einer Person können unter melodischem Gesichtspunkt zunächst die dominierende Sprechstimmlage und die Modulationsbreite wahrgenommen werden. Es bestehen Erwartungshaltungen und Hörmuster bezüglich dieser Parameter, wobei nicht die absolute, sondern immer die relative Tonhöhe bzw. Tonhöhenbewegung ausschlaggebend ist. Die „topline“ und „baseline“ sind abgeleitete Größen, die die oberen bzw. unteren Extrempunkte der Melodiebewegung verbinden. Die Bedeutung der relativen Wahrnehmung wird von diversen Forschern herausgestellt (Kranich 2016: 183). So konnte etwa Reinke in einer kontrastiven Untersuchung zwischen russischen und deutschen Hörer*innen nachweisen, dass die Rezeption der Tonhöhe bzw. Melodieführung auch sprachabhängig ist (Reinke 2010: 109). Dies führt zur Annahme, dass Tonsprachensprecher (z. B. chinesische Muttersprachler) in der Perzeption von Tonhöhendifferenzen auch fremder Sprachen genauere Aussagen treffen können, da sie in einer besonderen Weise für sprachliche Tonhöhenunterschiede sensibel sind. Indirekt nachweisen konnten das Pfordresher/Brown (2009: 1385ff.), die hinsichtlich der Wahrnehmung für musikalische Tonhöhenunterschiede eine bessere Diskriminierungsfähigkeit bei Tonsprachensprecher*innen feststellten. Prinzipiell können Ergebnisse von Perzeptionsuntersuchungen zum Melodieverlauf jeweils nur auf die untersuchte Sprache bezogen und nicht verallgemeinert werden (Kranich 183f.). So wie Tonsprachensprecher*innen wohl besonders sensibel für Tonhöhenunterschiede sind, so wird hier vermutet, dass auch Sprachexpert*innen besonders sensibilisiert sind für die einzelnen prosodischen Parameter und deren Diskriminierung. Aber selbst bei phonetisch ungeschulten Personen konnte Künzel (1989: 85ff.) die Bedeutung der Sprechstimmlage und der Modulationsbreite im Kontext der Sprechererkennung als relevante individuelle Indizes nachweisen. Bei der Analyse der Tonhöhenkontur finden sich allerdings auch Bewegungen, die auch bei bewusster Fokussierung durch geübte Hörer*innen

kaum perzipiert werden können. Diese Bewegungen, z. B. die vokalspezifische Grundfrequenz, werden als Mikroprosodie oder C-Prosodie bezeichnet (Kranich 2016: 184). In einem Experiment von Isačenko und Schädlich (1966: 11f.) konnte gezeigt werden, dass nicht die absoluten, sondern die relativen Tonhöhenunterschiede für die Bewertung ausschlaggebend sind und viele tonale Elemente aus dem Signal eliminiert werden und die Proband*innen trotzdem ein eindeutiges Wahrnehmungsergebnis erzielen können. Aus einer Untersuchung Niebuhrs (2007: 319f.) zur Wahrnehmungskonstruktion der Sprechmelodie ging hervor, dass die Faktoren Kontrastbildung, Verstärkung, Top-down-Wissen und Gestaltgesetze das Perzeptionsresultat bedingen. Hier werden also zahlreiche wahrnehmungspsychologische Aspekte berücksichtigt. Aus neurowissenschaftlicher Sicht gilt es festzuhalten, dass Veränderungen im Signalverlauf kortikal registriert werden, wobei gerade die Verarbeitung mikroprosodischer Merkmale Aktivitäten in der linken Hemisphäre bewirken. Dabei ist interessant, dass diese Merkmale normalerweise nicht bewusst werden, obgleich sie vom auditorischen System nachweislich erkannt und verarbeitet werden (Falk 2009: 39). Bei der Verarbeitung der Sprechmelodie hingegen spielt die rechte Hemisphäre eine zentrale Rolle, insbesondere die rechte superiore Temporalregion und da vor allem das Planum temporale (Meyer et al. 2002: 73ff.).

2.4.4 Wahrnehmung von Lautstärke

Forschungen zur Lautstärke, unabhängig ob aus Signal-, Sprecher*innen- oder Perzipient*innenperspektive sind in der Prosodieforschung sehr rar. Dies liegt in erster Linie an den schwierigen Untersuchungsbedingungen. So müssten Tonaufnahmen unter konstant gleichen Bedingungen in Bezug auf Mikrofonabstand, Kopfhörer und audiometrischem Abbild der Hörer*innen (Ausschluss von Anomalien) in einem akustisch abgeschirmten Raum durchgeführt werden, wobei Aufnahmen in derartigen Studiobedingungen kaum mehr Alltagsrelevanz zukommt und der Aufwand für die Experimente extrem groß ist. Selbst bei optimalen Aufnahmebedingungen ergibt sich aber eine weitere Problematik – die maximale Wahrnehmungsdistanz ist für Hörer*innen ohne Kopfhörer höchst unterschiedlich (Kranich 2016: 193f.). Vokale etwa werden über größere Distanzen als Konsonanten perzipiert und auch innerhalb der diversen Lautkategorien bestehen erhebliche Differenzen (v. Essen 1979: 132). Dieses Phänomen ist in der Alltagskommunikation irrelevant, da die Hörwahrnehmung derartige Unterschiede ausgleicht und Erwartungshaltungen hinsichtlich der Äußerungen bestehen (Kranich 2016: 194). Außerdem zeigt sich ein Zusammenhang zwischen der Perzeption der Lautstärke und der Tonhöhe, wobei in Wahrnehmungsexperimenten intensitätsstarke Silben von Hörer*innen zugleich als

höher gegenüber intensitätsschwachen empfunden werden (vgl. Stock 1996: 218). So ist es nicht verwunderlich, dass einerseits eine gewisse Skepsis in Bezug auf experimentelle Untersuchungen der Lautstärke existiert und andererseits die verstärkte Forderung, den Einfluss der Lautstärke für eine Modellierung der prosodischen Merkmale vermehrt ins Zentrum zu rücken (vgl. Niebuhr 2007: 371), auf keine konkreten Lösungsvorschläge trifft.

Das entscheidende Kriterium in Bezug auf die Lautheit gesprochener Äußerungen ist die kommunikative Angemessenheit, welche situations- und kulturabhängig ist. So werden Veränderungen der Lautstärke erst durch deren Abweichungen vom Hörmuster bewusst. Sehr leise oder sehr laute Sprechweisen sind schlecht verständlich. Bei zu geringer Intensität wird die Wahrnehmung konsonantischer Segmente beeinträchtigt, ein zu hoher Schalldruckpegel hingegen ist beim Hören physiologisch und sozial unangenehm und erschwert die Sinnerschließung des Gesagten (Kranich 2016: 194f.; Lazarus et al. 2007: 104). Auch Stock erörtert, dass sich der Hörer bzw. die Hörerin hinsichtlich der Lautstärke auf eine mittlere Lautheit einstellt und von diesem Mittelwert Abweichungen, z. B. Akzentabstufungen, in Relation setzt. Solche Werte sind jedoch weder auditiv noch apparativ eindeutig bestimmbar (Stock 1980: 141). Auch aus neurowissenschaftlicher Sicht gibt es keine derartigen Untersuchungen zur Lautstärke, wobei aus den bisher referierten Darstellungen aber grundsätzlich angenommen werden kann, dass verschiedene Areale und Strukturen im Gehirn bei der Perzeption der Lautstärke beteiligt sein dürften (Kranich 2016: 195).

2.4.5 Wahrnehmung von Sprechgeschwindigkeit und Pausen

Als direkte temporale Kriterien sind die Artikulationsrate und die Pausierung zu nennen, die in ihrer Relativität wahrgenommen werden. Wenn jemand sehr rasch und mit einer hohen bzw. niedrigen Artikulationsrate spricht, kommt es zum Vergleich mit dem eigenen inhärenten Maß, wobei bei bekannten Gesprächspartner*innen auch das individuelle Hörmuster aktiviert wird (Kranich 2016: 181). Dabei ist die Dauer der Sprachlaute von zahlreichen Faktoren abhängig, nämlich zum einen von linguistischen Faktoren wie Phrasierung und Fokussierung und so der Position der Laute im Satz und zum anderen von nicht-linguistischen Faktoren wie z. B. der situationsbedingten Sprechgeschwindigkeit und dem emotionalen Zustand des Sprechers bzw. der Sprecherin (Enterlein/Bartels/Sendlmeier 2005: 12). Es konnte herausgefunden werden, dass für die Emotionen Trauer und Ekel eine langsamere Sprechgeschwindigkeit vorherrscht, für Ärger und Freude hingegen eine rasche Sprechgeschwindigkeit (Kienast 2002). Außerdem

sind Sprechgeschwindigkeit und Pausierung nicht nur individuelle Größen, sondern auch kulturell geprägt (Uhmann 1997a: 78ff.). Daher erschien auch für die vorliegende Befragung die Betrachtung von Sprechgeschwindigkeit und Pausierung sinnvoll.

2.4.6 Wahrnehmung von Stimmfärbung

Die Klangqualität bzw. Stimmfärbung (die Begriffe sollen hier synonym verwendet werden) als prosodisches Merkmal ist bisher kaum in den Fokus empirischer Forschungen gerückt, vor allem unter perzeptivem Aspekt liegen kaum Untersuchungen vor. Dies liegt zum einen daran, dass es erhebliche Divergenzen in der Begriffsdefinition gibt, und zum anderen am mangelnden Konsens darüber, welche Merkmale für die Generierung der Stimmfärbung angenommen werden können. Interessante Arbeiten zum Thema aus dem Bereich der Musikwissenschaft können nur bedingt auf die gesprochene Sprache übertragen werden.

Zum einen zeigen sich in der Klangqualität die als Formanten bezeichneten Merkmale, die mit den physikalischen Lauteigenschaften zusammenhängen und die Sprachklangmerkmale bedingen (Kranich 2016: 187). Die Formanten sind kommunikativ relevante Anteile, die bei Menschen, aber auch in der Tierwelt eine große Bedeutung haben und vom Cortischen Organ über die gesamte Hörbahn rezipiert und im Bereich des Hirnstammes analysiert werden (vgl. Griffiths 2003: 47). Zum anderen zeigen sich hinsichtlich der Stimmfärbung die auf den phonatorischen Aspekten basierenden Merkmale der Stimmklangeigenschaften, die als indexikalisches Merkmal auf einen Sprecher bzw. eine Sprecherin verweisen und so dessen bzw. deren individuelle Sprechklangeigenschaften bedingen. Die drei genannten Manifestationen Sprache, Stimme und Sprechen mischen sich also in der Klangqualität, was zu Problemen in empirischen Studien führen kann. Um das Timbre bzw. die Stimmfärbung eines Sprechers zu analysieren, eignen sich Beschreibungen wie z. B. „dunkel“, „warm“, „klangarm“ etc., ohne dass sich hierfür eine eindeutige physikalisch messbare Entsprechung finden lässt (Kranich 2016: 188). Generell kann hier die enorme Bedeutung des Stimmklangs in der Kommunikation nur nochmals betont werden – über dieses Merkmal zieht der Hörer bzw. die Hörerin nämlich Rückschlüsse auf die physischen (z. B. Geschlecht, Alter, körperliche Konstitution, Aussehen inkl. Größe, Gewicht, Attraktivität, Gesundheitszustand, ethnische Zugehörigkeit, sexuelle Orientierung, Raucher*in bzw. Nicht-Raucher*in usw.), psychologischen (z. B. Stimmung bzw. emotionaler Zustand, Intelligenz, Persönlichkeit, Stress, Vertrauenswürdigkeit, Kompetenz usw.) und sozi-

alen (z. B. Bildung, berufliche Tätigkeit, regionale Herkunft, Rolle in der Kommunikationssituation, sozialer Status) Charakteristika des Sprechers bzw. der Sprecherin (Kreiman/Vanlancker-Sidtis/Gerratt 2005: 338f.).

2.4.7 Allgemeine Probleme bei Untersuchungen der Prosodieperzeption

Zunächst gilt festzuhalten, dass allgemeine Fragestellungen nach der Art und Weise der Prosodieperzeption für weiterführende Untersuchungen eher nicht produktiv sind – ein gesprochene Sprache wahrnehmender Mensch wird in verschiedenen Kontexten die prosodischen Parameter verschieden perzipieren und verarbeiten. Bis zu einer bestimmten Stufe der Hörbahn kann eine ähnliche Verarbeitung angenommen werden, doch Aufmerksamkeitsprozesse beeinflussen schon die Aktivität der äußeren Haarzellen im Cortischen Organ. Auch die präkortikalen Stationen der Hörbahn zeichnen sich durch Selektions- und Filterprozesse aus. Daher kann gerade auf höheren Verarbeitungsstufen eine durch die Hörintention geprägte und damit nicht unter identischen Bedingungen reproduzierbare Verarbeitung prosodischer Erscheinungen angenommen werden. Das erschwert also deren empirische Untersuchung erheblich, zeigt aber, dass nur umgrenzte Fragestellungen mit überschaubarem Fokus zu relevanten Ergebnissen führen. Diese wiederum können kaum den gesamten Prozess der Verarbeitung prosodischer Merkmale umfassen (Kranich 2016: 234). In dieser Studie kann anhand konkreter Hörbeispiele und Proband*innengruppen der Frage nachgegangen werden, welche Rolle die prosodischen Parameter bei der Perzeption emotionaler Sprechweisen (Ärger und Freude) spielen und inwieweit kultursprachbedingte, geschlechtsbezogene und berufsbedingte Faktoren die Wahrnehmung bedingen.

An dieser Stelle muss auf die Problematik hingewiesen werden, dass Untersuchungen zur Perzeption von Prosodie, aber auch von Sprache allgemein immer nur indirekt, also über den Umweg einer produktiven Indexfunktion erfassbar sind, da sich perzeptive Verarbeitungsleistungen und Sprachverstehen einer direkten Beobachtung entziehen. Dieses messmethodische Problem perceptionsseitiger Untersuchungen ist bis jetzt nicht lösbar.

Ein weiteres Problem bei der Untersuchung der Prosodieperzeption ist die Verbindung segmentaler und prosodischer Merkmale. Sprachliche Zeichen sind beim Sprechen für die Übertragung stets an den Sprechakt gebunden, d. h. eine Trennung segmentaler und prosodischer Erscheinungen (z. B. in Laborexperimenten) ist nur um den Preis einer unnatürlich wirkenden Sprechweise möglich. So ergibt sich das Dilemma, natürlich gesprochene Sprache untersuchen zu

wollen, die im kommunikativen Geschehen jedoch durch verschiedene Komponenten modifiziert wird. Während im Labor die Kontrollen dieser Variablen gut ermöglicht werden kann, sind die daraus resultierenden Ergebnisse jedoch kaum auf die alltagstypische Kommunikation übertragbar. Ein möglicher Mittelweg, der auch hier besprochen werden soll, ist eine Untersuchungskonzeption, die zum einen natürlich gesprochene Sprache berücksichtigt und zum anderen den Einfluss des segmentalen Materials minimiert. So wird durch eine geeignete Aufgabenstellung der Fokus auf die Verarbeitung inhaltlicher Aspekte reduziert und auf prosodische Merkmale konzentriert (Kranich 2016: 236f.).

Die nächste Problematik wurde bereits auch an anderen Stellen (Kapitel 2.2) diskutiert – prosodische Merkmale bilden einen Merkmalskomplex, der unter dem Begriff Prosodie subsumiert wird. Grundsätzlich ist davon auszugehen, dass prosodische Parameter in alltäglichen Kommunikationssituationen unbewusst als Komplexphänomen wahrnehmen und verarbeiten, wobei die Gewichtung der relevanten prosodischen Merkmale unbewusst erfolgt. Bewusst ist nur das Interpretationsergebnis, also z. B. die Wahrnehmung einer Stimme als „angenehm“. Da also auch in gewissen Situationen der Alltagskommunikation wahrnehmungsseitig von einer Separierung prosodischer Erscheinungen ausgegangen werden kann, kann dieses Vorgehen für empirische Untersuchungen als legitim betrachtet werden. Natürlich ist die Generalisierung experimenteller Ergebnisse auf alltagssprachliche Belange aber immer ein Analogieschluss, der kritisch zu hinterfragen ist, schließlich ist jede Untersuchung von Formmerkmalen gesprochener Sprache nicht mehr mit einer dem alltäglichen Kommunikationsverlauf entsprechenden Perzeptionsleistung zu vergleichen. In der Studie können die verschiedenen prosodischen Merkmale durch konkrete Aufgabenstellungen mehr oder weniger bewusst gemacht werden, wobei Merkmale zu Tonhöhereigenschaften wohl sicherer erkannt werden als jene zur Stimmfärbung. Hier ist es also sinnvoll, eine überschaubare Zahl von Wahlalternativen anzubieten, die aber dennoch über eine ausreichende Trennschärfe verfügen (Kranich 2016: 238f.).

Zuletzt soll hier nochmals das Problem thematisiert werden, dass sich die auditorische Wahrnehmung beim natürlichen Verarbeitungsvorgang gesprochener Sprache im Vergleich zu experimentellen Designs stets unterscheidet. Während in der Alltagskommunikation der Fokus auf dem Inhalt des Gesagten liegt und das rein analytische Hören (Wahrnehmung der Formseite der Prosodie) kaum eine Rolle spielt, wird in Perzeptionsexperimenten auch nach dem Formaspekt gefragt. Dieser aber ist in Alltagssituationen nicht relevant und nur dann bewusst, wenn deutliche Abweichungen zur Hörerwartung vorliegen. Daher stellt sich die Frage, inwieweit Ergebnisse solcher empirischer Untersuchungen generell auf die Alltagssprache und die Hörwahrnehmung zu übertragen sind. Durch bisherige Forschungsergebnisse (vgl. Lindner 1977:

164) lässt sich jedoch feststellen, dass gewisse bewusste Auseinandersetzungen mit prosodischen Merkmalen auch im natürlichen Sprachverarbeitungsprozess beobachtbar sind. Dadurch scheint eine Übertragung der Ergebnisse aus Untersuchungen zur Prosodieperzeption grundsätzlich möglich, wenngleich die Besonderheit der Untersuchungssituation berücksichtigt werden muss (Kranich 2016: 239f.)

2.5 Einflussfaktoren auf die Prosodieperzeption

In der vorliegenden Studie sollen die Kategorien Beruf, Geschlecht, Kultur und Sprache als Einflussfaktoren in der Perzeption emotionaler Sprechweisen durch prosodische Merkmale differenziert und untersucht werden. Diese Komponenten sollen im folgenden Abschnitt daher kurz einzeln betrachtet und als Einflussfaktoren in der Emotionsperzeption näher dargestellt werden. Es gilt zu beachten, dass die einzelnen Komponenten natürlich als Merkmalskomplex in die Emotionswahrnehmung einfließen, ebenso wie der paraverbale Ausdruck stets mit dem verbalen und nonverbalen Ausdruck gekoppelt ist. Die einzelnen prosodischen Parameter sind also miteinander verknüpft, sie können sich gegenseitig ergänzen, partiell vertreten und werden auch von Kultur zu Kultur verschieden realisiert bzw. perzipiert. Der Merkmalskomplex ist also multifaktoriell bedingt und kann nur als Realisationsmöglichkeit mit bestimmten Erscheinungstendenzen beschrieben werden (Hirschfeld/Stock 2016: 17).

2.5.1 Einflussfaktor Kultur

Die kulturelle und sprachliche Heterogenität und damit die interkulturelle Kommunikation gewinnen aufgrund der immer stärker eingreifenden Globalisierung, der wachsenden Multikulturalität und Transkulturalität (fremde Kulturelemente werden im interkulturellen Kontakt durch Einzelne oder Gruppen aufgenommen, Welsch 1997) und vor allem auch im Rahmen der aktuellen Flüchtlingsbewegung eine immer größere Bedeutung. In zahlreichen Wissenschaftszweigen, z. B. der Linguistik, Ethnologie, Anthropologie, Soziologie, Geschichte, Kulturwissenschaft, Psychologie, Pädagogik und Betriebswirtschaftslehre wird zur Interkulturellen Kommunikation geforscht (vgl. z. B. Lüsebrink 2016; Hofstede 2003; Moosmüller 2007; Heringer 2014; Helfrich 2013). Je nach Zusammenhang wird der Begriff der Interkulturellen Kommunikation also unterschiedlich gebraucht und jeweils spezifisch präzisiert. Eine allgemein akzeptierte und genügend differenzierte Begriffsdefinition liegt daher nicht vor, was auch zu Unsicherheiten führen kann (Gogolin/Krüger-Potratz 2006: 114). Für diese Untersuchung orientiere

ich mich an der sprechwissenschaftlichen Definition des Begriffs nach Hirschfeld und Stock (2016:14). So wird Interkulturelle Kommunikation verstanden als die „interpersonelle Sprechkommunikation zwischen Angehörigen verschiedener Kulturen, wobei die Interagierenden sich gegenseitig als fremd erleben, sich also bewusst sind, dass der „Andere“ anders ist (Hirschfeld/Stock 2016: 14).“ Die Face-to-Face-Kommunikation steht dabei im Vordergrund, aber auch die technisch und institutionell vermittelte, v. a. die massenmediale Sprechkommunikation rückt immer mehr ins Blickfeld. Natürlich ist im Rahmen dieses Projekts nur die eine Seite des Rezipienten bzw. der Rezipientin von Interesse, weil die Arbeit darauf fokussiert, die Wahrnehmung der Hörer*innen in Bezug auf mutter- und fremdsprachliche emotionale Äußerungen zu untersuchen. Das heißt, hier kann nur ein Teilaspekt der Interkulturellen Kommunikation erforscht werden.

2.5.1.1 Kulturbegriff

Auch der Kulturbegriff wird äußerst vielseitig verwendet und definiert. Hier soll an jener Denktradition angeschlossen werden, die Kultur bildhaft als eine Art Text versteht, dessen Komponenten (Lebens-, Verhaltens- und Äußerungsweisen, Mimik und Gestik, Kleidung etc.) erst im Gesamtzusammenhang Bedeutung erlangen. Gruppen von Menschen stellen mit so einem Text in ihrer natürlichen Umgebung und ihren Beziehungen Strukturiertheit und Sinnzusammenhänge her und erschaffen sich so eine Identität, die sie gegenüber Menschen anderer Kulturen abgrenzt. Als Kultur kann somit die in einer Gesellschaft bzw. Gruppe von Menschen tradierte Gesamtheit der kollektiven Denk-, Wahrnehmungs-, Wertungs- und Handlungsmuster bezeichnet werden (Hirschfeld/Stock 2016: 16). Kulturen sind nicht nur durch die geographischen Grenzen bedingt, sondern beziehen sich auf alle Formen vergesellschafteter Subjekte, welche sich als zusammengehörig empfinden. Demnach kann auch von einer Unternehmens-Kultur oder Stadtteil-Kultur gesprochen werden, sofern die Betrachtungsebene und kategoriale Reichweite zuvor klar definiert werden. Kulturen werden fortlaufend neu gebildet, indem Menschen Bisheriges übernehmen bzw. verändern. So entscheidet sich, an welche Regeln, Normen und Werte sich eine Gruppe zu einem bestimmten Zeitpunkt halten will und welche modifiziert werden sollen. Kulturen sind also keine statischen Gebilde, sondern Voraussetzung und Ergebnis sozialer Aushandlungsprozesse. Die eigenständige Zuordnung zu einer bestimmten kulturellen Gruppe wirkt identitätsstiftend und verstärkt das Zusammengehörigkeitsgefühl innerhalb der Gruppe, wobei sich die Identität einer Person aus einer Vielzahl von kulturellen Identitäten auseinandersetzt (z. B. Studierende, Sprecherzieher*in etc.) (Kammhuber 2011: 172). Auch in der vorliegenden Studie lassen sich nach diesem Gesichtspunkt diverse (kulturelle) Gruppen

ausmachen, da nach Geschlecht, Beruf, Alter und geographisch/sprachlicher Kultur gefragt wird.

In dieser Arbeit schließe ich mich außerdem der Denktradition des Sprechwissenschaftlers Stock und der Sprechwissenschaftlerin Hirschfeld an und gehe davon aus, dass auch die Kultur der Sprechkommunikation zentraler Bestandteil der Kultur einer Ethnie ist. Sie tritt abhängig von der Lebenssphäre oder Situation variabel in Erscheinung, ist interdependentiell mit der Gesamtkultur verbunden und schafft wie diese eine soziale Identität. Zu ihren Merkmalen zählen u. a. die Gesprächsstrategien, die diversen Reaktionsformen auf sprachliche Äußerungen unterschiedlicher Art sowie die verbalen, paraverbalen und nonverbalen Formen der verschiedenen Sprechweisen (vgl. Hirschfeld/Stock 2016: 16).

2.5.1.2 Interkulturelle Kommunikation

In Bezug auf das Gelingen interkultureller Kommunikationsprozesse ist es problematisch, dass die der eigenen Kultur zugrunde liegenden Regeln, Werte und Normen nicht bewusstseinspflichtig sind. In der Eigengruppe wird selbstverständlich davon ausgegangen, dass die anderen Gruppenmitglieder das eigene kulturelle System teilen. Erst in „kulturellen Überschneidungssituationen“ (Winter 1992), in welchen Personen aufeinandertreffen, die kulturbedingt verschiedene Interpretationen an eine Situation herantragen und dadurch bei der Verfolgung ihrer Handlungsziele positiv oder negativ voneinander überrascht werden, wird die eigene Kultur bewusst und reflektierbar. In einer interkulturellen Kommunikationssituation bringen die jeweils beteiligten Personen ihr spezifisches, im Laufe der individuellen Sozialisation entwickeltes kulturelles Orientierungssystem ein, welches den weiteren Verlauf des Gesprächs in Form von gegenseitigen kulturspezifischen Rollen- und Verhaltenserwartungen beeinflussen wird. Dabei können sich die handlungsleitenden Merkmale des jeweiligen Orientierungssystems ähneln oder voneinander abweichen und damit Missverständnisse mehr oder weniger wahrscheinlich machen. Auch haben die Gesprächspartner*innen schon im Vorfeld ein mehr oder minder stereotypes Fremdbild voneinander entwickelt und interagieren auf dessen Grundlage. Dies kann hinderlich oder förderlich für das Gespräch sein. Des Weiteren vollzieht sich die interkulturelle Kommunikationssituation in einem situativen Kontext, der durch Zeit, Ort und die in die Gesprächssituation eingebundenen Menschen bestimmt ist (Kammhuber 2011: 172ff.). Auch in der hier durchgeführten Studie sind die Resultate der Perzeptionsanalysen von diesem Kontext sowie den kulturellen Prägungen und Erwartungen geprägt. Das betrifft ebenso den Umgang und die Perzeption von Emotionen, welche stark von der jeweiligen Sprache bzw. Kultur

beeinflusst sind und in interkulturellen Kommunikationssituationen auch zu Missverständnissen führen können. Gelingt es, mit negativen Emotionen im interkulturellen Kontext gut umzugehen, kann es zu einer Erweiterung des eigenen Handlungsspielraums und der interkulturellen kommunikativen Kompetenz kommen. Dabei sind Offenheit, Flexibilität und ein kritischer Zugang zu den eigenen und fremden Wertesystemen hilfreich (Matsumoto/Hee Yoo/LeRoux 2007: 81ff.). Der Respekt vor dem kulturgeprägten Verhalten des Anderen und vor seinem Wertesystem ermöglicht gegenseitige Achtung und Verständigung. Das Kennenlernen der jeweils anderen Kultur und das Akzeptieren der Gleichwertigkeit etwaiger Unterschiede zum eigenen sprachlich-kulturellen Hintergrund führt quasi automatisch auch zu einem angemessenen Verhalten (Allhoff/Allhoff 2014: 243ff.).

Was die Interpretation der Ergebnisse betrifft, so ist also Sorgfalt geboten. Die Validität der Ergebnisse aus Wirkungsuntersuchungen hängt eben nicht nur von den zu bewertenden Ausdrucksformen ab, sondern auch von der Persönlichkeit der Teilnehmer*innen, also von deren Wertesystem, Motivation, internalisierten Denk- und Verhaltensmustern usw. Dabei spielt nicht nur die jeweilige Kultur, sondern auch die Sozialisation in der Gruppe eine relevante Rolle, weshalb in dieser Untersuchung Personen aus unterschiedlichen sozialen Gruppen (z. B. Studierende, Berufstätige usw.) an der Befragung teilnehmen (vgl. Hirschfeld/Stock 2016: 19ff.).

2.5.1.3 IK und Prosodie

Um das Kernproblem der Interkulturellen Kommunikation zu erfassen, sind Wirkungsuntersuchungen sinnvoll. So kann ermittelt werden, wie Angehörige der einen Kultur auf die Äußerungsweisen einer anderen Kultur reagieren. Dabei wird generell insbesondere nach Einschränkungen der phonetischen Verständlichkeit im interkulturellen Kontakt, nach Abweichungen von erwarteten Merkmalen der Rhetorizität, nach phonetischen Interferenzen und nach Besonderheiten auf der nonverbalen und paraverbalen Ebene gefragt. Im Rahmen dieser Studie soll dies in Bezug auf die Wahrnehmung emotional gefärbter Prosodie für die deutsche bzw. italienische Sprach- und Kulturgruppe in Form von Perzeptionsanalysen geschehen. Die hier vorliegende Untersuchung zur Wirkung der paraverbalen Komponenten im Gesprochenen ist natürlich nur ein sehr kleiner Baustein für die Beschreibung des Gesamtphänomens des interkulturellen Sprechens und Verstehens. Die Ergebnisse sollten aber zum Verständnis des untersuchten Phänomens beitragen und auch einen praktischen Nutzen haben, wobei eine Erweiterung des Untersuchungsbereichs auf andere Sprach- bzw. Kulturgruppen sicherlich wünschenswert

ist. So definieren Hirschfeld/Neuber (2010: 53) etwa den abschließenden Arbeitsschritt in Untersuchungen zur Interkulturellen Kommunikation explizit als didaktische Schlussfolgerungen für den Unterricht von Deutsch als Fremdsprache, da für die Vermittlung von interkultureller Kompetenz auch spezifische Kenntnisse und Fähigkeiten notwendig sind. Außerdem führen gerade unzureichende Kenntnisse der Fremdsprache oder Unsicherheiten in der Verwendung von Lexik, Syntax und Stil nicht selten dazu, dass Sprechende stärker als gewöhnlich aufgrund der paraverbalen Ebene beurteilt und als „fremd“ wahrgenommen werden. Die prosodischen Merkmale sind in vielen Situationen der Interkulturellen Kommunikation bei Weitem die wichtigste Grundlage für die Personenwahrnehmung und das Erfassen des Gemeinten und spielen daher gerade auch im interkulturellen und im didaktischen Kontext eine große Rolle. Paraverbale Merkmale wie Prosodie und Stimmklang sind nämlich ebenso wie nonverbale Merkmale kultur- und sprachbestimmt ausgeprägt und werden situativ variiert und kultur- bzw. sprachbestimmt eingesetzt (Hirschfeld/Stock 2010: 17).

2.5.2 Einflussfaktor Sprache

Die Sprachperzeption ist natürlich auch durch die phonologischen Eigenschaften der Erstsprache des Hörers bzw. der Hörerin bedingt. Was die Perzeption einer Fremdsprache betrifft, so ist das Wissen um die neurologischen Vorgänge sowie um die linguistischen und sensorischen Mechanismen dabei noch recht eingeschränkt, da dies erst seit jüngeren Jahren intensiver erforscht wird und sich viele Studien bisher nur mit der Perzeption von Segmenten (Vokalen oder Konsonanten) auseinandergesetzt haben (vgl. Strange 1995). Tatsächlich sind jedoch auch weitere Faktoren zu beachten. Verschiedene Sprachen „klingen“ unterschiedlich – bei der Perzeption wird das Sprachsignal in verschiedene linguistische Einheiten geteilt und aufgrund des phonologischen Vorwissens über Segmente, Silben, Betonungsschemata etc. zugeordnet. Dabei können verschiedene Sprachen in allen diversen Komponenten differieren. Dies gilt natürlich auch und besonders für die prosodischen Merkmale. Bei der Perzeption fremder Sprachen kann es zu „Sprechillusionen“ kommen – so kann es etwa passieren, dass die Hörer*innen keine Unterschiede zwischen verschiedenen Sprachäußerungen wahrnehmen können. Ein populäres Beispiel hierfür wäre das Phänomen, dass Japaner*innen die englischen Wörter „road“ und „load“ nicht unterscheiden können. Des Weiteren können Informationen hinzugefügt werden, die nicht im Sprachsignal vorhanden sind (spanische Muttersprachler*innen hören z. B. ein [e] am Beginn aller englischen Wörter), oder bestimmte Klänge in andere, der eigenen Sprache näheren Klänge umgewandelt werden (Gallés 2005: 546ff.).

Den prosodischen Parametern kommt bei der Analyse von Sprachäußerungen eine besondere Bedeutung zu – auch im fremdsprachigen Kontext scheint die Sprachperzeption nämlich in größeren (prosodischen) Einheiten zu erfolgen und nicht auf Basis einzelner Phoneme. Entgegen früherer Annahmen kann die Silbe nicht als universelle Einheit in der Sprachperzeption betrachtet werden, da in den unterschiedlichen Sprachgruppen relevante Unterschiede in der Wahrnehmung und Bildung der Silben vorliegen. So würden z. B. französisch- oder spanischsprachige Sprecher*innen die Sequenz /lemən/ in die zwei Silben /le/ und /mən/ unterteilen, während hingegen englischsprachige Sprecher die Einheit in /lem/ und /ən/ trennen würden. Wichtiger für die Sprachperzeption scheint der Rhythmus zu sein, welcher relevante Auswirkungen auf die Art und Weise hat, wie eine Sprache perzipiert wird (Gallés 2005: 552ff.). Pike (1945) hat als Erster in einem Modell zwischen „syllable-timed“ und „stress-timed“ Sprachen unterschieden. Sein Ansatz wurde von vielen Autor*innen aufgegriffen und weiterentwickelt (z. B. wurde noch eine dritte Gruppe, „mora-timed“, für Sprachen wie z. B. das Japanische, hinzugefügt), aber auch kritisiert (vgl. Dasher/Bolinger 1982; Ramus/Nespor/Mehler 1999). Besonders relevant beim Perzipieren einer anderen Sprache sind auch die Betonung und Akzentuierung, da diverse Betonungen in vielen Sprachen auch die Bedeutung der Äußerung verändern. In einigen Sprachen, in denen das nicht der Fall ist (z. B. im Französischen, hier werden alle Wörter auf der Endsilbe betont) kann es beim Perzipieren einer Fremdsprache auch zu Schwierigkeiten kommen, relevante Betonungen und Akzentuierungen überhaupt wahrzunehmen und zu erkennen (Gallés 2005: 555f.). Die jeweilige Muttersprache der Hörer*innen wirkt sich, neben anderen Faktoren, also auch auf die Perzeption einer fremdsprachigen Äußerung aus.

Sprachvergleich: Prosodie im Deutschen und Italienischen

In dieser Arbeit geht es um Perzeptionsanalysen deutscher bzw. italienischer Äußerungen durch deutsch- bzw. italienischsprachige Zuhörer*innen. Daher soll hier abschließend exemplarisch auf prosodische Merkmale der beiden Sprachen hingewiesen werden. Eine ausführliche sprachtypologische Beschreibung scheint für die Durchführung und Interpretation der Befragung nicht vonnöten, deshalb wird hier darauf verzichtet. Einen ausführlichen Forschungsüberblick und eine genaue Darstellung zu lautlichen Unterschieden in der Intonation im Deutschen und Italienischen finden sich in Rabanus 2001, eine prägnante linguistische Gegenüberstellung der beiden Sprachen bietet z. B. Ferrer 1999.

Im Deutschen und Italienischen gibt es zahlreiche Dialekte und Regionalsprachen – in dieser Arbeit beziehe ich mich allerdings nur auf die Standardvarietäten der jeweiligen Sprache. Diese weisen trotz ihrer sprachtypologischen Nähe ein deutlich verschiedenes klangliches Erscheinungsbild auf. Während das Deutsche häufig als „eckig“ oder „hart“ charakterisiert wird, gilt Italienisch allgemein als eine sehr melodische und musikalische Sprache. Diese Differenzen sind jedoch nur zu einem geringen Teil intonatorisch bedingt, sondern gehen primär auf Unterschiede in der metrischen Struktur, der Silben-Struktur und der Segment-Struktur zurück. Silben im Deutschen sind tendenziell komplexer als im Italienischen und unterliegen stärker Reduktions- und Assimilationsprozessen (Rabanus 2001: 63f.).

Die italienische Standardaussprache ist durch eine hohe Sprechspannung gekennzeichnet, was sich u. a. in den klaren gespannten Vokalen in unbetonter Stellung und in den Langkonsonanten zeigt. Das Spannungspotenzial ist dabei auf einer hohen Ebene recht ausgeglichen, während das Deutsche zwischen betonter und unbetonter Silbe große Spannungsunterschiede verzeichnet. Da im Italienischen klanglich die Vokale dominieren, gilt es als die Sprache des Gesangs. Es gibt eine Fülle an zwei- und mehrgliedrigen Vokalverbindungen, Konsonantenfolgen sind im Gegensatz zum Deutschen hingegen nur zweiteilig. Die klare und reine Artikulation der italienischen Vokale bleibt in allen Positionen erhalten und trifft auch auf die unbetonten Vokale zu, die immer gespannt sind, während im Deutschen gerade in akzentlosen Silben starke Reduktionserscheinungen auftreten können. Auch ist der vokalische Wortauslaut bei italienischen Wörtern die Regel, weshalb Italiener bei konsonantisch auslautenden Fremdwörtern möglicherweise ein [e] anhängen (Keßler 2013: 3).

Vokalismus: In nebetoniger Position bleiben im Deutschen von der in haupttoniger Position größeren Zahl an Vokalen nur noch Schwundstufen und vokalisierte Konsonanten, während im Italienischen nur die Opposition von offenen und geschlossenen Vokalen verschwindet. Das führt dazu, dass italienische Äußerungen in der metrischen Analyse deutlich mehr rhythmische Schläge erhalten als im Deutschen und daher insgesamt prominenter („vollklingender“) sind.

Konsonantismus: Im Deutschen sind die Merkmale „fortis“ und „lenis“ primär, im Italienischen die Merkmale „stimmhaft“ und „stimmlos“ (Rabanus 2001: 63f.). Dabei werden die stimmhaften Konsonanten im Italienischen auch tatsächlich stimmhaft gesprochen, im Deutschen nicht immer. Weil es im Italienischen nur nichtaspirierte Plosive gibt, werden diese mithilfe des Stimmhaftigkeitsmerkmals unterschieden. Das Phänomen der Langkonsonanz ist eine Besonderheit im Italienischen – fast alle italienischen Konsonanten können lang oder kurz vorkommen, wobei die langen Konsonanten deutlich länger und intensiver artikuliert werden (Keßler

2013: 3). Außerdem sind im Italienischen die Aspiration, die stimmlosen deutschen Konsonanten folgt, sowie der Glottisschlag vor haupttonigen Vokalen im Anlaut völlig unbekannt. Auch die Laute [h], [ç], [x] und die Affrikate [pf] kennt das Italienische nicht. Im deutschen System hingegen sind die Laute [ɲ], [ʎ], [w] und die Affrikate [dz] nicht enthalten (Rabanus 2001: 63f.).

Im Deutschen liegt ein akzentzählender Rhythmus, im Italienischen hingegen ein silbenzählender Rhythmus vor. Der harte, stoßende Rhythmus des Deutschen mit großen Intensitätsschwankungen zwischen akzentuierter und nicht akzentuierter Silbe steht also der italienische Rhythmus mit relativ regelmäßigem Wechsel zwischen betonter und unbetonter Silben und mit geringen Intensitätsunterschieden gegenüber. Dies hängt auch damit zusammen, dass das Deutsche eine Wort- und das Italienische eine Silbensprache ist. Während Silbensprachen durch die Optimierung der phonologischen Silbe der Sprachproduktion dienen und eine leichte Aussprache begünstigen, steht bei den Wortsprachen durch die Verdeutlichung des phonologischen Wortes hingegen die leichte Dekodierung im Vordergrund. Wortsprachen wie das Deutsche sind also eher hörerfreundlich, Silbensprachen wie das Italienische jedoch vorteilhaft für die Sprecher*innen (Nübling et al. 2006: 17).

Der Wortakzent liegt bei italienischen Wörtern meist auf der Pänultima (z. B. *andare*), seltener auf der letzten (z. B. *virtù*), dritt- oder viertletzten Silbe. Außerdem weist die Akzentuierung im Italienischen eine größere Regelmäßigkeit als im Deutschen auf. Der Wortakzent kann auch im Italienischen bedeutungsdistinktiv sein (z. B. *meta* – *metà*). Mehrsilbige Wörter mit drei oder mehr Silben können außerdem einen Nebenakzent tragen, wobei es zu einem relativ regelmäßigen Wechsel von nebetonigen und unbetonten Silben kommt (Keßler 2013: 4).

Deakzentuierungen treffen das Deutsche häufiger als das Italienische. Werden im Deutschen Wörter mit geringer kommunikativer Dynamik deakzentuiert, z. B. weil sie wiederholt sind, bleiben die Akzentverhältnisse im Italienischen normalerweise auch unter solchen Bedingungen unverändert. Stattdessen wird auf das syntaktische Mittel der Rechtsversetzung (*dislocazione a destra*) zurückgegriffen (vgl. dazu Berretta 1995, Berruto 1986) (Rabanus 2001: 63f.). Der dynamische- bzw. Druckakzent im Deutschen zeigt sich in einer Präferenz der Anfangs- bzw. Stammbetonung, in welcher sich die maximale Ausatmungsspannung manifestiert (z. B. *Úr.laub*, *án.fangen*). Dieses Merkmal, welches im melodischen Akzent des Italienischen fehlt, ermöglicht im Deutschen die Bildung einiger Wortpaare, die sich nur in der Akzentuierung semantisch unterscheiden, z. B. [*´um*]*gehen* versus *um*[*´gehen*]; [*´über*]*setzen* versus *über*[*´setzen*]. Die Intonation erfüllt eine pragmatische Funktion, indem sie die Informationen in der

jeweiligen Äußerung kodifiziert. So dient eine flache Kontur einer neutralen Aussage (z. B. *Hans arbeitet den ganzen Tag*), während eine steigende oder fallende Kontur Informationen hervorhebt bzw. weniger wichtig nimmt (Ferrer 1999: 37ff.). Im Italienischen herrscht in Aussagen, Fragen mit Fragewort, Befehlen und Ausrufen der fallende Melodieverlauf vor, wobei die Melodie meist allmählich zum Satzende hin abfällt (kein Fall bis in die Lösungstiefe). In Entscheidungsfragen steigt die Melodie (Keßler 2013: 4).

Die Bedeutung der einzelnen prosodischen Merkmale wurde bereits im Kapitel 2.1 „Prosodie und Sprechausdruck“ erläutert und gilt unabhängig von der jeweiligen Sprache – die konkrete Umsetzung der prosodischen Merkmale jedoch differiert je nach Sprecher*in, Situation und auch Sprache. Daher kann erwartet werden, dass die jeweilige Sprache – Deutsch bzw. Italienisch – auch als Einflussfaktor in den Perzeptionsanalysen fungiert.

2.5.3 Einflussfaktor Geschlecht

In dieser Studie wird eine Unterscheidung zwischen männlichen und weiblichen Hörer*innen vorgenommen, um etwaige geschlechtstypische Unterschiede in der Perzeption des Gesprochenen herauszufiltern und zu untersuchen bzw. um die Ergebnisse durch unscharfes Aufgliedern der Probandengruppen nicht zu verfälschen. Eine erschöpfende Darstellung der geschlechtstypischen Unterschiede im Kommunikationsverhalten oder der feministischen Theoriebildung ist für diese Arbeit nicht relevant. Sex und Gender sind mittlerweile recht gut beforschte Größen im Wissenschaftsdiskurs, wenngleich erst im 20. Jh., mit Beginn der feministischen Linguistik in den 60er-Jahren, das Interesse entstand, Redende als sprechende Männer und Frauen mit ihren Besonderheiten sichtbar zu machen (Heilmann 2011:162). Dabei fokussieren die meisten Forschungen allerdings auf das Sprechen und die Merkmale im direkt beobachtbaren Sprechverhalten, z. B. Körperausdruck, Turn-taking etc. und weniger auf die Perzeption sprachlicher Äußerungen. Nun sei auf eine terminologische Differenzierung verwiesen, die Degenhardt und Trautner (1979: 13) zur Unterscheidung von geschlechtstypischen (Merkmale im Sprech- und Gesprächsverhalten, die häufiger vom einen als vom anderen Geschlecht genutzt werden, prinzipiell aber bei beiden Geschlechtern anzutreffen sind) und geschlechtsspezifischen (Merkmale, die unabänderlich an das jeweilige Geschlecht gebunden sind und nur da vorkommen) Verhaltensweisen vornehmen. In dieser Arbeit werden die Begriffe „geschlechtstypisch“ und „geschlechtsspezifisch“ synonym verwendet, da hier nur die geschlechtstypischen Tendenzen im Perzeptionsvorgang eruiert werden. Besonderheiten des Stimmorgans oder geschlechtsbezogene Spezifika des Stimmklangs bzw. der Tonhöhe werden als gegeben vorausgesetzt und

nicht näher untersucht. Generell gilt auch hier bei der Auswertung der Ergebnisse zu beachten, dass das Geschlecht nur ein wichtiger Einflussfaktor im komplexen Bedingungsgefüge von Kommunikationsprozessen ist. Kotthoff spricht in ihrer These diesbezüglich von einer Relevanzgraduierung (Kotthoff 1996: 13). Diese ist nötig, weil nicht die Geschlechtszugehörigkeit selbst, sondern eher das Ausfüllen einer sozialen Rolle im Rahmen einer soziokulturell-psychischen Gender-Prägung im Zentrum steht. Das bedeutet gleichzeitig, dass Personen, die diese Rollen handelnd gestalten, das Bild von Männlichkeit und Weiblichkeit in der Gesellschaft eigenverantwortlich verändern können (Heilmann 2011: 170). In mehreren Studien konnte außerdem herausgefunden werden, dass weibliches und männliches Gesprächsverhalten gar nicht so sehr voneinander abweicht (vgl. Kowal et al. 1998; Kotthoff 1993; Heilmann 2002b), sondern dass sich vielmehr die Bedeutung der Sozialrolle (z. B. die des Moderators bzw. der Moderatorin) als ausgesprochen dominant, den anderen Gesprächsverhaltensregeln übergeordnet erweist (Heilmann 2011: 170). Daher wird für diese Untersuchung angenommen, dass das Merkmal Geschlecht im Rahmen der Perzeptionsanalysen einen deutlich geringeren Einfluss darstellt als die jeweilige situationsbedingte soziale Rolle bzw. die anderen Einflussfaktoren. Es wird kein relevanter geschlechtsbezogener Unterschied in den Ergebnissen erwartet.

2.5.4 Einflussfaktor Beruf

In dieser Studie spielt der Beruf als Unterscheidungskriterium bei den Perzeptionsanalysen eine wichtige Rolle, da davon ausgegangen wird, dass Stimmexpert*innen (z. B. Sänger*innen oder Logopäd*innen) gegenüber Laien über eine präzisere Wahrnehmungsfähigkeit der prosodischen Parameter verfügen. Alltags- und Erfahrungswissen zeigen, dass ein feines Ohr für sprachliche Nuancen, wie andere Fähigkeiten auch, das Ergebnis von Talent und oder Training ist. Dies zeigt u. a. auch eine britische Studie, in welcher Hirnscans von Sprachexpert*innen und Laien verglichen werden (Golestani et al. 2011: 4213).

Neben den bisher dargestellten und in dieser Arbeit untersuchten Einflussfaktoren auf die Perzeption emotional gefärbter Äußerungen durch prosodische Merkmale gibt es natürlich noch weitere Komponenten, die dabei eine Rolle spielen. So ist das Alter z. B. auch ein wesentliches Kriterium bei der Sprachperzeption. Dieses Merkmal wird in dieser Arbeit allerdings nicht weiter untersucht. Um ein durch breite Altersstreuung der Proband*innen möglicherweise unsauberes Ergebnis zu vermeiden, wurde also eine gewisse Altersspanne (20-30 Jahre) festgelegt,

innerhalb derer sich die Hörer*innen größtenteils bewegen. So können altersbedingte Unterschiede ausgeschlossen und die Einflussfaktoren besser voneinander abgegrenzt werden. Natürlich gilt aber festzuhalten, dass alle Faktoren – Sprache, Kultur, Geschlecht, Beruf, Alter, soziale Gruppe, individuelle Persönlichkeit der Hörer*innen und viele andere mehr – stets zusammenwirken und daher nur Tendenzen festgestellt werden können.

3 Empirische Untersuchung

3.1 Forschungsdesign

Das Forschungsdesign mit Thema, Forschungsfrage und Methodik wurde bereits in der Einleitung im ersten Teil der Arbeit vorweggenommen. Daher soll diesbezüglich ein zusammenfassender Überblick genügen. Vielmehr sollen hier Präzisierungen und auch methodenkritische Überlegungen zur Problematik von Prosodieuntersuchungen ihren Platz finden.

3.1.1 Fragestellung und Zielsetzung

Ziel der empirischen Studie ist es, herauszufinden, anhand welcher prosodischen Merkmale die emotionalen Sprechweisen erkannt werden, welche Relevanz sie für die Perzeption der emotionalen Äußerung haben und wie sie von den einzelnen Hörer*innen rezipiert werden. Dabei sollen bei der Analyse geschlechter- sprach-/kultur- und berufsspezifische (gemeint ist damit die durch entsprechende Ausbildung oder berufliche Tätigkeit verknüpfte Fähigkeit des funktionellen Hörens, wie sie etwa bei Sprecherzieher*innen, Sänger*innen oder Logopäd*innen auftritt) Aspekte berücksichtigt werden. Daraus ergeben sich folgende zentrale Forschungsfragen:

- Welche prosodischen Faktoren spielen beim Erkennen und Perzipieren einer natürlichen emotionalen Äußerung inwieweit eine Rolle?
- Wie werden die einzelnen prosodischen Merkmale der jeweiligen freudigen bzw. ärgerlichen Sprechweisen von den Proband*innen beschrieben?
- Gibt es bei oben genannten Fragen einen sprach-/kulturspezifischen, mit der Fähigkeit des funktionellen Hörens (durch entsprechende Ausbildung/berufliche Tätigkeit) verknüpften oder geschlechterspezifischen Unterschied?

Darüber hinaus ergeben sich weitere Fragestellungen, etwa:

- Wird der Ausprägungsgrad der Emotion von allen Proband*innen gleich empfunden oder gibt es dabei sprachlich/kulturelle, geschlechts- und berufsbezogene Unterschiede?
- Besteht eine Übereinstimmung der perzipierten prosodischen Eigenschaften der Hörbeispiele mit den für die jeweiligen emotionalen Sprechweisen vorliegenden typischen phonetischen Charakteristika (vgl. Kranich 2003)?

Für die empirische Studie in dieser Forschungsarbeit werden die Emotionen Ärger und Freude untersucht. Die Wahl fällt deshalb auf diese Kategorien, da in der bisherigen Forschungsliteratur diese Emotionen weitgehend als Basisemotionen klassifiziert sind, welche zumindest im europäischen Kontext übernationale Gültigkeit haben (vgl. Kranich 2003). Sie sollten sich bei der Analyse also gut im Deutschen und Italienischen für deutsche und italienische Hörer*innengruppen eignen.

3.1.2 Methodenkritik

Bei Untersuchungen zur Perzeption prosodischer Eigenschaften ergeben sich allgemeine Probleme, die auch bei dieser Arbeit berücksichtigt werden müssen.

1. Perzeption von Sprache als Untersuchungsgegenstand

Das Verstehen sprachlicher Elemente und perzeptive Verarbeitungsleistungen entziehen sich einer direkten Beobachtung, d. h. sie sind nur indirekt erfassbar. Auch Untersuchungen zu sprachproduktiven Leistungen weisen diese Problematik auf, nur umgekehrt – der interessierende Output kann zwar direkt bestimmt werden, der den sprachproduktiven Prozess auslösenden Input-Stimulus jedoch ist perzeptiver Natur. Dieses messmethodische Problem bei Untersuchungen zur Perzeption von Sprache ist nach wie vor nicht lösbar (Kranich 2016: 236f.). Diesem Umstand wird hier insofern Rechnung getragen, als dass keine allgemeinen Aussagen über den neurologischen oder innerpsychischen Vorgang der Prosodieperzeption angestrebt oder getätigt werden. Es wird lediglich mit dem operationalisierbaren Output der Proband*innen gearbeitet, wobei Vergleiche darüber Aufschluss geben sollen über die oben genannten Einflussfaktoren des Geschlechts, der Sprache und Kultur sowie der Fähigkeit des funktionalen Hörens (Beruf), ohne dabei auf die inneren Sprachverarbeitungsprozesse näher einzugehen.

2. Differenzierung prosodischer Merkmale von segmentalen Eigenschaften

Sprachliche Zeichen sind beim Sprechen für die Übertragung immer an den Sprechakt gebunden. Daher ist eine Separierung prosodischer und segmentaler Merkmale nur um den Preis einer unnatürlichen Sprechweise möglich, also z. B. in Maskierungsexperimenten zur Tilgung segmentaler Eigenschaften. Derartiges Vorgehen ist in Laborexperimenten denkbar, die daraus resultierenden Befunde lassen sich aber kaum auf natürlich gesprochene Sprache übertragen (vgl.

Kranich 2016: 237). In dieser Studie sollen dennoch natürlich gesprochene emotionale Äußerungen untersucht werden. Dem oben genannten Kritikpunkt kann das Argument vorgebracht werden, dass ansonsten Untersuchungen im natürlichen Kontext fehlen würden. In der Studie werden jedoch einige Maßnahmen in Bezug auf das beschriebene Problem umgesetzt, um den Einfluss des segmentalen Materials gering zu halten. So werden nur Hörbeispiele verwendet, die den Proband*innen durch mehrmaliges Hören bekannt gemacht und die zuvor im Rahmen einer Vorstudie selektiert und als adäquat eingestuft werden. Außerdem wird die Aufmerksamkeit der Hörer*innen durch geeignete Aufgabenstellungen auf die prosodischen Merkmale konzentriert, wodurch die inhaltlichen Aspekte des Gesagten in den Hintergrund rücken. Überdies wird der Einfluss des Textes, also der inhaltlichen Mitteilung, auch dadurch Rechnung getragen, dass er im Fragebogen, neben der prosodischen Merkmale, als ein von den Proband*innen mögliches auszuwählendes Element bei der Frage nach der Zuordnung der Sprechweise zur jeweiligen Emotion angeführt ist. Der Einfluss des Textes auf die Analyse der Hörbeispiele dürfte v. a. bei der Beurteilung der anderssprachigen Sequenzen jedoch äußerst gering sein, weil die Proband*innen die jeweils andere Sprache (das Deutsche bzw. Italienische) nicht verstehen.

3. Separierung einzelner Merkmale aus dem Komplexphänomen Prosodie

Diese Problematik wurde in Teilen bereits im Kapitel zur Prosodie thematisiert. Der Begriff Prosodie subsummiert einen Merkmalskomplex, der sich aus den einzelnen prosodischen Merkmalen zusammensetzt. Im Rahmen empirischer Untersuchungen ist es kaum möglich, alle diese Eigenschaften adäquat zu berücksichtigen. Dennoch ist die Dissolvierung einzelner prosodischer Elemente aus dem Komplexphänomen Prosodie in wissenschaftlichen Untersuchungen dazu notwendig, weshalb auch in dieser Studie die einzelnen Merkmale Akzent/Rhythmus, Tonhöhe/Melodieführung, Lautstärke, Sprechgeschwindigkeit, Pausen und Stimmfärbung (Timbre) extrahiert werden.

Natürlich kann bei einer Untersuchung von Formmerkmalen gesprochener Sprache, unabhängig davon, ob es sich um segmentale oder prosodische Elemente handelt, nicht mehr von einer dem normalen, alltagstypischen mündlichen Kommunikationsablauf vergleichbaren Perzeptionsleistung gesprochen werden. Die Generalisierung experimenteller Befunde auf alltags-sprachliche Situationen muss daher immer kritisch hinterfragt werden. Grundsätzlich wird angenommen, dass prosodische Merkmale von Personen in alltäglichen Kommunikationssi-

tuationen unbewusst als Komplexphänomen wahrgenommen und verarbeitet werden. Die Gewichtung der jeweils relevanten prosodischen Erscheinungen erfolgt normalerweise unbewusst und in Zusammenhang mit der Sinnerschließung des Gesagten, welche die besondere Relevanz einzelner prosodischer Merkmale für den Verstehensprozess festlegt. Bewusst ist hingegen nur das Interpretationsergebnis, also dass etwa die Stimme des Gesprächspartners als angenehm, warm, weich usw. empfunden wird. Da also auch in gewissen alltagstypischen Kommunikationssituationen wahrnehmungsseitig von einer Differenzierung prosodischer Merkmale ausgegangen werden kann, kann diese Prämisse auch für dieses und andere Untersuchungsdesigns als legitim betrachtet werden. Außerdem gilt als gesichertes Wissen, dass die einzelnen prosodischen Merkmale durch konkrete Aufgabenstellungen den Hörer*innen mehr oder weniger bewusst gemacht werden können. Dabei ist aufgrund der Erfahrungen der Prosodieforschung anzunehmen, dass Merkmale zu den Tonhöheneigenschaften sicherer erkannt werden als etwa solche zur Klangqualität (Kranich 2016: 238f.). Generell ist daher die adäquate Erstellung des Fragenkatalogs äußerst wichtig. Für die Erstellung des Fragebogens in diesem Untersuchungsdesign erscheint es also sinnvoll, in den Einzeluntersuchungen Wahlalternativen vorzugeben, die auf eine überschaubare Anzahl von Kategorien beschränkt sind und dennoch über die nötige Trennschärfe verfügen, um valide Aussagen hinsichtlich der Perzeption der emotionalen Sprechäußerungen zu ermöglichen. Außerdem haben die Proband*innen auch die Möglichkeit anzugeben, falls sie die jeweilige Frage zu einer Merkmalsausprägung eines prosodischen Elements nicht beantworten können – so sollen zum einen verzerrte Ergebnisse und zum anderen die potenzielle Überforderung der Proband*innen vermieden werden. Von einem je für die Laien und Expert*innen differenzierten Fragebogen wird abgesehen, da dieses Vorgehen den konstruktiven Vergleich zwischen den beiden Gruppen ausschließen würde.

4. Verschiedenheit der auditorischen Perzeption in experimentellen Designs im Vergleich zum natürlichen Verarbeitungsvorgang gesprochener Sprache

In alltäglichen Kommunikationssituationen nehmen die Zuhörer*innen „durch prosodische Erscheinungen hindurch“ (Kranich 2016: 239) den Inhalt des Gesagten wahr, während Untersuchungen zur Perzeption von Prosodie auch nach dem Formaspekt fragen, welche im Alltag kaum eine Rolle spielen. Die Problematik wirft die Frage auf, inwieweit Ergebnisse solcher Untersuchungen überhaupt auf die Alltagskommunikation und Hörwahrnehmung

übertragen werden können. Grundsätzlich gilt zu beachten, dass es auch bei geschulten Expert*innen nicht zu einem rein analytischen Hören kommt, sondern immer zusätzliche, z. B. semantische, situative oder individuelle Komponenten miteinfließen. Umgekehrt können auch ungeschulte Laien mit Einschränkungen analytisch hören, so kann fast jede Person etwa am Telefon die stimmliche Indisposition eines vertrauten Gesprächspartners feststellen. Das analytische Hören ist vor allem die Fähigkeit der Bewusstseinsbildung für die formseitigen Merkmale der gesprochenen Sprache und kann natürlich (in Studium und Beruf z. B.) auch trainiert werden. Abgesehen davon aber zeigt sich aus Perzeptionsexperimenten zur Prosodie (vgl. Lindner 1977), dass aufgrund des sich bereits im Untersuchungsverlauf einstellenden Lernprozesses der Proband*innen nicht mit einer stabilen individuellen Perzeptionsleistung derselben gerechnet werden kann. Diese Tatsache muss als solche akzeptiert werden und betrifft außerdem auch diverse andere Bereiche, da im Laufe der Individualentwicklung ja permanent Lernprozesse stattfinden. Dieser Vorgang bei der Untersuchung zur Perzeption prosodischer Merkmale lässt sich also auch unter alltäglichen Situationen beobachten.

Insgesamt lässt sich festhalten, dass Untersuchungen zur Perzeption prosodischer Merkmale zwar nicht mit alltagstypischen Kommunikationssituationen vergleichbar sind, sich die derart gewonnenen Ergebnisse jedoch in zumindest ähnlicher Form auf den natürlichen Zuhörprozess übertragen lassen (Kranich 2016: 239f.).

Fazit

Aus den oben erörterten Problemdarstellungen resultieren methodische Schwierigkeiten, die den unterschiedlichen empirischen Perzeptionsuntersuchungen von Prosodie anhaften. Die diversen Probleme lassen sich unter kontrollierten Bedingungen zwar zumindest verringern, aber nicht ganz ausschalten. Alle Untersuchungen zur Perzeption prosodischer Merkmale können daher nur Teilprozesse der Realität abbilden und insgesamt bestenfalls zu einem Gesamtverständnis des Komplexphänomens Prosodie führen (Kranich 2016: 241). Zu enge Fragestellungen der Studie können dieses Phänomen in ihrer Komplexität kaum berücksichtigen, aber für sich genommen wertvolle Teilaussagen generieren, zu weit gefasste Fragestellungen können die für alltagstypische Kommunikationssituationen grundlegenden Details übersehen. Daher muss Neubers Aussage bestätigt werden, hinsichtlich der Zielsetzungen für die Untersuchung prosodischer Merkmale größte Bescheidenheit anzusetzen (Neuber 2002). Bezieht sich diese Aussage zwar auf die signalseitigen Aspekte der prosodischen

Elemente, so gilt sie doch gerade auch unter dem Aspekt der Prosodieperzeption. Die vorliegende empirische Studie ist also im Bewusstsein der bereits thematisierten Problematiken in dem Bestreben entstanden, durch ein entsprechendes Untersuchungsdesign dennoch gültige Aussagen zu einem klar umrissenen Teilbereich des Perzeptionsprozesses zu generieren, um sich dem Themenkomplex der Perzeption prosodischer Merkmale unter dem Gesichtspunkt natürlicher emotionaler Sprechäußerungen wenigstens anzunähern.

3.1.3 Hypothesen

Aufgrund der im ersten Teil dargestellten theoretischen Grundlagen und bisherigen Untersuchungen zur Perzeption von Prosodie im Kontext emotionaler Sprechweisen werden für die vorliegende empirische Studie folgende Hypothesen angenommen:

1. Hörer*innen sind in der Lage, verschiedene natürlich gesprochene Äußerungen den emotionalen Kategorien Ärger oder Freude ausschließlich auf der Grundlage auditiver Präsentation zuzuordnen.
2. Hörer*innen vermögen die prosodischen Merkmale der einzelnen Äußerungen in unterschiedlichem Präzisionsgrad aus dem Gesamteindruck herauszufiltern und zu beschreiben.
3. Die Präzision der Zuordnung einer emotionalen Sprechausdrucksweise sowie die Analyse ihrer prosodischen Charakteristika sind sowohl produktionsseitig von der konkreten Sprechleistung und der jeweiligen Ausprägung prosodischer Merkmale als auch rezeptionsseitig von der Wahrnehmungsleistung der Hörer*innen abhängig.
4. Die Wahrnehmungsleistung der Hörer*innen unterscheidet sich je nach Beruf (Unterscheidungskategorie Laie vs. Expert*in für Stimme und Sprechen) bzw. Ausprägungsgrad ihrer Fähigkeit zum analytischen Hören, wobei die Expert*innen über eine präzisere Wahrnehmung der prosodischen Parameter verfügen.
5. Die Perzeption der prosodischen Merkmale differiert in Bezug auf die jeweilige Sprache und Kultur der Hörer*innen, wobei die Äußerungen aufgrund der jeweiligen sprachlich-kulturellen Charakteristika des Deutschen und Italienischen (vgl. Kapitel: Sprachvergleich. Prosodie im Deutschen und Italienischen) jeweils anders beurteilt werden.

6. Die Perzeption der prosodischen Merkmale unterscheidet sich nicht aufgrund des jeweiligen Geschlechts der Hörer*innen, sondern ist vielmehr sozial bedingt.
7. Die den verschiedenen emotionalen Sprechäußerungen zugeordneten prosodischen Charakteristika entsprechen in groben Zügen der für die jeweilige gesprochene Emotion typischen phonetischen Eigenschaften (nach Kranich 2003).
8. Die Wahrnehmung und Zuordnung der einzelnen emotionalen Äußerungen erfolgt ganzheitlich als komplexes Phänomen durch eine Zusammenwirkung aller sprachlichen und prosodischen Parameter mit verschiedenartiger Dominanz.
 - a) Die Kategorie Akzent/Rhythmus ist weniger relevant aufgrund der Schwierigkeit, sie an emotionalen Merkmalen festzumachen.
 - b) Die Merkmale Tonhöhe/Melodieführung verfügen über eine herausragende Bedeutung im Perzeptionsprozess, v. a. bei der Emotion Freude (hohe Sprechstimmlage, großer Tonumfang erwartet).
 - c) Text und Wortwahl spielen eine zentrale Rolle bei der Zuordnung des Gehörten, weil sie in allen alltagstypischen Kommunikationssituationen zunächst die primäre Sinnkonstitution erlauben, beeinträchtigen aufgrund der entsprechenden Aufgabenstellungen aber nicht die Perzeption der prosodischen Merkmale.
 - d) Die Lautstärke ist insgesamt weniger relevant, bei der Emotion Ärger jedoch etwas gewichtiger.
 - e) Die Sprechgeschwindigkeit spielt eine untergeordnete Rolle, sie ist bei beiden Emotionen hoch.
 - f) Pausen spielen in den vorliegenden Beispielen keine Rolle. Bei anderen Emotionen, z. B. Trauer, oder bei gefüllten Pausen (z. B. Lachen, Seufzen usw.) wären sie hingegen wesentlich.
 - g) Der Stimmfärbung (Timbre) wird eine herausragende Stellung zugesprochen. Sie unterscheidet sich bei den beiden Emotionen (Ärger: hell, scharf, klangarm, behaucht; Freude: hell, weich, warm, weit; Kranich 2003: 198), ist aber besonders für die Laien schwer zu erfassen.

3.2 Untersuchungsaufbau und –durchführung

3.2.1 Aufbereitung und Selektion des Audiomaterials

Auf die Relevanz eines adäquaten Untersuchungsaufbaus im Rahmen einer Forschung zur Prosodieperzeption wurde bereits mehrfach hingewiesen. Da die auditive Analyse von alltagstypischen emotionalen Sprechäußerungen als Ziel und Methode der Arbeit im Vordergrund steht, ist die Auswahl des entsprechenden Untersuchungsmaterials für die Generierung brauchbarer Ergebnisse von größter Bedeutung. Aus diesem Grund wurde eine Vorstudie durchgeführt, um das passende Audiomaterial für die in der eigentlichen Untersuchung zu analysierenden Hörbeispiele zu selektieren. Dabei wurde das Material in mehreren Schritten vorausgewählt, bearbeitet, durch je italienisch- und deutschsprachige Kontrollgruppen analysiert und schließlich ausgewählt.

Korpuserstellung

Im ersten Schritt musste ein Korpus passender mündlicher emotionaler Sprechäußerungen (in den Ausprägungen Ärger und Freude mit deutsch- und italienischsprachigen und je weiblichen bzw. männlichen Sprech*innen) im natürlichen Sprachgebrauch als Grundlage der empirischen Untersuchung erstellt werden. Auf bereits vorhandene Korpora konnte nicht zurückgegriffen werden, da die darin gesammelten Ausschnitte zumeist experimentellen Sprechsituationen entstammen, in denen die jeweiligen Emotionen induziert, d. h. von Schauspieler*innen auf der Grundlage entsprechender Vorgaben künstlich erzeugt und vorgetragen waren. Das sterile Umfeld im Tonstudio erlaubt hier zwar eine ungleich höhere technische sowie den Zwecken der jeweiligen Studie entsprechende Qualität, der Output kann jedoch nicht als natürliche emotionale Sprechweise betrachtet werden. Daher stellte sich hier die Herausforderung, entsprechendes natürliches Material in genügender Qualität und natürlich unter Berücksichtigung aller rechtlichen Aspekte zu sammeln. Ausgewählte, im Internet zugängliche Fernsehsendungen schienen diesen Anforderungen zu genügen, so wurden verschiedene Aufnahmen anhand bestimmter Kriterien vorausgewählt:

- Es muss sich um natürliche emotionale Sprechweisen handeln, d. h. die emotionale Äußerung muss „echt“ sein und darf nicht gespielt werden.

- Die sprechende Person soll im Idealfall nicht darüber informiert sein, dass sie aufgenommen wird bzw. die Aufnahmesituation darf nicht so vordergründig sein, dass sie die authentische Emotionsäußerung unterbindet.
- Die emotionalen Äußerungen sind in einer natürlichen Kommunikationssituation kurz, die ausgewählten Turns sollen daher einen Umfang von 7 – 12 Wörtern aufweisen. Dadurch wird auch ihre Vergleichbarkeit gewährleistet.
- Die Äußerungen sollen den Emotionen Freude und Ärger zugewiesen werden können und je von männlichen und weiblichen muttersprachlichen Sprecher*innen je deutscher und italienischer Sprache gesprochen sein.
- Es soll sich um möglichst standardsprachliche Äußerungen handeln, wobei individuelle regiolektale Anteile in einer authentischen Kommunikationssituation kaum zu vermeiden sind und das Ergebnis im Rahmen der vorliegenden Fragestellungen nicht beeinträchtigen.

Die Aufnahmen wurden der Plattform YouTube entnommen. Die Sprecher*innen bedienen sich im Wesentlichen der jeweiligen Standardaussprache. Bei einigen Sprecher*innen sind regiolektale Anklänge zu verzeichnen, die jedoch vernachlässigt werden können, weil hier nicht die standardaussprachliche Phonemrealisation untersucht werden soll. Von ausgesprochenen Dialektsprecher*innen kann nicht die Rede sein.

Eine genaue Übersicht mit Auflistung der Links, Angabe der Minuten und Textausschnitte findet sich im Anhang.

Alle Ausschnitte für die Emotion Ärger im Deutschen, für weibliche und männliche Sprecher*innen, sind der Fernsehsendung „Verstehen Sie Spaß“ entnommen. In dieser Sendung werden Menschen mit versteckter Kamera in diverse ärgerliche, komische oder außergewöhnliche Situationen gebracht und beobachtet. Da die Personen nicht wissen, dass sie unter Beobachtung stehen, sind die geäußerten emotionalen Sprechweisen zu hundert Prozent authentisch und das Format der Sendung daher prädestiniert für diese Studienzwecke.

Die Ausschnitte für die Emotion Freude im Deutschen für männliche und weibliche Sprecher*innen entstammen ebenso der Fernsehsendung „Verstehen Sie Spaß“, außerdem Aufnahmen emotionaler Live-Situationen der gefilmten Person (Erfahren von einer Schwangerschaft), Interviews (im Rahmen von Veranstaltungen und Sportevents) und dem Live-Kommentar der Fußball-Weltmeisterschaft im Jahr 2014. In letztgenannten Ausschnitten sind sich die sprechenden Personen zwar der Aufnahmesituation bewusst, aufgrund der spezifischen Situation

jedoch rückt das künstliche Setting mit Kamera und Mikrofon jedoch soweit in den Hintergrund, dass die Authentizität der emotionalen Äußerungen durchaus gewährleistet ist und daher zumindest von halbnatürlichen Sprechweisen die Rede sein kann.

Die Ausschnitte für die Emotion Ärger im Italienischen wurden für die weiblichen und männlichen Sprecher*innen der Reality-Show „Amici“ des Kanals „Real Time“ entnommen. In dieser Serie werden die Personen über einen längeren Zeitraum in ihrem Alltag gefilmt. Auch hier handelt es sich um eine den Sprecher*innen bekannte künstliche Aufnahmesituation. Dadurch, dass die Personen jedoch über längere Zeit hinweg damit konfrontiert sind, kann davon ausgegangen werden, dass sich ein gewisser Gewöhnungseffekt an die Aufnahmesituation eingestellt hat und die geäußerten emotionalen Aussagen authentisch sind.

Die Sequenzen für die Emotion Freude im Italienischen für männliche und weibliche Sprecher*innen hingegen entstammen wiederum der Reality-Show „Amici“ sowie der Reality-Show „Master Chef“ und „X-Factor Italia“, Interviews (Festival Sanremo und Sportevents), Aufnahmen emotionaler Live - Situationen der gefilmten Person sowie dem Live-Kommentar bei der Fußballweltmeisterschaft 2004. Auch hier gilt festzuhalten, dass die ausgewählten Situationen trotz des künstlichen Aufnahmesettings einen authentischen Emotionsausdruck sowie die Erfüllung der weiteren oben dargestellten Kriterien gewährleisten.

Technische Aufbereitung der Audiodateien

Die Audiodateien mussten natürlich entsprechend aufbereitet werden, nachdem die Videos als Ausgangsformat und der passende Ausschnitt daraus festgelegt waren.

Mit dem Programm Free YouTube to Mp3 Converter (Premium) von Digital Wave Ltd. wurde zunächst die gesamte Tonspur des Videos extrahiert und in das Audio-Format WAVE konvertiert. Dann erfolgten mit dem Programm WaveLab 8 der Firma Steinberg Media Technologies GmbH der Schnitt der Tonspur an den erforderlichen Stellen und die Aufbereitung des Tonmaterials. Dabei wurde die Lautstärke maximiert, damit alle Tonbeispiele auf den annähernd selben Pegel ausgesteuert sind.

Insgesamt wurden auf Grundlage des Kriterienkatalogs 68 Audiodateien aus den genannten Aufnahmen aus Internet und Fernsehen erstellt. 37 Dateien davon sind Äußerungen in deutscher Sprache, wobei 20 Äußerungen freudiger Natur sind und 17 Äußerungen der Emotion Ärger zugeordnet werden. 31 Dateien hingegen sind Äußerungen in italienischer Sprache, die sich aus

14 freudigen und 17 ärgerlichen Sprechweisen zusammensetzen. Die zahlenmäßig unterschiedliche Zusammensetzung der Hörbeispiele für das Deutsche und Italienische und für die jeweiligen Emotionen der Freude bzw. des Ärgers hat seine Ursache rein in der unterschiedlichen Verfügbarkeit geeigneten Materials. In jedem Fall ist das Korpus jedoch so bemessen, um zweifelsohne genügend Wahlmöglichkeiten pro Emotion, Sprache und Geschlecht der Sprecher*innen sicherzustellen.

An dieser Stelle sei auf zwei in Bezug auf das Untersuchungsdesign notwendige Entscheidungen hingewiesen, die sowohl die Vorstudie als auch die Hauptuntersuchung betreffen:

1. Verwendung der Audiodateien ohne Videofunktion

Vom ursprünglichen Ausgangsmaterial mit Video wurden nur Audiodateien ohne Bild oder Video verwendet. Natürlich muss beachtet werden, dass die Perzeption von sprachlichen Äußerungen nicht nur über das Gehör vollzogen wird. Es handelt sich vielmehr um einen multimodalen Prozess, in dem auch die visuelle Komponente eine bedeutende bzw. z. T. vorrangige Rolle spielt. Dies gilt insbesondere für einen Kommunikationskontext mit starkem Hintergrundlärm, mit hörbeeinträchtigten Personen, bei besonders komplizierten Sachverhalten und auch bei Gesprächspartnern mit starkem fremdländischem Akzent. Der Sehsinn ist auch im Rahmen der kindlichen Sprachentwicklung äußerst relevant – so werden bestimmte phonetische Distinktionen, die schwer zu hören, aber einfach zu sehen sind (z. B. /m/ versus /n/) vor allem visuell erlernt (Rosenblum 2005: 51f.). Als die „McGurk-Effekte“ (vgl. McGurk/MacDonald 1976) sind außerdem jene Phänomene bekannt, bei denen der Vorrang der visuellen vor der auditiven Komponente im Experiment nachgewiesen werden konnte. So wurden unterschiedliche visuelle (z. B. /va/) bzw. auditive (z. B. /ba/) Stimuli gleichzeitig vorgeführt, wobei sich die Proband*innen durch den visuellen Eindruck in der Perzeption täuschen ließen (sie „hörten“ z. B. das /va/, obgleich /ba/ gesprochen wurde) (Rosenblum 2005: 51f.). Um den Einfluss der visuellen Komponente auf die Perzeption emotionaler Sprechweisen durch prosodische Merkmale ausschließen und damit möglicherweise verfälschte Ergebnisse vermeiden zu können, wird in dieser Studie bewusst nur die auditive Komponente erfasst und betrachtet. Aus diesem Grund werden die sprachlichen Äußerungen nur im Audioformat angeboten, die Videofunktion wurde bewusst nicht miteinbezogen. Natürlich gilt jedoch zu beachten, dass in einer realen Gesprächssituation stets auch die visuelle Komponente und damit alle weiteren Aspekte der Kommunikation, z. B. die nonverbale Ebene, Kleidung, Umgebung etc., mitspielen. Diese Faktoren müssen für die Untersuchung hier jedoch ausgeklammert werden.

2. Verzicht auf instrumentalphonetische Analyse des Audiomaterials

Prosodische Eigenschaften können sowohl auf perzeptiver als auch auf akustischer Ebene beschrieben werden. In auditiven Analysen, wie sie im Rahmen dieser Studie angewendet werden, sind prosodische Merkmale qualitativ analysierbar. Hierbei gilt zu berücksichtigen, dass die Ergebnisse immer auch von den Fähigkeiten, Erfahrungen und Einstellungen der analysierenden Personen abhängen und keine präzisen Auswertungen ermöglichen. Diese etwaigen Nachteile können durch zusätzliche objektive akustische Messungen aufgefangen werden. Die gemessenen akustischen Größen wiederum können jedoch nicht linear auf die Wahrnehmung des Menschen bezogen werden, ja widersprechen sich sogar zum Teil mit dieser (Enterlein/Bartels/Sendlmeier 2005: 10). Darum und weil es in dieser Arbeit um die individuelle Perzeption der Proband*innen und nicht um objektive akustische Daten geht, wird hier auf akustische Messungen des Audiomaterials verzichtet. Sie würden für die hier vorliegende Forschungsfrage keinen relevanten Erkenntniszuwachs ermöglichen.

Vorstudie zur Selektion des Audiomaterials

Um aus den 68 ausgesuchten und aufbereiteten Audiodateien die acht Sequenzen auszuwählen, welche dann im Rahmen der empirischen Hauptuntersuchung analysiert werden sollen, wird eine Vorstudie durchgeführt, in der zwei Kontrollgruppen mit je deutsch- und italienischsprachigen Teilnehmer*innen die einzelnen Audiosequenzen bewerten sollen. Dabei sollen sie angeben, inwieweit die abgespielten emotionalen Äußerungen der jeweiligen Emotion Freude oder Ärger zugeordnet werden können. Anhand eines Fragebogens werden die Äußerungen in Bezug auf ihren emotionalen Erkennungswert beurteilt, wobei für jedes Hörbeispiel fünf Kategorien zur Verfügung stehen (1 = Emotion unklar; 2 = Emotion wenig klar; 3 = Emotion relativ klar; 4 = Emotion klar; 5 = Emotion sehr klar). Dieses Vorgehen hat den Vorteil, dass induktiv ein Beurteilungsraster erfasst werden kann, der es erlaubt, nur die eindeutig zuordenbaren emotionalen Sprechweisen auszuwählen und subjektive Verzerrungen zu vermeiden. Die Fragebögen sind je in deutscher und italienischer Sprache verfasst und wurden händisch von den Proband*innen ausgefüllt. Vor der Beurteilung der Audiosequenzen wurden einige demographische Daten ausgefüllt, wobei für die Auswertung der Ergebnisse nur das Geschlecht (m/w) und die jeweilige Muttersprache herangezogen werden.

Insgesamt 21 Personen, davon neun Männer und zwölf Frauen, haben an der Vorstudie teilgenommen. Sie wurden vor der Bewertung der Hörbeispiele entsprechend instruiert und es wurden alle Fragen zum Projekt geklärt. Die Personen befinden sich in einer recht homogenen

Altersrange (ca. 20 – 30 Jahre), da es sich um Student*innen handelt. Außerdem sind die Proband*innen entweder deutsche oder italienische Muttersprachler*innen ohne Kenntnisse der jeweiligen anderen Sprache. Es sei hier nochmals darauf hingewiesen, dass beide Gruppen (also deutsch- und italienischsprachige) alle Hörbeispiele, d. h. auch die Sequenzen der jeweils anderen Sprache, anhören und bewerten. Nur so kann die passende Auswahl des Audiomaterials sichergestellt werden. Was die Reihenfolge der Hörbeispiele betrifft, so beginnen alle Proband*innen zunächst mit der eigenen Muttersprache, da dieses Vorgehen für einfacher betrachtet wird. Die Abfolge ist aus Gründen der Vergleichbarkeit in beiden Gruppen dieselbe:

Eigene Muttersprache (dt./it.):

- Freude, weibliche Sprecherin
- Freude, männlicher Sprecher
- Ärger, weibliche Sprecherin
- Ärger, männlicher Sprecher

Andere Sprache (dt./it.):

- Freude, weibliche Sprecherin
- Freude, männlicher Sprecher
- Ärger, weibliche Sprecherin
- Ärger, männlicher Sprecher

Zur Durchführung der Vorstudie wurde je ein Seminarraum mit dem notwendigen technischen Equipment an der Universität benutzt, um die Audiosequenzen in genügender Tonqualität und beliebig oft abspielen zu können. Die Gruppe mit den italienischen Proband*innen setzt sich aus elf Personen (fünf Männer und sechs Frauen) zusammen, die dem Gebiet Venedig, Treviso und Umgebung entstammen und an der Università Ca' Foscari in Venedig humanistische Fächer studieren. Die Befragung der deutschsprachigen Proband*innen hingegen erfolgte an der Universität Regensburg mit insgesamt zehn Personen (vier Männer, sechs Frauen), die das Studienfach Sprechwissenschaft belegen und daher als Expert*innen der mündlichen Kommunikation gewertet werden können.

Die detaillierte Übersicht über die Ergebnisse und die Fragebögen (dt./it.) befinden sich im Anhang. Durch die Durchführung der Befragung konnten diejenigen acht emotionalen Sprechweisen ausgewählt werden, die von allen Proband*innen am besten den Kategorien Freude bzw. Ärger zugeordnet werden konnten:

- Hörbsp. 9: deutsch, Freude, weibliche Sprecherin
- Hörbsp. 15: deutsch, Freude, männlicher Sprecher
- Hörbsp., 30: deutsch, Ärger, weibliche Sprecherin
- Hörbsp. 36: deutsch, Ärger, männlicher Sprecher
- Hörbsp. 46: italienisch, Freude, weibliche Sprecherin
- Hörbsp. 47: italienisch, Freude, männlicher Sprecher
- Hörbsp. 57: italienisch, Ärger, weibliche Sprecherin
- Hörbsp. 67: italienisch, Ärger, männlicher Sprecher

Bei den oben genannten für die Untersuchung ausgewählten Hörbeispielen ist die Bewertung der beiden Gruppen und Geschlechter eindeutig und übereinstimmend, d. h. es lassen sich keine relevanten Unterschiede zwischen Männern und Frauen und italienisch- oder deutschsprachigen Personen erkennen. Insgesamt lassen sich bei der Auswertung der Vorstudie aber einige Tendenzen feststellen, für ein genaues und valides Ergebnis wäre jedoch eine Untersuchung in größerem Stil von Nöten. Da das Ziel hier nur die Selektion der Sequenzen für die Untersuchung war, wurde darauf verzichtet. Es werden folgend also nur kurz allgemeine Beobachtungen angeführt. Die Bewertung der deutschen Sprechäußerungen mit der Emotion Freude erfolgt tendenziell sehr ähnlich zwischen den deutsch- und italienischsprachigen Proband*innen und zwischen Männern und Frauen. Die deutschen Äußerungen mit der Emotion Ärger hingegen werden von den deutschsprachigen Proband*innen tendenziell eher höher bewertet als die italienischsprachige Gruppe, die Emotion wird also von der deutschsprachigen Gruppe klarer erkannt. Möglicherweise spielen hier das Textverständnis oder auch sprachlich-/kulturelle Faktoren (im Italienischen wird Ärger möglicherweise mit stärkerem emotionalen Transport bzw. anderen sprachlichen Mitteln ausgedrückt) eine Rolle. Die italienischen, v. a. von weiblichen Sprecherinnen getätigten Äußerungen mit der Emotion Freude werden hingegen von den italienischsprachigen Proband*innen deutlich höher bewertet bzw. klarer der Emotion zugeordnet als von der deutschsprachigen Gruppe, und zwar unabhängig vom Geschlecht der Proband*innen. Auch hier könnten das Textverständnis bzw. die sprachlich/kulturell unterschiedliche prosodische Realisation der Emotion eine Rolle spielen. Die italienischen Äußerungen mit weiblichen und männlichen Sprecher*innen mit der Emotion Ärger hingegen werden von allen Proband*innen unabhängig ihrer Kultur/Sprache und ihres Geschlechts sehr ähnlich bewertet, d. h. von allen homogen als mehr oder wenig klar betrachtet. Zusammenfassend lässt sich feststellen, dass die deutschen Äußerungen zur Emotion Freude sowie die italienischen Äußerungen zur

Emotion Ärger von allen Proband*innen größtenteils gleich bewertet wurden, während es hingegen zu Abweichungen kam bei der Bewertung von deutschen Äußerungen mit der Emotion Ärger und von italienischen, v. a. von weiblichen Sprecherinnen getätigten Äußerungen mit der Emotion Freude. Die Abweichungen sind hier nicht durch das Geschlecht der Proband*innen bedingt, sondern durch ihre kulturelle/sprachliche Herkunft und die damit wohl verknüpften unterschiedlichen emotionalen Sprechausdrucksweisen und Hörmuster, die sich bei der Perzeption diverser Äußerungen niederschlagen.

3.2.2 Fragebogenerhebung

Nach der im Rahmen der Vorstudie durchgeführten Selektion der acht emotionalen Äußerungen, die für die Analyse herangezogen werden sollen, erfolgt die Untersuchung der Perzeption dieser emotionalen Sprechweisen durch prosodische Merkmale. Hierfür wurde ein Fragebogen in deutscher und italienischer Sprache entworfen und vier Proband*innengruppen zur Durchführung vorgelegt:

- deutsche Muttersprachler*innen, Laien (keine Stimm-/Sprachexpert*innen)
- deutsche Muttersprachler*innen, Stimmexpert*innen (z. B. Sprecherzieher*in)
- italienische Muttersprachler*innen, Laien (keine Stimm-/Sprachexpert*innen)
- italienische Muttersprachler*innen, Stimmexpert*innen (z. B. Sprecherzieher*in)

In allen Gruppen vier Gruppen wird auf die Verteilung von Männern und Frauen geachtet und dieses Geschlechtsmerkmal bei der Auswertung entsprechend berücksichtigt. Es wurde beschlossen, in dieser Untersuchung nur die Kategorien Mann / Frau bei der Analyse zu unterscheiden, alle weiteren Geschlechter müssten in einer eigenen entsprechenden Studie einbezogen und differenziert ausgewertet werden. Diese Unterscheidung beruht angesichts der Forschungsfrage und des Umfangs der Untersuchung auf rein pragmatischen Gründen, ohne andere Geschlechter dabei diskriminieren zu wollen.

In Bezug auf den Beruf der Proband*innen wird zwischen Laien und Expert*innen in den Bereichen Stimme, Sprache und Sprechen und damit zwischen deren mehr oder minder ausgeprägten Fähigkeit des analytischen Hörens unterschieden.

Im Rahmen dieser Studie werden folgende Berufsgruppen zur Kategorie der Sprachexpert*innen gezählt: Sänger*innen, Logopäd*innen, Sprecherzieher*innen, Sprechwissenschaftler*innen.

Alle weiteren Berufsgruppen zählen zur Kategorie der Laien. Insgesamt nehmen 82 Personen an der Umfrage teil, davon 28 Männer und 54 Frauen, 48 deutsche, 33 italienischsprachige sowie 52 Laien und 29 Stimmexpert*innen. Bei der Auswertung der Ergebnisse werden alle gegebenen Antworten berücksichtigt, auch wenn einzelne Teilnehmer*innen nicht jede Frage beantwortet haben. In der folgenden Übersicht ist die genaue Zusammensetzung der einzelnen Teilnehmergruppen ersichtlich.

	Deutsch	Italienisch
Kein Stimmexperte	27: 6m 21w	25: 12m 13w
Stimmexperte	21: 9m 12w	8: 1m 7w
Geschlecht	48: 15m 33w	33: 13m 20w

Die Proband*innen gehören verschiedenen Nationalitäten (v. a. Italien, Deutschland, Österreich, Schweiz u. a.) an, Unterscheidungskriterium ist jedoch nicht die politische Zugehörigkeit zu einem Staat, sondern die Zugehörigkeit zur deutschen oder italienischen Sprachgruppe ohne Kenntnis der jeweils anderen untersuchten Sprache. Die Fragebögen sind im Anhang ersichtlich. Zu Beginn werden demographische Daten aufgenommen, die die nähere Differenzierung der Proband*innen erlaubt. Im Folgenden werden die Daten einbezogen und ausgewertet, die für die Beantwortung der Forschungsfragen relevant sind. Nicht näher eingegangen wird auf die Unterscheidung zwischen Dialektsprecher*innen und Personen, die nur die Standardsprache sprechen. Dies wäre in einer eigenen Studie nachzuprüfen unter Berücksichtigung der diversen deutschen und italienischen Dialekte. Ein Großteil der Befragten gab an, auch Dialekt zu sprechen. Dies spiegelt die vielfältige Sprachlandschaft im deutschsprachigen und italienischen Gebiet sehr gut wider, in dem fast alle Personen streng genommen gewisse regionale Anklänge aufweisen. Alle Proband*innen bewegen sich jedoch im akademischen Umfeld und betätigen sich beruflich v. a. der jeweiligen Standardsprache, daher wurde auf eine nähere Differenzierung in der Auswertung verzichtet.

Die Fragebogenerhebung erfolgte online, nachdem die vier Proband*innengruppen ausgewählt und in das Thema eingeführt worden waren, über einen Zeitraum von circa einem Semester. Bei den befragten Personen handelt es sich um Student*innen zwischen 20 und 30 Jahren der Universitäten Wien, Regensburg, Venedig und des Konservatoriums in Venedig. Die Stimmexpert*innen sind z. T. etwas älter und wurden, für die deutsche Gruppe, im berufsbegleitenden Masterstudiengang der Sprechwissenschaft und Rhetorik der Universität Regensburg zur Befragung eingeladen. Die italienischsprachigen Stimmexpert*innen hingegen wurden v. a. am

Konservatorium in Venedig (Studienzweig Gesang) gefunden. In Italien gibt es keinen vergleichbaren Studienzweig der Sprechwissenschaft, Sprecherziehung oder Rhetorik, wie sie in Deutschland an diversen Universitäten gelehrt werden. Außerdem gilt zu beachten, dass in den genannten Studienfächern der absolute Großteil der Studierenden aus Frauen besteht. Diese Tatsachen erklären die unterschiedliche Zusammensetzung und Größe der einzelnen Teilnehmergruppen. Die Befragung wurde online nach freier Zeiteilteilung der Proband*innen durchgeführt und erfolgte auf der Internet-Plattform UmfrageOnline (www.umfrageonline.com). Dies hat den Vorteil, dass die Proband*innen sich in eigenem Tempo den doch sehr aufwändigen Analysen widmen und die Hörbeispiele so oft abspielen können, wie sie wollen. Außerdem wird darauf hingewiesen, dass die einzelnen prosodischen Merkmale genau erklärt und beschrieben sind in einem verpflichtenden Tooltip zu Beginn der Analyse, um sicherzugehen, dass allen Personen die zu bewertenden Kategorien klar sind. Die Bezeichnungen sind ja, gerade für Laien, nicht für alle leicht verständlich oder intuitiv zugänglich.

3.3 Ergebnisse

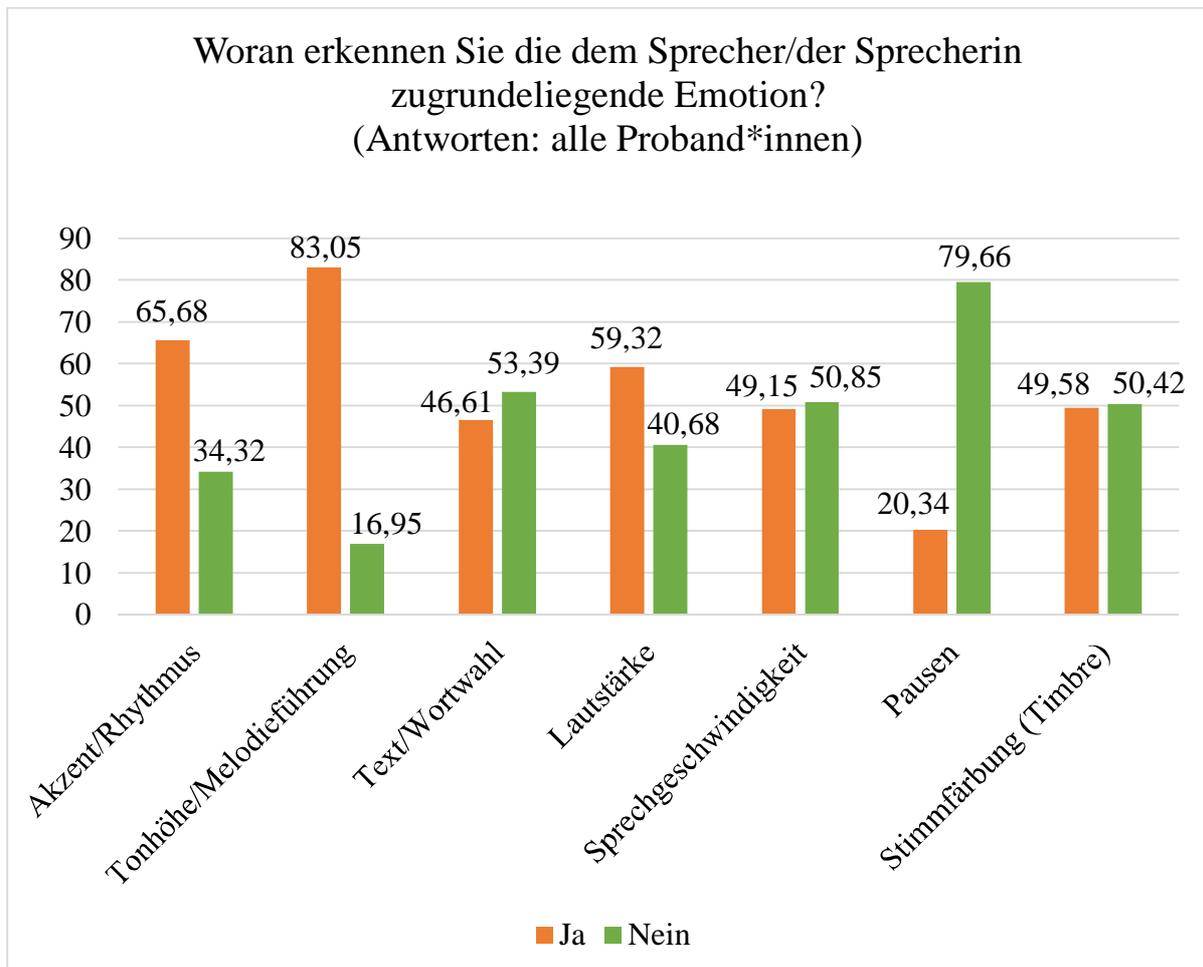
Im Folgenden werden die Ergebnisse aus der Untersuchung zur Perzeption emotionaler Äußerungen durch prosodische Merkmale im Deutschen und Italienischen vorgestellt. Die Daten wurden zunächst von der Plattform UmfrageOnline, welche nur einen ersten groben Überblick bietet, auf die Statistiksoftware IBM SPSS übertragen und hier näher ausgewertet. Die verschiedenen Proband*innengruppen und Einflussfaktoren wurden gegeneinander abgeglichen und durch zahlreiche Kombinationsmöglichkeiten verschiedene Zusammenhänge, Analogien und Unterschiede in der Perzeption herausgearbeitet. Der Anschaulichkeit wegen werden die wichtigsten Ergebnisse in Diagrammen dargestellt, wobei die Auswertung vom Großen hin zum Kleinen erfolgt, also vom allgemeinen Überblick hin zur detaillierteren Beobachtung. Die angegebenen Zahlen im Diagramm bilden die prozentuale Häufigkeit der jeweiligen Nennungen ab, wobei die Prozente auf zwei Kommazahlen gerundet wurden.

3.3.1 Darstellung der für die Emotionserkennung zentralen prosodischen Merkmale

Durch die erste inhaltliche Frage im Fragebogen soll herausgefunden werden, welche prosodischen Parameter überhaupt bei der Emotionserkennung der jeweiligen Äußerungen eine Rolle spielen. Auch der Text bzw. die Wortwahl werden als eigene Kategorie miteinbezogen, um ein

vollständiges Profil zu erhalten. Da v. a. die prosodischen Merkmale von Interesse sind, die insgesamt bei der Emotionserkennung relevant sind, und nicht die der einzelnen Äußerungen für sich genommen, erfolgt die Auswertung der Ergebnisse in Bezug auf alle acht Hörbeispiele. In einem zweiten Schritt werden die Ergebnisse auch in Bezug auf die jeweiligen emotionalen Sprechweisen Freude bzw. Ärger differenziert dargestellt. Es gilt zu beachten, dass bei der Bewertung Mehrfachantworten möglich waren. Alle Angaben sind in Prozent.

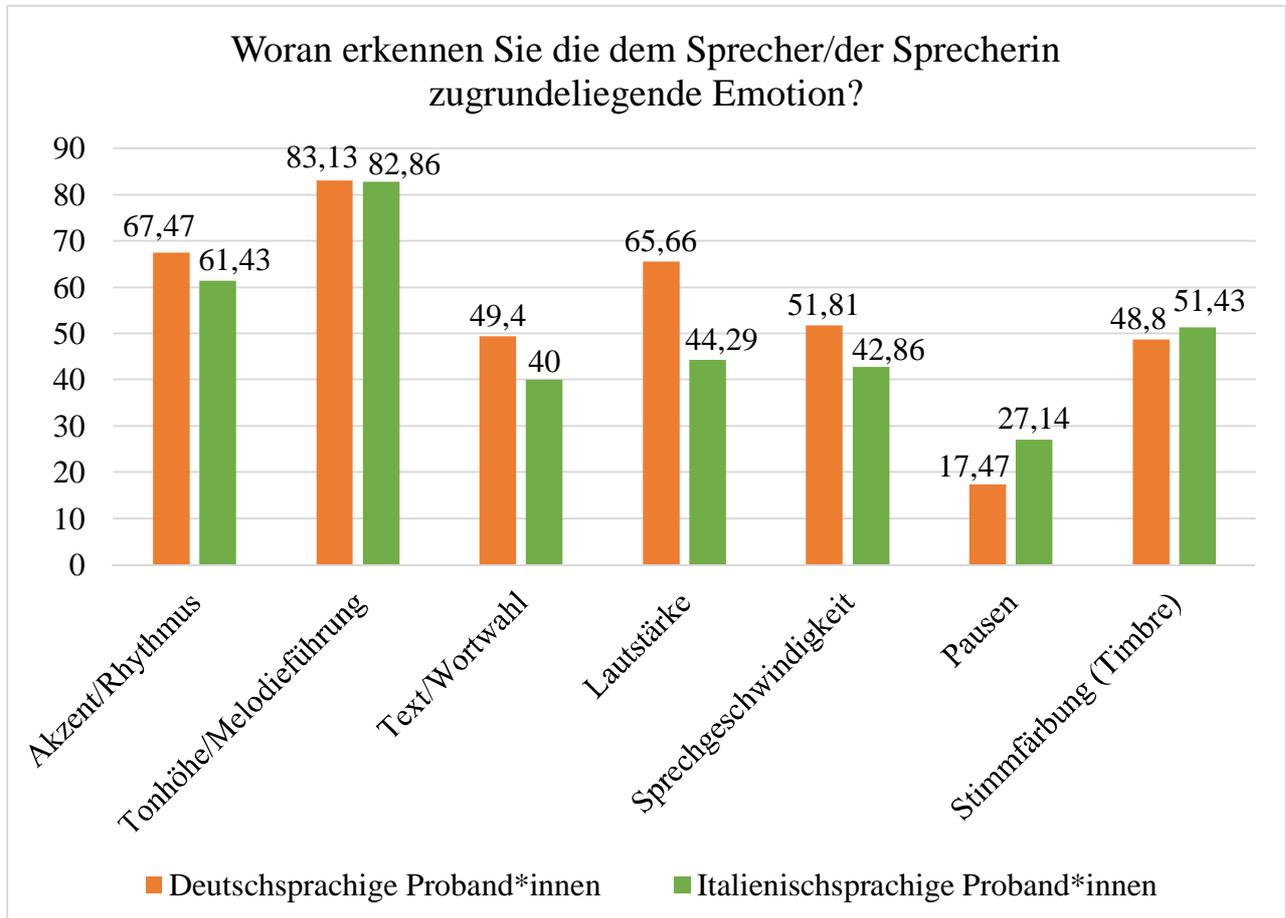
3.3.1.1 Gesamtüberblick aller Antworten



Zunächst soll ein Gesamtüberblick über die Nennungen aller Proband*innen darüber Aufschluss geben, welche prosodischen Merkmale generell in welcher Häufigkeit zur Emotionserkennung der acht Äußerungen herangezogen wurden. In absteigender Reihenfolge sind dies Tonhöhe/Melodieführung, Akzent/Rhythmus, Lautstärke, Stimmfärbung (Timbre), Sprechgeschwindigkeit, Text/Wortwahl und Pausen. Die grünen Balken im Diagramm geben die Zahl der Nennungen wieder, die orangenen die Zahl derjenigen, die dieses Merkmal nicht angegeben

haben. Differenzen dieser Angaben in Bezug auf Sprache/Kultur, Beruf und Geschlecht der Proband*innen lassen sich in den folgenden Diagrammen und Beschreibungen herauslesen.

3.3.1.2 Deutsch- und italienischsprachige Proband*innen



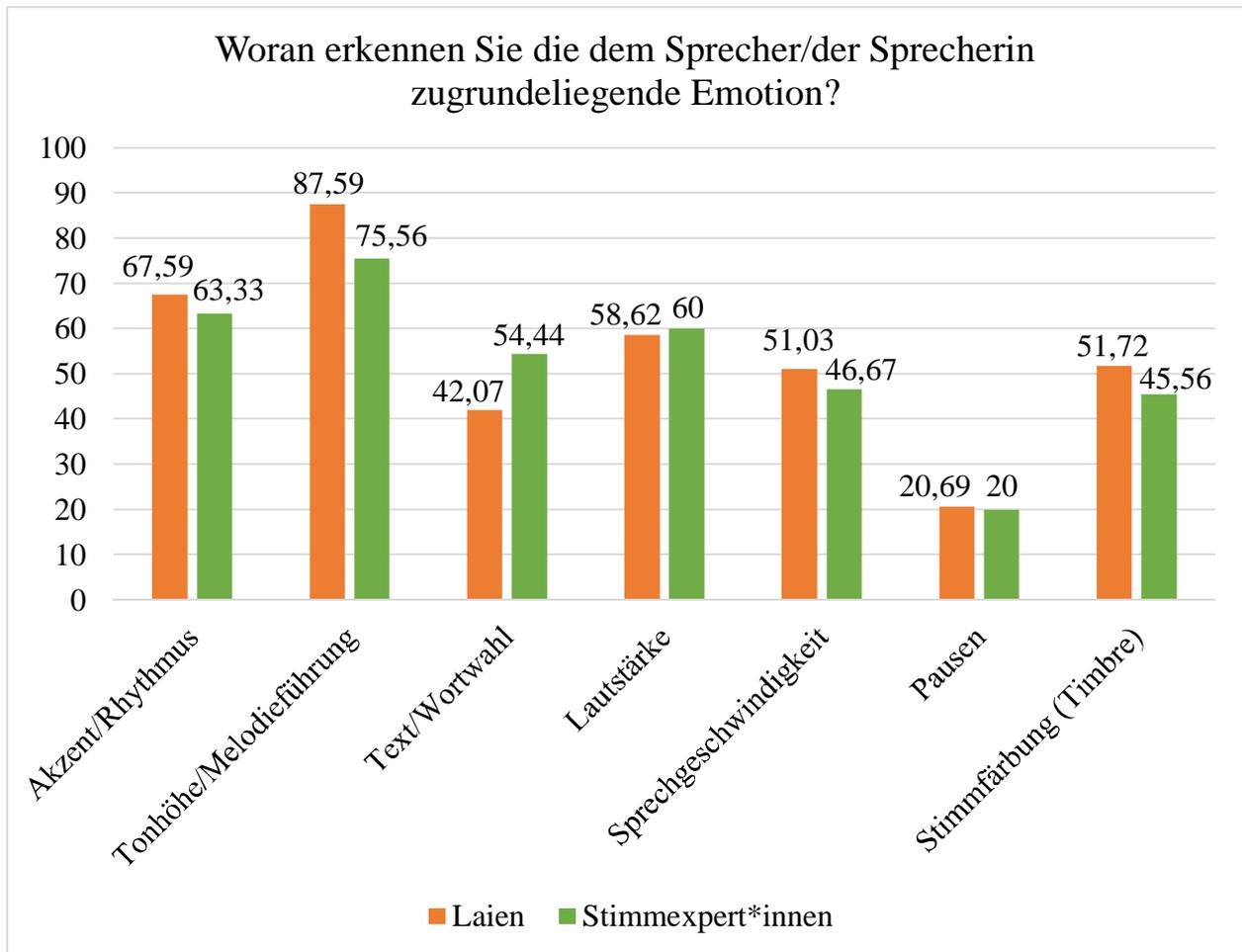
Die Zugehörigkeit zu einer bestimmten Sprach- bzw. Kulturgruppe wirkt als distinktives Merkmal, v. a. in Bezug auf die Merkmale Lautstärke, Text/Wortwahl, Sprechgeschwindigkeit und Pausen, welche (mit Ausnahme der Pausen, welche für die italienische Gruppe wichtiger war) von den deutschsprachigen Proband*innen prozentuell häufiger genannt wurden. Besonders auffallend ist der Unterschied beim Merkmal Lautstärke, das für 65,66 Prozent der deutschsprachigen, aber nur für 44,29 Prozent der italienischsprachigen Gruppe eine Rolle spielt. Die Merkmale Tonhöhe/Melodieführung und Stimmfärbung (Timbre) hingegen werden von beiden Gruppen fast gleich oft genannt, auch bei der Kategorie Akzent/Rhythmus ist der Unterschied eher gering. Die Tonhöhe/Melodieführung spielt eindeutig die größte Rolle bei der Emotionserkennung, gefolgt von den Merkmalen Akzent/Rhythmus, Lautstärke (für die dt. Gruppe),

Stimmfärbung (ital. Gruppe), Sprechgeschwindigkeit, Text/Wortwahl und zuletzt Pausen, welchen keine bedeutende Rolle zukommt. Die Reihenfolge der Parameter, die bei der Emotionserkennung in den beiden Gruppen relevant sind, wird hier in absteigender Häufigkeit angeführt:

Deutschsprachige Proband*innen: Tonhöhe/Melodieführung, Akzent/Rhythmus, Lautstärke, Sprechgeschwindigkeit, Text/Wortwahl, Stimmfärbung (Timbre), Pausen.

Italienischsprachige Proband*innen: Tonhöhe/Melodieführung, Akzent/Rhythmus, Stimmfärbung (Timbre), Lautstärke, Sprechgeschwindigkeit, Text/Wortwahl, Pausen.

3.3.1.3 Laien und Stimmexpert*innen

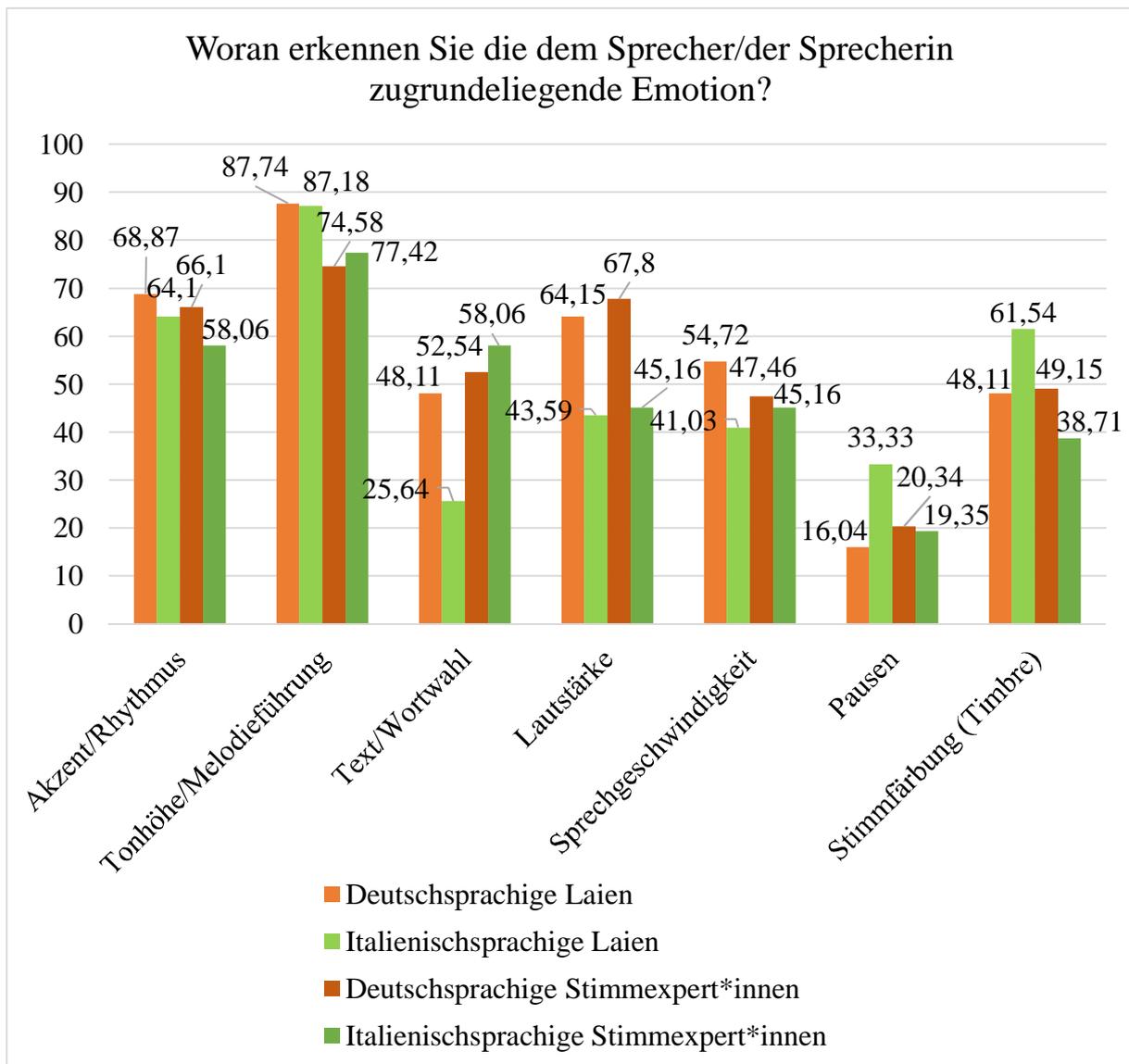


Bei der Unterscheidung zwischen Laien und Stimmexpert*innen insgesamt, also unabhängig von anderen Variablen, lassen sich keine so großen Unterschiede feststellen. Nur bei zwei Parametern, nämlich Tonhöhe/Melodieführung und Text/Wortwahl sind relevante Diskrepanzen zu verzeichnen, wobei Ersteres für die Laien als wichtiger befunden wird und Zweiteres von den Stimmexpert*innen. Die übrigen prosodischen Merkmale werden recht homogen bewertet mit leichten Schwankungen bei der Stimmfärbung, Sprechgeschwindigkeit und Akzent/Rhythmus, welche von den Laien je wichtiger eingeschätzt werden als von den Expert*innen. Bis auf die Kategorie Text/Wortwahl (und Lautstärke, hier ist der Unterschied jedoch vernachlässigbar) werden die prosodischen Merkmale tendenziell von den Laien als relevanter eingestuft als von den Stimmexpert*innen. Hier die Reihenfolge der prosodischen Merkmale mit abnehmender Relevanz für die jeweilige Gruppe:

Laien: Tonhöhe/Melodieführung, Akzent/Rhythmus, Lautstärke, Stimmfärbung (Timbre), Sprechgeschwindigkeit, Text/Wortwahl, Pausen;

Expert*innen: Tonhöhe/Melodieführung, Akzent/Rhythmus, Lautstärke, Text/Wortwahl, Sprechgeschwindigkeit, Stimmfärbung (Timbre), Pausen;

3.3.1.4 Deutsch- und italienischsprachige Laien und Stimmexpert*innen



Um herauszufinden, wie sich die Parameter Sprache/Kultur und Beruf (Laie vs. Stimmexpert*in) gegenseitig beeinflussen und in Bezug auf die einzelnen prosodischen Merkmale auswirken, werden die entsprechenden Gruppen zueinander in Beziehung gesetzt. Dabei lässt sich insgesamt Folgendes ablesen – bei manchen prosodischen Merkmalen ist die Sprache/Kultur (v. a. bei Lautstärke), bei anderen die Berufszugehörigkeit (v. a. bei Tonhöhe/Melodieführung) das entscheidende Kriterium und bei manchen führt der Beruf zu Differenzen innerhalb der Sprachgruppe, v. a. bei der italienischsprachigen (v. a. bei Text/Wortwahl).

Akzent/Rhythmus: Dieses Merkmal ist insgesamt für die deutschsprachigen Proband*innen etwas relevanter als für die italienischsprachigen und für die Laien etwas wichtiger als für die Stimmexpert*innen. Die Abstufungen verlaufen proportional. Dieses Merkmal wird jedoch von allen Proband*innen homogen als sehr wichtig für die Emotionserkennung angegeben.

Tonhöhe/Melodieführung: Hier liegt ein signifikanter Unterschied zwischen Laien und Stimmexpert*innen vor, der sich ohne Unterscheidung über beide Sprach-/Kulturgruppen zieht. So ist dieses Merkmal für die Laien (dt. u. it., circa 87 Prozent) wichtiger als für die Expert*innen (dt. u. it., circa 76 Prozent). Deutlich hervor geht jedenfalls die hohe Relevanz dieses Merkmals für die Emotionserkennung, es ist einstimmig das wichtigste für alle vier Gruppen (dt., it., Laien, Sprachexpert*innen).

Text/Wortwahl: Dieser Parameter ist innerhalb der deutschsprachigen Gruppe unabhängig vom Beruf für circa 50 Prozent der Befragten relevant. Innerhalb der italienischsprachigen Gruppe gibt es jedoch einen enormen Unterschied in Bezug auf den Beruf, wobei interessanterweise nur 25,64 Prozent der Laien, aber 58,06 Prozent der Expert*innen den Text bzw. die Wortwahl für relevant halten.

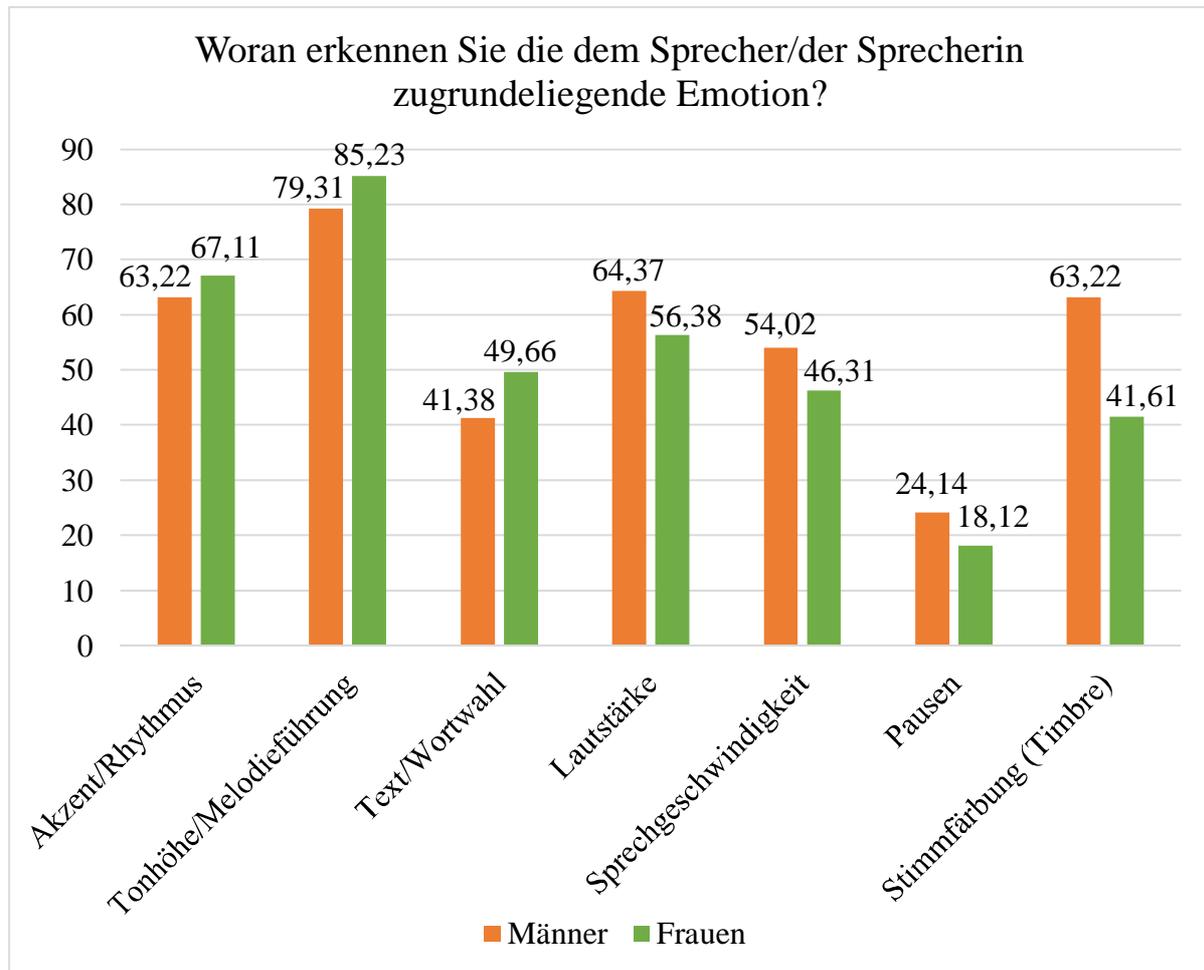
Lautstärke: Bei diesem prosodischen Merkmal gibt es einen sehr deutlichen Unterschied in Bezug auf die Sprache/Kultur, wobei ca. 66 Prozent der deutschsprachigen und nur ca. 44 Prozent der italienischsprachigen Gruppe es für relevant halten. Berufsbedingte Differenzen sind hier vernachlässigbar.

Sprechgeschwindigkeit: Hier zeigt sich wieder ein sprachlich/kulturell bedingter Unterschied zwischen den Gruppen, wobei die deutschsprachige dieses Merkmal öfters nennt als die italienischsprachige. Innerhalb der deutschsprachigen Gruppe lässt sich außerdem eine leichte Differenz zwischen Laien (häufigere Nennung) und Expert*innen erkennen.

Pausen: Insgesamt spielt dieses Merkmal keine zentrale Rolle bei der Emotionserkennung, unabhängig von der Gruppe. Nur ca. 18 Prozent der deutschsprachigen Proband*innen nennen dieses Merkmal, unabhängig vom Beruf. Kaum mehr sind es bei den italienischsprachigen Expert*innen. Aber die italienischsprachigen Laien halten dieses Parameter mit 33,33 Prozent für wesentlich wichtiger, hier ist also eine deutliche berufsbedingte Differenzierung zu beobachten.

Stimmfärbung (Timbre): Die Nennungen der deutschsprachigen Proband*innen sind unabhängig vom Beruf homogen und liegen zwischen 48-49 Prozent. Bei den italienischsprachigen Expert*innen liegt die Zahl jedoch nur bei 38,71 Prozent, während hingegen 61,54 Prozent der italienischsprachigen Laien dieses Merkmal angeben. Innerhalb der italienischsprachigen Gruppe klafft die Bewertung der verschiedenen Berufsgruppen also deutlich auseinander.

3.1.1.5 Männer und Frauen



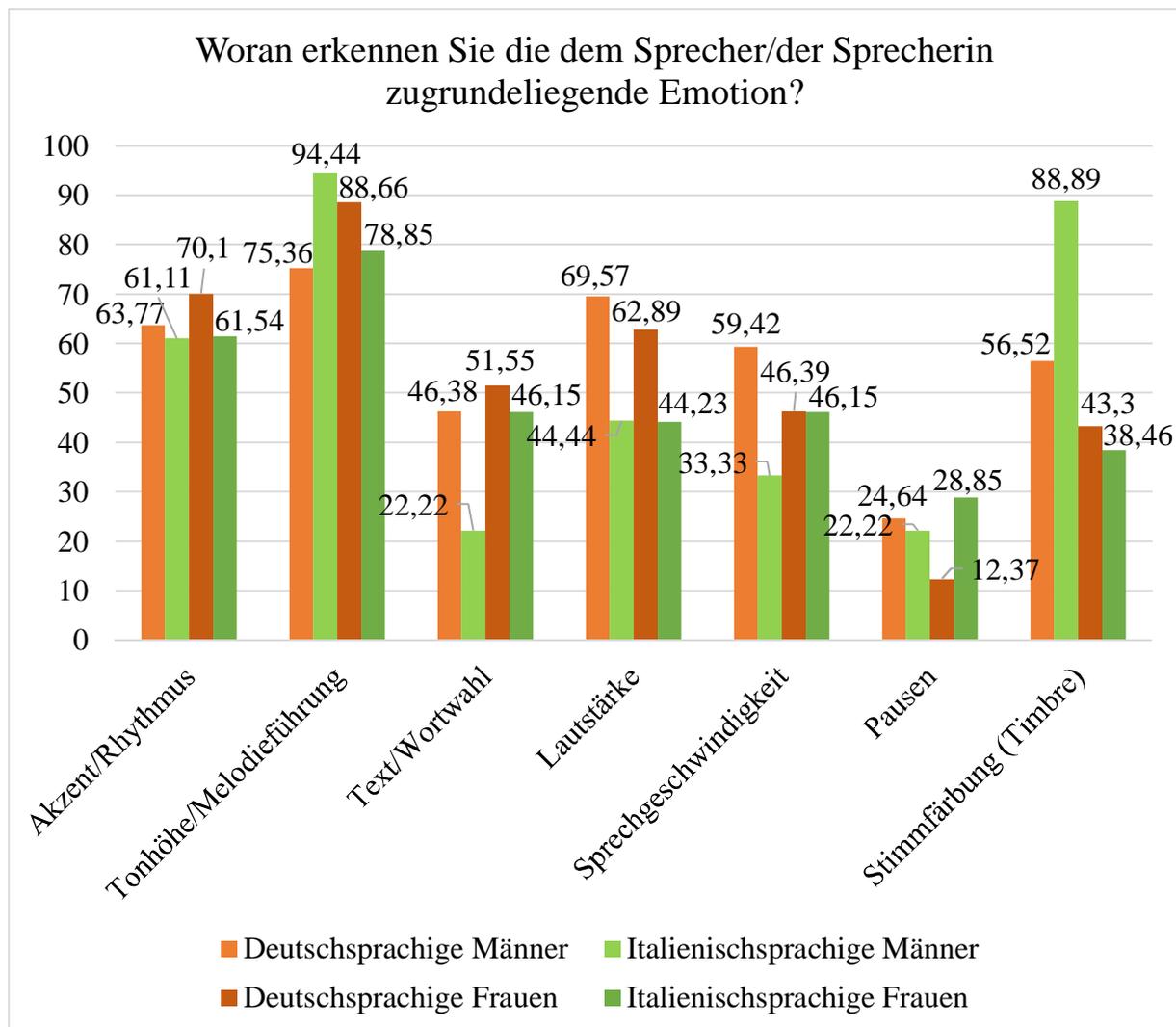
Was die Unterscheidung in Bezug auf das Geschlecht der Proband*innen anbelangt, so lassen sich bei allen Merkmalen leichte Unterschiede in der Häufigkeit ihrer Nennung ablesen, besonders ins Auge sticht aber der Unterschied bei der Stimmfärbung (Timbre). Dieses Merkmal halten 63,22 Prozent der Männer, aber nur 41,61 Prozent der Frauen für relevant bei der Emotionserkennung. Außer bei diesem Merkmal sind die Angaben relativ vergleichbar, leichte Unterschiede mit abnehmender Größe gibt es bei den Kategorien Text/Wortwahl (von Frauen öfters genannt als von Männern), Lautstärke, Sprechgeschwindigkeit (beide von Männern öfters angegeben als von Frauen), Tonhöhe/Melodieführung (von Frauen öfters genannt) und Pausen (von Männern öfters genannt). Hier nun die Reihenfolge der Nennungen für jede Gruppe, vom wichtigsten bis zum unwichtigsten Kriterium zur Emotionserkennung:

Männer: Tonhöhe/Melodieführung, Lautstärke, Stimmfärbung (Timbre) und Akzent/Rhythmus (beide 63,22 Prozent), Sprechgeschwindigkeit, Text/Wortwahl, Pausen;

Frauen: Tonhöhe/Melodieführung, Akzent/Rhythmus, Lautstärke, Text/Wortwahl, Sprechgeschwindigkeit, Stimmfärbung (Timbre), Pausen;

Einheitlich festgestellt werden kann auch hier wieder die Relevanz der Tonhöhe/Melodieführung an erster Stelle, gefolgt von Akzent/Rhythmus und Lautstärke. Das Schlusslicht bildet auch hier wieder das Merkmal Pausen.

3.1.1.6 Deutsch- und italienischsprachige Männer und Frauen



Die Ergebnisse bei der Gegenüberstellung der beiden Kategorien Geschlecht und Sprache/Kultur zeigen die bereits weiter oben herausgearbeiteten sprachlich/kulturellen Unterschiede (v. a. bei Lautstärke) auf, aber auch große geschlechterbedingte Unterschiede (v. a. bei Stimmfärbung und Text/Wortwahl der italienischsprachigen Proband*innen) innerhalb der jeweiligen Sprachgruppe.

Akzent/Rhythmus: Dieses Merkmal wird von allen Gruppen vergleichbar homogen und häufig genannt, wobei die deutschsprachigen Frauen es tendenziell öfters nannten als die deutschsprachigen Männer. Ansonsten sind keine relevanten Unterschiede in Bezug auf Geschlecht oder Sprache/Kultur zu verzeichnen.

Tonhöhe/Melodieführung: Dieses Merkmal zeichnet sich als das wichtigste aller Proband*innen für die Emotionserkennung ab, und zwar prinzipiell unabhängig aller Variablen. Sind die Nennungen der deutschsprachigen und italienischsprachigen Personen insgesamt fast gleich

häufig, so ergeben sich doch innerhalb der jeweiligen Sprach-/Kulturgruppe erhebliche geschlechterbedingte Unterschiede. So gibt es bei den deutschsprachigen Proband*innen mehr Nennungen bei den Frauen und bei den italienischsprachigen Proband*innen umgekehrt deutlich mehr Nennungen bei den Männern.

Text/Wortwahl: Hier zeigt sich das bereits bekannte Bild, dass dieses Merkmal insgesamt öfters von den deutschsprachigen Proband*innen und von Frauen angegeben wurde, allerdings zeigt sich eine relevante Diskrepanz zwischen den italienischsprachigen Männern (22 Prozent) und den italienischsprachigen Frauen (46,15 Prozent), welche das Merkmal viel öfters nennen. Der Unterschied zwischen deutschsprachigen Männern und Frauen ist hingegen nicht relevant.

Lautstärke: Dieses Merkmal zeigt eine klare sprach-/kulturbedingte Differenz auf, wohingegen die geschlechterbezogenen Unterschiede minimiert werden können bzw. nur für die deutschsprachige Gruppe gelten, in der mehr Männer als Frauen dieses Merkmal für relevant halten. Der große Unterschied liegt darin, dass zw. 66-67 Prozent der deutschsprachigen Proband*innen dieses Merkmal angeben, aber nur ca. 44 Prozent der italienischsprachigen. Innerhalb der italienischsprachigen Gruppe ist die Verteilung der Angaben von Männern und Frauen identisch.

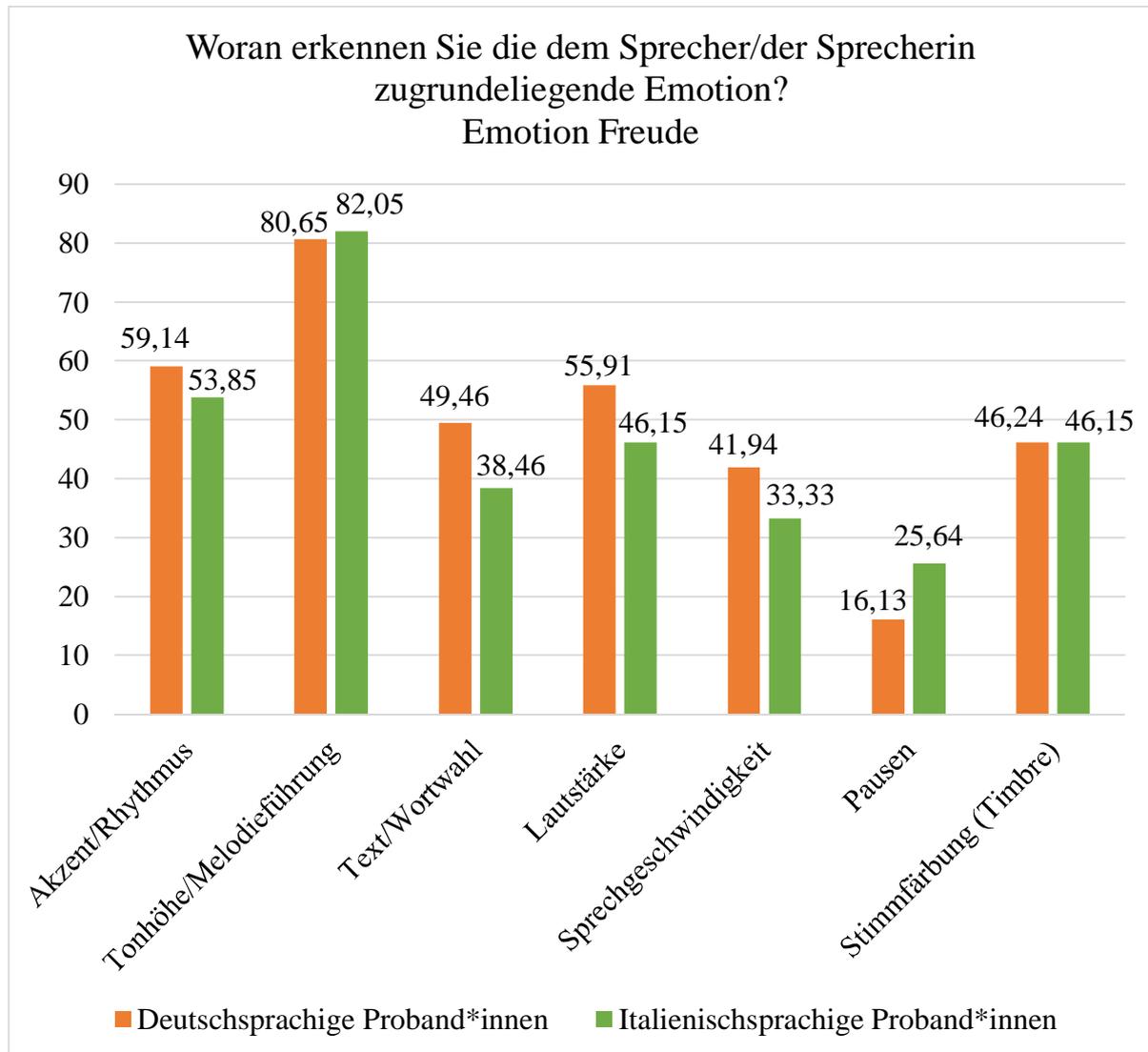
Sprechgeschwindigkeit: Bei diesem Merkmal gibt es innerhalb der beiden Sprachgruppen große Unterschiede, und zwar proportional umgekehrt – in der deutschsprachigen Gruppe gibt es von Seiten der Männer viel mehr Nennungen, in der italienischsprachigen Gruppe von den Frauen. Die Frauen (it. u. dt.) nennen dieses Merkmal unabhängig der Sprache/Kultur gleich oft. Bei den Männern ist der Unterschied enorm mit 59,42 Prozent bei den deutschsprachigen und 33,33 Prozent bei den italienischsprachigen Probanden.

Pausen: Auch bei diesem Merkmal vermischen sich die diversen Parameter, wobei hier die Männer (it. u. dt.) einen sehr ähnlichen Wert aufweisen, während hingegen die Zahlen bei den deutschsprachigen Frauen (12,37 Prozent) und italienischsprachigen Frauen (28,85 Prozent) deutlich auseinanderklaffen.

Stimmfärbung (Timbre): Insgesamt halten deutlich mehr Männer als Frauen dieses Merkmal für relevant in Bezug auf die Emotionsbestimmung. Allerdings muss hier näher nach deren sprachlich-kulturellem Hintergrund unterschieden werden. Die deutschsprachigen Männer nennen dieses Parameter mit 56,52 Prozent deutlich öfters als die deutschsprachigen Frauen (43,3 Prozent) und die italienischsprachigen Frauen (38,46 Prozent). Für die italienischsprachigen Männer ist dieses Merkmal jedoch ungleich relevanter. Es wird von 88,89 Prozent der italienischsprachigen Männer genannt und zählt damit nach der Tonhöhe/Melodieführung zu deren zweitwichtigstem Merkmal zur Emotionserkennung.

3.3.1.7 Ergebnisse in Bezug auf die freudigen Sprechweisen

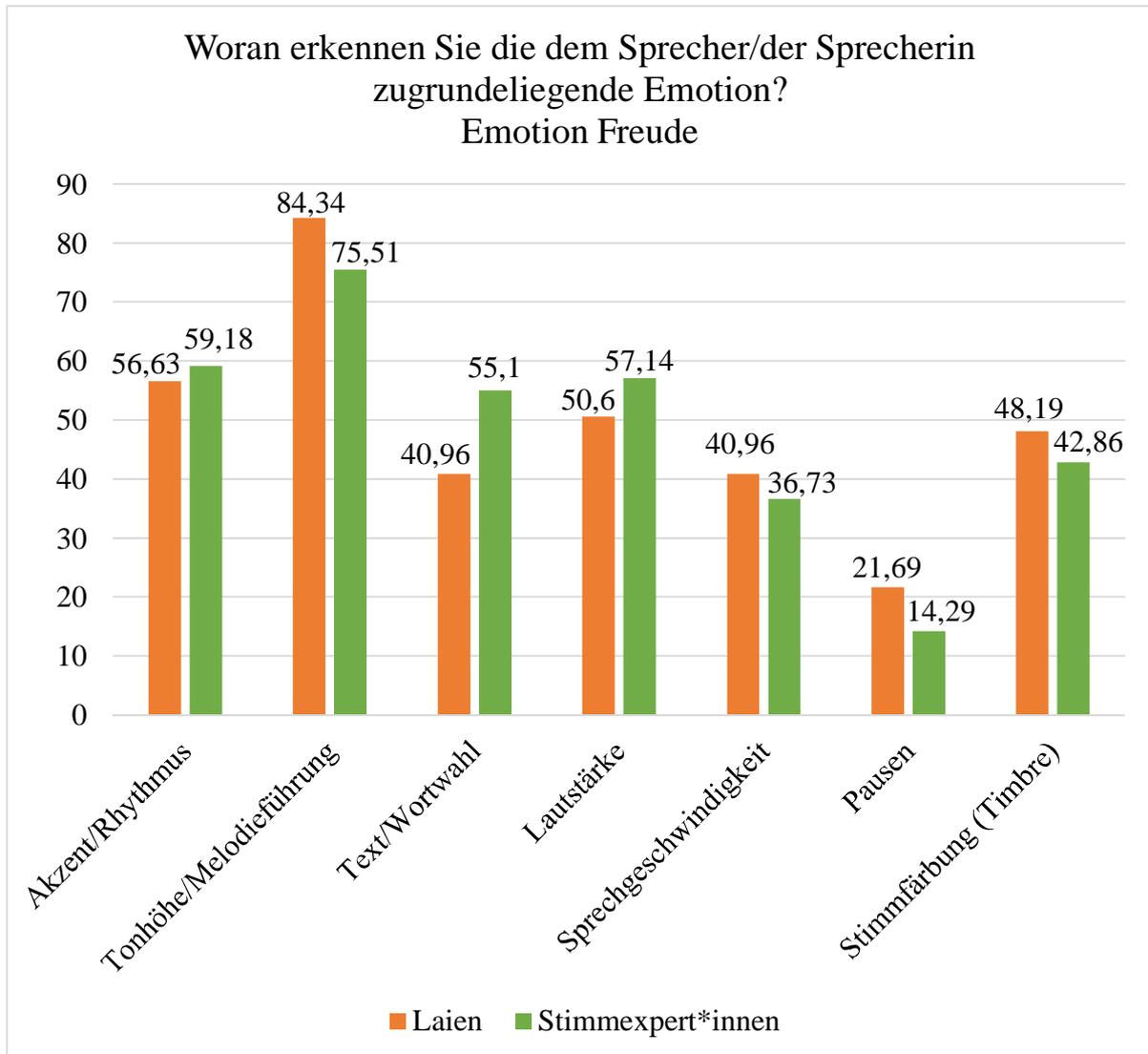
Deutsch- und italienischsprachige Proband*innen



Bei der Perzeption der freudigen Sprechweisen spielen manche prosodischen Merkmale eine wichtigere Rolle als andere, wobei hier ein klarer sprachlich/kultureller Unterschied erkennbar ist. Für beide Gruppen ist die Tonhöhe/Melodieführung mit ca. 81 Prozent der wichtigste Parameter. Auf dem zweiten Platz folgt der Akzent/Rhythmus, der tendenziell für die deutschsprachigen Proband*innen etwas relevanter ist. Die Lautstärke ist für beide Gruppen das drittwichtigste Element zur Emotionserkennung, allerdings spielt es für die deutschsprachigen Proband*innen (55,91 Prozent) eine größere Rolle als für die italienischsprachigen (46,15 Prozent). Letztere halten die Stimmfärbung (Timbre) außerdem für gleich wichtig wie die Lautstärke.

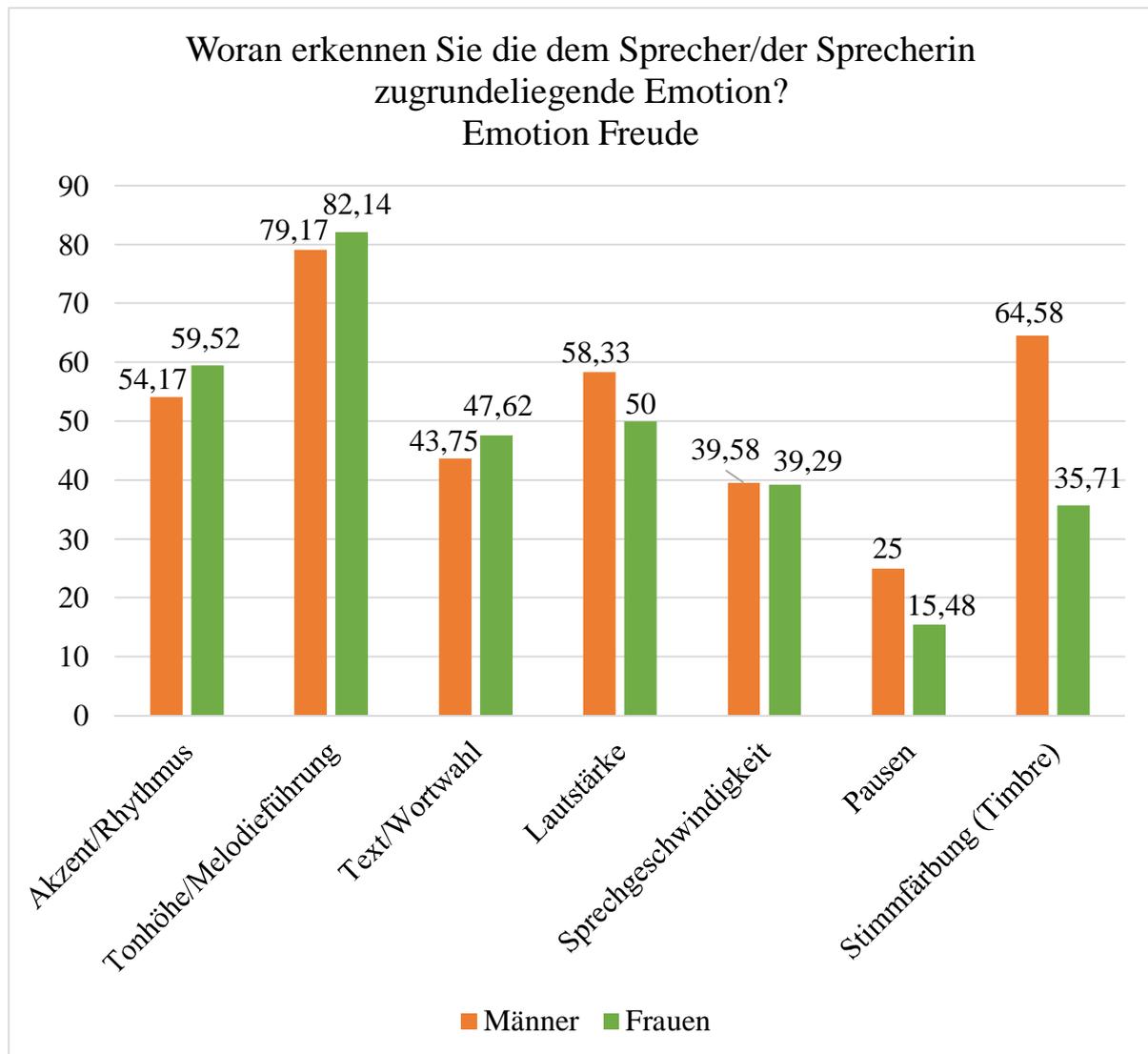
Gleich viele deutschsprachige wie italienischsprachige Proband*innen bedienen sich des Timbres zur Emotionserkennung (beide ca. 46 Prozent), nur landet bei den deutschsprachigen Proband*innen Text/Wortwahl im Ranking knapp weiter oben (49,46 Prozent). Als nächstes Merkmal wird die Sprechgeschwindigkeit genannt, welche wiederum eine relevantere Rolle spielt bei den deutschsprachigen Proband*innen. Das Schlusslicht bilden bei beiden Gruppen die Pausen, sie sind allerdings für die italienischsprachigen Proband*innen deutlich wichtiger.

Laien und Stimmexpert*innen



Auch bei der Unterscheidung der Bewertungen der Laien und Stimmexpert*innen zeigen sich z. T. deutliche Differenzen. An erster Stelle steht auch hier zwar die Tonhöhe/Melodieführung, allerdings wird sie von den Laien als relevanter betrachtet. Dies gilt, wenngleich in geringerer Ausprägung, auch für die Pausen und für die Stimmfärbung (Timbre). Die Stimmexpert*innen hingegen halten v. a. die Kategorie Text/Wortwahl für sehr viel relevanter als die Laien, auch die Lautstärke wird öfters von ihnen genannt. Ungefähr gleich häufig werden hingegen Akzent/Rhythmus und Sprechgeschwindigkeit angegeben.

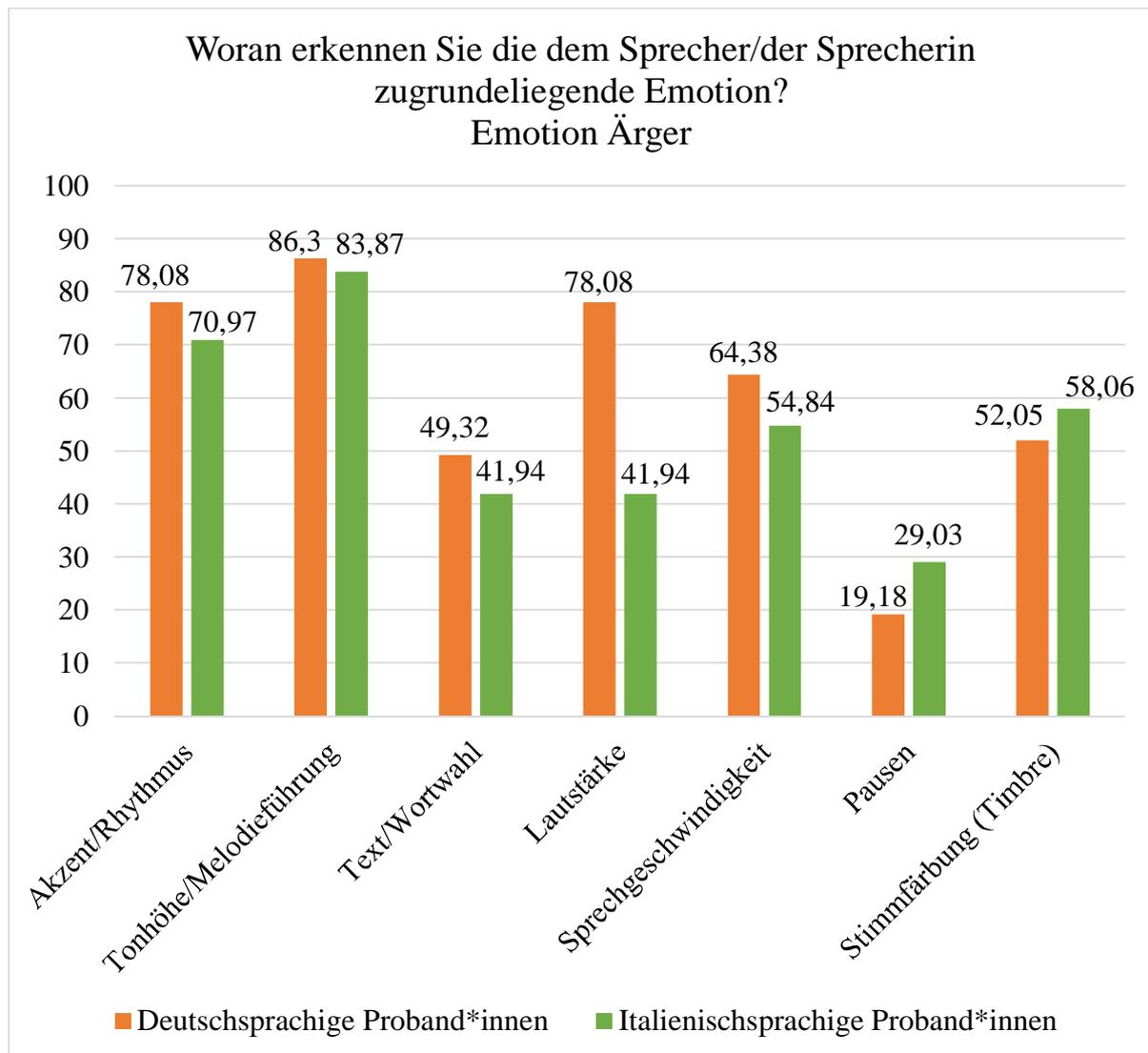
Männer und Frauen



Bei der Erkennung der freudigen Sprechweisen spielen die prosodischen Merkmale bei den Männern und Frauen eine je unterschiedlich gewichtete Rolle. Beinahe identisch sind die Nennungen in Bezug auf die Tonhöhe/Melodieführung, welches das wichtigste Merkmal darstellt. Der Akzent/Rhythmus landet bei den Frauen auf zweiter Stelle und wird von ihnen häufiger angegeben als von den Männern. Auf zweiter Stelle bei den Männern steht hingegen die Stimmfärbung (Timbre) mit 64,58 Prozent, während dieses Merkmal bei den Frauen (35,71 Prozent) eine untergeordnete Rolle spielt. Auf dem dritten Posten befindet sich bei beiden Gruppen die Lautstärke, die von Männern jedoch häufiger genannt wird. Text/Wortwahl und die Sprechgeschwindigkeit hingegen werden von beiden Gruppen vergleichbar häufig angegeben. Die Pausen spielen für die Männer eine deutlich größere Rolle als für die Frauen.

3.3.1.8 Ergebnisse in Bezug auf die Sprechweisen zu Ärger

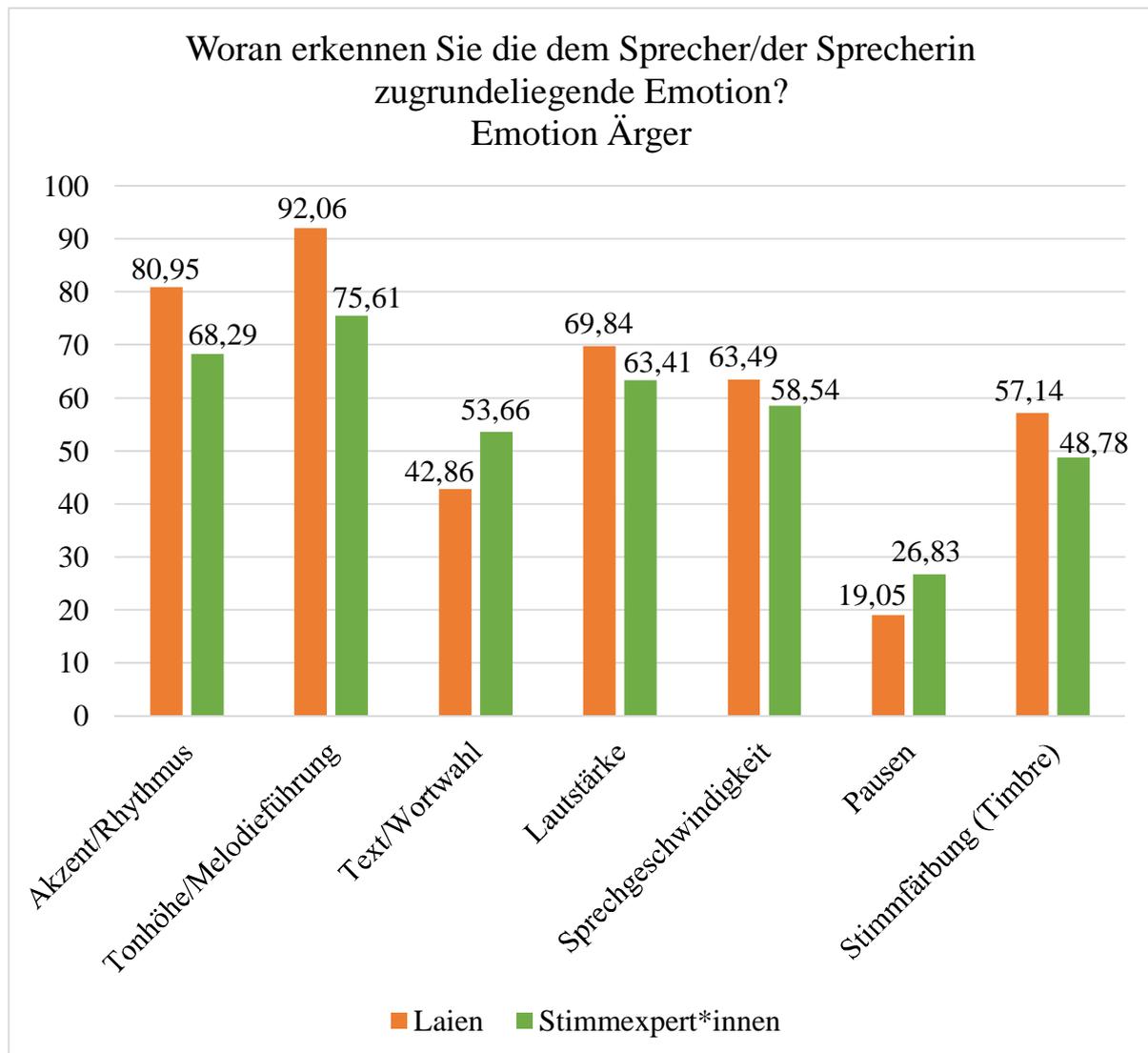
Deutsch- und italienischsprachige Proband*innen



Bei der Erkennung der ärgerlichen Sprechweisen spielt eindeutig die Tonhöhe/Melodieführung für beide Gruppen die wichtigste Rolle. In allen anderen Parametern gibt es deutlich differierende Nennungen, v. a. bei der Perzeption der Lautstärke. Für 78,08 Prozent der deutschsprachigen Proband*innen spielt die Lautstärke nach der Tonhöhe/Melodieführung eine äußerste relevante Rolle, dies gilt aber nur für 41,94 Prozent der italienischsprachigen Proband*innen. Dieser Unterschied ist enorm, es gibt also doppelt so viele Nennungen auf deutschsprachiger Seite. Auch alle anderen Parameter spielen eine größere Rolle für die deutschsprachigen Proband*innen, mit Ausnahme der Pausen, welche 10 Prozent häufiger von den italienischsprachigen Personen genannt wird. Im Vergleich mit den Ergebnissen zur Freude zeigt sich hier ein

sehr großer Unterschied bei der Lautstärke, welche von den deutschsprachigen Proband*innen hier über 20 Prozent häufiger angegeben wird (Freude: 55,91 Prozent, Ärger: 78,08 Prozent). Die Zahl der Nennungen der italienischsprachigen Gruppe beläuft sich bei Freude und Ärger auf ca. 44 Prozent.

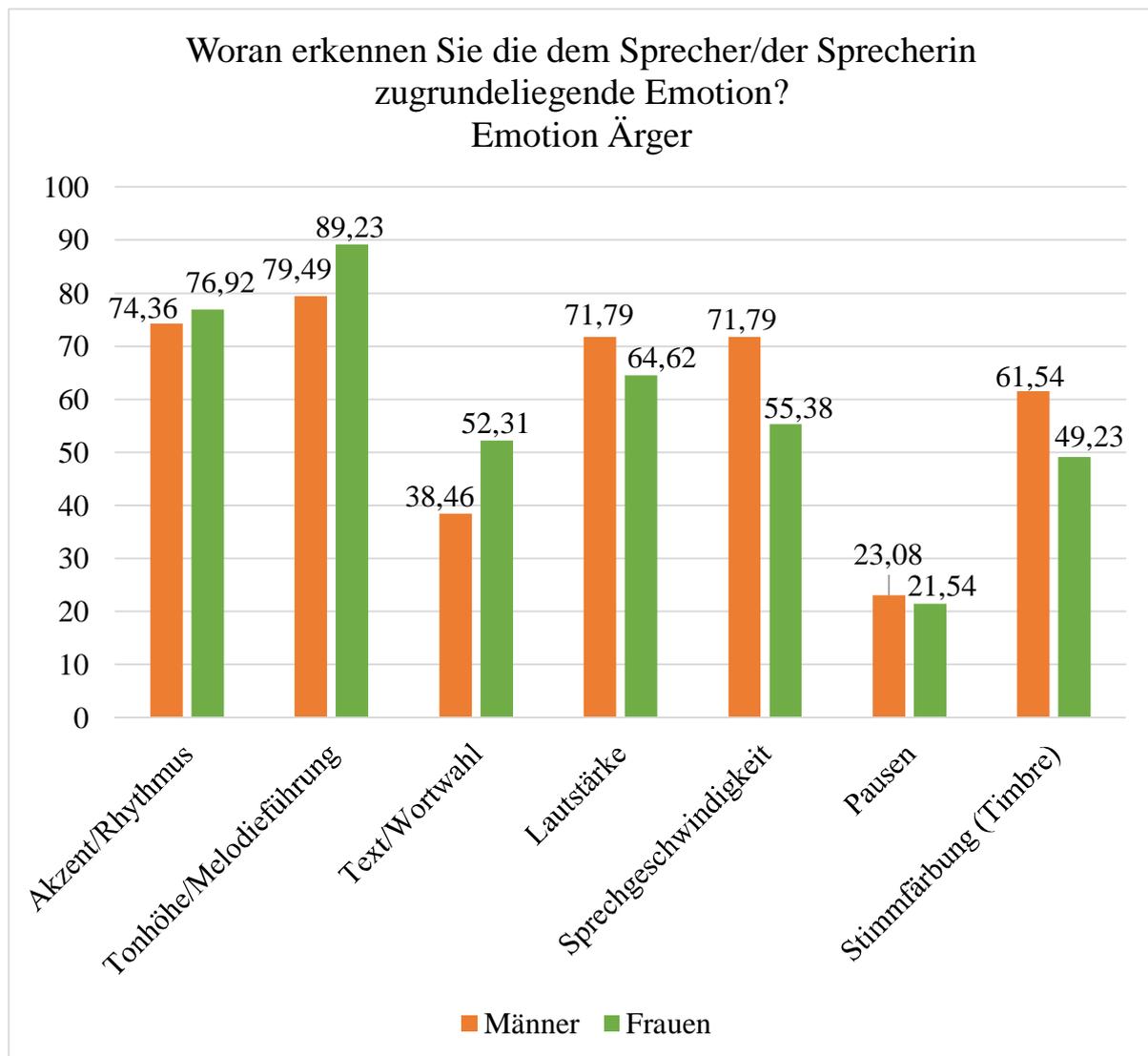
Laien und Stimmexpert*innen



Das Diagramm zeigt, welche prosodischen Merkmale für das Erkennen der ärgerlichen Sprechweisen für Laien und Stimmexpert*innen besonders relevant sind. Dabei zeichnet sich ein ähnliches Bild wie bei den Äußerungen zur Freude ab. Das Merkmal Tonhöhe/Melodieführung spielt eine wesentlich größere Rolle für die Laien, selbiges gilt für Akzent/Rhythmus, Stimmfärbung und, in geringerem Ausmaß, für die Lautstärke. Die Parameter Text/Wortwahl und die Pausen hingegen werden von den Stimmexpert*innen häufiger genannt. Waren die Pausen bei

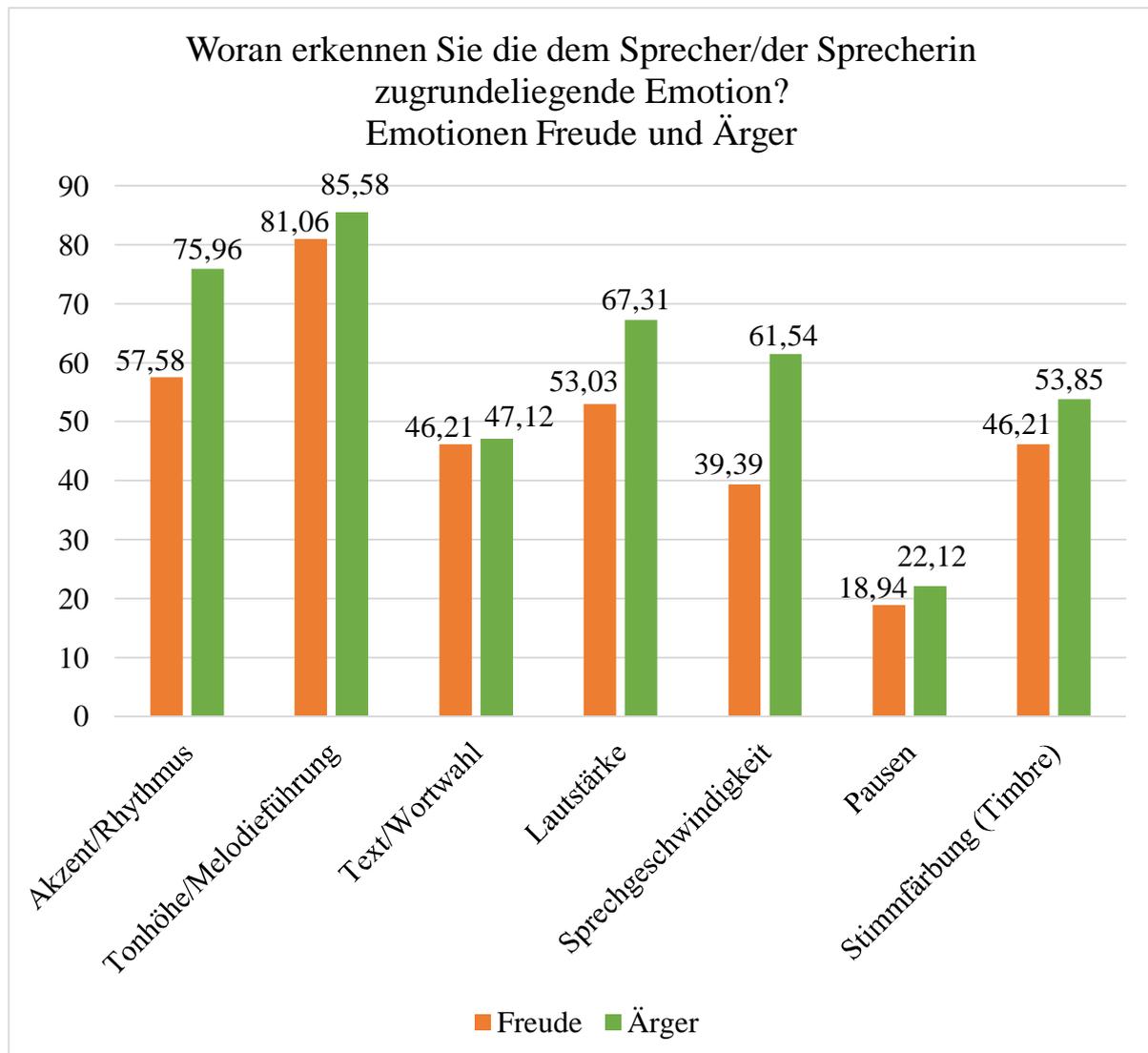
den Äußerungen zur Freude für die Laien relevanter, sind sie es hier für die Expert*innen. Annähernd gleich viele Angaben wurden zur Sprechgeschwindigkeit gemacht, wobei auch hier die Laien knapp überwiegen. Diese Relation entspricht auch dem Ergebnis zu den freudigen Sprechweisen mit dem Unterschied, dass die Sprechgeschwindigkeit bei Ärger um circa 20 Prozent öfters angegeben wird. Auch die Lautstärke spielt eine größere Rolle bei Ärger und bei den Laien – dieses Verhältnis war bei den Sprechweisen zur Freude umgekehrt.

Männer und Frauen



Bei den Ergebnissen zur Frage, anhand welcher prosodischen Parameter die Männer und Frauen die ärgerlichen Sprechweisen erkennen können, zeigt sich ein sehr unterschiedliches Bild zwischen den Geschlechtern. Die Stimmfärbung (Timbre) ist für die Männer wesentlich relevanter als für die Frauen, wenngleich hier die Differenz geringer ausfällt als bei den Sprechweisen zur Freude. Besonders groß, auch im Vergleich zu den Angaben zur Freude (beide Geschlechter ca. 39 Prozent), ist der Unterschied bei der Sprechgeschwindigkeit, welche viel häufiger von den Männern angegeben wird. Auch die Lautstärke, welche insgesamt hier wichtiger ist, spielt für die Männer eine größere Rolle als für die Frauen. Bei den Pausen hebt sich, auch im Unterschied zum Ergebnis zur Freude, der geschlechtsbezogene Unterschied auf. In den Nennungen zu Text/Wortwahl, Tonhöhe/Melodieführung und Akzent/Rhythmus überwiegt hier wie beim Ergebnis zur Freude der Anteil der Frauen.

3.3.1.9 Gegenüberstellung Freude – Ärger



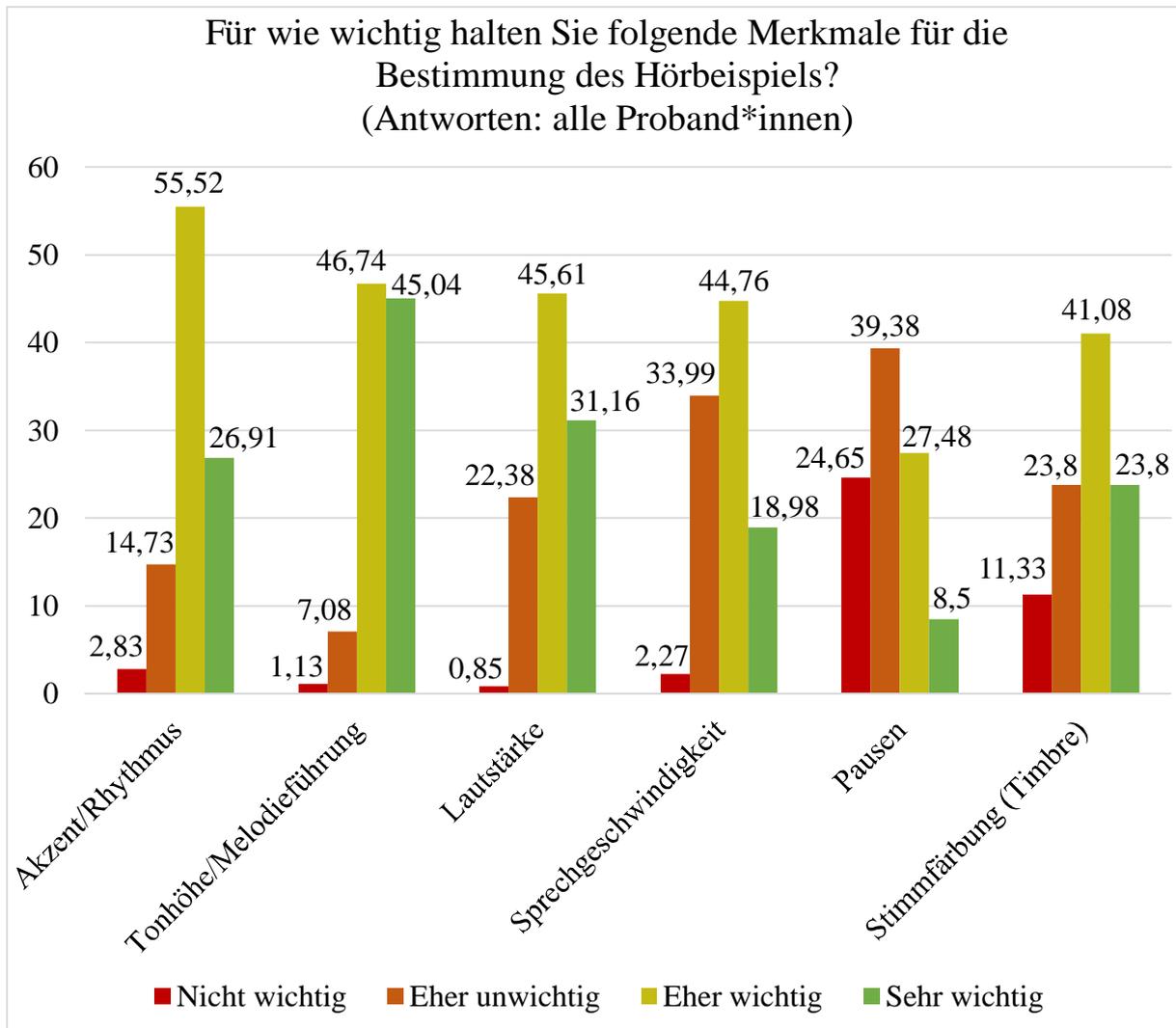
Werden nun die Ergebnisse zu den einzelnen Parametern insgesamt zwischen den freudigen und ärgerlichen Sprechweisen verglichen, so fällt auf, dass die einzelnen prosodischen Merkmale je nach Emotion eine verschieden große Rolle spielen bei der Perzeption und Zuordnung des Gehörten. Besonders groß ist der Unterschied bei der Sprechgeschwindigkeit – ist sie zur Erkennung der freudigen Sprechweisen für nur 39,39 Prozent der Proband*innen relevant, sind es bei den ärgerlichen Sprechweisen 61,54 Prozent. Ähnlich gewichtig ist die Differenz beim Akzent/Rhythmus, welcher um knapp 20 Prozent häufiger genannt wird bei der Erkennung der ärgerlichen Äußerungen. Gleich viele Angaben gibt es hingegen bei den Kategorien Text/Wortwahl (ca. 46,5 Prozent) und Pausen (ca. 20 Prozent), wobei hier z. T. starke sprachlich/kulturelle, geschlechter- und berufsbedingte Schwankungen innerhalb der Proband*innengruppen

vorliegen, wie in den Kapiteln zuvor aufgezeigt werden konnte. Die Lautstärke spielt eine deutlich wichtigere Rolle bei der Erkennung der ärgerlichen Sprechweisen, selbiges gilt in geringerem Ausmaß auch für die Stimmfärbung (Timbre). Dem Merkmal Tonhöhe/Melodieführung kommt in beiden Fällen und für alle Proband*innen in vergleichbarem Ausmaß die größte Bedeutung bei der Emotionsperzeption zu.

3.3.2 Relevanz der prosodischen Merkmale in der Perzeption der emotionalen Sprechweisen

In der zweiten Frage soll nun präzisiert werden, für wie relevant die Proband*innen die einzelnen prosodischen Merkmale im Rahmen der Emotionszuordnung des Gehörten halten. Es geht also nicht mehr nur darum, welche Parameter generell zur Emotionserkennung herangezogen werden, sondern welche Wichtigkeit ihnen dabei zugemessen wird. Auch hier werden zunächst alle Antworten aus den acht Fragebögen zusammengenommen ausgewertet, um ein aussagekräftiges Bild entwerfen zu können. In einem zweiten Schritt werden die Ergebnisse je nach emotionaler Sprechweise in Freude und Ärger getrennt dargestellt und miteinander verglichen. Der Faktor Sprache/Wortwahl wird nicht mehr miteinbezogen, da nur die prosodischen Merkmale untersucht werden sollen.

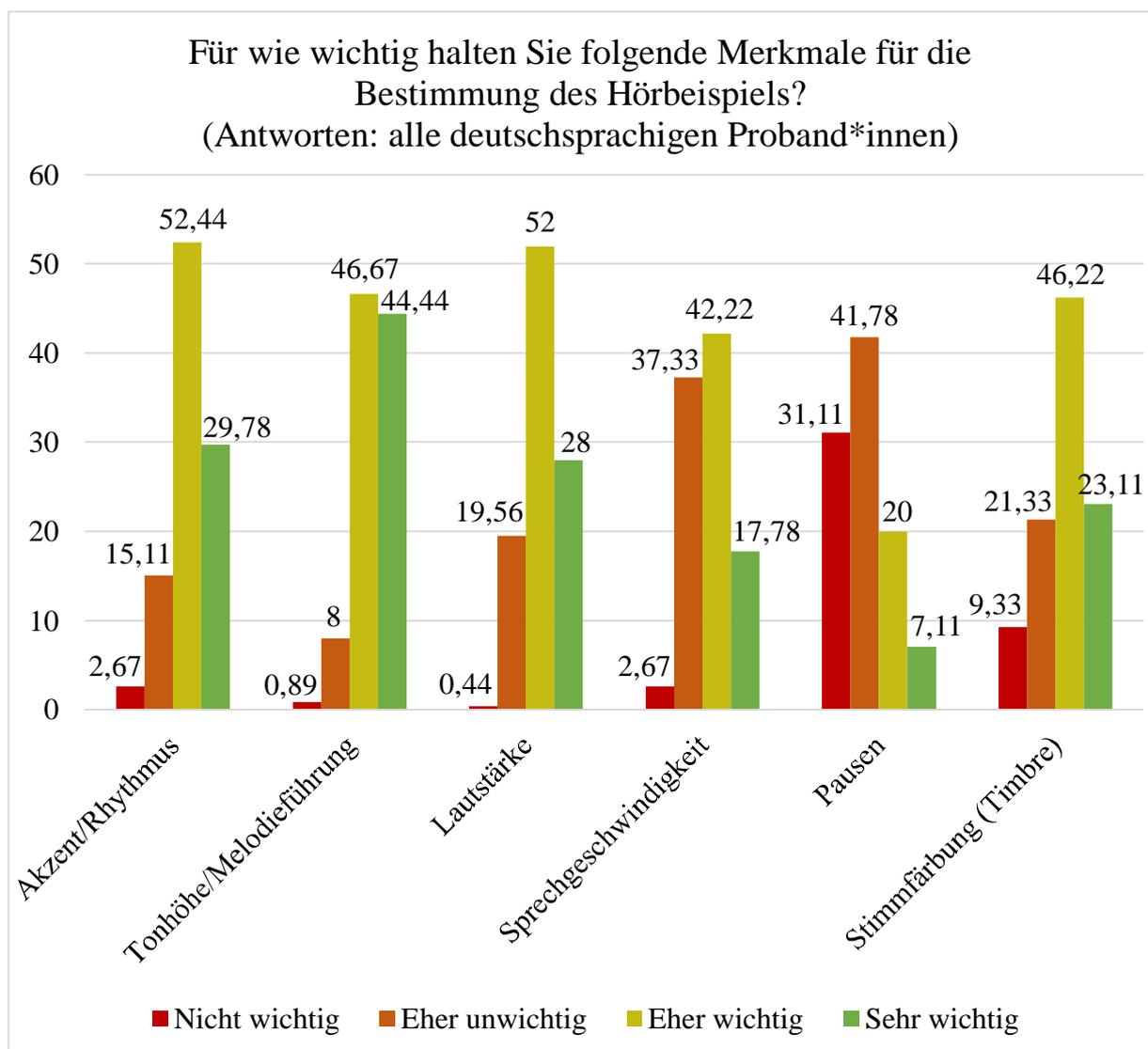
3.3.2.1 Gesamtüberblick aller Antworten



Dieses Diagramm zeigt die Relevanz der einzelnen prosodischen Merkmale für die Emotionserkennung aller Proband*innen auf und zeichnet ein klares Ranking ab, welches in der folgenden Tabelle mit absteigender Wichtigkeit veranschaulicht wird. Dabei beziehen sich die Zahlen auf die prozentualen Angaben in den beiden Kategorien „nicht wichtig/eher unwichtig“ und „eher wichtig/sehr wichtig“. Auf genauere sprach-/kulturbedingte, berufsbezogene oder geschlechterspezifische Differenzierungen wird im Folgenden näher eingegangen.

	Nicht wichtig / Eher unwichtig	Eher wichtig / Sehr wichtig
Tonhöhe/Melodieführung	8,21	91,78
Akzent/Rhythmus	17,56	82,43
Lautstärke	23,23	76,77
Stimmfärbung (Timbre)	35,13	64,88
Sprechgeschwindigkeit	36,26	63,74
Pausen	64,03	35,98

3.3.2.2 Deutschsprachige Proband*innen



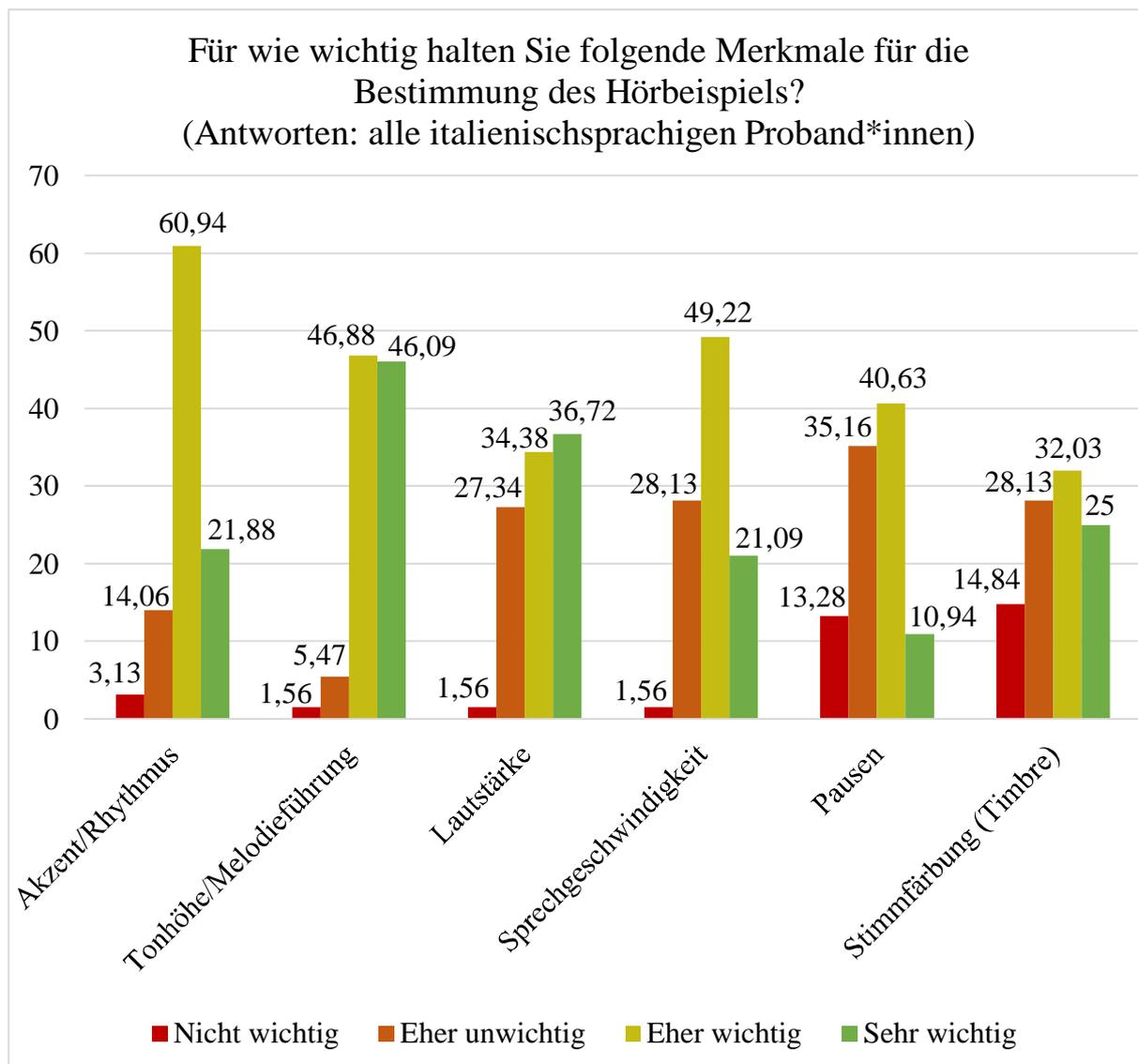
Das obige Diagramm zeigt die Bewertungen aller deutschsprachigen Proband*innen an und zeichnet ein recht klares Bild im Ranking der einzelnen prosodischen Merkmale. Fast alle Per-

sonen halten die Tonhöhe/Melodieführung für wesentlich, gefolgt von Akzent/Rhythmus, Lautstärke, Stimmfärbung (Timbre), Sprechgeschwindigkeit und zuletzt Pausen. Es fällt auf, dass die mittleren Angaben (eher unwichtig bzw. eher wichtig) dominieren, v. a. die Kategorie „eher wichtig“. Nur bei den Pausen spielt die Kategorie „nicht wichtig“ mit 31,11 Prozent eine zentrale Rolle.

In der folgenden Übersicht sind die Reihenfolge der Merkmale mit absteigender Wichtigkeit und die prozentualen Angaben in den zwei Kategorien „nicht wichtig/eher unwichtig“ und „eher wichtig/sehr wichtig“ zusammengefasst, um das Ergebnis zu veranschaulichen.

	Nicht wichtig / Eher unwichtig	Eher wichtig / Sehr wichtig
Tonhöhe/Melodieführung	8,89	91,11
Akzent/Rhythmus	17,78	82,22
Lautstärke	20	80
Stimmfärbung (Timbre)	30,66	69,33
Sprechgeschwindigkeit	40	60
Pausen	72,89	27,11

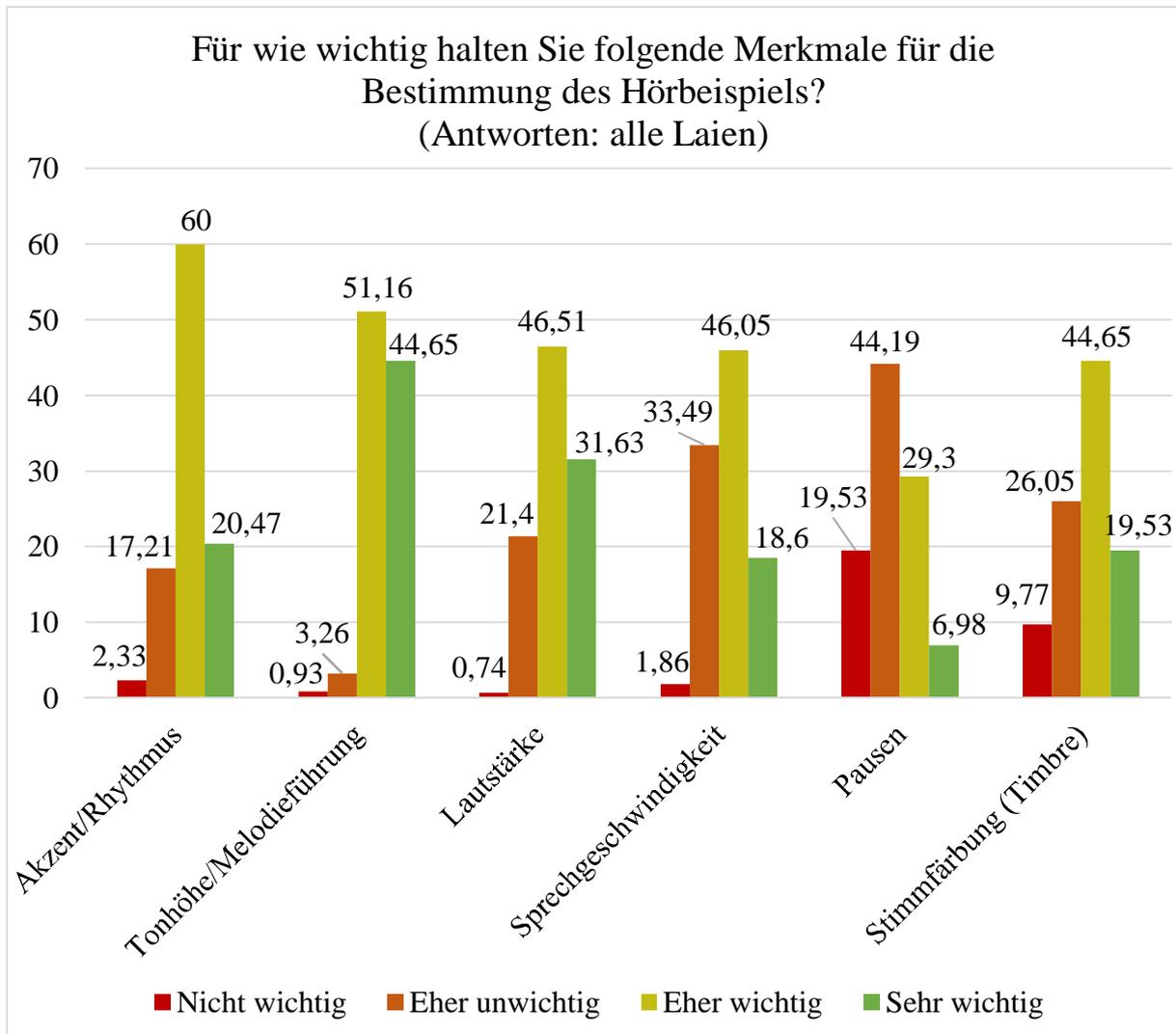
3.3.2.3 Italienischsprachige Proband*innen



Tatsächlich zeichnet sich hier ein recht unterschiedliches Bild im Vergleich zu den deutschsprachigen Proband*innen, nur die Angaben bei den Merkmalen Tonhöhe/Melodieführung und Akzent/Rhythmus sind praktisch identisch. Die Merkmale Lautstärke und Stimmfärbung (Timbre) sind für die deutschsprachigen Proband*innen wesentlich wichtiger als für die italienischsprachigen. Anders verhält es sich bei den Merkmalen Sprechgeschwindigkeit und v. a. Pausen, welche für die italienischsprachigen Proband*innen deutlich relevanter sind. Besonders groß ist der Unterschied bei den Pausen (51,57 Prozent it. vs. 27,11 Prozent dt.). Folgend zeigt die Übersicht das Ranking und die diversen Angaben in Prozent.

	Nicht wichtig / Eher unwichtig	Eher wichtig / Sehr wichtig
Tonhöhe/Melodieführung	7,03	92,97
Akzent/Rhythmus	17,19	82,82
Lautstärke	28,9	71,1
Sprechgeschwindigkeit	29,69	70,31
Stimmfärbung (Timbre)	42,97	57,03
Pausen	48,44	51,57

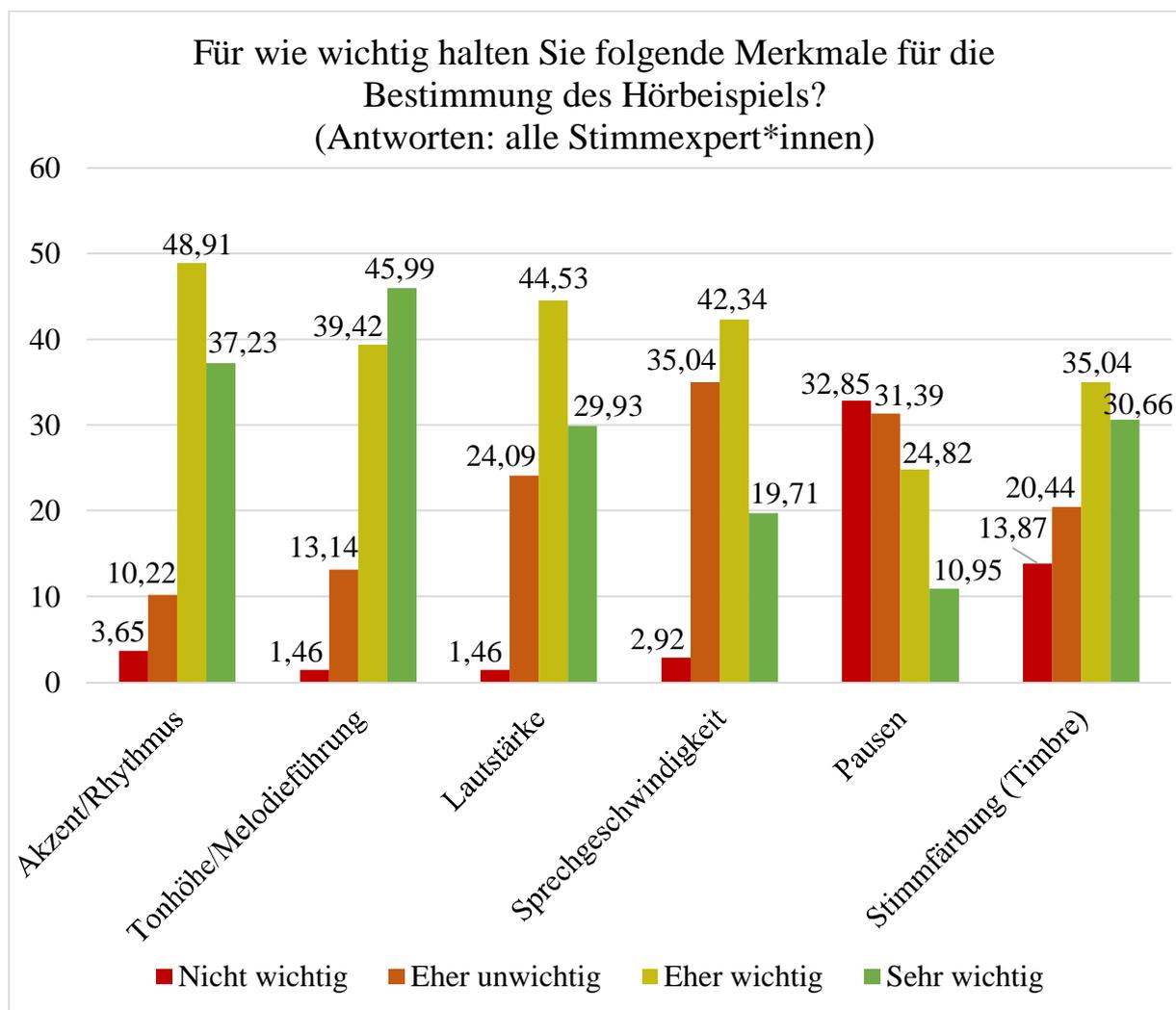
3.3.2.4. Alle Laien



Das obige Diagramm zeigt die Beurteilung der Relevanz der prosodischen Parameter bei der Emotionserkennung durch alle Laien. Die wichtigsten drei Merkmale sind hier auch wieder Tonhöhe/Melodieführung an erster Stelle, gefolgt von Akzent/Rhythmus und Lautstärke. Überwiegend oft erfolgt die Angabe „eher wichtig“, am wenigsten häufig wird die Kategorie „nicht wichtig“ angekreuzt. Das im Vergleich zu den vorigen Diagrammen leicht veränderte Ranking der prosodischen Merkmale und die prozentualen Angaben in den zwei Dimensionen wichtig/unwichtig sind hier zusammengefasst:

	Nicht wichtig / Eher unwichtig	Eher wichtig / Sehr wichtig
Tonhöhe/Melodieführung	4,19	95,81
Akzent/Rhythmus	19,54	80,47
Lautstärke	22,14	78,14
Sprechgeschwindigkeit	35,35	64,65
Stimmfärbung (Timbre)	35,82	64,18
Pausen	63,72	36,28

3.3.2.5 Alle Stimmexpert*innen

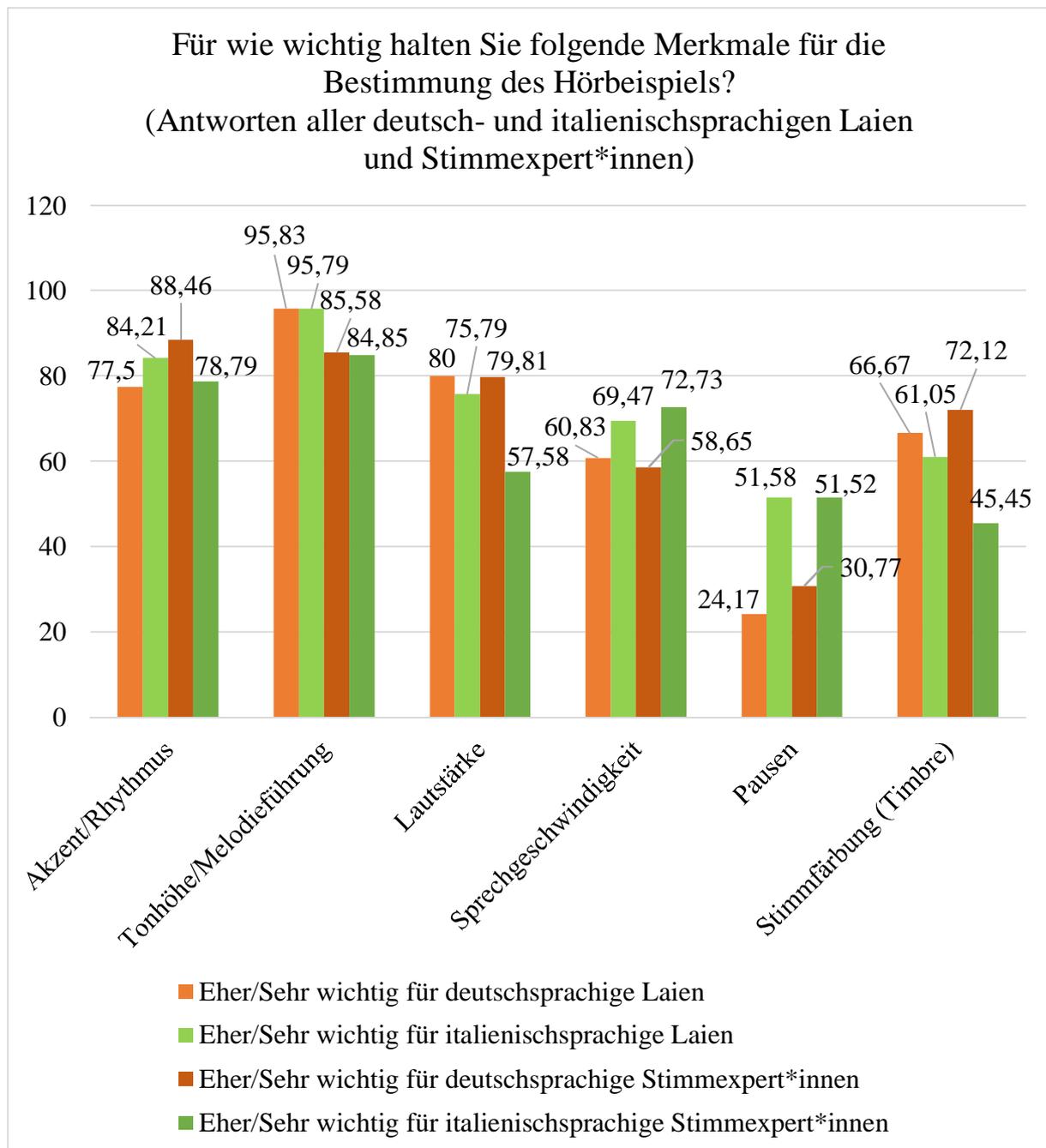


Die Beurteilungen der Stimmexpert*innen unterscheiden sich, insbesondere bei den prosodischen Merkmalen Tonhöhe/Melodieführung und Akzent/Rhythmus, deutlich im Vergleich zu denen der Laien. So halten die Expert*innen die oben genannten Merkmale insgesamt für wichtiger als die Laien, außerdem werden sie als fast gleich relevant bewertet und nicht, wie bei den

Laien, die Tonhöhe/Melodieführung deutlich dem Akzent/Rhythmus vorgezogen. Außerdem erfolgen die Beurteilungen nicht nur hauptsächlich anhand der Mittelwerte, sondern im Vergleich mit den Laien häufiger auch anhand der Kategorien „nicht wichtig“ und „sehr wichtig“, was auf eine sicherere Unterscheidung hindeutet. Bei den Merkmalen Lautstärke, Sprechgeschwindigkeit, Pausen und Stimmfärbung (Timbre) sind die Bewertungen innerhalb der Kategorien „unwichtig/wichtig“ insgesamt jedoch fast identisch mit jenen der Laien. Hier eine Übersicht zu den einzelnen Bewertungen:

	Nicht wichtig / Eher unwichtig	Eher wichtig / Sehr wichtig
Akzent/Rhythmus	13,87	86,14
Tonhöhe/Melodieführung	14,6	85,41
Lautstärke	25,55	74,46
Stimmfärbung (Timbre)	34,31	65,7
Sprechgeschwindigkeit	37,96	62,05
Pausen	64,24	35,77

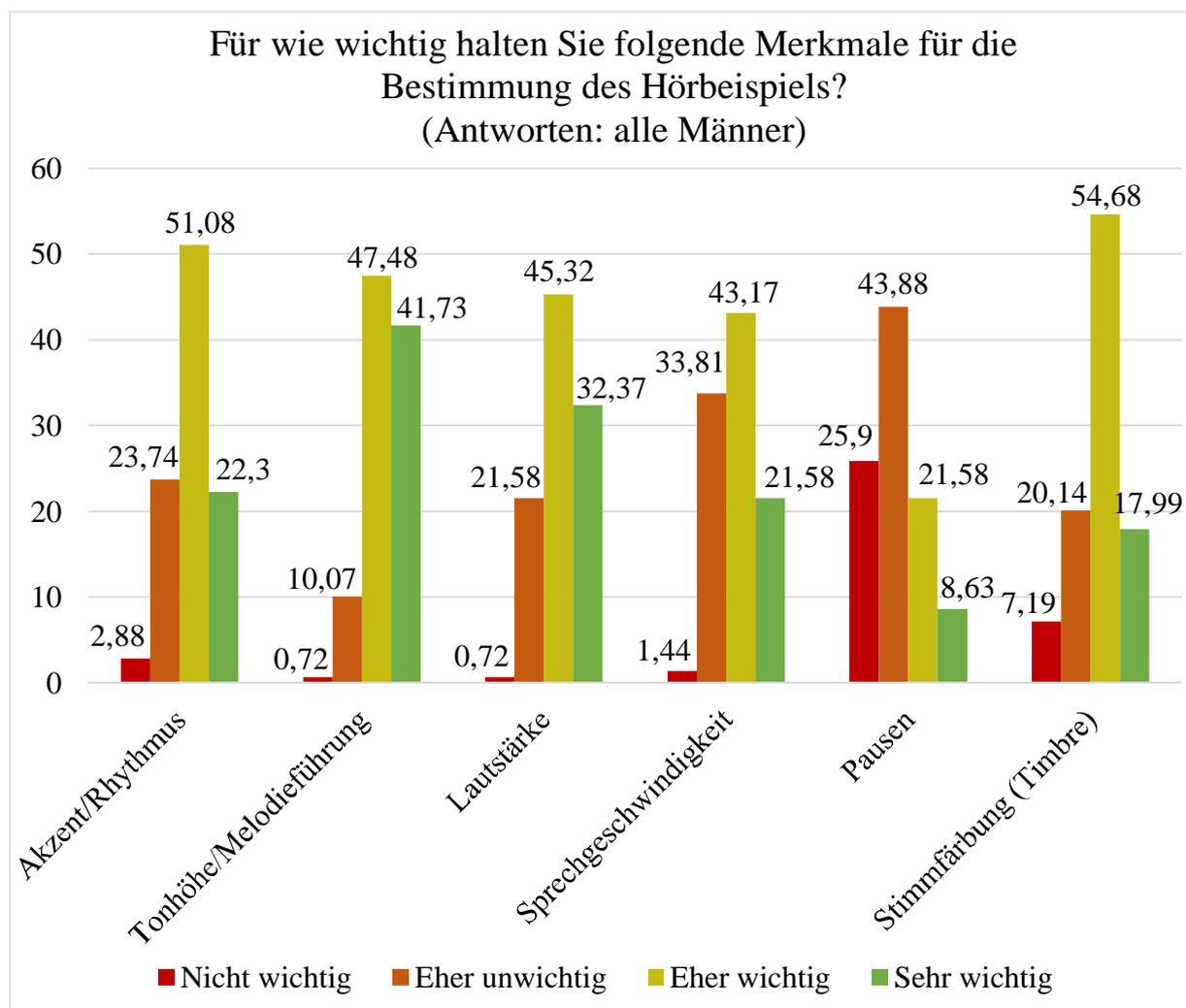
3.3.2.6 Deutsch- und italienischsprachige Laien und Stimmexpert*innen



In diesem Diagramm werden nur die Ergebnisse aus den Kategorien „eher wichtig“ oder „sehr wichtig“ dargestellt, dabei kommen die Faktoren Beruf (Laie bzw. Stimmexpert*in) und Sprache/Kultur (dt./it.) in der Darstellung der für die jeweiligen Gruppen als relevant betrachteten prosodischen Parameter zur Geltung. Dabei lässt sich feststellen, dass es zum einen sprachlich/kulturell bedingte Unterschiede gibt (die Pausen sind für die italienischsprachigen Proband*innen unabhängig vom Beruf relevanter als für die deutschsprachigen Proband*innen), berufsbezogene Unterschiede (Alle Laien halten die Tonhöhe/Melodieführung unabhängig von

Sprache/Kultur für wichtiger) und berufsbedingte Unterschiede innerhalb der jeweiligen Sprachgruppe. Diese betreffen das Merkmal Akzent/Rhythmus, welches die deutschsprachigen Stimmexpert*innen für relevanter halten als die deutschsprachigen Laien, das Merkmal Lautstärke, das die italienischsprachigen Laien für wichtiger halten als die italienischsprachigen Expert*innen, und das Merkmal Timbre, das die italienischsprachigen Laien wieder für relevanter halten als die gleichsprachlichen Expert*innen. Beim Merkmal Sprechgeschwindigkeit lässt sich unabhängig vom Beruf eine sprachlich/kulturell bedingte Divergenz dahingehend feststellen, dass es eine deutlich größere Rolle bei den italienischsprachigen als bei den deutschsprachigen Proband*innen spielt. Die entsprechenden sprachlich/kulturell oder berufsbedingten Differenzen müssen also konkret am jeweiligen prosodischen Merkmal festgemacht werden.

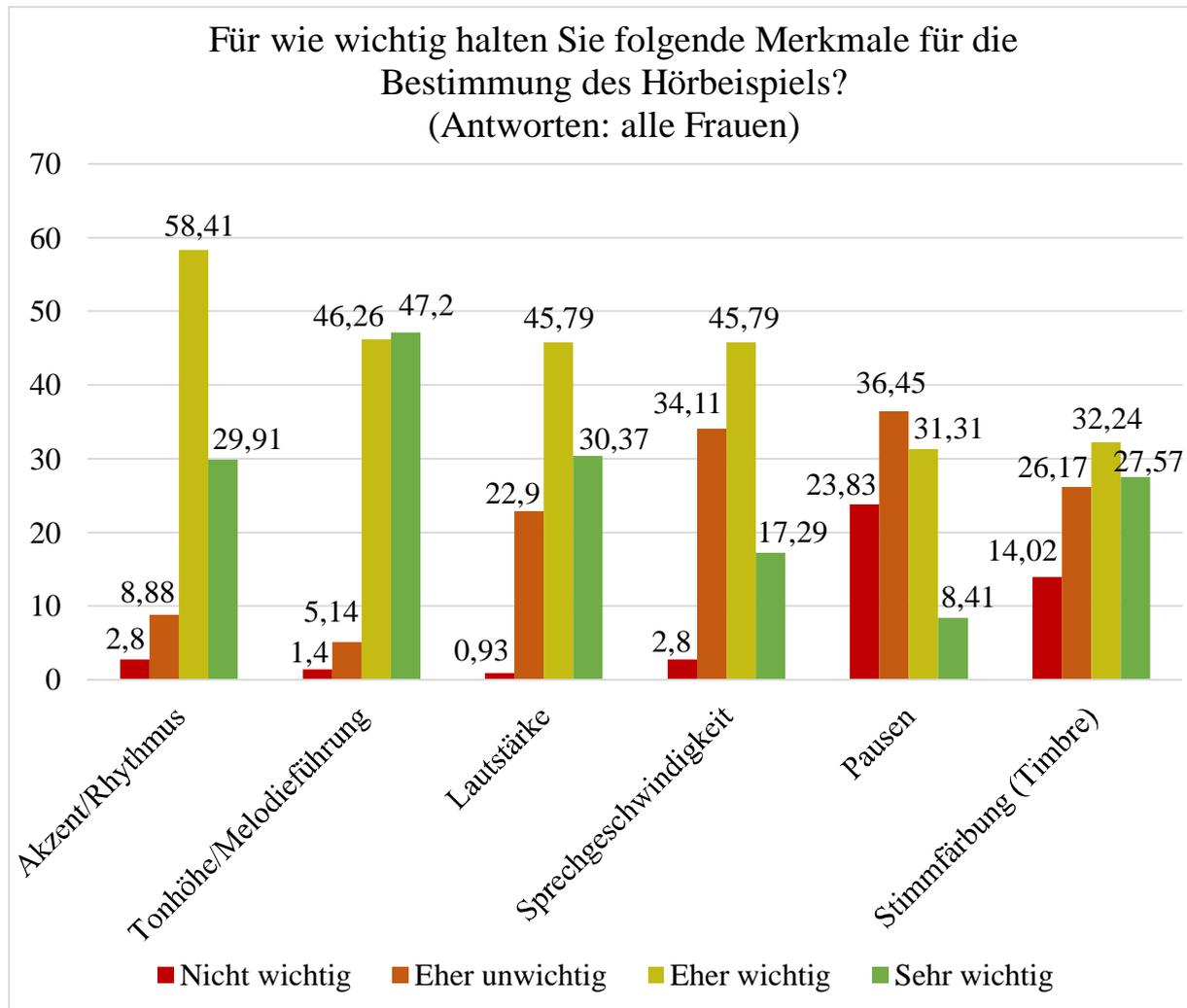
3.3.2.7 Männer



Nun soll die Kategorie Geschlecht in der Auswertung näher beleuchtet werden. Im Diagramm befinden sich die Antworten aller Männer, welche Tonhöhe/Melodieführung für besonders relevant halten, gefolgt von Lautstärke, Akzent/Rhythmus und Stimmfärbung (Timbre). Überdurchschnittlich oft wurde die Bewertung „eher wichtig“ gewählt. Die folgende Tabelle zeigt die jeweils angegebenen Werte (in Prozent) in den Kategorien „unwichtig/wichtig“.

	Nicht wichtig / Eher unwichtig	Eher wichtig / Sehr wichtig
Tonhöhe/Melodieführung	10,79	89,21
Lautstärke	22,3	77,69
Akzent/Rhythmus	26,62	73,38
Stimmfärbung (Timbre)	27,33	72,67
Sprechgeschwindigkeit	35,25	64,75
Pausen	69,78	30,21

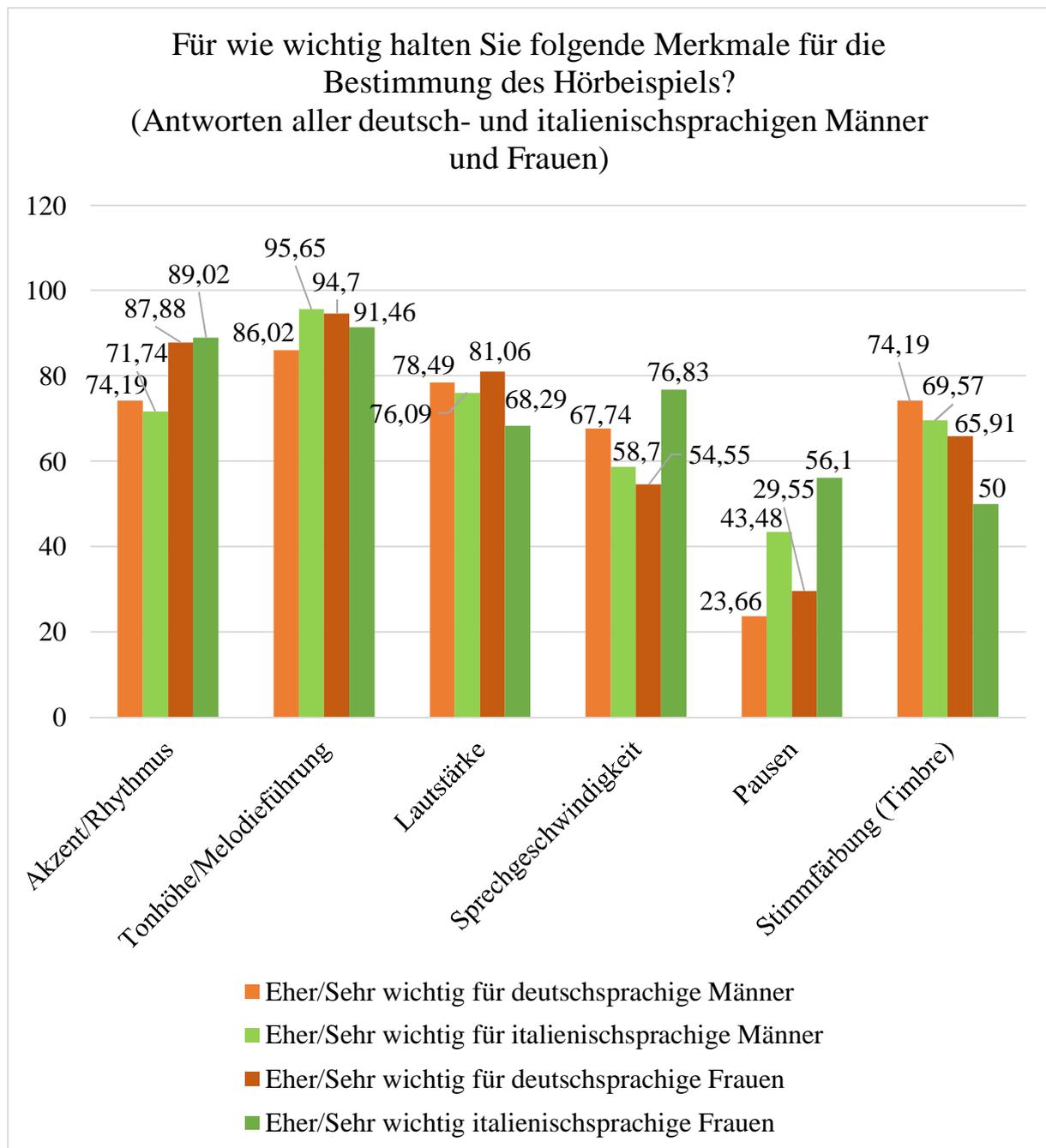
3.3.2.8 Frauen



Die Antworten der Frauen unterscheiden sich v. a. bei den Merkmalen Akzent/Rhythmus und Stimmfärbung (Timbre) deutlich von denen der Männer, bei der Bewertung der Pausen gibt es einen leichten Unterschied. Letztere werden von den Frauen tendenziell wichtiger eingeschätzt als von den Männern. Am größten sind die Unterschiede beim Akzent/Rhythmus, den 88,32 Prozent der Frauen, aber nur 73,38 Prozent der Männer für wichtig halten, und bei der Stimmfärbung (Timbre), welche deutlich mehr Relevanz aufweist für die Männer (72,67 Prozent) als für die Frauen (59,81 Prozent). Bei der Bewertung der Tonhöhe/Melodieführung, die auch bei den Frauen am wichtigsten ist, sowie bei jener der Lautstärke und Sprechgeschwindigkeit sind keine geschlechterbezogenen Unterschiede zu verzeichnen. Hier nun die Tabelle mit den zusammengefassten Daten:

	Nicht wichtig / Eher unwichtig	Eher wichtig / Sehr wichtig
Tonhöhe/Melodieführung	6,54	93,46
Akzent/Rhythmus	11,68	88,32
Lautstärke	23,83	76,16
Sprechgeschwindigkeit	36,91	63,08
Stimmfärbung (Timbre)	40,19	59,81
Pausen	60,28	39,72

3.3.2.9 Deutsch- und italienischsprachige Männer und Frauen



In diesem Diagramm erfolgt die Gegenüberstellung der Ergebnisse der verschiedenen Gruppen in Bezug auf ihr Geschlecht und ihre Sprache/Kultur. Diese beiden Parameter wirken sich in unterschiedlicher Zusammenwirkung in Bezug auf die jeweiligen prosodischen Merkmale aus. **Akzent/Rhythmus:** Hier zeigt sich deutlich nochmals die unterschiedliche Bewertung der Männer und Frauen unabhängig von ihrer Sprache/Kultur, für die Frauen (ca. 88 Prozent) ist dieses Merkmal deutlich relevanter als für die Männer (ca. 73 Prozent).

Tonhöhe/Melodieführung: Bei diesem Merkmal ist ein geschlechterbedingter Unterschied innerhalb der deutschsprachigen Gruppe zu beobachten, wobei die dt. Frauen es für etwas wichtiger halten als die dt. Männer. Sprachlich/kulturell bedingte Unterschiede gibt es keine, die Ergebnisse sind insgesamt recht einheitlich. Es steht fest, dass dieses Merkmal für alle das relevanteste darstellt.

Lautstärke: Hier zeigt sich ein sprachlich/kultureller Unterschied, wobei die deutschsprachige Gruppe dieses Merkmal für wichtiger hält als die italienischsprachige. Innerhalb der letztgenannten Gruppe zeigt sich auch ein leichtes geschlechterbedingtes Gefälle zwischen Männern, die dieses Merkmal tendenziell wichtiger einschätzen, und Frauen.

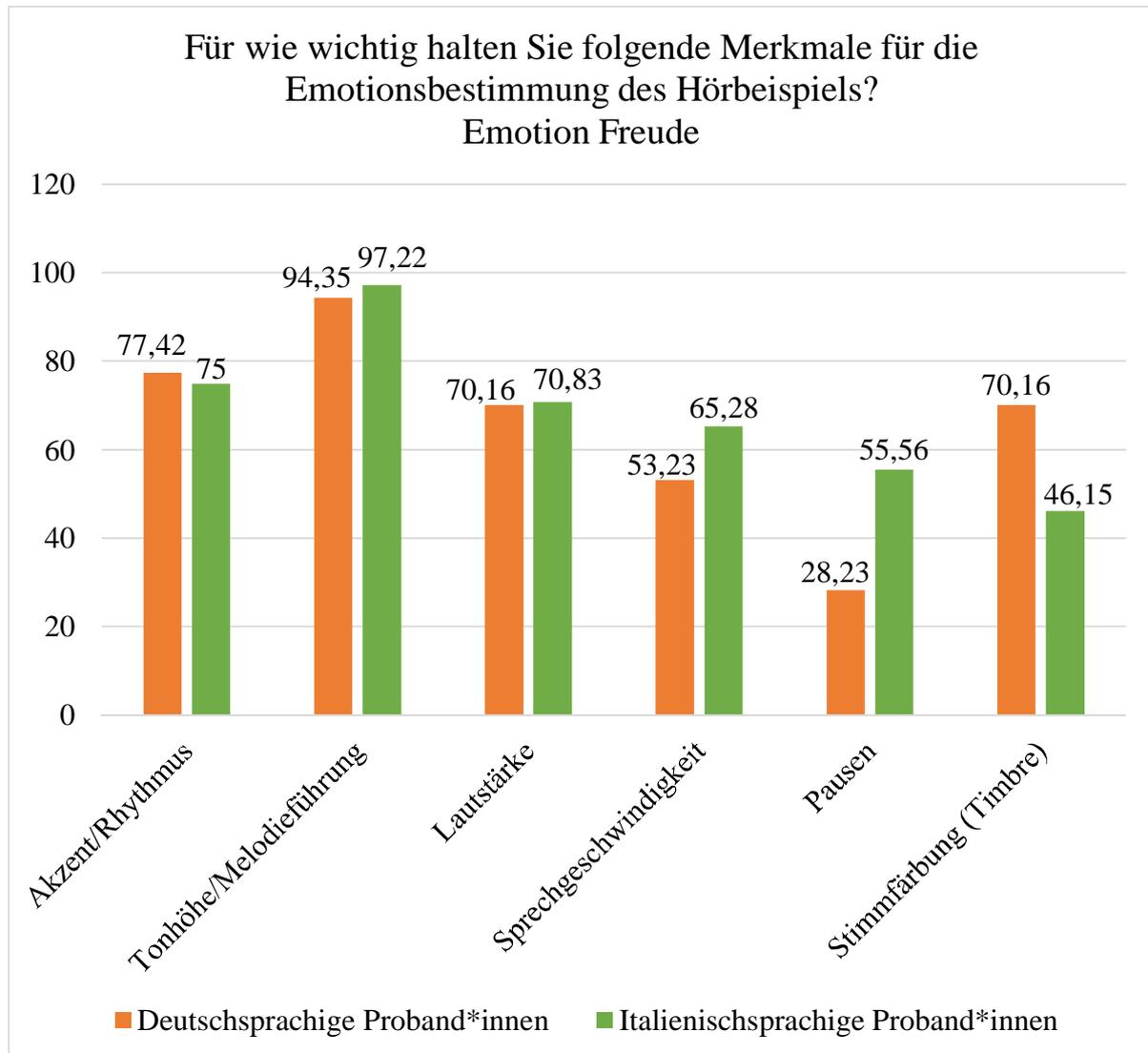
Sprechgeschwindigkeit: Bei diesem Merkmal zeichnet sich nicht nur ein sprachlich/kultureller Unterschied ab, sondern auch ein sehr deutlicher geschlechterbedingter Unterschied innerhalb der jeweiligen Sprachgruppe. In der deutschsprachigen Gruppe ist die Sprechgeschwindigkeit relevanter für die Männer (67,74 Prozent) als für die Frauen (54,55 Prozent), in der italienischen Gruppe verhält es sich jedoch umgekehrt. Hier halten 76,83 Prozent der Frauen, aber nur 58,7 Prozent der Männer dieses Merkmal für relevant.

Pausen: Hier zeigt sich eine starke sprachlich/kulturell bedingte Differenz – die italienischsprachigen Proband*innen halten dieses Merkmal für wichtiger als die deutschsprachigen. Außerdem zeigt sich in der italienischsprachigen Gruppe auch ein geschlechterbedingter Unterschied, da die Frauen dieses Merkmal für relevanter befinden als die Männer. Diese Tendenz ist auch innerhalb der deutschsprachigen Gruppe zu erkennen, wenngleich mit geringerer Ausprägung.

Stimmfärbung (Timbre): Dieses Merkmal ist für Männer insgesamt wichtiger als für Frauen, besonders stark ist der Unterschied allerdings zwischen den italienischsprachigen Männern (69,57 Prozent) und Frauen (50 Prozent). Auch zeichnet sich ein sprachlich/kulturell bedingtes Gefälle ab, wobei die deutschsprachige Gruppe dieses Merkmal für relevanter hält als die italienischsprachige.

3.3.2.10 Ergebnisse in Bezug auf die freudigen Sprechweisen

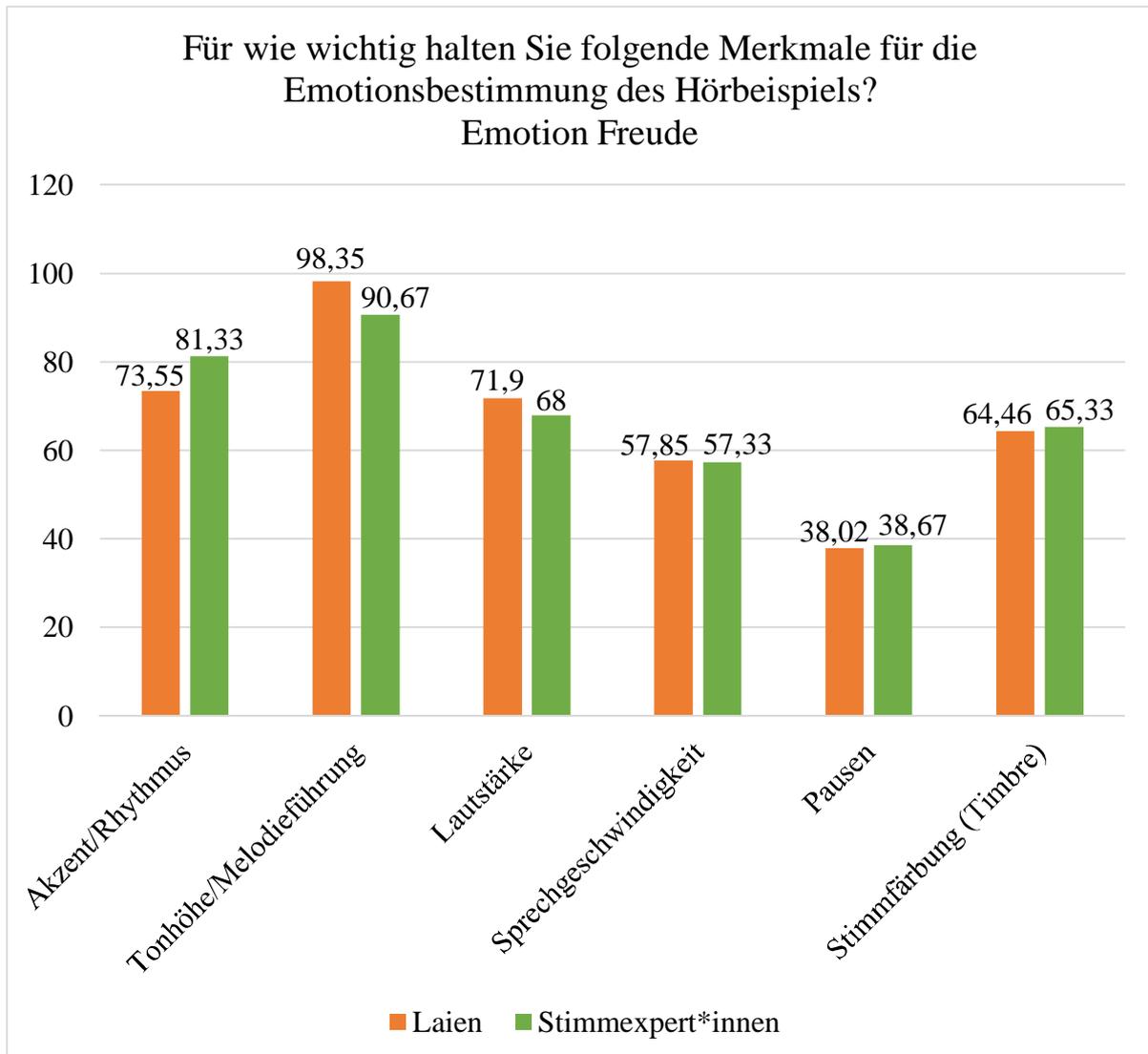
Deutsch- und Italienischsprachige Proband*innen



Bei der Frage, welche Relevanz den einzelnen prosodischen Parametern beim Indizieren der freudigen Sprechweisen zugemessen wird, fällt auf, dass es bei den Kategorien Stimmfärbung (Timbre), Pausen und Sprechgeschwindigkeit zu gewichtigen sprachlich/kulturellen Differenzen kommt. Während 70,16 Prozent der deutschsprachigen Proband*innen das Timbre für sehr wichtig halten, spielt es bei den italienischsprachigen Proband*innen mit 46,15 Prozent nur eine geringe Rolle. Anders verhält es sich mit den Pausen – fast 30 Prozent mehr italienischsprachige Personen halten dieses Merkmal für relevant (55,56 Prozent), bei den deutschsprachigen Proband*innen bildet dieser Parameter mit 28,23 Prozent das Schlusslicht im Ranking. Auch die Sprechgeschwindigkeit wird von den italienischsprachigen Personen für relevanter

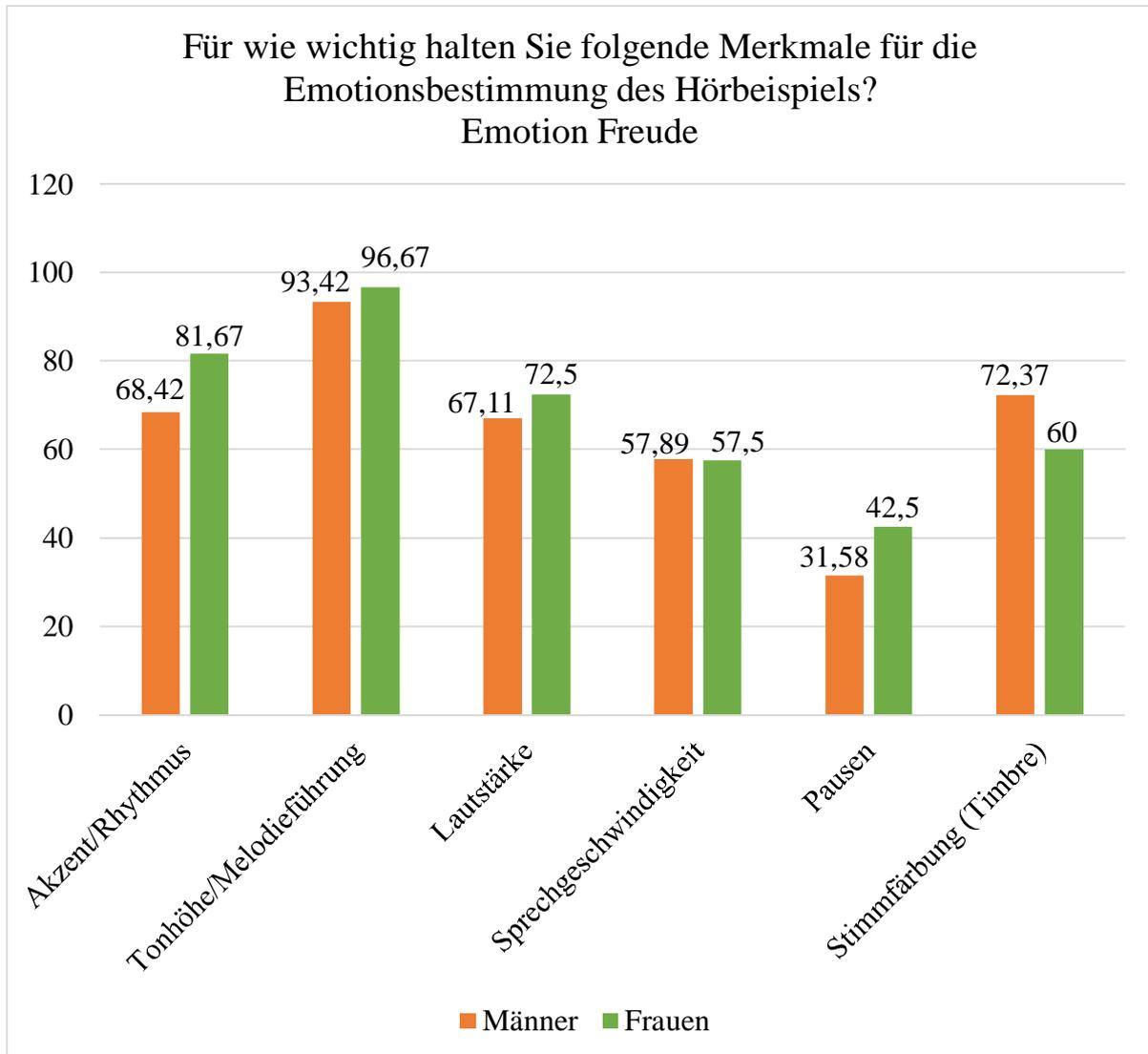
befunden als von den deutschsprachigen. Kaum Unterschiede lassen sich hingegen in der Beurteilung von Lautstärke, Tonhöhe/Melodieführung und Akzent/Rhythmus verzeichnen.

Laien und Stimmexpert*innen



Zieht man die jeweilige Beurteilung der Laien und Stimmexpert*innen heran, so fällt ein sehr ausgeglichenes Bild ins Auge. Die Angaben sind zu fast allen prosodischen Merkmalen beinahe identisch. Lediglich bei zwei Parametern lassen sich Differenzen verzeichnen. So wird der Akzent/Rhythmus von den Expert*innen als relevanter eingeschätzt als von den Laien. Umgekehrt halten die Laien die Tonhöhe/Melodieführung für wichtiger als die Stimmexpert*innen. Insgesamt sind die Beurteilungen jedoch sehr ausgewogen.

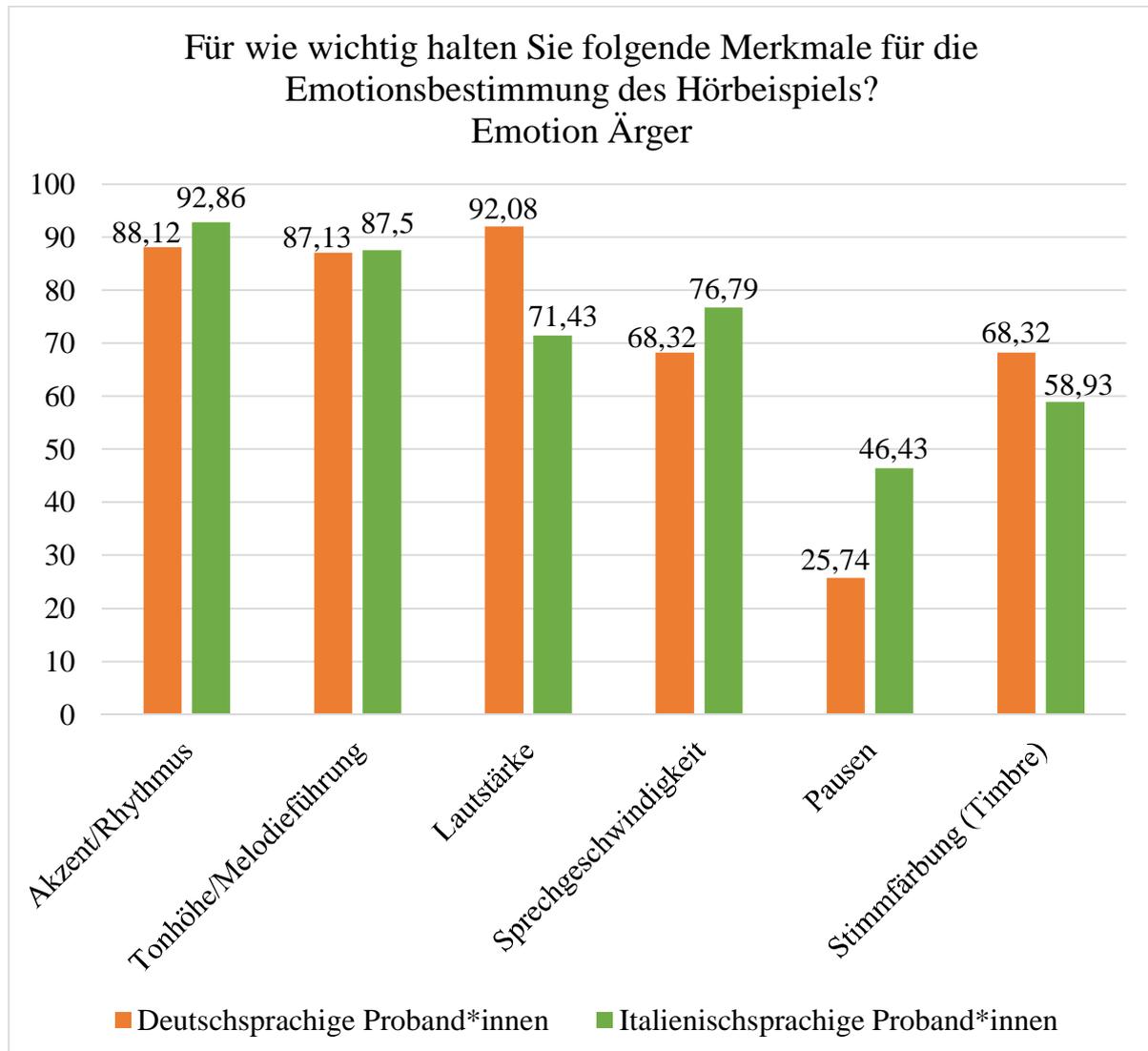
Männer und Frauen



Das obige Diagramm zeigt die unterschiedliche Bewertung der prosodischen Merkmale in Bezug auf die Emotionsbestimmung der freudigen Sprechweisen von Seiten der Männer und Frauen. Geschlechterspezifische Differenzen lassen sich dabei hinsichtlich der Parameter Stimmfärbung (Timbre), Pausen und Akzent/Rhythmus feststellen. Das Timbre wird von den Männern als deutlich relevanter eingeschätzt als von den Frauen. Umgekehrt verhält es sich bei den Pausen und dem Akzent/Rhythmus, welche für die Frauen eine größere Rolle bei der Emotionsperzeption der Freude spielen.

3.3.2.11 Ergebnisse in Bezug auf die Sprechweisen zu Ärger

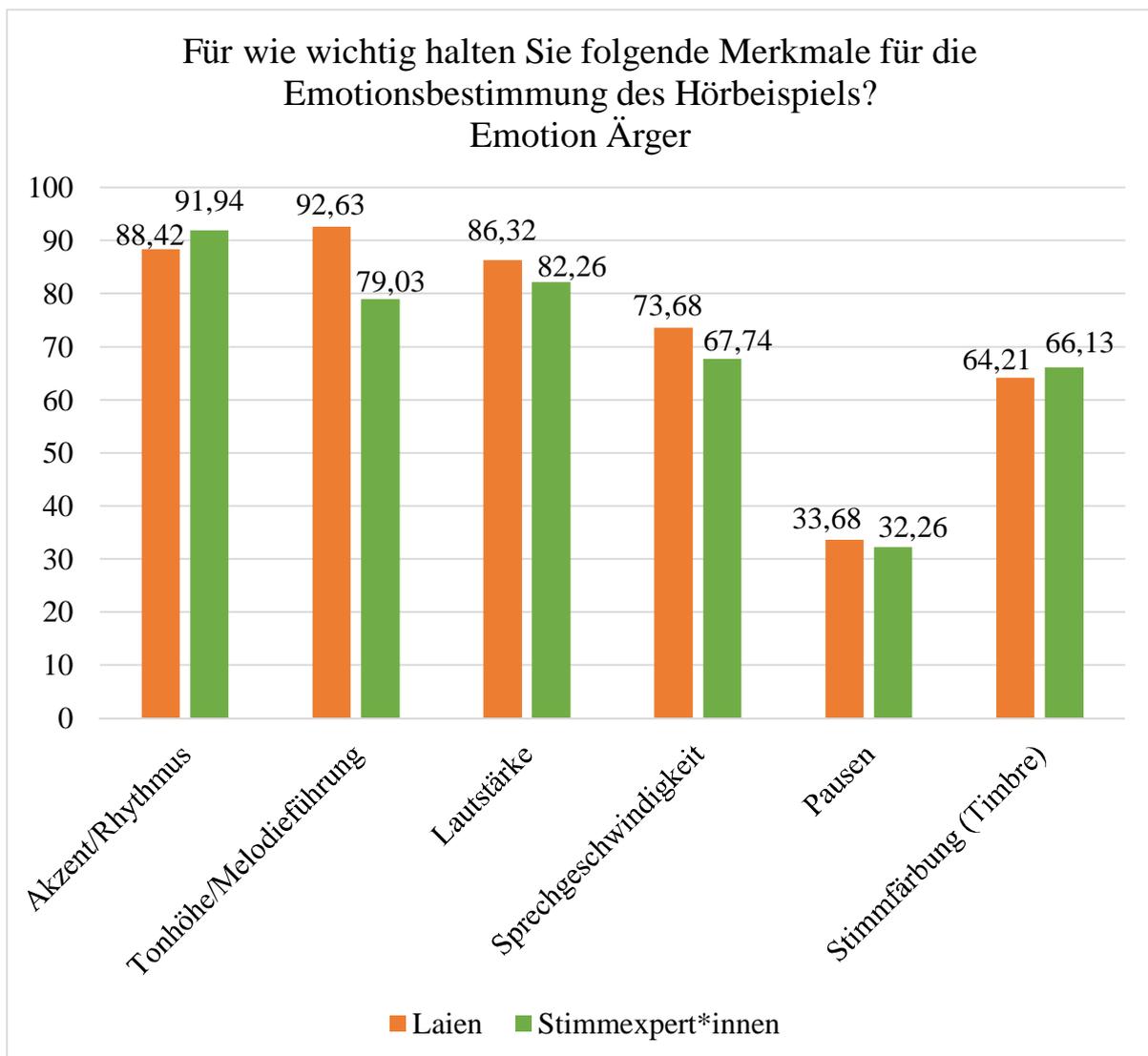
Deutsch- und italienischsprachige Proband*innen



Bei der Beurteilung der prosodischen Merkmale im Rahmen der ärgerlichen Sprechweisen zeichnet sich ein sehr ähnliches Bild der unterschiedlichen Gewichtung zwischen den beiden Sprachen/Kulturen wie bei der Beurteilung zur Freude ab. Abgesehen von der verschiedenen Gewichtung der Parameter insgesamt, auf welche unter Punkt 3.3.2.12 eingegangen wird, fällt v. a. die Beurteilung der Lautstärke auf. Sie wird bei der Erkennung von Ärger von 92,08 Prozent der deutschsprachigen Proband*innen für relevant gehalten, aber nur von 71,43 Prozent der italienischsprachigen und bei der Perzeption von Freude von beiden Gruppen ca. 70 Prozent. Auch das Timbre ist für die deutschsprachigen Proband*innen wieder relevanter als für

die italienischsprachigen, wengleich die Differenz hier im Vergleich zur Beurteilung der freudigen Sprechweisen geringer ausfällt. Bei den Pausen zeichnet sich auch bei der Bewertung der ärgerlichen Äußerungen wieder ein mit der Bewertung der freudigen Sprechweisen sehr ähnliches Bild ab – die italienischsprachigen Proband*innen halten dieses Merkmal für wesentlich relevanter als die deutschsprachigen.

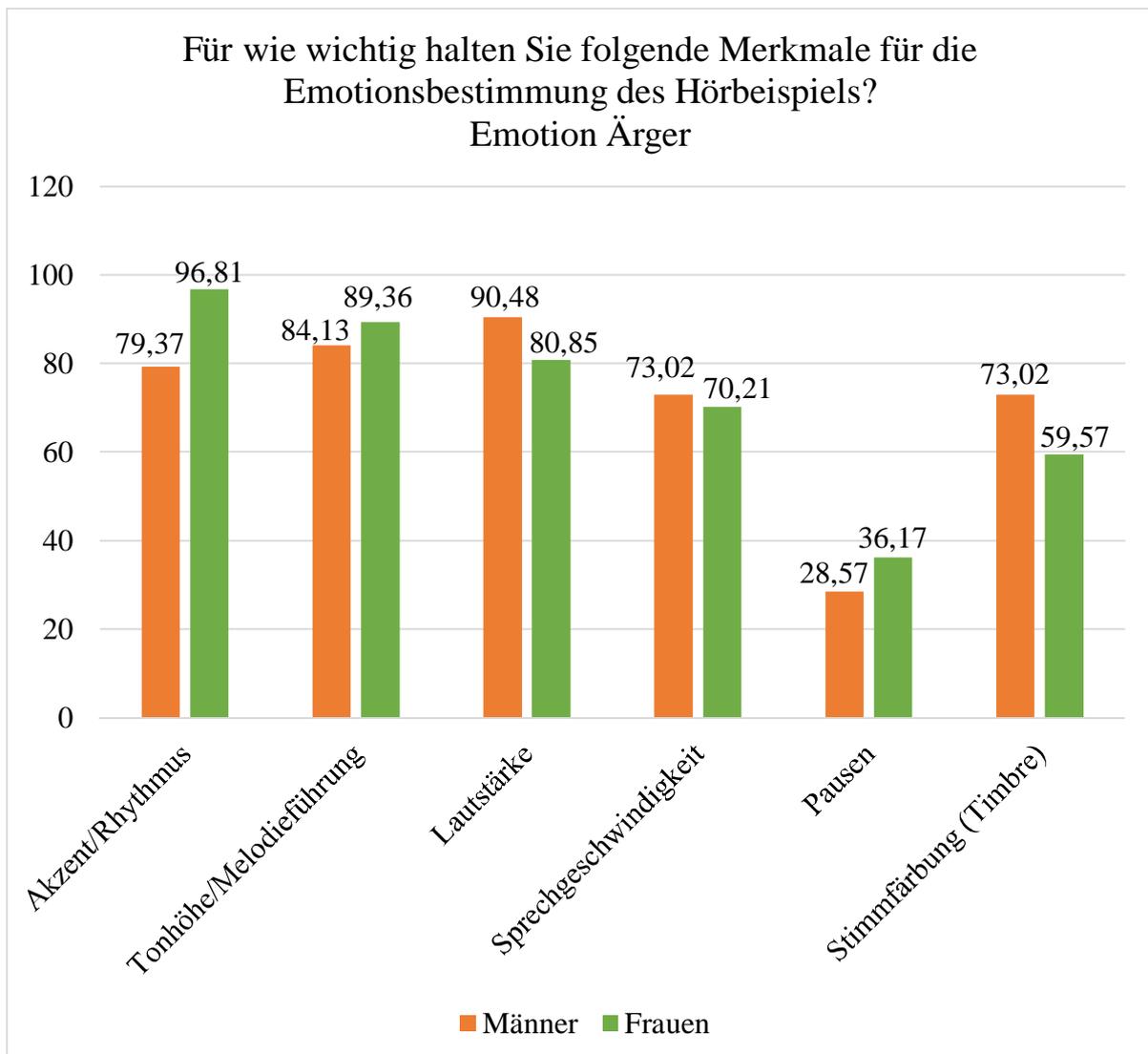
Laien und Stimmexpert*innen



Ähnlich wie beim Ergebnis der freudigen Sprechweisen zeigt sich auch in Bezug auf die ärgerlichen Äußerungen ein sehr einheitliches Bild bei der Beurteilung der Laien und Stimmexpert*innen. Die einzelnen prosodischen Parameter werden, bis auf Sprechgeschwindigkeit und Tonhöhe/Melodieführung, fast identisch bewertet. Das Merkmal Tonhöhe/Melodieführung

wird von den Laien (92,63 Prozent) für wesentlich relevanter befunden als von den Expert*innen (79,03 Prozent). Diese Differenz zeigt sich auf bei der Beurteilung in Bezug auf die freudigen Sprechweisen. Während dort jedoch auch die Bewertung von Akzent/Rhythmus differiert, zeigt sich hier ein Unterschied in Bezug auf die Sprechgeschwindigkeit. Diese wird von den Laien tendenziell als etwas wichtiger erachtet als von den Expert*innen. Bei der Beurteilung der freudigen Sprechweisen sind die Angaben zu diesem Merkmal hingegen identisch.

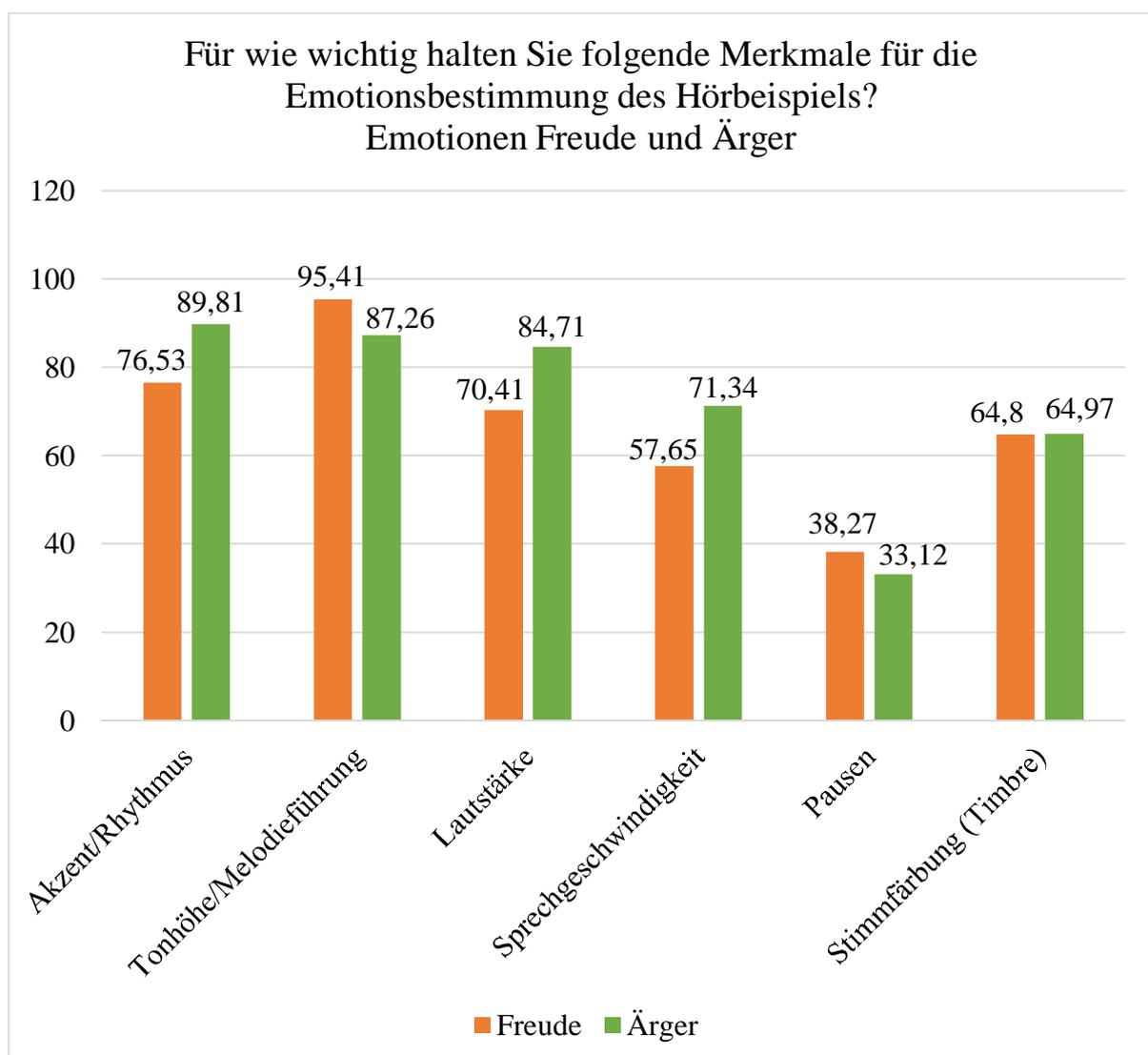
Männer und Frauen



Bei der Bewertung der Relevanz der einzelnen prosodischen Merkmale zur Emotionserkennung in Bezug auf die ärgerlichen Sprechweisen lassen sich einige geschlechterbedingte Differenzen feststellen. Ähnlich wie bei den Ergebnissen zu den freudigen Äußerungen spielen auch hier die Merkmale Akzent/Rhythmus und Pausen eine gewichtigere Rolle für die Frauen als für die

Männer. Umgekehrt ist auch bei der Erkennung der ärgerlichen Sprechweisen jedoch das Timbre für die Männer wesentlich relevanter als für die Frauen. Dieser Unterschied hatte sich auch bei den freudigen Sprechweisen gezeigt. Anders ist hier jedoch das Verhältnis der Bewertung des Merkmals Lautstärke, welches von den Männern als deutlich wichtiger eingeschätzt wird. Bei der Erkennung der freudigen Sprechweisen waren die Angaben zur Lautstärke hingegen ausgeglichen und sogar in leicht umgekehrter Tendenz (für Frauen wichtiger als für Männer).

3.3.2.12 Gegenüberstellung Freude – Ärger

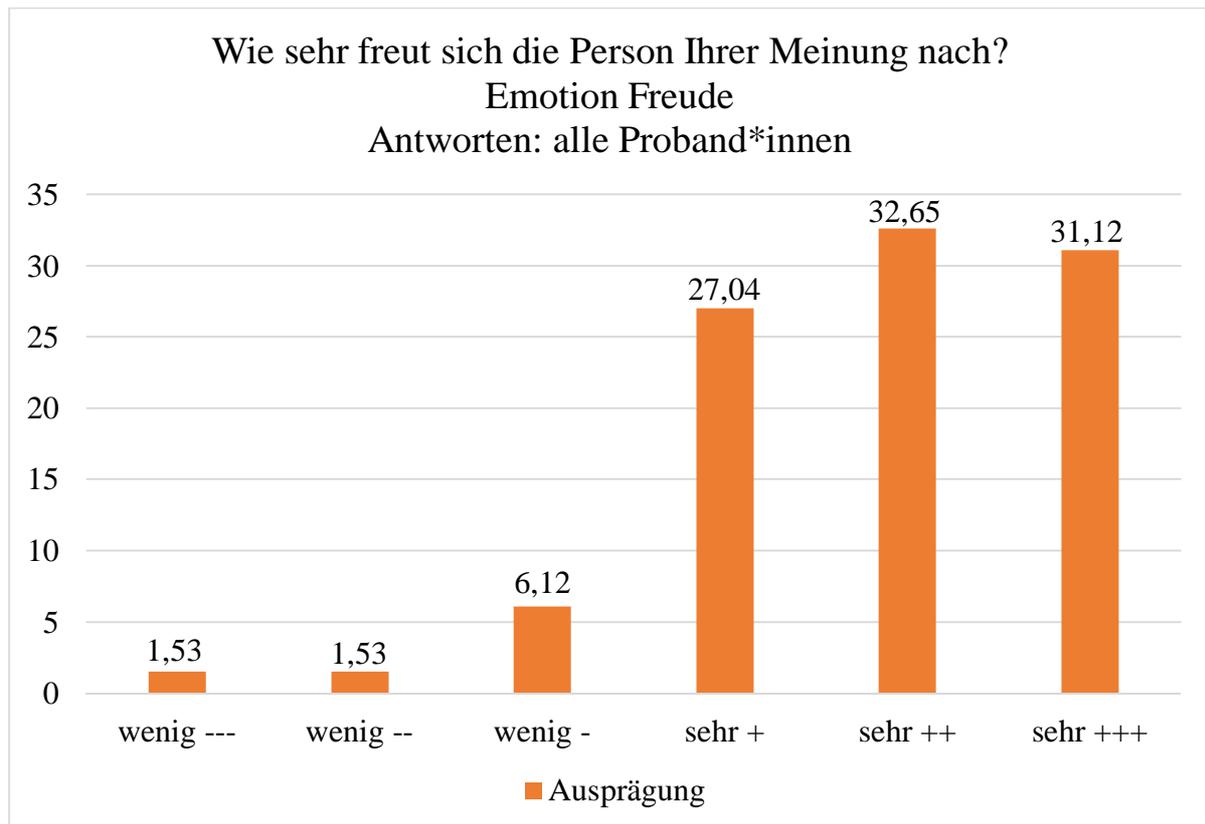


Betrachtet man die Ergebnisse zu den Bewertungen der einzelnen prosodischen Merkmale in Bezug auf die emotionalen Sprechweisen zu Freude bzw. Ärger, dann fallen in fast allen Parametern deutliche Unterschiede auf. Nur die Parameter Pausen und Stimmfärbung (Timbre) werden bei den freudigen und bei den ärgerlichen Sprechweisen für gleich wichtig bzw. unwichtig

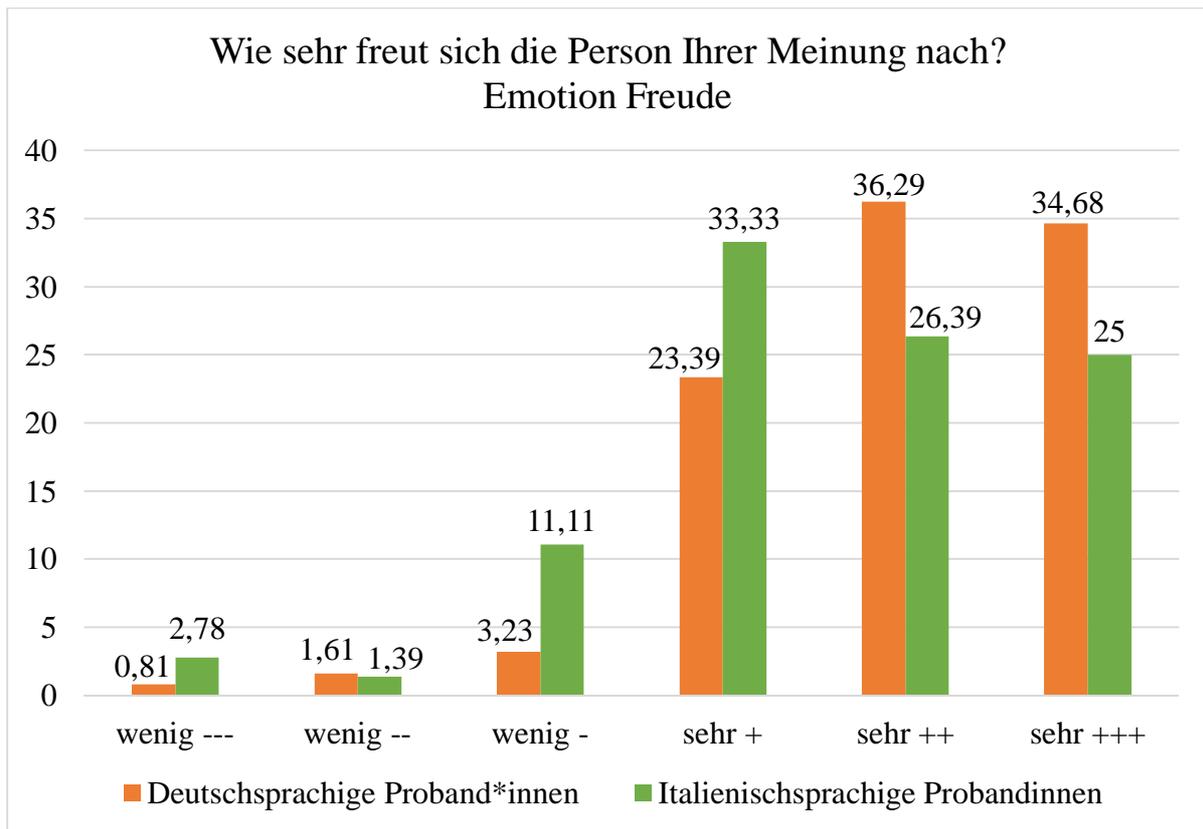
gehalten. Markante Differenzen lassen sich bei der Bewertung der Sprechgeschwindigkeit und der Lautstärke ablesen, welche für die Erkennung der ärgerlichen Sprechweisen ungleich wichtiger zu sein scheinen. Auch die Wichtigkeit des Merkmals Akzent/Rhythmus wird zur Bestimmung der ärgerlichen Äußerungen noch höher eingeschätzt. Der Parameter Tonhöhe/Melodieführung hingegen scheint tendenziell für die Zuordnung der freudigen Sprechweisen eine größere Rolle zu spielen als für die ärgerlichen, ist aber für beide Emotionen ein markantes Merkmal.

3.3.3 Ausprägungsgrad der jeweiligen Emotion

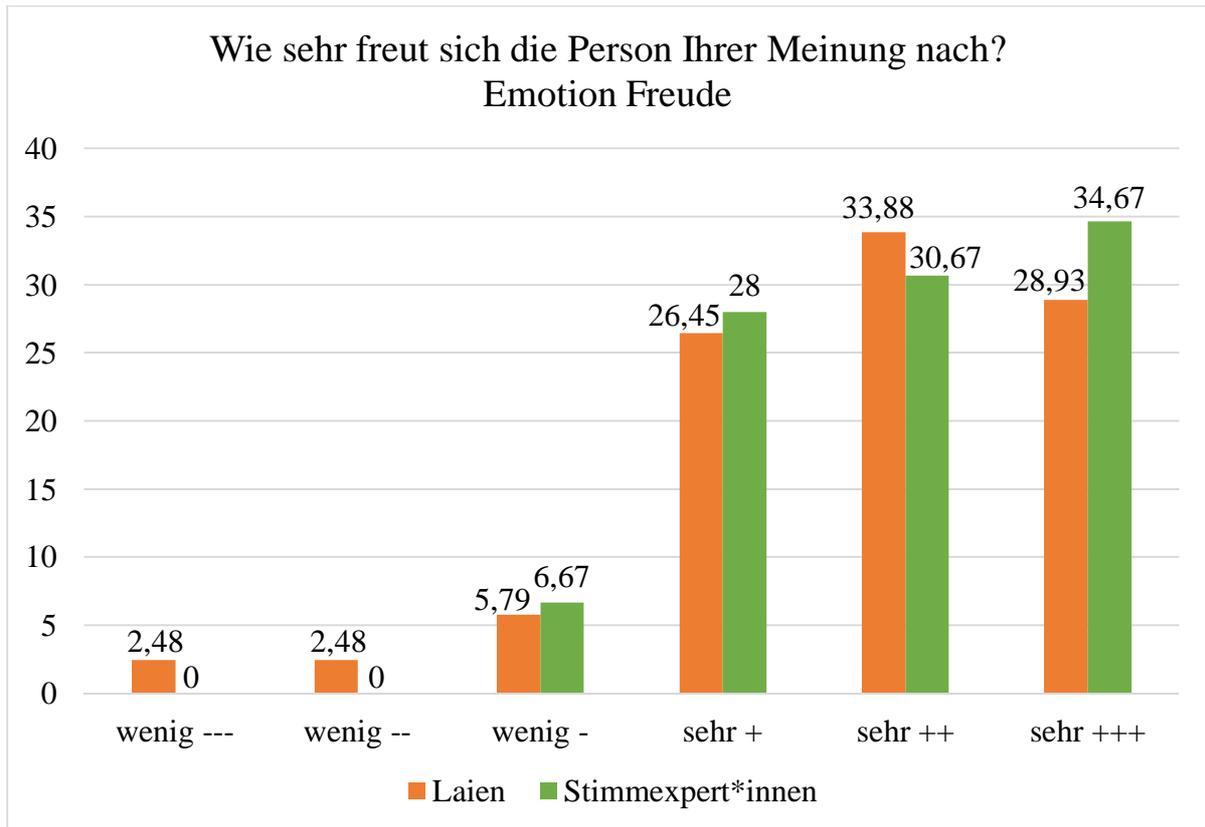
Im Rahmen der Untersuchung wurde auch nach dem Ausprägungsgrad der den jeweiligen Sprechweisen zugrunde liegenden Emotionen gefragt. Diese Frage dient der Rückkopplung an die Vorstudie, in der ja nur Sprechäußerungen ausgewählt wurden, die von allen Proband*innen als klar freudig oder ärgerlich wahrgenommen werden. Hier geht es jedoch um den zusätzlichen Aspekt des perzipierten Ausprägungsgrades, um zumindest tendenziell der Frage nachzuspüren, ob es hierbei auch sprach-/kulturbezogene, geschlechterspezifische oder berufsbedingte Differenzen gibt. Aufgrund eines Systemfehlers bei der Durchführung der Befragung in Bezug auf die ärgerlichen Sprechweisen, der leider erst im Nachhinein erkannt wurde, sind die Ergebnisse zur Emotion Ärger nicht gültig. Daher können im Folgenden nur die Ergebnisse zum perzipierten Ausprägungsgrad der Emotion Freude dargelegt werden. Die Ergebnisse hierzu sollten allerdings genügen, um zumindest die Möglichkeit von Wechselwirkungen und Differenzen durch die Faktoren Sprache/Kultur, Geschlecht und Beruf auch im Rahmen der Perception des Ausprägungsgrades der Emotion anzuzeigen. Auch wenn sich die Ergebnisse leider nur auf eine emotionale Sprechweise (Freude) beziehen, kann eindeutig erkannt werden, dass die oben genannten Faktoren, insbesondere Sprache/Kultur, auch die Wahrnehmung der Ausprägung der Emotion beeinflussen.



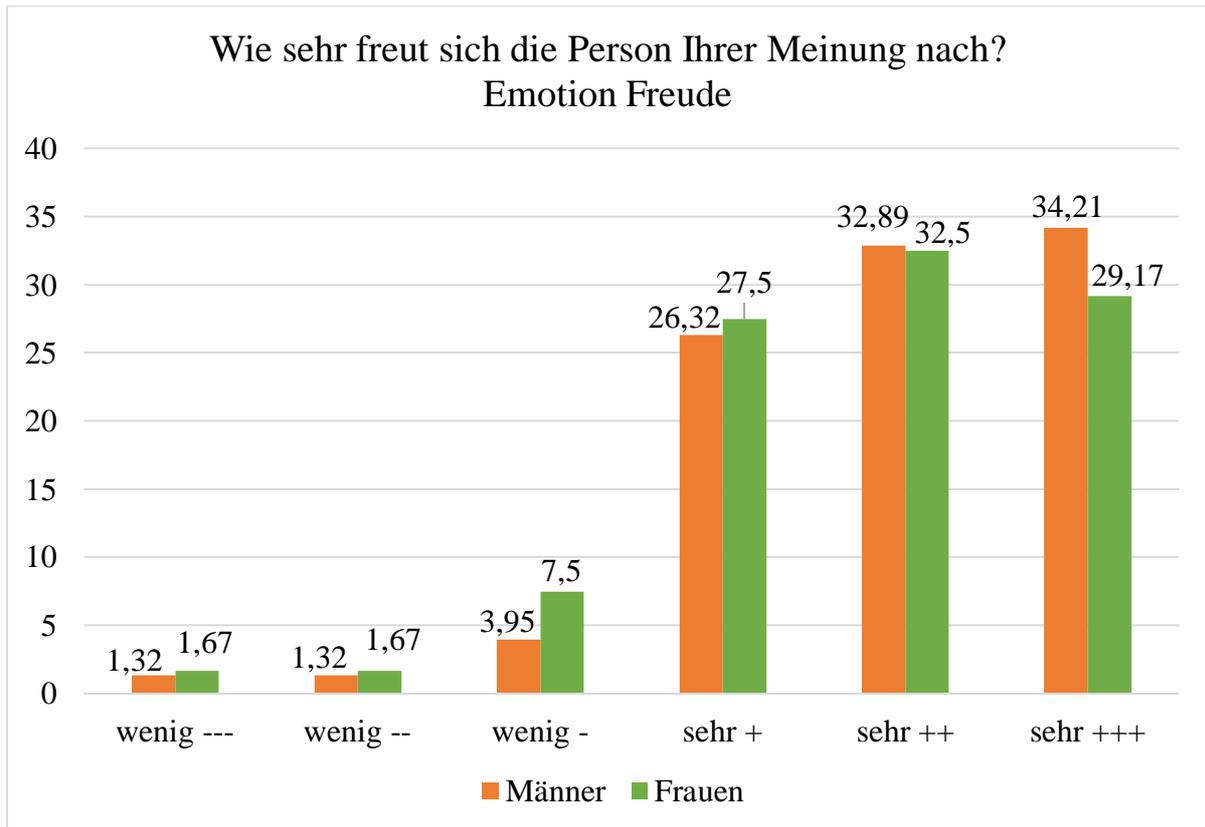
In einem ersten Schritt soll die Einschätzung aller Proband*innen veranschaulicht werden. Im Diagramm sind die sechs Abstufungen mit Plus- und Minuszeichen gekennzeichnet. Es geht klar hervor, dass eine starke Ausprägung der Emotion Freude in den gehörten Sprechweisen perzipiert wurde. Dies liegt auch daran, dass nur diejenigen Äußerungen im Rahmen der Vorstudie ausgewählt wurden, welche von den Proband*innen eindeutig einer Emotion zugeordnet werden konnten. Im Normalfall handelte es sich dabei um eher stark ausgeprägte Emotionen, die einen intensiven Gebrauch prosodischer Merkmale aufweisen.



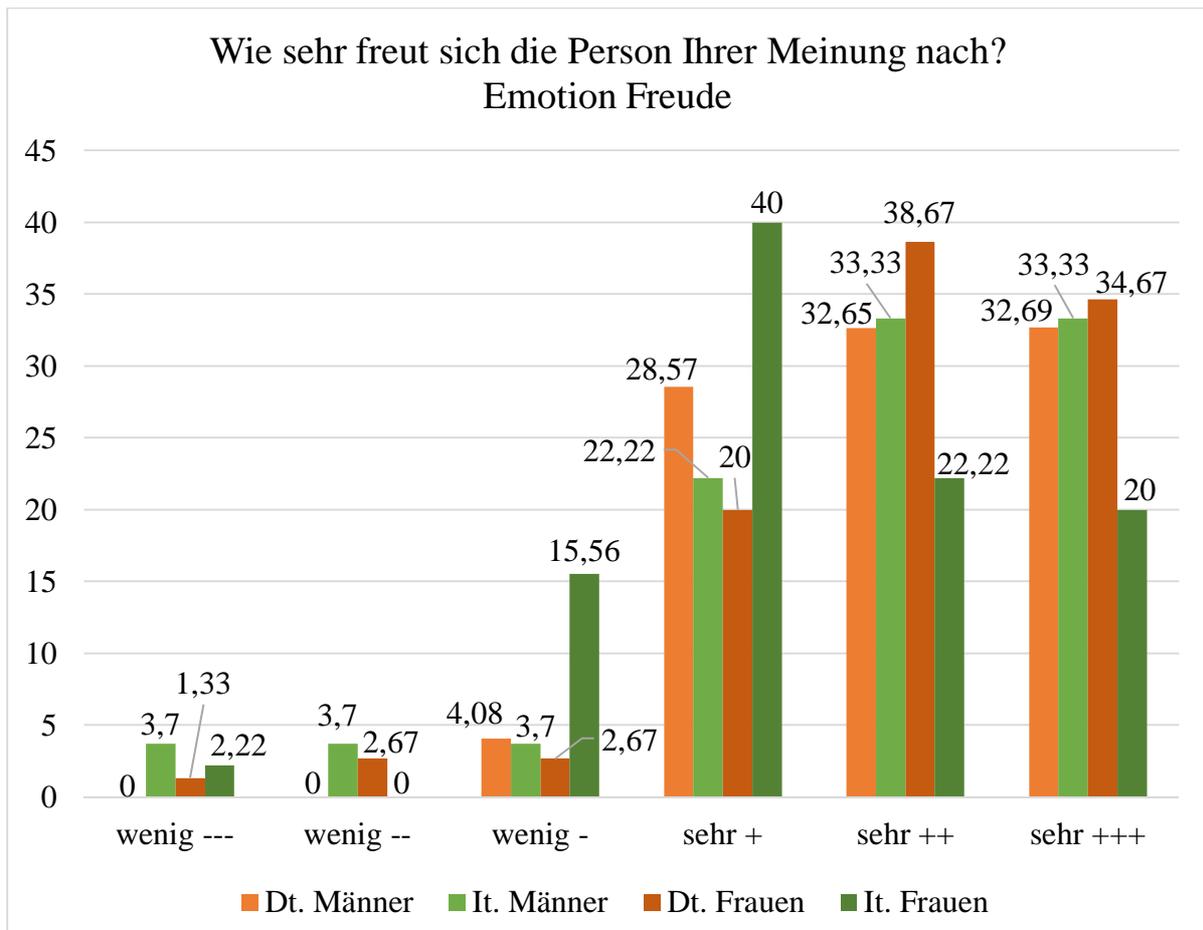
Differenziert man die Angaben nach Sprache/Kultur, so zeigt sich ein interessantes Bild – die deutschsprachigen Proband*innen nehmen die Emotion insgesamt als stärker ausgeprägt wahr als die italienischsprachigen Personen. Dieser Unterschied könnte darauf zurückgeführt werden, dass die Emotionen im italienischsprachigen Kontext anders bzw. unter stärkerem Gebrauch bestimmter prosodischer Mittel geäußert werden.



Bei der Einschätzung der Laien und Stimmexpert*innen in Bezug auf den Ausprägungsgrad der Emotion Freude zeigen sich leichte Unterschiede in den jeweiligen Zuordnungen, generell lässt sich aber eine klare Tendenz ablesen. So halten die Stimmexpert*innen die Emotion für stärker im Vergleich zu den Laien. 34,67 Prozent der Expert*innen geben die höchste Ausprägung an, aber nur 28,93 Prozent der Laien, außerdem nimmt niemand der Expert*innen Bewertungen in den unteren beiden Kategorien (wenig ---/wenig --) vor.



Bei Betrachtung der Ergebnisse zur Wahrnehmung des Ausprägungsgrads der Emotion Freude von Seiten der Männer und Frauen, lässt sich insgesamt kein relevanter geschlechtsbedingter Unterschied feststellen. Tendenziell halten die Männer die Emotion für etwas stärker als die Frauen, die Differenz ist jedoch nicht relevant. Zumindest in dieser generellen Aufstellung scheint der Faktor Geschlecht in der Wahrnehmung der Ausprägung der Emotion hier also kaum eine Rolle zu spielen.



Betrachtet man die Ergebnisse der Proband*innen in etwas differenzierterer Weise, so lassen sich doch z. T. sehr große geschlechterbezogene Unterschiede in der Perzeption innerhalb einer Sprach-/kulturgruppe aufweisen. So nehmen die italienischsprachigen Männer die Emotion Freude als deutlich stärker wahr als die italienischsprachigen Frauen. Innerhalb der deutschsprachigen Gruppe zeigt sich eine etwas geringer ausgeprägte umgekehrte Tendenz – die Frauen halten die Emotion für stärker ausgeprägt als die Männer.

3.3.4 Perzeption und Analyse der prosodischen Merkmale

Der dritte Bereich der Untersuchung ist der Perzeption und Analyse der einzelnen prosodischen Merkmale der emotionalen Sprechweisen zur Freude bzw. zum Ärger gewidmet und in mehrere Abschnitte und Teilfragen untergliedert. Jedes prosodische Merkmal wird aufgrund seiner wichtigsten Charakteristika analysiert. Dabei sind die Fragen so formuliert, dass sie sowohl das jeweilige prosodische Merkmal in ausreichender Genauigkeit beschreiben, als auch die Analysekompetenz aller Proband*innen (Laien und Stimmexpert*innen) berücksichtigen. So erfolgt eine genügende Einschätzung der prosodischen Merkmale, ohne die befragten Personen zu

überfordern und damit das Ergebnis zu verzerren. Um dem vorzubeugen und auch um die Durchführbarkeit der Analyse selbst zu testen, wurde außerdem die Kategorie „weiß ich nicht“ eingeführt, derer sich die Befragten bedienen können, wenn sie durch die Fragestellung überfordert sind oder das Herausfiltern der gefragten Informationen nicht möglich ist. Nur Einfachantworten sind gültig, bei den Zahlen handelt es sich um Angaben in Prozent. Im Folgenden werden die Ergebnisse der Analyse nach Emotionen getrennt (Freude – Ärger) dargestellt, da sich die charakteristische Ausprägung der einzelnen prosodischen Merkmale je nach Emotion (vgl. Kranich 2003) deutlich unterscheidet. Hier interessieren also die jeweiligen Angaben aller Hörbeispiele, die der Emotion Freude bzw. Ärger zugeordnet sind. Untersucht wird die Wahrnehmung der wichtigsten Charakteristika des jeweiligen prosodischen Merkmals in folgender Reihenfolge und in den entsprechenden angeführten Kategorien:

Akzent/ Rhythmus	Stärke	Häufigkeit	Sprechrhythmus		
Tonhöhe/ Melodieführung	Sprechstimmlage	Tonhöhenverlauf	Tonhöhenumfang		
Lautstärke	Lautstärke	Lautheitsverlauf	Lautheitsverlauf am Ende der Phrase		
Sprechgeschwindigkeit	Geschwindigkeit	Geschwindigkeitsverlauf	Geschwindigkeitsverlauf am Ende der Phrase		
Pausen	Häufigkeit	Geräuschvolle Pausen			
Stimmfärbung (Timbre)	Klangfarbe	Klangfülle	Faukale Distanz	Stimmklang	Stimmansatz

Die jeweiligen Kategorien, die überblicksartig in den Spalten rechts neben der Bezeichnung des jeweiligen prosodischen Merkmals stehen, sind so erklärt und beschrieben, dass sie von allen Personen verstanden und damit beurteilt werden können. Bei der Erstellung des Analyserasters wurde auf den Katalog zur auditiven Merkmalsbeschreibung von Sprechausdrucksmustern von Bose (Bose 2001) zurückgegriffen. Das Raster wurde für die vorliegenden Untersuchungszwecke entsprechend modifiziert.

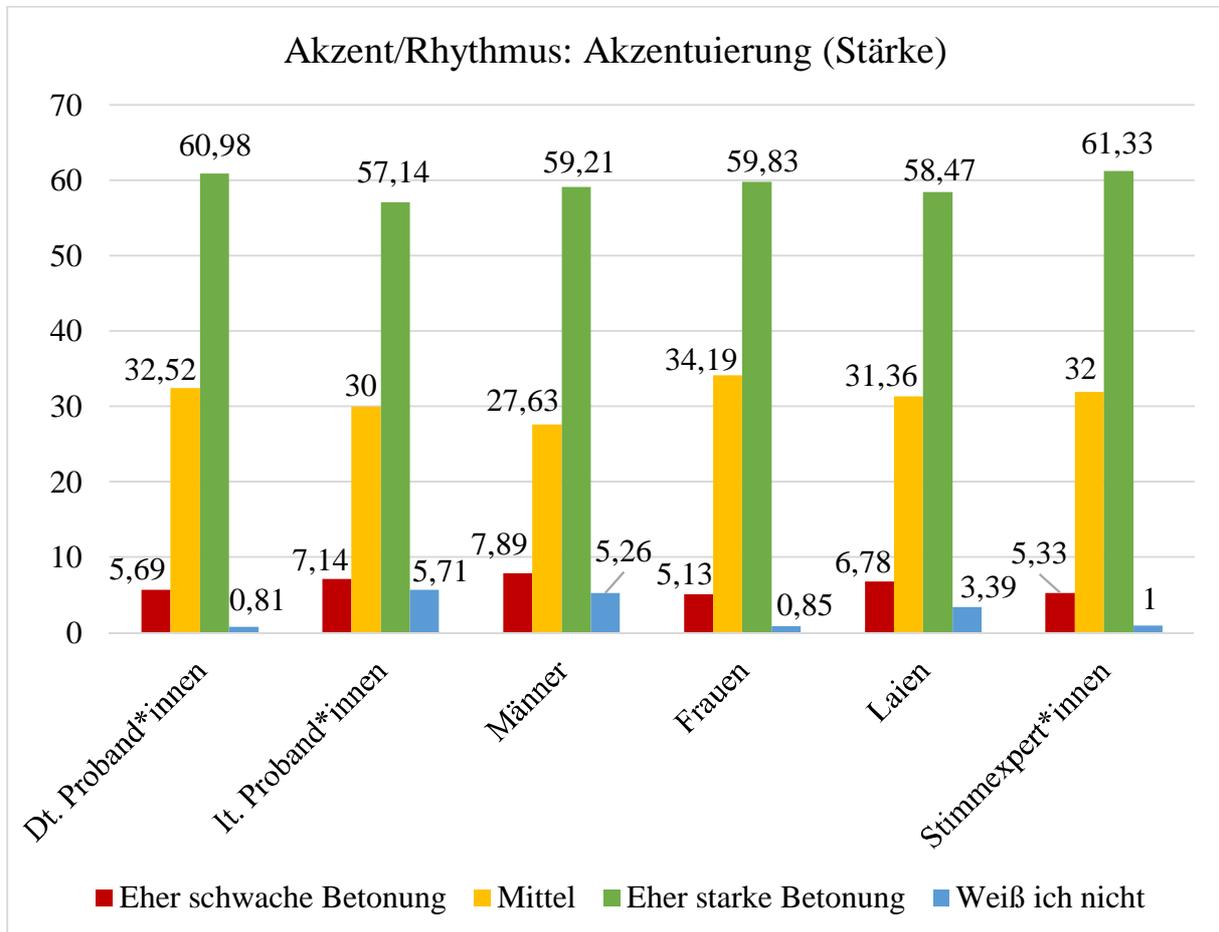
Zum einen soll dargestellt werden, wie die einzelnen prosodischen Merkmale der jeweiligen freudigen bzw. ärgerlichen Sprechweisen von den Proband*innen beschrieben werden. Zum anderen sollen dabei auch die Einflussfaktoren Sprache/Kultur, Beruf und Geschlecht berücksichtigt werden. Der besseren Übersicht wegen werden die Ergebnisse zur dritten Frage nicht genauer differenziert (also etwa Sprache/Kultur kombiniert mit dem Geschlecht der Proband*innen), sondern nur in die Kategorien Sprache/Kultur, Beruf und Geschlecht unterteilt. Dadurch können zwar wertvolle Detailinformationen verloren gehen, allerdings begünstigt dieses Vorgehen die deutlichere Darstellung der durch die genannten Faktoren bedingten Differenzen. Eine weitere Aufgliederung brächte diesbezüglich keinen relevanten Mehrwert. Überdies lässt sich das Zusammenwirken der einzelnen Faktoren durch die gewählte Darstellung (alle Faktoren sind in einem Diagramm veranschaulicht) trotzdem tendenziell ablesen. Eine weitere Untergliederung ist für die Beantwortung der Forschungsfrage nicht notwendig. Im Folgenden werden die Ergebnisse zur Perzeption und Analyse der freudigen und dann der ärgerlichen Sprechweisen dargestellt. In einem dritten Schritt werden die Ergebnisse zu den beiden emotionalen Ausprägungen schließlich miteinander verglichen.

3.3.4.1 Perzeption und Analyse der prosodischen Merkmale – Freude

Die Auswertung der Ergebnisse erfolgt in der Reihenfolge, in der die einzelnen prosodischen Merkmale von den Proband*innen beurteilt wurden (siehe Tabelle oben). Zur Veranschaulichung der Daten wurden Diagramme erstellt, welche die Analyseergebnisse der Proband*innen nach Sprache/Kultur, Geschlecht und Beruf differenziert umfassen und abgleichen.

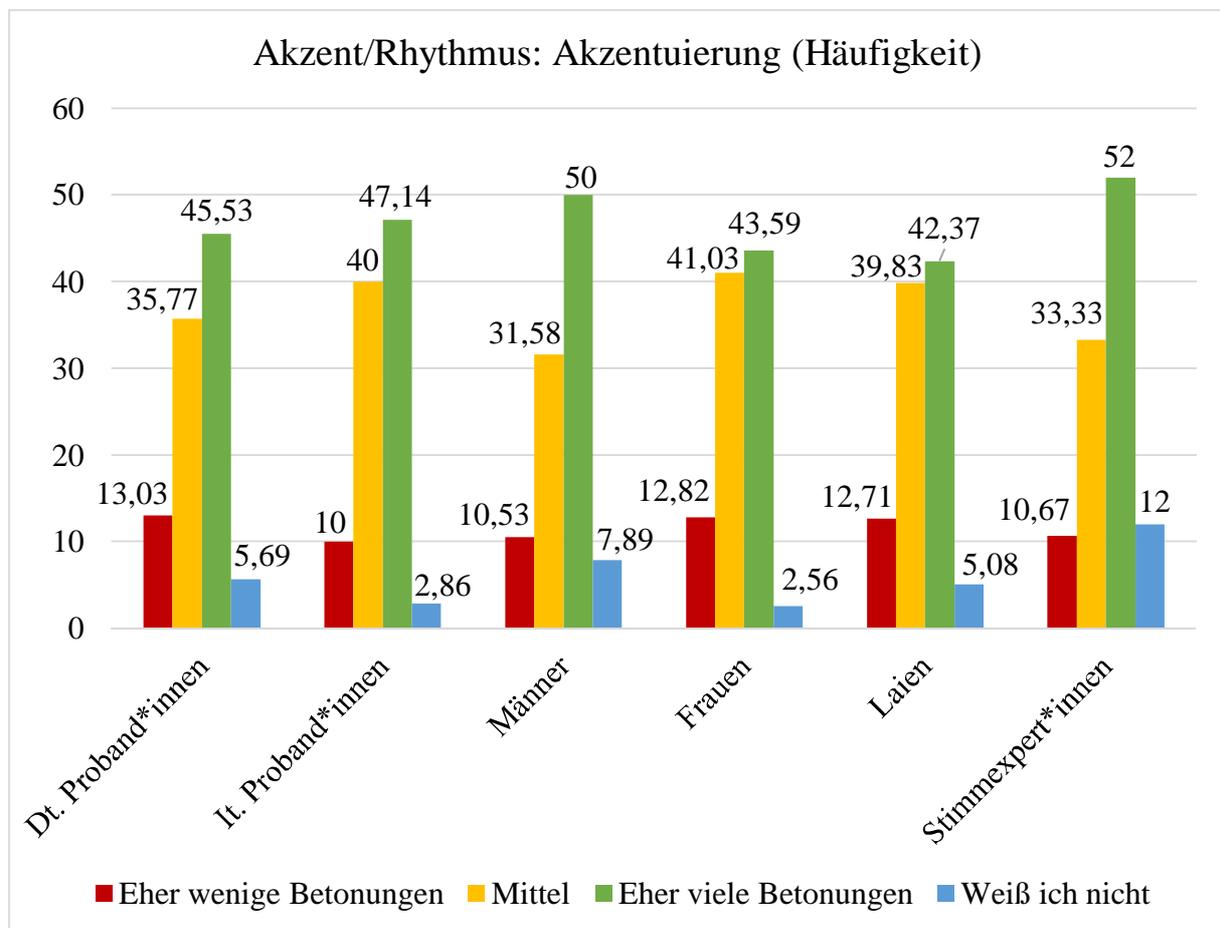
3.3.4.1.1 Akzent/Rhythmus

Akentuierung (Stärke)



Die Wahrnehmung der Stärke der Akzentuierungen in den freudigen Sprechweisen ist sehr homogen. Circa 59 Prozent der Proband*innen perzipieren eher starke Betonungen, ca. 31 Prozent mittlere Betonungen und nur ca. 6 Prozent eher schwache Betonungen. Die Zahl derer, die diese Frage nicht beantworten können („weiß ich nicht“) ist verschwindend gering, am höchsten jedoch bei den italienischsprachigen Proband*innen und bei den Männern. Hier kann also ein sehr klares Ergebnis abgelesen werden, das sich kaum durch die Faktoren Sprache/Kultur, Geschlecht oder Beruf unterscheidet.

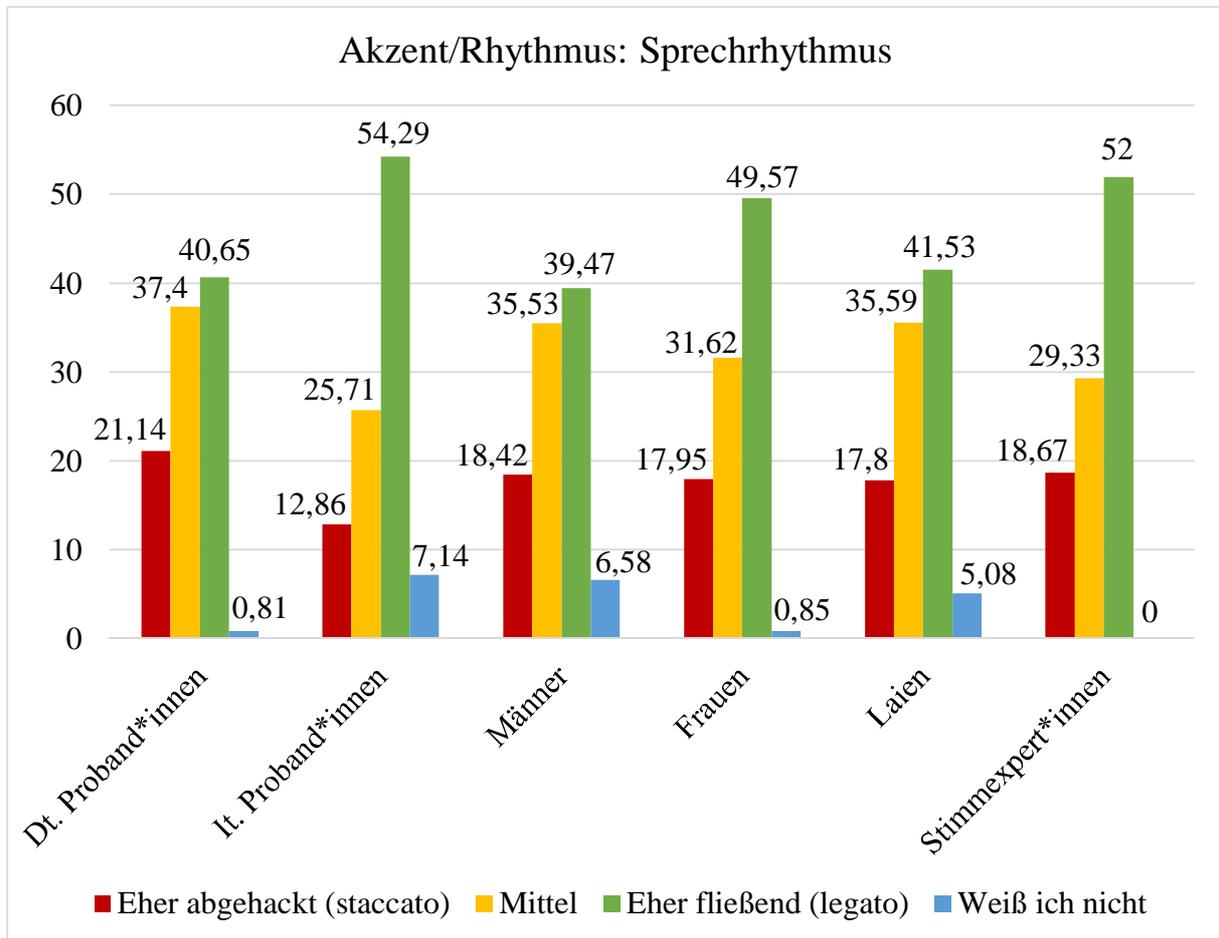
Akzentuierung (Häufigkeit)



Die zweite Dimension der Akzentuierung, die hier untersucht wurde, ist die Häufigkeit. Die Mehrheit der Proband*innen perzipiert eher viele Betonungen, gefolgt von mittlerer Häufigkeit und eher wenigen Betonungen. Die wenigsten Angaben befinden sich in der Kategorie „weiß ich nicht“, hier allerdings am häufigsten von den Stimmexpert*innen und von den Männern, am wenigsten von den italienischsprachigen und weiblichen Proband*innen verwendet. Das Ergebnis ist hier nicht mehr so ähnlich wie bei den Angaben zur Stärke der Akzentuierungen im ersten Diagramm. Tatsächlich lassen sich hier nun einige Unterschiede in der Perzeption innerhalb der einzelnen Gruppierungen feststellen. Bei den Stimmexpert*innen sind die Unterschiede zwischen den einzelnen Angaben sehr deutlich - 52 Prozent perzipieren eher viele Betonungen, 33,33 Prozent eine mittlere Häufigkeit, 10,67 Prozent eher wenige Betonungen und 12 Prozent wissen es nicht. Ähnlich klar abgegrenzt sind die Antworten der Männer. Besonders schwammig wird sie hingegen bei den Frauen und bei den Laien, wo es keinen relevanten Unterschied zwischen häufig und mittelmäßig häufig perzipierten Akzentuierungen gibt. Etwas klarer tritt diese Unterscheidung bei den deutsch- und italienischsprachigen Proband*innen auf,

wobei kein bedeutender sprach-/kulturbedingter Unterschied in dieser Bewertung ersichtlich ist.

Sprechrhythmus



Größere kultur-/sprachbedingte, berufs- und geschlechtsspezifische Unterschiede in der Perzeption und Analyse des Merkmals Akzent/Rhythmus gibt es hier im Vergleich zu den Kategorien Akzentuierung. Der Sprechrhythmus wird vom Großteil der Proband*innen als eher fließend (legato) wahrgenommen. Dies gilt insbesondere für die italienischsprachigen Proband*innen, die Stimmexpert*innen und die Frauen. Hier zeigt sich eine deutliche Abgrenzung zu den anderen Bewertungsmöglichkeiten („mittel“, „eher abgehackt“, „weiß ich nicht“). Diese klare Unterscheidung fehlt bei den Laien, den Männern und den deutschsprachigen Proband*innen. Die Faktoren Geschlecht (klarere Beurteilung der Frauen) und Beruf (wesentlich klarere Beurteilung der Expert*innen) spielen hier also eine große Rolle. Auch der sprachlich-/kulturelle Unterschied fällt hier in allen Variablen sehr groß aus, wobei die Analyse der italienisch-

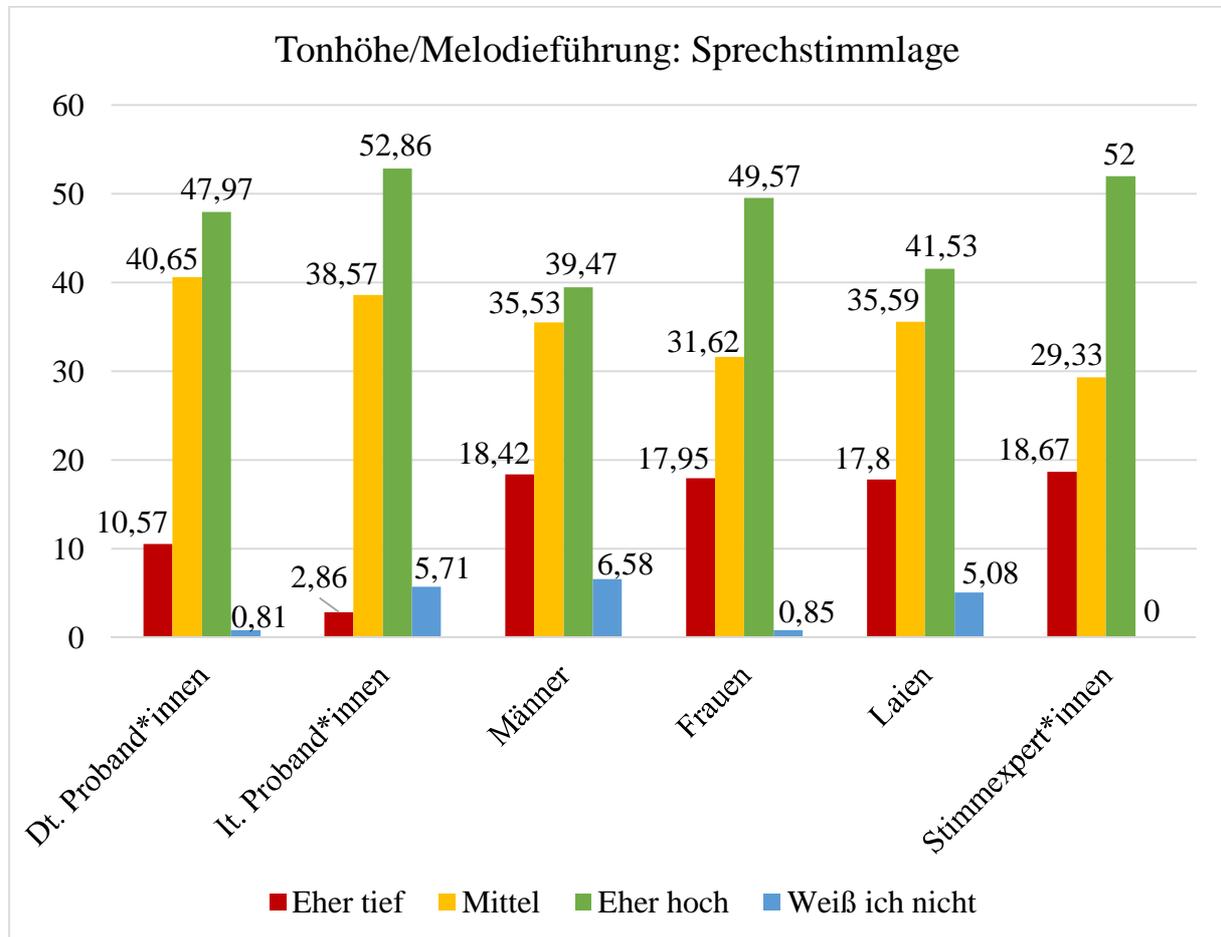
sprachigen Proband*innen wesentlich differenzierter ist als jene der deutschsprachigen. Gleichzeitig geben jedoch auch 7,14 Prozent der italienischsprachigen Gruppe an, die Frage nicht beantworten zu können, gefolgt von Männern und Laien.

Fazit

In Bezug auf die Akzentuierung werden eher starke und eher viele Betonungen, in Bezug auf den Sprechrhythmus ein eher fließender Rhythmus (*legato*) wahrgenommen. Bei der Beurteilung der Stärke der Akzentuierung gibt es keine wesentlichen Unterschiede hinsichtlich der Sprache/Kultur, des Geschlechts oder Berufes der Proband*innen, dieses Kriterium scheint am klarsten und einfachsten zu beurteilen sein. Die anderen beiden Eigenschaften des Merkmals Akzent/Rhythmus konnten zwar auch eindeutig bestimmt werden, weisen allerdings größere Differenzen in den Gruppierungen untereinander auf. Die Faktoren Sprache/Kultur, Geschlecht und Beruf spielen bei der Beurteilung der Häufigkeit der Akzentuierung und des Sprechrhythmus eine bedeutende Rolle, wobei die Stimmexpert*innen in allen Kategorien die differenziertere Beurteilung vornehmen. Bei der Analyse der Häufigkeit der Akzentuierung beurteilen die Männer differenzierter als die Frauen, ein sprachlich/kulturell relevanter Unterschied lässt sich nicht feststellen. Was die Perzeption und Beurteilung des Sprechrhythmus betrifft, so wird dieser von italienischsprachigen Proband*innen, den Stimmexpert*innen und den Frauen häufiger als „*legato*“ eingestuft als von den anderen.

3.3.4.1.2 Tonhöhe/Melodieführung

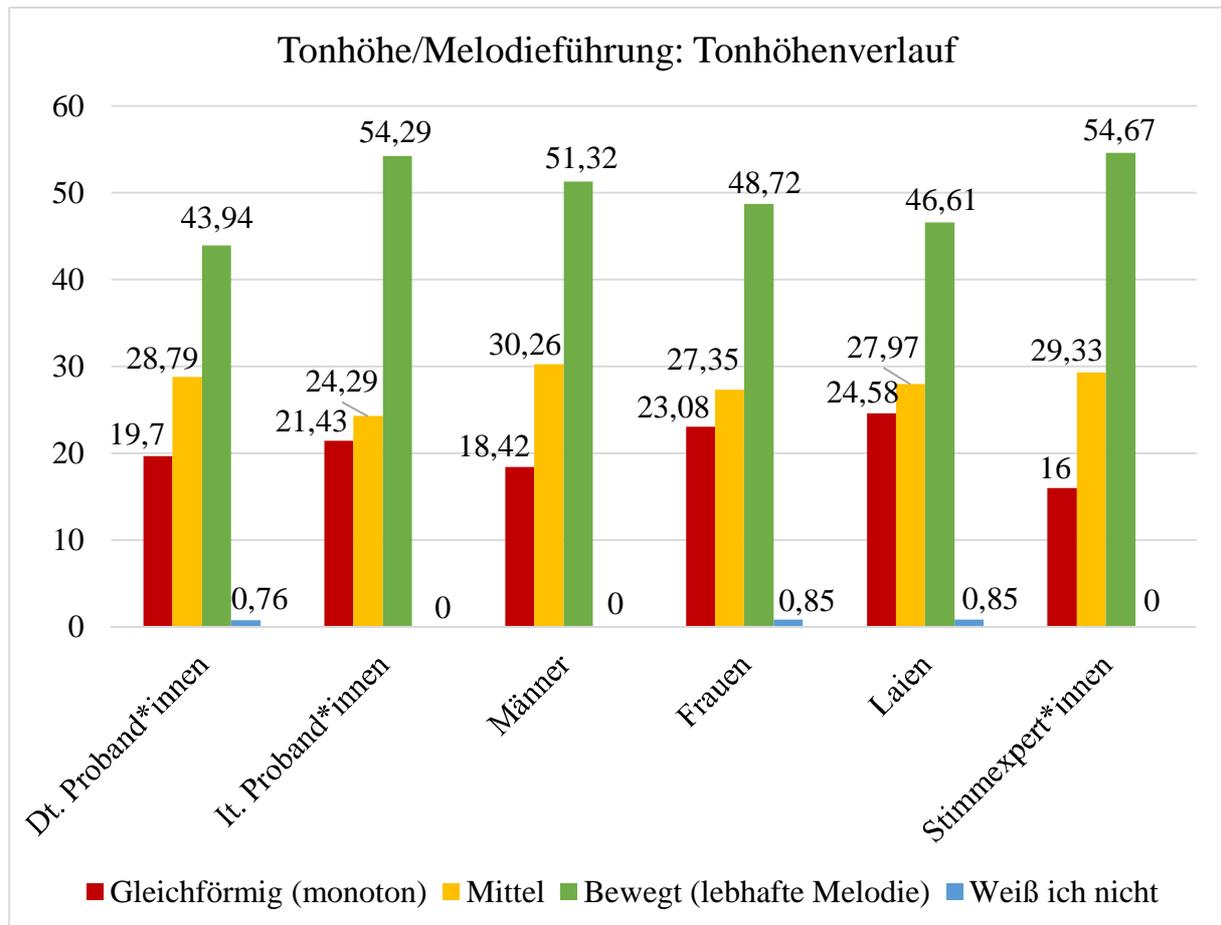
Sprechstimmlage



Das zweite prosodische Merkmal, das analysiert wurde, ist die Tonhöhe/Melodieführung. Dabei wurde zuerst die Sprechstimmlage eingeschätzt, da sie als zentrale Eigenschaft unmittelbar auf die Hörer*innen wirkt. Die bei den freudigen Sprechweisen verwendete Sprechstimmlage wird insgesamt als „eher hoch“ wahrgenommen, insbesondere von Seiten der italienischsprachigen Proband*innen und Stimmexpert*innen, gefolgt von den Frauen und deutschsprachigen Proband*innen. Innerhalb der italienischsprachigen Gruppe sind sich die Personen einig – die Sprechstimmlage ist eher hoch (52,86 Prozent) oder mittel (38,57 Prozent), nur 2,86 Prozent nehmen sie als „eher tief“ wahr. Dies ist ein deutlich klareres Ergebnis als bei den anderen Gruppierungen, in denen 10,57 Prozent (dt. Personen) bzw. ca. 18 Prozent (alle anderen Gruppierungen) diese Bewertung vornehmen. Sicher in der Beurteilung zeigen sich die Stimmexpert*innen, welche im Unterschied v. a. zu den italienischsprachigen Proband*innen und den Männern alle diese Frage beantworten können. Das Geschlecht spielt bei der Analyse also auch

eine Rolle, insgesamt weisen die Frauen eine klarere und sicherere Analyse auf als die Männer. Selbiges gilt für die Stimmexpert*innen gegenüber den Laien. Geringer fallen hingegen die sprachlich-/kulturell bedingten Differenzen aus.

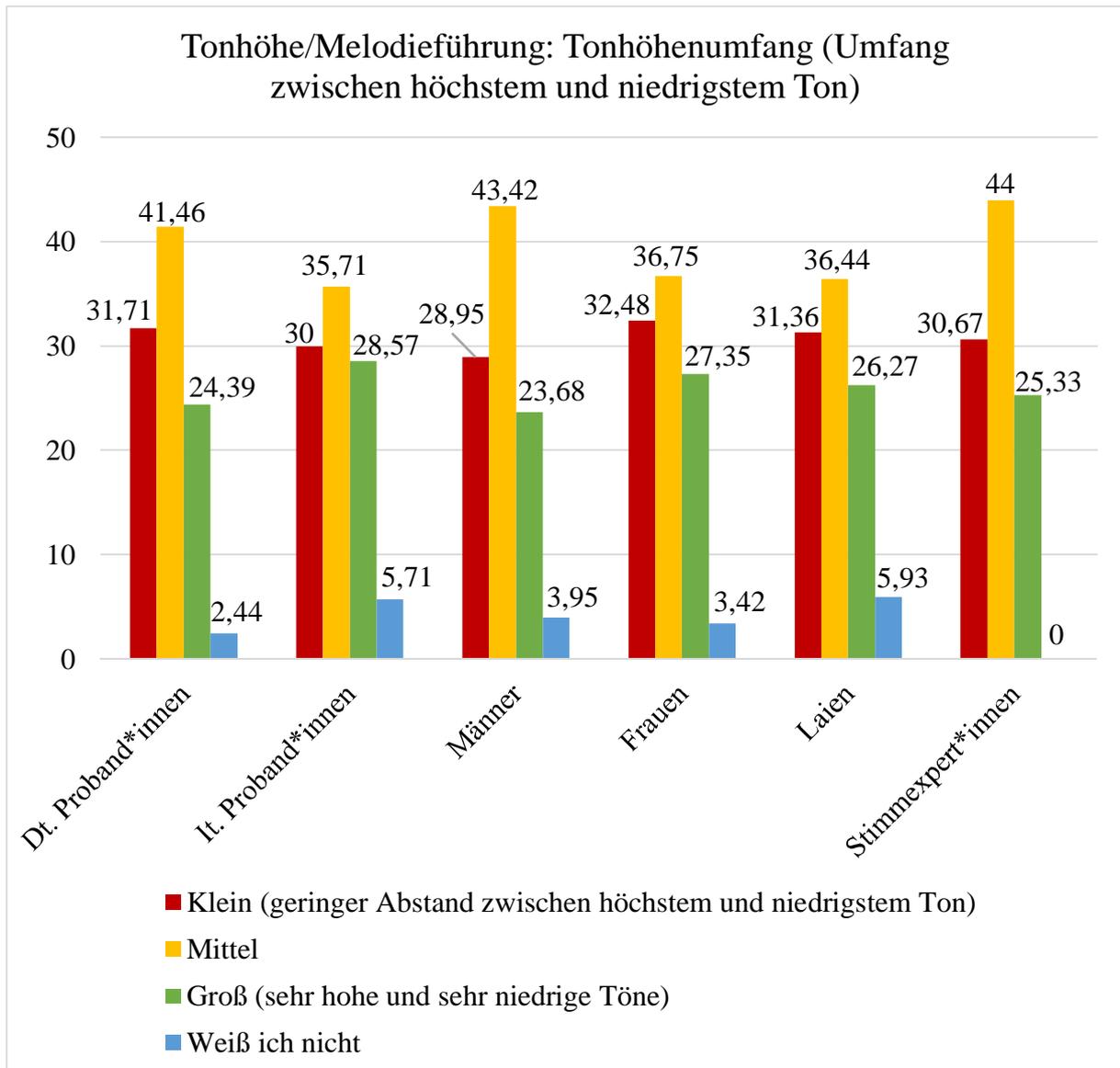
Tonhöhenverlauf



In Bezug auf diese Eigenschaft ist das Ergebnis aller Proband*innen sehr ähnlich und klar. Die überwiegende Mehrheit der Personen definiert den Tonhöhenverlauf der freudigen Sprechweisen als bewegt (lebhaft Melodie). Insgesamt scheint diese Eigenschaft sehr gut perzipiert worden zu sein und nur ein verschwindend geringer Prozentsatz der Proband*innen konnte die Frage nicht beantworten. Am eindeutigsten ist das Ergebnis der Stimmexpert*innen, gefolgt von denen der italienischsprachigen Personen und den Männern. Der Unterschied zwischen den Geschlechtern ist jedoch vernachlässigbar. Etwas relevanter sind die Differenzen in Bezug auf die Sprache/Kultur (54,29 Prozent der italienischsprachigen, aber nur 43,94 Prozent der deutschsprachigen Proband*innen halten den Tonhöhenverlauf für bewegt) und den Beruf (das

Ergebnis der Stimmexpert*innen ist differenzierter als jenes der Laien). Insgesamt sind die Ergebnisse aber recht vergleichbar und eindeutig.

Tonhöhenumfang



Als dritte Eigenschaft des Merkmals Tonhöhe/Melodieführung wurde der Tonhöhenumfang, also der Umfang zwischen dem höchsten und dem niedrigsten Ton, analysiert. Der Tonhöhenumfang ist deutlich schwieriger zu perzipieren und daher zu analysieren, was sich auch in den Ergebnissen widerspiegelt. Insgesamt wurde ein mittlerer Tonhöhenumfang wahrgenommen, gefolgt von der Bewertung „klein“. Die Abgrenzungen zwischen den jeweils anderen Bewertungen sind jedoch gering. Am klarsten ist die Analyse der Stimmexpert*innen, gefolgt von jener der Männer. Zwischen den deutsch- und italienischsprachigen Proband*innen zeichnet

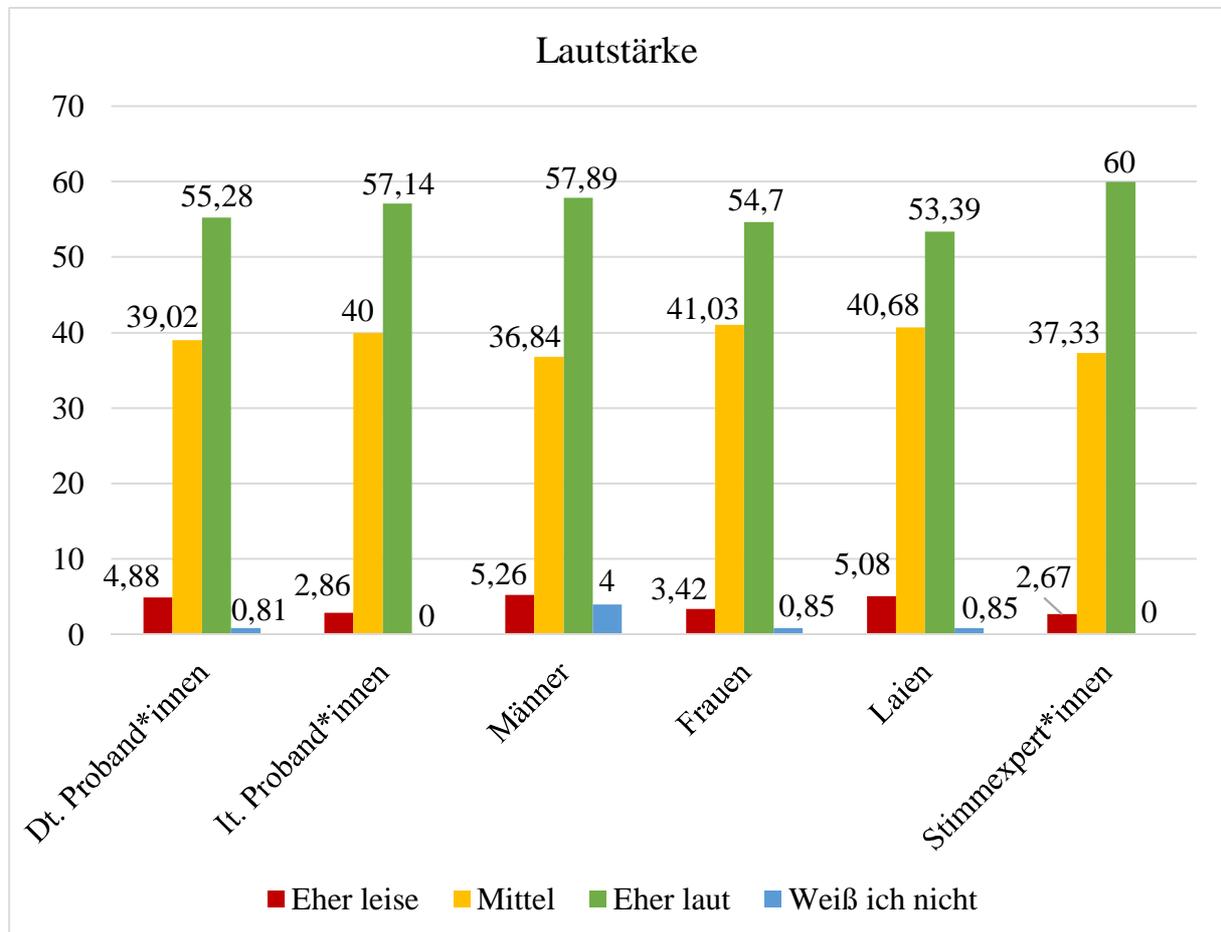
sich ein leichter sprachlich-/kultureller Unterschied ab, wobei erstere ein etwas klareres Ergebnis aufweisen. Selbiges gilt auch für die Männer im Vergleich mit den Frauen, wobei der geschlechterbedingte Unterschied kaum relevant ausfällt. Etwas mehr ins Gewicht fällt, wie bereits erwähnt, der Faktor Beruf. Während alle Stimmexpert*innen die Frage beantworten können, tun dies knapp 6 Prozent der Laien nicht.

Fazit

Was die Tonhöhe/Melodieführung der freudigen Sprechweisen betrifft, so wurden die Sprechstimmlage insgesamt als eher hoch, der Tonhöhenverlauf als bewegt und der Tonhöhenumfang als mittelmäßig perzipiert. Sprach-/kulturspezifische, geschlechter- und berufsbezogene Differenzen in der Analyse zeigen sich in unterschiedlichem Ausmaß je nach zu bewertender Kategorie. Ist die Bewertung der Frauen in Bezug auf die Wahrnehmung der Sprechstimmlage klarer als jene der Männer, dreht sich der Spieß in Bezug auf die Analyse des Tonhöhenumfangs um. Auch sprach-/kulturbedingte Unterschiede lassen sich nicht in jeder Kategorie gleich finden und treten v. a. bei der Perzeption des Tonhöhenverlaufs auf, der von den italienischsprachigen Proband*innen eindeutig öfters als bewegt wahrgenommen wurde. Am relevantesten ist der Unterschied zwischen den Stimmexpert*innen und Laien bzw. anderen Gruppierungen. Die Analyseergebnisse der Expert*innen sind eindeutig differenzierter in allen drei Kategorien zur Beschreibung der Tonhöhe/Melodieführung. Der Faktor Beruf dürfte hier also die größte Rolle spielen. Der Tonhöhenverlauf konnte, im Vergleich mit den anderen Eigenschaften, am klarsten von allen Proband*innen analysiert werden.

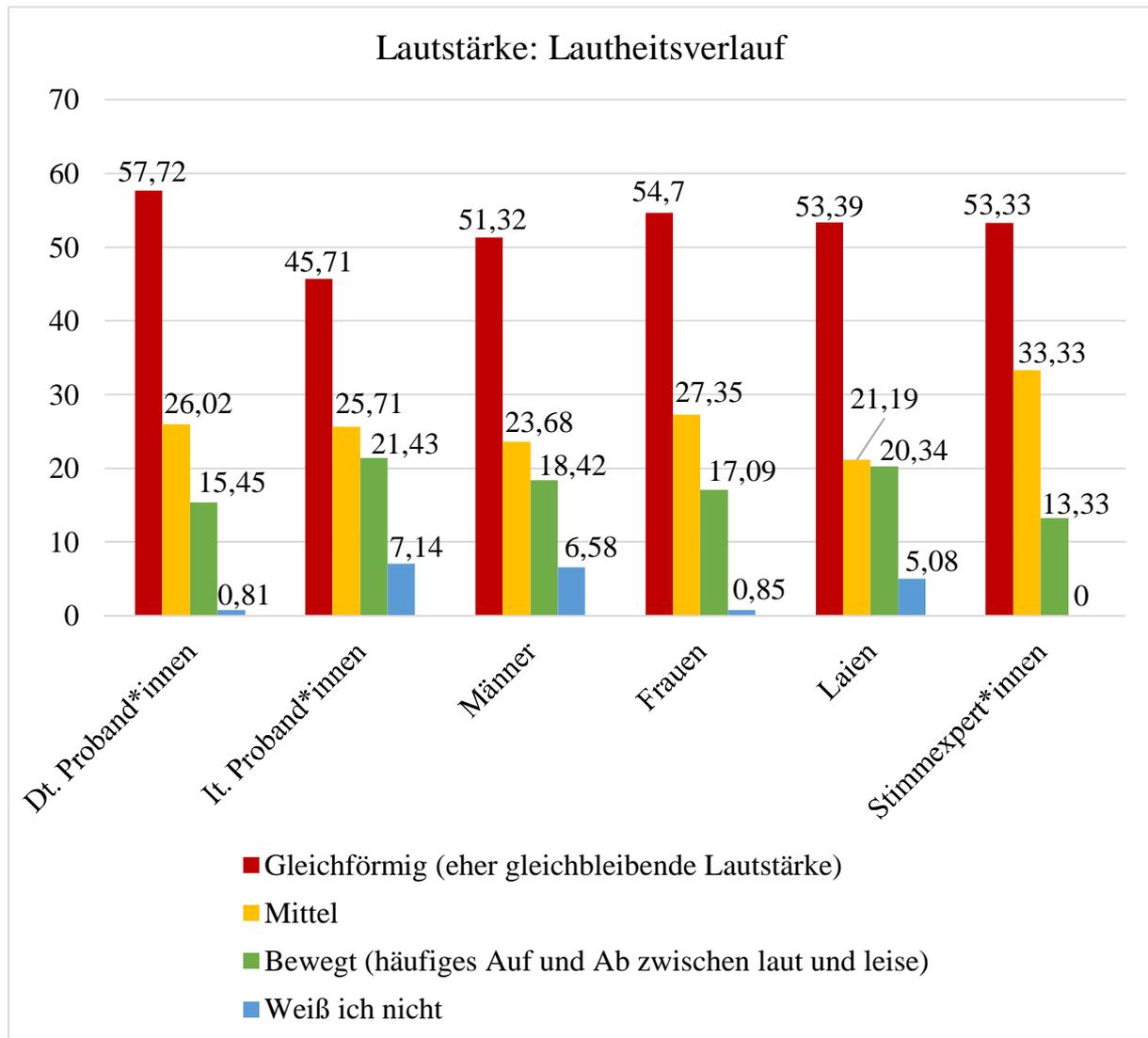
3.3.4.1.3 Lautstärke

Lautstärke



Zur Perzeption der Lautstärke gilt zunächst festzuhalten, dass die freudigen Sprechweisen von allen als eher laut empfunden wurden. Dies gilt insbesondere für die Stimmexpert*innen, weniger für die Laien. Der berufsbedingte Unterschied lässt sich hier an einem insgesamt klareren Ergebnis von Seiten der Expert*innen ablesen, stellt jedoch nur eine Tendenz dar, da er sehr gering ist. Tatsächlich sind die Ergebnisse aller Proband*innen absolut vergleichbar und eindeutig. Bis auf einen leichten berufsbezogenen Unterschied lassen sich keine relevanten Differenzen (Sprache/Kultur oder Geschlecht) in den verschiedenen Gruppierungen ablesen.

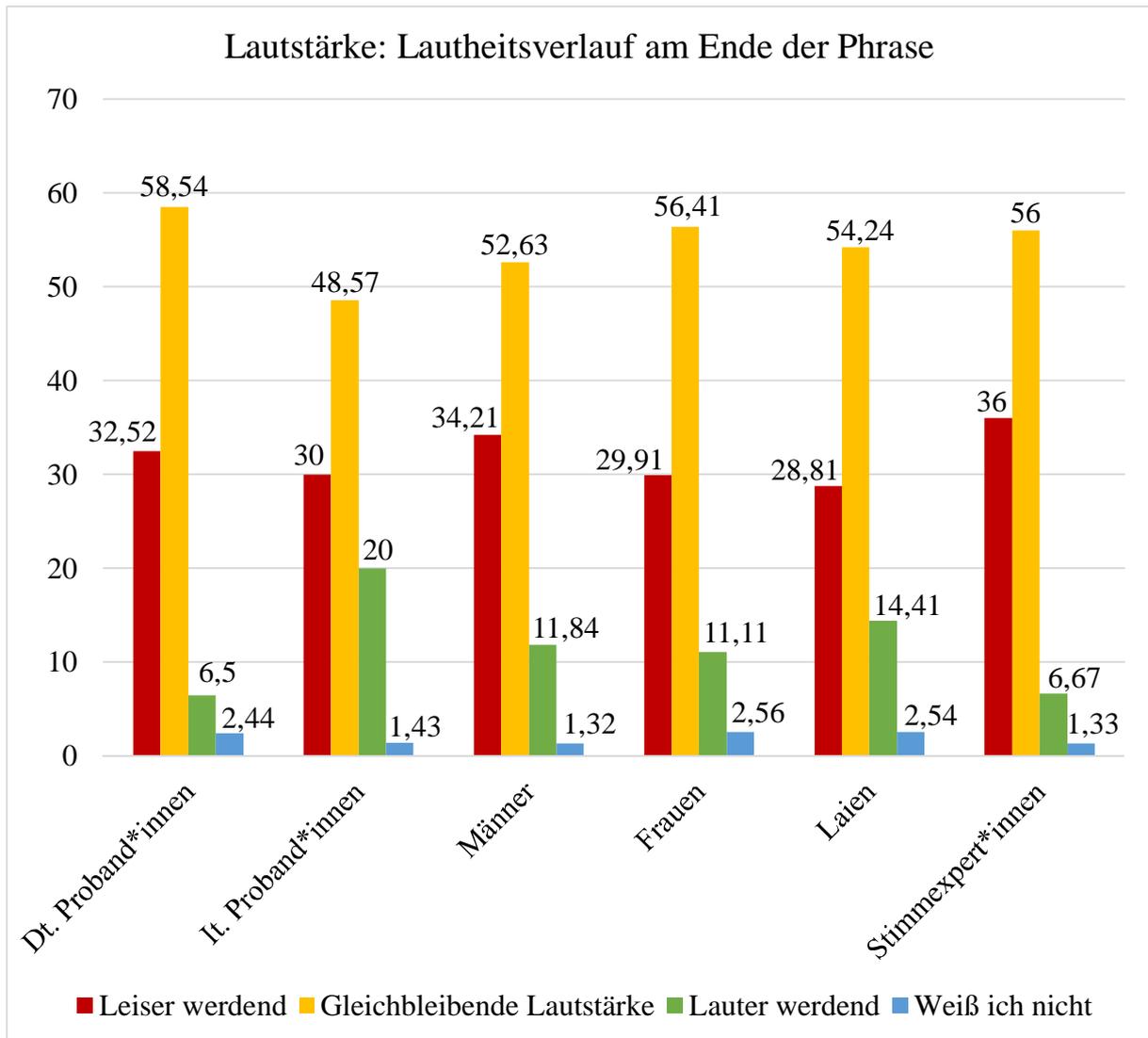
Lautheitsverlauf



Der Lautheitsverlauf der freudigen Sprechweisen wurde insgesamt als gleichförmig (eher gleichbleibende Lautstärke) wahrgenommen. Hier lassen sich jedoch einige Unterschiede in der Analyse der einzelnen Proband*innen ablesen. Während ca. 58 Prozent der deutschsprachigen Personen den Lautheitsverlauf als gleichförmig perzipieren, sind es in der italienischsprachigen Gruppe hingegen nur ca. 46 Prozent. Hier wird von mehreren Proband*innen auch die Bewertung „bewegt“ oder „weiß ich nicht“ gewählt. Es liegt aber nicht nur ein sprachlich/kultureller Unterschied vor. Auch der Faktor Geschlecht spielt eine, wenngleich geringe, Rolle – so lässt sich feststellen, dass mehr Männer Schwierigkeiten darin hatten, die Frage zu beantworten, als Frauen. Selbiges gilt auch für die Laien im Vergleich zu den Stimmexpert*innen. Die Analyse dieser beiden Gruppen ergibt aber noch relevantere Unterscheidungen. Während gleich viele Laien den Lautheitsverlauf als „mittel“ oder „bewegt“ wahrnehmen, gibt es bei den

Expert*innen eine deutlich höhere Bewertung für „mittel“ (33,33 Prozent) und nur 13,33 Prozent nehmen den Lautheitsverlauf als bewegt wahr.

Lautheitsverlauf am Ende der Phrase



Ergänzend zur Beurteilung des Lautheitsverlaufs über die gesamte Sprechäußerung hinweg, wurde hier der Fokus auf den Lautheitsverlauf am Ende der Phrase gelegt. Dieser wird insgesamt so wahrgenommen, dass die Lautstärke gleich bleibt. Mit geringerer Häufigkeit wird er als leiser werdend und nur von wenigen als lauter werdend bezeichnet. Allerdings lassen sich sehr große sprachlich/kulturelle Differenzen und ein relevanter berufsbezogener Unterschied in der Perzeption feststellen. Deutlich mehr deutschsprachige Proband*innen empfinden eine gleichbleibende Lautstärke am Ende der Phrase (58,54 Prozent dt. vs. 48,57 Prozent it.). Dafür nehmen die italienischsprachigen Proband*innen den Lautheitsverlauf viel öfters als „lauter

werdend“ wahr (20 Prozent it. vs. 6,5 Prozent dt.). Hier lässt sich ein Zusammenhang mit der Beurteilung des Lautheitsverlaufs über die gesamte Äußerung hinweg erkennen, wie er im Diagramm zuvor beschrieben worden war. Die 20 Prozent der italienischsprachigen Proband*innen, die den Lautheitsverlauf am Ende der Phrase als lauter werdend bezeichnen, stehen auch in starkem Kontrast zur Einschätzung der Stimmexpert*innen (6,67 Prozent). Der berufsbezogene Unterschied drückt sich generell in einer etwas differenzierteren Analyse aus. 36 Prozent der Expert*innen nehmen den Lautheitsverlauf als „leiser werdend“ wahr und stehen damit im Kontrast zu den Laien, von denen nur 28,81 Prozent diese Beurteilung vornehmen.

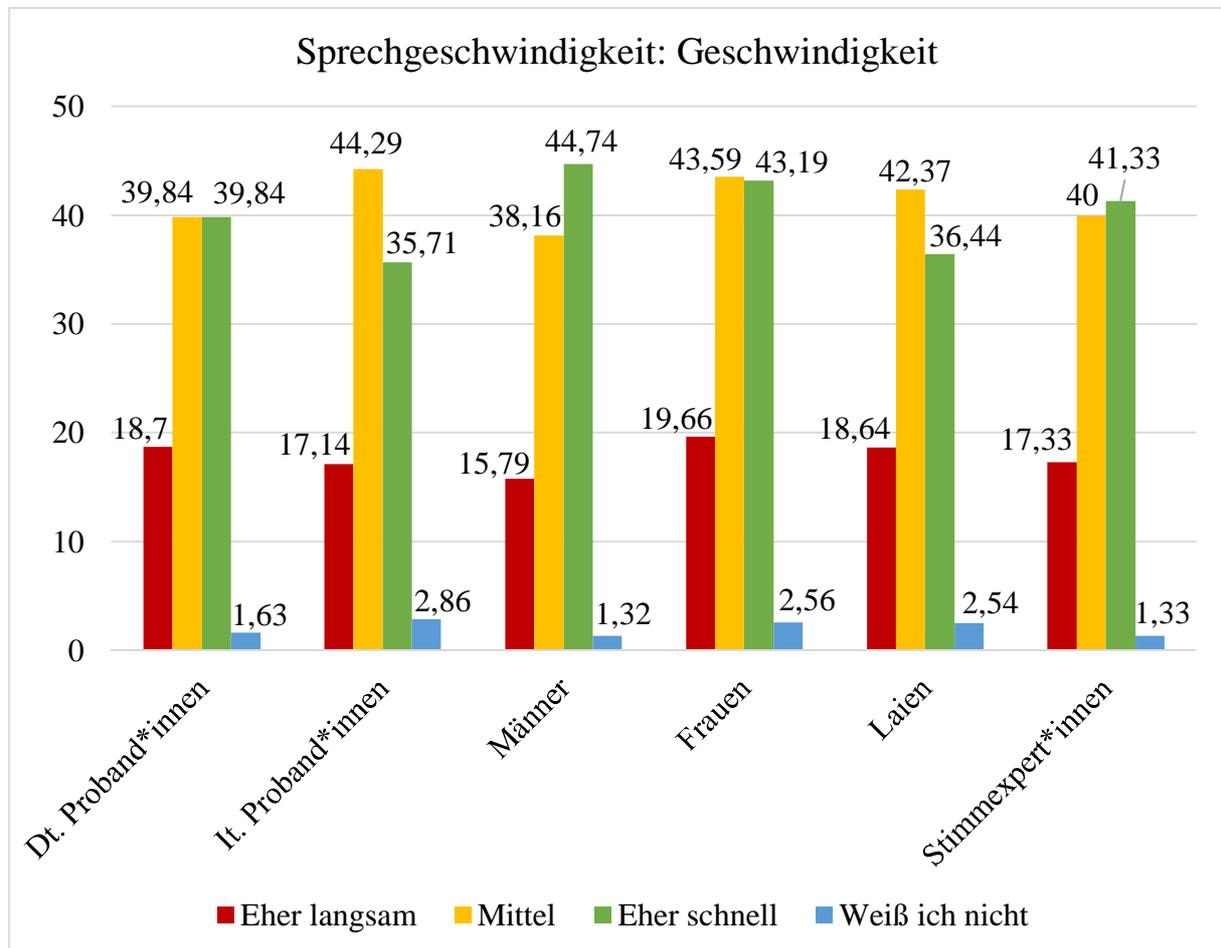
Ein geschlechtsspezifischer Unterschied in der Perzeption lässt sich hingegen nicht feststellen.

Fazit

Die Lautstärke als solche konnte sehr gut und klar von allen Proband*innen beurteilt werden, hier lassen sich keine sprachlich/kulturellen oder geschlechterbezogenen Unterschiede in der Perzeption feststellen und auch der berufsbedingte Unterschied (Stimmexpert*innen weisen eine etwas differenziertere Beurteilung auf als Laien) ist eher gering. Diese Eigenschaft des prosodischen Merkmals Lautstärke wurde also von allen Proband*innen klar wahrgenommen und vergleichbar beurteilt. Anders ist es bei den Eigenschaften des Lautheitsverlaufs und des Lautheitsverlaufs am Ende der Phrase. Hier sticht besonders der sprachlich/kulturelle Unterschied ins Auge. Die italienischsprachigen perzipieren im Vergleich zu den deutschsprachigen Proband*innen öfters ein bewegteres Auf und Ab zwischen laut und leise und einen lauter werdenden Verlauf gegen Ende der Phrase. Auch ein berufsbedingter Unterschied ist ersichtlich, wobei die Stimmexpert*innen insgesamt eine differenziertere Beurteilung aufweisen. Eher gering ist der geschlechtsspezifische Unterschied in der Perzeption, hier kann maximal die Tendenz beobachtet werden, dass die Frauen eine klarere Differenzierung vornehmen als die Männer.

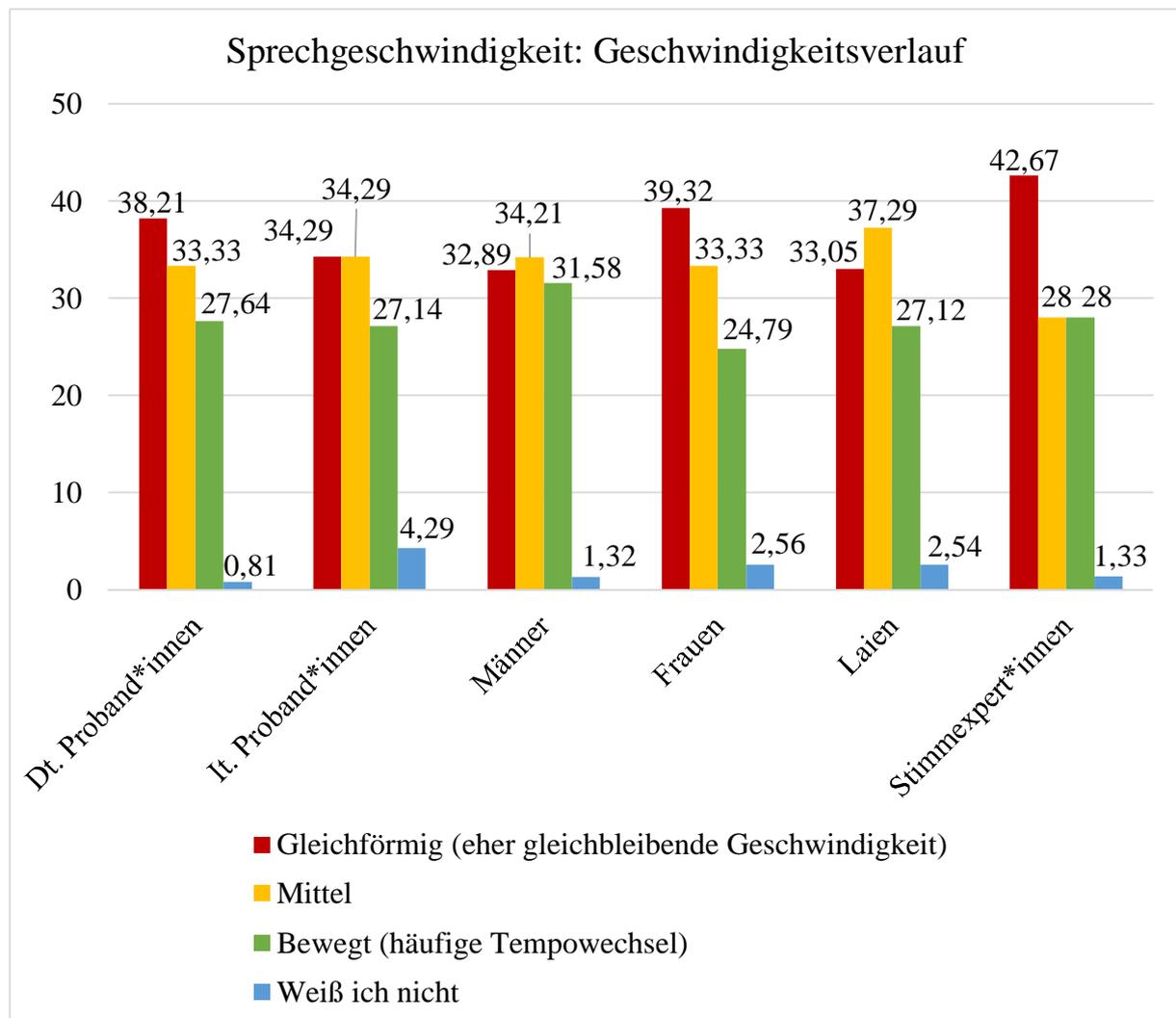
3.3.4.1.4 Sprechgeschwindigkeit

Geschwindigkeit



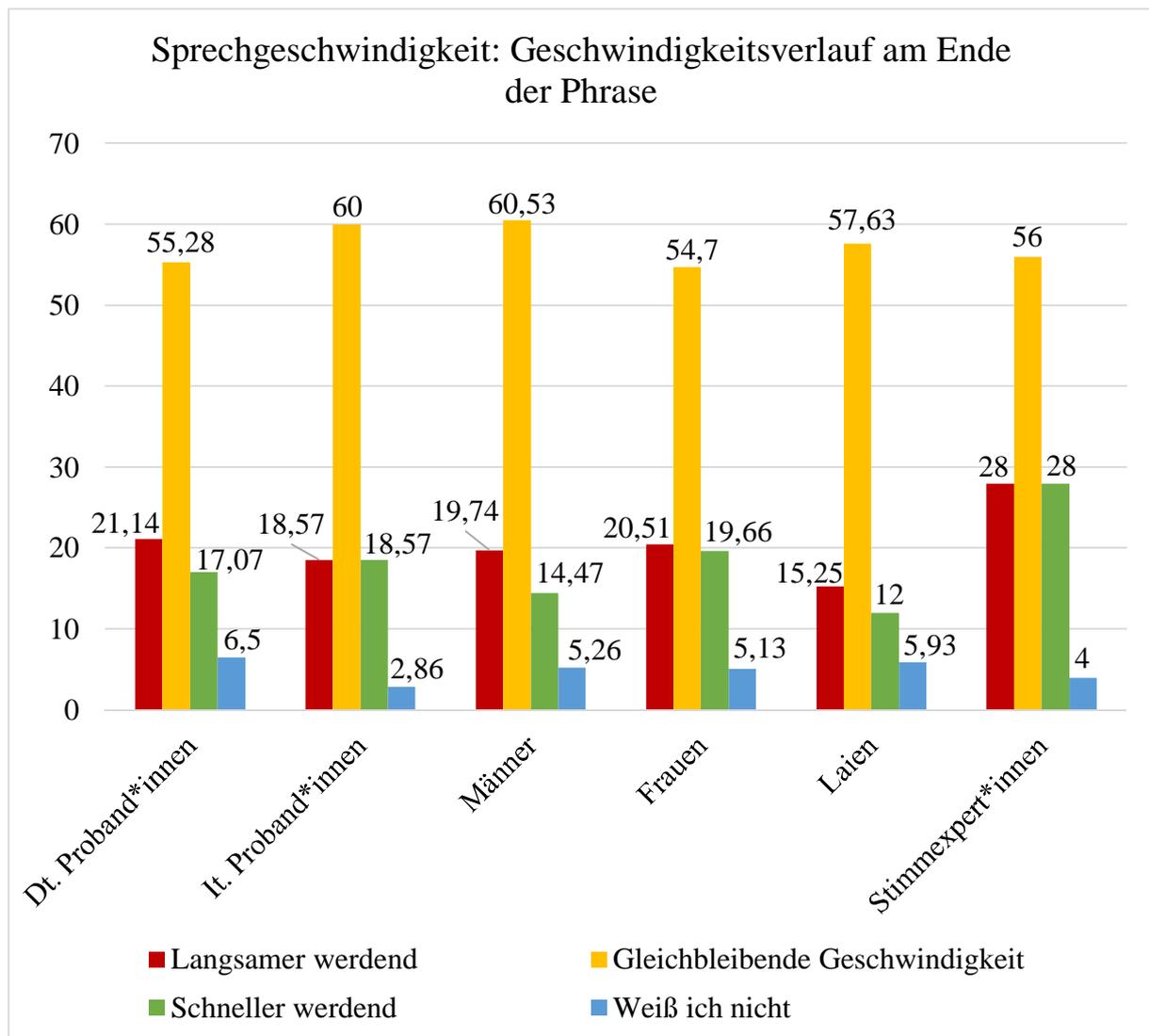
Bei der Beurteilung der Sprechgeschwindigkeit der freudigen Sprechweisen liegt kein eindeutiges Ergebnis vor, so wird diese Eigenschaft insgesamt als „mittel“ bzw. „eher schnell“ perzipiert. Keine klare Unterscheidung dieser Kategorien findet sich in der Analyse der Stimmexpert*innen, der Frauen und der deutschsprachigen Proband*innen. Die Mehrheit der italienischsprachigen Proband*innen und der Laien jedoch hält die Geschwindigkeit für mittelmäßig, die Mehrheit der Männer hingegen für eher schnell. Nur ca. 17 Prozent aller Befragten halten die Geschwindigkeit für eher langsam, hier zeichnen sich auch keine relevanten Unterschiede durch Sprache/Kultur, Geschlecht oder Beruf ab.

Geschwindigkeitsverlauf



Auch was den Geschwindigkeitsverlauf betrifft, so liegt kein eindeutiges Ergebnis vor. Insgesamt wird er als gleichförmig (mit gleichbleibender Geschwindigkeit) wahrgenommen, bei den Laien und den Männern überwiegt jedoch knapp die Beurteilung des Geschwindigkeitsverlaufs als mittelmäßig. Sie nehmen ihn also tendenziell als etwas bewegter wahr. Die Unterscheidung zwischen „mittel“ und „bewegt“ ist nicht sehr ausgeprägt und überlappt sich bei den Stimmexpert*innen sogar. Die Wahrnehmung des Geschwindigkeitsverlaufs scheint individuell sehr unterschiedlich zu sein. Eine relevante sprachlich/kulturell bedingte Differenz lässt sich nicht herauslesen. Das Geschlecht scheint eher prägnant zu sein, hier nehmen die Frauen den Geschwindigkeitsverlauf als etwas differenzierter wahr als die Männer. Selbiges gilt auch für die Stimmexpert*innen gegenüber den Laien.

Geschwindigkeitsverlauf am Ende der Phrase



Einig sind sich alle Proband*innen hingegen über den Geschwindigkeitsverlauf am Ende der Phrase, welcher insgesamt eindeutig als gleichbleibend perzipiert wird. Ein viel geringerer Prozentsatz von Personen nimmt ihn als langsamer werdend und noch weniger Personen als schneller werdend wahr. Diese Eigenschaft der Sprechgeschwindigkeit kann im Vergleich mit den anderen beiden (Geschwindigkeit und Geschwindigkeitsverlauf) von mehr Personen nicht beurteilt werden, so beantworten etwa 6,5 Prozent der deutschsprachigen Proband*innen die Frage mit „weiß ich nicht“. Insgesamt kristallisiert sich bei der Perzeption keine klare sprach-/kulturbedingte Differenz heraus. Auch der Faktor Geschlecht spielt nur eine geringe Rolle, wobei das Ergebnis der Männer etwas klarer ist als jenes der Frauen. Bei den Laien und Stimmexpert*innen verhält es sich hier so, dass wesentlich mehr Expert*innen den Geschwindigkeitsverlauf am Ende der Phrase für „langsamer werdend“ bzw. „schneller werdend“ halten (je 28

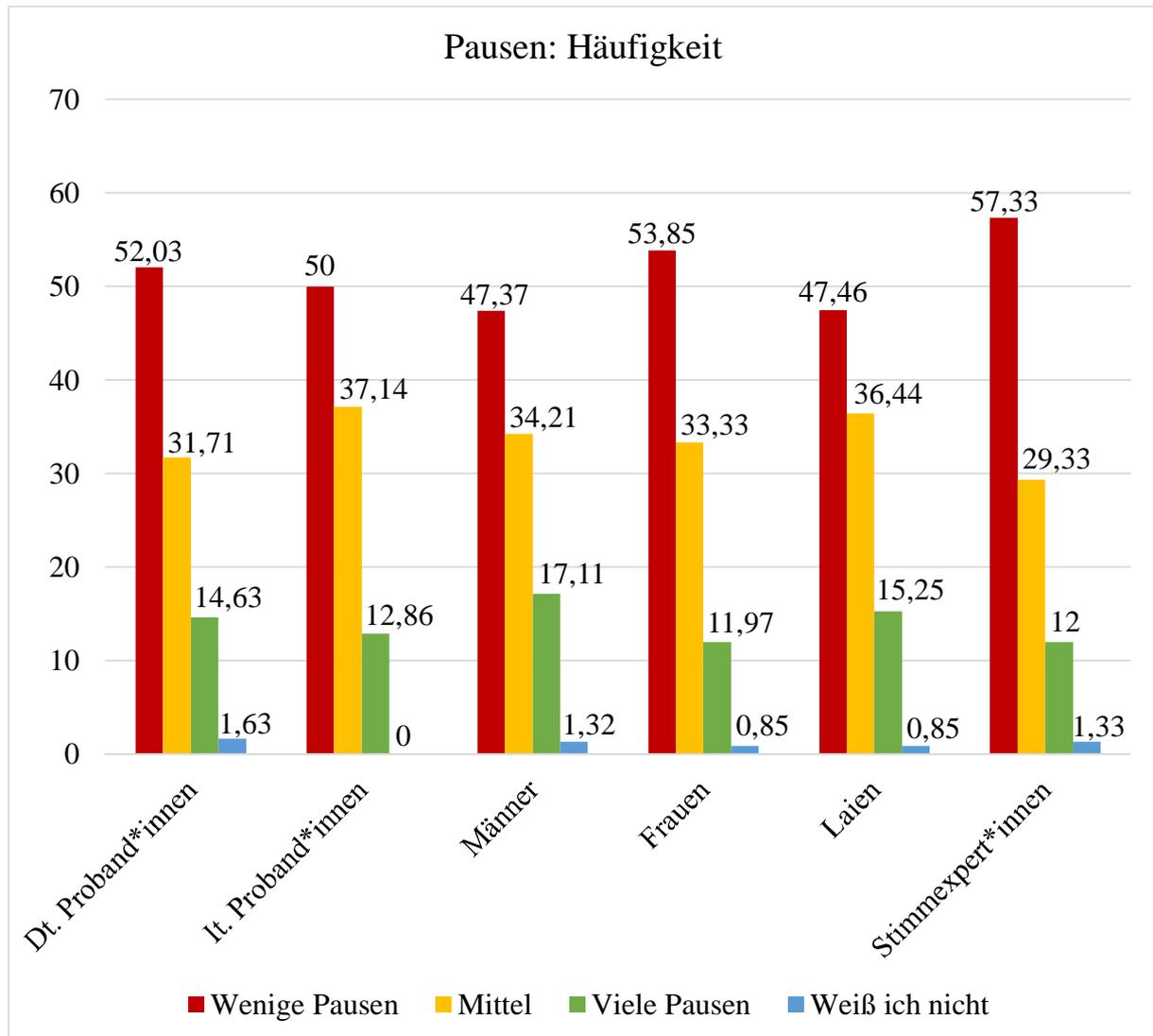
Prozent). Somit zeigen die Laien hier erstmalig eine eindeutig präzisere Beurteilung auf als die Expert*innen.

Fazit

Insgesamt werden die freudigen Sprechäußerungen als eher schnell bzw. mittelmäßig schnell und mit einem gleichförmigen bzw. mittelmäßigen Geschwindigkeitsverlauf wahrgenommen. Ein eindeutiges Ergebnis zeigt die Beurteilung des Geschwindigkeitsverlaufs am Ende der Phrase, hier wird von allen Proband*innen eine gleichbleibende Geschwindigkeit festgestellt. Eine relevante sprachlich/kulturell geprägte Differenz in der Perzeption der Sprechgeschwindigkeit lässt sich nicht erkennen, allerdings weisen die italienischsprachigen Proband*innen eine tendenziell klarere Beurteilung auf. Auch ein geschlechterbedingter Unterschied ist nicht eindeutig feststellbar – die Geschwindigkeit und der Geschwindigkeitsverlauf am Ende der Phrase werden tendenziell von den Männern differenzierter beurteilt als von den Frauen, umgekehrt verhält es sich bei der Analyse des Geschwindigkeitsverlaufs. Was den Faktor Beruf betrifft, so lassen sich hier keine allgemeinen Schlüsse ziehen, da sich einmal das Ergebnis der Expert*innen (Geschwindigkeitsverlauf), einmal jenes der Laien (Geschwindigkeitsverlauf am Ende der Phrase) als differenzierter darstellt.

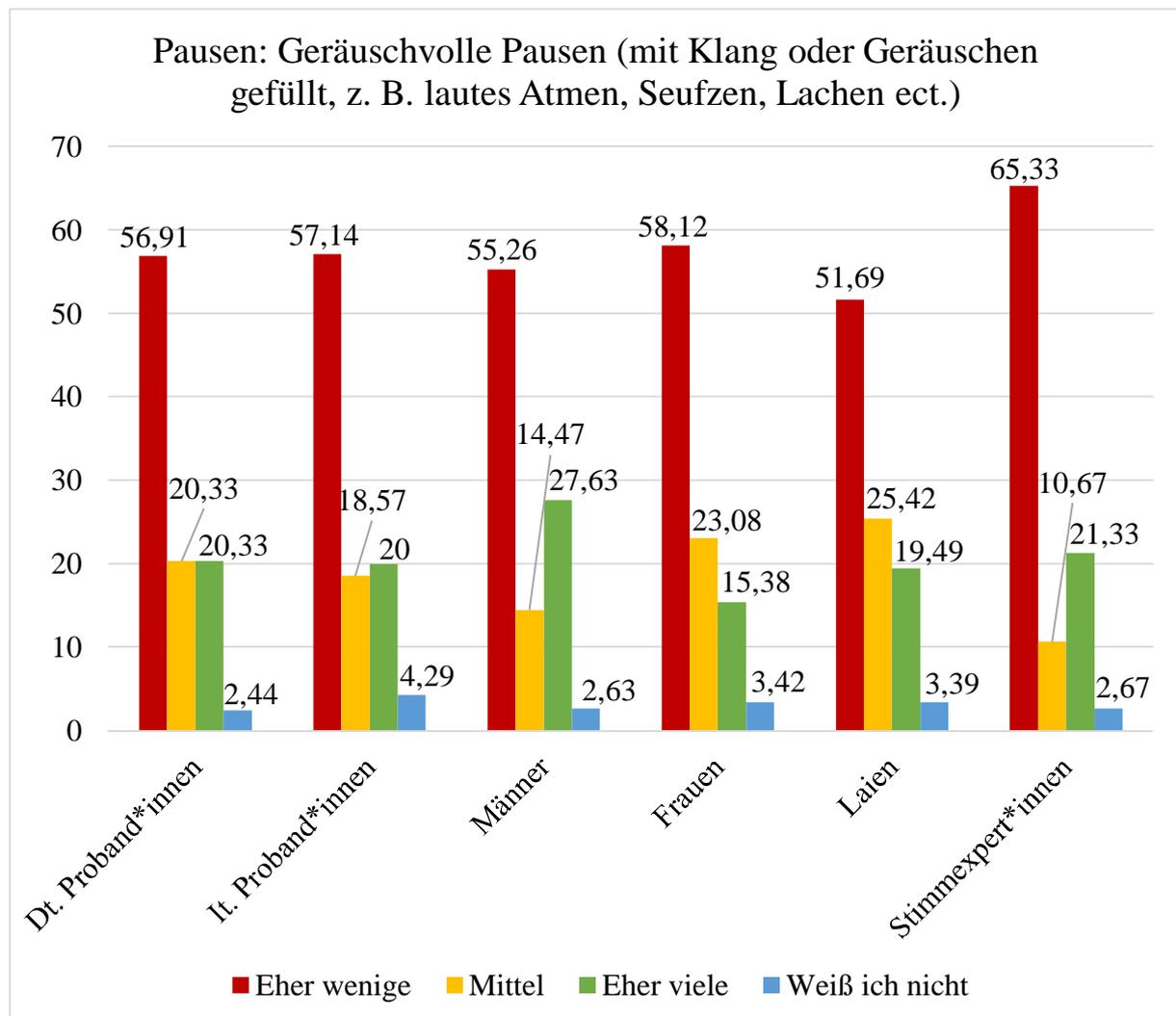
3.3.4.1.5 Pausen

Häufigkeit



Insgesamt wurde die Häufigkeit der Pausen einheitlich als gering wahrgenommen. Circa 29 bis 37 Prozent der Personen geben eine mittlere Häufigkeit an und nur 12 bis 17 Prozent vernehmen viele Pausen. Besonders klar zum Ausdruck kommt dieses Ergebnis bei den Stimmexpert*innen. Weniger differenziert zeigt es sich hingegen bei den Laien. Neben dem deutlich ausgeprägten berufsbedingten Unterschied zeigt sich auch eine leichte geschlechterbezogene Differenz, wobei die Analyse der Frauen etwas differenzierter ist als jene der Männer. Ein relevanter sprach-/kulturspezifischer Unterschied bei der Perzeption und Analyse dieser Eigenschaft der Pausen lässt sich hingegen nicht ausmachen.

Geräuschvolle Pausen



Als zweite Eigenschaft der Pausen wurde deren Geräuschanteil (Klänge und Geräusche diverser Art, z. B. Lachen, lautes Atmen, Seufzen usw.) analysiert, den der Großteil der Proband*innen als gering wahrnimmt. Bei den weiteren Abstufungen, nämlich mittlere Häufigkeit und „eher viele“ geräuschvolle Pausen gibt es kein klares Ranking. Während die Männer (27,63 Prozent) und die Stimmexpert*innen (21,33 Prozent) öfters die Angabe „eher viele“ als „mittlere“ machen, nehmen die Frauen und die Laien öfters eine „mittlere“ Häufigkeit wahr. So ist das Ergebnis der Frauen in sich etwas stimmiger als jenes der Männer. Bei den deutschsprachigen und italienischsprachigen Proband*innen hingegen lässt sich kein Unterschied feststellen, die Angaben sind fast identisch. Auch von einer berufsbedingten differenzierteren Einschätzung kann kaum gesprochen werden, da zwar 65,33 Prozent der Expert*innen im Vergleich zu nur 51,69 Prozent der Laien „eher wenige“ geräuschvolle Pausen perzipieren, sich aber dafür danach die Einschätzung umdreht (zweithäufigste Angabe bei den Expert*innen: „eher viele“).

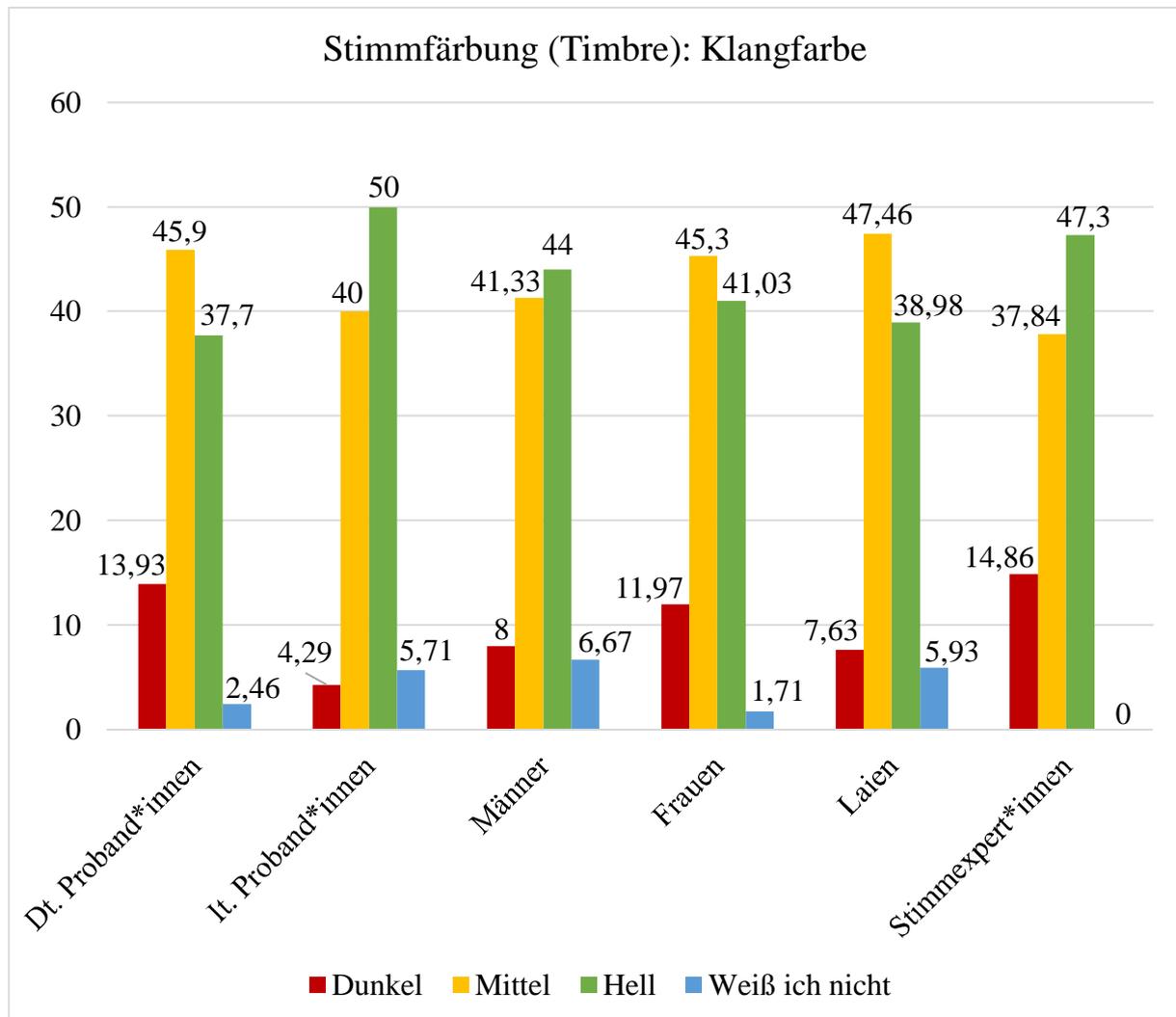
danach erst „mittel“ – bei den Laien ist umgekehrt die zweithäufigste Angabe „mittel“ und danach erst „eher viele“).

Fazit

Insgesamt werden eine geringe Häufigkeit der Pausen und eher wenige geräuschvolle Pausen perzipiert. Dieses Ergebnis ist eindeutig und klar ablesbar. In der genauen Einschätzung gibt es dann Differenzen in den unterschiedlichen Gruppierungen, wobei in beiden Fällen eine leichte geschlechterspezifische Tendenz auffällt. Das Analyseergebnis der Frauen ist etwas stimmiger und klarer als jenes der Männer. Selbiges kann nur im Fall Bestimmung der Häufigkeit der Pausen auch von den Stimmexpert*innen im Vergleich mit den Laien genannt werden. Sprachlich/kulturell bedingte Perzeptionsunterschiede sind jedoch keine auszumachen.

3.3.4.1.6 Stimmfärbung (Timbre)

Klangfarbe



Die Stimmfärbung als komplexes prosodisches Merkmal umfasst mehrere Eigenschaften. Zunächst wurde die Klangfarbe der freudigen Sprechweisen analysiert. Das Ergebnis unterscheidet sich sehr aufgrund der Faktoren Sprache/Kultur, Beruf und Geschlecht der Proband*innen. Insgesamt wird die Klangfarbe (Timbre) als „mittel“ bzw. „hell“ wahrgenommen. Im Folgenden sollen die Differenzen in Bezug auf oben genannte Einflussfaktoren herausgearbeitet werden sollen:

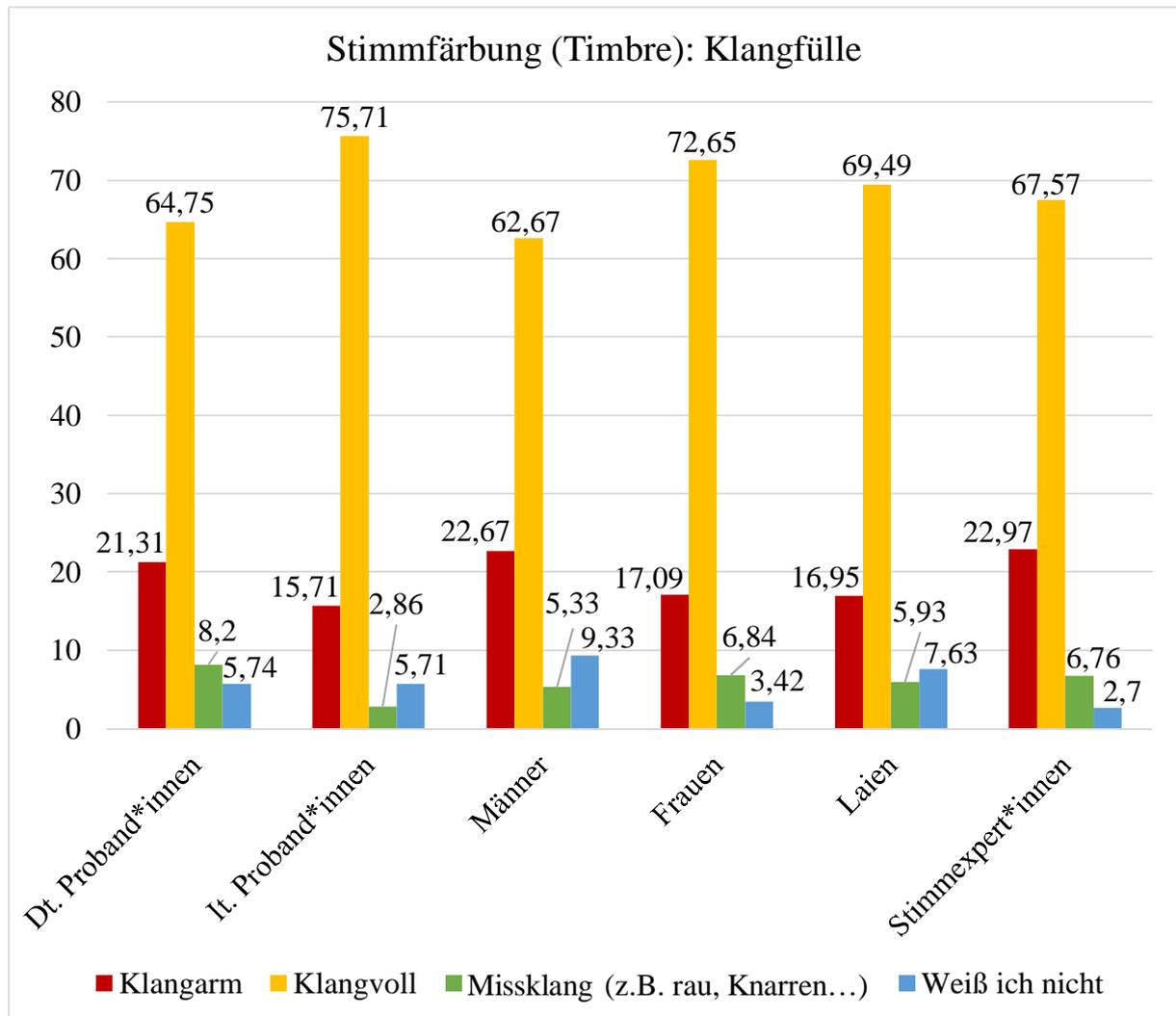
Sprache/Kultur: Der Großteil der deutschsprachigen Proband*innen nimmt die Klangfarbe als „mittel“ wahr, nur 37,7 Prozent als hell. Bei den italienischsprachigen Proband*innen hingegen nimmt der Großteil die Klangfarbe als hell wahr (50 Prozent). Nur 4,29 Prozent derselben empfindet die Klangfarbe als dunkel im Vergleich zu den knapp 14 Prozent der deutschsprachigen

Proband*innen. Insgesamt lässt sich also feststellen, dass die italienischsprachige Gruppe die Klangfarbe als deutlich heller empfindet als die deutschsprachige Gruppe.

Geschlecht: Zwar zeigt sich hier ein leichter Unterschied, dass die Männer die Klangfarbe insgesamt als etwas heller beurteilen als die Frauen, aber die Differenzen sind kaum relevant. Zudem können mehr Männer (6,67 Prozent) als Frauen (1,71 Prozent) diese Frage nicht beantworten.

Beruf: Ein relevanter Unterschied lässt sich in der Beurteilung der Stimmexpert*innen im Vergleich zu jener der Laien erkennen. Erstere nehmen die Klangfarbe als deutlich heller wahr, gleichzeitig gibt aber auch ein größerer Teil (14,86 Prozent) der Expert*innen im Vergleich zu den Laien (7,63 Prozent) an, die Klangfarbe als „dunkel“ zu empfinden. Die Bewertung „hell“ haben 47,3 Prozent der Expert*innen und 38,98 Prozent der Laien vorgenommen, die Bewertung „mittel“ hingegen 47,46 Prozent der Laien und nur 37,84 Prozent der Expert*innen. Mehr Laien als Expert*innen können die Frage außerdem nicht beantworten.

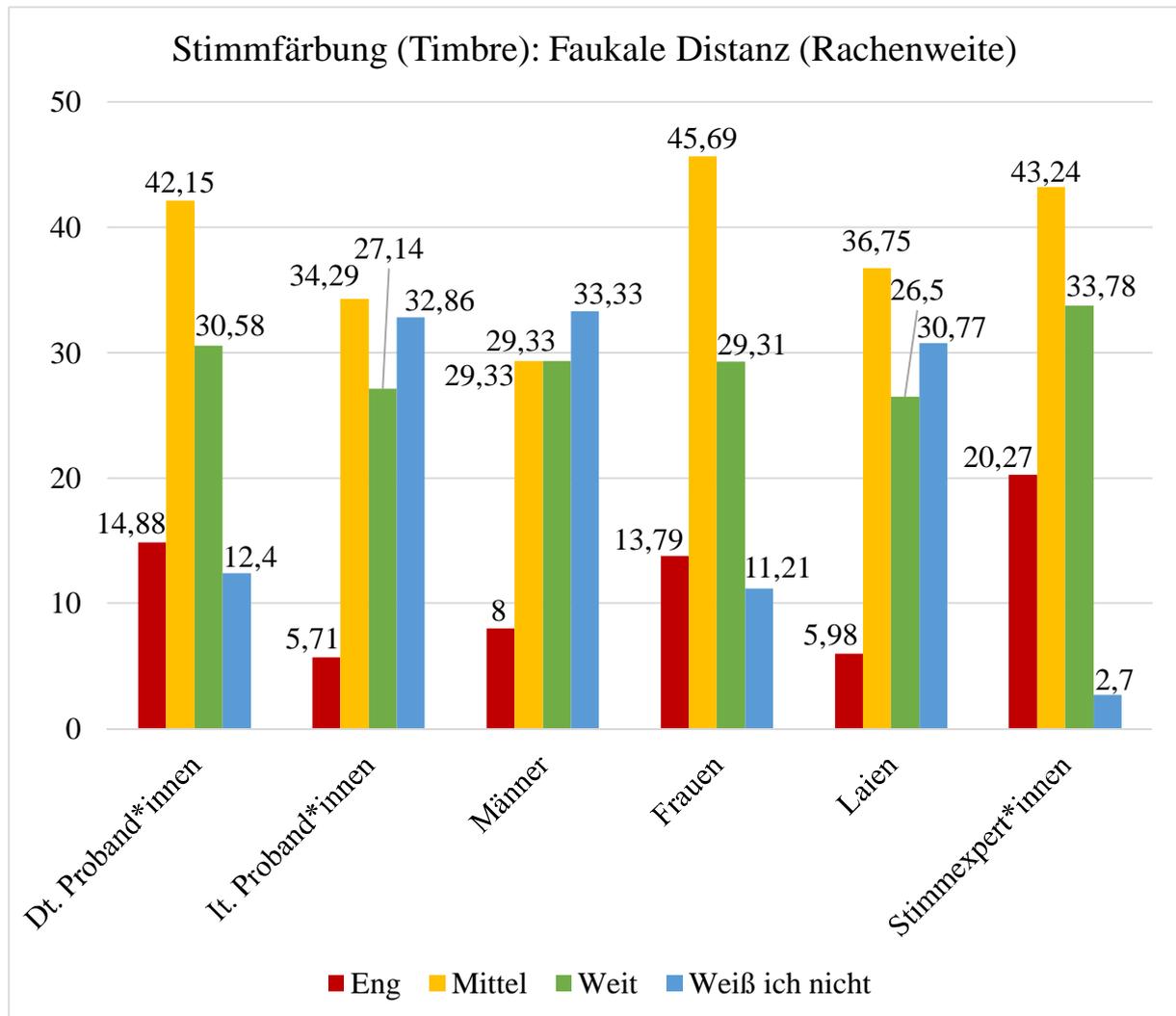
Klangfülle



Bei der Beurteilung der Klangfülle ist das Ergebnis eindeutig – die absolute Mehrheit der Personen bewertet sie als „klangvoll“, nur circa 20 Prozent als „klangarm“ und noch weniger als „Missklang“. Die Quote derer, die die Frage nicht beantworten können, ist besonders bei den Männern mit 9,33 Prozent recht hoch. Differenzen in der Analyse zeigen sich hier besonders aufgrund der Sprache/Kultur und des Geschlechts.

Die italienischsprachigen empfinden die Äußerungen als deutlich klangvoller als die deutschsprachigen Proband*innen. Dies zeigt sich nicht nur im Vergleich der in der Kategorie „klangvoll“ vorgenommenen Bewertungen, sondern auch bei den anderen Abstufungen. Selbiges zeigt sich als geschlechtsbezogener Unterschied, hier nehmen die Frauen das Timbre als deutlich klangvoller wahr als die Männer. Keine relevanten Differenzen gibt es hingegen bei der Perzeption der Stimmexpert*innen und Laien. Insgesamt kann von einem von allen Gruppen recht einheitlichen Analyseergebnis gesprochen werden.

Faukale Distanz (Rachenweite)



Bei der Beurteilung der faukalen Distanz (Rachenweite) kommt es zu großen Differenzen zwischen den Proband*innengruppen, wobei besonders die hohe Zahl derjenigen ins Auge sticht, welche die Frage nicht beantworten können. Insgesamt lässt sich festhalten, dass eine mittlere Rachenweite festgestellt wurde. Circa 29 Prozent beurteilen die faukale Distanz recht einheitlich als weit. Die Bewertung „eng“ hingegen fällt unterschiedlich hoch aus, von knapp 6 Prozent der italienischsprachigen Proband*innen bis zu 20,27 Prozent der Expert*innen. Des Weiteren lassen sich folgende Differenzen herauslesen:

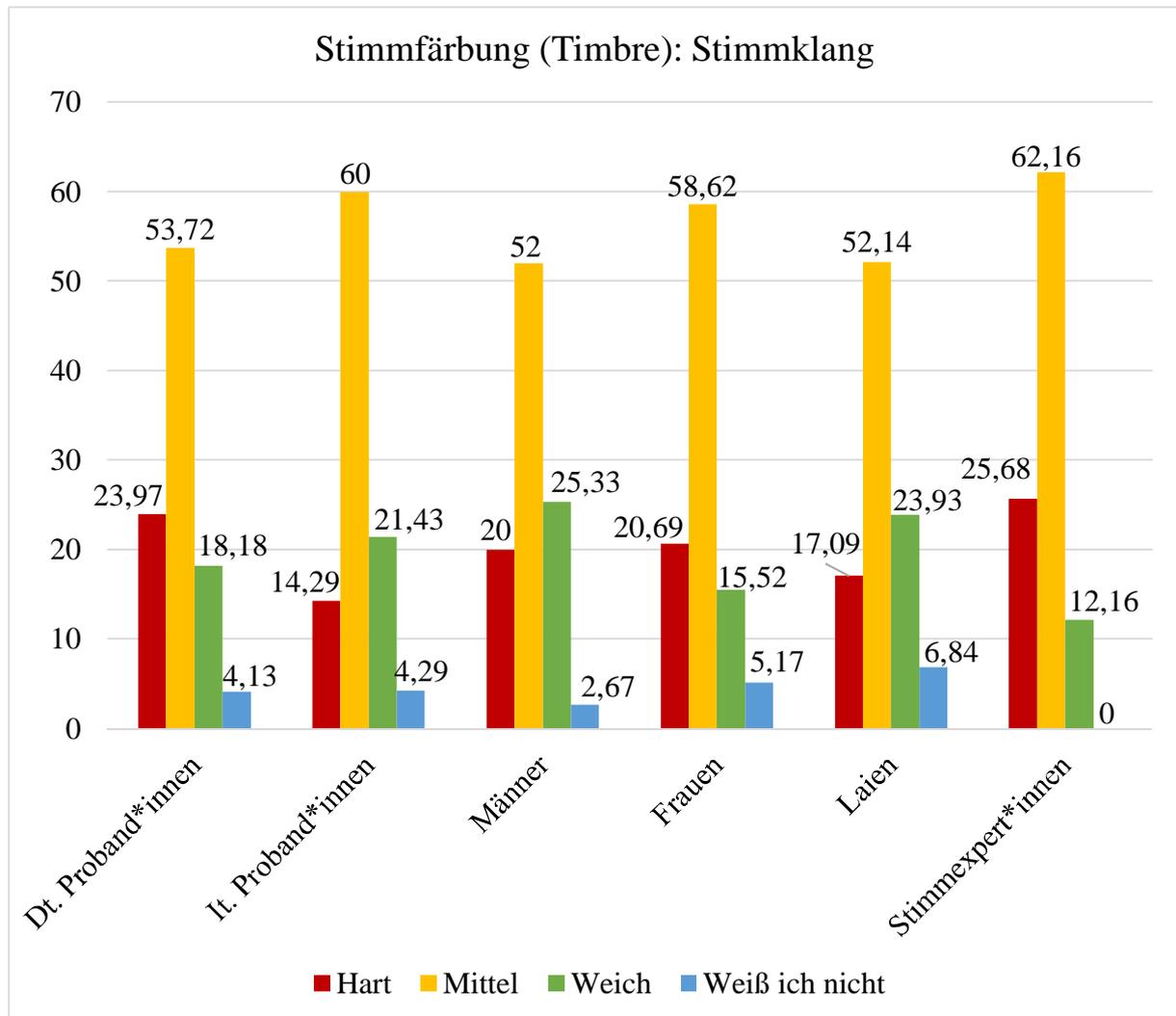
Sprache/Kultur: Das Ergebnis der deutschsprachigen Gruppe ist deutlich differenzierter als jenes der italienischsprachigen. Circa 42,15 Prozent der deutschsprachigen Proband*innen nehmen die Bewertung „mittel“ vor, aber nur 34,29 Prozent der italienischsprachigen. Die Angabe „weit“ wird von beiden Gruppen fast identisch gemacht (ca. 29 Prozent), die Angabe „eng“ von mehr deutsch- als italienischsprachigen Personen. Der größte Unterschied aber liegt darin, dass

32,86 der italienischsprachigen Gruppe die Frage nicht beantworten können, im Vergleich zu nur 12,4 Prozent der deutschsprachigen.

Geschlecht: Der wohl größte Unterscheidungsfaktor scheint hier das Geschlecht zu sein, wobei die Frauen eine deutlich differenziertere und klarere Analyse vornehmen als die Männer, von denen 33,33 Prozent die Frage (im Vergleich zu 11,21 Prozent bei den Frauen) gar nicht beantworten können. Die Bewertung „mittel“ wird von 45,69 Prozent der Frauen, aber nur von 29,33 Prozent der Männer vorgenommen. Die Angabe „weit“ wird von gleich vielen Personen gemacht.

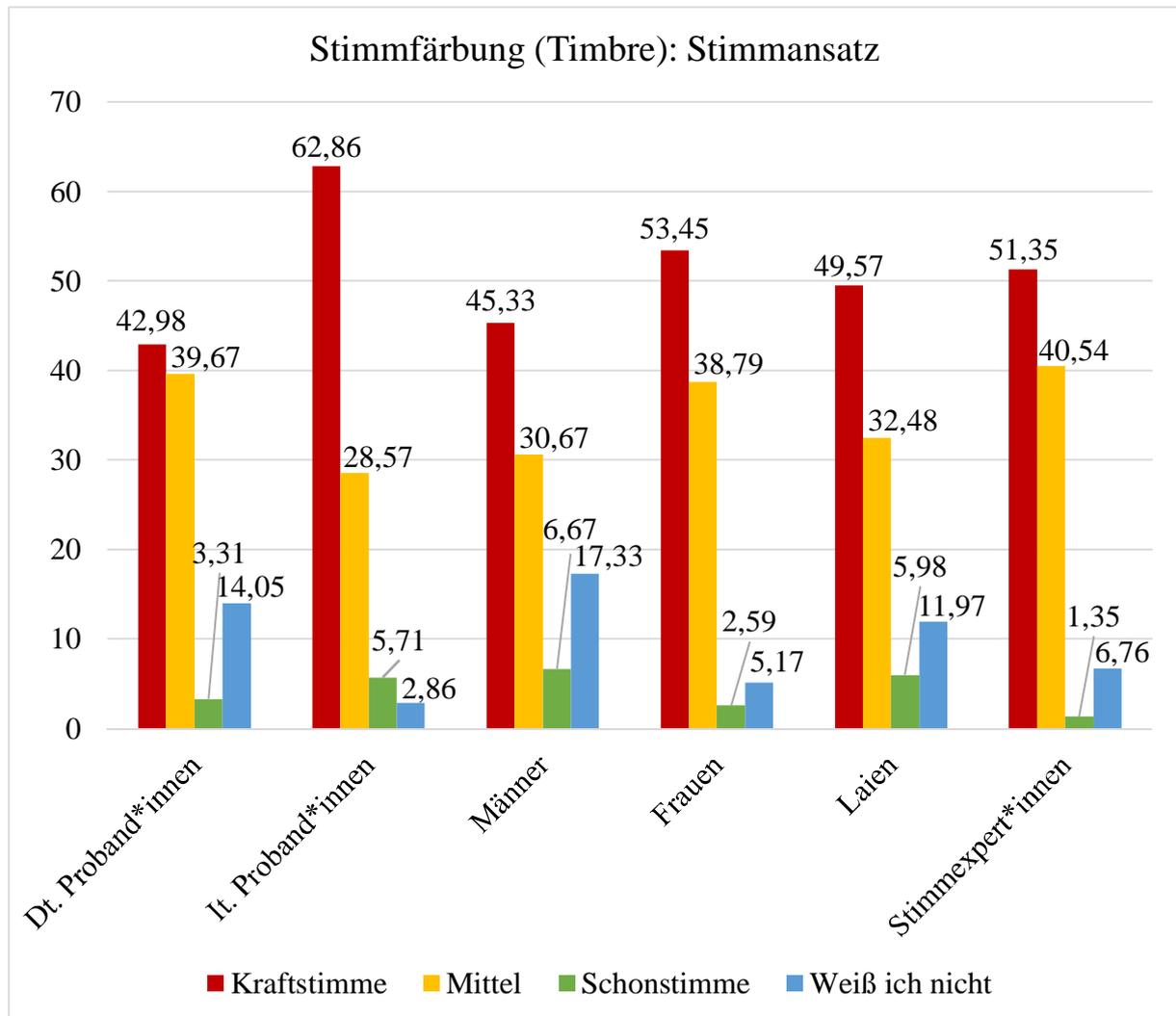
Beruf: Eine enorme Differenz in Bezug auf den Faktor Beruf zeigt sich in der Zahl derer, die die Frage nicht beantworten können – 30,77 Prozent der Laien versus nur 2,7 Prozent der Stimmexpert*innen. Im Ranking sind sich beide Gruppen einig, die meisten Personen perzipieren eine mittlere, etwas weniger eine weite Rachenweite. Die Angabe „eng“ wird von 20,27 Prozent der Expert*innen, aber nur von 5,98 Prozent der Laien gemacht.

Stimmklang



Der Stimmklang wird von allen Proband*innen eindeutig als „mittel“ eingestuft, nicht klar zu unterscheiden sind die Zuordnungen zu „weich“ und „hart“, hier gibt es kein klares Ranking. Insgesamt geht aber aus der Analyse klar hervor, dass ein mittlerer Stimmklang perzipiert wird, wobei dennoch diverse sprachlich/kulturelle, geschlechterbezogene und berufsspezifische Differenzen erkennbar sind. Tendenziell lässt sich feststellen, dass die italienischsprachigen Personen den Stimmklang insgesamt als weicher wahrnehmen als die deutschsprachigen. Selbiges gilt auch für die Laien im Vergleich zu den Stimmexpert*innen und für die Männer im Vergleich zu den Frauen.

Stimmansatz



Was die Analyse des Stimmansatzes betrifft, so nimmt die Mehrheit der Proband*innen den Gebrauch der Kraftstimme wahr, nur max. 6,67 Prozent (Männer) hingegen sprechen von der Verwendung der Schonstimme. Das Ranking ist also klar und eindeutig, wenngleich sich zwischen den Proband*innengruppen einige Differenzen feststellen lassen. Die Sprache/Kultur scheint eine große Rolle zu spielen – insgesamt nehmen die italienischsprachigen Personen den Stimmansatz als sehr viel kräftiger wahr als die deutschsprachigen (Bewertung „Kraftstimme“: 62,86 Prozent it. vs. 42,98 Prozent dt.; Bewertung „mittel“: 28,57 Prozent it. vs. 3,31 Prozent dt.). In nicht ganz so starkem Ausmaß lässt sich diese Unterscheidung auch in Bezug auf das Geschlecht treffen. Hier sind es die Frauen, welche den Stimmansatz als kräftiger empfinden als die Männer. Außerdem können deutlich mehr Männer (17,33 Prozent) als Frauen (5,17 Prozent) diese Frage nicht beantworten. In Bezug auf den Beruf lässt sich nur die Tendenz erkennen, dass die Stimmexpert*innen den Stimmansatz als etwas kräftiger wahrnehmen als die Laien,

von denen auch ein größerer Prozentsatz die Frage nicht beantworten kann. Diese Differenz ist jedoch nur gering.

Fazit

Da die Stimmfärbung (Timbre) als prosodisches Merkmal sehr komplex ist, wurden je mehrere Eigenschaften analysiert, um zu einem Gesamteindruck zu gelangen. Insgesamt wurde die Klangfarbe als „mittel“ bzw. „hell“, die Klangfülle als „klangvoll“ und die faukale Distanz (Rachenweite) als „mittel“ bis weit definiert. In Bezug auf den Stimmklang wurde die Bewertung „mittel“ vorgenommen und bezüglich des Stimmansatzes die „Kraftstimme“ ausgewählt. Es zeigt sich, dass es unterschiedlich schwierig war, die einzelnen Eigenschaften zu beurteilen. Besonders eindeutig und klar ist die Analyse der Klangfülle und des Stimmklangs, besonders schwammig jene der faukalen Distanz. Hier zeigt sich deutlich ein berufsbezogener Unterschied – die Stimmexpert*innen hatten im Gegensatz zu den anderen Proband*innen kaum Schwierigkeiten damit, die Frage zu beantworten. Bei allen Beurteilungen zeigen sich hingegen, teilweise sogar sehr große, sprachlich/kulturelle Differenzen in der Perzeption und Analyse der Stimmfärbung der freudigen Sprechweisen. Die italienischsprachigen Proband*innen nehmen das Timbre z. B. insgesamt als heller, klangvoller und weicher wahr als die deutschsprachigen Personen und empfinden den Stimmansatz als deutlich kräftiger. Auch geschlechterbedingte Unterschiede lassen sich feststellen. So nehmen die Frauen die Stimmfärbung etwa tendenziell als etwas klangvoller und kräftiger wahr als die Männer und zeigen ein etwas differenzierteres Ergebnis auf.

3.3.4.1.7 Gesamtergebnis zur Perzeption und Analyse der prosodischen Merkmale – Freude

In der folgenden Tabelle sind die zentralen Eigenschaften der einzelnen prosodischen Merkmale, wie sie von der Mehrheit aller Proband*innen wahrgenommen werden, überblicksartig zusammengefasst.

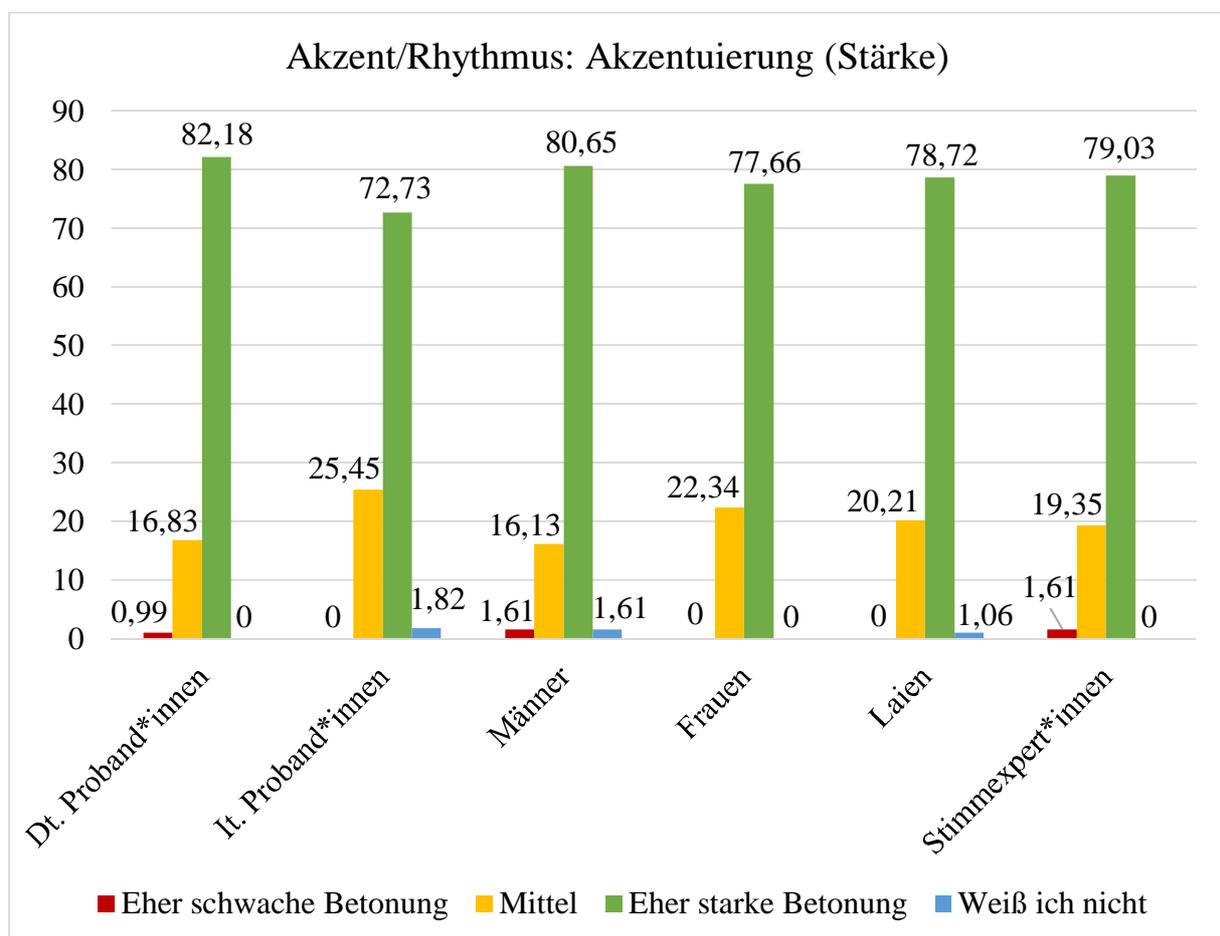
		Gesamtergebnis zur Perception und Analyse der freudigen Sprechweisen		
Akzent/ Rhythmus	Akzentuierung (Stärke) : eher starke Betonung	Akzentuierung (Häufigkeit) : eher viele Betonungen	Sprechrhythmus : eher fließend (legato)	
Tonhöhe/ Melodieführung	Sprechstimmlage : eher hoch	Tonhöhenverlauf: bewegt (lebhaft Melodie)	Tonhöhenumfang : mittel	
Lautstärke	Lautstärke : eher laut	Lautheitsverlauf : gleichförmig	Lautheitsverlauf am Ende der Phrase : gleichbleibende Lautstärke	
Spreche- schwindigkeit	Geschwindigkeit : mittel/eher schnell	Geschwindigkeitsverlauf : gleichförmig/mittel	Geschwindigkeitsverlauf am Ende der Phrase : gleichbleibende Geschwindigkeit	
Pausen	Häufigkeit : wenige Pausen	Geräuschvolle Pausen : eher wenige		
Stimmfärbung (Timbre)	Klangfarbe : hell/mittel	Klangfülle : klangvoll	Faukale Distanz : mittel	Stimmklang : mittel Stimmansatz : Kraftstimme

3.3.4.2 Perzeption und Analyse der prosodischen Merkmale – Ärger

Im Folgenden werden die Ergebnisse zur Perzeption und Analyse der prosodischen Merkmale in Bezug auf die ärgerlichen Sprechweisen ausgeführt. Die einzelnen prosodischen Parameter werden nacheinander und in ihren verschiedenen Eigenschaften analysiert und dargestellt. Dabei werden die Antworten der Proband*innengruppen so ausgewertet, dass die Faktoren Sprache/Kultur, Geschlecht und Beruf zum Tragen kommen und verglichen werden können.

3.3.4.2.1 Akzent/Rhythmus

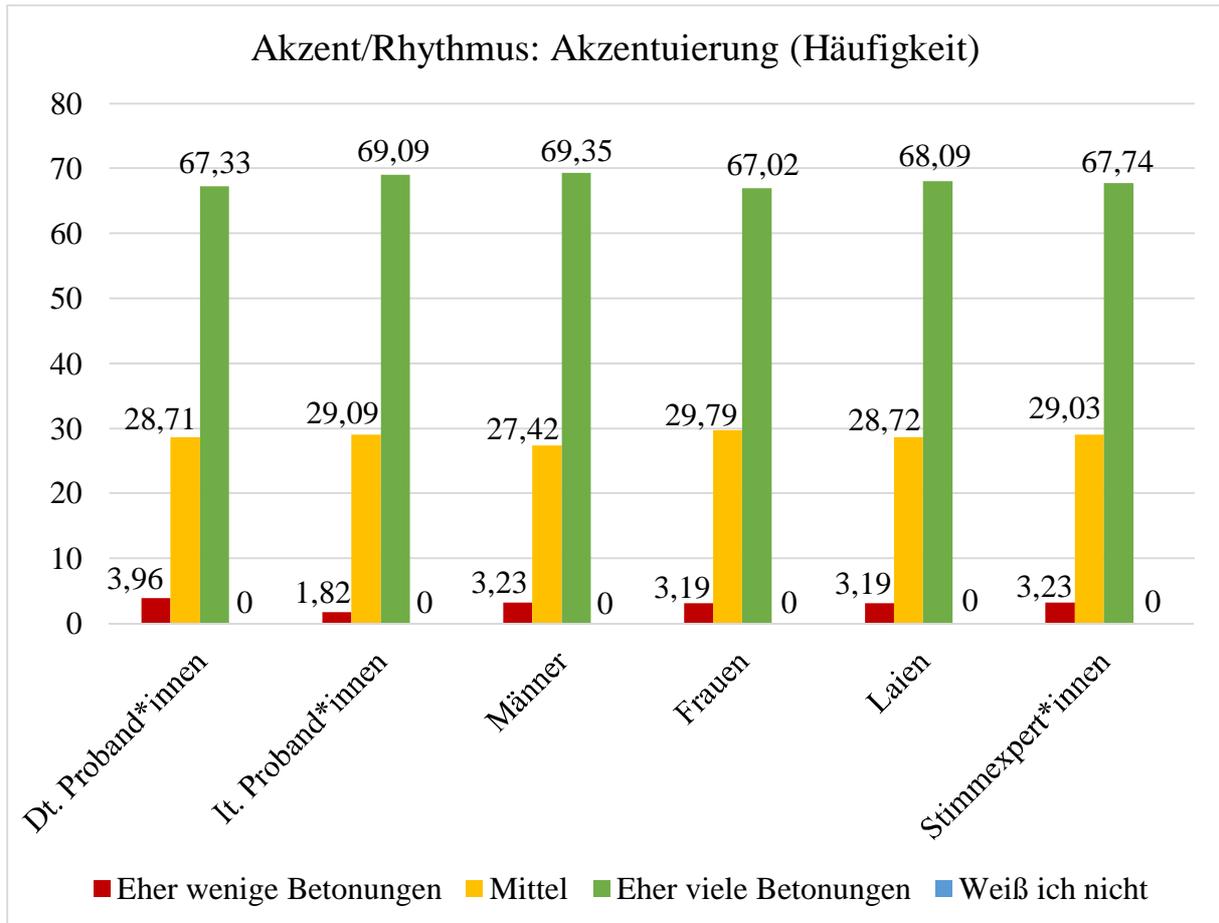
Akzentuierung (Stärke)



Die Beurteilung der Akzentuierung (Stärke) als Eigenschaft des prosodischen Merkmals Akzent/Rhythmus erfolgt sehr einheitlich von Seiten der Proband*innen. Es lässt sich ein eindeutiges Ergebnis ablesen – die absolute Mehrheit der Personen nimmt eher starke Betonungen wahr, nur ca. 20 Prozent geben eine mittlere Stärke an und fast niemand kann die Frage nicht

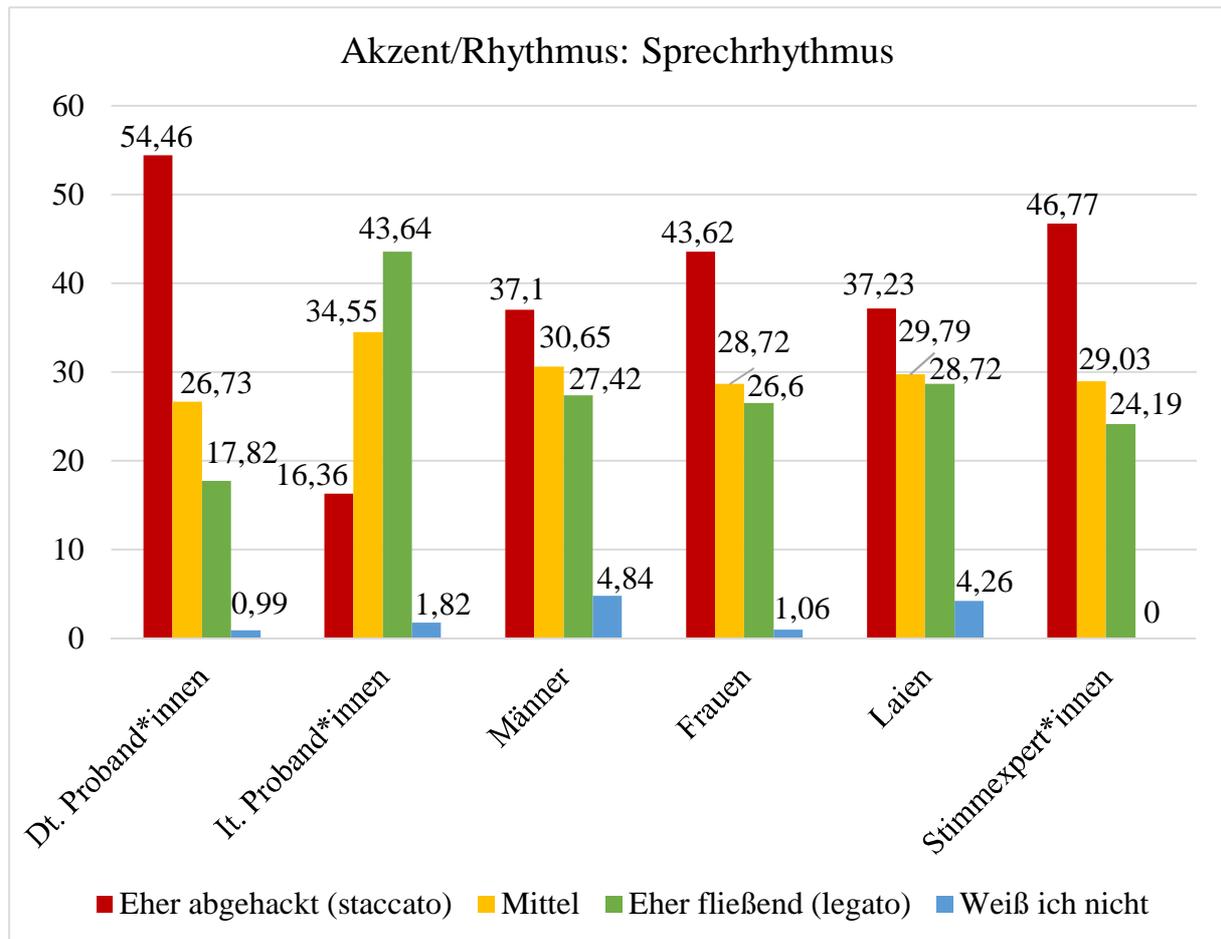
beantworten. Die Faktoren Geschlecht und Beruf scheinen bei der Beurteilung hier keine Rolle zu spielen. Es lässt sich allerdings eine sprach-/kulturbedingte Differenz feststellen, wobei die deutschsprachigen Proband*innen die Akzentuierung als stärker perzipieren als die italienischsprachigen.

Akzentuierung (Häufigkeit)



Auch die Analyse der Häufigkeit der Akzentuierung ergibt ein eindeutiges Ergebnis. Circa 68 Prozent der Proband*innen nehmen eher viele Betonungen wahr, circa 28 Prozent geben eine mittlere Häufigkeit an und alle können diese Frage beantworten. Die Bewertungen der einzelnen Gruppen sind identisch, d. h. es lassen sich hier keine sprachlich/kulturellen, geschlechts- oder berufsbezogenen Differenzen beobachten.

Sprechrhythmus



Der Sprechrhythmus der ärgerlichen Sprechweisen wird insgesamt als eher abgehackt (staccato) wahrgenommen. Dabei fallen jedoch ein großer sprachlich/kultureller, ein relevanter berufsbedingter und auch ein leichter geschlechtsbezogener Unterschied auf. Die deutschsprachigen Proband*innen empfinden den Sprechrhythmus als sehr viel abgehackter als die italienischsprachigen (Bewertung „eher abgehackt“: 54,46 Prozent dt. vs. 16,36 Prozent it.), welche ihn sogar als eher fließend (legato) perzipieren (Bewertung „eher fließend“: 43,64 Prozent it. vs. 17,82 Prozent dt.). Auch der Beruf spielt eine Rolle – mehr Stimmexpert*innen als Laien bezeichnen den Sprechrhythmus als eher abgehackt. Beim Faktor Geschlecht zeigt sich, dass die Männer den Rhythmus als fließender betrachten als die Frauen.

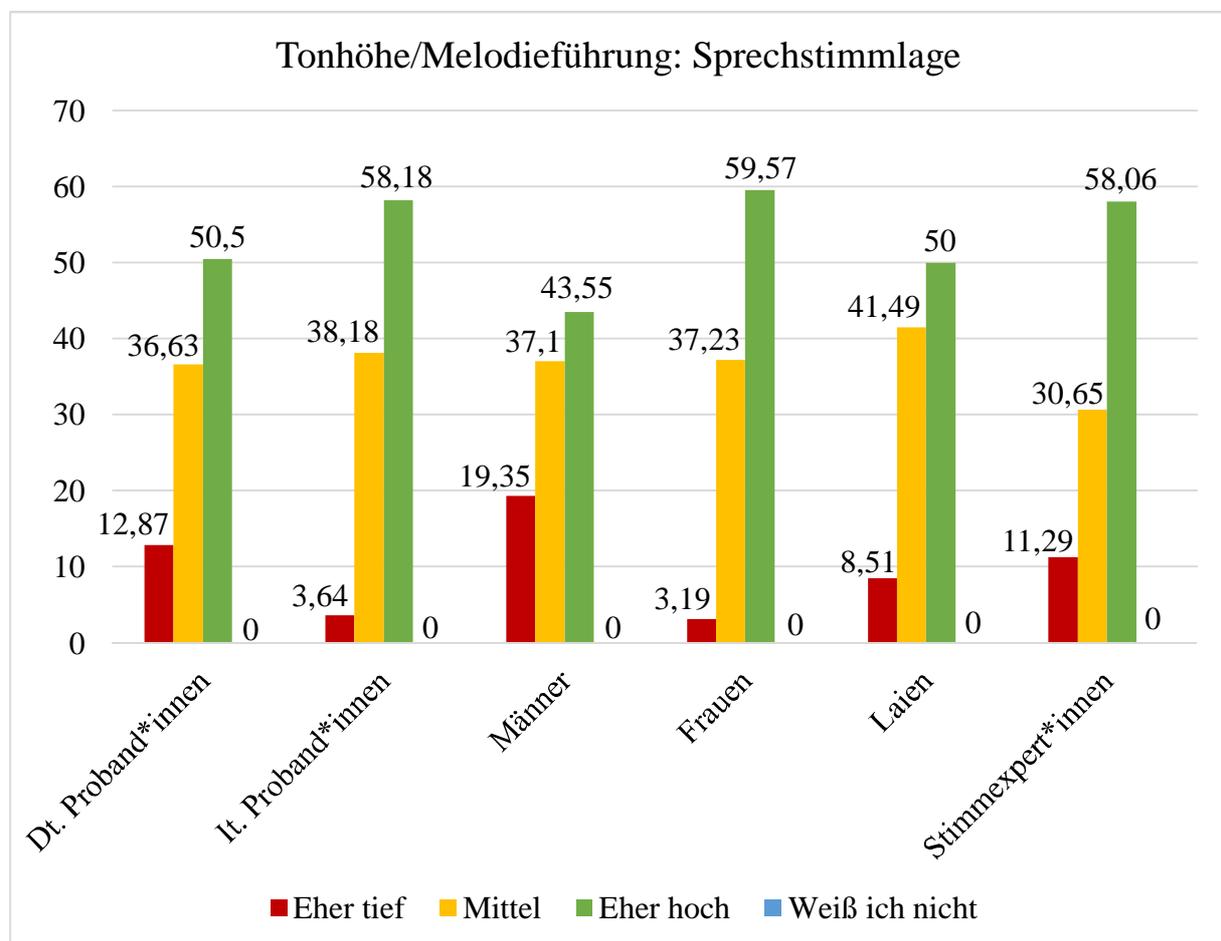
Fazit

Insgesamt werden in Bezug auf das prosodische Merkmal Akzent/Rhythmus eher starke und eher viele Betonungen und ein eher abgehackter Sprechrhythmus wahrgenommen. Die Ergebnisse zur Akzentuierung sind klar und eindeutig und es lässt sich nur bei der Beurteilung der

Stärke der Akzentuierung ein sprachlich/kultureller Unterschied feststellen (die deutschsprachigen Proband*innen empfinden die Betonungen als stärker als die italienischsprachigen), bei der Perzeption der Häufigkeit der Akzentuierung gibt es keine Differenzen in Bezug auf Sprache/Kultur, Geschlecht und Beruf. Anders ist dies bei der Beurteilung des Sprechrhythmus, der von den deutschsprachigen Personen sehr viel abgehackter perzipiert wird als von den italienischsprachigen, die ihn sogar als eher fließend bezeichnen. Die Stimmexpert*innen und die Frauen nehmen den Sprechrhythmus häufiger als abgehackt wahr als die Laien und die Männer.

3.3.4.2.2 Tonhöhe/Melodieführung

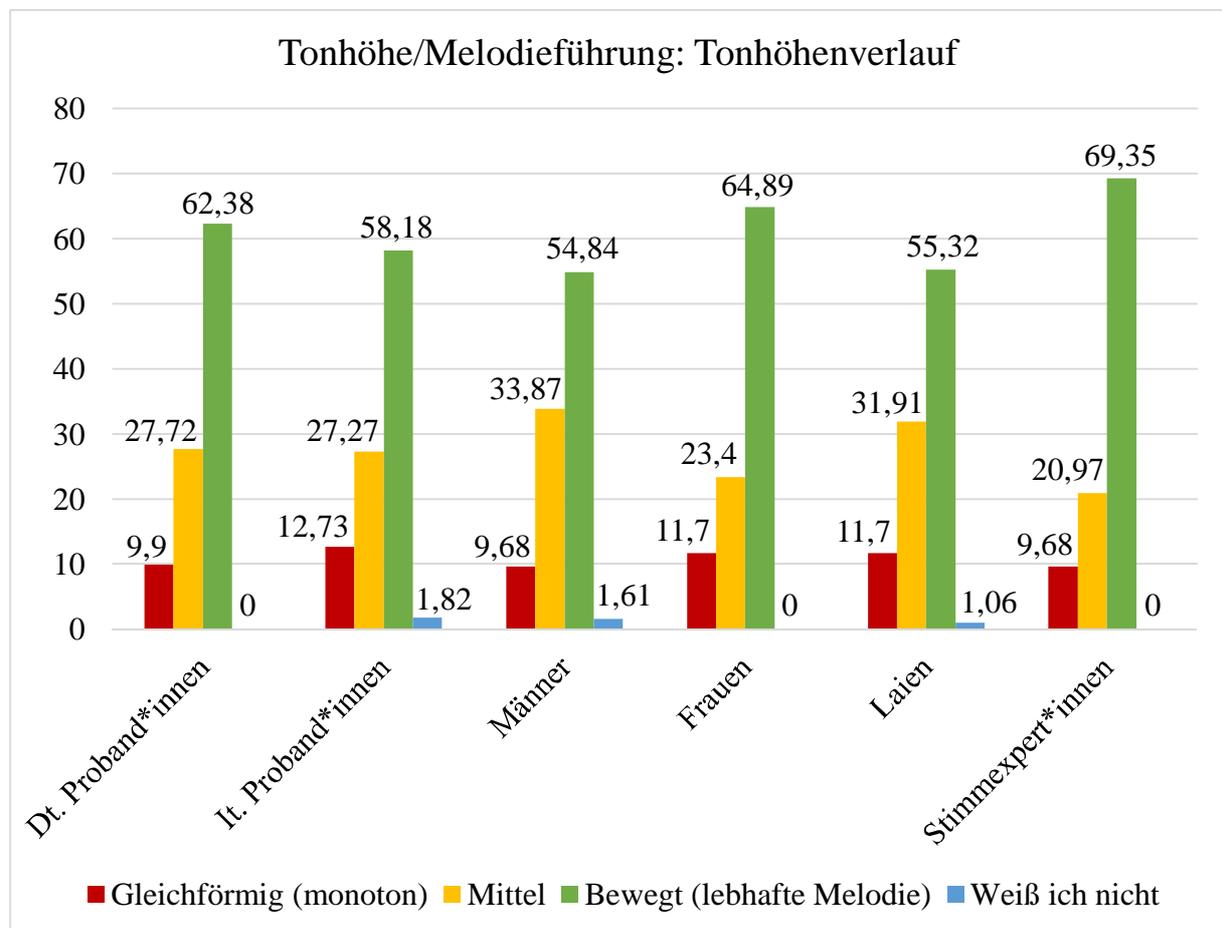
Sprechstimmlage



Im Rahmen der Analyse der Tonhöhe/Melodieführung der ärgerlichen Sprechweisen wird zunächst die Eigenschaft der Sprechstimmlage bewertet. Sie wird insgesamt als eher hoch eingeschätzt, vor allem von Seiten der Stimmexpert*innen, der Frauen und der italienischsprachigen Proband*innen. Die Faktoren Sprache/Kultur, Beruf und v. a. auch Geschlecht spielen bei der

Perzeption hier also eine Rolle. Besonders deutlich tritt die Differenz zwischen Männern und Frauen zutage, wobei letztere die Sprechstimmlage als höher betrachten (Bewertung „eher hoch“: 59,57 Prozent w vs. 43,55 Prozent m; Bewertung „eher tief“: 3,19 Prozent w vs. 19,35 Prozent m). Auch die Stimmexpert*innen nehmen die Sprechstimmlage als höher wahr im Vergleich zu den Laien. Selbiges gilt für die italienischsprachigen Proband*innen im Vergleich zu den deutschsprachigen. Die Differenzen sind hier aber nicht mehr so groß.

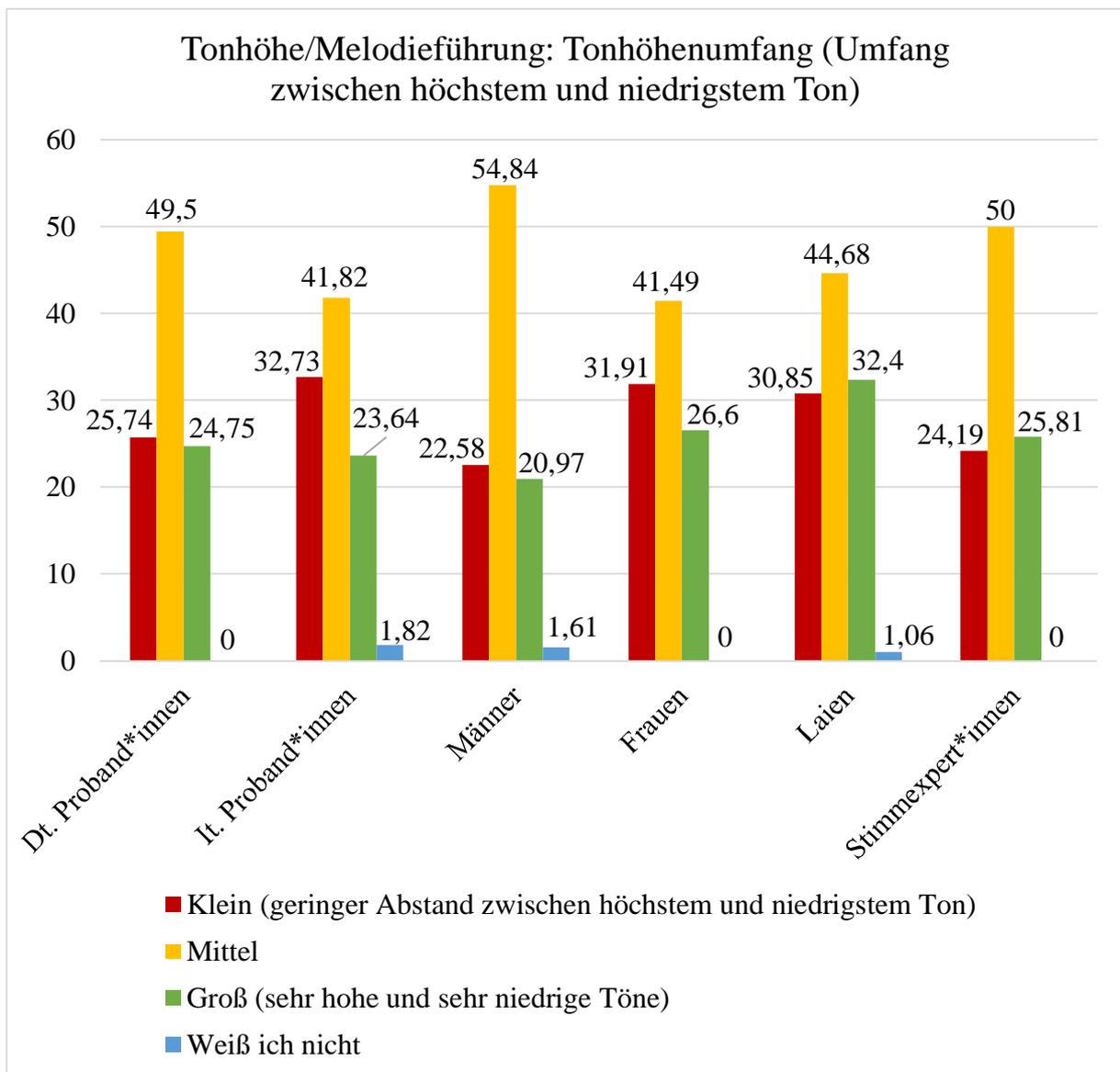
Tonhöhenverlauf



Der Tonhöhenverlauf der ärgerlichen Sprechweisen wird von der Mehrheit der Proband*innen einheitlich als bewegt (lebhaft Melodie) perzipiert. Besonders klar ist die Bewertung von Seiten der Expert*innen und der Frauen. Vergleicht man die Aussagen der deutschsprachigen und der italienischsprachigen Proband*innen, so lässt sich kein relevanter sprachlich/kultureller Unterschied feststellen. Eine deutliche Differenz zeigt sich hingegen zwischen der Analyse der

Laien und jener der Stimmexpert*innen, wobei letztere eine differenziertere Bewertung vornehmen. Selbiges gilt für die Frauen im Vergleich mit den Männern, welche den Tonhöhenverlauf als weniger bewegt einstufen als die Frauen.

Tonhöhenumfang



Die dritte zu analysierende Eigenschaft der Tonhöhe/Melodieführung ist der Tonhöhenumfang, d. h. der Umfang zwischen dem höchsten und dem niedrigsten Ton. Dieser wird von den Proband*innen mehrheitlich als mittelmäßig angegeben. Das klarste Ergebnis weisen die Männer auf, gefolgt von den Stimmexpert*innen und den deutschsprachigen Proband*innen. Die Faktoren Sprache/Kultur, Beruf und insbesondere Geschlecht, scheinen hier also alle eine Rolle zu

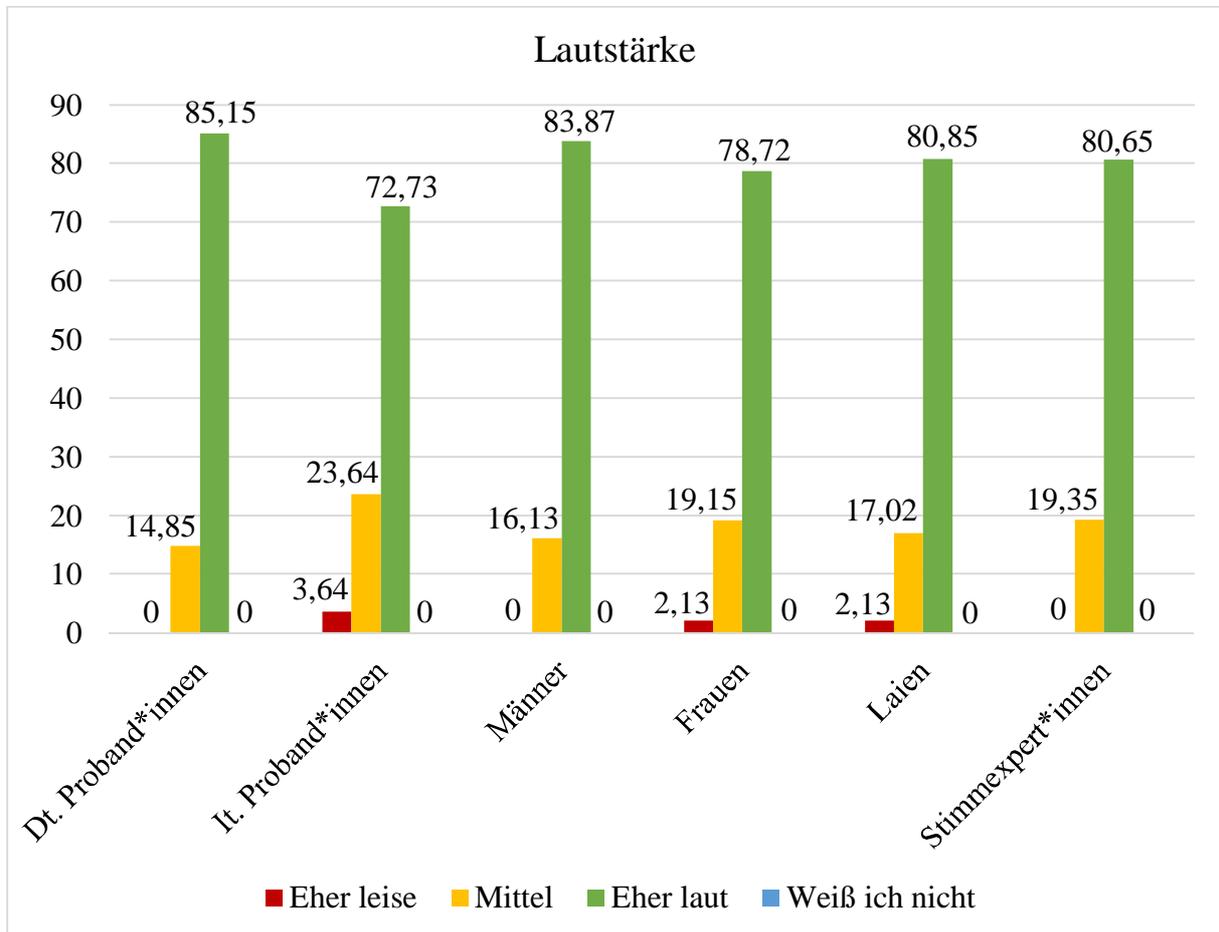
spielen. Bei den weiteren Bewertungsmöglichkeiten, „klein“ und „groß“, ist das Verhältnis unklar, wobei nach dem mittleren eher ein kleinerer Tonhöhenumfang wahrgenommen wird. Die italienischsprachigen Proband*innen perzipieren einen geringeren Tonhöhenumfang als die deutschsprachigen, die Frauen einen geringeren als die Männer und die Laien einen geringeren als die Stimmexpert*innen.

Fazit

Das prosodische Merkmal Tonhöhe/Melodieführung wird in Bezug auf die ärgerlichen Sprechäußerungen wie folgt wahrgenommen: Die Sprechstimmlage wird als eher hoch, der Tonhöhenverlauf als bewegt (lebhaft Melodie) und der Tonhöhenumfang als mittelmäßig perzipiert. Dabei spielen die Faktoren Sprache/Kultur, Geschlecht und Beruf bei allen Fragen in unterschiedlicher Ausprägung eine Rolle. Die Stimmexpert*innen weisen in allen Fällen eine differenziertere Bewertung auf als die Laien, die Analyse des Tonhöhenverlaufs und der Sprechstimmlage ist von Seiten der Frauen klarer als von Seiten der Männer. Bei der Bewertung des Tonhöhenumfangs jedoch ist das Ergebnis der Männer differenzierter als jenes der Frauen. Die Sprechstimmlage wird von der italienischsprachigen Gruppe als höher wahrgenommen als von der deutschsprachigen, letztere perzipieren den Tonhöhenumfang als etwas größer als die italienischsprachigen Proband*innen. In Bezug auf den Tonhöhenverlauf können keine relevanten sprachlich/kulturellen Differenzen festgestellt werden, auch insgesamt fallen sie eher gering aus. Wesentlich ausgeprägter sind die Unterschiede hinsichtlich des Geschlechtes und des Berufes der Teilnehmer*innen. Insgesamt konnte das Merkmal Tonhöhe/Melodieführung sehr gut erkannt und perzipiert werden, v. a. die Sprechstimmlage und der Tonhöhenverlauf.

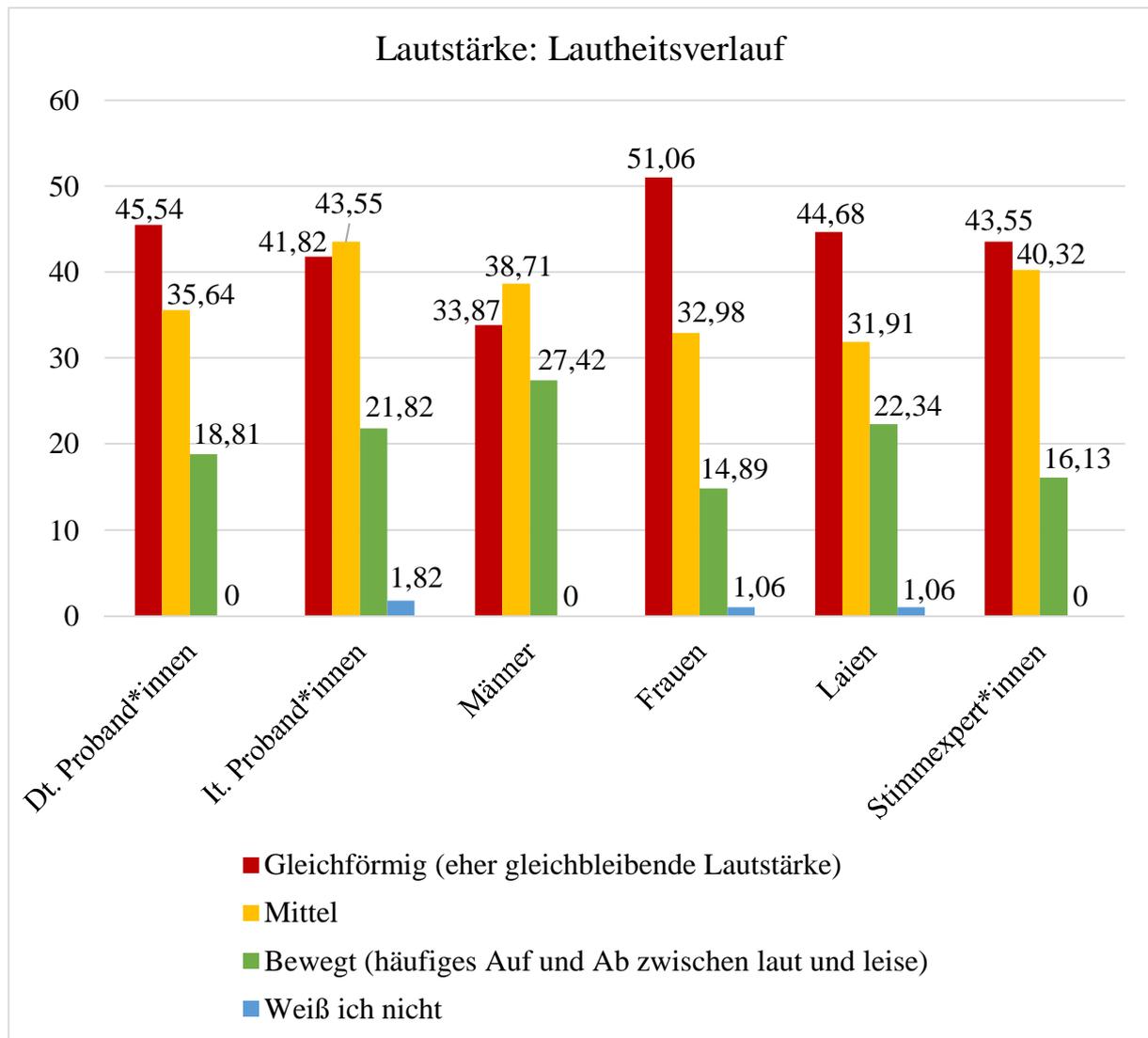
3.3.4.2.3 Lautstärke

Lautstärke



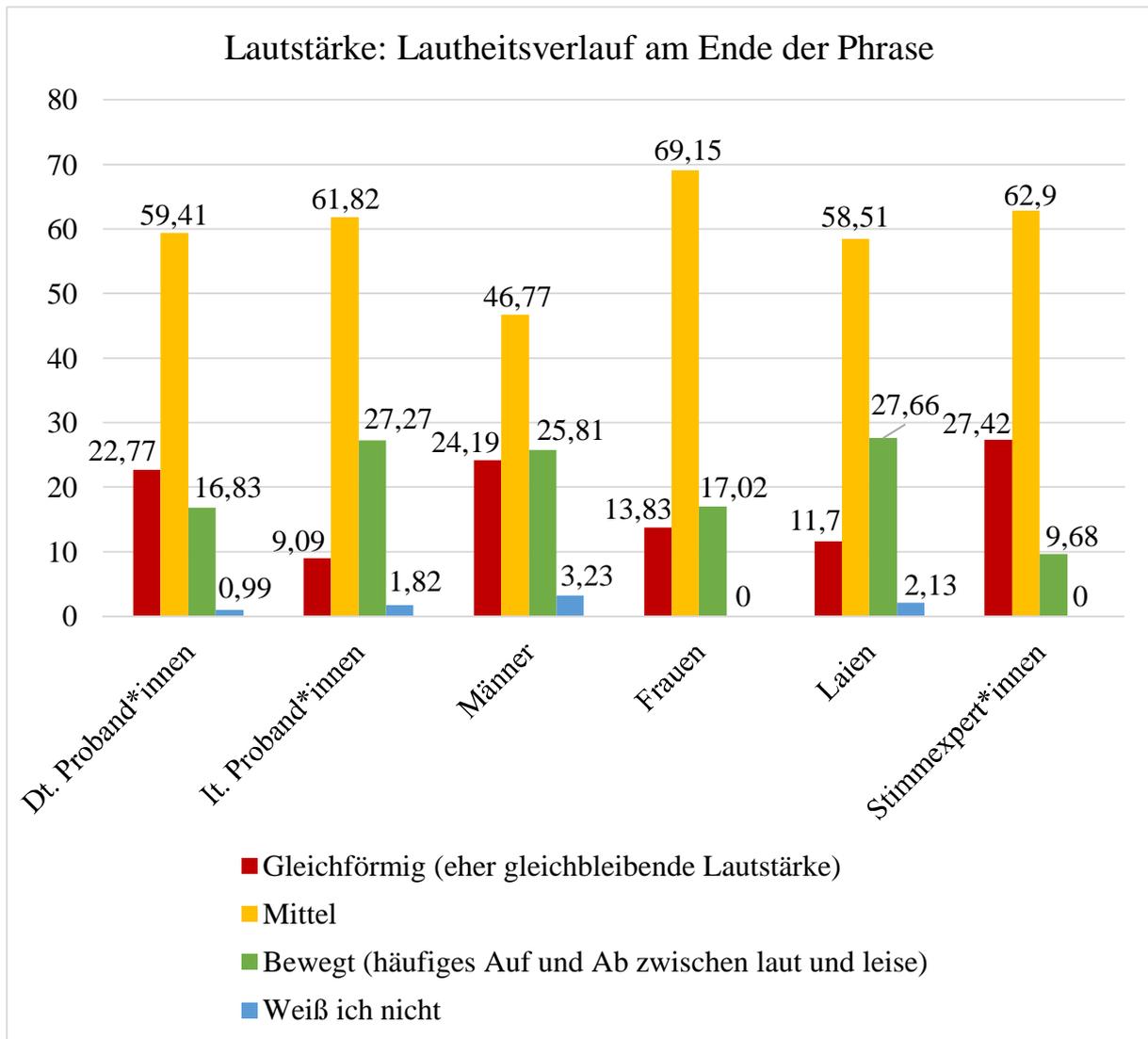
Das dritte prosodische Merkmal, das analysiert wurde, ist die Lautstärke. Was die Eigenschaft der Lautstärke betrifft, so lässt sich ein eindeutiges Ergebnis erkennen – die große Mehrheit der Proband*innen nimmt die Lautstärke als eher laut wahr (ca. 80 Prozent), nur ca. 18 Prozent als mittelmäßig. Niemand konnte diese Frage nicht beantworten. Ein berufsbedingter Unterschied in der Analyse lässt sich nicht ablesen, da die Angaben von Expert*innen und Laien beinahe identisch sind. Auch von einer relevanten geschlechterbedingten Differenz kann nicht gesprochen werden, wenngleich die Männer die Lautstärke tendenziell als etwas lauter einstufen. Ein deutlicher Unterschied trifft jedoch in Bezug auf die Sprache/Kultur auf, wobei die deutschsprachigen Personen die ärgerlichen Sprechweisen als viel lauter wahrnehmen als die italienischsprachigen.

Lautheitsverlauf



Was den Lautheitsverlauf betrifft, so wird dieser von der Mehrheit der Proband*innen insgesamt als gleichförmig angesehen. Auf dem zweiten Posten ist die Bewertung „mittel“ und nur ca. 20 Prozent nehmen den Lautheitsverlauf als bewegt wahr. Zudem fällt auf, dass fast alle diese Frage beantworten konnten. Vergleicht man die Proband*innengruppen untereinander, lassen sich sprach-/kulturbezogene, geschlechter- und berufsabhängige Differenzen beobachten. Am differenziertesten sind die Analysen der deutschsprachigen Proband*innen, der Frauen und der Laien. Bei den Stimmexpert*innen, den Männern und den italienischsprachigen Personen sind die Abstufungen zwischen den Bewertungen nicht klar bzw. nehmen diese Gruppen den Lautheitsverlauf als etwas bewegter wahr als die anderen.

Lautheitsverlauf am Ende der Phrase



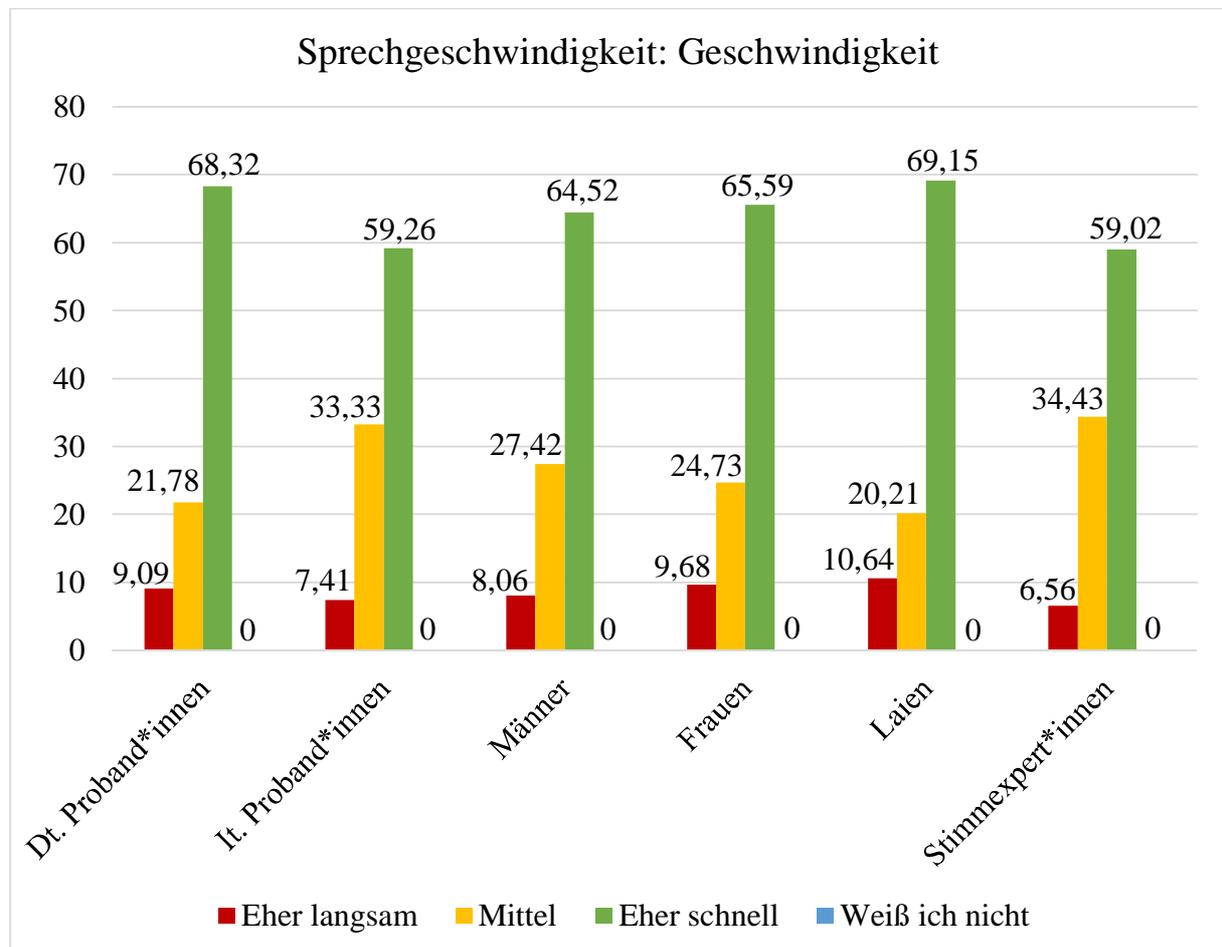
Als dritte Eigenschaft der Lautstärke wurde der Lautheitsverlauf analysiert, der in Bezug auf die ärgerlichen Sprechweisen als gleichbleibend beurteilt wird. Diese Bewertung wird relativ gleichmäßig in allen Gruppen von ca. 60 Prozent der Personen vorgenommen, bei den Männern sind es mit 46,77 Prozent jedoch deutlich weniger, am meisten sind es bei den Frauen (69,15 Prozent). Das Geschlecht scheint hier also eine relevante Rolle zu spielen. Das Ergebnis der Männer ist insgesamt am unklarsten. Was die anderen beiden Bewertungsmöglichkeiten, nämlich „leiser werdend“ und „lauter werdend“ betrifft, so gibt es hier keine klare Festlegung. Tatsächlich unterscheidet sich die Einstufung drastisch je nach Sprache/Kultur und Beruf. Die deutschsprachigen Proband*innen perzipieren den Lautheitsverlauf als deutlich leiser als die italienischsprachigen Personen. Selbiges trifft auch auf die Stimmexpert*innen im Vergleich mit den Laien zu.

Fazit

Die Lautstärke als prosodisches Merkmal konnte recht einheitlich von den Proband*innen analysiert werden, dies gilt in besonderem Maße für die Eigenschaft der Lautstärke. Insgesamt nehmen die Proband*innen die ärgerlichen Sprechweisen als eher laut und mit gleichförmigem Lautheitsverlauf, auch am Ende der Phrase, wahr. Die Faktoren Sprache/Kultur, Geschlecht und Beruf beeinflussen die Bewertungen. Die deutschsprachigen Teilnehmer*innen perceive eine größere Lautstärke und einen gleichförmigeren und gegen Ende der Phrase hin leiseren Lautheitsverlauf als die italienischsprachigen Proband*innen. Die Frauen weisen ein differenzierteres Ergebnis auf als die Männer und die Stimmexpert*innen perceive den Lautheitsverlauf als etwas bewegter und am Ende der Phrase als wesentlich leiser im Vergleich zu den Laien.

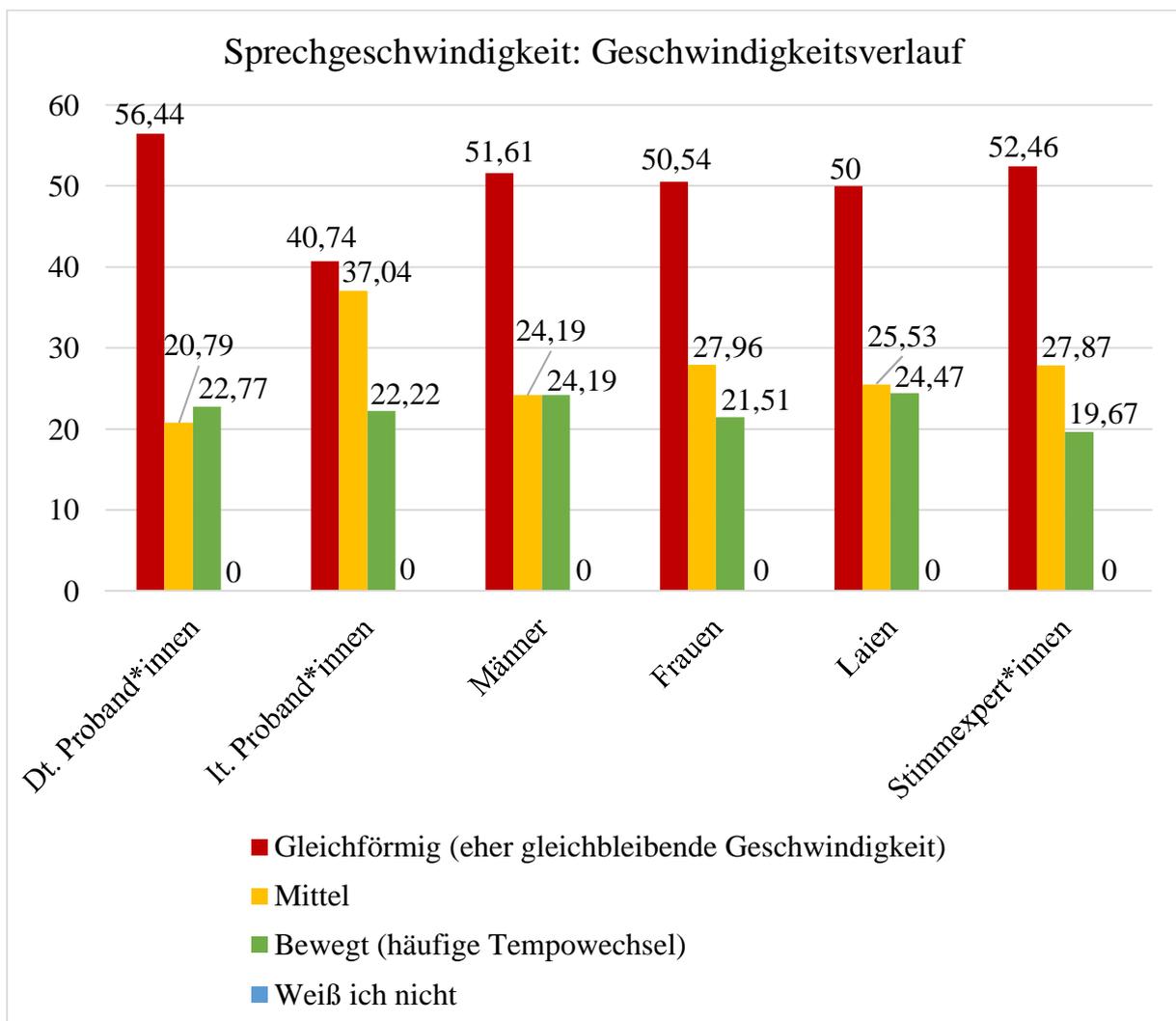
3.3.4.2.4 Sprechgeschwindigkeit

Geschwindigkeit



Zunächst wurde im Rahmen der Analyse der Sprechgeschwindigkeit die Geschwindigkeit untersucht. Sie wird eindeutig von ca. 64 Prozent der Proband*innen als eher schnell eingestuft. Dies gilt in besonderem Maße für die deutschsprachigen Personen und die Laien, welche die Geschwindigkeit für schneller halten als die italienischsprachigen Proband*innen und die Stimmexpert*innen. Die Faktoren Sprache/Kultur und Beruf scheinen hier also eine relevante Rolle zu spielen. Eine nennenswerte geschlechtsbedingte Differenz lässt sich hingegen nicht feststellen. Insgesamt ist das Ergebnis sehr einheitlich und klar, außerdem konnten alle Personen diese Frage beantworten.

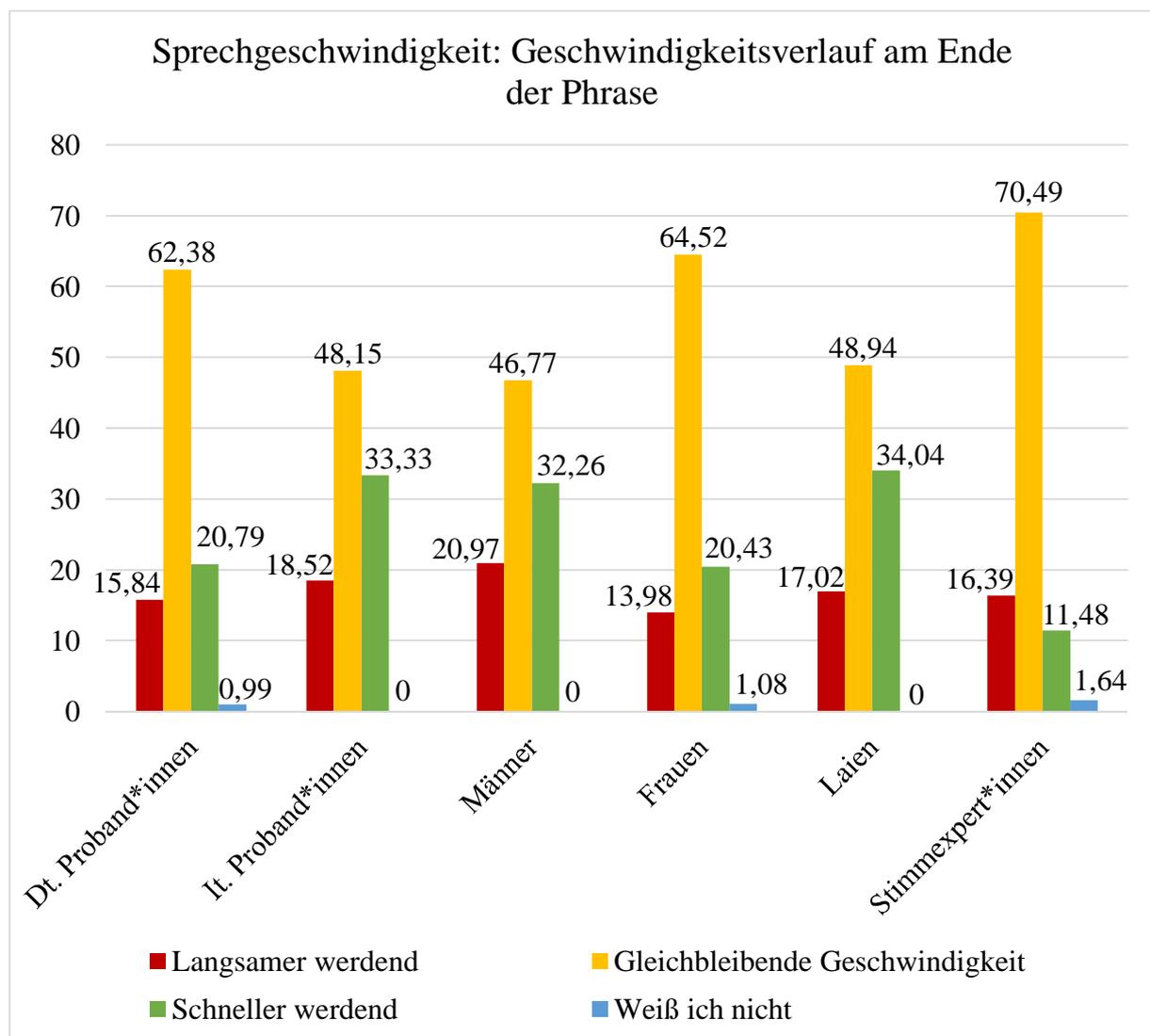
Geschwindigkeitsverlauf



Als weitere Eigenschaft der Sprechgeschwindigkeit wurde in einem zweiten Schritt der Geschwindigkeitsverlauf analysiert. Circa 50 Prozent der Befragten perceive ihn als gleichförmig (d. h. eher gleichbleibende Geschwindigkeit). Bei den Kategorien „mittel“ und „bewegt“

fehlt z. T. eine klare Abgrenzung. Lediglich die Stimmexpert*innen, die italienischsprachigen Proband*innen und tendenziell auch die Frauen nehmen hier eine klare Differenzierung vor. Ein geschlechterbedingter Unterschied fällt ansonsten, beim Vergleich der Zahlen zu den einzelnen Bewertungen, nicht auf. Selbiges gilt auch für den Faktor Beruf – sind die Nennungen zu den einzelnen Kategorien von beiden Gruppen zwar fast gleich häufig, so liegt innerhalb der Expert*innengruppe doch eine klarere Abstufung vor. Besonders ins Gewicht fällt hier allerdings vor allem der sprachlich/kulturelle Unterschied – die deutschsprachigen Proband*innen empfinden den Geschwindigkeitsverlauf als deutlich gleichförmiger als die italienischsprachigen, welche eher einen mittleren Verlauf perzipieren.

Geschwindigkeitsverlauf am Ende der Phrase



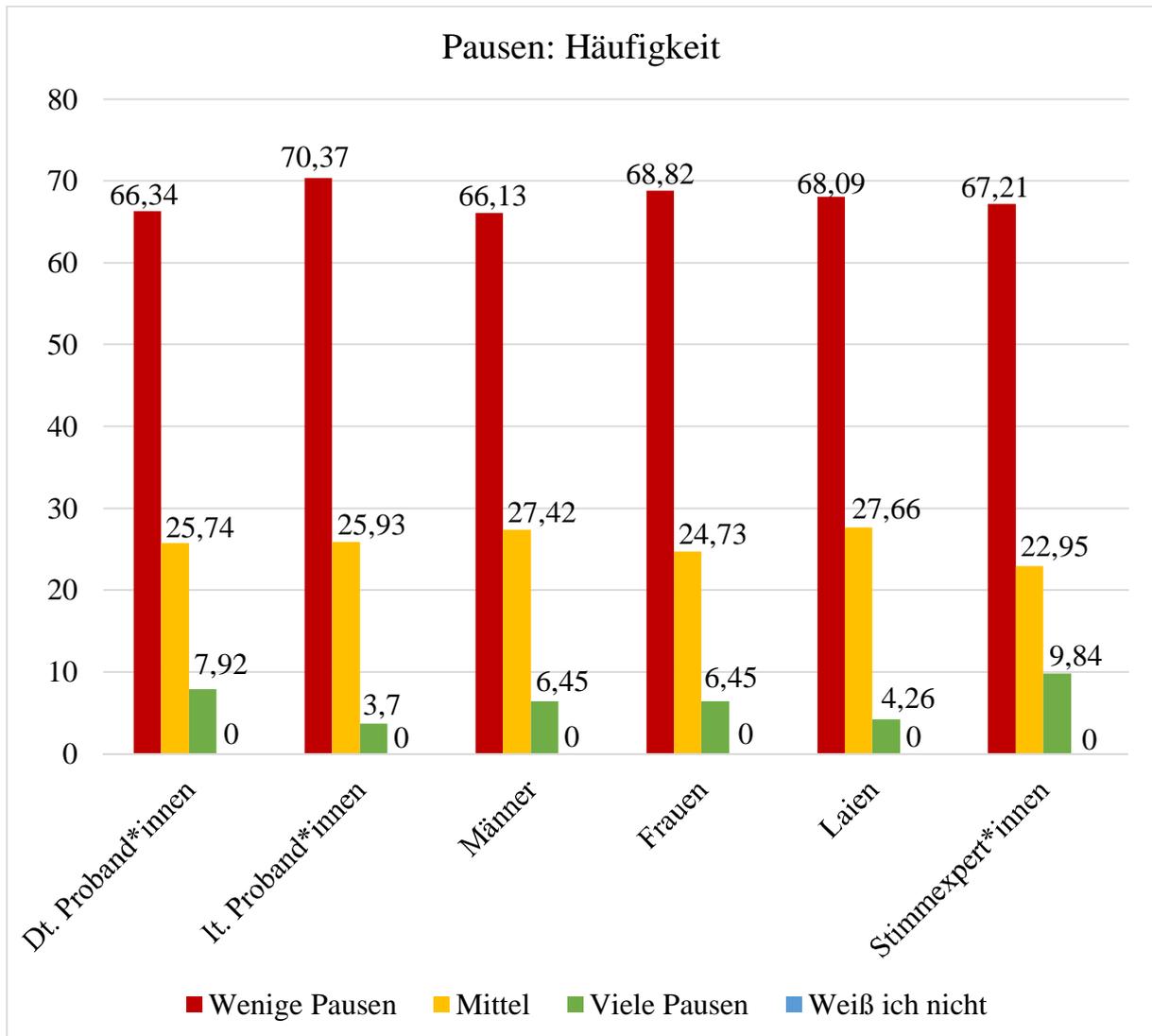
Im Rahmen der Analyse zur Sprechgeschwindigkeit sollte auch der Geschwindigkeitsverlauf am Ende der Phrase untersucht werden. Die Proband*innen perzipieren mehrheitlich eine gleichbleibende Geschwindigkeit, wobei starke sprachlich/kulturelle, geschlechter- und berufsbezogene Schwankungen vorliegen. Besonders klar und eindeutig zeigt sich das Ergebnis der Stimmexpert*innen, von denen 70,49 Prozent eine gleichbleibende Geschwindigkeit wahrnehmen. Im Vergleich dazu nehmen nur 48,94 Prozent der Laien diese Bewertung vor. Letztere perzipieren insgesamt einen eher schneller werdenden Verlauf am Ende der Phrase. Selbiges gilt für die Männer im Vergleich zu den Frauen und für die deutschsprachigen Proband*innen im Vergleich zu den italienischsprachigen. Insgesamt kann der Geschwindigkeitsverlauf am Ende der Phrase der ärgerlichen Sprechweisen also als gleichbleibend bezeichnet werden, mit Tendenz zum Schnellerwerden.

Fazit

Die Sprechgeschwindigkeit konnte von allen Proband*innen klar perzipiert und analysiert werden. Es fällt auf, dass fast niemand die Fragen nicht beantworten konnte. Besonders das Ergebnis zur Geschwindigkeit ist eindeutig, aber auch die anderen Eigenschaften konnten klar zugeordnet werden. Insgesamt wird eine eher schnelle und gleichförmige Geschwindigkeit vernommen und auch gegen Ende der Phrase wird die Geschwindigkeit als gleichbleibend beurteilt. In allen Fragen spielen sprachlich/kulturelle Differenzen eine sehr große Rolle, so nehmen die italienischsprachigen Proband*innen die Geschwindigkeitsverläufe als bewegter und die Geschwindigkeit selbst als weniger rasch als die deutschsprachigen Personen wahr. Auch der Faktor Geschlecht spielt eine Rolle, insbesondere beim Beurteilen des Geschwindigkeitsverlaufs am Ende der Phrase, den die Männer insgesamt undifferenzierter und als schneller bewerteten. Deutliche Differenzen zeigen sich auch zwischen Laien und Stimmexpert*innen, letztere beurteilen die Geschwindigkeit als weniger schnell und den Geschwindigkeitsverlauf gegen Ende der Phrase als gleichförmiger als die Laien und weisen ein etwas differenzierteres Ergebnis auf.

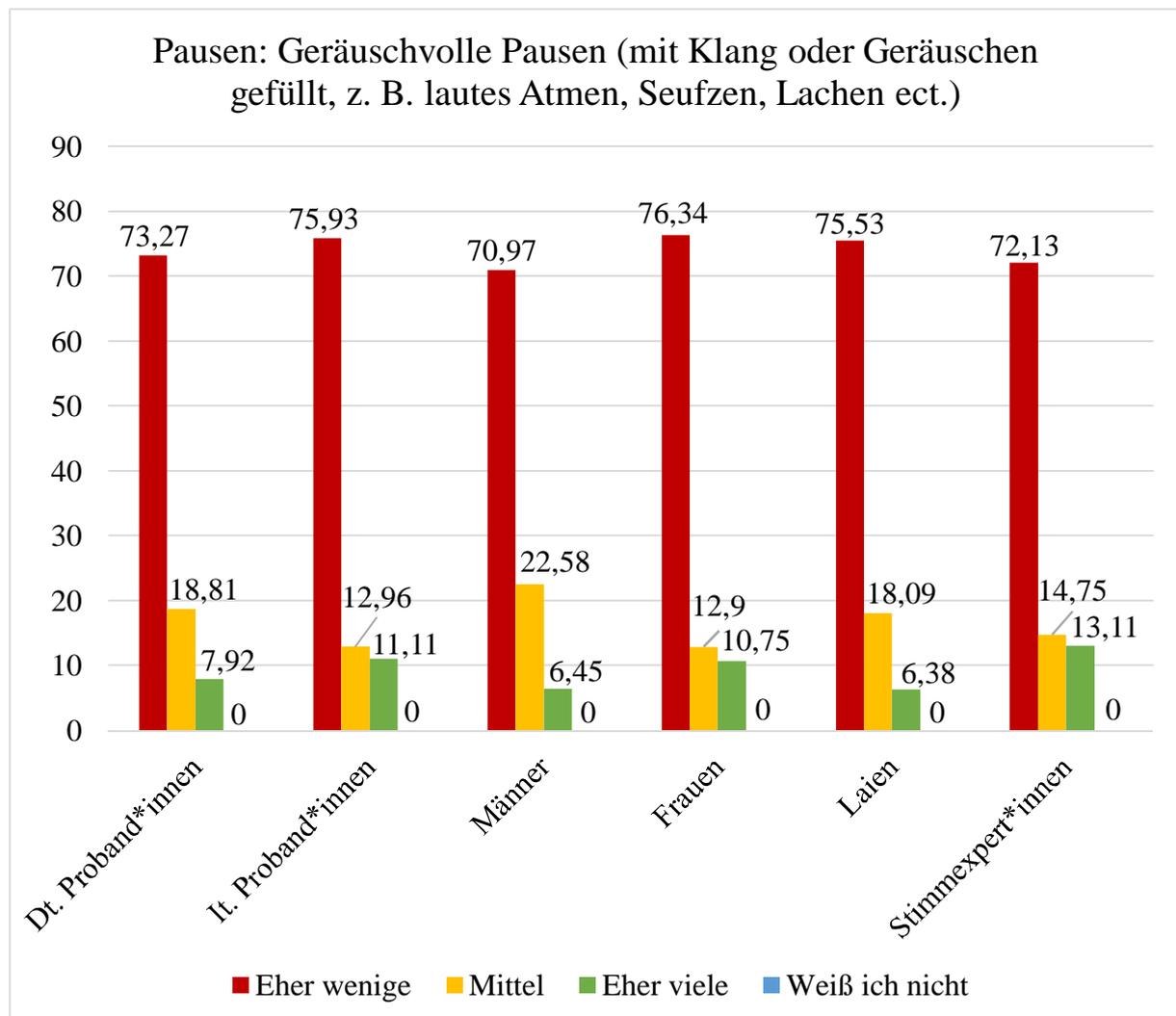
3.3.4.2.5 Pausen

Häufigkeit



Das fünfte analysierte prosodische Merkmal sind die Pausen. Die absolute Mehrheit, ca. 68 Prozent der Befragten, nehmen die Bewertung „wenige Pausen“ vor, ca. 25 Prozent nennen eine mittlere Häufigkeit und ca. 6 Prozent „viele Pausen“. Alle können die Frage beantworten, das Ergebnis ist klar und eindeutig. Eine relevante geschlechterbezogene Differenz lässt sich nicht ablesen, ebenso wenig eine sprachlich/kulturelle, wenngleich die deutschsprachigen Proband*innen tendenziell ein klein wenig mehr Pausen vernehmen. Etwas deutlicher ist die berufsbedingte Differenz zwischen Laien und Stimmexpert*innen, wobei letztere mehr Pausen wahrnehmen. Insgesamt sind die Unterschiede zwischen den Gruppen jedoch gering.

Geräuschvolle Pausen



Noch klarer zeigt sich das Ergebnis zur Analyse der geräuschvollen Pausen im Rahmen der ärgerlichen Sprechweisen – die absolute Mehrheit der Proband*innen spricht sich dafür aus, dass es „eher wenige“ geräuschvolle Pausen sind, alle können die Frage beantworten. Leichte Differenzen in der Perzeption lassen sich in Bezug auf das Geschlecht und den Beruf feststellen. Tendenziell nehmen die Männer und die Stimmexpert*innen mehr geräuschvolle Pausen wahr als die Frauen und die Laien.

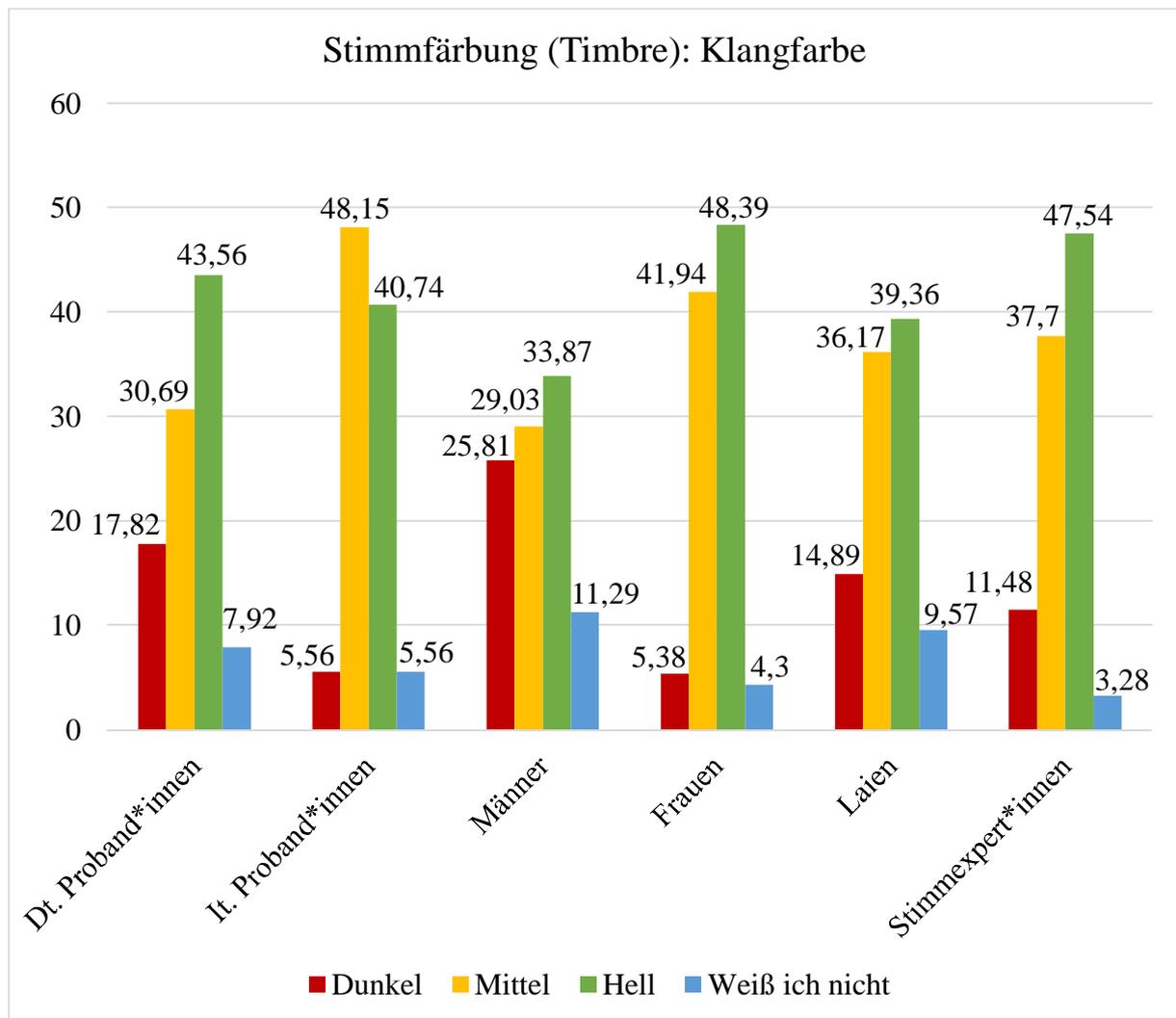
Fazit

Das prosodische Merkmal Pause kann klar perzipiert und eindeutig analysiert werden. Insgesamt werden eher wenige Pausen und eher wenige geräuschvolle Pausen wahrgenommen. Sprachlich/kulturelle Differenzen lassen sich in der Analyse kaum feststellen. Das Geschlecht und der Beruf hingegen scheinen eine Rolle zu spielen, wenngleich die Unterschiede gering

sind. Die Stimmexpert*innen nehmen insgesamt mehr Pausen und auch mehr geräuschvolle Pausen wahr als die Laien, selbiges gilt in Bezug auf die geräuschvollen Pausen auch für die Männer im Vergleich zu den Frauen.

3.3.4.2.6 Stimmfärbung (Timbre)

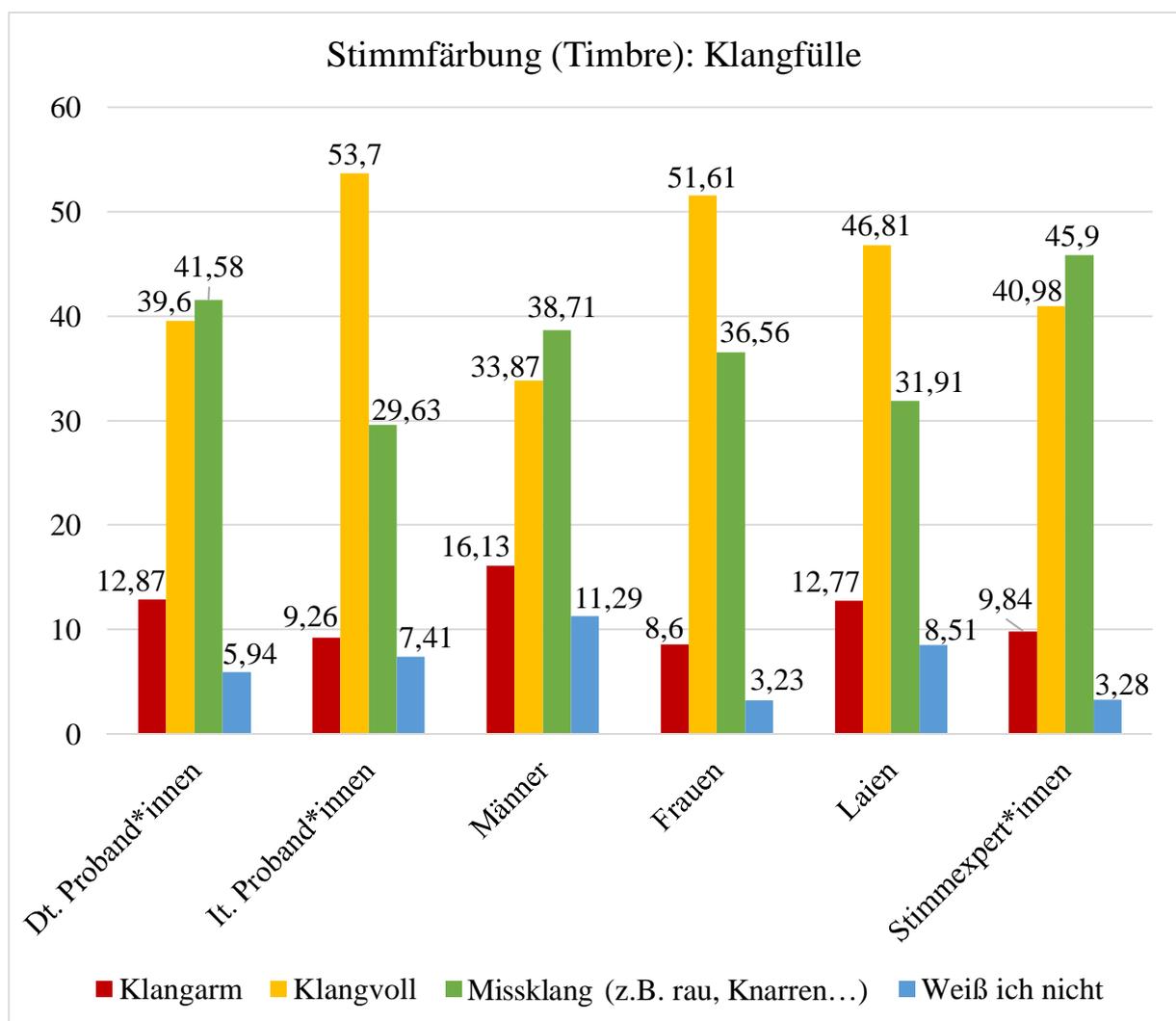
Klangfarbe



Bei der Analyse der Stimmfärbung (Timbre) wurde zunächst die Klangfarbe der ärgerlichen Sprechweisen untersucht. Die Antworten ergeben kein klares Ergebnis, wobei die meisten Proband*innen insgesamt eine helle Klangfarbe angeben. Auffallend groß sind auch die Differenzen zwischen den einzelnen Gruppen in Bezug auf die Faktoren Sprache/Kultur, Geschlecht und Beruf. Innerhalb der italienischsprachigen Gruppe nehmen die Proband*innen im Unterschied zu den deutschsprachigen Personen mehrheitlich eine eher mittlere Klangfarbe wahr,

allerdings ist die Zahl derer geringer, die eine dunkle Klangfarbe angeben. Somit ist kein eindeutig auslegbares Ergebnis möglich. Klar ersichtlich scheint in Bezug auf das Geschlecht hingegen die Tatsache zu sein, dass die Frauen die Klangfarbe als sehr viel heller einstufen als die Männer, welche zudem in größerer Häufigkeit die Frage nicht beantworten können. Selbiges gilt auch für die Stimmexpert*innen im Vergleich mit den Laien. Insgesamt fällt auf, dass ein recht großer Prozentsatz der Proband*innen (11,29 Prozent bei den Männern) die Frage nicht beantworten können.

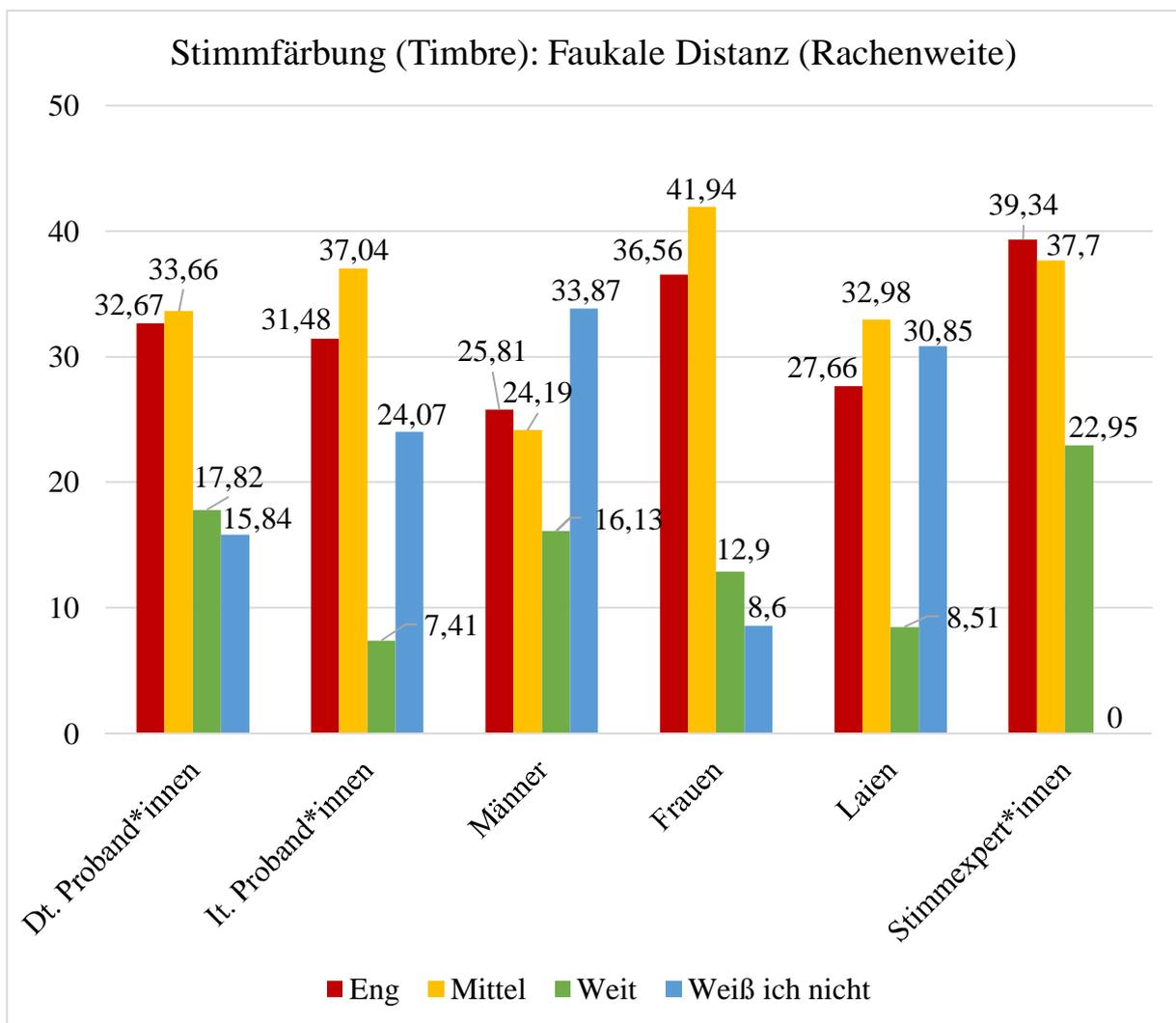
Klangfülle



Bei der Bestimmung der Klangfülle der ärgerlichen Sprechweisen zeigt sich auch kein klares Ergebnis – insgesamt nehmen die meisten Personen zwar ein klangvolles Timbre wahr, viele perceive aber auch einen Missklang. Auch die Zahl derer, die diese Frage nicht beantworten

können, ist relativ hoch, insbesondere bei den Männern (11,29 Prozent). Die Faktoren Sprache/Kultur, Geschlecht und Beruf spielen hier eine überaus bedeutende Rolle. Besonders ins Auge sticht zunächst der Unterschied zwischen den deutsch- und italienischsprachigen Proband*innen, wobei letztere das Timbre als deutlich klangvoller wahrnehmen. Selbiges gilt auch für die Frauen im Unterschied zu den Männern, welche zudem häufiger die Bewertung „klangarm“ vornehmen. Auch die Laien erleben das Timbre als wesentlich klangvoller als die Stimmexpert*innen. Letztere nehmen am häufigsten auch einen Missklang wahr.

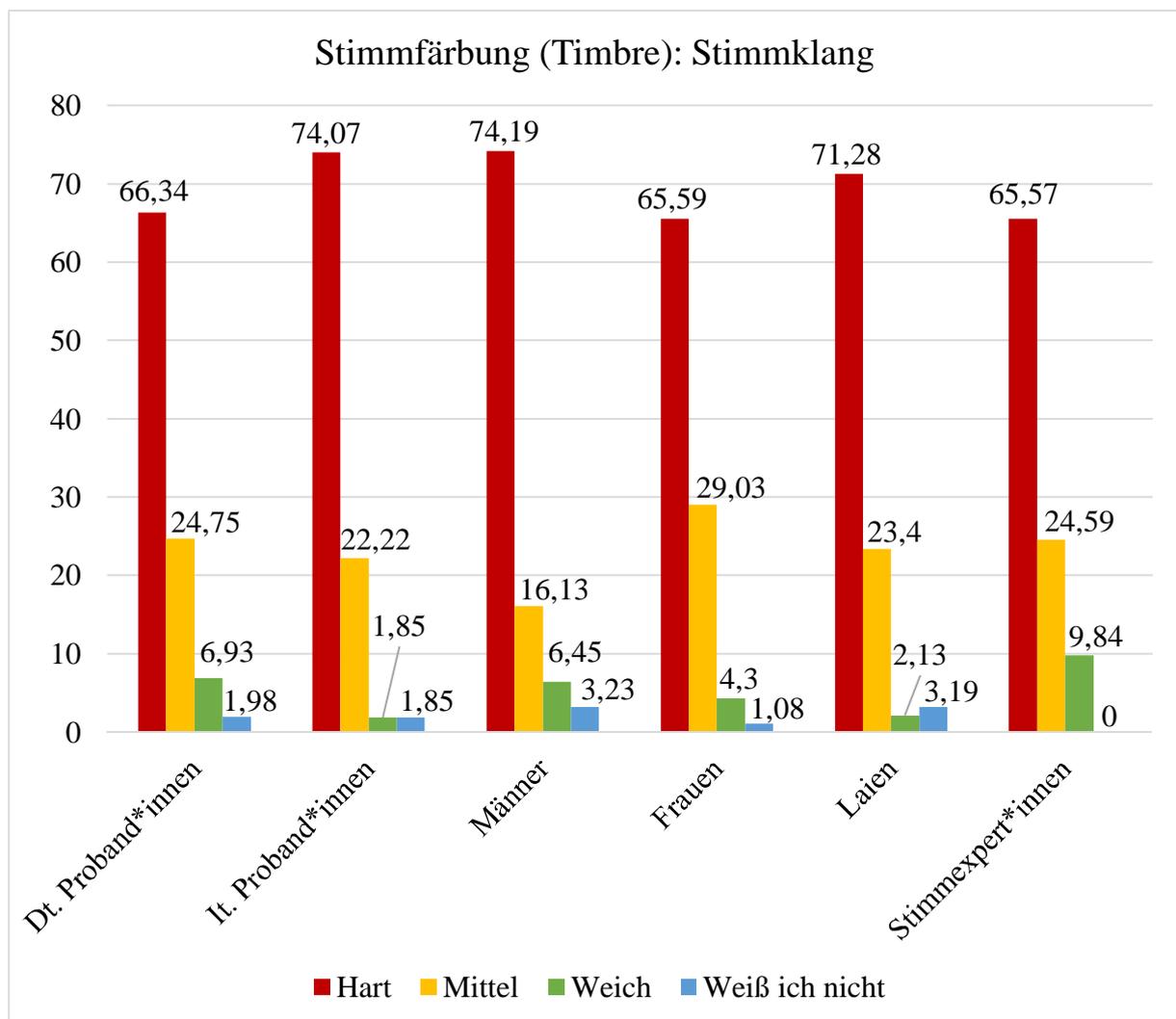
Faukale Distanz (Rachenweite)



Als besonders schwierig erweist sich auch hier, ähnlich wie bei der Analyse der freudigen Sprechweisen, die Beurteilung der faukalen Distanz (Rachenweite). Insgesamt wird eine mittlere Rachenweite wahrgenommen, die eher zu einer Enge tendiert. Eine sehr große Zahl der Personen kann diese Frage nicht beantworten, dies trifft v. a. auf die Laien und auf die Männer zu. Alle

Stimmexpert*innen hingegen beantworten diese Frage. Sie perzipieren im Unterschied zu den Laien auch eine eher enge Rachenweite, wobei sich aber gleichzeitig 22,95 Prozent für eine weite Rachenweite entscheiden. Die Frauen und die Männer nehmen eine mittlere bis enge Rachenweite wahr, 33,87 Prozent der Männer können die Frage aber nicht beantworten. Auch deutlich mehr italienischsprachige als deutschsprachige Proband*innen beantworten die Frage mit „weiß ich nicht“ und beurteilen die Rachenweite als etwas enger als die deutschsprachige Gruppe.

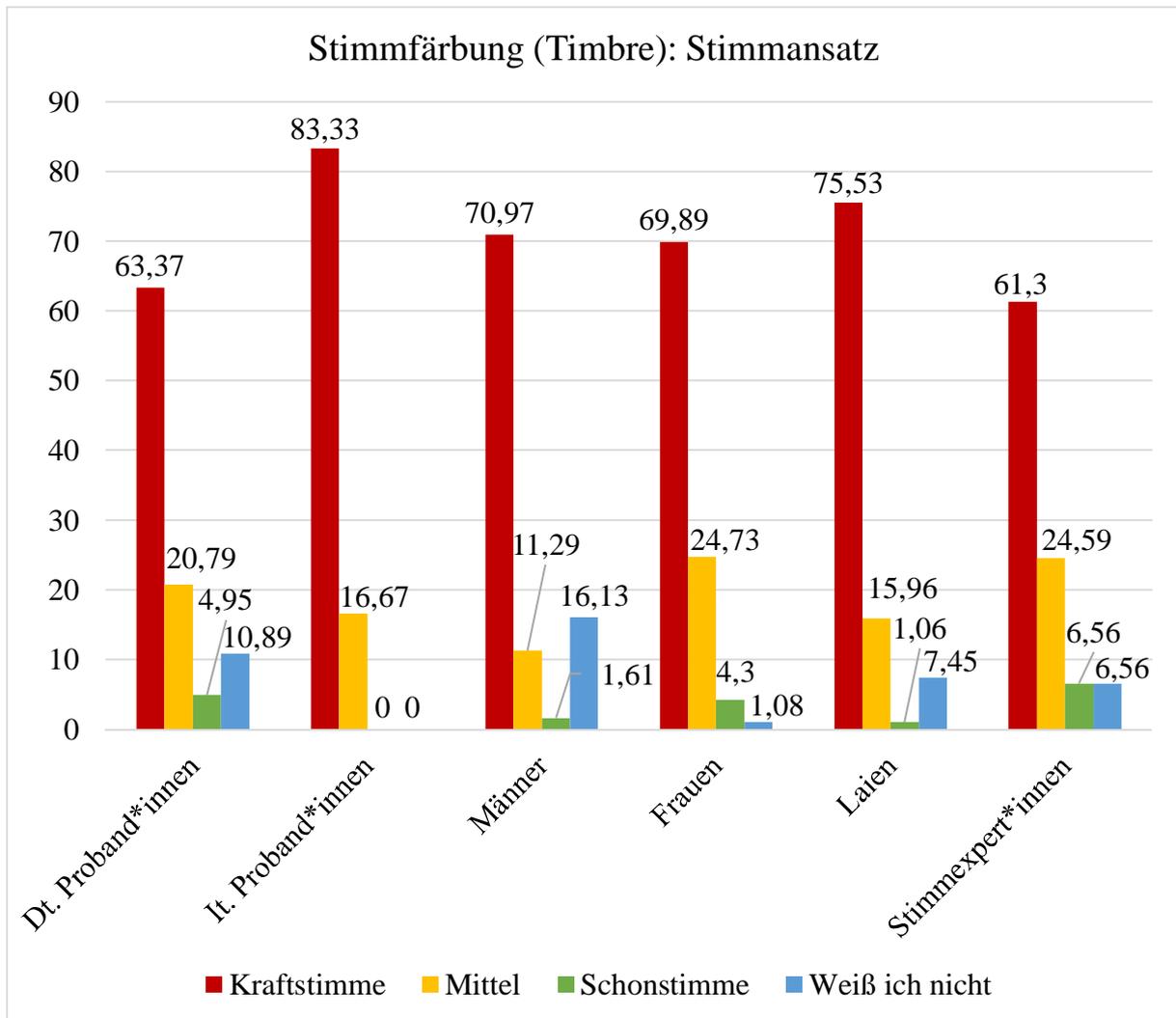
Stimmklang



Die Eigenschaft des Stimmklanges der ärgerlichen Sprechweisen konnte eindeutig und klar definiert werden. Der Großteil der Proband*innen nimmt einen harten Stimmklang wahr, nur etwa 22 Prozent einen mittleren und kaum jemand einen weichen. Außerdem konnten sehr we-

nige Personen diese Frage nicht beantworten. Sprach-/kulturbedingte, geschlechter- und berufsbezogene Differenzen finden sich in geringem Ausmaß. So perzipieren die italienischsprachigen Proband*innen, die Männer und die Laien den Stimmklang als härter als die deutschsprachigen Personen, die Frauen und die Stimmexpert*innen.

Stimmansatz



Als letztes wurde der Stimmansatz beurteilt. Die breite Mehrheit der Proband*innen gibt an, die Verwendung der Kraftstimme zu perzipieren. Kaum jemand nennt den Gebrauch der Schonstimme. Allerdings können recht viele Personen die Frage nicht beantworten, v. a. die Männer (16,13 Prozent). Folgende sprachlich/kulturelle, geschlechterbedingte und berufsbezogene Differenzen lassen sich feststellen: 83,33 Prozent der italienischsprachigen, aber nur 63,37 Prozent der deutschsprachigen Proband*innen perzipieren die Verwendung der Kraftstimme, nehmen den Stimmansatz also als kräftiger wahr als die deutschsprachigen Personen. Mehr Männer als

Frauen können die Frage hingegen nicht beantworten, auch sprechen sich weniger Männer für einen mittleren Stimmansatz aus. Wichtig ist auch die Differenz zwischen der Analyse der Laien und der Stimmexpert*innen, welche einen weicherer Stimmansatz perzipieren.

Fazit

Insgesamt wird das prosodische Merkmal Stimmfärbung (Timbre) wie folgt perzipiert. Die Klangfarbe wird als hell, das Timbre als klangvoll mit mittlerer bis hin zu enger Rachenweite, mit hartem Stimmklang und dem Gebrauch der Kraftstimme wahrgenommen. Dieses prosodische Merkmal ist ein sehr vielschichtiges und komplexes, was sich auch im Rahmen der Ergebnisse zeigt, die nicht immer klar und eindeutig sind. Besonders einfach zu beurteilen waren der Stimmklang und der Stimmansatz, besonders schwierig die faukale Distanz. Bei der Beurteilung aller Eigenschaften zeigten sich z. T. sehr große sprachlich/kulturelle, geschlechterspezifische und berufsbedingte Differenzen. Die Analysen der Männer und jene der Laien sind bei der Beurteilung der Rachenweite wesentlich undifferenzierter als die Analysen der Frauen und der Stimmexpert*innen. Die italienischsprachigen Proband*innen perzipieren das Timbre als klangvoller, härter und kräftiger als die deutschsprachigen Proband*innen.

3.3.4.2.7 Gesamtergebnis zur Perzeption und Analyse der prosodischen Merkmale – Ärger

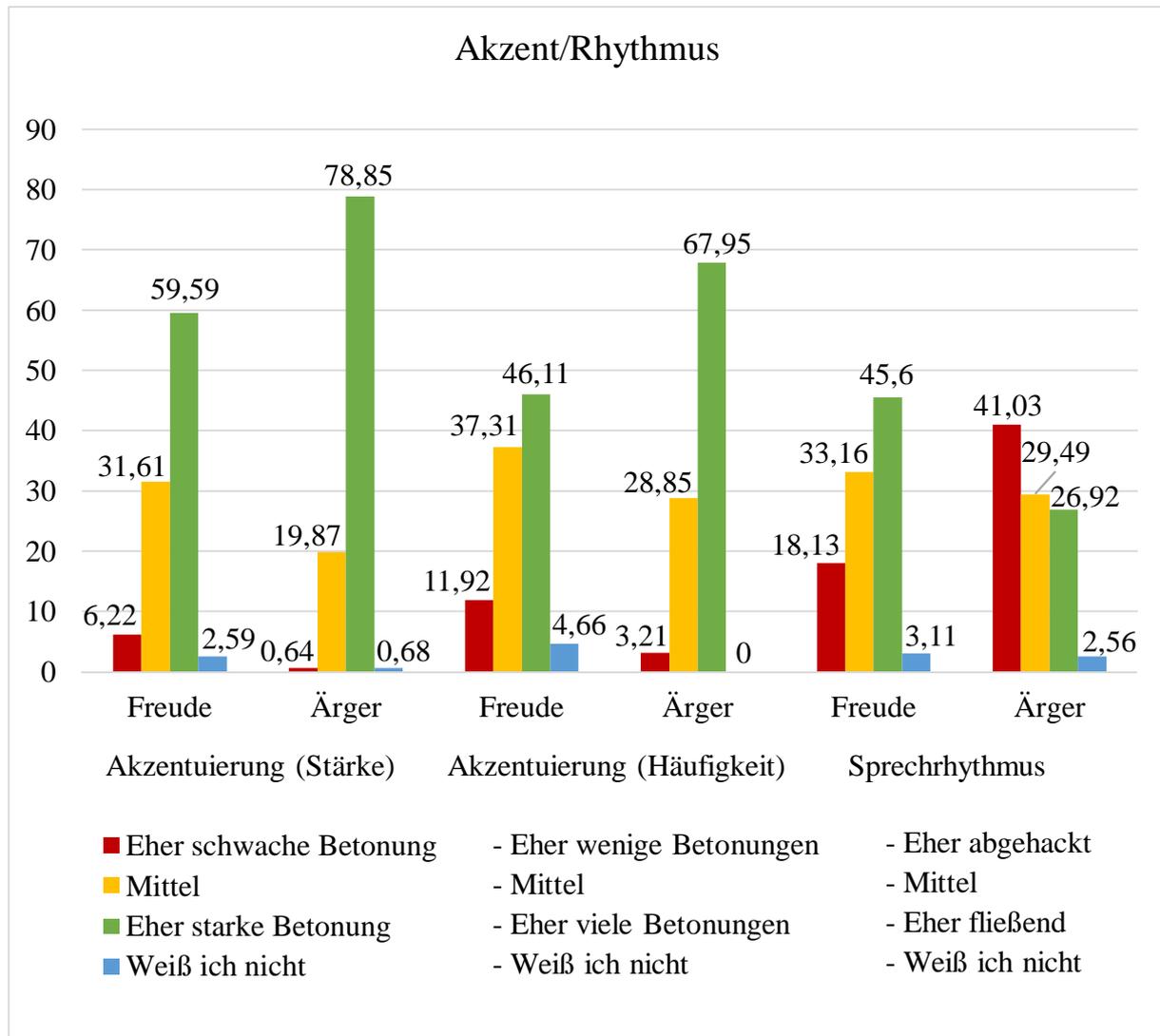
In der folgenden Tabelle sind die zentralen Eigenschaften der einzelnen prosodischen Merkmale, wie sie von der Mehrheit aller Proband*innen in Bezug auf die ärgerlichen Sprechweisen rezipiert werden, überblicksartig zusammengefasst.

Gesamtergebnis zur Perzeption und Analyse der ärglichen Sprechweisen			
Akzent/ Rhythmus	Akzentuierung (Stärke) : eher starke Betonung	Akzentuierung (Häufigkeit) : eher viele Betonungen	Sprechrhythmus : eher abgehackt (staccato)
Tonhöhe/ Melodieführung	Sprechstimmlage : eher hoch	Tonhöhenverlauf: bewegt	Tonhöhenumfang : mittel
Lautstärke	Lautstärke : eher laut	Lautheitsverlauf : gleichförmig	Lautheitsverlauf am Ende der Phrase : gleichbleibende Lautstärke
Spreche- schwindigkeit	Geschwindigkeit : eher schnell	Geschwindigkeitsverlauf: gleichförmig	Geschwindigkeitsverlauf am Ende der Phrase: gleichbleibende Geschwindigkeit
Pausen	Häufigkeit : wenige Pausen	Geräuschvolle Pausen : eher wenige	
Stimmfärbung (Timbre)	Klangfarbe : hell	Klangfülle : klangvoll	Faukale Distanz : mittel
			Stimmklang : hart
			Stimmansatz : Kraftstimme

3.3.4.3 Gegenüberstellung Freude – Ärger

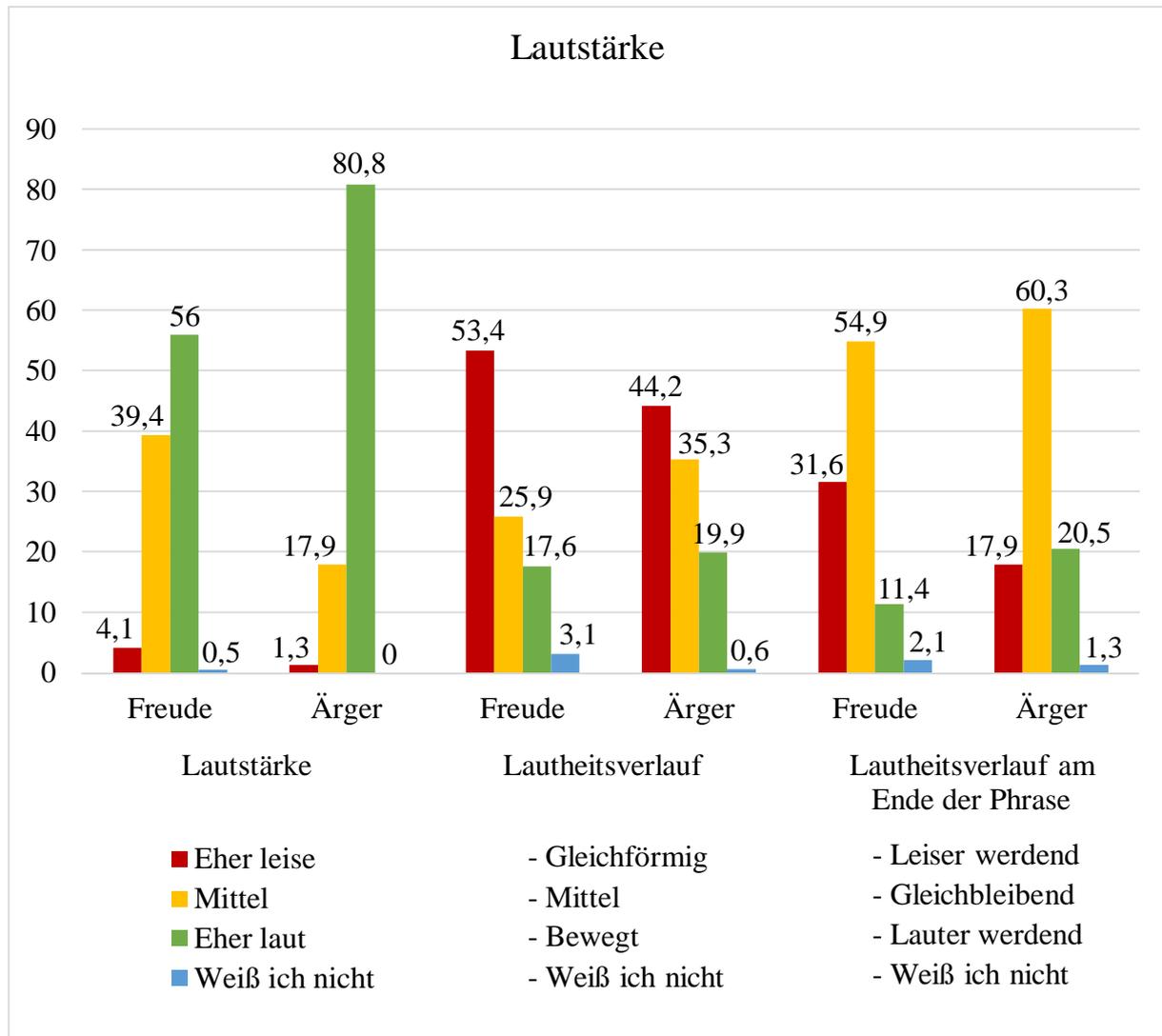
Im Folgenden sollen die Ergebnisse aller Proband*innen, in den beiden emotionalen Ausprägungen Freude und Ärger, zusammengefasst und miteinander verglichen werden. So können die Unterschiede in der Perzeption und der Ausprägung der prosodischen Merkmale und Eigenschaften in Bezug auf die jeweilige Emotion herausgearbeitet und veranschaulicht werden. Konnte in den bisherigen Aufstellungen eine differenzierte Übersicht über die Unterschiede in Bezug auf die Sprache/Kultur, das Geschlecht und den Beruf sowie in den Tabellen am jeweiligen Kapitelende eine generelle Übersicht über die perzipierten Merkmalseigenschaften geleistet werden, dient die folgende Aufstellung der präzisen Gegenüberstellung des jeweiligen Perzeptionseindrucks und der entsprechenden prosodischen Ausprägungen. Alle Angaben sind in Prozent. Pro Diagramm sind, mit Ausnahme der Stimmfärbung, alle drei bzw. zwei Eigenschaften des jeweiligen prosodischen Merkmals dargestellt.

3.3.4.3.1 Akzent/Rhythmus



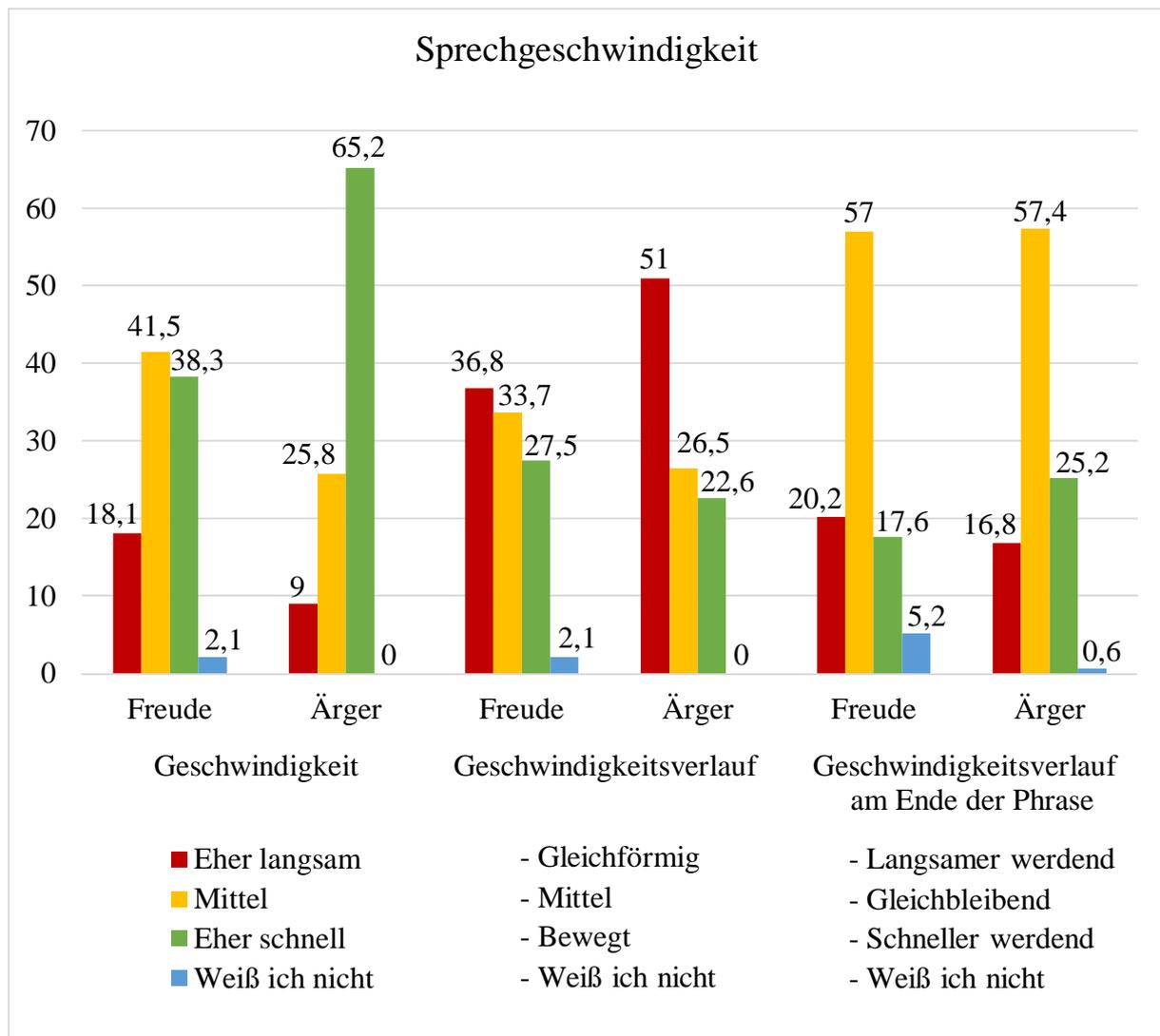
Was die Perzeption der Stärke der Akzentuierung betrifft, so werden sei es bei den freudigen als auch bei den ärgerlichen Sprechweisen eher starke Betonungen vernommen. Trotzdem lässt sich ein klarer Unterschied zwischen den beiden emotionalen Ausprägungen beobachten – bei den freudigen Sprechweisen werden im Vergleich zu den ärgerlichen die Betonungen als insgesamt viel schwächer perzipiert. Ein ähnliches Ergebnis zeigt sich bei der Analyse der Häufigkeit der Akzentuierung als zweite Eigenschaft des Akzentes/Rhythmus. Werden bei beiden emotionalen Ausprägungen zwar insgesamt eher viele Betonungen wahrgenommen, gilt dies doch in besonderem Maße für die ärgerlichen Sprechweisen. Eine völlig konträre Bewertung nahmen die Proband*innen bei der Beurteilung des Sprechrhythmus vor. Werden die freudigen Äußerungen als eher fließend perzipiert, werden die ärgerlichen mit einem eher abgehackten

3.3.4.3.3 Lautstärke



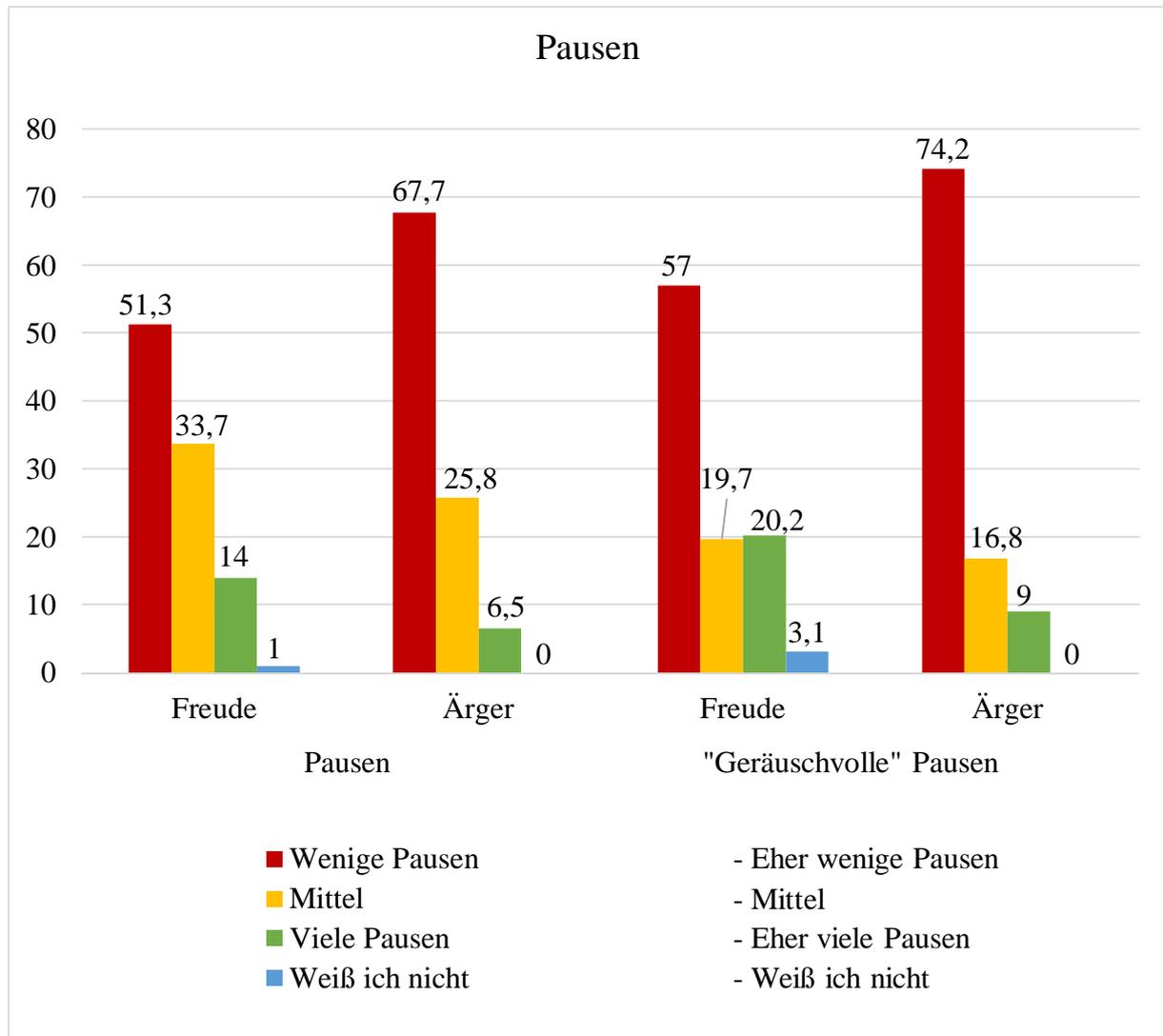
Das prosodische Merkmal Lautstärke konnte klar von den Proband*innen herausgehört und analysiert werden und fast niemand konnte die Fragen nicht beantworten. Die emotionalen Sprechweisen zu Freude und Ärger werden beide als eher laut perzipiert, wobei jedoch ein großer Unterschied zwischen den Emotionen ersichtlich ist (die ärgerlichen werden als deutlich lauter empfunden als die freudigen Sprechweisen). Der Lautheitsverlauf wird in beiden Fällen als gleichförmig bewertet, wobei die ärgerlichen Sprechweisen im Vergleich zu den freudigen als etwas bewegter eingestuft werden. Auch der Lautheitsverlauf am Ende der Phrase gilt als gleichbleibend. Hier werden die freudigen Äußerungen jedoch insgesamt als leiser perzipiert als die ärgerlichen.

3.3.4.3.4 Sprechgeschwindigkeit



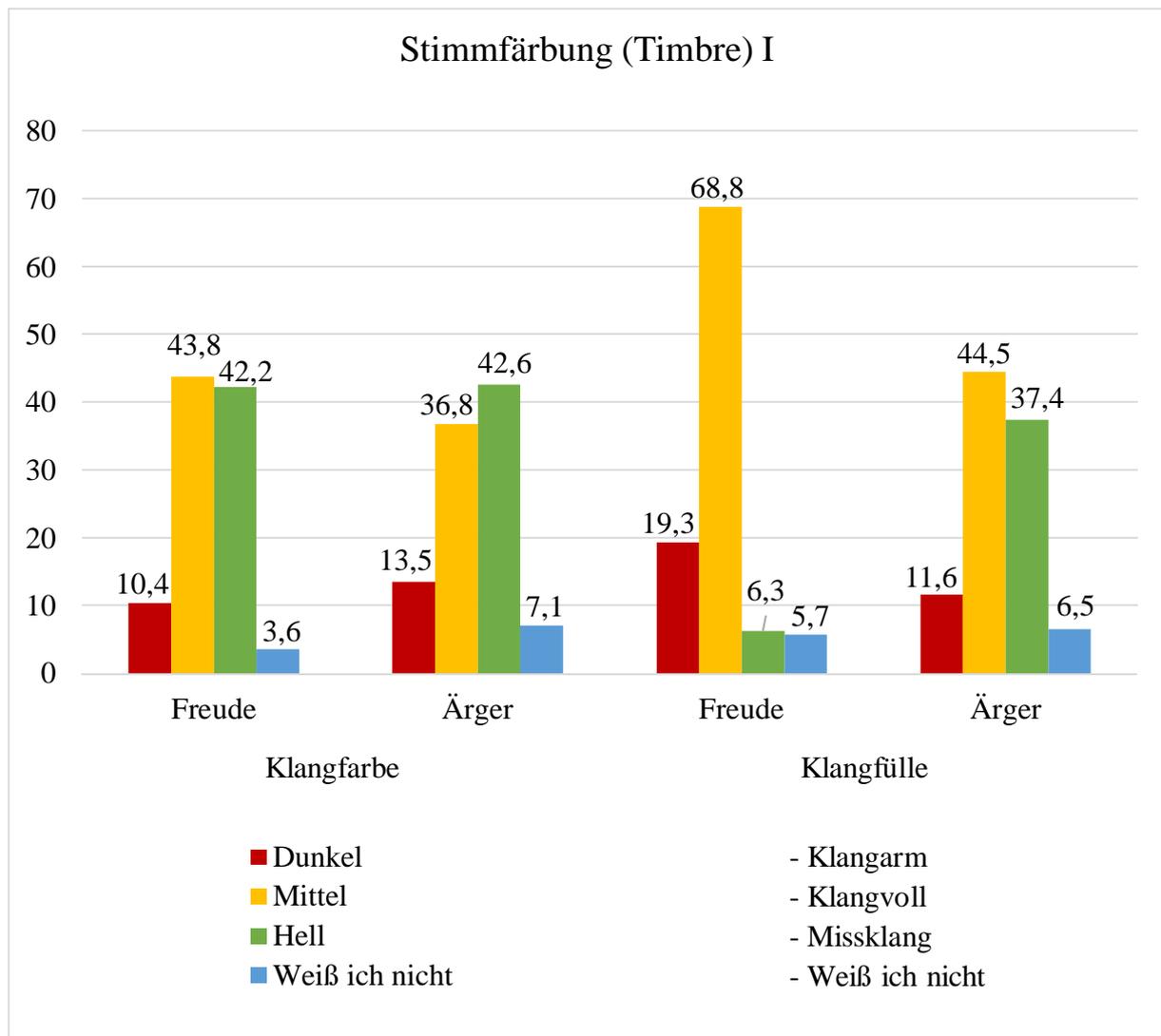
Große Unterschiede zwischen den beiden emotionalen Ausprägungen lassen sich hier v. a. bei der Perzeption der Geschwindigkeit ausmachen. Während die freudigen Sprechweisen als mittelschnell bis schnell wahrgenommen werden, beurteilen die Proband*innen die ärgerlichen Sprechweisen als viel schneller. Der Geschwindigkeitsverlauf wird bei letzteren als gleichförmiger betrachtet als bei den freudigen Sprechweisen, bei denen ein gleichförmiger bis mittlerer Geschwindigkeitsverlauf perzipiert wird. Am Ende der Phrase ist der Geschwindigkeitsverlauf in beiden emotionalen Sprechweisen gleichbleibend, allerdings wird er im Unterschied zur Freude bei der Emotion Ärger eher als schneller werdend betrachtet. Nur ein geringer Prozentsatz an Personen konnte die Fragen nicht beantworten, wobei dies hauptsächlich für die Beurteilung der ärgerlichen Sprechäußerungen zutrifft.

3.3.4.3.5 Pausen

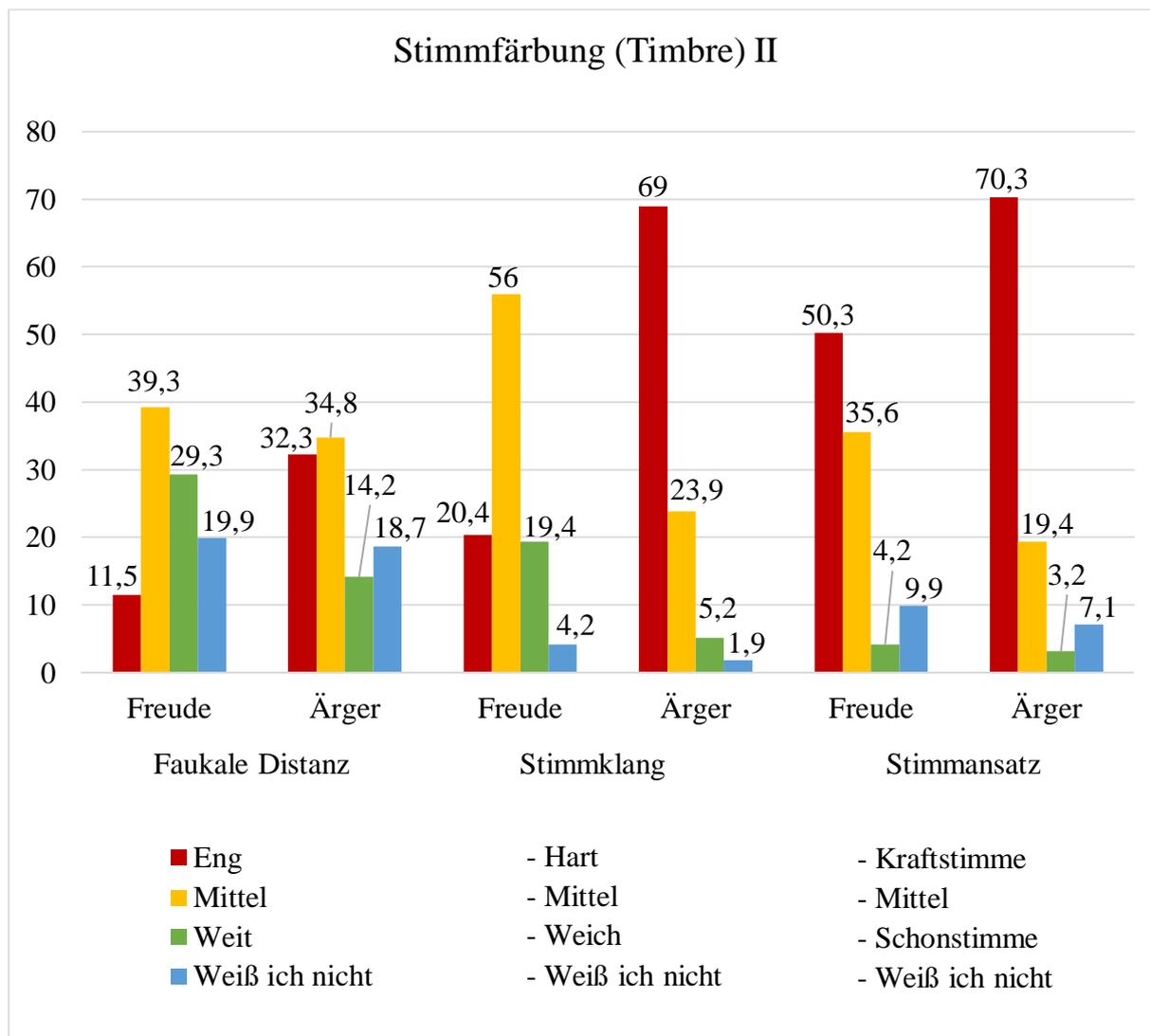


In Bezug auf das prosodische Merkmal der Pause wurden die Häufigkeit der Pausen und jene der geräuschvollen Pausen analysiert. Darunter versteht man Pausen, welche mit Klang oder Geräuschen gefüllt sind, z. B. lautes Atmen, seufzen, lachen usw. Obwohl die Proband*innen bei beiden Sprechweisen dieselbe Bewertung vorgenommen haben, ist doch ein deutlicher Unterschied zwischen den emotionalen Ausprägungen ersichtlich. Bei den ärgerlichen Sprechweisen werden insgesamt viel weniger Pausen und auch viel weniger geräuschvolle Pausen perzipiert als bei den freudigen Äußerungen. Nur wenige Personen konnten die Frage außerdem nicht beantworten. Die Ergebnisse sind sehr klar und eindeutig.

3.3.4.3.6 Stimmfärbung (Timbre)



Um das prosodische Merkmal Stimmfärbung (Timbre) adäquat zu beschreiben, wurden mehrere Eigenschaften analysiert. Zunächst wird die Klangfarbe untersucht, welche bei den freudigen Sprechweisen als mittel bis hell und bei den ärgerlichen Sprechweisen als eher hell wahrgenommen wird. In Bezug auf die Beurteilung der Klangfülle ist trotz derselben Angabe – das Timbre wird in beiden Fällen als klangvoll perzipiert – ein deutlicher Unterschied zwischen den ärgerlichen und freudigen Sprechweisen erkennbar. So werden die Äußerungen zur Freude als deutlich klangvoller wahrgenommen, während bei jenen zum Ärger sehr häufig ein Missklang (z. B. rau, Knarren) vernommen wird. Im Vergleich zu den anderen prosodischen Merkmalen und Eigenschaften ist hier die Zahl derer, die die Frage mit „weiß ich nicht“ beantworten, recht hoch.



Auch bei der Beurteilung der weiteren Eigenschaften des Merkmals Stimmfärbung (Timbre) ist der Prozentsatz der Personen, die die Frage nicht beantworten können, z. T. sehr hoch. Dies gilt insbesondere für die Eigenschaft der faukalen Distanz. Diese wird in Bezug auf die freudigen Sprechweisen als „mittel“ und in Bezug auf die ärgerlichen Sprechweisen als mittel bis eng beurteilt. Der Unterschied tritt hier sehr klar hervor – die freudigen Sprechweisen werden tendenziell mit einer viel weiteren Rachenweite perzipiert als die ärgerlichen. Während der Stimmklang der ärgerlichen Äußerungen als hart empfunden wird, wird jener der freudigen Äußerungen als mittelmäßig und viel weicher beurteilt. In Bezug auf den Stimmansatz wird bei beiden emotionalen Sprechweisen die Verwendung der Kraftstimme angegeben, insgesamt werden die ärgerlichen Äußerungen jedoch als deutlich härter und kräftiger wahrgenommen.

3.3.4.4 Anmerkungen (optional)

Am Ende des Fragebogens hatten die Proband*innen die Möglichkeit, eigene Anmerkungen zur Untersuchung oder zu den Hörbeispielen zu machen. Die Frage war optional. Außerdem bestand die Option, auch zur Frage sieben eigene Kommentare zu machen oder andere Parameter zu nennen, die nicht als Auswahlmöglichkeit im Fragebogen genannt waren. Im Folgenden werden die relevanten Aussagen nach Emotion geordnet im direkten Zitat wiedergegeben. Es wird darauf hingewiesen, dass etwaige orthographische Fehler nicht korrigiert wurden, um die Zitate nicht zu verfälschen.

Anmerkungen zu den freudigen Sprechweisen:

- „Risata di sottofondo finale“ (Zusatzfeld „andere“ bei Frage 7)
- „ist mir nicht so sympathisch“
- „Betonung von „richtig gscheid“
- „Wortbetonung „richtig“
- „Lacht während des Sprechens“
- „rumori di sottofondo. Nessuna particolare“ (Zusatzfeld „andere“ bei Frage 7)
- „siccome sono italiana percepisco questo brano come normale (non troppo emotivo), mentre magari per i tedeschi ha una connotazione più emotiva.“
- „regionale Färbung“
- „Ho colto molte differenze tra la prosodia italiana e quella tedesca: infatti, nonostante la lingua tedesca abbia delle parole molto "secche" e "dure", la musicalità non cambia. La ragazza che parla ha comunque un tono dolce, nonostante il parlato sia molto duro.“
- „lang gezogene Vokale“
- „finde des Sprechers Freude kann noch mit Begeisterung beschrieben werden, man merkt beruflichen Kontext“
- „sembra una cantante lirica“
- „Una voce felice!“
- „klingt sympathisch und irgendwie überwältigt“
- „Dà l'impressione di star sorridendo“
- „Lachen erleichtert auch die Interpretation“

Die Anmerkungen zeigen die persönliche Wahrnehmung der Person zu bestimmten Aspekten wie Stimmklang, Sympathie oder Assoziationen auf, es wird auf z. T. regionale Anklänge der

Sprecher*innen und auf Hintergrundgeräusche, bestimmte Betonungen und Lachen hingewiesen. Eine italienischsprachige Probandin teilt ihre Vermutung zu einem sprachlich/kulturellen Unterschied in der Emotionsinterpretation mit.

Anmerkungen zu den ärgerlichen Sprechweisen:

Hier wurde des Öfteren darauf hingewiesen, dass die Frage zur Emotionsausprägung nicht korrekt angezeigt wird. Im Kapitel 3 wurde erläutert, dass hier ein technischer Fehler im System aufgetreten ist. Die einzelnen Anmerkungen hierzu werden folgend nicht aufgelistet.

- Ausruf „Ev“ schwer mit einzuordnen und leichter Missklang“
- „voce isterica“
- „Stimme weinerlich, sich überschlagend“
- „aggressive Stimme; starke Betonung am Ende wird bei dem Wort "me" durch hartes Klopfen auf die Brust unterstützt“
- „Sembra disperato, estremamente arrabbiato“
- „leichter inspiratorischer Stridor“

Die Anmerkungen zu den ärgerlichen Sprechweisen sind hier deutlich weniger als jene zu den freudigen Sprechweisen, was allerdings damit in Verbindung steht, dass öfters auf den Systemfehler zur Frage zur Emotionsausprägung verwiesen wurde. Auf weitere Bemerkungen wurde dann wohl eher verzichtet. Die obigen Aussagen zeigen wieder persönliche Beobachtungen und Interpretationen deutsch- und italienischsprachiger Proband*innen.

4 Ergebnisinterpretation und Hypothesenbeantwortung

4.1 Forschungsfrage 1

1. Welche prosodischen Faktoren spielen beim Erkennen und Perzipieren einer natürlichen emotionalen Äußerung inwieweit eine Rolle?

Um diese Forschungsfrage adäquat beantworten zu können, wurden in der empirischen Studie zwei Aspekte untersucht – zum einen sollte herausgefunden werden, anhand welcher Parameter die jeweilige Emotion von den Proband*innen erkannt wird und zum anderen sollte untersucht werden, für wie relevant die Proband*innen die einzelnen prosodischen Merkmale für die Perception der emotionalen Äußerungen halten. Im Folgenden werden also, zur Beantwortung der oben genannten ersten Forschungsfrage, die Ergebnisse beider Teilfragen erörtert.

Als herausragendes Merkmal bei der Emotionsperzeption wurde, unabhängig von der jeweiligen emotionalen Ausprägung Freude bzw. Ärger, die Tonhöhe/Melodieführung genannt. In absteigender Relevanz gaben die Proband*innen in Frage eins des Fragebogens insgesamt folgende Parameter zur Emotionserkennung an: Tonhöhe/Melodieführung (83,05 Prozent), Akzent/Rhythmus (65,68 Prozent), Lautstärke (59,32 Prozent), Stimmfärbung (Timbre) (49,58 Prozent), Sprechgeschwindigkeit (49,15 Prozent), Text/Wortwahl (46,61 Prozent) und Pausen (20,34 Prozent).

Als Antwort auf die zweite Frage, für wie wichtig jedes einzelne Merkmal zur Bestimmung der Äußerung gehalten wird, wurde folgendes Ranking aufgestellt: Die Tonhöhe/Melodieführung wird von 91,78 aller Proband*innen als relevant erachtet, gefolgt von Akzent/Rhythmus mit 82,43 Prozent. An dritter Stelle steht die Lautstärke mit 76,77 Prozent, 64,88 Prozent nennen die Stimmfärbung (Timbre), 63,74 Prozent die Sprechgeschwindigkeit und 35,98 Prozent die Pausen. Diese Übersicht ist konform mit den Ergebnissen zu Frage eins, wenngleich die Prozentangaben differieren. Das Ranking bleibt hingegen dasselbe.

Es ist also klar ersichtlich, dass insbesondere die Tonhöhe/Melodieführung zentral ist, aber auch alle anderen Kategorien bei der Zuordnung des Gehörten eine Rolle spielen. Dieses Ergebnis stimmt überein mit einer Untersuchung zur automatischen Spracherkennung, in welcher ebenfalls die Tonhöhe bzw. der Tonhöhenverlauf als relevantestes Merkmal zur Erkennung von natürlichen Emotionen in Kombination mit anderen prosodischen Merkmalen genannt wurde (Busso/Bulut/Lee/Narayanan 2009). Auch der Inhalt des Gesagten (Kategorie Text/Wortwahl)

spielt eine Rolle, wenngleich keine allzu wichtige. Dies ist darauf zurückzuführen, dass zum einen die Turns so ausgewählt wurden, dass nicht der Inhalt allein ausschlaggebend für die Emotionserkennung ist, und zum anderen auch die Fragestellung den Fokus auf die prosodischen Merkmale lenkt, welche im Zentrum der Untersuchung stehen.

Bei der Emotionszuordnung spielen aber nicht nur die prosodischen Merkmale und deren Ausprägung als solche, sondern auch die zugrundeliegende jeweilige Emotion eine Rolle. So konnten für die Sprechweisen zu Freude im Vergleich mit jenen zu Ärger z. T. deutliche Differenzen in Bezug auf die erste Frage festgestellt werden, woran die Proband*innen die jeweilige den Sprechweisen zugrundeliegende Emotion erkennen. Dass der Akzent/Rhythmus beim Erkennen der ärgerlichen Sprechweisen eine größere Rolle spielt als bei den freudigen, wird auf die jeweils deutlich stärker perzipierte Betonung und auf die größere Zahl derselben zurückgeführt. Es wird vermutet, dass auch der eher abgehackte Rhythmus, der die ärgerlichen Sprechweisen im Vergleich zu dem als fließend perzipierten Sprechrhythmus der freudigen Sprechweisen ausmacht, als markanteres Merkmal eher die Aufmerksamkeit der Hörer*innen steigert. Eine ähnliche Erklärung wird für die Lautstärke angenommen, welche beim Erkennen der Sprechäußerungen zum Ärger eine größere Rolle spielt als bei jenen zur Freude. Erstere werden als deutlich lauter perzipiert als die freudigen Sprechweisen und wirken dadurch wohl stärker auf die Zuhörer*in ein. Selbiges gilt für die Sprechgeschwindigkeit – sie ist beim Erkennen der ärgerlichen Äußerungen viel relevanter als beim Erkennen der freudigen Sprechweisen, was wohl daran liegt, dass die ärgerlichen Äußerungen von den Proband*innen als viel schneller wahrgenommen werden als die freudigen. Tendenziell gilt auch die Stimmfärbung (Timbre) als wichtiger bei den ärgerlichen als bei den freudigen Sprechweisen. Dies wird damit erklärt, dass im Rahmen der ärgerlichen Äußerungen deutlich öfters ein Missklang verzeichnet wurde und das Timbre insgesamt als härter und kräftiger perzipiert wird. Das weichere, weitere und klagvollere Timbre des freudigen Ausdrucks wirkt wahrscheinlich angenehmer auf die Hörer*in, aber weniger stimulierend. Entwicklungsbiologisch betrachtet ergibt es Sinn, Ärger rascher als Freude zu erkennen, um gegebenenfalls entsprechend reagieren zu können. Der Mensch als soziales Wesen hat im Laufe der Evolution auch gelernt, Gefühlsäußerungen des Gegenübers zu erkennen und adäquat darauf zu reagieren. Im Rahmen dieser Studie zeigt sich, dass die emotionalen Sprechweisen zu Ärger insgesamt deutlicher und klarer erkannt und analysiert werden können als jene zu Freude.

4.1.1 Forschungsfrage 1a

1.a Gibt es bei oben genannter Frage einen sprach/kulturbedingten, mit der Fähigkeit des funktionellen Hörens (durch entsprechende Ausbildung/berufliche Tätigkeit) verknüpften oder geschlechterspezifischen Unterschied?

In einem zweiten Schritt soll die Forschungsfrage, welche prosodischen Faktoren beim Erkennen und Perzipieren einer natürlichen emotionalen Äußerung auf welche Weise eine Rolle spielen, in Bezug auf die Einflussfaktoren Sprache/Kultur, Geschlecht und Beruf beantwortet werden. Dabei werden wieder die Ergebnisse beider Teilfragen referiert, um ein aussagekräftiges Bild darüber zu erhalten, anhand (1) welcher Parameter die emotionalen Sprechweisen überhaupt erkannt wurden und (2) für wie relevant die prosodischen Merkmale zur Perzeption derselben von Seiten der Proband*innen eingeschätzt werden.

4.1.2 Akzent/Rhythmus

1. Die Rolle des Merkmals Akzent/Rhythmus im Perzeptionsprozess der emotionalen Sprechweisen

Sprache/Kultur

67,47 Prozent der deutschsprachigen und 61,43 Prozent der italienischsprachigen Proband*innen geben an, anhand dieses Merkmals die Emotion der Äußerung zu perzipieren. Damit kommt diesem Parameter, nach der Tonhöhe/Melodieführung, die zweitwichtigste Rolle im Perzeptionsprozess zu. Dies gilt für beide Sprach-/Kulturgruppen, wobei der Akzent/Rhythmus jedoch für die deutschsprachigen Proband*innen eine etwas größere Rolle spielt als für die italienischsprachigen, sei es bei den Äußerungen zu Freude als auch bei jenen zu Ärger in gleichem Ausmaß.

Die hohe Relevanz dieses Merkmals für den Perzeptionsprozess natürlicher emotionaler Sprechweisen wird u. a. darauf zurückgeführt, dass die Betonung und Akzentuierung in vielen Sprachen auch die Bedeutung der Äußerung verändern (Gallés 2005). Dies gilt auch für das Deutsche und Italienische, hier kann der Wortakzent bedeutungsdistinktiv sein (z. B. *meta – metà*). Im Deutschen können Wortpaare gebildet werden, die sich nur in der Akzentuierung semantisch unterscheiden, z. B. *übersetzen* versus *übersetzen* (Ferrer 1999). Dass der Akzent/Rhythmus für die deutschsprachige Gruppe relevanter ist als für die italienischsprachige,

wird damit erklärt, dass das Deutsche über einen akzentzählenden Rhythmus verfügt, der durch einen harten, stoßenden Rhythmus gekennzeichnet ist, während das Italienische durch einen silbenzählenden Rhythmus mit geringen Intensitätsunterschieden gekennzeichnet ist (Keßler 2013).

Beruf

Aufgrund der Tatsache, dass das Merkmal Akzent/Rhythmus bedeutungsdistinktiv sein kann, spielt es in beiden Sprachen eine relevante Rolle in der Sprachperzeption. Aus diesem Grund verfügen auch die nicht ausgebildeten Laienhörer*innen über eine gut ausgebildete Fähigkeit, dieses Merkmal zu perzipieren. So lassen sich in Bezug auf den Akzent/Rhythmus im Rahmen dieser Untersuchung keine relevanten berufsbezogenen Differenzen feststellen. In der italienischsprachigen Gruppe halten die Laien dieses Merkmal für etwas wichtiger als die Stimmexpert*innen.

Geschlecht

In Bezug auf die Frage, woran die Proband*innen die emotionalen Sprechweisen erkennen, gibt es beim Merkmal Akzent/Rhythmus generell keine großen Unterschiede. Innerhalb der deutschsprachigen Gruppe wird dieses Merkmal jedoch von den Frauen öfters genannt als von den Männern. Bei den italienischsprachigen Proband*innen gibt es keinen geschlechterbedingten Unterschied.

2. Relevanz des Merkmals Akzent/Rhythmus in der Perzeption der emotionalen Sprechweisen

Sprache/Kultur

82,22 Prozent der deutschsprachigen und 82,82 Prozent der italienischsprachigen Proband*innen halten das prosodische Merkmal Akzent/Rhythmus insgesamt für eher bis sehr wichtig. Auch in den Ergebnissen zu Freude und Ärger gibt es kaum Unterschiede. Nennenswerte sprach-/kulturbezogene Differenzen lassen sich in der Bewertung der Relevanz dieses Parameters also nicht feststellen. Dies liegt wohl daran, dass dieses Merkmal aufgrund seiner bedeutungsdistinktiven Eigenschaft in beiden Sprachen für beide Sprach-/Kulturgruppen eine wichtige Rolle spielt.

Beruf

80,47 Prozent der Laien und 86,14 Prozent der Stimmexpert*innen bewerten das Merkmal Akzent/Rhythmus insgesamt mit eher bis sehr wichtig. Diese Differenz zeigt sich insbesondere bei der Beurteilung der freudigen Sprechweisen (Laien: 73,55 Prozent vs. Expert*innen: 81,33 Prozent). Allerdings muss hier genauer differenziert werden, der berufsbezogene Unterschied liegt nämlich v. a. innerhalb des jeweiligen Sprach-/Kulturkreises. So ist dieses Merkmal für die deutschsprachigen Expert*innen (88,46 Prozent) wichtiger als für die deutschsprachigen Laien (77,5 Prozent), aber nicht für die italienischsprachigen Expert*innen (78,79 Prozent) im Vergleich mit den italienischsprachigen Laien (84,21 Prozent). Hier verhält es sich sogar umgekehrt. Der Faktor Beruf spielt also eine Rolle, aber nur innerhalb und in Verbindung mit der jeweiligen Sprache/Kultur. Die Differenzen werden damit erklärt, dass die Expert*innen über ein größeres Wissen über die Akzent- und rhythmische Struktur der eigenen Sprache verfügen und dementsprechend die deutschsprachigen Expert*innen den Akzent/Rhythmus für das akzentzählende, härtere Deutsche für relevanter erachten als die Laien, während diesem Merkmal für das weichere, silbenzählende Italienische umgekehrt eine etwas geringere Bedeutung beigemessen wird.

Geschlecht

Bei der Beurteilung der Relevanz des Merkmals Akzent/Rhythmus lässt sich ein relevanter Unterschied zwischen den Geschlechtern feststellen. 88,32 Prozent der Frauen, aber nur 73,38 Prozent der Männer halten es für eher bis sehr wichtig bei der Perzeption der emotionalen Sprechweisen. Bei der Verteilung zwischen den Beurteilungen zu Freude oder zu Ärger lassen sich dabei kaum Unterschiede erkennen. Auch macht es keinen Unterschied, ob es sich um deutsch- oder italienischsprachige Männer bzw. Frauen handelt. Die geschlechterbedingte Differenz zeigt sich gleichmäßig in beiden Sprach-/Kulturgruppen.

Die Geschlechtsunterschiede werden damit erklärt, dass die Perzeption emotionaler Prosodie in Abhängigkeit verschiedener Persönlichkeitseigenschaften variieren kann. Nicht nur die jeweilige Sprache bzw. Kultur oder das Alter üben einen Einfluss auf die Proband*innen und deren Perzeption emotionaler Prosodie aus, sondern auch das Geschlecht als wesentliches Merkmal der einzelnen Personen (Brück/Kreifelts/Ethofer/Wildgruber 2013). Forschungsergebnisse zu geschlechtsspezifischen Unterschieden in Bezug auf emotional-kommunikative Fähigkeiten lassen sich nicht so klar darstellen, wenngleich es viele Hinweise dazu aus bisherigen empirischen Untersuchungen gibt. Diese berichten vielfach darüber, dass die Frauen im Bereich

Sprache und Kommunikation bessere Fähigkeiten aufweisen als Männer. So sprechen etwa Collignon et al. (2010) von einer weiblichen Überlegenheit im Dekodieren diverser emotionaler Hinweisreize. Die Autoren konnten beobachten, dass audiovisuell, visuell und rein auditiv präsentierte Stimuli von Frauen mit größerer Genauigkeit und höherer Geschwindigkeit korrekt hinsichtlich der zugrundeliegenden Emotion (Ärger oder Ekel) identifiziert werden. Gleichzeitig werden diese Untersuchungsergebnisse oft kritisch betrachtet, da es sich häufig um kurze Experimente oder unzulässige Verallgemeinerungen handelt (Rossell et al. 2002).

Es wird hier davon ausgegangen, dass geschlechtsspezifische Unterschiede in der Perzeption emotionaler Sprechweisen nicht nur auf die Verhaltensebene beschränkt sind, sondern auch auf neurowissenschaftlichen Grundlagen beruhen. Als neuronale Grundlage der Prosodieperzeption wird ein weitreichendes Netzwerk kortikaler und subkortikaler Strukturen angenommen (Brück/Kreifelts/Ethofer/Wildgruber 2013). Bei gleicher Intelligenz unterscheiden sich die Struktur und Form des Gehirns von Männern und Frauen (vgl. Einstein 2007). Dabei wurde meist das Kriterium der Asymmetrie/Symmetrie der Sprachfunktion zur Erklärung der hypothetisierten Geschlechterdifferenz herangezogen, die besagt, dass Männer Sprache hemisphärisch links lateral und Frauen Sprache hemisphärisch bilateral verarbeiten. Diese Hypothese, die auf klinischen Beobachtungen zu Aphasien infolge von Läsionen in der linken Hemisphäre beruht, die bei Männern häufiger auftreten als bei Frauen, konnte in der Folge nicht bestätigt werden. Manche Studien konnten geschlechtsspezifische Lateralitätsunterschiede nachweisen (Harasty et al. 1997), andere zeigten auf, dass die Sprachlateralisierung nicht konstant ist, sondern sich über die Zeit (z. B. im Alter) verändern kann (Ferstl/Kaiser 2013). Die Studienlage zum Thema ist z. T. sehr konträr, was an der Schwierigkeit der inter- bzw. transdisziplinären Geschlechterforschung liegt. Dies liegt daran, dass eine große Kluft besteht zwischen sozial- und kulturwissenschaftlich orientierter Theoriebildung einerseits und naturwissenschaftlich orientierter quantitativer Forschung andererseits. Auch unterschiedliche kognitive Strategien spielen bei der Sprachverarbeitung eine Rolle (Ferstl/Kaiser 2013). Auf eine detaillierte Erörterung des Forschungsstandes zum Thema Geschlecht und Sprache muss hier verzichtet werden. Abschließend sei darauf hingewiesen, dass die Geschlechtsunterschiede in der vorliegenden Untersuchung auf ein Zusammenspiel von biologischen, kulturellen, sozialen, kognitiven u. a. Prozessen zurückgeführt werden und jegliche Schlussfolgerungen nur vorsichtig formuliert werden müssen, da in jeder Untersuchung Störvariablen wirken können (z. B. sozio-kulturelles Umfeld, Metakognition usw.), die mit dem Geschlecht konfundiert sind.

4.1.3 Tonhöhe/Melodieführung

1. Die Rolle des Merkmals Tonhöhe/Melodieführung im Perzeptionsprozess der emotionalen Sprechweisen

Sprache/Kultur

Das Merkmal Tonhöhe/Melodieführung spielt bei den deutschsprachigen als auch bei den italienischsprachigen Proband*innen eine enorme Rolle, ohne dass hier ein sprachlich/kultureller Unterschied auszumachen wäre. 83,13 Prozent der deutschsprachigen und 82,86 Prozent der italienischsprachigen Personen geben an, dass dieses prosodische Merkmal im Rahmen der Emotionserkennung eine zentrale Rolle für sie spielt. Dies gilt bei den Sprechweisen zur Freude ebenso wie bei jenen zum Ärger. Das Ergebnis zur hohen Relevanz dieses Merkmals stimmt überein mit jenem aus einer Studie von Kranich (2003), der die herausragende Rolle der Tonhöhe/Melodieführung im Perzeptionsprozess postuliert hatte. Die Tonhöhe/Melodieführung stellt also in beiden Sprachen/Kulturen ein relevantes Indiz für die Perzeption natürlicher emotionaler Sprechweisen dar.

Beruf

87,59 Prozent der Laien und nur 75,56 Prozent der Stimmexpert*innen halten dieses Merkmal für zentral im Prozess der Emotionserkennung. Dieser relevante berufsbezogene Unterschied gilt unabhängig von der jeweiligen Sprache/Kultur der Personen und ist in Bezug auf die ärgerlichen Sprechweisen besonders ausgeprägt. Es wird angenommen, dass dieses Ergebnis die Tatsache widerspiegelt, dass die Tonhöhe/Melodieführung als wichtigstes aller Merkmale bei der Emotionsperzeption auch den Laien besonders zugänglich ist. So konnte etwa Künzel (1989) die Bedeutung dieses Parameters für den Perzeptionsprozess auch bei phonetisch ungeschulten Personen nachweisen.

Geschlecht

79,31 Prozent der Männer und 85,23 Prozent der Frauen bedienen sich insgesamt der Tonhöhe/Melodieführung als Parameter zur Emotionserkennung der Sprechäußerungen. Allerdings muss hier nach Sprach-/Kulturgruppe differenziert werden – während mehr deutschsprachige Frauen (88,66 Prozent) als Männer (75,36 Prozent) dieses Merkmal für relevant halten, sind es innerhalb der italienischsprachigen Gruppe deutlich mehr Männer (94,44 Prozent) als Frauen

(78,85 Prozent). Der Faktor Geschlecht spielt also eine wesentliche Rolle, allerdings v. a. innerhalb der jeweiligen Sprach-/Kulturgruppe. Die geschlechtsspezifischen Unterschiede werden hier und im Folgenden, wie im vorigen Abschnitt zu Akzent/Rhythmus dargestellt, auf ein komplexes Zusammenspiel von biologischen, kulturellen, sozialen, kognitiven u. a. Prozessen zurückgeführt. Das Geschlecht wird als ein relevantes Persönlichkeitsmerkmal der Proband*innen verstanden, das von vielen weiteren Faktoren beeinflusst ist und sich auch auf die Perzeption und Analyse gesprochener Sprache auswirkt.

2. Relevanz des Merkmals Tonhöhe/Melodieführung in der Perzeption der emotionalen Sprechweisen

Sprache/Kultur

Auch bei der Beurteilung der Relevanz des Merkmals Tonhöhe/Melodieführung gibt es keinen sprachlich/kulturellen Unterschied. 91,11 Prozent der deutschsprachigen und 92,97 Prozent der italienischsprachigen Proband*innen halten diesen Parameter für wichtig, dabei gibt es auch keine relevante Unterscheidung in den Ergebnissen zu den freudigen oder ärgerlichen Sprechweisen. Die hohe Relevanz dieses Merkmals wurde bereits diskutiert und auch in anderen Untersuchungen bestätigt.

Beruf

Bei der Bewertung der Relevanz zeigt sich ein ähnliches Bild wie bei der ersten Frage, die darauf abzielte zu erfassen, ob und inwieweit dieses Merkmal für die Proband*innen überhaupt eine Rolle spielt bei der Emotionserkennung. Auch hier halten die Laien (95,81 Prozent halten es für eher bis sehr wichtig) dieses Merkmal für wichtiger als die Stimmexpert*innen (nur 85,41 Prozent nehmen diese Bewertung vor), besonders groß ist der Unterschied wieder bei den ärgerlichen Sprechweisen. Diese scheinen in Bezug auf die Tonhöhe/Melodieführung eingängiger zu sein als die freudigen Äußerungen und scheinen daher auch den Laien intuitiv leichter zugänglich zu sein. Der Faktor Beruf spielt hier unabhängig von der Sprache/Kultur eine Rolle.

Geschlecht

Leichte geschlechterbedingte Differenzen lassen sich hier, v. a. innerhalb der deutschsprachigen Gruppe, verzeichnen. Insgesamt halten 89,21 Prozent der Männer und 93,46 Prozent der Frauen dieses Merkmal für relevant. Allerdings handelt es sich um deutlich mehr deutschspra-

chige Frauen (94,7 Prozent) als deutschsprachige Männer (86,02 Prozent), während in der italienischsprachigen Gruppe keine relevanten geschlechterbedingten Differenzen vorliegen. Auch die jeweilige Emotion spielt eine Rolle – im Rahmen der ärgerlichen Sprechweisen wird die Tonhöhe/Melodieführung von den Frauen (89,36 Prozent) für etwas wichtiger erachtet als von den Männern (84,13 Prozent). Der Faktor Geschlecht spielt hier also im Zusammenhang mit der jeweiligen Sprach-/Kulturgruppe eine Rolle. Es wird ein komplexes Zusammenspiel von geschlechtsspezifischen, kulturellen und anderen Faktoren angenommen, wobei Geschlecht als Persönlichkeitseigenschaft der jeweiligen Proband*innen verstanden wird, das im Wechselspiel und in gegenseitiger Beeinflussung mit anderen Persönlichkeitseigenschaften (in diesem Fall z. B. deutsche Sprache und Kultur) wirkt.

4.1.4 Text/Wortwahl

1. Die Rolle des Elements Text/Wortwahl im Perzeptionsprozess der emotionalen Sprechweisen

Sprache/Kultur

Wie bereits dargestellt, wird dieses Element zwar von 49,4 Prozent der deutschsprachigen und 40 Prozent der italienischsprachigen Hörer*innen zur Emotionserkennung verwendet, allerdings liegt es im Ranking in beiden Gruppen auf vorletzter Stelle. Die wahrscheinlichen Gründe dafür wurden bereits erörtert. Warum es zu einem sprachlich/kulturellen Gap kommt im Ergebnis, ist unklar. Die Differenz ist v. a. bei den freudigen Sprechweisen besonders ausgeprägt. Es wird vermutet, dass die deutschsprachigen Personen mehr den Text fokussieren und die italienischsprachigen hauptsächlich auf die Melodie und Tonhöhe des Gesagten achten, da dieser Aspekt im melodischen Italienischen besonders nahe liegt. Diese These wird auch gestützt durch die Tatsache, dass es kein prosodisches Element gibt, das die italienischsprachige Gruppe öfters als die deutschsprachige Gruppe zur Emotionserkennung herangezogen hat.

Beruf

42,07 Prozent der Laien und 54,44 Prozent der Stimmexpert*innen geben an, sich dieses Elements zur Emotionserkennung zu bedienen. Dieser Unterschied ist schon sehr erstaunlich, hätte man sich doch aufgrund der hypothetisierten besseren Wahrnehmungsfähigkeit der Prosodie von Seiten der Expert*innen eher ein umgekehrtes Ergebnis erwartet. Allerdings steht der Faktor Beruf hier in engem Zusammenhang mit jenem der Sprache/Kultur – die berufsbezogene

Differenz innerhalb der deutschsprachigen Gruppe fällt sehr gering aus (ca. 4 Prozent), beachtlich ist jedoch der Unterschied zwischen den italienischsprachigen Laien (25,64 Prozent) und den italienischsprachigen Expert*innen (58,06 Prozent). Es kann nur vermutet werden, dass dieses Ergebnis mit einer besseren Wahrnehmungsfähigkeit des sprachlichen Elements der Turns von Seiten der Expert*innen einhergeht, welche sich v. a. in der italienischsprachigen Gruppe niederschlägt, gerade weil dort Sprache/Text im Vergleich zur deutschsprachigen Gruppe insgesamt eine geringere Rolle spielt. Tatsächlich ist ein gewisses Maß an Perzeptionsfähigkeit vonnöten, die Sprache bzw. den Text der kurzen Turns, welche ja eher aufgrund ihrer prosodischen Charakteristika ausgewählt wurden, herauszuhören. Insofern ergibt es Sinn, dass die Expert*innen mehr mit dem Element Sprache/Text anzufangen wissen als die Laien, besonders in der italienischsprachigen Gruppe, in der dieser Parameter generell eine noch geringere Rolle spielt als in der deutschsprachigen.

Geschlecht

41,38 Prozent der Männer und 49,66 Prozent der Frauen verwenden das Element Text/Wortwahl zur Emotionserkennung, der geschlechterbezogene Unterschied zeigt sich v. a. bei den ärgerlichen Sprechweisen sehr deutlich (38,46 Prozent m vs. 52,31 Prozent w) und ist auch hier wieder unter dem Aspekt der Sprache/Kultur zu betrachten. Zwar zeigt sich, sei es innerhalb der deutschsprachigen als auch innerhalb der italienischsprachigen Gruppe, die Tendenz, dass für die Frauen dieses Element eine größere Rolle spielt als für die Männer, so gilt dies doch in besonderem Maße für die italienischsprachigen Proband*innen (22,22 Prozent der it. Männer, aber 46,15 Prozent der it. Frauen bedienen sich des Merkmals Text/Wortwahl zur Emotionserkennung). Warum eine geschlechterbedingte Differenz v. a. in der italienischsprachigen Gruppe vorliegt, muss hier offen bleiben. Es wird angenommen, dass das Persönlichkeitsmerkmal Geschlecht mit anderen Persönlichkeitsmerkmalen, etwa zu Sprache/Kultur, auf eine spezifische Weise interferiert. Hierzu müssten weitere Studien erfolgen.

Für die zweite Frage zur Relevanz der einzelnen Parameter zur Bestimmung der emotionalen Sprechweisen wurde das Element Text/Wortwahl nicht miteinbezogen, da dies für die Beantwortung der Forschungsfragen nicht relevant ist und der Fokus auf die prosodischen Merkmale gelegt werden sollte.

4.1.5 Lautstärke

1. Die Rolle des Merkmals Lautstärke im Perzeptionsprozess der emotionalen Sprechweisen

Sprache/Kultur

Der sprachlich/kulturelle Unterschied ist in Bezug auf das Merkmal Lautstärke besonders stark ausgeprägt – in keinem der anderen prosodischen Parameter fällt die Differenz so prägnant aus. 65,66 Prozent der deutschsprachigen und 44,29 Prozent der italienischsprachigen Proband*innen bedienen sich der Lautstärke zur Emotionserkennung. Besonders im Rahmen der ärgerlichen Sprechweisen ist der Unterschied enorm (78,08 Prozent dt. vs. 41,94 Prozent it.). Diese große Differenz wird auf ein alltagssprachliches Phänomen zurückgeführt. So wird im italienischsprachigen Raum vermutlich tendenziell lauter gesprochen als im deutschsprachigen, weshalb deutschsprachige Hörer*innen vermutlich sensibler auf Lautstärkeschwankungen, v. a. emotional bedingte, reagieren. Gerade ärgerliche Sprechweisen werden also von den deutschsprachigen Hörer*innen wahrscheinlich viel stärker aufgrund deren Lautheit erkannt als von italienischsprachigen Proband*innen, welche an ein generell höheres Lautstärkepegel in diversen emotionalen Situationen gewöhnt sind. Die Perzeption der Lautstärke ist stark vom jeweiligen Hörmuster geprägt und damit situations- und kulturabhängig (Kranich 2016).

Beruf

58,62 Prozent der Laien und 60 Prozent der Stimmexpert*innen verwenden das Merkmal Lautstärke zur Emotionserkennung, weitestgehend unabhängig von der jeweiligen Sprache/Kultur der Proband*innen. Ein Unterschied lässt sich nur in Hinblick auf die Unterscheidung zwischen freudigen und ärgerlichen Sprechweisen verzeichnen – während die Lautstärke im Rahmen der Perzeption der freudigen Äußerungen eine größere Rolle spielt für die Expert*innen (plus 7 Prozent im Vergleich zu den Laien), ist sie bei der Perzeption der ärgerlichen Sprechweisen relevanter für die Laien (plus 6 Prozent). Dies wird damit in Verbindung gebracht, dass die Lautstärke im Rahmen der ärgerlichen Äußerungen stärker ins Gewicht fällt als bei den freudigen, und damit gerade für die Laien einfacher zugänglich ist für die Emotionserkennung.

Geschlecht

Das Geschlecht spielt bei der Beantwortung der ersten Frage eine Rolle, aber nur in Verbindung mit dem Faktor Sprache/Kultur. Insgesamt nehmen 64,37 Prozent der Männer und 56,38 Prozent der Frauen die emotionalen Sprechweisen anhand des Kriteriums Lautstärke wahr. Die geschlechterbedingte Differenz gilt aber ausschließlich innerhalb der deutschsprachigen Gruppe. Auf die unterschiedlichen Hörmuster der beiden Sprach-/Kulturgruppen in Bezug auf die Lautstärke wurde bereits hingewiesen. Dieses Element könnte zusammen mit dem Persönlichkeitsmerkmal Geschlecht eine Rolle spielen und als Ursache dieser Differenz verstanden werden. Generell gibt es wenige Studien zur Perzeption von Lautstärke (Kranich 2016), hier wären v. a. in Verbindung mit den diversen hier genannten Einflussfaktoren weitere Untersuchungen vonnöten.

2. Relevanz des Merkmals Lautstärke in der Perzeption der emotionalen Sprechweisen

Sprache/Kultur

Gemäß der Antwort auf die erste Frage ist auch hier die Lautstärke für die deutschsprachigen Proband*innen (80 Prozent geben an, das Merkmal für eher bis sehr wichtig zu halten) etwas relevanter als für die italienischsprachigen Personen (71,1 Prozent der Personen halten es für eher bis sehr wichtig). Während bei der Beurteilung der freudigen Sprechweisen beide Sprach-/Kulturgruppen (je ca. 70 Prozent) die Lautstärke für gleich wichtig erachten, schätzen die deutschsprachigen Proband*innen (92,08 Prozent) die Relevanz dieses prosodischen Merkmals für deutlich größer ein als die italienischsprachigen Proband*innen (71,43 Prozent). Die Erklärung hierfür wird, wie bei Frage 1, in Bezug auf die unterschiedlichen, kulturell geprägten Hörmuster zur Lautstärke gefunden.

Beruf

Tendenziell wird die Lautstärke von den Laien als etwas relevanter eingeschätzt als von den Stimmexpert*innen, der Unterschied ist insgesamt aber gering. So geben 78,14 Prozent der Laien und 74,46 Prozent der Expert*innen an, dieses Merkmal für eher bis sehr wichtig zu erachten. Bei näherer Betrachtung in Bezug auf den Faktor Sprache/Kultur zeigte sich, dass es einen relevanten berufsbezogenen Unterschied nur innerhalb der italienischsprachigen Gruppe gibt (75,79 Prozent der it. Laien vs. 57,58 Prozent der it. Expert*innen halten die Lautstärke für relevant). Auch die jeweilige Emotion spielt eine Rolle, die Differenz ist im Rahmen der ärgerlichen Sprechweisen größer als bei den freudigen. Es wird vermutet, dass die Unterschiede

auf die Tatsache zurückzuführen sind, dass die Lautheitseigenschaften v. a. im Rahmen der ärgerlichen Sprechweisen gerade auch für phonetisch ungeübte Hörer*innen leichter zu perzipieren und damit für gewichtiger erachtet werden. Dass es innerhalb der deutschsprachigen Gruppe keine berufsbezogenen Differenzen gibt, wird hingegen eher auf die Prävalenz kultureller Hörmuster zurückgeführt, welche sich gerade in Bezug auf die Lautstärke sehr zwischen den beiden Sprach-/Kulturgruppen unterscheiden.

Geschlecht

77,69 Prozent der Männer und 76,16 Prozent der Frauen halten das Merkmal Lautstärke insgesamt für eher bis sehr wichtig im Perzeptionsprozess der emotionalen Sprechweisen. Dabei ist die Lautstärke im Rahmen der ärgerlichen Äußerungen für die Männer relevanter als für die Frauen, bei den freudigen Äußerungen ist es, in geringerem Ausmaß, tendenziell umgekehrt. Allerdings muss auch hier der Faktor Sprache/Kultur miteinbezogen werden, um zu einem gültigen Ergebnis zu kommen – tatsächlich lassen sich v. a. innerhalb der italienischsprachigen Gruppe relevante geschlechtsbezogene Differenzen feststellen, wobei die Lautstärke für die italienischsprachigen Männer eine größere Rolle spielt als für die italienischsprachigen Frauen. Das Zusammenwirken der Persönlichkeitsmerkmale Geschlecht und Sprache/Kultur müsste in weiteren Studien näher differenziert werden, um hierzu zu aussagekräftigen Erklärungen zu gelangen.

4.1.6 Sprechgeschwindigkeit

1. Die Rolle des Merkmals Lautstärke im Perzeptionsprozess der emotionalen Sprechweisen

In der ersten Frage im Fragebogen ging es darum herauszufinden, ob und inwieweit die aufgelisteten Parameter überhaupt zur Emotionserkennung der Äußerungen herangezogen werden. Die Sprechgeschwindigkeit befindet sich im Ranking an fünfter Stelle.

Sprache/Kultur

Es gibt einen klar sprachlich/kulturellen Unterschied, da 51,81 Prozent der deutschsprachigen, aber nur 42,86 der italienischsprachigen Proband*innen angaben, sich bei der Emotionsperzeption dieses Merkmals zu bedienen. Die Unterschiede sind in den Angaben zu den freudigen als

auch in jenen zu den ärgerlichen Sprechweisen ersichtlich. Es wird vermutet, dass diese Differenz auf die unterschiedlichen, kulturell geprägten Hörmuster der Proband*innen zurückzuführen ist. Es gilt als Stereotyp, dass die italienische Sprechweise auf deutschsprachige Personen rascher wirkt als die eigene. Möglicherweise spielt dies in das Ergebnis mit hinein. Auf jeden Fall sind weitere Untersuchungen hierzu notwendig.

Beruf

51,03 Prozent der Laien und 46,67 Prozent der Stimmexpert*innen erkennen die Emotion anhand der Sprechgeschwindigkeit, diese Differenz findet sich quasi unabhängig von der Ausprägung der Emotion in Freude oder Ärger. Der Faktor Beruf steht aber in engem Zusammenhang zur Sprache/Kultur der Proband*innen, so sind relevante berufsbezogene Differenzen nur in der deutschsprachigen Gruppe ersichtlich. 54,72 Prozent der deutschsprachigen Laien, aber nur 47,46 Prozent der deutschsprachigen Expert*innen geben an, sich der Sprechgeschwindigkeit zur Emotionserkennung zu bedienen. Möglicherweise ist der berufsbedingte Unterschied darauf zurückzuführen, dass die in beiden emotionalen Ausprägungen eher raschen Turns intuitiv recht zugängliche Größen auch für phonetisch ungeschulte Hörer*innen sind.

Geschlecht

Auch das Geschlecht spielt eine Rolle, so nehmen 54,02 Prozent der Männer und 46,31 Prozent der Frauen die emotionalen Sprechweisen anhand der Sprechgeschwindigkeit wahr. Während es bei der Perzeption zu den freudigen Äußerungen keine Unterschiede in den Angaben zwischen den Geschlechtern gibt, so ist die Differenz im Rahmen der ärgerlichen Sprechweisen besonders groß (71,79 Prozent m vs. 55,38 Prozent w). Der Faktor Geschlecht interagiert jedoch auf spezifische Art und Weise auch mit jenem der Sprache/Kultur. Während die Sprechgeschwindigkeit innerhalb der deutschsprachigen Gruppe für die Männer eine größere Rolle spielt als für die Frauen, ist diese Tendenz innerhalb der italienischsprachigen Gruppe umgekehrt. Wie sich die beiden Faktoren Geschlecht und Sprache/Kultur bei der Emotionsperzeption durch Prosodie gegenseitig beeinflussen, müsste in weiteren Studien näher untersucht werden.

2. Relevanz des Merkmals Sprechgeschwindigkeit in der Perzeption der emotionalen Sprechweisen

Sprache/Kultur

Im Unterschied zum Ergebnis der weiter oben ausgeführten Frage 1, zeigt das Ergebnis zu dieser Frage, dass die italienischsprachigen Personen die Sprechgeschwindigkeit für relevanter halten zur Bestimmung der Sprechweisen als die deutschsprachigen Proband*innen. 60 Prozent der deutschsprachigen und 70,31 Prozent der italienischsprachigen Proband*innen halten dieses Merkmal für eher bis sehr wichtig. Diese Differenz zeigt sich sei es bei den freudigen als auch bei den ärgerlichen Sprechweisen. Möglicherweise halten die italienischsprachigen Personen dieses Merkmal für relevanter, da sie ihm in ihrem Sprachgebrauch eine größere distinktive Bedeutung zumessen. Diese Hypothese müsste jedoch durch weitere Studien verifiziert werden.

Beruf

In Bezug auf den Beruf lässt sich bei der Beurteilung der Relevanz des Merkmals Sprechgeschwindigkeit keine ausschlaggebende Differenz feststellen. 64,65 Prozent der Laien und 62,05 Prozent der Stimmexpert*innen halten dieses Merkmal für relevant. Im Rahmen der ärgerlichen Sprechweisen spielt die Sprechgeschwindigkeit für die Laien eine größere Rolle als für die Expert*innen. Dies könnte daran liegen, dass die ärgerlichen Turns aufgrund deren als rascher perzipierten Geschwindigkeit v. a. auch den Laien schneller „ins Ohr stechen“. Es lassen sich jedenfalls unabhängig der jeweiligen Sprache/Kultur keine relevanten Unterschiede zwischen der Einschätzung des Merkmals der Laien und jener der Expert*innen feststellen.

Geschlecht

Bei der Einschätzung der Relevanz des Merkmals Sprechgeschwindigkeit liegt ein relevanter geschlechtsspezifischer Unterschied innerhalb der jeweiligen Sprach-/Kulturgruppe vor. 64,75 Prozent der Männer und 63,08 Prozent der Frauen insgesamt halten dieses Merkmal für wichtig, unabhängig davon, ob es sich um freudige oder ärgerliche Sprechweisen handelt. Allerdings sind es innerhalb der deutschsprachigen Gruppe die Männer (67,74 Prozent halten dieses Merkmal für eher bis sehr wichtig), welche die Sprechgeschwindigkeit für deutlich relevanter einstufen als die Frauen (54,55 Prozent), und innerhalb der italienischsprachigen Gruppe hingegen die Frauen (76,83 Prozent), welche dieses Merkmal für sehr viel wichtiger erachten als die

Männer (58,7 Prozent). Innerhalb der italienischsprachigen Gruppe ist der geschlechterbezogene Unterschied (knapp 20 Prozent) also deutlich größer als innerhalb der deutschsprachigen Gruppe (ca. 13 Prozent). Es wird ein komplexes Zusammenspiel diverser, v. a. geschlechtsspezifischer und kultureller Faktoren angenommen, das jedoch weiterer Untersuchungen bedarf.

4.1.7 Pausen

1. Die Rolle des Merkmals Pause im Perzeptionsprozess der emotionalen Sprechweisen

Sprache/Kultur

Ogleich die Pausen von beiden Sprach-/Kulturgruppen verhältnismäßig wenig genannt werden, gibt es doch einen deutlichen sprachlich/kulturellen Unterschied. 27,14 Prozent der italienischsprachigen Proband*innen geben an, dieses Merkmal zur Emotionserkennung heranzuziehen, aber nur 17,47 Prozent der deutschsprachigen Hörer*innen. Dieser Gap zeigt sich bei den freudigen und ärgerlichen Sprechweisen gleichermaßen. Es wird davon ausgegangen, dass die Differenz auf kulturell geprägte Hörmuster der Proband*innen zurückzuführen ist. Inwieweit sich die jeweilige Muttersprache in Bezug auf das Merkmal Pause im Perzeptionsprozess der emotionalen Sprechweisen genau auswirkt, müsste näher untersucht werden.

Beruf

Berufsbezogene Differenzen lassen sich in Bezug auf die jeweilige Emotion der Äußerungen und in Bezug auf die Sprache/Kultur der Proband*innen feststellen. Insgesamt bedienen sich 20,69 Prozent der Laien und 20 Prozent der Stimmexpert*innen der Pausen für die Emotionserkennung. Im Rahmen der freudigen Sprechweisen spielt dieses prosodische Merkmal eine wesentlich größere Rolle für die Laien (21,69 Prozent) als für die Expert*innen (14,29 Prozent), im Rahmen der ärgerlichen Sprechweisen verhält es sich hingegen umgekehrt (19,05 Prozent Laien vs. 26,83 Prozent Expert*innen). Außerdem macht sich der berufsbezogene Unterschied v. a. innerhalb der italienischsprachigen Gruppe bemerkbar, in welcher viel mehr Laien (33,33 Prozent) als Expert*innen (19,35 Prozent) die Pausen zur Emotionsbestimmung verwenden. Ein äußerst komplexes Zusammenspiel aus sprachlich/kulturellen und berufsbezogenen Faktoren scheint hier also eine Rolle zu spielen.

Geschlecht

24,14 Prozent der Männer und 18,12 Prozent der Frauen geben an, anhand der Pause die emotionalen Sprechweisen zu perzipieren. Insbesondere im Rahmen der freudigen Äußerungen gilt, dass dieses Merkmal für die Männer wichtiger ist als für die Frauen. Der Faktor Geschlecht steht aber nicht nur in Zusammenhang mit der jeweiligen Emotion, sondern auch mit der Sprache/Kultur der Proband*innen, wobei es in beiden Gruppen zu geschlechterbedingten Differenzen kommt. Innerhalb der deutschsprachigen Gruppe verwenden die Männer (24,64 Prozent) die Pausen häufiger als die Frauen (12,37 Prozent), innerhalb der italienischsprachigen Gruppe verhält es sich genau umgekehrt (22,22 Prozent Männer vs. 28,85 Prozent Frauen). Auch hier gilt wieder die Annahme, dass diese Differenzen durch das komplexe Zusammenspiel mehrerer Faktoren bedingt sind, die in weiteren Untersuchungen näher ausgeführt werden müssten.

2. Relevanz des Merkmals Pause in der Perzeption der emotionalen Sprechweisen

Sprache/Kultur

Bei der Beurteilung der Relevanz der Pausen lässt sich ein enormer sprachlich/kultureller Unterschied feststellen. Während 51,57 Prozent der italienischsprachigen Proband*innen dieses Merkmal für eher bis sehr wichtig halten, sind es nur 27,11 Prozent der deutschsprachigen Personen. Diese Differenz ist bei beiden emotionalen Ausprägungen vorhanden, bei den freudigen Sprechweisen aber besonders ausgeprägt. Durch welche sprachlichen und kulturellen Faktoren die diversen Hörmuster zu Stande kommen, und inwieweit diese den großen sprachlich/kulturellen Unterschied in der Beurteilung der Relevanz der Pause erklären können, müsste in weiteren Studien näher untersucht werden.

Beruf

36,28 Prozent der Laien und 35,77 Prozent der Stimmexpert*innen halten dieses Merkmal für relevant. Dabei gibt es keinen Unterschied in Bezug auf die freudigen oder ärgerlichen Sprechweisen. Allerdings spielt auch der Faktor Sprache/Kultur eine relevante Rolle. So schätzen die deutschsprachigen Laien die Pausen für weniger wichtig ein als die deutschsprachigen Stimmexpert*innen. Innerhalb der italienischsprachigen Gruppe gibt es hingegen keinen berufsbezogenen Unterschied. Dieses Resultat könnte damit einhergehen, dass die Stimmexpert*innen mehr Pausen und auch mehr geräuschvolle Pausen wahrnehmen als die Laien (siehe Kapitel 4.2.5) und sie deshalb auch für relevanter halten im Rahmen der Emotionsperzeption.

Geschlecht

Deutlich hervor tritt bei der Frage nach der Beurteilung der Relevanz der Pausen im Perzeptionsprozess ein geschlechtsbedingter Unterschied, wobei dieses Merkmal von den Frauen als relevanter eingestuft wird als von den Männern. 30,21 Prozent der Männer und 39,72 Prozent der Frauen halten dieses Merkmal für eher bis sehr wichtig für die Perzeption der emotionalen Sprechweisen, unabhängig davon, ob es sich um freudige oder ärgerliche Äußerungen handelt. Dieser geschlechtsspezifische Unterschied zeigt sich in etwa gleicher Ausprägung in beiden Sprach-/Kulturgruppen. Interessant ist, dass dieses Ergebnis nicht mit einer besseren oder anderen Wahrnehmung und Analyse der Pausen (siehe Kapitel 4.2.5) von Seiten der Frauen einhergeht. Von welchen Faktoren die geschlechtsbedingte Differenz im Speziellen abhängt, muss an dieser Stelle offen bleiben.

4.1.8 Stimmfärbung (Timbre)

1. Die Rolle des Merkmals Stimmfärbung (Timbre) im Perzeptionsprozess der emotionalen Sprechweisen

Sprache/Kultur

48,8 Prozent der deutschsprachigen und 51,43 Prozent der italienischsprachigen Proband*innen bedienen sich des Merkmals Stimmfärbung (Timbre), um die Emotion der Äußerungen zu erkennen. Insgesamt zeigt sich hierbei also kein relevanter sprachlich/kultureller Unterschied. In Bezug auf die ärgerlichen Sprechweisen liegt jedoch eine Differenz vor, da 58,06 Prozent der italienischsprachigen und nur 52,05 Prozent der deutschsprachigen Personen die Stimmfärbung als Indiz für die entsprechende Emotion nutzen. Das Merkmal Stimmfärbung ist ein sehr komplexes und umfasst verschiedene Eigenschaften und Aufgaben, weshalb es schwierig ist festzumachen, welche sprachlich/kulturellen Aspekte der Stimmfärbung hier in Bezug auf die sprachlich/kulturelle Differenz zum Tragen kommen.

Beruf

Auch ein berufsbedingter Unterschied innerhalb der italienischsprachigen Gruppe liegt vor. Insgesamt verwenden 51,72 Prozent der Laien und 45,56 Prozent der Stimmexpert*innen das Merkmal Stimmfärbung zur Emotionserkennung. In beiden emotionalen Ausprägungen, aber v. a. bei den ärgerlichen Sprechweisen, spielt das Timbre für die Laien eine relevantere Rolle als für die Expert*innen. Dies könnte sich dadurch erklären, dass genau bei den ärgerlichen

Äußerungen deutliche Missklänge perzipiert wurden, welche auch den ungeschulten Hörer*innen schnell ins Ohr springen. Allerdings gelten die berufsbedingten Differenzen nur für die italienischsprachige Gruppe, innerhalb derer sich 61,54 Prozent der Laien, aber nur 38,71 Prozent der Expert*innen des Merkmals Stimmfärbung zur Emotionsbestimmung bedienen. Dieses Ergebnis ist überraschend und könnte darauf zurückgeführt werden, dass die Timbreigenschaften in ihrer starken emotional bedingten Ausprägung besonders die italienischsprachigen Laien beeindrucken, welche an einen sonst eher melodischen, singenden Klang ihrer Sprache gewöhnt sind.

Geschlecht

Kein anderes prosodisches Merkmal wirft eine solch große geschlechterbedingte Differenz auf wie die Stimmfärbung (Timbre). Das Ergebnis, dass dieses Merkmal für die Männer (63,22 Prozent geben an, sich dieses Parameters zur Emotionserkennung zu bedienen) ungleich wichtiger ist als für die Frauen (41,61 Prozent) ist überraschend und kann mit dem vorliegenden Wissensstand nicht beantwortet werden. Zwar gilt diese Differenz in beiden emotionalen Sprechweisen, allerdings ist sie bei den freudigen Äußerungen um ein Vielfaches gesteigert. Auch der Faktor Sprache/Kultur spielt eine Rolle – ist die Tatsache, dass das Timbre von den Männern wesentlich häufiger zur Emotionserkennung herangezogen wird als von den Frauen, in beiden Sprach-/Kulturgruppen gegeben, so gilt sie doch in besonderem Maße für die italienischsprachigen Proband*innen. 88,89 Prozent der italienischsprachigen Männer, aber nur 38,46 Prozent der italienischsprachigen Frauen bedienen sich dieses Merkmals für die Emotionserkennung. Über die Gründe für diese geschlechterspezifische Differenz könnte nur spekuliert werden, fest steht, dass auch hier wohl eine Kombination an geschlechterbezogenen, sprachlich/kulturellen u. a. Faktoren zusammenwirkt.

2. Relevanz des Merkmals Stimmfärbung (Timbre) in der Perzeption der emotionalen Sprechweisen

In der zweiten Frage im Fragebogen sollte die Relevanz der einzelnen prosodischen Merkmale für die Perzeption der jeweiligen emotionalen Äußerungen angegeben werden. Hier werden die Antworten in Bezug auf die Einflussfaktoren Sprache/Kultur, Beruf und Geschlecht dargestellt.

Sprache/Kultur

69,33 Prozent der deutschsprachigen und 57,03 Prozent der italienischsprachigen Proband*innen halten dieses Merkmal für relevant im Perzeptionsprozess der emotionalen Sprechweisen. Der sprachlich/kulturelle Unterschied gilt für beide emotionale Ausprägungen, ist aber bei den freudigen Sprechweisen besonders groß (70,16 Prozent der deutschsprachigen, aber nur 46,15 Prozent der italienischsprachigen Personen halten die Stimmfärbung für wichtig). Welche sprachlich/kulturellen Faktoren für diese Differenz genau verantwortlich sind, müsste in weiteren Studien eruiert werden.

Beruf

Der Beruf spielt in Kombination mit der Sprache/Kultur eine große Rolle in der Bewertung der Relevanz dieses prosodischen Merkmals. Insgesamt geben 64,18 Prozent der Laien und 65,7 Prozent der Stimmexpert*innen an, die Stimmfärbung für eher bis sehr wichtig zu halten. Die Bewertung im Rahmen der beiden Emotionen Freude und Ärger fällt dabei annähernd gleich aus. Eine wichtige Rolle spielt hingegen der Faktor Sprache/Kultur – während in der deutschsprachigen Gruppe die Stimmexpert*innen (72,12 Prozent halten das Merkmal für eher bis sehr wichtig) das Timbre für etwas wichtiger einschätzen als die Laien (66,67 Prozent), verhält es sich in der italienischsprachigen Gruppe in größerem Ausmaß genau umgekehrt. Hier geben 61,05 Prozent der Laien an, das Merkmal für wichtig zu halten, aber nur 45,45 Prozent der Expert*innen. Dieses Ergebnis stimmt mit jenem zu Frage eins überein, nach dem sich in der italienischsprachigen Gruppe mehr Laien als Expert*innen dieses Merkmals zur Emotionserkennung bedienen. Dieses differenzierte Ergebnis zur Einschätzung der Relevanz des Timbres geht wohl auf eine Vielzahl an Ursachen zurück, zu denen zum einen die Komplexität des prosodischen Merkmals an sich und zum anderen das Zusammenwirken der Faktoren Geschlecht und Sprache/Kultur mit möglicherweise weiteren Elementen gezählt werden.

Geschlecht

Wie schon bei Frage eins zur Rolle des Merkmals Stimmfärbung im Perzeptionsprozess wird auch dessen Relevanz deutlich unterschiedlich bewertet von Männern und von Frauen, wobei erstere das Timbre für wesentlich relevanter befinden. 72,67 Prozent der Männer geben an, dieses Merkmal für eher bis sehr wichtig zu halten, dieselbe Bewertung wird aber nur von 59,81 Prozent der Frauen vorgenommen. Diese geschlechterbezogene Differenz wirkt unabhängig von der jeweiligen zugrundeliegenden Emotion (Freude bzw. Ärger) und in beiden Sprach-

/Kulturkreisen, aber v. a. innerhalb der italienischsprachigen Gruppe. Dieses Ergebnis entspricht jenem zu Frage eins und wird auf eine komplexe Wechselwirkung diverser geschlechterspezifischer und sprachlich/kultureller Faktoren zurückgeführt.

4.1.9 Fazit

Es konnte aufgezeigt werden, welche Parameter zur Emotionserkennung generell und in Bezug auf die Sprechweisen zu Freude bzw. Ärger inwieweit eine Rolle spielen und für wie relevant die Proband*innen die einzelnen prosodischen Merkmale für den Perzeptionsprozess halten. Dabei wurden sprachlich/kulturelle, geschlechtsspezifische und berufsbezogene Differenzen herausgestellt und thematisiert. Hierzu kann insgesamt festgestellt werden, dass sich keine generalisierten Aussagen zu den Faktoren Sprache/Kultur, Beruf und Geschlecht hinsichtlich ihrer Einwirkung auf den Perzeptionsprozess der emotionalen Sprechweisen treffen lassen. Vielmehr spielen diese Elemente zwar eine zentrale Rolle bei der Perzeption und es ist zweifelsohne wichtig und aufschlussreich, sie bei der Auswertung zu berücksichtigen und miteinzubeziehen, aber sie verhalten sich in einem komplexen Wechselspiel zueinander, d. h. die Faktoren sind nicht einzeln jeder für sich zu betrachten, sondern immer im Gesamtzusammenhang mit den anderen Aspekten und in Bezug auf die einzelnen prosodischen Merkmale. Erst in der Kombination aller Elemente lassen sich interessante und valide Aussagen über den Einfluss der Sprache/Kultur, des Berufes und Geschlechts generieren, wobei jegliche Verallgemeinerungen und allgemeinen Rückschlüsse jedoch nur mit äußerster Vorsicht zu treffen sind. Zudem bedarf es noch einer Vielzahl an weiteren Untersuchungen, um das spezifische Zusammenwirken der einzelnen Faktoren und dessen Ursachen adäquat zu beschreiben. Hier könnten und müssten künftige Studien ansetzen, um eine genügende Erklärung für die sprachlich/kulturell bedingten, geschlechtsspezifischen oder berufsbezogenen Divergenzen liefern zu können. Selbstverständlich können auch weitere Persönlichkeitsmerkmale (z. B. das Alter) den Perzeptionsprozess beeinflussen. Dies gilt es bei der Durchführung und Auswertung der Ergebnisse stets zu berücksichtigen.

Des Weiteren fällt bei der Auswertung der zweiten Frage im Fragebogen auf, dass die Kategorie „nicht wichtig“ am wenigsten verwendet wurde, die Kategorien „eher wichtig“ am häufigsten und „sehr wichtig“ am zweithäufigsten. Auch dadurch ist einmal mehr ersichtlich, dass beim Erkennen der Emotion und bei der Perzeption des Gehörten nicht einzelne Parameter verwendet werden, sondern alle in unterschiedlichem Ausmaß. Es ist also das Zusammenwirken aller prosodischen Merkmale von Bedeutung.

In Bezug auf die Faktoren Sprache/Kultur, Beruf und Geschlecht kann zusammenfassend festgestellt werden, dass sie sich in relevantem Ausmaß und in diverser Kombination auf die Perzeption der emotionalen Sprechweisen durch Prosodie auswirken. Sie wirken sich in Bezug auf die einzelnen prosodischen Merkmale und auf die jeweilige Emotion (Freude bzw. Ärger) je unterschiedlich aus und treten auch in Kombination miteinander (z. B. berufsbezogener Unterschied nur in einer Sprach-/Kulturgruppe) auf. Daher kann ein komplexes Wechselspiel dieser Faktoren festgestellt werden, welches durch weitere Untersuchungen noch näher differenziert werden müsste. Des Weiteren hat sich herausgestellt, dass zwar insgesamt ein klares Ranking in der Bewertung der Relevanz der einzelnen prosodischen Parameter für den Perzeptionsprozess vorliegt, dieses aber innerhalb der Vergleichsgruppen (dt.-it.; m-w; Laie- Expert*in) durchaus leicht variieren kann. Es ist erstaunlich, dass die Proband*innen aufgrund ihrer sprachlich/kulturellen, beruflichen und geschlechtsspezifischen Differenzen zwar zum selben Ergebnis kommen (die Äußerungen werden als freudig bzw. ärgerlich perzipiert – dies wurde schon in der Vorstudie sichergestellt, da es ansonsten auch hier zu Unterschieden kommen könnte), dies aber auf unterschiedlichem Wege tun (so verwenden etwa deutlich mehr italienischsprachige als deutschsprachige Personen die Pausen zur Emotionserkennung, dafür spielt die Lautstärke für die deutschsprachigen Proband*innen eine größere Rolle).

4.2 Forschungsfrage 2

2. Wie werden die einzelnen prosodischen Merkmale der jeweiligen freudigen bzw. ärgerlichen Sprechweisen von den Proband*innen beschrieben?

Um diese Forschungsfrage zu beantworten, wurden im dritten Teil des Fragebogens die prosodischen Merkmale von den Proband*innen analysiert. Dazu wurden die Merkmale in ihre charakteristischsten und relevantesten Eigenschaften untergliedert und beschrieben. Im Folgenden werden die Ergebnisse der Übersichtlichkeit wegen gebündelt stichpunktartig festgehalten. In einem zweiten Schritt werden die Ergebnisse hinsichtlich der Faktoren Sprache/Kultur, Beruf und Geschlecht untersucht. Vorab gilt festzuhalten, dass die Ergebnisse zur Analyse der ärgerlichen Sprechweisen klarer sind als jene der freudigen Sprechweisen, hier haben auch mehr Personen die Kategorie „weiß ich nicht“ angegeben. Dies könnte daran liegen, dass im Rahmen der ärgerlichen Sprechweisen bestimmte prosodische Mittel stärker eingesetzt werden (z. B. Lautstärke) als bei den freudigen Äußerungen, außerdem wurden die ärgerlichen Sprechweisen

nach den freudigen analysiert, weshalb möglicherweise schon ein gewisser Lerneffekt in der Beurteilung der prosodischen Parameter aufgetreten ist.

Die prosodischen Merkmale der freudigen bzw. ärgerlichen Sprechweisen wurden von den Proband*innen wie folgt beschrieben:

Akzent/Rhythmus

1) Freude

- Akzentuierung (Stärke): eher starke Betonung (59,59 Prozent)
- Akzentuierung (Häufigkeit): eher viele Betonungen (46,11 Prozent)
- Sprechrhythmus: eher fließend (legato) (45,6 Prozent)

2) Ärger

- Akzentuierung (Stärke): eher starke Betonung (78,85 Prozent)
- Akzentuierung (Häufigkeit): eher viele Betonungen (67,95 Prozent)
- Sprechrhythmus: eher abgehackt (staccato) (41,03 Prozent)

Im Rahmen der ärgerlichen Sprechweisen werden mehr und stärkere Betonungen perzipiert als bei den freudigen Äußerungen, außerdem wird der Sprechrhythmus einmal als fließend (Freude) und einmal als abgehackt (Ärger) beschrieben.

Tonhöhe/Melodieführung

1) Freude

- Sprechstimmlage: eher hoch (49,7 Prozent)
- Tonhöhenverlauf: bewegt (lebhaft Melodie) (49,7 Prozent)
- Tonhöhenumfang: mittel (39,4 Prozent)

2) Ärger

- Sprechstimmlage: eher hoch (53,2 Prozent)
- Tonhöhenverlauf: bewegt (lebhaft Melodie) (60,9 Prozent)
- Tonhöhenumfang: mittel (46,8 Prozent)

Der Tonhöhenverlauf der ärgerlichen Sprechweisen wird als deutlich bewegter wahrgenommen als jener der freudigen Sprechweisen. In Bezug auf die anderen Eigenschaften sind die Beurteilungen vergleichbar.

Lautstärke

1) Freude

- Lautstärke: eher laut (56 Prozent)
- Lautheitsverlauf: gleichförmig (53,4 Prozent)
- Lautheitsverlauf am Ende der Phrase: gleichbleibende Lautstärke (54,9 Prozent)

2) Ärger

- Lautstärke: eher laut (80,8 Prozent)
- Lautheitsverlauf: gleichförmig (44,2 Prozent)
- Lautheitsverlauf am Ende der Phrase: gleichbleibende Lautstärke (60,3 Prozent)

Auch wenn die Angaben zu den Eigenschaften der Lautstärke bei Freude und Ärger dieselben sind, werden die ärgerlichen Sprechweisen als deutlich lauter, etwas bewegter und am Ende der Phrase als lauter perzipiert als die freudigen Sprechweisen.

Sprechgeschwindigkeit

1) Freude

- Geschwindigkeit: mittel (41,5 Prozent) bis eher schnell (38,3 Prozent)
- Geschwindigkeitsverlauf: gleichförmig (36,8 Prozent) bis mittel (33,7 Prozent)
- Geschwindigkeitsverlauf am Ende der Phrase: gleichbleibende Geschwindigkeit (57 Prozent)

2) Ärger

- Geschwindigkeit: eher schnell (65,2 Prozent)
- Geschwindigkeitsverlauf: gleichförmig (51 Prozent)
- Geschwindigkeitsverlauf am Ende der Phrase: gleichbleibende Geschwindigkeit (57,4 Prozent)

Die ärgerlichen Sprechweisen werden im Vergleich zu den freudigen als deutlich schneller und den Geschwindigkeitsverlauf betreffend als gleichförmiger und am Ende der Phrase als tendenziell rascher wahrgenommen.

Pausen

1) Freude

- Häufigkeit: wenige Pausen (51,3 Prozent)
- Geräuschvolle Pausen (mit Klang oder Geräuschen gefüllt): eher wenige Pausen (57 Prozent)

2) Ärger

- Häufigkeit: wenige Pausen (67,7 Prozent)
- Geräuschvolle Pausen (mit Klang oder Geräuschen gefüllt): eher wenige Pausen (74,2 Prozent)

Bei den ärgerlichen Sprechweisen werden insgesamt weniger Pausen und weniger geräuschvolle Pausen perzipiert als bei den freudigen Äußerungen.

Stimmfärbung (Timbre)

1) Freude

- Klangfarbe: mittel (43,8 Prozent) bis hell (42,2 Prozent)
- Klangfülle: klangvoll (68,8 Prozent)
- Faukale Distanz (Rachenweite): mittel (39,3 Prozent), Tendenz hin zu weit (29,3 Prozent)
- Stimmklang: mittel (56 Prozent)
- Stimmansatz: Kraftstimme (50,3 Prozent)

2) Ärger

- Klangfarbe: hell (42,6 Prozent)
- Klangfülle: klangvoll (44,5 Prozent), Tendenz hin zu Missklang (37,4 Prozent)
- Faukale Distanz (Rachenweite): mittel (34,8 Prozent) bis eng (32,3 Prozent)
- Stimmklang: hart (69 Prozent)

- Stimmansatz: Kraftstimme (70,3 Prozent)

Die freudigen Sprechweisen werden als klangvoller, weicher und mit weiterem Rachenraum perzipiert als die ärgerlichen Sprechweisen. Diese werden als härter, kräftiger, enger und mit einer starken Tendenz hin zum Missklang beschrieben.

4.2.1 Forschungsfrage 2a

2.a Gibt es bei oben genannter Frage einen sprach/kulturbedingten, mit der Fähigkeit des funktionellen Hörens (durch entsprechende Ausbildung/berufliche Tätigkeit) verknüpften oder geschlechterspezifischen Unterschied?

4.2.2 Sprache/Kultur

Sprachlich/kulturell bedingte Differenzen in der Perzeption und Analyse der prosodischen Merkmale ziehen sich durch alle Parameter in unterschiedlichem Ausmaß durch. Im Folgenden werden diejenigen prosodischen Merkmale und Eigenschaften dargestellt, in welchen relevante Unterschiede in der Wahrnehmung und Beschreibung zwischen der deutsch- und der italienischsprachigen Proband*innengruppe auftreten. Die überblicksartige Beschreibung ist als zusammenfassende Übersicht der wichtigsten Unterschiede zu verstehen, für eine detaillierte Darstellung siehe Kapitel 3.3.

Akzent/Rhythmus

1) Freude

- **Sprechrhythmus:** Die italienischsprachigen Proband*innen perzipieren den Sprechrhythmus als deutlich fließender (54,29 Prozent) als die deutschsprachigen Proband*innen (40,65 Prozent).

2) Ärger

- **Akzentuierung (Stärke):** Die deutschsprachige Gruppe (82,18 Prozent) nimmt die Betonungen als stärker wahr als die italienischsprachige (72,73 Prozent).

- **Sprechrhythmus:** Hier ist der Unterschied besonders auffallend – 54,46 Prozent der deutschsprachigen Proband*innen beschreibt den Rhythmus als eher abgehackt, 43,64 Prozent der italienischsprachigen Personen hingegen als fließend.

Die sprachlich/kulturell bedingten Unterschiede in Bezug auf das Merkmal Akzent/Rhythmus werden auf die unterschiedlichen Hörmuster der Proband*innen zurückgeführt. Handelt es sich beim akzentzählenden Deutschen eher um eine harte, stoßende Sprache, ist das silbenzählende Italienisch viel melodischer und sanglicher. Im Ergebnis zur Perzeption und Beschreibung des Akzent/Rhythmus zeigt sich, dass die Vertreter der jeweiligen Sprach-/Kulturgruppe je die für die eigene Sprache charakteristischen Merkmale heraushören.

Tonhöhe/Melodieführung

1) Freude

- **Sprechstimmlage:** Die italienischsprachigen Proband*innen beurteilen die Sprechstimmlage als höher (52,86 Prozent) als die deutschsprachigen Proband*innen (47,97 Prozent).
- **Tonhöhenverlauf:** Die italienischsprachige Gruppe nimmt ihn als bewegter (54,29 Prozent) wahr als die deutschsprachige (43,94 Prozent).

2) Ärger

- **Sprechstimmlage:** Auch hier perzipieren die italienischsprachigen Personen die Sprechstimmlage als höher (58,18 Prozent) als die deutschsprachigen (50,5 Prozent).
- **Tonhöhenumfang:** Der Umfang zwischen dem höchsten und dem niedrigsten Ton wird von deutschsprachigen Proband*innen als größer eingeschätzt als von den italienischsprachigen Personen.

Zwar kommen die italienisch- und die deutschsprachigen Proband*innen bei der Beurteilung dieses, von beiden Gruppen als relevantestes Merkmal eingeschätzten Parameters, zum selben Schluss, aber die Abstufungen sind z. T. sehr unterschiedlich und werden auf die jeweiligen diversen Hörmuster der Personen zurückgeführt.

Lautstärke

1) Freude

- **Lautheitsverlauf:** Diese Eigenschaft wird zwar von beiden Gruppen als gleichförmig beurteilt, aber die italienischsprachigen Proband*innen beurteilen ihn als bewegter als die deutschsprachigen Proband*innen.
- **Lautheitsverlauf am Ende der Phrase:** Beide Gruppen nehmen eine gleichbleibende Lautstärke wahr, für die italienischsprachige Gruppe ist der Lautheitsverlauf gegen Ende der Phrase jedoch lauter.

2) Ärger

- **Lautstärke:** Sie wird von beiden Gruppen als eher laut beurteilt, wird von den deutschsprachigen Proband*innen jedoch für lauter befunden als von den italienischsprachigen.
- **Lautheitsverlauf:** Die italienischsprachigen Proband*innen halten ihn für bewegter als die deutschsprachigen Personen.
- **Lautheitsverlauf am Ende der Phrase:** Er wird von beiden Gruppen als gleichbleibend perzipiert, für die für die italienischsprachige Gruppe ist der Lautheitsverlauf gegen Ende der Phrase jedoch auch hier lauter.

Die Lautstärke konnte recht klar und einheitlich von beiden Gruppen umrissen werden, trotzdem zeigen sich klare sprachlich/kulturell bedingte Differenzen in beiden emotionalen Ausprägungen. Die Hörmuster in Bezug auf die Lautstärke differieren in den diversen Sprach-/Kulturgruppen und wirken sich zweifelsohne auf die Beurteilung dieses Merkmals aus.

Sprechgeschwindigkeit

1) Freude

- **Geschwindigkeit:** Für die italienischsprachigen Proband*innen klingen die freudigen Sprechweisen etwas langsamer als für die deutschsprachigen Proband*innen.

2) Ärger

- **Geschwindigkeit:** Auch die ärgerlichen Sprechweisen klingen für die italienischsprachigen Personen deutlich langsamer als für die deutschsprachige Gruppe.

- **Geschwindigkeitsverlauf:** Dieser wird von den italienischsprachigen Proband*innen als deutlich bewegter wahrgenommen.
- **Geschwindigkeitsverlauf am Ende der Phrase:** Er wird zwar von beiden als gleichbleibend beurteilt, aber für die italienischsprachige Gruppe klingt er rascher als für die deutschsprachige.

Auch die Beurteilung der Sprechgeschwindigkeit ist an kulturell geprägte Sprech- und Hörmuster gekoppelt und unterscheidet sich demzufolge zumindest teilweise zwischen den italienischsprachigen und den deutschsprachigen Proband*innen.

Pausen

Bei der Beurteilung der Pausen liegen keine relevanten sprach-/kulturspezifischen Differenzen vor. Beide Gruppen nehmen unabhängig von der Emotion Freude oder Ärger in gleichem Ausmaß wenige Pausen und wenige geräuschvolle Pausen wahr.

Stimmfärbung (Timbre)

1) Freude

- **Klangfarbe:** Hier liegt ein großer sprachlich/kultureller Unterschied vor. Die Klangfarbe wird von der deutschsprachigen Gruppe als mittelmäßig bewertet, von der italienischsprachigen Gruppe als hell.
- **Klangfülle:** Die italienischsprachigen Proband*innen beurteilen die freudigen Sprechweisen als deutlich klangvoller als die deutschsprachigen Personen.
- **Faukale Distanz (Rachenweite):** Sie wird von beiden Gruppen als mittel bis weit wahrgenommen, allerdings können deutlich mehr italienischsprachige als deutschsprachige Personen diese Frage nicht beantworten.
- **Stimmklang:** Er wird von den italienischsprachigen Proband*innen als weicher eingeschätzt als von den deutschsprachigen.
- **Stimmansatz:** Hier liegt auch ein sehr großer Unterschied vor – nur 42,98 Prozent der deutschsprachigen, aber 62,86 Prozent der italienischsprachigen Proband*innen geben die Verwendung der Kraftstimme an. Die italienischsprachige Gruppe beurteilt das Timbre also als kräftiger.

2) Ärger

- **Klangfarbe:** Anders als bei den freudigen Sprechweisen empfinden die deutschsprachigen Proband*innen die Klangfarbe der ärgerlichen Äußerungen als heller als die italienischsprachigen.
- **Klangfülle:** Auch hier empfinden die italienischsprachigen Personen die Äußerungen als deutlich klangvoller als die deutschsprachigen Proband*innen, welche deutlich mehr Missklang perzipieren.
- **Faukale Distanz (Rachenweite):** Beide Gruppen beurteilen die Rachenweite als mittel bis eng, wobei sie von den italienischsprachigen Personen noch enger bewertet wird. Außerdem können mehr italienischsprachige als deutschsprachige Proband*innen die Frage nicht beantworten.
- **Stimmklang:** Anders als bei den freudigen Sprechweisen wird er hier von den italienischsprachigen Personen als härter beurteilt.
- **Stimmansatz:** 63,37 Prozent der deutschsprachigen und 83,33 Prozent der italienischsprachigen Personen geben an, die Verwendung der Kraftstimme wahrzunehmen, d. h. auch hier perzipieren die italienischsprachigen Proband*innen ein kräftigeres Timbre.

Das prosodische Merkmal Stimmfärbung (Timbre) ist sehr komplex und umfasst viele Eigenschaften. Zwar kommen die beiden Gruppen größtenteils zum selben Ergebnis, aber doch lassen sich in allen Bereichen z. T. relevante sprachlich/kulturelle Differenzen feststellen. Auch die unterschiedliche Perzeption dieses Merkmals lässt sich auf verschiedene Hörmuster der Proband*innen zurückführen.

4.2.3 Beruf

Auch berufsbezogene Unterschiede in der Perzeption und Analyse der prosodischen Merkmale lassen sich in unterschiedlichem Ausmaß in allen Kategorien finden und sollen im Folgenden überblicksartig dargestellt werden. Dabei werden die wichtigsten Differenzen zusammengefasst, für eine detaillierte Präsentation wird auf das Kapitel 3.3 verwiesen.

Akzent/Rhythmus

1) Freude

- **Akzentuierung (Häufigkeit):** Die Stimmexpert*innen nehmen deutlich mehr Betonungen wahr (52 Prozent) als die Laien (Bewertung „eher viele Betonungen“: 42,37 Prozent; Bewertung „mittel“: 39,83 Prozent).
- **Sprechrhythmus:** Die Stimmexpert*innen halten den Rhythmus für deutlich fließender als die Laien.

2) Ärger

- **Sprechrhythmus:** Die Stimmexpert*innen beurteilen den Rhythmus als deutlich abgehackter im Vergleich zu den Laien, welche keine klare Differenzierung vornehmen.

Beim Vergleich der Analysen wird klar, dass die Stimmexpert*innen das prosodische Merkmal Akzent/Rhythmus in beiden emotionalen Ausprägungen differenzierter und stimmiger beurteilen als die Laien. Dies wird auf die größere Erfahrung der Expert*innen und deren geschulteres Wahrnehmungsvermögen zurückgeführt.

Tonhöhe/Melodieführung

1) Freude

- **Sprechstimmlage:** Die Stimmexpert*innen schätzen die Sprechstimmlage als höher ein als die Laien, welche weniger differenziertere Abstufungen vornehmen.
- **Tonhöhenverlauf:** Die Stimmexpert*innen halten den Tonhöhenverlauf für bewegter als die Laien.
- **Tonhöhenumfang:** Beide Gruppen geben einen mittleren Tonhöhenumfang an, das Ergebnis der Stimmexpert*innen ist aber differenzierter. Knapp 6 Prozent der Laien können im Gegensatz zu den Expert*innen die Frage nicht beantworten.

2) Ärger

- **Sprechstimmlage:** Auch hier schätzen die Stimmexpert*innen die Sprechstimmlage als höher ein als die Laien.
- **Tonhöhenverlauf:** Der Tonhöhenverlauf wird von den Stimmexpert*innen (69,35 Prozent) als deutlich bewegter perzipiert als von den Laien (55,32 Prozent).

- **Tonhöhenumfang:** Hier zeigt sich dasselbe Ergebnis wie bei den freudigen Sprechweisen, d. h. beide Gruppen nehmen den Umfang zwischen dem höchsten und dem niedrigsten Ton für mittelgroß wahr, das Ergebnis der Stimmexpert*innen ist jedoch klarer.

Zwar kommen Laien wie Expert*innen insgesamt zur selben Beurteilung der einzelnen Eigenschaften der Tonhöhe/Melodieführung, die Analyse der Stimmexpert*innen ist jedoch deutlich differenzierter, was wohl mit dem differenzierteren Wahrnehmungs- und Beurteilungsvermögen der Expert*innen in Zusammenhang gebracht werden kann.

Lautstärke

1) Freude

- **Lautstärke:** Die freudigen Sprechweisen werden von den Stimmexpert*innen als lauter beurteilt als von den Laien.
- **Lautheitsverlauf am Ende der Phrase:** Er wird von beiden Gruppen als gleichbleibend perzipiert, allerdings von den Expert*innen als leiser wahrgenommen als von den Laien.

2) Ärger

- **Lautheitsverlauf:** Die Laien nehmen den Lautheitsverlauf als gleichförmig wahr (Bewertung „gleichförmig“: 44,68 Prozent, Bewertung „mittel“: 31,91 Prozent), die Stimmexpert*innen als etwas bewegter (Bewertung „gleichförmig“: 43,55 Prozent; Bewertung „mittel“: 40,32 Prozent).
- **Lautheitsverlauf am Ende der Phrase:** Beide Gruppen perzipieren zwar eine gleichbleibende Lautstärke, für die Stimmexpert*innen ist der Lautheitsverlauf gegen Ende der Phrase jedoch deutlich leiser.

Die Analyseergebnisse zur Lautstärke sind zwar ähnlich, aber doch z. T. auch recht unterschiedlich. Dabei spielen berufsbezogene Aspekte stets im Zusammenhang mit anderen Faktoren (z. B. Sprache/Kultur) eine Rolle.

Sprechgeschwindigkeit

1) Freude

- **Geschwindigkeitsverlauf:** Er wird von den Laien (Bewertung „mittel“: 37,29 Prozent, Bewertung „gleichförmig“: 33,05 Prozent) als bewegter wahrgenommen als von den Stimmexpert*innen (Bewertung „gleichförmig“ 42,67 Prozent).
- **Geschwindigkeitsverlauf am Ende der Phrase:** Beide Gruppen vernehmen eindeutig eine gleichbleibende Geschwindigkeit. Dies gilt in besonderem Maße für die Laien, während der Prozentsatz der Stimmexpert*innen, die einen langsamer bzw. schneller werdenden Geschwindigkeitsverlauf perzipieren, im Vergleich deutlich höher ist.

2) Ärger

- **Geschwindigkeit:** In Bezug auf diese Eigenschaft gibt es einen klaren Unterschied – 69,15 Prozent der Laien, aber nur 59,02 Prozent der Expert*innen beurteilen die Geschwindigkeit als eher schnell.

Die Bewertungen der Sprechgeschwindigkeit der beiden Gruppen sind insgesamt dieselben. Dass die Ergebnisse doch in ihrer Ausprägung differieren, wird v. a. auf die diversen beruflich bedingten Hör- und Analysegewohnheiten der Personen zurückgeführt. Diese wirken natürlich stets im Zusammenhang mit anderen persönlichkeitsbildenden Faktoren.

Pausen

1) Freude

- **Häufigkeit:** Obgleich beide Gruppen wenige Pausen perzipieren, tritt diese Bewertung doch häufiger bei den Stimmexpert*innen auf.
- **Geräuschvolle Pausen:** Pausen, die mit Klang oder Geräuschen (z. B. lautes Atmen, Seufzen, Lachen usw.) gefüllt sind, kommen laut Laien und Expert*innen sehr wenige vor. Letztere nehmen diese Bewertung jedoch deutlich öfters vor (Bewertung „eher wenige“: 51,69 Prozent der Laien vs. 65,33 Prozent der Expert*innen).

2) Ärger

- **Häufigkeit:** Die Angaben sind fast identisch (wenige Pausen), tendenziell nehmen die Laien etwas weniger Pausen wahr als die Expert*innen.

- **Geräuschvolle Pausen:** Auch hier ist das Ergebnis sehr ähnlich, wobei die Laien etwas weniger Pausen wahrnehmen als die Stimmexpert*innen.

Es ist interessant, dass die Ergebnisse hier v. a. in Bezug auf die jeweilige Emotion differieren – im Rahmen der freudigen Sprechweisen nehmen die Stimmexpert*innen im Vergleich mit den Laien weniger Pausen wahr, bei den ärgerlichen Sprechweisen mehr. Insgesamt sind die Ergebnisse jedoch sehr ähnlich.

Stimmfärbung (Timbre):

1) Freude

- **Klangfarbe:** Hier liegt ein großer Unterschied vor – während die Laien (47,46 Prozent) die Klangfarbe als „mittel“ beurteilen, beschreiben die Stimmexpert*innen (47,3 Prozent) sie als hell.
- **Faukale Distanz (Rachenweite):** Hier zeigt sich eine große berufsbedingte Differenz. Beide Gruppen beurteilen die Rachenweite als mittel bis weit, allerdings können deutlich mehr Laien (30,77 Prozent) als Expert*innen (2,7 Prozent) die Frage nicht beantworten.
- **Stimmklang:** Die Laien wahrnehmen ihn als weicher als die Expert*innen.
- **Stimmansatz:** Auch der Stimmansatz wird von den Laien tendenziell als weicher wahrgenommen.

2) Ärger

- **Klangfarbe:** Auch die ärgerlichen Sprechweisen werden von den Stimmexpert*innen als heller beurteilt als von den Laien. Mehr Laien als Expert*innen können die Frage nicht beantworten.
- **Klangfülle:** Hier liegt ein deutlicher Unterschied vor – während die Laien das Timbre als klangvoll bezeichnen, wahrnimmt die Mehrheit der Stimmexpert*innen einen Missklang.
- **Faukale Distanz (Rachenweite):** Große Unterschiede zeigen sich auch hier v. a. darin, dass 30,85 Prozent der Laien die Frage nicht beantworten konnten. Alle Stimmexpert*innen hingegen beantworteten sie.
- **Stimmklang:** Anders als bei den freudigen Sprechweisen wird der Stimmklang bei den ärgerlichen Äußerungen von den Laien als härter empfunden.

- **Stimmansatz:** Hier wird auch der Stimmansatz von den Laien als kräftiger eingestuft.

Die Stimmfärbung (Timbre) ist ein komplexes Merkmal. Gerade hier zeigen sich große berufsbedingte Differenzen, wobei die Stimmexpert*innen v. a. in Bezug auf die Beurteilung der fakalen Distanz eindeutig auf ihren berufsbedingten Erfahrungsschatz und ihr phonetisches Wissen und Wahrnehmungsvermögen zurückgreifen können. Hier kommt dieser Unterschied besonders klar zur Geltung.

4.2.4 Geschlecht

Akzent/Rhythmus

1) Freude

- **Akzentuierung (Häufigkeit):** Die Männer nehmen deutlich mehr Betonungen wahr als die Frauen.
- **Sprechrhythmus:** Die Frauen perzipieren öfters einen fließenden Rhythmus als die Männer.

2) Ärger

- **Sprechrhythmus:** Für die Frauen klingt der Rhythmus der ärgerlichen Sprechweisen stärker abgehackt als für die Männer.

Es ist hier und im Folgenden klar ersichtlich, dass sich das Geschlecht auf die Perzeption und Analyse der Äußerungen auswirkt. Es wird, wie bereits zu Forschungsfrage eins ausgeführt, als Persönlichkeitsmerkmal verstanden, das biologische, soziale, gesellschaftliche u. a. Implikationen auf den Menschen hat. Die Gründe für die unterschiedlichen Ergebnisse der Männer und Frauen müssten also in weiteren Untersuchungen geklärt werden. Hier und in Bezug auf die im Folgenden beschriebenen prosodischen Parameter sollen nur die konkreten Differenzen dargestellt werden.

Tonhöhe/Melodieführung

1) Freude

- **Sprechstimmlage:** Sie wird von den Frauen (49,57 Prozent) als deutlich höher wahrgenommen als von den Männern (39,47 Prozent).
- **Tonhöhenumfang:** Der Umfang zwischen höchstem und niedrigstem Ton wird von beiden Gruppen als mittelmäßig angegeben, das Ergebnis der Männer ist jedoch klarer differenziert.

2) Ärger

- **Sprechstimmlage:** Auch bei den ärgerlichen Sprechweisen schätzen die Frauen (Bewertung „eher hoch“: 59,57 Prozent) die Sprechstimmlage höher ein als die Männer (43,55 Prozent).
- **Tonhöhenverlauf:** Die Frauen (64,89 Prozent) perzipieren den Tonhöhenverlauf als bewegter als die Männer (54,84 Prozent).
- **Tonhöhenumfang:** Auch hier liegt ein großer Unterschied vor. 54,84 Prozent der Männer, aber nur 41,49 Prozent der Frauen bewerten den Tonhöhenumfang als mittelmäßig, die weiteren Abstufungen (klein bzw. groß) sind bei beiden Gruppen unklar. Trotzdem ist das Ergebnis der Männer differenzierter als das der Frauen.

Während die Frauen die Sprechstimmlage bei Freude und Ärger stets höher beurteilen als die Männer, so nehmen diese eine genauere Differenzierung in Bezug auf den Tonhöhenumfang vor als die Frauen. Jegliche Generalisierung dieser Ergebnisse wäre, ohne weitere empirische Grundlagen, jedoch nicht haltbar.

Lautstärke

1) Freude

- **Lautheitsverlauf:** Eine leichte Differenz zeigt sich hier darin, dass 6,58 Prozent der Männer, aber nur 0,85 Prozent der Frauen die Frage nicht beantworten können.

2) Ärger

- **Lautheitsverlauf:** Hier zeigt sich ein großer Unterschied – während die Männer den Lautheitsverlauf als mittel (38,71 Prozent) bis gleichförmig (33,87 Prozent) wahrnehmen, perzipieren ihn die Frauen (51,06 Prozent) v. a. als gleichförmig.
- **Lautheitsverlauf am Ende der Phrase:** Noch größer ist die Differenz hier - 46,77 Prozent der Männer, aber 69,15 Prozent der Frauen nehmen eine gleichbleibende Lautstärke wahr.

Interessant ist, dass sich die geschlechterbezogenen Differenzen in den Analyseergebnissen v. a. in Bezug auf den Lautheitsverlauf zeigen. Die anderen Eigenschaften der Lautstärke werden sehr ähnlich beurteilt.

Sprechgeschwindigkeit

1) Freude

- **Geschwindigkeitsverlauf:** Er wird von den Männern als bewegter wahrgenommen als von den Frauen. Zugleich ist das Ergebnis der Männer aber auch undifferenzierter (circa gleich viele Angaben in allen Kategorien).

2) Ärger

- **Geschwindigkeitsverlauf am Ende der Phrase:** 46,77 Prozent der Männer, aber 64,52 Prozent der Frauen perzipieren eine gleichbleibende Geschwindigkeit am Ende der Phrase. Das Ergebnis der Männer ist undifferenzierter als das der Frauen.

Auch hier zeigen sich geschlechterbedingte Differenzen, wie bei der Analyse der Lautstärke, nur in Bezug auf den Geschwindigkeitsverlauf, wobei das Ergebnis der Männer undifferenzierter ist als jenes der Frauen. Natürlich gilt auch hier, dass diese Feststellung keinesfalls verallgemeinert werden kann.

Pausen

1) Freude

- **Häufigkeit:** Die Männer nehmen etwas mehr Pausen wahr als die Frauen.

- **Geräuschvolle Pausen:** Die Männer nehmen etwas mehr mit Klang oder Geräuschen (z. B. lautes Atmen, Seufzen, Lachen usw.) gefüllte Pausen wahr als die Frauen.

2) Ärger

- **Häufigkeit:** Auch im Rahmen der ärgerlichen Sprechweisen perzipieren die Männer etwas mehr Pausen als die Frauen.

Insgesamt nehmen Männer und Frauen wenige Pausen und wenige geräuschvolle Pausen wahr, wenngleich die Männer tendenziell mehr Pausen perzipieren als die Frauen.

Stimmfärbung (Timbre)

1) Freude

- **Klangfarbe:** Für die Männer klingt die Klangfarbe der freudigen Sprechweisen etwas heller als für die Frauen.
- **Klangfülle:** Die Frauen (72,65 Prozent) nehmen das Timbre als deutlich klangvoller wahr als die Männer (62,67 Prozent).
- **Faukale Distanz (Rachenweite):** Hier liegt ein sehr großer Unterschied vor. Die Männer nehmen die faukale Distanz als mittel bzw. weit wahr (je 29,33 Prozent), die Frauen v. a. als mittel (45,69 Prozent), außerdem können mehr Männer (33,33 Prozent) als Frauen (11,21 Prozent) die Frage nicht beantworten.
- **Stimmklang:** Er wird von den Männern als weicher perzipiert als von den Frauen.
- **Stimmansatz:** Die Frauen nehmen einen deutlich kräftigeren Stimmansatz wahr.

2) Ärger

- **Klangfarbe:** Anders als bei den freudigen Sprechweisen nehmen hier die Frauen (48,39 Prozent) die Klangfarbe als deutlich heller wahr als die Männer (33,87 Prozent).
- **Klangfülle:** Hier ist ein relevanter Unterschied zwischen den Geschlechtern ersichtlich. Das Timbre wirkt auch hier für die Frauen (51,61 Prozent) viel klangvoller als für die Männer (33,87 Prozent).
- **Faukale Distanz (Rachenweite):** Beide Gruppen perzipieren die Rachenweite als mittel bis eng, doch 33,87 Prozent der Männer können im Vergleich zu 8,6 Prozent der Frauen die Frage nicht beantworten.

- **Stimmklang:** Er wird im Unterschied zu den freudigen Sprechweisen von den Männern als härter empfunden als von den Frauen.
- **Stimmansatz:** Auch hier können deutlich mehr Männer (16,13 Prozent) als Frauen (1,08 Prozent) die Frage nicht beantworten.

Auch in Bezug auf die Perzeption und Analyse des prosodischen Merkmals der Stimmfärbung (Timbre) lassen sich in diversen Bereichen z. T. sehr große geschlechtsspezifische Differenzen beobachten. Generell scheinen die Frauen etwas sicherer zu sein in der Beurteilung der Parameter (v. a. ersichtlich bei der Beschreibung der falkalen Distanz) als die Männer und das Timbre in beiden Emotionsausprägungen als klangvoller zu beurteilen. Generalisierungen lassen sich aber auch hier keine treffen – die geschlechterspezifischen Differenzen müssen konkret am jeweiligen prosodischen Merkmal bzw. an seinen Eigenschaften und an der jeweiligen Emotion festgemacht werden.

4.2.5 Fazit

Es konnte dargelegt werden, wie die einzelnen prosodischen Merkmale der freudigen bzw. ärgerlichen Sprechweisen von den Proband*innen insgesamt beschrieben wurden. Die Analyseergebnisse wurden in einem zusammenfassenden Überblick dargestellt und erörtert. Es kann festgehalten werden, dass die prosodischen Merkmale klar analysiert und beschrieben werden konnten, die Ergebnisse sind eindeutig ablesbar und unterscheiden sich z. T. sehr je nach emotionaler Ausprägung (Freude bzw. Ärger). In einem zweiten Schritt wurde der Einfluss der Faktoren Sprache/Kultur, Beruf und Geschlecht auf die Perzeption und Analyse der prosodischen Parameter dargestellt. Hierbei kann Folgendes festgehalten werden:

Sprache/Kultur

Sprachlich/kulturell bedingte Differenzen betreffen alle prosodischen Merkmale, bis auf die Pausen, in unterschiedlichem Ausmaß und in Bezug auf die Ausprägung der Emotion (Freude bzw. Ärger). Daraus lässt sich folgern, dass keine verallgemeinernden Aussagen getroffen werden können, sondern nur im Hinblick auf die jeweilige zugrundeliegende Emotion und auf das konkrete prosodische Merkmal und dessen Eigenschaften. Die Sprache/Kultur stellt ein sehr relevanter Einflussfaktor auf die Perzeption und Analyse der prosodischen Parameter bzw. auf die Proband*innen dar. Die unterschiedlichen Bewertungen werden auf die diversen sprachlich/kulturell geprägten Hörmuster der Proband*innen zurückgeführt. Natürlich muss beachtet

werden, dass alle Faktoren stets miteinander interagieren. Insofern müssen alle Einflussfaktoren zusammen als komplexes Merkmalsbündel verstanden werden, das auf das Individuum einwirkt.

Beruf

Auch die berufsbezogenen Fähigkeiten und das berufsbezogene Wissen stellen in der Perzeption und Analyse der prosodischen Merkmale einen relevanten Faktor dar. Dabei geht es hier und bei der Diskussion der Ergebnisse der anderen Gruppen nicht darum, wer „recht hat“. Aus diesem und anderen Gründen wurde auch auf eine phonetische Analyse der Hörbeispiele verzichtet, zumal sich außerdem das phonetisch messbare Ergebnis oftmals nicht mit dem individuellen Höreindruck der Proband*innen deckt. Um doch eine gewisse Richtschnur zu haben, wurden die Äußerungen von einer erfahrenen Sprechwissenschaftlerin als Referenzhörer*in analysiert. Ihre Analyse deckt sich im Allgemeinen mit dem Gesamtergebnis der Proband*innen. Vielmehr geht es aber darum aufzuzeigen, dass die hier untersuchten Faktoren (Sprache/Kultur, Beruf, Geschlecht) eine Rolle in der Perzeption und Analyse spielen, und darzustellen, inwieweit und auf welche Art und Weise sie sich darauf auswirken. Dabei gilt auch hier, dass alle Aussagen hierzu nicht generalisiert werden können, sondern konkret an die jeweilige Emotion und an das jeweilige prosodische Merkmal festgemacht werden müssen. Wenngleich nicht für jedes prosodische Merkmal in gleichem Ausmaß gültig, so zeigt sich doch, dass die Stimmspezialist*innen im Vergleich zu den Laien insgesamt eine differenziertere Analyse mit klareren Abstufungen vornehmen und demzufolge ein stimmigeres Ergebnis aufweisen.

Geschlecht

Das Geschlecht zeigt sich ebenfalls als relevante Größe im Perzeptions- und Analyseprozess der prosodischen Merkmale. Ausführungen dazu, was hier unter Geschlecht verstanden wird, finden sich im Abschnitt zur ersten Forschungsfrage und werden an dieser Stelle nicht noch einmal thematisiert. Wie bei den anderen Faktoren gilt auch hier festzuhalten, dass geschlechtsspezifische Differenzen keinesfalls verallgemeinert werden dürfen, sondern vielmehr nur im Hinblick auf die jeweilige Emotion und die konkreten prosodischen Eigenschaften betrachtet werden können. Außerdem stehen die Faktoren in wechselseitiger Beeinflussung. Fest steht jedenfalls, dass sich das Geschlecht relevant auf die Perzeption und Analyse der freudigen bzw. ärgerlichen Sprechweisen auswirkt und daher auch in Folgestudien unbedingt miteinbezogen werden sollte.

Allein die Tatsache, dass Männer und Frauen und italienisch- und deutschsprachige Personen alle zum Ergebnis kommen, die jeweilige emotionale Äußerung als freudig bzw. ärgerlich wahrzunehmen, ist nicht selbstverständlich, denn auch hier liegen z. T. große interkulturelle Unterschiede vor (vgl. Chen et al. 2004). In diesem Bereich gibt es noch viel Forschungsbedarf. Das Spannende in dieser Untersuchung, in der die Sprechweisen ja vorausgewählt wurden, so dass sie für alle Proband*innen eindeutig emotional zuordenbar sind, sind aber die unterschiedliche Wahrnehmung und Einschätzung der für die emotionale Bestimmung so relevanten prosodischen Merkmale. Es ist faszinierend, wie sehr sich die Perzeption je nach Sprache/Kultur, Beruf und Geschlecht der Proband*innen unterscheidet. Die der Äußerung zugrundeliegende Emotion wird in dieser Untersuchung zwar von allen gleich als Freude oder Ärger interpretiert, aber viele der anderen Parameter, die diese Äußerung charakterisieren und daher auch deren Aussage für den Hörer beeinflussen, unterscheiden sich z. T. sehr zwischen den Proband*innen. So lassen sich auch Missverständnisse in der Kommunikation zwischen deutsch- und italienischsprachigen Personen, Männern und Frauen und auch Laien und Expert*innen erklären.

4.3 Forschungsfrage 3

Neben den zentralen beiden Hauptforschungsfragen, die in den obigen Abschnitten beantwortet wurden, soll im Folgenden noch auf die zwei Nebenforschungsfragen eingegangen werden.

3. Wird der Ausprägungsgrad der Emotion von allen Proband*innen gleich empfunden oder gibt es dabei sprachlich/kulturelle, geschlechts- und berufsbezogene Unterschiede?

Mit dieser Frage wurde untersucht, wie stark die Proband*innen die Emotion einschätzen und ob es hierbei Differenzen aufgrund der Faktoren der Sprache/Kultur, Beruf und Geschlecht gibt. Die Antworten beziehen sich, wie im Kapitel 3.3.3 dargestellt, nur auf die Emotion Freude, und zeigen auf, dass v. a. die Einflussfaktoren Sprache/Kultur und Beruf, aber auch Geschlecht, auf die Perzeption des Ausprägungsgrades der Emotion einwirken. Alle Proband*innen nehmen die Emotion insgesamt als recht stark ausgeprägt wahr – die deutschsprachigen Proband*innen empfinden sie jedoch als deutlich stärker als die italienischsprachigen und die Stimmexpert*innen als stärker als die Laien. Der geschlechtsspezifische Unterschied ist eng an die jeweilige Sprache/Kultur der Proband*innen gebunden und hat innerhalb der jeweiligen Sprach-/Kulturgruppe relevante Implikationen – die italienischsprachigen Männer perzipieren die Emotion

Freude als deutlich stärker als die italienischsprachigen Frauen, innerhalb der deutschsprachigen Gruppe zeigt sich eine geringer ausgeprägte umgekehrte Tendenz. Am ausgeprägtesten ist die sprachlich/kulturell bedingte Differenz – sie wird auf die verschiedenen Sprech- und Hörmuster der Proband*innen zurückgeführt. Es wird vermutet, dass Emotionen im italienischsprachigen Raum generell mit größerem Impetus und stärkerem Gebrauch der prosodischen Mittel (z. B. Lautstärke) verbalisiert werden. Dass die Stimmexpert*innen die Emotion als stärker ausgeprägt beurteilen als die Laien, könnte möglicherweise auf deren geschulteres Gehör für den Einsatz emotional gefärbter Prosodie zurückgeführt werden. So ist die Perzeption der prosodischen Parameter ausschlaggebend für die richtige Einordnung des Gehörten und notwendigerweise mit einer mehr oder minder ausgeprägten Fähigkeit des funktionellen Hörens verknüpft. Alle Aussagen zu dieser Fragestellung müssen aber im Kontext der entsprechenden Emotion (Freude) und der gegenseitigen Wechselwirkung aller Faktoren betrachtet werden und können daher nicht automatisch auf andere Emotionen oder Situationen übertragen werden.

4.4 Forschungsfrage 4

3. Besteht eine Übereinstimmung der perzipierten prosodischen Eigenschaften der Hörbeispiele mit den für die jeweiligen emotionalen Sprechweisen vorliegenden typischen phonetischen Charakteristika (vgl. Kranich 2003)?

Wie bereits an anderer Stelle erörtert wurde, liegt das Ziel der Befragung und Analyse nicht darin, festzustellen, ob die Einschätzung der Proband*innen „richtig“ ist, also mit den objektiven Daten der instrumentalphonetisch durchgeführten Messung der Sprechäußerungen übereinstimmt. Die subjektiven Höreindrücke unterscheiden sich oftmals auch davon, außerdem lassen sich etwa für die Beschreibung (z. B. „hell“, „klangvoll“ etc.) des Merkmals Stimmfärbung (Timbre) gar keine eindeutigen physikalisch messbaren Entsprechungen finden (Kranich 2016). Die obige Forschungsfrage ist also als Unterfrage zu Forschungsfrage zwei anzusehen und soll lediglich überprüfen, ob und inwieweit die Höreindrücke der Proband*innen mit den von Kranich (2003) postulierten phonetischen Merkmalen der emotionalen Sprechweisen übereinstimmen. Dabei gilt zu beachten, dass in Kranichs Studie simulierte expressive Sprechweisen von Schauspieler*innen, Sänger*innen und Lehrer*innen untersucht wurden (Kranich 2003: 79) und die Gültigkeit der eruierten phonetischen Charakteristika der emotionalen Sprechweisen also nicht problemlos auch für natürliche emotionale Sprechweisen angenom-

men werden kann. Außerdem hat Kranich nicht alle Kategorien untersucht, die den Proband*innen in dieser Studie zur Analyse vorgelegt wurden. Daher werden nur jene Merkmale angeführt, zu denen ein Ergebnis von Seiten der messphonetischen Untersuchung von Kranich vorliegt. Zudem gibt Kranich zu bedenken, dass neben den unten angeführten eindeutigen Resultaten z. T. auch völlig divergente phonetische Eigenschaften herausgearbeitet wurden. Die Messergebnisse werden hier als Richtwerte verstanden, die mit den subjektiven Höreindrücken der Proband*innen verglichen werden.

Im Folgenden werden, um diese Forschungsfrage beantworten zu können, die suprasegmentalen Eigenschaften der analysierten Sprechausdrucksweisen von Kranich zusammengefasst und mit den Ergebnissen der hier durchgeführten Untersuchung gegenübergestellt.

Freudige Sprechweisen

Temporalität

- **Phonetisches Messergebnis (Kranich):** hohe Sprechgeschwindigkeit
- **Gesamtergebnis aus dieser Studie:** mittlere bis hohe Sprechgeschwindigkeit

In Bezug auf die Temporalität lässt sich also feststellen, dass das Ergebnis ähnlich ist, wobei die Sprechgeschwindigkeit von den Proband*innen in dieser Untersuchung nicht eindeutig als hoch eingeschätzt wurde.

Tonhöeneigenschaften

- **Phonetisches Messergebnis (Kranich):** viele Tonschritte, große Tonschritte, hohe Sprechstimmlage
- **Gesamtergebnis aus dieser Studie:** Tonhöhenverlauf bewegt, mittlerer Tonhöhenumfang, hohe Sprechstimmlage

Die perzipierten und phonetischen Charakteristika sind hier beinahe deckungsgleich, lediglich der Tonhöhenumfang wird unterschiedlich bewertet (hier: mittlerer Tonhöhenumfang vs. großer Tonhöhenumfang nach Kranich).

Stimmfärbung (Timbre)

- **Phonetisches Messergebnis (Kranich):** heller, weicher, warmer Stimmklang, faukale Weite

- **Gesamtergebnis aus dieser Studie:** mittlere bis helle Klangfarbe, klangvoll, mittlere bis weite faukale Distanz

Wenngleich nicht alle Eigenschaften zum Timbre abgedeckt sind und manche Ergebnisse nicht deckungsgleich sind, so lässt sich hier dennoch ein breiter Konsens der beiden Untersuchungen feststellen.

Ärgerliche Sprechweisen

Temporalität

- **Phonetisches Messergebnis (Kranich):** hohe Sprechgeschwindigkeit
- **Gesamtergebnis aus dieser Studie:** hohe Sprechgeschwindigkeit

Die Ergebnisse beider Untersuchungen zur Temporalität sind deckungsgleich.

Tonhöhereigenschaften

- **Phonetisches Messergebnis (Kranich):** großer Tonraum, mehr Abwärtsschritte
- **Gesamtergebnis aus dieser Studie:** bewegter Tonhöhenverlauf, mittlerer Tonhöhenumfang

Zu den Tonhöhereigenschaften bei Kranich liegen nur obige Ergebnisse vor – der Tonhöhenumfang wurde von den Proband*innen hier als geringer eingeschätzt. Über den Verlauf im Sinne einer Auf- oder Abwärtsbewegung kann hier keine Aussage getroffen werden.

Intensität

- **Phonetisches Messergebnis (Kranich):** geringe Dynamikbreite
- **Gesamtergebnis aus dieser Studie:** gleichförmiger Lautheitsverlauf

Der Lautheitsverlauf wird in beiden Studien als gleichförmig beurteilt.

Stimmfärbung (Timbre)

- **Phonetisches Messergebnis (Kranich):** heller, scharfer, klangamer Stimmklang, Behauchung

- **Gesamtergebnis aus dieser Studie:** helle Klangfarbe, klangvoll bis hin zu Missklang, mittlere bis enge faukale Distanz

Zur Stimmfärbung wurden z. T. leicht unterschiedliche Bewertungskategorien herangezogen. Auch deshalb gibt es Differenzen in den Ergebnissen. Aber es lassen sich auch durchaus Gemeinsamkeiten in der Beurteilung dieses prosodischen Merkmals finden (z. B. in Klangfarbe und Stimmklang).

Fazit

Die verwendeten Beurteilungskategorien unterscheiden sich z. T. sehr zwischen den beiden Untersuchungen, außerdem wurden hier Elemente analysiert, welche in Kranichs Studie nicht aufgegriffen sind. Außerdem gilt es natürlich zu beachten, dass die beiden Untersuchungen je unterschiedliche Vorgehensweisen und Ziele haben. Auch dies führt, gemeinsam mit den im Einleitungstext zur Forschungsfrage vier formulierten Gründen, zu Differenzen in der Beurteilung der prosodischen Parameter und zu einer nur bedingten Vergleichbarkeit. Trotzdem lassen sich die Ergebnisse der beiden Untersuchungen hier auf fruchtbare Art und Weise gegenüberstellen. Die von Kranich analysierten prosodischen Eigenschaften der freudigen und ärgerlichen Sprechweisen sind insgesamt absolut vergleichbar mit den Resultaten aus dieser Untersuchung. Manche Ergebnisse sind deckungsgleich, andere leicht divergent. Insgesamt kann jedenfalls festgestellt werden, dass der Überschneidungsbereich zwischen den phonetischen Charakteristika der emotionalen Äußerungen und den von den Proband*innen perzipierten Höreindrücken sehr groß ist.

4.5 Hypothesenbeantwortung

1. Hörer*innen sind in der Lage, verschiedene natürlich gesprochene Äußerungen den emotionalen Kategorien Ärger oder Freude ausschließlich auf der Grundlage auditiver Präsentation zuzuordnen.

Diese Hypothese kann eindeutig bestätigt werden. Wir Menschen sind spezialisiert darauf, auch nur rein auditive Eindrücke (beispielsweise am Telefon) wahrzunehmen und zu interpretieren. In der Vorstudie wurden diejenigen Hörbeispiele für die Untersuchung ausgewählt, die eindeutig von allen Personen den emotionalen Kategorien Ärger bzw. Freude zugeordnet wurden.

2. Hörer*innen vermögen die prosodischen Merkmale der einzelnen Äußerungen in unterschiedlichem Präzisionsgrad aus dem Gesamteindruck herauszufiltern und zu beschreiben.

Diese Aussage kann verifiziert werden. Wie im Kapitel vier näher ausgeführt, können einzelne prosodische Merkmale und deren Eigenschaften (z. B. Lautstärke) besser herausgehört werden als andere (z. B. faukale Distanz). Dies hängt mit der individuellen bzw. berufsbedingten Fähigkeit des funktionellen Hörens der Proband*innen zusammen. Insgesamt konnten aber alle prosodischen Merkmale gelungen perzipiert und analysiert werden.

3. Die Präzision der Zuordnung einer emotionalen Sprechausdrucksweise sowie die Analyse ihrer prosodischen Charakteristika sind sowohl produktionsseitig von der konkreten Sprechleistung und der jeweiligen Ausprägung prosodischer Merkmale als auch rezeptionsseitig von der Wahrnehmungsleistung der Hörer*innen abhängig.

Diese Annahme kann bestätigt werden. Für die Untersuchung wurden deshalb nur Sprechweisen ausgewählt, die produktionsseitig zum einen eindeutig emotional gefärbt sind und zum anderen einen besonders intensiven Gebrauch der prosodischen Mittel aufweisen. Rezeptionsseitig ist die präzise Analyse der prosodischen Merkmale untrennbar verknüpft mit der Wahrnehmungsleistung der Hörer*innen. Diese kann sich aufgrund diverser Faktoren von Person zu Person unterscheiden, wie im Kapitel 4 erörtert werden konnte.

4. Die Wahrnehmungsleistung der Hörer*innen unterscheidet sich je nach Beruf (Unterscheidungskategorie Laie vs. Expert*in für Stimme und Sprechen) bzw. Ausprägungsgrad ihrer Fähigkeit zum analytischen Hören, wobei die Expert*innen über eine präzisere Wahrnehmung der prosodischen Parameter verfügen.

Diese Hypothese kann insgesamt bestätigt werden, obgleich dies nicht für alle prosodischen Merkmale bzw. Eigenschaften und nicht bei allen Bewertungen in gleichem Ausmaß gilt. Wie in Kapitel 4 aufgezeigt werden konnte, sind die Ergebnisse der Stimmexpert*innen nicht immer klarer oder eindeutiger (über die „Richtigkeit“ des Höreindrucks wird hier ja nicht diskutiert) als jene der Laien. Bei den meisten, besonders den komplexen Fragestellungen, wie der zur faukalen Distanz, zeigt sich allerdings klar der Vorteil der Expert*innen, die im Vergleich zu

den Laien über ein phonetisches Vorwissen und ein geschultes Ohr verfügen und so insgesamt die Fragen leichter und auf differenziertere Weise beantworten können.

5. Die Perzeption der prosodischen Merkmale differiert in Bezug auf die jeweilige Sprache und Kultur der Hörer*innen, wobei die Äußerungen aufgrund der jeweiligen sprachlich-kulturellen Charakteristika des Deutschen und Italienischen (vgl. Kapitel: Sprachvergleich. Prosodie im Deutschen und Italienischen) jeweils anders beurteilt werden.

Dass die Perzeption und Analyse z. T. sehr stark aufgrund der sprachlich/kulturellen Unterschiede differieren, kann eindeutig bestätigt werden. Die Gründe hierfür werden auf die jeweiligen sprachlichen Charakteristika und vor allem auf die damit verknüpften sprachlich/kulturellen Hörmuster der Proband*innen zurückgeführt. Sprachlich/kulturell geprägte charakteristische Sprechausdrucksweisen, auch und gerade in Bezug auf die Basisemotionen, wirken auf die Hörerwartung der Proband*innen und damit deren Wahrnehmung und Beurteilung des Gehörten ein.

6. Die Perzeption der prosodischen Merkmale unterscheidet sich nicht aufgrund des jeweiligen Geschlechts der Hörer*innen, sondern ist vielmehr sozial bedingt.

Diese Hypothese ist falsch. Auch das Geschlecht spielt eine zentrale Rolle bei der Wahrnehmung und Analyse der emotionalen Sprechweisen, es wirkt sich zusammen mit den anderen Faktoren und Persönlichkeitsmerkmalen wesentlich auf den Perzeptionsprozess aus.

7. Die den verschiedenen emotionalen Sprechäußerungen zugeordneten prosodischen Charakteristika entsprechen in groben Zügen der für die jeweilige gesprochene Emotion typischen phonetischen Eigenschaften (nach Kranich 2003).

Diese Hypothese kann bestätigt werden. Wie im Abschnitt zu Forschungsfrage vier näher dargestellt, ist der Konsens zwischen den beiden Ergebnissen trotz einiger Differenzen und Schwierigkeiten durchaus gegeben.

8. Die Wahrnehmung und Zuordnung der einzelnen emotionalen Äußerungen erfolgt ganzheitlich als komplexes Phänomen durch eine Zusammenwirkung aller sprachlichen und prosodischen Parameter mit verschiedenartiger Dominanz.

Dies wurde schon von anderen Autor*innen erläutert und kann in dieser Untersuchung bestätigt werden. Erst alle Parameter zusammen ermöglichen den spezifischen Höreindruck des Gesagten. Dabei spielen manche prosodische Merkmale jedoch eine wichtigere Rolle als andere. Hier ist die Reihenfolge der Parameter, die für die Perzeption herangezogen werden, vom wichtigsten bis zum weniger wichtigen Element: Tonhöhe/Melodieführung, Akzent/Rhythmus, Lautstärke, Stimmfärbung (Timbre), Sprechgeschwindigkeit, Text/Wortwahl und Pausen.

- a) Die Kategorie Akzent/Rhythmus ist weniger relevant aufgrund der Schwierigkeit, sie an emotionalen Merkmalen festzumachen.**

Diese Hypothese stellt sich als falsch heraus. Der Akzent/Rhythmus ist nach der Tonhöhe/Melodieführung das wichtigste Element im Perzeptionsprozess der emotionalen Sprechweisen und konnte auch von den Proband*innen klar analysiert werden.

- b) Die Merkmale Tonhöhe/Melodieführung verfügen über eine herausragende Bedeutung im Perzeptionsprozess, v. a. bei der Emotion Freude (hohe Sprechstimmlage, großer Tonumfang erwartet).**

Diese Annahme ist korrekt, tatsächlich ist die Tonhöhe/Melodieführung das wichtigste Merkmal im Perzeptionsprozess. Dies gilt nicht nur für die freudigen, sondern auch für die ärgerlichen Sprechweisen. Des Weiteren wurde eine hohe Sprechstimmlage, aber kein großer, sondern ein mittlerer Tonumfang von den Proband*innen festgestellt.

- c) Text und Wortwahl spielen eine zentrale Rolle bei der Zuordnung des Gehörten, weil sie in allen alltagstypischen Kommunikationssituationen zunächst die primäre Sinnkonstitution erlauben, beeinträchtigen aufgrund der entsprechenden Aufgabenstellungen aber nicht die Perzeption der prosodischen Merkmale.**

Der zweite Teil der hier formulierten Hypothese kann bestätigt werden – durch die entsprechende Konzeption des Fragebogens konnte die Aufmerksamkeit der Hörer*innen bewusst weg vom Text und hin zu den prosodischen Merkmalen gelenkt werden. Der Text stellte keinen Störfaktor im Perzeptions- und Analyseprozess der Prosodie dar. Der erste Teil der Hypothese kann hingegen nicht bestätigt werden. Es hat sich gezeigt, dass der Parameter Text/Wortwahl

in der vorliegenden Untersuchung eine eher untergeordnete Rolle spielt. Die Gründe hierfür sind auch in der Auswahl der Hörbeispiele zu finden, in welchen die jeweilige Emotion nicht allein durch den Text, sondern v. a. anhand der Prosodie transportiert wird.

d) Die Lautstärke ist insgesamt weniger relevant, bei der Emotion Ärger jedoch etwas gewichtiger.

Diese Hypothese kann so nicht bestätigt werden. Es hat sich gezeigt, dass die Lautstärke eine wesentliche Rolle spielt im Perzeptionsprozess und zur Emotionserkennung. Dieses Merkmal steht tatsächlich an dritter Stelle im Ranking. Der zweite Teil der oben formulierten Hypothese ist korrekt, die Lautstärke spielt im Rahmen der Perzeption der ärgerlichen Sprechweisen eine relevantere Rolle als bei den freudigen Äußerungen, was darauf zurückgeführt wird, dass jene durch eine größere Lautstärke geprägt sind und dadurch leichter auch anhand dieses Merkmals der Emotion Ärger zugeordnet werden können.

e) Die Sprechgeschwindigkeit spielt eine untergeordnete Rolle, sie ist bei beiden Emotionen hoch.

Diese Annahme ist nicht grundsätzlich falsch, die Antwort darauf muss aber präzisiert werden. Im Ranking steht die Sprechgeschwindigkeit tatsächlich erst an fünfter Stelle und spielt damit eine weniger zentrale Rolle wie die davor genannten prosodischen Merkmale, trotzdem halten 63,74 Prozent der Proband*innen die Sprechgeschwindigkeit für wichtig im Perzeptionsprozess. Das Ranking gibt Aufschluss über die unterschiedliche Relevanz der einzelnen Parameter, aus der Untersuchung geht jedoch klar hervor, dass alle prosodischen Merkmale im Perzeptionsprozess zusammenwirken. Der zweite Teil der Hypothese kann teilweise bestätigt werden – die Sprechgeschwindigkeit der freudigen Sprechweisen wird als mittel bis schnell und jene der ärgerlichen Äußerungen als schnell wahrgenommen.

f) Pausen spielen in den vorliegenden Beispielen keine Rolle. Bei anderen Emotionen, z. B. Trauer, oder bei gefüllten Pausen (z. B. Lachen, Seufzen usw.) wären sie hingegen wesentlich.

Diese Hypothese ist falsch. Die Pausen spielen eine untergeordnete Rolle in den untersuchten emotionalen Sprechweisen und stehen im Ranking daher an letzter Stelle. Dennoch sind sie

nicht völlig unwichtig, sondern werden im Gegenteil von 20,34 Prozent der Proband*innen zur Emotionsbestimmung herangezogen und von 35,98 Prozent hinsichtlich des Perzeptionsprozesses für wichtig befunden. Außerdem wurden bei den optionalen Anmerkungen am Ende des Fragebogens mehrmals geräuschvolle Pausen angesprochen, also auf Hintergrundgeräusche und Lachen als Hilfestellung für die Emotionszuordnung verwiesen.

- g) Der Stimmfärbung (Timbre) wird eine herausragende Stellung zugesprochen. Sie unterscheidet sich bei den beiden Emotionen (Ärger: hell, scharf, klangarm, behaucht; Freude: hell, weich, warm, weit; Kranich 2003: 198), ist aber besonders für die Laien schwer zu erfassen.**

Diese Hypothese hat sich nur teilweise als richtig herausgestellt. Die Stimmfärbung (Timbre) befindet sich im Ranking zur Relevanz der prosodischen Parameter nur an vierter Stelle und nimmt keine herausragende Stellung ein. 49,58 Prozent der Personen bedienen sich dieses Merkmals zur Emotionsbestimmung und 64,88 Prozent der Proband*innen halten es für relevant im Perzeptionsprozess. Insgesamt nimmt die Stimmfärbung nicht so eine herausragende Stellung ein wie angenommen, spielt aber zusammen mit den anderen prosodischen Merkmalen eine wichtige Rolle beim Indizieren der emotionalen Sprechweisen. Die Annahme, dass sich die Stimmfärbung in den beiden Ausprägungen Freude und Ärger unterscheidet, konnte bestätigt werden. Auch die Vermutung, dass dieses Merkmal besonders für die Laien schwer zu analysieren ist, hat sich als richtig herausgestellt. Tatsächlich ist die Stimmfärbung das Merkmal, bei dem die meisten Schwierigkeiten in der Analyse aufgetreten sind. Das liegt v. a. an der Komplexität dieses Parameters und an der Schwierigkeit, v. a. für phonetisch ungeschulte Hörer*innen, die Timbreeigenschaften herauszuhören und zu kategorisieren.

4.6 Zusammenfassung

In der vorliegenden Arbeit konnten theoretische Kenntnisse zur Prosodieperzeption in Bezug auf emotionale Sprechweisen vertieft und empirische Ergebnisse hierzu generiert werden. Es wurde in den Kapiteln eins und zwei des zweiten Teiles der Arbeit aufgezeigt, welche prosodischen Parameter bei der Perzeption und Zuordnung des Gehörten eine Rolle spielen und welche Relevanz ihnen für den Perzeptionsprozess zukommt. Dabei wurde festgestellt, dass die Parameter Tonhöhe/Melodieführung, Akzent/Rhythmus, Lautstärke, Stimmfärbung (Timbre), Sprechgeschwindigkeit, Text/Wortwahl und Pausen in dieser Reihenfolge (vom wichtigsten

zum weniger wichtigen Element) relevant sind für den Perzeptionsprozess, wobei alle prosodischen Merkmale stets als Merkmalskomplex zusammenwirken. Die Sprache/Kultur, der Beruf und das Geschlecht der Proband*innen wirken sich darauf aus, welche Parameter zur Emotionserkennung herangezogen und inwieweit diese für den Perzeptionsprozess für relevant befunden werden.

Des Weiteren konnten im vierten Kapitel des zweiten Teiles die einzelnen prosodischen Merkmale der freudigen und ärgerlichen Sprechweisen beschrieben und die Faktoren Sprache/Kultur, Beruf und Geschlecht als relevante Einflussgrößen auf die Perzeption und Analyse der emotionalen Äußerungen herausgearbeitet werden. Auch der Ausprägungsgrad der Emotion Freude konnte in einem kürzeren Abschnitt in Bezug auf oben genannte Faktoren analysiert werden, wobei sich sprachlich/kulturelle, berufsbezogene und auch geschlechtsspezifische Differenzen zeigen. Im dritten Teil der Arbeit wurden schließlich die Forschungsfragen und Hypothesen beantwortet und die durch die Fragebogenerhebung gewonnenen Beschreibungen der prosodischen Merkmale mit den aus instrumentellen Messungen resultierenden phonetischen Charakteristika (Kranich 2003) der emotionalen Sprechweisen verglichen, wobei ein breiter Konsens zwischen den instrumentalphonetisch gewonnenen Daten und dem subjektiven Höreindruck der Proband*innen festgestellt werden konnte.

Da die hier vorliegende Untersuchung über die Prosodieforschung hinaus von interdisziplinärem Interesse sein dürfte, besteht die Hoffnung, dass die hier gewonnenen Ergebnisse eine anregende Basis für weiterführende Forschungen bilden.

5 Literaturverzeichnis

Abercrombie, David (1967): Elements of general Phonetics. Edinburgh: Univ. Press.

Allhoff, Dieter-W. / Allhoff, Waltraud (2014): Rhetorik & Kommunikation. Ein Lehr- und Übungsbuch. 16., aktual. Aufl. München: Ernst Reinhardt Verlag.

Alter, Kai (2003): Prosodie. In: Rickheit, Gert / Herrmann, Theo / Deutsch, Werner (Hg.). Psycholinguistik – Ein internationales Handbuch. Psycholinguistics – An international handbook. Berlin: Walter de Gruyter. S. 339-346.

Altmann, Hans / Batlinger, Anton / Oppenrieder, Wilhelm (Hgg.): Zur Intonation von Modus und Fokus im Deutschen. Tübingen 1989: Niemeyer. (Ling. Arbeiten 234)

Arnold, Magda B. (1960): Emotion and personality: Vol 1. Psychological aspects. New York, NY: Academic Press.

Baldauf, Heike (2002): Knappes Sprechen. Tübingen: Niemeyer.

Bannert, Robert (1985a): Towards a Model for German Prosody. In: Folia Linguistica XIX/3-4, 1985, S. 321-341.

Barfield, Thomas (1997): The dictionary of anthropology. Oxford, UK: Blackwell.

Barker, Geoffrey (2005): Intonation Patterns in Tyrolean German. An Autosegmental-Metrical Analysis. New York: Peter Lang. (Berkeley insights in linguistics and semiotics Vol. 57)

Bolinger, Dwight (1983): Intonation and gesture. In: American Speech. A Quarterly of Linguistics Usage 58, 156-174.

Borod, Joan (Ed.) (2000): The neuropsychology of emotion. Oxford. New York: Oxford University Press.

Bose, Ines (1994): Zur temporalen Struktur frei gesprochener Texte. Frankfurt a. M.: Hector.

Bose, Ines (2001): Methoden der Sprechausdrucksbeschreibung am Beispiel kindlicher Spielkommunikation. In: Gesprächsforschung – Online Zeitschrift zur verbalen Interaktion, Ausgabe 2 (2001), S. 262-303.

Bose, Ines (2003): dóch da sín ja´ nur m Bose, Ines (2003): dóch da sín ja´ nur muster. Kindlicher Sprechausdruck im sozialen Rollenspiel. Frankfurt, München: Peter Lang. (Hallesche Schriften zur Sprechwissenschaft und Phonetik)

Bose, Ines / Hirschfeld, Ursula / Neuber, Baldur / Stock, Eberhard (2013): Einführung in die Sprechwissenschaft. Phonetik, Rhetorik, Sprechkunst. Tübingen: Narr Francke Attempto Verlag. (narr Studienbücher)

Bose, Ines / Hirschfeld, Ursula / Neuber, Baldur / Stock, Eberhard (2016): Einführung in die Sprechwissenschaft. Phonetik, Rhetorik, Sprechkunst. 2., überarb. u. erw. Aufl. Tübingen: Narr Francke Attempto Verlag. (narr Studienbücher)

Brück, Carolin / Kreifelts, Benjamin / Ethofer, Thomas / Wildgruber, Dirk (2013): Affektive Prosodie. In: Schröger, Erich / Birbaumer, Koelsch, Stefan (Hg.). Enzyklopädie der Psychologie: Themenbereich C, Theorie und Forschung, Serie 2, Kognition, Bd.5: Affektive und kognitive Neurowissenschaft. Göttingen, Bern, Toronto, Seattle: Hogrefe, Verlag für Psychologie. S. 254 – 280.

Busso, Carlos / Bulut, Murtaza / Lee, Sungbok / Narayanan, Shrikanth (2009): Fundamental Frequency Analysis for Speech Emotion Processing. In: Sylvie Hancil (Hg.): Linguistic Insights, 97. The Role of Prosody in Affective Speech. Bern: Peter Lang, S. 309-337.

Chen, Aoju / Gussenhoven, Carlos / Rietveld, Toni (2004): Language-Specificity in the Perception of Paralinguistic Intonational Meaning. In: Language and Speech, 47 (4), S. 311-349.

Collignon, O. / Girard, S. / Gosselin, F. / Saint-Amour, D. / Lepore, F. / Lassonde, M. (2010): Women process multisensory emotion expressions more efficiently than men. In: Neuropsychologia, 48. S. 220–225.

Cowie, Roddy / Douglas-Cowie, Ellen (2009): Prosodic and Related Features that signify Emotional Colouring in Conversational Speech. In: Sylvie Hancil (Hg.): Linguistic Insights, 97. The Role of Prosody in Affective Speech. Bern: Peter Lang, S. 213-240.

Damasio, Antonio (2005): Descartes' Error: Emotion, reason and the human brain. 4. Aufl. London: Penguin Books.

Darwin, Charles (1998): The expression of emotion in man and animals (3rd ed., P. Ekman, Ed.). Oxford, UK: Oxford University Press. (Original work published 1872).

Dasher, Richard / Bolinger, Dwight (1982): On pre-accentual lengthening. In: Journal of the International Phonetic Association, 12, S. 58-69.

Degenhardt, Annette / Trautner, Hans M. (Hg.): Geschlechtstypisches Verhalten. Mann und Frau in psychologischer Sicht. München 1979: Beck.

Dietrich, Rainer (2007): Psycholinguistik. Stuttgart: Metzler.

Dogil, Grzegorz (2003): Understanding Prosod. In: Rickheit, Gert / Herrmann, Theo / Deutsch, Werner (Hg.). Berlin: Walter de Gruyter, S. 544-565.

Drach, Erich (1926): Die redenden Künste. Leipzig: Quelle & Meyer.

Drescher, Martina (2003): Sprachliche Affektivität. Darstellung emotionaler Beteiligung am Beispiel von Gesprächen aus dem Französischen. Tübingen: Max Niemeyer Verlag. (Linguistische Arbeiten 486)

Einstein, Gillian (Ed.) (2007): Sex and the brain. Cambridge, Massachusetts Institute of Technology: The MIT Press.

Ekman, Paul (1992): An argument for basic emotions. In: Cognition and Emotion 6 (3/4), S. 169-200.

Ekman, Paul / Friesen, Wallace V. (1975): Unmasking the face. A guide to recognizing emotions of Personality and Affect. In: *Journal of Personality and Social Psychology* 38 (2), S. 270-277.

Ekman, Paul / Friesen, Wallace V. (1969): The repertoire of nonverbal behavior – Categories, origins, usage, and coding. *Semiotica*, 1, S. 49-98.

Enterlein, Ines / Bartels, Astrid / Sendlmeier, Walter (2005): Prosodische Merkmale als Indikatoren der Sprechereinstellung. In: Sendlmeier, Walter / Bartels, Astrid (Hg.): *Stimmlicher Ausdruck in der Alltagskommunikation. Mündliche Kommunikation Bd. 4*. Berlin: Logos Verlag, S. 9-38.

Erickson, Frederick (1992): „They Know All The Lines“. Rhythmic Organization and Contextualization in a Conversational Listing Routine. In: Auer/di Luzio (eds.), S. 365-397.

Essen, Otto v. (1957): *Allgemeine und Angewandte Phonetik*. 2. Aufl. Berlin: Akademie-Verlag.

Essen, Otto v. (1979): *Allgemeine und Angewandte Phonetik*. Berlin: Akademie-Verlag.

Fährmann, Rudolf (1982): Elemente der Stimme und Sprechweise. In: Scherer, Klaus R. (Hg.): *Vokale Kommunikation. Nonverbale Aspekte des Sprachverhaltens*. Weinheim und Basel: Beltz Verlag, S. 138-163.

Falk, Simone (2009): *Musik und Sprachprosodie. Kindgerichtetes Singen im frühen Spracherwerb*. Berlin: Walter de Gruyter.

Ferrer, Eduardo Blasco (1999): *Italiano e tedesco. Un confronto linguistico*. Turin: Paravia scriptorium.

Ferstl, Evelyn C. / Kaiser, Anelis (2013): *Sprache und Geschlecht. Wie quantitative Methoden aus der Experimental- und Neuropsychologie einen Beitrag zur Geschlechterforschung leisten können*. Humboldt-Universität zu Berlin: 2013. <https://doi.org/10.25595/557>

Féry, Caroline (1993): *German Intonational Patterns*. Tübingen: Max Niemeyer Verlag. (Linguistische Arbeiten 285)

Fiehler, Reinhard / Barden, Birgit / Elstermann, Mechthild / Kraft, Barbara (2004): *Eigenschaften gesprochener Sprache*. Tübingen: Gunter Narr Verlag. (Studien zur Deutschen Sprache, Bd. 30)

Fiehler, Reinhard (1990b): *Kommunikation und Emotion*. Berlin/New York: de Gruyter. (Grundlagen der Kommunikation und Kognition)

Fischinger, Timo (2009): *Zur Psychologie des Rhythmus. Präzision und Synchronisation bei Schlagzeugern*. Kassel: University Press.

Fox, Anthony (2002): *Prosodic features and prosodic structure. The phonology of suprasegmentals*. Univ. Press Oxford.

Friederici, Angela D. (2006): *Neurobiologische Grundlagen der Sprache*. In: Karnath, Hans-Otto / Thier, Peter (Hgg.): *Neuropsychologie*. Berlin/Heidelberg: Springer Verlag, S. 346-355.

Fry, Dennis B. (1985): *Experiments in the perception of stress*. In: *Language and Speech* 1, S. 126-152.

Gadler, Hanspeter (2006): *Praktische Linguistik. Eine Einführung in die Linguistik für Logopäden und Sprachheillehrer*. 4., erw. u. aktual. Aufl. Tübingen: Narr Francke Attempo Verlag. (UTB 1411)

Gilles, Peter (2005): *Regionale Prosodie im Deutschen. Variabilität in der Intonation von Abschluss und Weiterweisung*. Berlin: Walter de Gruyter. (Linguistik – Impulse & Tendenzen, Bd. 6)

Gogolin, Ingrid / Krüger-Portratz, Marianne (2006): *Einführung in die Interkulturelle Pädagogik*. Oplanden u. a.: Verlag Barbara Budrich. (UTB 8246)

Golestani, Narly / Price, Cathy J. / Scott, Sophie K. (2011): Born with an Ear for Dialects? Structural Plasticity in the Expert Phonetician Brain. In: Journal of Neuroscience 31 (11), S. 4213-4220.

Grassegger, Hans (2010): Phonetik, Phonologie. Schulz-Kirchner Idstein.

Griffiths, Timothy D. (2003): Functional imaging of pitch analysis. In: Annuals N.Y. Acad. Science 999, S. 40-49.

Günther, Carsten (1999): Prosodie und Sprachproduktion. Tübingen: Max Niemeyer Verlag. (Linguistische Arbeiten 401)

Gutenberg, Norbert (2001): Einführung in Sprechwissenschaft und Sprecherziehung. Frankfurt a. M.: Lang.

Gutenberg, Norbert (1998): Einzelstudien zu Sprechwissenschaft und Sprecherziehung. Arbeiten in Teilfeldern. Göppingen: Kümmerle.

Hagemann, Jörg/Henle, Julia (2014): Transkribieren nach GAT 2 (Minimal- und Basistranskript) – Schritt für Schritt. [https://www.ph-freiburg.de/fileadmin/dateien/mitarbeiter/hagemannfr/Transkribieren_nach_GAT_2.pdf]

Hagen, Mechthild (2003): Förderung des Hörens und Zuhörens in der Schule. Begründung, Entwicklung und Evaluation eines Handlungsmodells. Phil. Diss. München. (unveröffentl. Mskr.)

Hammarström, Göran (1963b): Prosodeme und Kontureme. In: Phonetica: International Journal of Speech Science 10, S. 194-202.

Harasty, Jenny / Double, Kay / Halliday, Gary / Kril, Jillian / McRitchie, Donna (1997): Language Associated Cortical Regions Are Proportionally Larger in the Female Brain. In: Archives of Neurology, 54(2). S. 171–176.

Harf, Rainer (2011): Der Pulsschlag der Moderne. In: GEO kompakt 27. Das Rätsel Zeit. S. 16-128.

Hausendorf, Heiko / Quasthoff, Uta (1995): Discourse and Oral Contextualization: Vocal Cues. In: Quasthoff (ed.), S. 220-255.

Heilmann, Christa M. (2011): Geschlechtstypische Unterschiede. In: Pabst-Weinschenk, Marita (Hg.): Grundlagen der Sprechwissenschaft und Sprecherziehung. 2. überarb. Aufl. München: Ernst Reinhardt Verlag. S. 162-171.

Heilmann, Christa M. (2011): Körpersprache richtig verstehen und einsetzen. 2. Aufl. München: Ernst Reinhardt Verlag.

Heilmann, Christa M. (2002): Interventionen im Gespräch: Neue Ansätze der Sprechwissenschaft. Niemeyer Verlag Tübingen.

Heilmann, Christa M. (2002b): Wer schreit gewinnt? Neuere Ergebnisse empirischer Genderforschung zur nonverbalen Kommunikation. In: Faschingbauer (2002a), S.115-136.

Helfrich, Hede (2013): Kulturvergleichende Psychologie. Wiesbaden: Springer.

Helmholtz, Hermann (1877): Die Lehre von den Tonempfindungen. 2. Aufl. Braunschweig: Friedrich Vieweg.

Heringer, Hans-Jürgen (2014): Interkulturelle Kommunikation. Grundlagen und Konzepte. Tübingen: Francke.

Herrmann, Theo (2005): Sprache verwenden. Funktionen – Evolution – Prozesse. Stuttgart: Kohlhammer Verlag.

Hielscher, Martina (2003): Sprachrezeption und emotionale Bewertung. In: Rickheit, Gert / Herrmann, Theo / Deutsch, Werner (Hg.). Psycholinguistik: Ein internationales Handbuch. Berlin: Walter de Gruyter. S. 677 – 707.

Hirschfeld, Julia / Rosenberg, Andrew (2009): Charisma perception from text and speech. In: *Speech Communication*, 51, 7, S. 640-655.

Hirschfeld, Ursula / Neuber, Baldur (2010): Prosodie im Fremdsprachenunterricht Deutsch – ein Überblick über Terminologie, Merkmale und Funktionen. In: *Deutsch als Fremdsprache* (1), S. 10-16.

Hirschfeld, Ursula / Neuber, Baldur (2010): Methodische Überlegungen zur Untersuchung phonetischer und rhetorischer Parameter im interkulturellen Wissenschaftsdiskurs. In: *sprechen* 27:49, S. 45-55.

Hirschfeld, Ursula / Stock, Eberhard (2016): Positionen und Probleme sprechwissenschaftlicher Untersuchungen zur interkulturellen Kommunikation. In: Hirschfeld, Ursula / Lange, Friederike / Stock, Eberhard (Hg.): *Phonetische und rhetorische Aspekte der interkulturellen Kommunikation*. Frank & Timme Verlag Berlin, S. 13-22. (Schriften zur Sprechwissenschaft und Phonetik Bd. 7)

Hirschfeld, Ursula / Stock, Eberhard (2013a): Phonologische Grundlagen des Deutschen. In: Bose, Ines / Hirschfeld, Ursula / Neuber, Baldur / Stock, Eberhard (Hg.): *Einführung in die Sprechwissenschaft. Phonetik, Rhetorik, Sprechkunst*. Tübingen: Narr Verlag, S. 27-80.

Hirschfeld, Ursula / Stock, Eberhard (2010): Untersuchungen zur interkulturellen sprechsprachlichen Kommunikation. In: Hirschfeld, Ursula / Stock, Eberhard (Hg.): *Sprechwissenschaftlich-phonetische Untersuchungen zur interkulturellen Kommunikation Russisch – Deutsch*. Peter Lang Verlag Frankfurt a. M., S. 9-28. (Hallesche Schriften zur Sprechwissenschaft und Phonetik 33)

Hofstede, Geert (2003): *Lokales Denken globales Handeln. Interkulturelle Zusammenarbeit und globales Management*. München: DTV.

Imhof, Margarete (2003): *Zuhören. Psychologische Aspekte auditiver Informationsverarbeitung*. Göttingen: Vandenhoeck & Ruprecht.

Isačenko, Alexander V. / Schädlich, Hans-Joachim (1966): Untersuchungen über die Satzintonation. In: Bierwisch, Manfred (Hg.): Untersuchungen über Satzakzent und Intonation im Deutschen. Berlin: Akademie Verlag, S. 7-68. (Studia Grammatica VII)

Isen, Alice. M. (1984): Toward understanding the role of affect in cognition. In: R. S. Wyer & T. K. Srull (Eds.), Handbook of social cognition (Vol. 3, pp. 179-236). Hillsdale, NJ: Erlbaum.

Iversen, John R. / Aniruddh D. Patel / Kengo Ohgushi (2008): Perception of rhythmic grouping depends on auditory experience. In: JASA 124, S. 2263-2271.

Izard, Carroll Ellis (1971): The face of emotion. New York, NY: Appleton-Century-Crofts.

James, William (1884): What is an Emotion? In: Mind, 9, S. 188-205.

Janney, Richard W. (1996): Speech and Affect. Emotive Uses of English. München: o.V.

Johnson-Laird, Philip N. / Oatleys, Keith (1989): The language of emotions: An analysis of a semantic field. In: Cognition and Emotion 3, S. 81-123.

Jussen, Heribert / Kloster-Jensen, Martin / Wisotzki, Karl Heinz (1994): Lautbildung bei Hörgeschädigten. Abriss einer deutschen Sprachlautlehre, 3. überarb. Aufl. Berlin: Ed. Marhold.

Kallmeyer, Werner (1981): Gestaltungsorientiertheit in Alltagserzählungen. In: Rolf Kloepfer/Gisela Janetzke-Dillner (Hgg.): Erzählung und Erzählforschung im 20. Jahrhundert. Stuttgart: Kohlhammer. S. 409-429.

Kammhuber, Stefan (2011): Interkulturelle Aspekte. In: Pabst-Weinschenk, Marita (Hg.): Grundlagen der Sprechwissenschaft und Sprecherziehung. 2. überarb. Aufl. München: Ernst Reinhardt Verlag. S. 171-180.

Kehrein, Roland (2002): Prosodie und Emotionen. Tübingen: Max Niemeyer Verlag. (Reihe Germanistische Linguistik 231)

Keim, Inken (2001): Die Powergirls. Aspekte des kommunikativen Stils einer Migrantinnen-
gruppe aus Mannheim. In: Jakobs/Rothkegel (Hgg.), S. 375-400.

Kelter, Stephanie (2003): Mentale Modelle. In: Rickheit, Gert / Herrmann, Theo / Deutsch,
Werner (Hgg.): Psycholinguistik. Ein internationales Handbuch. An International Handbook.
Bd. 24. (Handbücher zur Sprach- und Kommunikationswissenschaft), Berlin: Walter de Gruy-
ter, S. 505-517.

Keßler, Christian (2013): Italienisch. In: Hirschfeld, U. / Kelz, H. P. / Müller, U. (Hg.): Phone-
tik international. Grundwissen von Afrikaans bis Zulu. Kontrastive Studien für Deutsch als
Fremdsprache. www.phonetik-international.de

Kienast, Miriam (2002): Phonetische Veränderungen in Emotionaler Sprechweise. Disserta-
tion, techn. Universität Berlin. Aachen: Shaker Verlag.

Kohler, Klaus J. (1995): Einführung in die Phonetik des Deutschen. Berlin: Schmidt Verlag.

Kotthoff, Helga (1997): Erzählstile von mündlichen Witzen. Zur Erzielung von Komikeffekten
durch Dialoginszenierungen und die Stilisierung sozialer Typen im Witz. In: Selting/Sandig
(Hgg.): Sprech- und Gesprächsstile. Berlin/New York: de Gruyter, S. 123-169.

Kotthoff, Helga (1996): Die Geschlechter in der Gesprächsforschung. Hierarchien, Theorien,
Ideologien. In: Der Deutschunterricht 1: Differenzen – diesseits und jenseits von Geschlechter-
fixierungen, S.9-15.

Kotthoff, Helga (1993): Kommunikative Stile, Asymmetrie und „doing gender“. Fallstudien
zur Inszenierung von Expert(inn)entum in Gesprächen. In: Feministische Studien 2, S.79-96.

Kowal, S. / Barth, H.-C. / Egemann, H. / Galusic, G. / Kögel, C. / Lippold, N. / Pfeil, A. /
O'Connel, D. (1998): Unterbrechungen in Medieninterviews – Geschlechtstypisches Ge-
sprächsverhalten? In: Schoenthal (1998a), S.279-299.

Kranich, Wieland (2016): Sprechwissenschaftliche Grundlagen der Prosodieperzeption. Berlin: Frank & Timme GmbH. (Schriften zur Sprechwissenschaft und Phonetik Bd. 4)

Kranich, Wieland (2015): Perzeption prosodischer Merkmale: Probleme ihrer empirischen Untersuchung. In: Teuchert, Brigitte (Hg.). Aktuelle Forschungstendenzen in der Sprechwissenschaft: Normen, Werte, Anwendung. Baltmannsweiler: Schneider Verlag Hohengehren. (Sprache und Sprechen Bd. 48)

Kranich, Wieland (2014a): Überlegungen zur Beschreibung der Klangqualität gesprochener Sprache. In: Sprechen 58, S.21-38.

Kranich, Wieland (2003): Phonetische Untersuchungen zur Prosodie emotionaler Sprechausdrucksweisen. Frankfurt am Main: Peter Lang Europäischer Verlag der Wissenschaften. (Halle'sche Schriften zur Sprechwissenschaft und Phonetik Bd. 11)

Kreimann, Jody / Vanlancker-Sidtis, Diana / Gerratt, Bruce R. (2005): Perception of Voice Quality. In: Pisoni, David B; Remez, Povert E. (Hgg.): The Handbook of Speech Perception. Cornwall: Blackwell. S. 338-362.

Künzel, Hermann J. (1989): Phonetische Untersuchungen zur Sprecher-Erkennung durch linguistisch naive Personen. Stuttgart: Franz Steiner Verlag.

Ladd, Robert (2008): Intonational phonology. 2nd edn. Cambridge: Cambridge University Press.

Lazarus, Hans / Sust, Charlotte A. / Steckel, Rita / Kurtz, Patrick (2007): Akustische Grundlagen sprachlicher Kommunikation. Berlin: Springer Verlag.

Lazarus, Richard S. (1968): Emotions and adaption. Conceptual and empirical relations. In: W. J. Arnold (Ed.), Nebraska Symposium on Motivation (Vol. 16, S. 175-270). Lincoln, NE: University of Nebraska Press.

Lazarus, Richard S. / Coyne, James / Folkman, Susan (1984): Cognition, emotion and motivation. The doctoring of 'humpty-dumpty'. In: K. R. Scherer / P. Ekman (Eds.), Approaches to emotion (S. 221-237). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lindner, Gerhard (1977): Hören und Verstehen. Phonetische Grundlagen der auditiven Lautsprachperzeption. Berlin: Akademie.

Lieberman, Alvin M. / Mattingly, Ignatius G (1985): The motor theory of speech perception revised. *Cognition* 21, S. 1-36.

Lieberman, Alvin M. / Cooper, Franklin S. / Harris, Katherine S. / MacNeilage, Peter (1963): A motor theory of speech perception. *Proceedings of Speech Communication Seminar Vol.2*. Stockholm.

Lipkowski, Eva / Schüller, Liane (2017): Perspektiven für den inklusiven Deutschunterricht. Literatur- und sprachdidaktische Praxisbeispiele zum Thema Hören und Höreinschränkungen. Münster: Waxmann.

Lüdtke, Ulrike (2002a): Die Sprache der Gefühle – Gefühle in der Sprache. Kommunikation und Regulation von Emotionen in Therapie und Unterricht. In: Kolberg, T., Otto, K., Wahn, C. (Hg.): *Phänomen Sprache. Laut- und Schriftsprachstörungen unter veränderten Kommunikationsbedingungen* (394-404). Rimpf: von Freisleben.

Lüdtke, U. (2004a): Emotionen im Unterricht. Theorie und Praxis einer Relationalen Didaktik im Förderschwerpunkt Sprache. In: Grohnfeldt, M. (Hg.): *Lehrbuch der Sprachheilpädagogik und Logopädie, Bd. 5: Bildung, Erziehung und Unterricht* (106-126). Stuttgart: Kohlhammer.

Lüdtke, U. (2004b): Emotionen in Therapie und Unterricht – Grundlagen einer Relationalen Didaktik. In: Lüdtke, U. (Hg.): *Fokus: MENSCH. Subjektzentrierte Unterrichts- und Therapiemodelle in der Sprachbehindertenpädagogik* (187-201). Würzburg: von Freisleben.

Lüdtke, U. (2005): Sprache und Emotion: Vom Logos zum Dialog. Zur Konstruktion einer Relationalen Theorie der Sprachbehindertenpädagogik: Erkenntnistheorie – Sprachtheorie – Didaktiktheorie. Unveröffentlichte Habilitationsschrift, Bremen: Universität Bremen.

Lüdtke, U. (2006a): Learning how to mean: Semiolinguistic aspects of intersubjective communication. In: Frank, B., Marwick, H. (Eds.): Psychological, sociological and linguistic aspects of intersubjectivity in early childhood.

Lüdtke, U. (2006b): Sprache und Emotion: Linguistische, neurowissenschaftliche und didaktische Zusammenhänge. In: Bahr, R., Iven, C. (Hg.): Sprache – Emotion – Bewusstheit. Beiträge zur Sprachtherapie in Schule, Praxis, Klinik (17-26). Idstein: Schulz-Kirchner.

Lüdtke, U. (2007). „Unterricht“ als intersubjektive Konstruktion: Zur emotionalen Regulation sprachlicher Lehr-Lern-Prozesse an der Schnittstelle von Individuellem, Sozialem und Kulturellem. Sprachtherapeutische Förderung im Unterricht. Kolberg, T. (Hg.). Festschrift zur Emeritierung von Prof. Braun. Stuttgart: Kohlhammer, 161-182.

Lüsebrink, Hans-Jürgen (2016): Interkulturelle Kommunikation. Interaktion, Fremdwahrnehmung, Kulturtransfer. Stuttgart: Metzler.

Mangold, Max (1984): Hören ohne Verstehen. In: Gutenberg, Norbert (Hg.): Hören und Beurteilen. Scriptor Frankfurt a. M., S. 149-158, (Sprache und Sprechen 12).

Matsumoto, David / Hee Yoo, Seung / LeRoux, Jeffrey A. (2007): Emotion and intercultural adjustment. In: Kotthoff, Helga / Spencer-Oatey, Helen (Hgg.): Handbooks of Applied Linguistics. Handbook of Intercultural Communication, 7. Berlin: Mouton de Gruyter, S. 77-97.

McDougall, William (1928/1969): An outline of psychology (4th ed.). London: Methuen.

Meinhold, Gottfried (1995): Zeitparameter gesprochener Sprache. Forschungsbericht. Universität Jena. Lehrstuhl für Phonetik und Sprechwissenschaft. Unveröffentlichtes Typoskript.

Meise, Katrin (1996): Une forte absence. Schweigen in alltagsweltlicher und literarischer Kommunikation. Tübingen: Narr.

Meyer, Martin / Alter, Kai / Friederici, Angela D. / Lohmann, Gabriele / v. Cramon, Yves D. (2002): FMRI Reveals Brain Regions Mediating Slow Prosodic Modulations in Spoken Sentences. In: *Human Brain Mapping* 17, S. 73-88.

Meyer, Wulf U. / Schützwohl, Achim / Reisenzein, Rainer (1993): *Einführung in die Emotionspsychologie* (Bd. 1). Bern: Huber.

Meyer, Wulf U. / Schützwohl, Achim / Reisenzein, Rainer (1997): *Einführung in die Emotionspsychologie* (Bd. 2). Bern: Huber.

Miller, George A. / Johnson-Laird, Philip N. (1976): *Language and Perception* Cambridge/Massachusetts: The Belknap Press of Harvard University Press.

Missaglia, Federica (2001): Die Erkennung simulierter Emotionen durch deutsche und italienische Muttersprachler – Implikationen für den DaF-Unterricht in Italien. In: *Zeitschrift für Angewandte Linguistik* 35, 81-96.

Moosmüller, Alois (Hg.) 2007: *Interkulturelle Kommunikation. Konturen einer wissenschaftlichen Disziplin*. Münster: Waxmann.

Neuber, Baldur (2002): *Prosodische Formen in Funktion. Leistungen der Suprasegmentalia für das Verstehen, Behalten und die Bedeutungs(re)konstruktion*. Frankfurt am Main: Peter Lang. Europäischer Verlag der Wissenschaften. (Hallesche Schriften zur Sprechwissenschaft und Phonetik Bd. 7)

Neuber, Baldur (2001): Bedeutung des Schalls – Gedanken zur Semiotik akustischer Ereignisse. In: Bräunlich, Margret / Neuber, Baldur / Rues, Beate (Hg.): *Gesprochene Sprache – transdisziplinär*. Festschrift zum 65. Geburtstag von Gottfried Meinhold. Peter Lang Verlag Frankfurt a. M., S. 103-109. (Hallesche Schriften zur Sprechwissenschaft und Phonetik 5)

Niebuhr, Oliver (2007): *Perzeption und kognitive Verarbeitung der Sprechmelodie. Theoretische Grundlagen und empirische Untersuchungen*. Berlin: Walter de Gruyter.

Noel Aziz Hanna, Patrizia (2003): Sprachrhythmus in Metrik und Alltagssprache. Untersuchungen zur Funktion des neuhochdeutschen Nebenakzents. München: Fink Verlag.

Nübling, Damaris / Dammel, Antje / Duke, Janet / Szczepaniak (2006): Historische Sprachwissenschaft des Deutschen. Eine Einführung in die Prinzipien des Sprachwandels. Tübingen: Günter Narr Verlag.

Oppenrieder, Wilhelm (1988): Intonation und Identifikation. Kategorisierungstests zur kontextfreien Identifikation von Satzmodi. In: Altmann, Hans (Hg.): Intonationsforschungen. Niemeyer Verlag Tübingen, S. 153-168.

O'Shaughnessy, Douglas (1996): Speech communication. Human and machine. Mass: Addison-Wesley Reading.

Otto, Jürgen / Euler, Harald A. / Mandl, Heinz (2000): Begriffsbestimmungen. In J. Otto, H. A. Euler und H. Mandl (Hg.), Handbuch Emotionspsychologie (S. 11-18). Weinheim: Beltz, PsychologieVerlagsUnion.

Paeschke, Astrid (2003): Prosodische Analyse emotionaler Sprechweise. Berlin: Logos Verlag.

Pabst-Weinschenk, Marita (Hg.): Grundlagen der Sprechwissenschaft und Sprecherziehung. 2., überarb. Aufl. München 2011: Ernst Reinhardt Verlag. (UTB 8294)

Panksepp, Jaak (2003): At the interface of the affective, behavioral, and cognitive neurosciences: Decoding the emotional feelings of the brain. Brain Cognition 52, 4-14.

Paul, Ingwer (2000): Gerahmte Kommunikation. Die Inszenierung ost-westdeutscher Kommunikationserfahrungen im Mediendiskurs. In: Peter Auer/Heiko Hausendorf (Hgg.): Kommunikation in gesellschaftlichen Umbruchsituationen. Tübingen: Niemeyer. S. 113-150.

Pétursson, Magnús / Neppert, Joachim M. H. (2002): Elementarbuch der Phonetik. 3. bearb. Aufl. Hamburg: Helmut Buske Verlag.

Pétursson, Magnús / Neppert, Joachim Buste (1991): Elementarbuch der Phonetik. Hamburg: Helmut Buske Verlag.

Pfordresher, Peter Q. / Brown, Steven (2009): Enhand production and perception of musical pitch in tone language speakers. In: *Attention, Perception, and Psychophysics*, 71 (6), S. 1385-1398.

Pheby, John (1981): Intonation. In: Heidolph, Karl E./Fläming, Walter (Hg.): *Grundzüge einer deutschen Grammatik*. Akademie Verlag Berlin, S. 839-897.

Pike, Kenneth L. (1945): *The Intonation of American English*. Ann Arbor, Michigan: University of Michigan Press.

Pompino-Marschall, Bernd (1995): *Einführung in die Phonetik*. Berlin/New York: Walter de Gruyter. (De-Gruyter-Studienbuch)

Pompino-Marschall, Bernd (2009): *Einführung in die Phonetik*. 3., durchgesehene Aufl. Berlin/New York: Walter de Gruyter. (De-Gruyter-Studienbuch)

Prestin, Elke (2003): Theorien und Modelle der Sprachrezeption. In: Rickheit, Gert / Herrmann, Theo / Deutsch, Werner (Hgg.): *Psycholinguistik. Ein internationales Handbuch. An International Handbook*. Bd. 24. (Handbücher zur Sprach- und Kommunikationswissenschaft), Berlin: Walter de Gruyter, S. 491-505.

Rabanus, Stefan (2001): *Intonatorische Verfahren im Deutschen und Italienischen. Gesprächsanalyse und autosegmentale Phonologie*. Tübingen: Max Niemeyer Verlag.

Ramus, Franck / Nespor, Marina / Mehler, Jacques (1999): Correlates of linguistic rhythm in the speech signal. In: *Cognition*, 73, S. 265-92.

Redecker, Beate (2006): *Persuasion und Prosodie. Eine empirische Untersuchung zur Perzeption prosodischer Stimuli in der Werbung (Hallesche Schriften zur Sprechwissenschaft und Phonetik 25)*. Frankfurt a. M.: Lang

Reinke, Kerstin (2010): Überblick über phonetische Forschungsarbeiten zur interkulturellen Kommunikation (Schwerpunkt: emotionale Sprechweise). In: Hirschfeld, Ursula / Stock, Eberhard (Hg.): Sprechwissenschaftlich-phonetische Untersuchungen zur interkulturellen Kommunikation Russisch-Deutsch. Frankfurt a. M.: Peter Lang Verlag, S. 99-114. (Hallesche Schriften zur Sprechwissenschaft und Phonetik 33)

Rickheit, Gert / Sichelschmidt, Lorenz / Strohner, Hans (2007): Psycholinguistik. Die Wissenschaft vom sprachlichen Verhalten und Erleben. Tübingen: Stauffenburg Verlag.

Rickheit, Gert / Strohner, Hans (1993): Grundlagen der kognitiven Sprachverarbeitung. Modelle, Methoden, Ergebnisse. Tübingen/Basel: Francke.

Rosenblum, Lawrence D. (2005): Primacy of Multimodal Speech Perception. In: Pisoni, David B. / Remez, Robert E. (Hg.): The Handbook of Speech Perception. Cornwall: Blackwell. S. 51-78.

Rossell, Susan L. / Bullmore, Edward T. / Williams, Steve C.R. / David, Anthony S. (2002): Sex differences in functional brain activation during a lexical visual field task. In: Brain and Language 80. S. 97-105.

Rummer, Ralf / Schweppe, Judith / Schlegelmilch, René / Grice, Martine (2014): Mood Is Linked to Vowel Type: The role of Articulatory Movements. In: American Psychological Association, Emotion, Vol. 14, (2), S. 246-250.

Sadownik, Barbara (2014): Die Modularität als Schlüssel zum Verständnis des Sprachverarbeitungssystems – kognitive Modelle und ihre neuronale Basis. In: GLOTTODIDACTICA XLI/1 (2014), S. 51-71.

Sandig, Barbara / Selting, Margret (1997): Einleitung. In: Selting, Margret / Sandig, Barbara (Hgg.): Sprech- und Gesprächsstile. Berlin: de Gruyter, S. 1-8.

Schade, Ulrich (2003): Kognitionswissenschaftliche Beiträge zu Sprachproduktion und Sprachrezeption. In: Rickheit, Gert (Hg.): Psycholinguistik. Ein internationales Handbuch. An International Handbook. Bd. 24. (Handbücher zur Sprach- und Kommunikationswissenschaft), Berlin: Walter de Gruyter, S. 80-91.

Scherer, Klaus R. / Clark-Polner, Elizabeth / Mortillaro, Marcello (2011): In the eye of the beholder? Universality and cultural specificity in the expression and perception of emotion. In: International Journal of Psychology, 2011, 46 (6), S. 401-435.

Scherer, Klaus R. (1993): Neuroscience projections to current debates in emotion psychology. Cognition and Emotion, 7, S. 1-41.

Scherer, Klaus R. (1990): Theorien und aktuelle Probleme der Emotionspsychologie. In K. R. Scherer (Hrsg.), Enzyklopädie der Psychologie (C, IV, 3). Psychologie der Emotion 1-38. Göttingen: Hogrefe.

Scherer, Klaus R. (1989): Von den Schwierigkeiten im Umgang mit den Emotionen oder: Terminologische Verwirrungen. Psychologische Rundschau, 40, S. 209-216.

Scherer, Klaus R. (1989b): Vocal correlates of emotional arousal and affective disturbance. In: Wagner, L. / Manstead, A. (Hgg.): Handbook of Psychophysiology. Emotion and social behaviour. London, S. 165-197.

Scherer, Klaus R. (1986b): Emotion experiences across European cultures: A summary statement. In: Scherer, K. R. / Wallbott, H. G. / Summerfield, A. B. (Hgg.): Experiencing emotion: A cross-cultural study. Cambridge, S. 173-190.

Scherer, Klaus R. (1982): Vokale Kommunikation. Nonverbale Aspekte des Sprachverhaltens. Weinheim und Basel: Beltz Verlag.

Scherer, Klaus R. (1978): Personality inference from voice quality: The loud voice of extraversion. European Journal of Social Psychology, 8, S. 467-487.

Schmidt-Atzert, Lothar (1987): Zur umgangssprachlichen Ähnlichkeit von Emotionswörtern. In: Psychologische Beiträge 29, S. 140-163.

Schmidt, Jürgen Erich (2001): Bausteine der Intonation? In: Germ. Ling. 157/158, S. 9-32.

Schnieders, Guido (2005): Reklamationsgespräche. Eine diskursanalytische Studie. Tübingen: Gunter Narr Verlag. (Kommunikation und Institution 25)

Schönherr, Beatrix (1997): Syntax – Prosodie – nonverbale Kommunikation. Empirische Untersuchungen zur Interaktion sprachlicher und parasprachlicher Ausdrucksmittel im Gespräch. Tübingen: Niemeyer (Reihe Germanistische Linguistik 182)

Schröder (2004): Speech and Emotion Research. An Overview of Research Frameworks and a Dimensional Approach to Emotional Speech Synthesis. In: Barry, W. J. (Hg.): Phonus 7, Saarbrücken: Institut für Phonetik Universität des Saarlandes. Dissertation. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.69.2894&rep=rep1&type=pdf> – 17.10.17

Schuller, Björn / Wöllmer, Martin / Eyben, Florian (2009): Prosodic, Spectral or Voice Quality? Feature Type Relevance for the Discrimination of Emotion Pairs. In: Sylvie Hancil (Hg.): Linguistic Insights, 97. The Role of Prosody in Affective Speech. Bern: Peter Lang, S. 285-307.

Schwarz-Friesel, Monika (2013): Sprache und Emotion. 2., aktual. u. erw. Aufl. Tübingen, Basel: A. Francke Verlag. (UTB)

Schwitalla, Johannes (2012): Gesprochenes Deutsch. Eine Einführung. 4. Neu bearb. u. erw. Aufl. Berlin: Erich Schmidt Verlag. (Grundlagen der Germanistik 33)

Sebastián-Gallés, Núria (2005): Cross-Language Speech Perception. In: Pisoni, David B; Remez, Povert E. (Hgg.): The Handbook of Speech Perception. Cornwall: Blackwell. S. 546-566.

Selting Margret [u. a.] (2009): Gesprächsanalytisches Transkriptionssystem 2 (GAT 2). In: Gesprächsforschung – Onlinezeitschrift zur verbalen Interaktion. Ausgabe 10, S. 353-402. (www.gespraechsforschung-ozs.de-28.09.2017)

Selting, Margret (1998): TCUs and TRPs: The Construction of Units in Conversational Talk. In: InLiSt – Interaction and Linguistic Structures, No.4. Potsdam. <http://www.inlist.uni-bayreuth.de/issues/4/inlist4.pdf> - 30.06.2016

Selting, Margret (1995): Prosodie im Gespräch. Aspekte einer interaktionalen Phonologie der Konversation. Tübingen: Niemeyer. (Linguistische Arbeiten 329)

Selting, Margret (1994): Empathetic speech style – with special focus on the prosodic signalling of heightened emotive involvement in conversation. In: Journal of Pragmatics, 22. S. 375 - 408.

Sievers, Eduard (1976): Grundzüge der Phonetik zur Einführung in das Studium der Lautlehre der indogermanischen Sprachen. Nachdruck der 5. Aufl. Leipzig 1901. Hildesheim/New York: Olms. [ursprünglich: Grundzüge der Lautphysiologie 1876]

Stock, Eberhard (1991): Grundfragen der Sprechwirkungsforschung. In: Krech, E./Richter, G./Stock, E./Suttner, J.: Sprechwirkung. Grundfragen, Methoden und Ergebnisse ihrer Forschung. Berlin: Akademie-Verlag.

Stock, Eberhard (1996): Text und Intonation. Sprachwissenschaft 21, S. 211-240.

Stock, Eberhard (1987): Probleme und Ergebnisse der Wirkungsforschung zu Intonation und Artikulation. In: Krech, Eva-Maria / Suttner, Jutta / Stock, Eberhard (Hgg.): Ergebnisse der Sprechwirkungsforschung. Halle (Saale): Wissenschaftliche Beiträge der Martin-Luther-Universität, S. 50-124.

Stock, Eberhard (1980): Untersuchungen zu Form, Bedeutung und Funktion der Intonation im Deutschen. Berlin: Akademie Verlag.

Stock, Eberhard / Suttner, Jutta (1991): Wirkungen des Stimm- und Sprechausdrucks. In: Krech/Richter/Stock/Suttner. Berlin: Akademie-Verlag. S. 59-142.

Stock, Eberhard / Velickova, Ludmila (2002): Sprechrhythmus im Russischen und Deutschen. Frankfurt a. M.: Peter Lang Verlag. (Hallesche Schriften zur Sprachwissenschaft und Phonetik 8)

Strange, Winifred (Hg.): *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore 1995: York Press.

Studdert-Kennedy, Michael / Liberman, Alvin M. / Harris, Katherine S. / Cooper, Franklin S. (1970): *Motor theory of speech perception: A reply to Lanes' critical review*. *Psychological Review* 77, S. 234-249.

Tatham, Mark / Morton, Katherine (2011): *A Guide to Speech Production and Perception*. Edinburgh: University Press.

Tatham, Mark / Morton, Katherine (2006): *Speech production and perception*. Basingstoke: Palgrave Macmillan Basingstoke etc.

Teuchert, Brigitte (Hg.): *Aktuelle Forschungstendenzen in der Sprechwissenschaft. Normen, Werte, Anwendung*. Baltmannsweiler 2015: Schneider Verlag Hohengehren. (Sprache und Sprechen Bd. 48)

Tomkins, Silvan S. (1962): *Affect imagery consciousness: Volume 1, The positive affects*. London, UK: Tavistock.

Trevarthen, Colwyn (2005): *Action and emotion in development of the human self, its sociability and cultural intelligence: Why infants have feelings like ours*. In: Nadel, J., Muir, D. (Eds.): *Emotional development* (61-91). Oxford: Oxford University Press.

Trojan, Felix (1975): *Biophonetik*. Mannheim/Wien/Zürich: B.I.-Wissenschaftsverlag.

Trubetzkoy, Nicolai S. (1989): *Grundzüge der Phonologie*. Vandenhoeck & Ruprecht Göttingen.

Uhmann, Susanne (1991): *Fokusphonologie. Eine Analyse deutscher Intonationskonturen im Rahmen der nicht-linearen Phonologie*. Tübingen: Niemeyer.

Uhmann, Susanne (1992): *Contextualizing Relevance: On Some Forms and Functions of Speech Rate Changes in Everyday Conversation*. In: Auer/di Luzio (eds.), S. 297-336.

Uhmann, Susanne (1996): On rhythm in everyday German conversation: beat clashes in assessment utterances. In: Couper-Kuhlen/Selting (eds.), S. 303-365.

Uhmann, Susanne (1997a): Grammatische Regeln und konversationelle Strategien. Fallstudien aus Syntax und Phonologie. Tübingen: Niemeyer.

Wendemuth, Andreas (2004): Grundlagen der stochastischen Sprachverarbeitung. Oldenbourg/München: Walter de Gruyter.

Welsch, Wolfgang (1997): Transkulturalität. Zur veränderten Verfassung heutiger Kulturen. In: Schneider, Irmela / Thomson, Christian W. (Hg.): Hybridkultur. Medien, Netze, Künste. Köln: Wienand, S. 67-90.

Wendt, Beate (2007): Analysen emotionaler Prosodie. Frankfurt am Main: Peter Lang Internationaler Verlag der Wissenschaften. (Hallesche Schriften zur Sprechwissenschaft und Phonetik Bd. 20)

Wiese, Richard (2011): Phonetik und Phonologie. Paderborn: Wilhelm Fink. (UTB 3354)

Williams, Carl E. / Stevens, Kenneth N. (1982): Akustische Korrelate diskreter Emotionen. In: Klaus R. Scherer (Hg.): Vokale Kommunikation. Nonverbale Aspekte des Sprachverhaltens. Weinheim/Basel: Beltz Verlag, S. 307-325.

Winter, Gerhard (1992): Was ist eigentlich eine kulturelle Überschneidungssituation. In: Thomas, Alexander (Hg.): Psychologie und multikulturelle Gesellschaft. Hogrefe, Göttingen, S. 221-227.

Wolff, Werner (1942): Zur Deutung der Sprechweise. Zeitschrift für Psychologie, 1942, 152.

Wunderlich, Dieter (1988): Der Ton macht die Melodie – Zur Phonologie der Intonation des Deutschen. In: Hans Altmann (Hg.): Intonationsforschungen. Tübingen, S. 1-40.

Wunderli, Peter [u. a.] (1978): Französische Intonationsforschung: kritische Bilanz und Versuch einer Synthese. Tübingen: Narr. (Tübinger Beiträge zur Linguistik 92)

Wundt, Wilhelm (1910): Grundzüge der physiologischen Psychologie. Bd. 2. 6. Aufl. Leipzig: Engelmann.

Wundt, Wilhelm (1903): Grundzüge der physiologischen Psychologie. Leipzig: Wilhelm Engelmann, Bd. 3.

Zheng, Xiaojun / Pierrehumbert, Janet B. (2010): The effects of prosodic prominence and serial position on duration perception. In: Acoustical Society of America, 128, S. 851-859.

Anhang

Vorstudie: Hörbeispiele

In der folgenden Übersicht sind die Links, Texte und Angabe der Minuten in Bezug auf die Hörbeispiele aufgelistet, die im Rahmen der Vorstudie als Audiomaterial aufbereitet und analysiert wurden.

Emotion Freude, Deutsch

Frau

1) 1:13 – 1:14

<https://www.youtube.com/watch?v=vdexRPVoOe8>

„Ja ja, ist immer schön“ (5 Wörter)

2) 1:56 – 1:59

<https://www.youtube.com/watch?v=LdfuFabQ2rs>

„Hoh, oh Gott ej, jetzt mach ich diesen Test“ (9 Wörter)

3) 2:10 – 2:12

<https://www.youtube.com/watch?v=LdfuFabQ2rs>

„Ich weiß gar nicht, ob das gut oder schlecht wär“ (10 Wörter)

4) 1:51 – 1:53

<https://www.youtube.com/watch?v=i2Wfs6dvwEc>

„Ich kann das bis heute kaum glauben“ (7 Wörter)

5) 8:23 – 8:25

<https://www.youtube.com/watch?v=i2Wfs6dvwEc>

„S´ist Wahnsinn, also ich glaub immer noch, dass ich träum“

6) 1:21 – 1:24

<https://www.youtube.com/watch?v=R06G8ZKEV68>

„Drei Tore bei diesem Spiel zu sehen, gigantisch“ (8 Wörter)

7) 1:24 – 1:26

<https://www.youtube.com/watch?v=R06G8ZKEV68>

„Gott sei Dank haben wir diese Karten bekommen“ (8 Wörter)

8) 0:58 – 1:01

https://www.youtube.com/watch?v=FQP6l_QTqps

„Als es dann vorbei war, hab i dacht Wahnsinn, Wahnsinn“ (10 Wörter)

9) 2:05 – 2:07

https://www.youtube.com/watch?v=FQP6l_QTqps

„Heit werma moi richtig gscheid feiern“ (6-7 Wörter)

10) 4:20 – 4:25

https://www.youtube.com/watch?v=FQP6l_QTqps

„Des is ähm für mich so so ein kleines persönliches Highlight irgendwie“ (11 Wörter ca)

Mann

1) 3:32 – 3:38

<https://www.youtube.com/watch?v=8aoHq3JWdBE>

„Das ist seine Theorie, das ist seine Theorie“ (8 Wörter)

2) 4:39 – 4:42

<https://www.youtube.com/watch?v=8aoHq3JWdBE>

„Du sprichst ja wie Professor Börne“ (6 Wörter)

3) 2:29 – 2:33

<https://www.youtube.com/watch?v=35bmWc7IUZ0>

„Geht ab heute, Kärntner-Platz, Glockenbachviertel“ (5 Wörter)

4) 0:15 - 0:18

<https://www.youtube.com/watch?v=oY7bEXd3Vek>

„Ja, ich glaube es ist unbeschreiblich“ (6 Wörter)

5) 0:18 – 0:23

<https://www.youtube.com/watch?v=Ahjz8ehTr7U>

„Mach ihn, mach ihn, er macht ihn, Mario Götze“ (9 Wörter)

6) 19:09 – 19:20

<https://www.youtube.com/watch?v=lZeXv138jDo>

„Tooor! Tor für Deutschland! Mario Götze macht das 1 zu 0!“ (11 Wörter)

7) 22:26 – 22:31

<https://www.youtube.com/watch?v=lZeXv138jDo>

„Jetzt ist es vollbracht, ja, ja, und nochmal ja!“ (9 Wörter)

8) 3:41 - 3:44

<https://www.youtube.com/watch?v=i2Wfs6dvwEc>

„Weil mir des Essen bei uns so gut schmeckt“ (9 Wörter)

9) 6:57 – 7:01

<https://www.youtube.com/watch?v=i2Wfs6dvwEc>

„es ist so ein Kinderpreis irgendwie, also für mich, also so eine Kindheitssache auch“ (14 Wörter)

10) 56:03 – 5:06

<https://www.youtube.com/watch?v=HhQGA2WzkZo>

„Die ham das super gemacht und schon ne tolle Leistung“ (10 Wörter)

11) 01:38:48 – 01:38:50

<https://www.youtube.com/watch?v=HhQGA2WzkZo>

„Wo habt ihr die Wahnsinnigen her, also des“ (8 Wörter)

Emotion Freude, Italienisch

Frau

1) 0:56 – 1:00

https://www.youtube.com/watch?v=QqDyJn_1qlQ

„Guarda, ci sono anche le m..., guarda!“ (ca. 7 Wörter)

2) 1:06 – 1:08

https://www.youtube.com/watch?v=QqDyJn_1qlQ

“Ma ci sono delle donne che sono ... come me” (9 Wörter)

3) 0:8 – 0:13

https://www.youtube.com/watch?v=a_b6eo7mT7Q

“Ma dai...tipo comunque...la squadra dolciosa questa” (8 Wörter)

4) 0:18-0:20

<https://www.youtube.com/watch?v=z0SicZIV-d4>

“Beh è un orgoglio anche mio farne parte” (8 Wörter)

5) 1:59 – 2:04

https://www.youtube.com/watch?v=W7TCe_-dpl0

“A sastri....bum bum bum” (5 Wörter)

6) 1:03 – 1:06

https://www.youtube.com/watch?v=MvK_sTfID9I

“Ma è buonissimo per davvero, e l'ho fatto io” (9 Wörter)

7) 1:20 – 1:24

<https://www.youtube.com/watch?v=vnlllJEw4fo>

“Lei è forte, lei ha dentro una maledizione” (8 Wörter)

8) 8:28 – 8:31

https://www.youtube.com/watch?v=N__QNEIG04U

„Ma non ci credo, oh Dio“ (6 Wörter)

9) 1:11 – 1:15

<https://www.youtube.com/watch?v=UNdz51wAoBM>

“no, non ci credo, ...tu” (5 Wörter)

Mann

1) 01:10 – 01:15

<https://www.youtube.com/watch?v=j8e5cMvScPw>

“Qua c'è la faccia mia, ma...com'è possibile, ma che dite” (12 Wörter)

2) 0:17 – 0:21

<https://www.youtube.com/watch?v=i0ichjoL4tA>

„Perchè Sanremo, ragazzi, è stancante dopo qualche giorno“ (8 Wörter)

3) 0:44 – 0:51

<https://www.youtube.com/watch?v=i0ichjoL4tA>

In generale, i miei testi, le mie canzoni vengono in realtà da una espressione spontanea.” (15 Wörter)

4) 2:26 – 2:28

<https://www.youtube.com/watch?v=iw5krHBaTVk>

“Balla.....liscio!” (2 Wörter)

5) 1:30 – 1:34

<https://www.youtube.com/watch?v=EzHBHzvo174>

“Abbracciamoci forte e vogliamoci tanto bene!” (6 Wörter)

Emotion Ärger, Deutsch

Mann

1) 5:28 – 5:31

https://www.youtube.com/watch?v=hO7GoB_TU9M

„dann würde ich jetzt auch wie alle andern am Hunger nagen“ (11 Wörter)

2) 6:31-6:33

https://www.youtube.com/watch?v=hO7GoB_TU9M

„meinen Sie so einfach kann man sich einigen, also“ (9 Wörter)

3) 2:24 – 2:26

<https://www.youtube.com/watch?v=cvBRFzk-rUA>

„hier auf dem Grundstück definitiv nicht“ (6 Wörter)

4) 3:54 – 3:56

<https://www.youtube.com/watch?v=cvBRFzk-rUA>

„also soll das scheiß Ding da etwa die ganze Zeit hier rumtuckern“ (12 Wörter)

5) 5:14 - 5:16

<https://www.youtube.com/watch?v=cvBRFzk-rUA>

„also ne, soviel können Sie mir gar nicht bezahlen“ (9 Wörter)

6) 6:03 – 6:05

<https://www.youtube.com/watch?v=5-Op6hMjLjE>

„Ej ich will meine Karre jetzt wieder haben“ (8 Wörter)

7) 8:21 – 8:26

<https://www.youtube.com/watch?v=TLbpiwiDKxw>

„Die brauchn jetzt net auspackn...des...also des...was soll des jetzt“ (10 Wörter)

Frau

1) 4:30 – 4:33

<https://www.youtube.com/watch?v=gHW1jItt7ww>

„Dann machens das bitte woandersch, weil äh mir müssen das genauso zahlä“ (12 Wörter)

2) 5:06 – 5:08

<https://www.youtube.com/watch?v=gHW1jItt7ww>

„Haben Sie mich verstanden, sowas gibt’s bei mir nicht“ (9 Wörter)

3) 7:28 – 7:30

<https://www.youtube.com/watch?v=UqINXtc91Xg>

„Ja, logisch, aber da ist ja auch ein Schild“ (9 Wörter)

4) 2:50 - 53

https://www.youtube.com/watch?v=_li3zB5qF4E

„Des isch ein wildfremder Mann, der hat sie jetzt angeleimt“ (10 Wörter)

5) 4:55 – 4:58

https://www.youtube.com/watch?v=_li3zB5qF4E

„Weil wir hier stehen seit etwa 10 Minuten“ (8 Wörter)

6) 5:43 – 5:46

https://www.youtube.com/watch?v=_li3zB5qF4E

„Wir sind seit 10 Stunden heute unterwegs“ (7 Wörter)

7) 4:18 – 4:20

<https://www.youtube.com/watch?v=N1GdhxXPT3k>

„Sie denken doch, dass wir jetzt Ihren Wohnwagen klauen wollen“ (10 Wörter)

8) 4:23 – 4:24

<https://www.youtube.com/watch?v=N1GdhxXPT3k>

„Ja das waren wir doch nicht“ (6 Wörter)

9) 4:25 – 4:28

<https://www.youtube.com/watch?v=N1GdhxXPT3k>

„Wir waren ja in der Wirtschaft drinnen, meine Güte ej“ (10 Wörter)

10) 4:29 – 4:31

<https://www.youtube.com/watch?v=N1GdhxXPT3k>

„Das waren wir nicht“ (4 Wörter)

Emotion Ärger, Italienisch

Frau

1) 1:48 – 1:52

<https://www.youtube.com/watch?v=Zr9Im8lhdFA>

„Ma no, ma mai, ve lo giuro per la mia madre che muore in questo momento.“ (16 Wörter)

2) 3:42 – 3:44

<https://www.youtube.com/watch?v=h10lvPTaUOQ>

„No, no, no non me ne frega un cazzo, vai vai (11 Wörter)

3) 0:12 – 0:15

https://www.youtube.com/watch?v=0SGP_T4irtY

“Cosa avevo detto? Domani torno a fare lezione.” (8 Wörter)

4) 2:03 – 2:05

https://www.youtube.com/watch?v=0SGP_T4irtY

“Ve lo sto dicendo anche” (5 Wörter)

5) 5:15 – 5:17

https://www.youtube.com/watch?v=0SGP_T4irtY

“Il problema è che ti sveglio”

6) 6:15 - 6:18

https://www.youtube.com/watch?v=0SGP_T4irtY

“Perché qua dentro voi vi dovete sudare qualcosa, non io” (10 Wörter)

7) 6:21 – 6:25

https://www.youtube.com/watch?v=0SGP_T4irtY

“E tutti giorni mi sveglio, ok, e me lo continuo a prendere.” (12 Wörter)

8) 3:53 – 3:57

https://www.youtube.com/watch?v=QL_Zj_nLnes

“Sono entrata in questo momento, fammi sistemare le cose” (9 Wörter)

Mann

1) 0:44 – 0:49

<https://www.youtube.com/watch?v=Zr9Im8lhdFA>

„Teo un uomo ignorante soltanto volgare se cel 'hai fatta lui possono farcela tutti” ?? (14 Wörter)

2) 3:16 – 3:18

<https://www.youtube.com/watch?v=MHgtnr5uRN0>

“Ma sembro io de cosa” (5 Wörter)

3) 3:47 - 3:52

<https://www.youtube.com/watch?v=MHgtnr5uRN0>

“Eh ho capito Toni ma guarde rimuovere ...dal momento che ...” (10 Wörter)

4) 4:11 – 4:14

<https://www.youtube.com/watch?v=MHgtnr5uRN0>

C'è quello sono io (ca 9 Wörter)

5) 5:59 – 6:01

<https://www.youtube.com/watch?v=MHgtnr5uRN0>

“nella mia vita non ho mai fatto qualcosa del genere” (10 Wörter)

6) 6:10 – 6:12

<https://www.youtube.com/watch?v=MHgtnr5uRN0>

“Mi procuri il filmato quello si vede che non sono io” (11 Wörter)

7) 4:18 – 4:21

https://www.youtube.com/watch?v=QL_Zj_nLnes

“Ti ho portato le rose....tu sei stata fuori a fare shopping” (12 Wörter)

8) 5:18 – 5:22

https://www.youtube.com/watch?v=QL_Zj_nLnes

“Chi è che messaggiava con un'altra persona davanti a me” (11 Wörter)

9) 8:30 – 8:32

https://www.youtube.com/watch?v=QL_Zj_nLnes

“Io ti ho lasciata libera tre settimane fa” (8 Wörter)

Tabelle 1: Auswertung Hörprobe 1, Vorstudie, deutsche Proband*innen

Anmerkung: gerade gedruckte Zahlen geben die Anzahl der Männer, kursiv gedruckte Zahlen die Anzahl der Frauen wieder, die die entsprechende Bewertung vorgenommen haben.

Dt.= Deutsche/r Sprecher/in

It.= Italienische/r Sprecher/in

W= weibliche Sprecherin

M= männlicher Sprecher

Fr.= Freude, d. h. freudige Sprechäußerung

Ärg.= Ärger, d. h. verärgerte Sprechäußerung

Nummern 1 – 68: Bezeichnung des jeweiligen Hörbeispiels

Zahlen von 1 – 5: Erkennungsgrad der jeweiligen emotionalen Ausprägung:

1= Emotion unklar; 2= Emotion wenig klar; 3= Emotion relativ klar; 4= Emotion klar;

5= Emotion sehr klar;

Bezeichnung Hörbsp. ↓	Erkennungsgrad der jeweiligen emotionalen Ausprägung (1-5)				
	1	2	3	4	5
1: dt., Fr., w	1		1 1	3 3	1
2: dt., Fr., w		3 2	1 1	2	1
3: dt., Fr., w	1 1	1 3	1 2	1	
4: dt., Fr., w		2 2	2 3		1
5: dt., Fr., w	1 1	3 2	1	1	1
6: dt., Fr., w	1	2	3	1 1	1
7: dt., Fr., w		2 3	1	2	2
8: dt., Fr., w	1	1	2 3	1 1	1
9: dt., Fr. w	1		1 1	1 2	2 2
10: dt., Fr., m	2	1	2	2 1	1 1
11: dt., Fr., m	3 5	1 1			
12: dt., Fr., m	1		2	2 4	1
13: dt., Fr., m		1 1	1 1	1 1	1 3
14: dt., Fr., m		2	1		4 3
15: dt., Fr., m		1		1	4 4
16: dt., Fr., m		1	1		4 4
17: dt., Fr., m	1 1	2 1	1 1	3	
18: dt., Fr., m		1 2	3 1	1	2
19: dt., Fr., m		2	2	3 2	1
20: dt., Fr., m	1 5	1 1	1	1	

21: dt., Ärg., w			1 1	1 3	2 2
22: dt., Ärg., w		1	1 1	2 2	3
23: dt., Ärg., w		2 1	1	1 1	1 3
24: dt., Ärg., w	2 3	1 1	2	1	
25: dt., Ärg., w	1	1 2	1	2 2	1
26: dt., Ärg., w		1	2 3	2	1 1
27: dt., Ärg., w		3 2	3	1	2
28: dt., Ärg., w			1	3 1	5
29: dt., Ärg., w			1	1 1	2 5
30: dt., Ärg., w				2 1	2 5
31: dt., Ärg., m	1	1 1	2 1	2	2
32: dt., Ärg., m	1	3 2	2	1	1
33: dt., Ärg., m		2	2 3	2	1
34: dt., Ärg., m		3 1	1	1 2	2
35: dt., Ärg., m		2	1	2	1 4
36: dt., Ärg., m			1	1	2 6
37: dt., Ärg., m	1 1	2	1 4		1
38: it., Fr., w	1 5	3	1		
39: it., Fr., w	5	2 1	2		
40: it., Fr., w		1	2	4 2	1
41: it., Fr., w	1	2 3	2	1	1
42: it., Fr., w	1 1	1 1		2 3	1
43: it., Fr., w	2 2	2 1	2		1
44: it., Fr., w	1 2	1 1	1 2	1 1	
45: it., Fr., w	1 4		1 1	1 1	1
46: it., Fr., w			1 1	1	3 4
47: it., Fr., m				2 1	2 5
48: it., Fr., m	2 1	2 3	1		1
49: it., Fr., m		1	3	1 3	2
50: it., Fr., m	2		1	1 1	2 3
51: it., Fr., m	2	1	1	1	3 2
52: it., Ärg., w		3 3	1 1	1	1
53: it., Ärg., w	1	1	1	1 4	2
54: it., Ärg., w	2 1	1 1	1 3	1	
55: it., Ärg., w	1 1	1 3	1 1	1	1
56: it., Ärg., w		1 1		2	1 5
57: it., Ärg., w			1		3 6
58: it., Ärg., w			1	1	2 6
59: it., Ärg., w	1			2 1	2 4
60: it., Ärg., m		3 3	1 2	1	
61: it., Ärg., m	1 1	2 1	2	1 1	1
62: it., Ärg., m			2 2	1 2	1 2
63: it., Ärg., m	1 1	1	1 2	1 4	
64: it., Ärg., m		3	2	1 1	3
65: it., Ärg., m		1	1 1	2 1	4
66: it., Ärg., m				2 1	2 5
67: it., Ärg., m			1	1 1	2 5
68: it., Ärg., m			1 1		3 5

Tabelle 2: Auswertung Hörprobe 2, Vorstudie, italienische Proband*innen

Anmerkung: gerade gedruckte Zahlen geben die Anzahl der Männer, kursiv gedruckte Zahlen die Anzahl der Frauen wieder, die die entsprechende Bewertung vorgenommen haben.

Dt.= Deutsche/r Sprecher/in

It.= Italienische/r Sprecher/in

W= weibliche Sprecherin

M= männlicher Sprecher

Fr.= Freude, d. h. freudige Sprechäußerung

Ärg.= Ärger, d. h. verärgerte Sprechäußerung

Nummern 1 – 68: Bezeichnung des jeweiligen Hörbeispiels

Zahlen von 1 – 5: Erkennungsgrad der jeweiligen emotionalen Ausprägung:

1= Emotion unklar; 2= Emotion wenig klar; 3= Emotion relativ klar; 4= Emotion klar;

5= Emotion sehr klar;

Bezeichnung Hörbsp. ↓	Erkennungsgrad der jeweiligen emotionalen Ausprägung (1-5)				
	1	2	3	4	5
1: dt., Fr., w		1	2 3	3 1	1
2: dt., Fr., w		1 1	1 1	3 3	1
3: dt., Fr., w		2	4 3	1 1	
4: dt., Fr., w	1 4	2 1	1 1	1	
5: dt., Fr., w	2 4	3 1	1		
6: dt., Fr., w		2 2	2	2 2	1
7: dt., Fr., w	1	1 4	4 1		
8: dt., Fr., w			1 2	3 2	2
9: dt., Fr., w	1 1	1	1	4 3	
10: dt., Fr., m				3 1	2 5
11: dt., Fr., m	3 4	1 1	1 1		
12: dt., Fr., m	2	2 1	1 2	2 1	
13: dt., Fr., m		1	3 3	2 1	1
14: dt., Fr., m					5 6
15: dt., Fr., m					5 6
16: dt., Fr., m				2 1	3 5
17: dt., Fr., m	1	2 1	1 2	2 3	
18: dt., Fr., m	3	2 1	2 2	1	
19: dt., Fr., m	1 4	1 1	1 1	1	1
20: dt., Fr., m	2 3	1 3	2		

21: dt., Ärg., w	1	4 5		1	
22: dt., Ärg., w	1 2	1 3	2 1	1	
23: dt., Ärg., w	1	1	1 4	2 1	1
24: dt., Ärg., w	2 2	2 2	1 2		
25: dt., Ärg., w			2 1	1 2	2 3
26: dt., Ärg., w		1 1	3 4	1	1
27: dt., Ärg., w		2 1	2	2 3	1
28: dt., Ärg., w				2 1	3 5
29: dt., Ärg., w			2	1 2	2 4
30: dt., Ärg., w				3 1	2 5
31: dt., Ärg., m	1	3 4	2 1		
32: dt., Ärg., m	1 3	2 2	2 1		
33: dt., Ärg., m	3 6	1		1	
34: dt., Ärg., m	1 4	3 2		1	
35: dt., Ärg., m		1 1	3 5	1	
36: dt., Ärg., m	1	1 1	1 1	1	2 3
37: dt., Ärg., m	3 5	1 1	1		
38: it., Fr., w	2	1 4	1 2	1	
39: it., Fr., w	1 3	3 2	1 1		
40: it., Fr., w			1 1	1 2	3 3
41: it., Fr., w	1	2 3	2	1 2	
42: it., Fr., w			1	3 2	1 4
43: it., Fr., w		1	1 1	2 1	2 3
44: it., Fr., w		1	1 1	3 1	1 3
45: it., Fr., w		1	1	2 1	2 4
46: it., Fr., w				1	4 6
47: it., Fr., m		1		3 2	2 3
48: it., Fr., m	1 1	2 4	2 1		
49: it., Fr., m		3 2	3	2 1	
50: it., Fr., m		1 1		1 1	3 4
51: it., Fr., m			1	1	4 5
52: it., Ärg., w	1	3 4	1	2	
53: it., Ärg., w		1 1	2	1	4 2
54: it., Ärg., w	1 2	1 3	2	1 1	
55: it., Ärg., w	2	1 2	3 1	1 1	
56: it., Ärg., w			2	1 2	2 4
57: it., Ärg., w			1		4 6
58: it., Ärg., w			1	1	3 6
59: it., Ärg., w					5 6
60: it., Ärg., m	1 2	1 1	1 3	2	
61: it., Ärg., m	1	2 2	1 3	1 1	
62: it., Ärg., m	1	1 1	2 1	1 2	1 1
63: it., Ärg., m	1	1 2	3 3		
64: it., Ärg., m		1 1	3 1	1 3	1
65: it., Ärg., m		1	2 2	2 2	2
66: it., Ärg., m		1	1		4 5
67: it., Ärg., m				1	4 6
68: it., Ärg., m			297	3	2 6

Fragebogen 1: Vorstudie, deutsche Proband*innen

Barbara Ladurner, a1004200, Universität Wien

Erste Hörprobe, Universität Regensburg, MA Sprechwissenschaft und Rhetorik

Projekt: Zur Perzeption emotionaler Sprechweisen durch prosodische Merkmale im Deutschen und Italienischen

Name (optional):

Geschlecht: m w

Geburtsort:

Alter:

Höchste abgeschlossene Ausbildung:

Beruf/Studium:

Muttersprache:

Andere Sprachkenntnisse:

1) DEUTSCH, FREUDE

Ihr hört jetzt kurze Audiosequenzen in **Deutsch**. Bitte gebt an, inwieweit diese der Emotion **“FREUDE”** zugeordnet werden können. Kreuzt bitte das entsprechende Kästchen an.

1=Emotion unklar; 2=Emotion wenig klar; 3=Emotion relativ klar; 4=Emotion klar; 5=Emotion sehr klar;

Sprecherin FRAU

Branco A	<input type="checkbox"/>				
Branco B	<input type="checkbox"/>				
Branco C	<input type="checkbox"/>				
Branco D	<input type="checkbox"/>				
Branco E	<input type="checkbox"/>				
Branco F	<input type="checkbox"/>				
Branco G	<input type="checkbox"/>				
Branco H	<input type="checkbox"/>				
Branco I	<input type="checkbox"/>				

 Sprecher MANN

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5
Brano H	1	2	3	4	5
Brano I	1	2	3	4	5
Brano J	1	2	3	4	5
Brano K	1	2	3	4	5

2) DEUTSCH, ÄRGER

Ihr hört jetzt kurze Audiosequenzen in **Deutsch**. Bitte gebt an, inwieweit diese der Emotion “**ÄRGER**” zugeordnet werden können. Kreuzt bitte das entsprechende Kästchen an.

1=Emotion unklar; 2=Emotion wenig klar; 3=Emotion relativ klar; 4=Emotion klar; 5=Emotion sehr klar;

 Sprecherin FRAU

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5
Brano H	1	2	3	4	5
Brano I	1	2	3	4	5
Brano J	1	2	3	4	5

 Sprecher MANN

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5

3) ITALIENISCH; FREUDE

Ihr hört jetzt kurze Audiosequenzen in **Italienisch**. Bitte gebt an, inwieweit diese der Emotion **“FREUDE”** zugeordnet werden können. Kreuzt bitte das entsprechende Kästchen an.

1=Emotion unklar; 2=Emotion wenig klar; 3=Emotion relativ klar; 4=Emotion klar; 5=Emotion sehr klar;

 Sprecherin FRAU

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5
Brano H	1	2	3	4	5
Brano I	1	2	3	4	5

 Sprecher MANN

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5

4) ITALIENISCH, ÄRGER

Ihr hört jetzt kurze Audiosequenzen in **Italienisch**. Bitte gebt an, inwieweit diese der Emotion “**ÄRGER**” zugeordnet werden können. Kreuzt bitte das entsprechende Kästchen an.

1=Emotion unklar; 2=Emotion wenig klar; 3=Emotion relativ klar; 4=Emotion klar; 5=Emotion sehr klar;

Sprecherin FRAU

Brano A	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano B	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano C	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano D	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano E	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano F	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano G	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano H	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>

Sprecher MANN

Brano A	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano B	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano C	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano D	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano E	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano F	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano G	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano H	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
Brano I	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>

Herzlichen Dank für Eure wertvolle Unterstützung!

Fragebogen 2: Vorstudie, italienische Proband*innen

Barbara Ladurner, a1004200, Università di Vienna

Prima prova di ascolto, Università Ca' Foscari di Venezia

Progetto: Perception of emotional speech through prosody in German and Italian

Nome (opzionale):

Sesso: m w

Luogo di nascita:

Età:

Titolo di studio:

Professione:

Madrelingua:

Altre lingue:

5) ITALIANO, GIOIA

Adesso sentirete dei brevi brani in **italiano** – indicate, per favore, se secondo voi rientrano più o meno nella categoria della “**GIOIA**”. Fate la crocetta a seconda del vostro parere.

1=emozione non chiara; 2=emozione poco chiara; 3=emozione abbastanza chiara; 4=emozione chiara; 5=emozione molto chiara

Parlante DONNA

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5
Brano H	1	2	3	4	5
Brano I	1	2	3	4	5

 Parlante UOMO

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5

6) ITALIANO, RABBIA

Adesso sentirete dei brevi brani in **italiano** – indicate, per favore, se secondo voi rientrano più o meno nella categoria della “**RABBIA**”. Fate la crocetta a seconda del vostro parere.

1=emozione non chiara; 2=emozione poco chiara; 3=emozione abbastanza chiara; 4=emozione chiara; 5=emozione molto chiara

 Parlante DONNA

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5
Brano H	1	2	3	4	5

 Parlante UOMO

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5
Brano H	1	2	3	4	5

Brano I 1 2 3 4 5

7) TEDESCO, GIOIA

Adesso sentirete dei brevi brani in **tedesco** – indicate, per favore, se secondo voi rientrano più o meno nella categoria della “**GIOIA**”. Fate la crocetta a seconda del vostro parere.

1=emozione non chiara; 2=emozione poco chiara; 3=emozione abbastanza chiara; 4=emozione chiara; 5=emozione molto chiara

Parlante DONNA

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5
Brano H	1	2	3	4	5
Brano I	1	2	3	4	5

Parlante UOMO

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5
Brano H	1	2	3	4	5
Brano I	1	2	3	4	5
Brano J	1	2	3	4	5
Brano K	1	2	3	4	5

8) TEDESCO, RABBIA

Adesso sentirete dei brevi brani in **tedesco** – indicate, per favore, se secondo voi rientrano più o meno nella categoria della “**RABBIA**”. Fate la crocetta a seconda del vostro parere.

1=emozione non chiara; 2=emozione poco chiara; 3=emozione abbastanza chiara; 4=emozione chiara; 5=emozione molto chiara

Parlante DONNA

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5
Brano H	1	2	3	4	5
Brano I	1	2	3	4	5
Brano J	1	2	3	4	5

Parlante UOMO

Brano A	1	2	3	4	5
Brano B	1	2	3	4	5
Brano C	1	2	3	4	5
Brano D	1	2	3	4	5
Brano E	1	2	3	4	5
Brano F	1	2	3	4	5
Brano G	1	2	3	4	5

Grazie tante per la Vostra preziosa partecipazione!

Fragebogen 3: Untersuchung zur Perzeption emotionaler Sprechweisen durch Prosodie – Deutsch

FRAGEBOGEN zum Thema „Perzeption emotionaler Sprechweisen durch prosodische Merkmale“, Hörbeispiel 1

Bitte wählen Sie eine Sprache aus. / Veuillez s'il vous plaît choisir une langue. / Please choose a language. / Selezionare una lingua. / Lütfen bir dil seçin. / Por favor, seleccione un idioma. / Por favor escolha um idioma. / Kies een taal. / Välj ett språk. / Vennligst velg et språk. / Vælg sprog. / Valitse kieli. / الرجاء اختيار لغتك.

- Deutsch
- Italiano

Einleitung

Beschreibung & Instruktion

In diesem Fragebogen bitte ich Sie, einige Fragen zu Ihrer Wahrnehmung eines ausgewählten Hörbeispiels zu beantworten. In meiner Dissertation im Fach Deutsche Philologie an der Universität Wien (Betreuer: ao. Prof. Dr. Peter Ernst) interessieren mich Zusammenhänge zwischen der Emotionsperzeption verschiedener Gruppen und prosodischer Merkmale.

Bitte nehmen Sie diesen Fragebogen zur Hand, wenn Sie mindestens eine halbe Stunde ungestört sind, und nehmen Sie sich ruhig Zeit für die Antworten. Sie können das Hörbeispiel so oft anhören, wie Sie wollen. Fast alle Fragen sollen Sie auf einer beschrifteten Skala beantworten.

Vielen Dank für Ihre wertvolle Unterstützung!

Datenschutzhinweis:

Dieser Fragebogen wurde erstellt von Barbara Ladurner. Alle Angaben in diesem Fragebogen werden anonym erfasst und verarbeitet. Die Daten werden ausschließlich genutzt für eine Dissertation im Fach Deutsche Philologie an der Universität Wien, sie werden nur von unmittelbar an der Studie beteiligten Personen genutzt und nicht an Dritte weitergegeben. Für weitere Informationen wenden Sie sich bitte an Barbara Ladurner, barbara.ladurner@univie.ac.at.

Angaben zur Person

1. Alter: * _____

2. Geschlecht: *

3. Nationalität: * _____

4. Muttersprache und aufgewachsen/sozialisiert in: *
 - Deutsch
 - Italienisch

5. Dialektsprecher: *
 - Ja
 - Nein

6. Beruf: *
 - Kein Stimmexperte
 - Stimmexperte (z. B. Sänger, Logopäde, Sprecherzieher)
 - Bitte genaue Berufsangabe: _____

Bitten hören Sie sich folgendes Hörbeispiel an: Deutsch, weibliche Sprecherin, Freude

Nehmen Sie sich Zeit und hören Sie gerne so oft Sie wollen das folgende Hörbeispiel an, um die nächsten Fragen zu beantworten. Es ist sicherlich hilfreich, die Sequenz immer wieder vor den einzelnen Fragen anzuhören. Etwaige Hintergrundgeräusche im Hörbeispiel können Sie einfach ignorieren.

Vielen Dank dafür, dass Sie durch Ihre akkuraten Antworten eine valide Auswertung ermöglichen!

<https://soundcloud.com/user-536792313/female-german-happyness/s-seOsW>

Analyse Hörbeispiel 1: FREUDE, weibliche Sprecherin, Deutsch

7. Woran erkennen Sie die dem Sprecher/der Sprecherin zugrundeliegende Emotion? Bitte die entsprechenden Kästchen ankreuzen, es können auch alle oder nur eines für Sie zutreffen. Beachten Sie die Erklärungen zu den einzelnen Wörtern im "Tooltip".* 

- Akzent/Rhythmus
- Tonhöhe/Melodieführung
- Text/Wortwahl
- Lautstärke
- Sprechgeschwindigkeit
- Pausen (Schweigen...)
- Stimmfärbung (Timbre)
- Weiß ich nicht
- Sonstiges: _____

8. Für wie wichtig halten Sie folgende Merkmale für die Emotionsbestimmung des Hörbeispiels?*

	Nicht wichtig	Eher unwichtig	Eher wichtig	Sehr wichtig
Akzent/Rhythmus	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tonhöhe/Melodieführung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Lautstärke	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sprechgeschwindigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pausen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Stimmfärbung (Timbre)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. Ausprägung der Emotion: Wie sehr freut sich die Sprecherin Ihrer Meinung nach? Bitte kreuzen Sie eines der Kästchen an.*

	Wenig					Sehr	
Geringe Ausprägung	<input type="radio"/>	Starke Ausprägung					

Wie schätzen Sie die Ausprägung der folgenden Merkmale im Hörbeispiel ein? Bitte entscheiden Sie sich für eine Antwort, indem Sie eines der vier Kästchen anklicken.

10. Akzent/Rhythmus *

	Eher schwache Betonung	Mittel	Eher starke Betonung	Weiß ich nicht
Akzentuierung (Stärke)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. Akzent/Rhythmus

	Eher wenige Betonungen	Mittel	Eher viele Betonungen	Weiß ich nicht
Akzentuierungen (Häufigkeit)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

12. Akzent/Rhythmus

	Eher abgehackt (staccato)	Mittel	Eher fließend (legato)	Weiß ich nicht
Sprechrhythmus	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie schätzen Sie die Ausprägung der folgenden Merkmale im Hörbeispiel ein?

13. Tonhöhe/Melodieführung

	Eher tief	Mittel	Eher hoch	Weiß ich nicht
Sprechstimmlage	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14. Tonhöhe/Melodieführung

	Gleichförmig (monoton)	Mittel	Bewegt (leb- hafte Melodie)	Weiß ich nicht
Tonhöhenverlauf	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

15. Tonhöhe/Melodieführung

	Klein (geringer Abstand zwi- schen höchstem und niedrigstem Ton)	Mittel	Groß (sehr hohe und sehr niedrige Töne)	Weiß ich nicht
Tonhöhenumfang (Umfang zwischen höchstem und nied- rigstem Ton)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie schätzen Sie die Ausprägung der folgenden Merkmale im Hörbeispiel ein?

16. Lautstärke

	Eher leise	Mittel	Eher laut	Weiß ich nicht
Lautstärke	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

17. Lautstärke

	Gleichförmig (e- her gleichblei- bende Laut- stärke)	Mittel	Bewegt (häufi- ges Auf und Ab zwischen laut und leise)	Weiß ich nicht
Lautheitsverlauf	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

18. Lautstärke

	Leiser werdend	Gleichbleibende Lautstärke	Lauter werdend	Weiß ich nicht
Lautheitsverlauf am Ende der Phrase	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie schätzen Sie die Ausprägung der folgenden Merkmale im Hörbeispiel ein?

19. Sprechgeschwindigkeit

	Eher langsam	Mittel	Eher schnell	Weiß ich nicht
Geschwindigkeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

20. Sprechgeschwindigkeit

	Gleichförmig (eher gleichbleibende Geschwindigkeit)	Mittel	Bewegt (häufige Tempowechsel)	Weiß ich nicht
Geschwindigkeitsverlauf	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

21. Sprechgeschwindigkeit

	Langsamer werdend	Gleichbleibende Geschwindigkeit	Schneller werdend	Weiß ich nicht
Geschwindigkeitsverlauf am Ende der Phrase	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie schätzen Sie die Ausprägung der folgenden Merkmale im Hörbeispiel ein?

22. Pausen

	Wenige Pausen	Mittel	Viele Pausen	Weiß ich nicht
Pausen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

23. Pausen

	Eher schwache Betonung	Mittel	Eher starke Betonung	Weiß ich nicht
„Geräuschvolle“ Pausen (mit Klang oder Geräuschen gefüllt, z. B. lautes Atmen, Seufzen, Lachen ect.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie schätzen Sie die Ausprägung der folgenden Merkmale im Hörbeispiel ein?

24. Stimmfärbung (Timbre)

	Dunkel	Mittel	Hell	Weiß ich nicht
Klangfarbe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

25. Stimmfärbung (Timbre)

	Klangarm	Klangvoll	Missklang (z. B. rau, Knarren...)	Weiß ich nicht
Klangfülle	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

26. Stimmfärbung (Timbre)

	Eng	Mittel	Weit	Weiß ich nicht
Faukale Distanz (Rachenweite)	O	O	O	O

27. Stimmfärbung (Timbre)

	Hart	Mittel	Weich	Weiß ich nicht
Stimmklang	O	O	O	O

28. Stimmfärbung (Timbre)

	Kraftstimme	Mittel	Schonstimme	Weiß ich nicht
Stimmansatz	O	O	O	O

29. Persönliche Eindrücke (optional): _____

Herzlichen Dank für Ihre wertvolle Mitarbeit!

Um über die weitere Auswertung und das Ergebnis informiert zu werden, wenden Sie sich gerne an barbara.ladurner@univie.ac.at. Danke und herzliche Grüße,

Barbara Ladurner

Fragebogen 4: Untersuchung zur Perzeption emotionaler Sprechweisen durch Prosodie – Italienisch

QUESTIONARIO relativo alla tematica: „Percezione del parlato emozionale attraverso la prosodia“, brano audio 1

Bitte wählen Sie eine Sprache aus. / Veuillez s'il vous plaît choisir une langue. / Please choose a language. / Selezionare una lingua. / Lütfen bir dil seçin. / Por favor, seleccione un idioma. / Por favor escolha um idioma. / Kies een taal. / Välj ett språk. / Vennligst velg et språk. / Vælg sprog. / Valitse kieli. / الرجاء اختيار لغتك.

- Deutsch
- Italiano

Introduzione

Descrizione & istruzioni

In questo questionario Le chiedo cortesemente di rispondere ad alcune domande che riguardano la Sua percezione di un brano audio. Queste informazioni servono per la mia tesi di dottorato in germanistica presso l'università di Vienna (il mio tutor è il professor Dr. Peter Ernst), in cui esamino il collegamento tra la percezione di emozioni di diversi gruppi di persone ed i parametri prosodici.

La prego di compilare il seguente questionario con calma quando ha almeno mezz'ora di tempo. Può ascoltare il brano audio quante volte vuole. La ringrazio sentitamente per il Suo prezioso sostegno!

Informativa sulla Privacy:

Questo questionario è stato compilato da Barbara Ladurner. Tutte le informazioni raccolte nel questionario vengono rilevate ed elaborate in modo anonimo e vengono utilizzate solo per una tesi di dottorato in germanistica all'università di Vienna e solo dalle persone strettamente coinvolte in questo progetto. I dati non saranno trasmessi a terzi. Per maggiori informazioni si rivolga a Barbara Ladurner, barbara.ladurner@univie.ac.at.

Dati personali

1. Et : * _____
2. Sesso: *
3. Cittadinanza: * _____
4. Lingua madre e socializzato in: *
 - o Tedesco
 - o Italiano
5. Parla un dialetto: *
 - o s 
 - o no
6. Professione: *
 - o Nessun esperto di voce
 - o Esperto di voce (ad es. cantante, logopedista, vocalcoach)
 - o Per favore specifichi la Sua professione: _____

Ascolti per favore il seguente brano audio: Tedesco, parlante donna, gioia

Si prenda del tempo e ascolti il brano quante volte ritiene necessario, per rispondere alle prossime domande.   sicuramente utile riascoltare il brano prima di rispondere alle singole domande. Pu  semplicemente ignorare eventuali rumori di sottofondo presenti nel brano.

La ringrazio, perch  con le Sue risposte accurate rende possibile un'analisi valida dei dati!

<https://soundcloud.com/user-536792313/female-german-happyness/s-seOsW>

Analisi brano audio 1: GIOIA, parlante donna, Tedesco

7. Da cosa riconosce l'emozione del parlante? Si prega di segnare l'apposita casella – anche più di una – con una croce. La prego di prendere visione delle spiegazioni delle parole nel “Tooltip” prima di procedere.* 

- Accento/ritmo
- Intonazione/altezza della voce
- Testo/scelta lessicale
- Volume
- Velocità di dizione
- Pause (silenzi...)
- Timbro di voce
- Non so
- Altro: _____

8. Quanto importante ritiene i seguenti parametri per l'individuazione dell'emozione?*

	Non importante	Poco importante	Piuttosto importante	Molto importante
Accento/ritmo	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Intonazione/altezza della voce	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Volume	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Velocità di dizione	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Pause	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Timbro di voce	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. Intensità dell'emozione: Secondo Lei, quanto è contento/a il/ la parlante? Si prega di mettere la crocetta su una delle caselle. *

	Poco					Molto	
Intensità bassa	<input type="radio"/>	Intensità alta					

Secondo Lei, com'è l'intensità dei seguenti parametri in questo brano audio? Per favore scelga una risposta e segni la relativa casella!

10. Accento/ritmo *

	Accentuazione piuttosto debole	Media	Accentuazione piuttosto intensa	Non so
Accentuazione (intensità)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. Accento/ritmo

	Poche accentuazioni	Medie	Molte accentuazioni	Non so
Accentuazione (intensità)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

12. Accento/ritmo

	Piuttosto stentato	Medio	Piuttosto fluido	Non so
Ritmo di dizione	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Secondo Lei, com'è l'intensità dei seguenti parametri in questo brano audio?

13. Intonazione/altezza della voce

	Piuttosto basso	Medio	Piuttosto alto	Non so
Registro vocale	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

14. Intonazione/altezza della voce

	Piuttosto uni- forme	Medio	Piuttosto mosso (melodia vivace)	Non so
Andamento (della melodia)	O	O	O	O

15. Intonazione/altezza della voce

	Piuttosto ridotta (piccolo intervallo tra il tono più alto e quello più basso)	Media	Piuttosto grande (toni molto alti e toni molto bassi)	Non so
Estensione tra i toni più alti ed i toni più bassi	O	O	O	O

Secondo Lei, com'è l'intensità dei seguenti parametri in questo brano audio?

16. Volume

	Piuttosto piano	Medio	Piuttosto forte	Non so
Volume	O	O	O	O

17. Volume

	Piuttosto uniforme	Medio	Piuttosto difforme (notevole alternanza tra toni alti e toni bassi)	Non so
L'andamento nella successione dell'intensità della voce	O	O	O	O

18. Volume

	Diventa più piano	Rimane costante	Diventa più forte	Non so
Andamento verso la fine della frase	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Secondo Lei, com'è l'intensità dei seguenti parametri in questo brano audio?

19. Velocità di dizione

	Piuttosto lenta	Media	Piuttosto veloce	Non so
Velocità	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

20. Velocità di dizione

	Piuttosto uniforme	Medio	Piuttosto mosso (frequenti cambi di velocità)	Non so
L'andamento della velocità di dizione	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

21. Velocità di dizione

	Diventa più lento	uniforme	Diventa più veloce	Non so
Andamento alla fine della frase	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Secondo Lei, com'è l'intensità dei seguenti parametri in questo brano audio?

22. Pause (silenzi)

	Poche	Normali	Molte	Non so
Pause	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

23. Pause (silenzi)

	Poche	Normali	Molte	Non so
“Pause piene” (colme di rumori, respiri intensi, di sospiri ecc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Secondo Lei, com'è l'intensità dei seguenti parametri in questo brano audio?

24. Timbro di voce

	Oscuro	Medio	Chiaro	Non so
Timbro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

25. Timbro di voce

	Poco sonoro	Molto sonoro	Dissonante (ad es. roco)	Non so
Sonorità	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

26. Timbro di voce

	Stretta	Media	Larga	Non so
Faringe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

27. Timbro di voce

	Duro	Medio	Dolce	Non so
Tono della voce	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

28. Timbro della voce

	Voce energica	Media	Voce riposante	Non so
Uso della voce	O	O	O	O

29. Impressioni personali (opzionale): _____

Grazie mille per la Sua preziosa analisi!

Se vuole rimanere informato/a del risultato dell'indagine La invito di contattarmi su barbara.ladurner@univie.ac.at.

Grazie, cordiali saluti,

Barbara Ladurner

Abstract

Die menschliche Kommunikation umfasst die Fähigkeit der Produktion und jene der Perzeption sprachlicher Informationen. Gerade in Bezug auf den Emotionsausdruck kommt dabei den prosodischen Merkmalen eine zentrale Bedeutung zu. In diversen Wissenschaftszweigen wurde der Zusammenhang zwischen Prosodie und emotionalen Sprechweisen verschiedentlich thematisiert, wobei v. a. die Fokussierung der Rezipient*innenperspektive und die Untersuchung natürlicher emotionaler Äußerungen noch viele Forschungsdesiderate aufweisen. Diese Aspekte wurden im Rahmen der vorliegenden empirischen Untersuchung aufgegriffen, in welcher anhand einer interkulturell angelegten Fragebogenerhebung zur Perzeption und Analyse freudiger und ärgerlicher Äußerungen folgende Themen bearbeitet werden konnten. Zum einen konnte herausgefunden werden, anhand welcher suprasegmentalen Eigenschaften diverse Emotionsausdrücke erkannt werden und welche Relevanz ihnen für den Perzeptionsprozess zukommt, zum anderen konnten die freudigen und ärgerlichen Sprechweisen anhand der prosodischen Merkmale auditiv von den Proband*innen analysiert und beschrieben werden. Hierbei konnten außerdem sprach-/kulturbedingte, berufsbezogene und geschlechtsspezifische Differenzen herausgearbeitet werden. Dazu wurden natürliche emotionale Äußerungen in den beiden Ausprägungen Freude und Ärger von deutsch- und italienischsprachigen Proband*innengruppen analysiert, die sich aus Männern und Frauen sowie Laien und Stimmexpert*innen auseinandersetzen. Es stellte sich heraus, dass die Faktoren Sprache/Kultur, Beruf und Geschlecht einen relevanten Einfluss auf die Perzeption emotionaler Sprechweisen der Proband*innen ausüben.

Human communication includes the ability to produce and perceive linguistic information. Especially when it comes to the expression of emotions, the prosodic characteristics are of central importance. In various scientific branches the connection between prosody and emotional ways of speaking has been discussed in different ways. The focus on the recipient's inner perspective and the investigation of natural emotional expressions left much room for further research. These aspects were taken up in the context of the present empirical study. On the basis of an interculturally designed questionnaire survey on the perception and analysis of joyful and angry utterances, the following topics could be worked on: On the one hand, it was possible to find out on the basis of which suprasegmental characteristics various emotional expressions are recognized and what relevance they have for the process of perception, on the other hand, the joyful and angry ways of speaking could be analysed and described by the test persons on the

basis of prosodic characteristics. It was also possible to identify language/cultural, occupational and gender-specific differences. For this purpose, natural emotional expressions of both joy and anger were analysed in German- and Italian-speaking groups of test persons, consisting of men and women as well as laypersons and voice experts. It turned out that the factors language/culture, profession and gender have a relevant influence on the perception of emotional ways of speaking.