

Inhaltsverzeichnis

11	Experimentalreihe 3: Perzeptionsuntersuchungen zu prosodischen Merkmalen in Monologisierungen	1
11.1	Experiment 3A: Experiment zur Perzeptionshierarchie prosodischer Merkmale.....	3
11.1.1	Fragestellungen zum Experiment.....	3
11.1.2	Hypothesen.....	4
11.1.3	Versuchsaufbau und Versuchsdurchführung	5
11.1.4	Untersuchungsergebnisse	8
11.1.5	Ergebnisinterpretation und Hypothesenbeantwortung.....	23
11.2	Experiment 3B: Experiment zur relativen Wahrnehmung prosodischer Merkmale..	30
11.2.1	Fragestellungen zum Experiment.....	30
11.2.2	Hypothesen.....	32
11.2.3	Versuchsaufbau und Versuchsdurchführung	33
11.2.4	Untersuchungsergebnisse	37
11.2.5	Ergebnisinterpretation und Hypothesenbeantwortung.....	45

11 Experimentalreihe 3: Perzeptionsuntersuchungen zu prosodischen Merkmalen in Monologisierungen

Mit der Experimentalreihe 3 wird der bisher eingeschlagene Untersuchungsweg fortgesetzt, indem zu gesprochenen Einheiten übergegangen wird, die sich über Äußerungslänge hinaus erstrecken. Im Zentrum dieser Untersuchungsreihe stehen monologische Gesprächssequenzen, die wiederum aus dem Untersuchungsmaterial des Telekommunikationskorpus des Seminars für Sprechwissenschaft der Universität Halle erstellt wurden. Auch hier bildete die Basis das Vorlesen der Allgemeinen Geschäftsbedingungen, wodurch sich eine nahezu identische Textgrundlage ergab, die durch verschiedene Sprecher realisiert wurde. Diese Sequenzen sind zwar in Gespräche eingebettet, formal sind dennoch grundlegende Kriterien einer *Rede* erfüllt. Im Folgenden werden diese Sequenzen als „Monologisierungen“ bezeichnet, da es sich um monologische Abschnitte innerhalb eines Dialogs handelt. Monologisierungen setzen sich aus einer Menge zusammenhängender Äußerungen eines Kommunikanten innerhalb eines Gesprächs zusammen, die inhaltlich miteinander verbunden sind, während der Gesprächspartner üblicherweise mittels Turnhaltesignalen reagiert. Im hier verwendeten Untersuchungsmaterial wurden die gelegentlichen Bestätigungssignale der Kunden – zumeist Interjektionen – entfernt. Bedingt durch die Länge derartiger Monologisierung muss der Perzipient verschiedene Mechanismen und Wissensbestände zur Sinnerschließung aktivieren. Insofern ist es wiederum naheliegend, dass bei der Perzeption prosodischer Merkmale von Monologisierungen Gesetzmäßigkeiten auftreten, die nicht allein durch Perzeptionsbefunde auf Äußerungsebene erklärbar sind.

Das verwendete Material weist eine Besonderheit auf: Beim Vorlesen der Allgemeinen Geschäftsbedingungen handelt es sich um eine Form reproduzierenden Textsprechens, die sich deutlich von frei vorgetragenen Beiträgen unterscheidet. Rothe untersuchte den hier verwendeten Korpus und stellte hinsichtlich des Gebrauchs prosodischer Merkmale teils erhebliche Unterschiede zwischen den beiden Modalitäten *Reproduzieren* und *Spontansprechen* fest. So ergab sich für das reproduzierende Sprechen ein größerer Tonhöhenumfang, eine gleich-

förmigere Tonhöhenkontur, primär mittelstarke Akzente, höhere Artikulationsraten, wenig Tempowechsel, sehr lange Sprechphrasen und nur wenige gliedernde Pausen. Frei formulierte Sequenzen zeigen hingegen variabelere Melodieverläufe und Akzentabstufungen, eine langsamere Sprechgeschwindigkeit mit kürzeren Phrasen sowie eine deutlichere Gliederung durch Pausen (2009, 119 f.). Der Wechsel zwischen beiden Modalitäten wird vom Perzipienten deutlich erkannt, jedoch kollidiert die reproduzierende Form, die ja im Gespräch wie „frei vorgetragen“ wirken soll, mitunter mit der Erwartungshaltung des Hörers. Aber gerade diese Nichterfüllung der Erwartungshaltung lenkt die Aufmerksamkeit auf prosodische Merkmale, die Gegenstand der hier vorgestellten Untersuchungsreihe ist. Insofern eignet sich dieses Material außerordentlich gut für Perzeptionsuntersuchungen zur Prosodie.

Im natürlichen Kommunikationsverlauf spielt die Sinnerfassung des Gesagten die entscheidende Rolle. Insofern kann nicht von der Sensitivität für bestimmte prosodische Merkmale ausgegangen werden, wie das beispielsweise in der Experimentalreihe 1 bei der Interjektion /hm/ bzw. /hmhm/ der Fall war. Daher muss im Folgenden die Wahrnehmung der Probanden auf bestimmte prosodische Merkmale gelenkt werden, um spezifische Perzeptionsresultate zu gewinnen. In den zwei Untersuchungen dieser Experimentalreihe wird überprüft, ob eine Perzeptionshierarchie prosodischer Merkmale angenommen werden kann und welche prosodischen Merkmale bezogen auf die Relativität der Wahrnehmung besonders beeinflussbar sind. Aus methodischer Sicht muss hier jedoch nochmals auf die Verschiedenheit der bewerteten Modalitäten verwiesen werden: Für die ursprünglich angerufenen Telefonkunden handelt es sich um Monologisierungen, wie diese eben eingeführt wurden. Für die Probanden, die die folgenden Perzeptionsexperimente durchlaufen haben, war der Sachverhalt des nach Leitfaden vorgetragenen Gesprächsteils bekannt. Insofern wurde in diesem Fall tatsächlich reproduziertes Textsprechen bewertet. Es wird jedoch davon ausgegangen, dass trotzdem hinreichend reproduzierbare Untersuchungsergebnisse zu erwarten sind, die sich auf die rhetorische Form einer „Rede“ anwenden lassen.

11.1 Experiment 3A: Experiment zur Perzeptionshierarchie prosodischer Merkmale

11.1.1 Fragestellungen zum Experiment

In der Untersuchungsreihe 2 wurde der Frage der *Wahrnehmungsgenauigkeit von globalen und innersilbischen Tonhöhenverläufen*, dem *Phänomen der subjektiven Dauerempfindung* in Äußerungen und der *Übereinstimmung bei der Perzeption verschiedener Akzentstufen* nachgegangen. Beim monologischen Sprechen kann die Empfindung der Dauer je nach Interesse eine wesentliche Rolle spielen. Bezüglich der Tonhöhenverläufe und Akzente dürften jedoch deren übergeordnete Merkmale wie z. B. *Stimmlage, Modulationsbreite, sprecherische Variabilität* neben primär inhaltlichen Kriterien entscheidend sein. Mit diesem Experiment wird außerdem das inhaltliche Verständnis des Gesagten untersucht, da sich aus der Wechselwirkung zur Erkennung prosodischer Merkmale interessante Erkenntnisse hinsichtlich der Perzeptionsontogenese ergeben könnten. Insofern kann in Bezug auf die Ausführungen im Theorieteil der Arbeit der Kreis der hier zutreffenden Modelle erweitert werden. Gerade unter ungünstigen akustischen Bedingungen fließt beim Sprachverstehen das Weltwissen im Sinne von *Top-down-Prozessen* (s. 5.2) maßgeblich ein. Da die Erkennung von Schlüsselbegriffen für das Gesamtverstehen von besonderer Bedeutung sind, kommen auch eine Kohortenbildung im Sinne der *COHORT-Modelle* (s. 5.2.1.9) bzw. bei „Hörfehlern“ der *Wortüberlegenheitseffekt* (s. 5.1.2) in Betracht. Auch der sprachpsychologische Ansatz Herrmanns (s. 5.2.2) kann in Beziehung zu diesem experimentellen Design gesetzt werden.

Mit dem Experiment zur Perzeptionshierarchie prosodischer Merkmale soll folgenden Fragen nachgegangen werden:

1. Welche prosodischen Merkmale werden am sichersten erkannt?
2. Wie genau werden die verschiedenen prosodischen Merkmale erkannt?
3. Welche Wortart wird unter akustisch ungünstigen Hörbedingungen am besten erkannt?
4. Welche Arten von „Hörfehlern“ können unter akustisch ungünstigen Hörbedingungen festgestellt werden?

5. Gibt es einen Zusammenhang zwischen erkanntem verbalen Material und dem subjektiv empfundenem Verständlichkeitsgrad des Gesagten unter akustisch ungünstigen Hörbedingungen?

Zur Untersuchung der Forschungsfragen wird im Folgenden der in der Akustik verbreitete Begriff *Signal-Rausch-Verhältnis* verwendet, welcher im Folgenden wie in der Akustik üblich mit *SRV* abgekürzt wird (vgl. Kapitel 6.3).

11.1.2 Hypothesen

Für den anstehenden Untersuchungsschritt wurden folgende Hypothesen postuliert:

Hypothese 1:

Prosodische Merkmale werden unter einem Maskierungssignal sicherer wahrgenommen als Lexeme.

Hypothese 2:

Mit zunehmendem Signalanteil gesprochener Sprache in einem SRV verändert sich die Qualität der Bewertung prosodischer Merkmale. Wird unter ungünstigen akustischen Bedingungen zunächst die Kategorie *mittel* erkannt, differenziert sich die Bewertung bei zunehmend günstigeren SRV hinsichtlich der tatsächlichen Ausprägung aus.

Hypothese 3:

Bevor in einem SRV prosodische Merkmale dezidiert erfasst werden, wird die Tatsache des Sprechens wahrgenommen.

Hypothese 4:

Unter einem Maskierungssignal verschiebt sich die Wahrnehmung zugunsten der Lexeme der offenen Klasse.

Hypothese 5:

Mit zunehmendem Alter der Probanden nimmt die Gewissheit des inhaltlichen Verstehens unter ungünstigem akustischen Bedingungen ab.

11.1.3 Versuchsaufbau und Versuchsdurchführung

Um die aufgestellten Hypothesen zu überprüfen, wurden wiederum Materialien aus dem Untersuchungskorpus zur Telekommunikation verwendet. Dabei wurden vier Gespräche ausgewählt, in denen Bestandskunden ein Vertrag zur Zusatzversicherung von Familienmitgliedern verkauft wurde, worauf die Allgemeinen Geschäftsbedingungen vorgelesen wurden. Unter Punkt 7.3 wurde für die Untersuchung prosodischer Merkmale eine Einheitlichkeit auf segmentaler Ebene favorisiert, um zu interpretierbaren Ergebnissen hinsichtlich der Perzeption prosodischer Merkmale zu gelangen. Andererseits würde durch eine völlige Konstanz der Textgrundlage die Erwartungshaltung der Probanden im Verlaufe der Untersuchung mehrerer Sprecher nachhaltig beeinflusst werden, was sich an präziseren Wahrnehmungsergebnissen am Ende einer Untersuchung zeigen könnte. Insofern erschien für diese Untersuchung eine geringe Variabilität des verbalen Materials bei an sich gleichbleibendem Inhalt günstig. Mit der Auswahl der vorgelesenen Allgemeinen Geschäftsbedingungen als Untersuchungsmaterial sollte beiden Aspekten Rechnung getragen werden: Es wurden ähnliche Abschnitte dieser Monologisierungen ausgewählt, wodurch die Texte inhaltlich vergleichbar sind, hinsichtlich des verbalen Materials jedoch nicht völlig identisch waren. Zur Untersuchung kamen die Monologisierungen zweier Sprecherinnen und zweier Sprecher. Rückfragen, Äußerungen und Bestätigungssignale der Kunden wurden entfernt.

Der genaue Wortlaut der vier Darbietungen ist im Anhang D wiedergegeben. Die auf diese Weise erstellten Audiodateien wurden pro Beitrag jeweils viermal präsentiert, wobei vom ersten bis zum vierten Untersuchungsdurchgang des jeweiligen Sprechers ein zusätzliches Maskierungsrauschen stufenweise reduziert wurde. Verwendet wurde dafür ein Bandpassrauschen zwischen 0,5 und 2 kHz (vgl. Kapitel 6.3). Während das Sprachsignal in allen Bedingungen konstant blieb, wurde das präsentierte Rauschen vom ersten bis zum vierten Durchgang jedes Sprechers von Null beginnend um jeweils 4 dB abgesenkt (0 dB; -4 dB; -8 dB; -12 dB). Damit sank das SRV stufenweise ab und das Sprachsignal wurde zunehmend deutlicher erkennbar.

Wahrnehmungsexperimente unter einem SRV erfordern besondere Aufmerksamkeit hinsichtlich der Untersuchungsdurchführung. Zum einen ist zu beachten, dass die Probanden ein normales Hörvermögen aufweisen, da bei beeinträchtigtem Hörvermögen die Filterprozesse des auditorischen Systems hinsichtlich des SRV anders wirksam sind. Zum anderen ist wichtig, dass die Probanden unter gleichen Bedingungen das Untersuchungsmaterial präsentiert bekommen. Für die Erstellung akustischer Testverfahren werden idealerweise Kopfhörer mit geschlossenem System verwendet. Nachteil dieses Verfahrens ist, dass nur mit erheblichem Aufwand eine hinreichende Anzahl an Probanden berücksichtigt werden kann. Um diesen beiden Aspekten Rechnung zu tragen wurde ein Mittelweg favorisiert: Zum einen wurden nur Versuchspersonen mit gesundem Hörvermögen berücksichtigt (vgl. Kapitel 7.3). Da die Probanden Studierende der Sprecherziehung der Universität Regensburg und Schüler der Staatlichen Logopädieschule Regensburg waren, lagen für deren Eignung zu den entsprechenden Studien- bzw. Ausbildungsgängen Audiogramme vor. Auch wenn deren Erstellungsdatum bei unauffälligem Befund im ungünstigsten Fall drei Jahre zurücklag, wurde das als hinreichendes Kriterium für die Eignung als Proband an dieser Untersuchung gewertet. Problematischer war zum anderen der Aspekt der konstanten akustischen Bedingungen. Aufgrund technischer Voraussetzungen war es nicht möglich, das Experiment unter Kopfhörerbedingungen durchzuführen. Jedoch wurde die Möglichkeit geschaffen, die Untersuchung unter gleichen räumlichen Bedingungen mit einem qualitativ hochwertigen Boxen-System zu realisieren, sodass der Abstand zwischen Schallquelle und Proband zwischen 2,0 und 2,5 m lag. Dafür wurden mehrere Lausprechersysteme des gleichen Typs parallel verwendet. Dieses Manko in der Durchführung wurde bewusst in Kauf genommen, da somit eine entsprechend hohe Anzahl an Probanden getestet werden konnte. Da es sich bei diesem Experiment nicht um die Normierung eines Testverfahrens handelte, sondern um die Frage, unter welchen SRV-Bedingungen sich prosodische Merkmale abbilden, wurde dieses Vorgehen als legitim betrachtet.

Die Untersuchung wurde mittels Bewertungsbogen durchgeführt, wobei auf jeder der vier Stufen für alle vier Darbietungen folgende Aspekte zu bewerten waren:

1. Wird gesprochen (im Folgenden „Sprechen?“)?
 - nein – wahrscheinlich – ja
2. Sprechergeschlecht:
 - weiblich – männlich – Kinderstimme
3. Artikulationsgeschwindigkeit:
 - fünfstufige Likert-Skala
 - sehr langsam bis sehr schnell
4. Sprechstimmlage:
 - fünfstufige Likert-Skala
 - sehr tief bis sehr hoch
5. Modulationsbreite:
 - dreistufige Likert-Skala
 - monoton bis sehr variabel
6. Wortverständlichkeit:
 - maximal fünf Wörter, die erkannt wurden
7. Textverständlichkeit:
 - weniger 50 % – zwischen 50 und 90 % – zwischen 90 und 100 %

Die ersten fünf Kriterien waren hinsichtlich der Qualität zusätzlich wie folgt einzuschätzen:

- nicht beurteilbar
- wahrscheinlich
- sicher

Der Bewertungsbogen der Untersuchung ist ausschnittsweise im Anhang D wiedergegeben. An der Untersuchung nahmen 76 Probanden teil (59 Studierende der Sprecherziehung und 17 Schülerinnen der Logopädie). Das Alter der 16 männlich und 60 weiblichen Probanden lag zwischen 19 und 54 Jahre (Medianwert: 23,5 Jahre).

11.1.4 Untersuchungsergebnisse

11.1.4.1 Ergebnisse zur Perzeptionshierarchie der untersuchten Merkmale

Im Vordergrund der Untersuchung stand zum einen die Frage, welche prosodischen Merkmale mit welcher Sicherheit auf welcher Stufe des SRV wahrgenommen werden. Die Sicherheit lässt sich allerdings nur indirekt über eine Befragung erschließen, als *wie sicher* der Proband sein Urteil einschätzt. Diese Grundproblematik wurde bereits unter Punkt 7.2 diskutiert. Jedoch ist hier nicht auszuschließen, dass eine Versuchsperson ein prosodisches Merkmal tatsächlich wahrnimmt, sich jedoch nicht sicher ist, ob dieses z. B. *mittelgradig* oder *hochgradig* ausgeprägt ist. Zwar wurde bei der Einführung des Experimentes auf die Intention der Untersuchung hingewiesen, tatsächlich kann jedoch nicht ausgeschlossen werden, dass das Unsicherheitsurteil in der qualitativen Wahrnehmung die eigentliche Intention der Untersuchung beeinträchtigt haben könnte.

Einen ersten Überblick über die Ergebnisse ermöglichen die Tabellen 1-4 im Anhang C. Hier werden die Items als Prozentwerte bezogen auf ein beurteiltes Merkmal dargestellt. Zur besseren Übersichtlichkeit werden die vier Darbietungen mit S1 bis S4 (S für Sprecher/in) in der Reihenfolge der Untersuchung angegeben. Die Spalten 2-5 berücksichtigen die vier Präsentationsstufen der Untersuchung, bei denen graduell das Hintergrundrauschen reduziert wurde. Zunächst fällt bei der Betrachtung der verschiedenen Ergebnisse auf, dass mit einer zunehmenden Reduktion des Geräuschanteils die Sicherheit für die Bewertungen zunimmt. Das betrifft die Fragestellung nach dem *Sprechen an sich*, dem *Sprechergeschlecht*, der *Artikulationsgeschwindigkeit*, der *Stimmlage*, der *Modulationsbreite*, dem *Grad des Textverstehens* und der *Zunahme der verstandenen und erfassten Lexeme*.

Dieses an sich wenig überraschende Ergebnis soll im Folgenden konkreter ausgeführt werden. Da mit den Perzeptionsgewohnheiten die Sicherheit der Wahrnehmungsstufe im Vordergrund steht, wurden die Fragestellungen mit den verschiedenen Graden der Wahrnehmungssicherheit gesondert berücksichtigt und wie folgt kodiert:

- 0: Merkmal nicht wahrgenommen bzw. *Textverständlichkeit* von weniger als 50 %
- 1: Merkmal wahrscheinlich bzw. *Textverständlichkeit* zwischen 50 % und 90 %
- 2: Merkmal sicher wahrgenommen bzw. *Textverständlichkeit* zwischen 90 % und 100 %

Diese Bewertungen, die jedem Sprecher zugeordnet wurden, sind in den Tabellen 5-8 des Anhangs C aufgeführt. Aufgrund der günstigeren Darstellung wurden die Werte einer bedingten Formatierung unterzogen. Daher erscheinen die sicheren Einschätzungen *grün* und die nicht wahrgenommenen Merkmale *rot* hinterlegt. Die Bewertungen wurden über das entsprechende Merkmal gemittelt. Ebenso wurden für die vier Wahrnehmungsstufen die Mittelwerte berechnet. Die absoluten Werte als gemittelttes Ergebnis über alle Probanden geben die Abbildungen 1-24 des Anhangs C wieder. Hier lässt sich deutlich erkennen, dass die *Tatsache des Sprechens* schon zwischen der Stufe 1 und 2 zu einem sicheren Wahrnehmungsergebnis führt und der Aspekt der *Sicherheit des Textverstehens* erst zwischen Stufe 3 und 4 sicherer wird. Die anderen vier Kriterien liegen zwischen diesen beiden Extremen.

	Sprechen?	Geschlecht	Geschwindigkeit	Stimmlage	Modulationsbreite	Textverstehen
Sprecher 1						
Summe Mittelwerte	149,25	135,25	120,75	117,75	111,75	37,75
Prozent	98,19%	88,98%	79,44%	77,47%	73,52%	24,84%
Rangplatz	1	2	3	4	5	6
Sprecherin 2						
Summe Mittelwerte	149,5	136,75	115,25	118,75	111,5	30,25
Prozent	98,36%	89,97%	75,82%	78,13%	73,36%	19,90%
Rangplatz	1	2	4	3	5	6
Sprecher 3						
Summe Mittelwerte	149,25	136	112,75	117	103,25	27,5
Prozent	98,19%	89,47%	74,18%	76,97%	67,93%	18,09%
Rangplatz	1	2	4	3	5	5
Sprecherin 4						
Summe Mittelwerte	150	119,75	109,75	112	100,25	21
Prozent	98,68%	78,78%	72,20%	73,68%	65,95%	13,82%
Rangplatz	1	2	4	3	5	6

Tab. 11-1: Übersicht über die gemittelten Häufigkeiten, Prozentwerte und Rangplätze aller vier Präsentationen

Die Ergebnisse der Bewertungen sind in Tabelle 11-1 zusammengefasst. Berücksichtigt sind hier die *Summe der Mittelwerte*, die entsprechenden *Prozentwerte* und der *Rangplatz der Erkennungsgenauigkeit* innerhalb einer Präsentation

on. Diese Ergebnisse werden nachfolgend grafisch in den Abbildungen 11-1 bis 11-4 wiedergegeben.

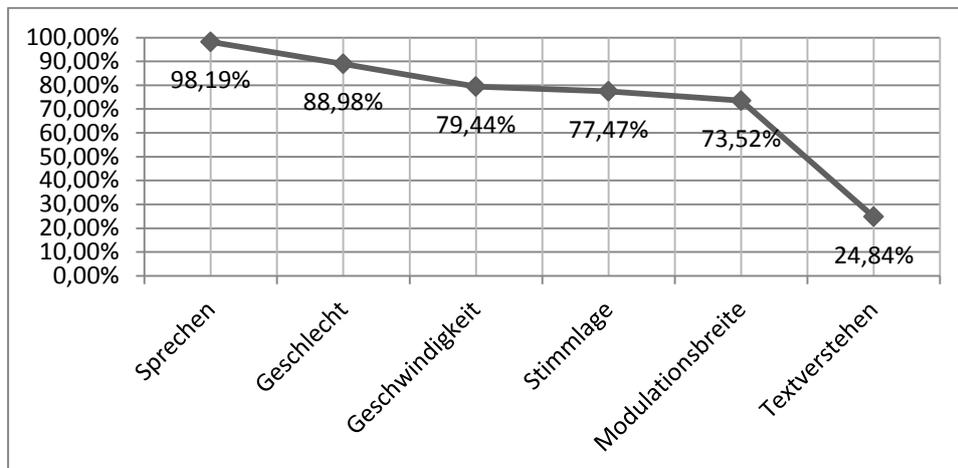


Abb. 11-1: Perzeptionssicherheit über die vier Präsentationsstufen des Sprechers 1

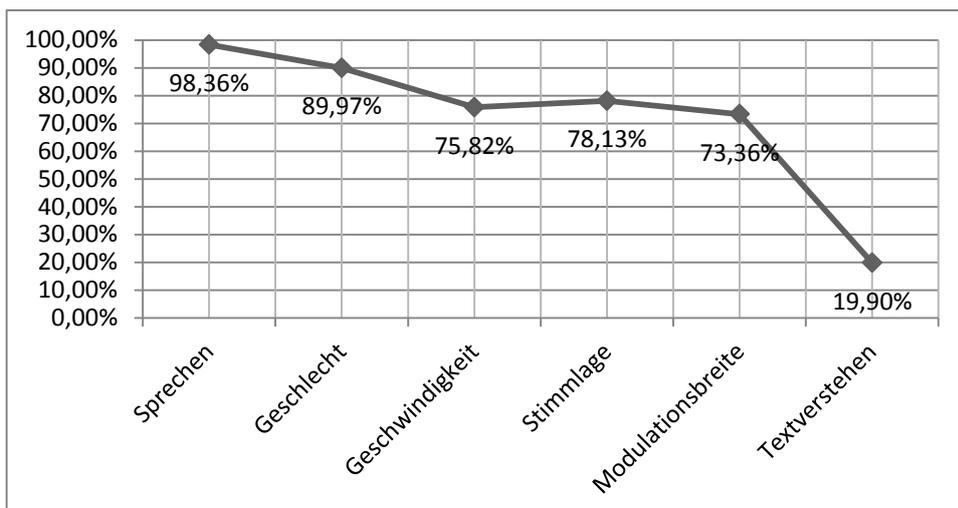


Abb. 11-2: Perzeptionssicherheit über die vier Präsentationsstufen der Sprecherin 2

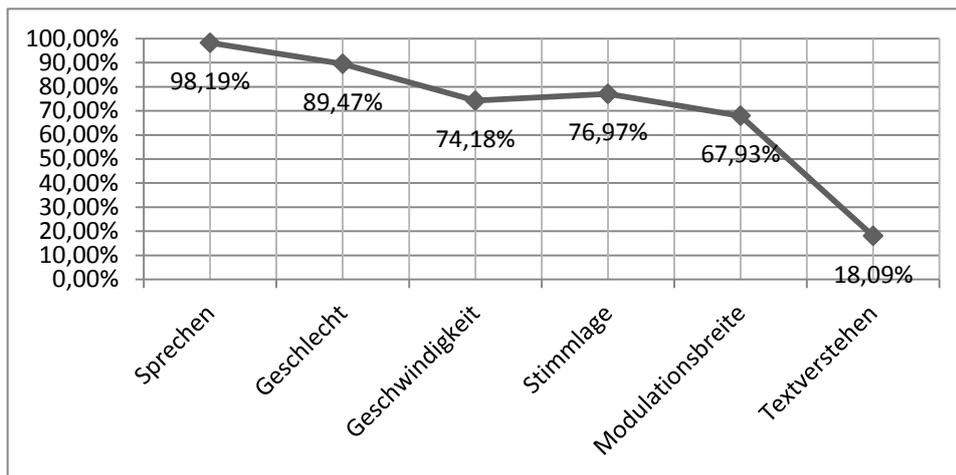


Abb. 11-3: Perzeptionssicherheit über die vier Präsentationsstufen des Sprechers 3

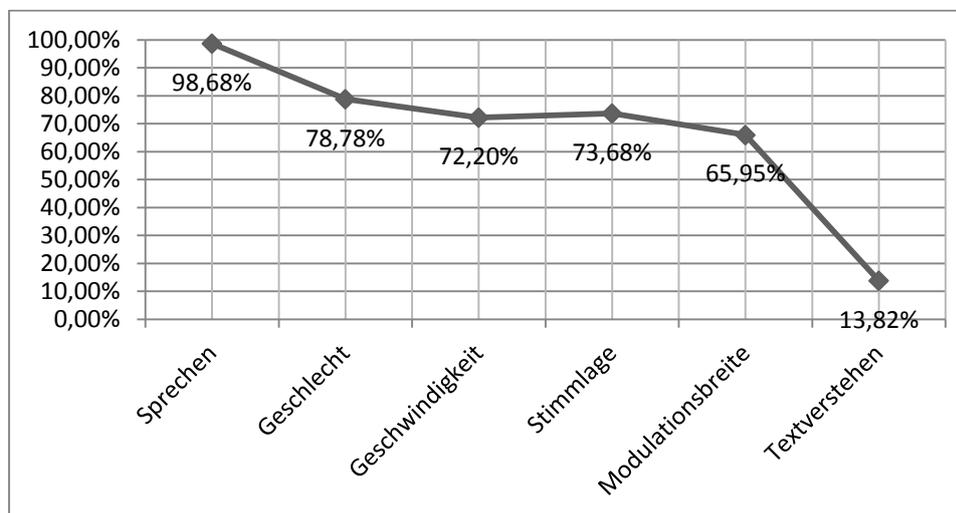


Abb. 11-4: Perzeptionssicherheit über die vier Präsentationsstufen der Sprecherin 4

Bei allen vier untersuchten Sprecherdarbietungen zeigt sich tendenziell ein gleichartiger Kurvenverlauf, dessen Mittelwerte, Prozente und Rangplätze in Tabelle 11-2 und Abbildung 11-5 (Mittelwerte) dargestellt sind.

	Sprechen?	Geschlecht	Geschwindigkeit	Stimmlage	Modulationsbreite	Textverstehen
Gesamtübersicht						
Summe Mittelwerte	598	527,75	458,5	465,5	426,75	116,5
Prozent	98,36%	86,80%	75,41%	76,56%	70,19%	19,16%
Rangplatz	1	2	4	3	5	6

Tab. 11-2: Gesamtübersicht über die gemittelten Häufigkeiten, Prozentwerte und Rangplätze

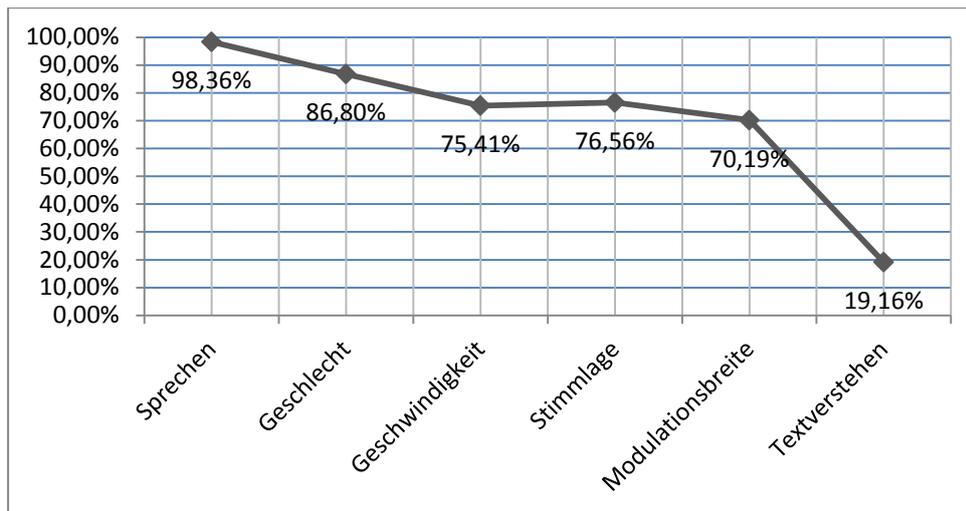


Abb. 11-5: Perzeptionssicherheit über die vier Präsentationsstufen aller gemittelten Darbietungen

Alle Abbildungen zeigen einen annähernd einheitlichen Kurvenverlauf auf. Das mit größter Sicherheit deutlich vor anderen erkannte Merkmal ist die *Tatsache des Sprechens* mit 98,36 % gefolgt von der *Sicherheit des Sprechergeschlechtes* mit 86,8 %. Unsicher ist die *Sicherheit des Textverstehens* mit 19,16 %. *Sprechgeschwindigkeit*, *Stimmlage* und *Modulationsbreite* liegen auf ungefähr gleichem Niveau und werden um ca. 10 % unsicherer eingeschätzt als das *Sprechergeschlecht*, aber deutlich höher als die *Sicherheit des Textverstehens*. Der gemittelte Verlauf gibt die Kontur wieder, die für alle untersuchten Sprecher ermittelt wurde (s. Abbildung 11-5). Auch die Rangplätze der Erkennungsgenauigkeit ist bis auf Sprecher 1 identisch. Während bei den Sprechern 2 bis 4 die *Stimmlage* etwas sicherer als die *Artikulationsgeschwindigkeit* erkannt wird, sind diese zwei Kriterien beim Sprecher 1 in umgekehrter Folge als *sicher* wahrgenommen worden. Tatsächlich liegen aber die Mittelwerte der *Artikulationsgeschwindigkeit* und *Stimmlage* so dicht beieinander, dass von einer annähernd identischen Erkennungssicherheit ausgegangen werden kann. Was anhand dieser Werte nicht zum Ausdruck kommt ist der Ausprägungsgrad des bewerteten Merkmals. Dieser soll im folgenden Abschnitt erörtert werden.

11.1.4.2 Ergebnisse der Perzeptionsgenauigkeit prosodischer Merkmale

Im Folgenden soll der Frage nachgegangen werden, wie sich die Qualitäten der Bewertungen über die verschiedenen Stufen des SRV verändern. Im Mittelpunkt

stehen die drei Kriterien *Artikulationsgeschwindigkeit*, *Sprechstimmlage* und *Modulationsbreite*. Dazu wird zum einen auf die Veränderung der Kriterien in der Bewertung der Gesamtgruppe eingegangen, in einem zweiten Schritt wird die Beurteilung dreier phonetisch geschulter Experten als Referenzgröße hinzugezogen.

Die Prozentwerte der Bewertungen finden sich in den Tabellen 1-4 des Anhangs C, die grafischen Darstellungen diese Werte finden sich in den Abbildungen 25-36 des Anhangs C. Hierbei waren die Angaben für *nicht beurteilbar* mit zu berücksichtigen. Es zeigen sich folgende Veränderungen über die vier Stufen:

Artikulationsgeschwindigkeit (bezieht sich auf die reine Artikulationsrate, d. h. es erfolgt keine Berücksichtigung der Pausierung):

- Sprecher 1:
 - *mittel* und *schnell* bei Stufe 1 und 4 etwa gleich hoch
 - Stufe 2 und 3: *schnell* stärker, *mittel* schwächer
- Sprecherin 2:
 - Stufe 1: *mittel* als stärkste Kategorie, die stetig abnimmt
 - *schnell* nimmt stetig zu, verliert jedoch auf Stufe 4
- Sprecher 3:
 - mit Reduktion für *nicht beurteilbar* steigt das Merkmal *schnell* stetig an
 - ab Stufe 2 relativ stabile Ausprägung
- Sprecherin 4:
 - *schnell* von Anfang an stärkste Ausprägung, die bis Stufe 3 zunimmt
 - insgesamt stabil ab Stufe 2

Sprechstimmlage:

- Sprecher 1:
 - *mittel* von Anfang an stärkste Kategorie, sinkt ab Stufe 2
 - *hoch* nimmt stetig zu
 - *tief* nimmt leicht ab

- Sprecherin 2:
 - *hoch* von Anfang an stärkste Kategorie, sinkt zwischen Stufe 3 und 4 zugunsten von *sehr hoch*
 - *mittel* relativ stabil
- Sprecher 3:
 - *tief* von Anfang an stärkste Kategorie, nimmt ab Stufe 2 leicht ab
 - *mittel* nimmt leicht zu
 - insgesamt stabile Form
- Sprecherin 4:
 - *hoch* von Anfang an stärkste Kategorie, nimmt eher zu
 - *mittel* nimmt leicht zu
 - insgesamt stabile Form

Modulationsbreite:

- Sprecher 1:
 - *mittel* nimmt stetig zu
 - *sehr variabel* mit starkem Anstieg, sinkt ab Stufe 2 stetig ab
 - *nicht beurteilbar* anfangs sehr häufig
- Sprecherin 2:
 - *mittel* nimmt stetig zu
 - *monoton* und *sehr variabel* nehmen etwas zu
 - *nicht beurteilbar* anfangs sehr häufig
- Sprecher 3:
 - *mittel* nimmt stetig zu
 - *sehr variabel* nimmt stetig zu
 - *nicht beurteilbar* anfangs sehr häufig und bis Stufe 4 auftretend
- Sprecherin 4:
 - *mittel* nimmt stetig zu
 - *sehr variabel* nimmt stetig zu
 - *nicht beurteilbar* anfangs sehr häufig

Bei der Betrachtung dieser qualitativen Ergebnisse ist zu berücksichtigen, dass mit einer stetigen Abnahme von *nicht beurteilbar* andere Bewertungen zunehmen. In Anbetracht dessen zeigen sich jedoch folgende allgemeine Tendenzen:

- **Artikulationsgeschwindigkeit:** Es zeigen sich dezente Merkmalsverschiebungen. Eine generelle Dominanz von *mittel*, die als ein Zeichen von Bewertungsunsicherheit angesehen werden könnte, kann nicht festgestellt werden.
- **Sprechstimmlage:** Es fallen die relativ stabilen Verteilungen über alle Stufen auf, die kaum nennenswerte Veränderungen erfahren.
- **Modulationsbreite:** Für dieses Merkmal ist auffällig, dass die Kategorie *nicht beurteilbar* sehr häufig bis teilweise zur Stufe 4 vermerkt ist. Das hängt nicht nur mit dem SRV zusammen, sondern mit der allgemeinen Unsicherheit bei der Beurteilung dieses Merkmals. Insofern ist es auch nicht verwunderlich, dass es mit zunehmender Stufe zu einer Merkmalsverschiebung zugunsten von *mittel* kommt. Diese kann bei diesem Kriterium durchaus als ein Zeichen von Bewertungsunsicherheit angesehen werden.

Die prosodischen Merkmale der vier Sprecher wurden von drei phonetisch geschulten Experten eingeschätzt, wobei deren Beurteilung ohne ein SRV erfolgte. Die Ergebnisse finden sich in Tabelle 11-3.

	Sprecher 1	Sprecherin 2	Sprecher 3	Sprecherin 4
Artikulationsgeschwindigkeit	mittel	mittel	schnell	mittel
Sprechstimmlage	mittel	mittel	mittel	mittel
Modulationsbreite	sehr variabel	mittel	mittel	mittel

Tab. 11-3: Experteneinschätzung der Sprechermerkmale

Es zeigt sich demnach, dass die Merkmalsausprägung *mittel* bei allen drei Kriterien stets überwiegt. Tabelle 11-4 setzt die Bewertungen der Probanden über alle vier Stufen in Abhängigkeit zur Experteneinschätzung. Die Zahlen in der Tabelle beziehen sich dabei auf die Häufigkeitsprozente aus den Tabellen 1-4

des Anhangs C. Im Fettdruck hervorgehoben sind dabei die von der Versuchspersonengruppe jeweils am häufigsten bewertete Kategorie.

Artikulationsgeschwindigkeit					
Experten:		Stufe 1	Stufe 2	Stufe 3	Stufe 4
mittel	Sprecher 1	42,11 %	36,84 %	35,53 %	44,74 %
mittel	Sprecherin 2	39,47 %	43,42 %	30,26 %	34,21 %
schnell	Sprecher 3	14,47 %	42,11 %	47,37 %	51,30 %
mittel	Sprecherin 4	21,05 %	18,42 %	23,68 %	23,68 %
Sprechstimmlage					
Experten:		Stufe 1	Stufe 2	Stufe 3	Stufe 4
mittel	Sprecher 1	46,05 %	61,84 %	51,32 %	48,68 %
mittel	Sprecherin 2	38,16 %	32,89 %	28,95 %	34,21 %
mittel	Sprecher 3	15,79 %	17,11 %	25,00 %	28,95 %
mittel	Sprecherin 4	17,11 %	19,74 %	23,68 %	25,00 %
Modulationsbreite					
Experten:		Stufe 1	Stufe 2	Stufe 3	Stufe 4
sehr variabel	Sprecher 1	38,16 %	48,68 %	46,05 %	39,47 %
mittel	Sprecherin 2	42,11 %	48,68 %	48,68 %	53,95 %
mittel	Sprecher 3	46,05 %	48,68 %	48,68 %	47,37 %
mittel	Sprecherin 4	34,21 %	43,42 %	47,37 %	51,32 %

Tab. 11-4: Prozentuale Übereinstimmung der Sprechermerkmale mit der jeweiligen von den Experten bewerteten Kategorie

Für die Interpretation der Tabelle 11-4 ist zu berücksichtigen, dass die Experten die Bewertung unter anderen Bedingungen als die Versuchspersonengruppe vorgenommen haben. Jedoch ist ersichtlich, dass sich mit zunehmender Präsentationsstufe die Probanden tendenziell der Bewertung der Experten angenähert haben. Deutlich ist das bei der Bewertung der *Artikulationsgeschwindigkeit* von Sprecher 3 und Sprecherin 4, der *Sprechstimmlage* von Sprecher 3 und Sprecherin 4 sowie der *Modulationsbreite* von Sprecherin 2 und Sprecherin 4 zu erkennen. Natürlich gibt es auch einige Sprünge über eine benachbarte Stufe wie z. B. bei der *Modulationsbreite* des Sprechers 1.

11.1.4.3 Ergebnisse der Perzeptionsgenauigkeit des verbalen Materials

Im Rahmen dieses Experiments war für die Sprachperzeption die anzunehmende Abhängigkeit von *verbalen* zu *prosodischen Merkmalen* von besonderem Interesse. Die Probanden wurden daher gebeten, auf allen Stufen jeder Sprecherpräsentation bis maximal fünf erkannte Wörter anzugeben. Hierbei wurden bewusst

keine Vorgaben gemacht, ob es sich um Wörter der *geschlossenen* oder *offenen* Klasse handeln sollte. Erwartungsgemäß wurde mit der Reduktion des SRV auch die Anzahl der Einträge bei allen Probanden höher, wobei einige Probanden schon auf einer sehr frühen Stufe verbales Material (vermeintlich) erkannten, bei anderen hingegen wurden erst relativ spät Einträge vorgenommen. Ein Nichteintrag ist natürlich kein Nachweis, dass nicht doch Wörter erkannt wurden. Sehr häufig wurden auch Syntagmen angegeben. In der Tabelle 11-5 sind getrennt nach allen vier Sprecherdarbietungen die jeweils am häufigsten zwölf erkannten Wörter angegeben.

Sprecher 1		Sprecherin 2		Sprecher 3		Sprecherin 4	
Informationen	76	kostenlos	81	kostenlos	76	Studenten	74
Unterlagen	67	Unterlagen	81	Unterlagen	57	Versicherungsschutz	73
Änderungen	53	jederzeit	40	EU	43	jederzeit	33
Mitarbeit	41	Versicherungsschutz	39	Versicherungsschutz	38	Kinder	32
12 Euro 95	31	Versicherung	35	Ländern	32	bitte	30
vielen Dank	26	angekommen	26	jederzeit	31	EU	26
Konto	23	eintreffen	25	angekommen	25	kostenlos	22
Ablauf	21	Tagen	21	einverstanden	24	Ablauf	21
rechtliche	18	bitte	20	verschiedene	21	Inhalt	17
nach Ablauf	17	wenigen	19	Genehmigung	20	Konto	17
verschiedenen	17	Ausschlüsse	18	60 Tage	19	monatliche	17
eingezogen	15	in wenigen Tagen	14	Nordamerika	18	Unterlagen	17

Tab. 11-5: Absolute Häufigkeiten für die zwölf am häufigsten genannten Wörter, separiert nach Sprechern und summiert über alle vier Darbietungsstufen

Bestimmte Wörter wurden demnach von vielen Probanden gleichermaßen erkannt, wobei es sich fast ausschließlich um Repräsentanten der *offenen* Klasse handelt. Einige Begriffe wurden besonders häufig vermerkt wie z. B. *kostenlos* und *Unterlagen* der Sprecherin 2 oder *Studenten* und *Versicherungsschutz* der Sprecherin 4. Da es sich um das Vorlesen der Allgemeinen Geschäftsbedingungen handelte sind natürlich auch einige erkannte Wörter bei mehreren Sprechern gleichermaßen anzutreffen. Dazu gehören z. B. *Unterlagen*, *kostenlos*, *Versicherungsschutz* und *jederzeit*.

Die zwölf am häufigsten erkannten Wörter der *geschlossenen* Klasse und deren absolute Häufigkeit summiert über alle vier Darbietungen sind folgende:

ihnen/ihr/ihrer 18
nicht 16

sie	16
wenn	14
nicht	14
sind	11
mit	11
sind	7
auch	6
in	6
sofort	6
und	6

Im Verhältnis zur Gesamtzahl der Einträge ergeben sich folgende Werte für die Wörter der *geschlossenen* vs. *offenen* Klasse (s. Tabelle 11-6):

	Geschlossene Klasse	Offene Klasse
Sprecher 1	12,10 %	87,90 %
Sprecherin 2	11,67 %	88,33 %
Sprecher 3	8,38 %	91,62 %
Sprecherin 4	7,65 %.	92,35 %

Tab. 11-6: Gegenüberstellung der erkannten Wörter der geschlossenen vs. offenen Wortklasse

Es ist deutlich erkennbar, dass die *geschlossene* Klasse bei weitem nicht an die Häufigkeit der offenen heranreicht. Die am häufigsten erkannten Redewendungen waren:

Vielen Dank

Vielen Dank für ihre Mitarbeit

Bitte lesen Sie

Man kann davon ausgehen, dass derartige Redewendungen bzw. Floskeln in der Wahrnehmung sehr stabil sind und daher auch unter ungünstigen akustischen Bedingungen relativ gut erkannt werden.

Zwangsläufig ergaben sich auch *Hörfehler*, d. h. die vermerkten Lexeme traten in dieser Form nicht in den Monologisierungen auf. Um die Art dieser Fehler zu quantifizieren, wurden alle Einträge mit der Textvorlage verglichen, die fehlerhaften Stellen vermerkt und klassifiziert. Eine Übersicht über die Anzahl der

Einträge und den Prozentsatz der darin enthaltenen Hörfehler ermöglicht Tabelle 11-7.

	Sprecher 1	Sprecherin 2	Sprecher 3	Sprecherin 4
Anzahl der Einträge	860	882	883	797
Fehlbenennungen	115	145	131	109
Fehlbenennungen (%)	13,37 %	16,44 %	14,84 %	13,68 %

Tab. 11-7: Häufigkeiten für Einträge perzipierter Wörter und Fehlbenennungen

Als Mittelwert über alle Sprecherdarbietungen aller vier Stufen ergibt sich ein Fehleranteil von 14,58 % aller Einträge. Nach der Art der Fehlbenennungen lassen sich folgende Fehlertypen ausmachen:

- *phonematisch* zu interpretierende Fehler: 44,44 %
- *morphologisch* zu interpretierende Fehler: 22,22 %
- *semantisch* zu interpretierende Fehler: 25,25 %
- *nicht klassifizierbare* Fehler: 8,08 %

Der hohe Anteil phonematischer Fehler überrascht nicht, da sich unter ungünstigen akustischen Bedingungen die Wahrnehmung für Laute nachhaltig verschlechtert. Auf diese Weise entstehen primär durch Phonemverwechslungen klangähnliche Wörter. Erstaunlich ist hingegen, dass in dem Bemühen um die Erkennung von Lexemen phonematische Fehler entstanden, die bei einem internen Monitoring hinsichtlich der Sinnhaftigkeit der Äußerung durch Korrekturmechanismen sicher behoben worden wären. Zum Beispiel ist der Begriff *Installation* statt dem tatsächlich vorhandenem *Information* im vorgegebenen Kontext der Allgemeinen Geschäftsbedingungen nicht sinntragend und wäre unter natürlichen Bedingungen mit der Intention der Sinnerschließung wahrscheinlich bemerkt worden.

Ein weiterer Aspekt scheint hinsichtlich der Wahrnehmungsfehler bemerkenswert: Es wurden teilweise Angaben bei Wörtern gemacht, die keinerlei Beziehung zum vorgegebenen Textmaterial aufwiesen, aber aus der Situation – sei sie biografisch oder auch bezogen auf die Untersuchung – erklärbar sein könnten. Beispielsweise wurde in einem Fall aus der Subgruppe der Logopädieschülerinnen das Wort *Klinikum* erkannt. Tatsächlich ist die Arbeit am Universitäts-

klinikum für die Logopädieschülerinnen ein wesentlicher Inhalt der Ausbildung und insofern individuell sehr bedeutsam. Eine vergleichbare Tatsache ist, dass die Studierenden wesentlich häufiger die Begriffe *Fachschule* oder *Fachhochschule* richtigerweise erkannt haben, während dies bei den Subgruppen der Logopädieschülerinnen und des Weiterbildungsmasters in dieser Ausprägung nicht der Fall war. Gerade bei der letzten Gruppe, die sich aus Teilnehmern zusammensetzt, die fest in Familie und Beruf verankert sind, zeigte sich dafür häufiger die Erkennung von *Versicherungspolice*. Ein interessanter Fall ergab sich auch bei der Präsentation der Sprecherin 4: Diese wurde, wie im Anhang C Tabelle 4 zu erkennen, in den ersten Wahrnehmungsstufen gelegentlich mit einer Kinderstimme verwechselt. Tatsächlich hat hier eine Versuchsperson das Wort *Hausaufgabe* wahrgenommen. Das lässt sich evtl. damit erklären, dass Hausaufgaben besonders mit Schulkindern in Verbindung gebracht werden.

11.1.4.4 Aussagen zur Perzeptionsgenauigkeit von Untergruppen

Mit der folgenden Auswertung der Ergebnisse soll nach Gruppenfaktoren gesucht werden, um spezifische Aussagen zur Perzeptionsfähigkeit zu erhalten. Ein interessantes Maß ist die Gegenüberstellung der *mittleren Fehlerquotienten* und der *Gewissheit des Textverstehens*. Der Fehlerquotient ergibt sich aus der Anzahl der gesamten Bewertungen im Verhältnis zu den Hörfehlern bezüglich der Erkennung von Wörtern. Die Gewissheit des Textverstehens findet sich in den Tabellen 5-8 des Anhangs C. Eine *Gewissheit des Textverstehens* mit einem Relativwert von 1,0 entsteht, wenn ein Proband schon ab der ersten Stufe eine Stufe des Textverstehens von über 90 % angegeben hätte. Die Gegenüberstellung dieser zwei Größen könnte als „Risikofreudigkeit“ der Probanden beim Erkennen von Wörtern interpretiert werden, die in Bezug zur Selbsteinschätzung hinsichtlich der verstandenen Inhalte der Texte gesetzt werden. Die Mittelwerte einiger Untergruppen sind in Tabelle 11-8 aufgeführt.

	Mittlerer Fehlerquotient	Gewissheit Textverstehen gemittelt
Frauen gemittelt	13,70%	0,41
Männer gemittelt	15,51%	0,29
Altersgruppe bis 25	10,90%	0,43
Altersgruppe ab 25	20,98%	0,28
Logopäden-Schülerinnen	10,69%	0,47
Direktstudierende	9,74%	0,45
Weiterbildungsmaster	30,67%	0,08

Tab. 11-8: Mittlere Fehlerquotienten und gemittelte Gewissheit des Textverstehens der Subgruppen

Schon bei der Gegenüberstellung der *Frauen* und *Männer* zeigt sich eine Überraschung. Während der mittlere Fehlerquotient durchaus vergleichbar ist, sind die männlichen Probanden häufiger der Meinung, dass sie den Inhalt weniger erfasst hätten. Ebenso überraschend ist die Einteilung hinsichtlich der zwei Altersgruppen. Der Fehlerquotient der jüngeren Probanden ist deutlich niedriger. Außerdem sind sich die Probanden dieser Gruppe darüber hinaus wesentlich sicherer, dass sie den Inhalt als solchen erfasst haben. Schließlich seien die drei Subgruppen der *Logopädieschülerinnen*, *Direktstudierenden* und der *Angehörigen des Weiterbildungsmasters* betrachtet: Die Logopädieschülerinnen und Direktstudierenden unterscheiden sich hinsichtlich Fehlerquotient und Gewissheitsgrad praktisch nicht. Deutliche Abweichungen zeigen sich bei den Teilnehmern des Weiterbildungsmasters. Zum einen liegt hier der mittlere Fehlerquotient nahezu dreimal höher, andererseits zeigt die Gewissheit des Textverstehens im Mittelwert an, dass bis zur Stufe 4 eine starke Unsicherheit hinsichtlich der Erfassung des Inhalts besteht. Dieses schlägt sich auch in der Bewertung der Altersgruppe *ab 26 Jahren* nieder, in der die Angehörigen des Weiterbildungsmasters zumeist liegen. Da untersuchungsmethodische Aspekte als Fehlerquelle ausscheiden, müssen andere Gründe angenommen werden. Bemerkenswert erscheinen diese Befunde dahingehend, dass jüngere und beruflich bzw. familiär noch nicht so stark gebundene Menschen wesentlich risikofreudiger und gleichzeitig auch voreingenommener sein könnten. Anhand der Alltagserfahrung lässt sich eine derartige Beziehung durchaus nachvollziehen.

Die Beziehung zwischen beiden Größen lässt sich auch grafisch im Streudiagramm darstellen (s. Abbildung 11-6). Hier wird der *Gewissheitsgrad des Textverstehens* in Abhängigkeit zum *mittleren Fehlerquotienten* gesetzt, wobei die

Ziffern die einzelnen Probanden repräsentieren. Zusätzlich ist für die Gesamtgruppe die Trendlinie eingezeichnet.

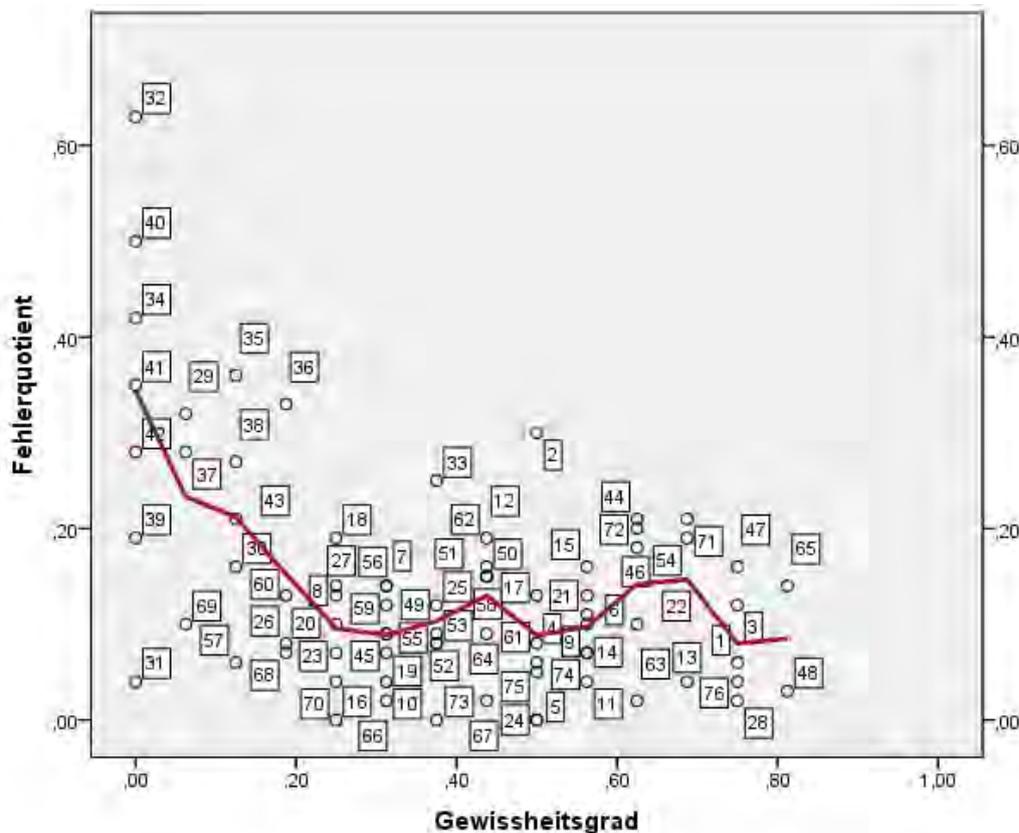


Abb. 11-6: Streudiagramm des Gewissheitsgrades des Textverstehens gegenüber dem mittleren Fehlerquotienten

Anhand der Punktwolke lässt sich ablesen, dass sich zum einen Probanden hinsichtlich der inhaltlichen Texterfassung unsicher sind, andererseits jedoch sehr wenig Hörfehler machen. Der gegenteilige Trend, d. h. ein hoher Gewissheitsgrad des Verstehens gegenüber sehr vielen Hörfehlern ist hingegen nicht zu verzeichnen. Die Trendlinie weist am Anfang einen stark fallenden Charakter auf, um dann mit Ausnahme zweier leichter Erhebungen annähernd waagrecht zu verlaufen. Im ersten Teil der fallenden Gerade befindet sich die Mehrzahl der Subgruppe des Weiterbildungsmasters. Die Mehrzahl der Bewertungen liegt im Bereich des Gewissheitsgrades zwischen 0,4 und 0,6 bei einem Fehlerquotienten von ca. 10 %.

Schließlich sei auf die Frage nach der Sicherheit des Erkennens für die verschiedenen Subgruppen eingegangen (s. Tabelle 11-9). Berücksichtigt sind hier die gemittelten Prozentwerte über alle vier Darbietungen aller Sprecher.

	Sprechen?	Geschlecht	Geschwindigkeit	Stimmlage	Modulationsbreite	Textverstehen
Frauen gemittelt	98,91%	87,45%	76,93%	77,24%	71,98%	20,47%
Männer gemittelt	96,29%	84,38%	69,73%	74,02%	63,48%	14,26%
Altersgruppe bis 25	99,04%	89,24%	76,80%	77,34%	71,75%	21,57%
Altersgruppe ab 26	96,88%	81,51%	72,40%	74,87%	66,74%	13,93%
Logopäden-Schülerinnen	98,90%	89,89%	79,78%	82,17%	78,68%	23,71%
Direktstudierende	99,08%	88,57%	75,78%	76,70%	69,46%	22,59%
Weiterbildungsmaster	95,63%	78,13%	69,38%	69,79%	62,71%	3,96%

Tab. 11-9: Übersicht über die gemittelten Prozentwerte aller vier Präsentationen nach Untergruppen

Interessant ist hierbei zunächst, dass die Rangfolge der Sicherheit, wie sie oben bestimmt wurde, in allen Fällen erhalten bleibt. Die *Sicherheit des Sprechens* steht vor der Erkennungssicherheit des *Geschlechtes*, während *Artikulationsgeschwindigkeit*, *Stimmlage* und *Modulationsbreite* ein breites Mittelfeld bilden. In allen Fällen ist die *Gewissheit des Textverstehens* deutlich schlechter ausgeprägt. Überraschend ist jedoch, dass auch in der Gruppe der geringsten allgemeinen Sicherheiten (Weiterbildungsmaster) mit 95,63 % eine sehr hohe Gewissheit hinsichtlich des Sprechens auftritt, während die Sicherheit des Sprechergeschlechts mit einer Differenz von 17,5 % schlechter ausfällt. Die geringste Differenz hinsichtlich dieser zwei Werte ergibt sich bei den Logopädieschülerinnen mit 9,01 %. Alle anderen Werte der einzelnen Gruppen sind durchaus vergleichbar, wenn man die Differenz zur jeweils vorherigen Spalte als Referenzwert annimmt. Insgesamt liegen mit einem Gesamtdurchschnitt aller Sicherheitsurteile mit 75,52 % die Logopädieschülerinnen etwas vor der Gruppe der Direktstudierenden (72,03 %), während sich die Gruppe des Weiterbildungsmasters mit einem Mittelwert von 63,27 % am wenigsten sicher ist.

11.1.5 Ergebnisinterpretation und Hypothesenbeantwortung

Die *Tatsache des Sprechens* an sich wird mit sehr großer Sicherheit auch bei starkem Maskierungssignal erkannt. Die Sicherheit des *Sprechergeschlechtes*, die *Artikulationsgeschwindigkeit*, die *Stimmlage* und *Modulationsbreite* folgen dicht aufeinander. Lediglich die Sicherheit der *Textverständlichkeit* wird deutlich schlechter eingeschätzt. Aufgrund der Ergebnisse dieser Untersuchung kann man schlussfolgern, dass prosodische Merkmale mit deutlich größerer Sicherheit wahrgenommen werden als Elemente der verbalen Ebene.

In **Hypothese 1** wird vermutet, dass prosodische Merkmale unter einem Signal-Rausch-Verhältnis sicherer als Lexeme wahrgenommen werden. **Hypothese 3** besagt, dass prosodische Merkmale unter einem Signal-Rausch-Verhältnis später erkannt werden als die *Tatsache des Sprechens*. Einen Eindruck hinsichtlich dieser Fragestellung gewinnt man aus Tabelle 11-2 und Abbildung 11-5, wo die Mittelwerte über die Beurteilungen aller vier Sprecher berücksichtigt sind. Hier zeigt sich eindeutig eine große Differenz zum einen zwischen der Erkennung der *Modulationsbreite* und der Gewissheit des *Textverstehens* (Hypothese 1) und zwischen der *Tatsache des Sprechens* und der Sicherheit der Bewertung des Sprechergeschlechts (Hypothese 3).

Zur statistischen Überprüfung dieser Fragestellung bietet sich der Wilcoxon-Vorzeichen-Rangsummentest an. Testet man die jeweils benachbarten Kategorien, kann außer zwischen der Sicherheit der Bewertung der *Stimmlage* und *Artikulationsgeschwindigkeit* die Nullhypothese abgelehnt werden, wonach zwischen benachbarten Kategorien der Medianwert der Unterschiede gleich Null ist. Lediglich zwischen der Sicherheit der Bewertung der *Stimmlage* und *Artikulationsgeschwindigkeit* errechnet sich ein Korrelationskoeffizient bei zweiseitiger Fragestellung von ,154. Berücksichtigt man die Sprecherdarbietungen im Einzelnen, so finden sich die Ergebnisse der statistischen Berechnung auch hier (Ausnahme ist die Sprecherin 3, bei der sich die Sicherheit der Bewertung der *Stimmlage* und *Artikulationsgeschwindigkeit* knapp unterscheiden). Insofern können die Hypothesen 1 und 3 angenommen werden. Darüber hinaus kann verallgemeinert werden, dass hinsichtlich der Sicherheit der Wahrnehmung folgende Rangfolge angenommen werden kann:

1. *Tatsache des Sprechens*
2. *Erkennung des Sprechergeschlechts*
3. *Erkennung der Stimmlage + der Artikulationsgeschwindigkeit*
4. *Erkennung der Modulationsbreite*
5. *Erkennung verbalen Materials*

Bemerkenswert ist die Konstanz der Rangfolge dieser Merkmale, die sich auch für alle Subgruppen in dieser Form gezeigt hat (außer in einem Fall des Sprechers 1, bei dem die *Artikulationsgeschwindigkeit* vor der *Stimmlage* er-

kannt wurde). Weiterhin liegen bezogen auf die die Mittelwerte aller vier Sprecherdarbietungen zwischen der Tatsache des *Erkennens des Sprechens* und der des *Sprechergeschlechtes* 11,65 %. Dieser große Differenzwert verweist auf eine „Grauzone“, in der hinsichtlich der Sprechermerkmale offenbar keine spezifischen Urteile möglich sind. Zwangsläufig ergibt sich die Frage, worauf sich eine derartige Sicherheit gründet, obwohl nichts Konkretes zu Sprechermerkmalen geschweige denn zum Inhalt gesagt werden kann. Schon Humboldt postulierte einen für den Menschen charakteristischen Sinn, den er den *Sprachsinn* nannte (1836, CCCXIII f.). Dieser an sich bisher nur in der Sprachphilosophie thematisierte Sinn gewinnt unter Berücksichtigung die hier vorliegenden Ergebnisse eine neue, am ehesten psychologisch zu begründende Dimension. Näheres dazu wird in der Abschlussdiskussion (Kapitel 13) ausgeführt.

Eine weitere interessante Tatsache ergibt sich aus der Qualität der Bewertungen der prosodischen Merkmale. Hier hat sich gezeigt, dass die *Stimmlage* von Sprechern schon unter ungünstigen SRV ziemlich stabil perzipiert wird. Ebenso erfährt die Bewertung der *Artikulationsrate* nur geringe Veränderungen, während hinsichtlich der *Modulationsbreite* bis zu einem günstigeren SRV Unsicherheiten bestehen. Die Tatsache der Unsicherheit in der Bewertung dieser Merkmale muss zum einen bei den Hörern gesucht werden, wenn sie hinsichtlich der formseitigen Bewertung von Sprecherleistungen ungeübt sind. Tatsächlich haben alle Probanden dieser Untersuchung an Seminaren teilgenommen, die die Diagnostik von Stimme und Sprechweise zum Inhalt hatten. Insofern kann nicht ausgeschlossen werden, dass die Bewertung der *Stimmlage* tatsächlich auch unter ungünstigen akustischen Bedingungen einfacher ist als die der *Artikulationsrate* und besonders der *Modulationsbreite*.

Mit **Hypothese 2** wurde postuliert, dass mit der Abnahme des Hintergrundrauschens die Bewertung der prosodischen Merkmale der Sprechermerkmale zunehmend aus einer mittleren (Unbestimmtheits-)Kategorie zur tatsächlichen Ausprägung tendiert. Als Bezugsgröße wurde dabei die Bewertung der Experten zugrunde gelegt. Die Antwort auf diese Hypothese gestaltet sich dennoch problematisch. Zum einen hätten sich die Versuchspersonen hinsichtlich der Bewertung relativ sicher sein müssen. Jedoch ist zu vermuten, dass in der allgemeinen

Unsicherheit der Bewertung von Sprechermerkmalen – namentlich der Modulationsbreite – auch die Rubrik *wahrscheinlich* markiert wurde, ohne dass dies auf die Unsicherheit hinsichtlich der Bewertung aus dem SRV erfolgte. Zum anderen hätten für eine derartige Fragestellung die *Artikulationsraten*, *Stimmlagen* und *Modulationsbreiten* der Sprecherinnen und Sprecher deutliche extremere Ausprägungen aufweisen müssen. Jedoch wiesen bis auf zwei Ausnahmen alle Sprecher *mittelgradige* Ausprägungen der untersuchten prosodischen Merkmale nach den Einschätzungen der Experten auf. Insofern ist die Ableitung nicht zulässig bzw. möglich, wonach von einer unter ungünstigem SRV beginnenden *mittleren* Ausprägungsstufe der Bewertungen zu den für die Sprecher spezifischen Merkmalen eine Annäherung stattfindet. Somit kann aber anhand dieser Untersuchung nicht konstatiert werden, dass unter ungünstigen Präsentationsbedingungen eines SRV eine allgemeine *mittlere* Bewertungsstufe der Sprechermerkmale auftritt, die als Zeichen von Bewertungsunsicherheiten angesehen werden kann. In Anbetracht der nicht hinreichend kontrollierten Bedingungen muss die Hypothese 2 demnach unbeantwortet bleiben.

In **Hypothese 4** wurde postuliert, dass unter ungünstigen akustischen Bedingungen der Wahrnehmungsfokus der auf die sinntragenden Wörter der *offenen* Klasse ausgerichtet ist. Hinsichtlich der zu erkennenden Lexeme wurden den Versuchspersonen keine Vorgaben gemacht. Dass primär Wörter der *offenen* Klasse angegeben wurden, beweist, dass die Wahrnehmung gesprochener Sprache immer auf einen Sinn gerichtet ist. Tatsächlich sind die Wörter der *geschlossenen* Klasse viel kürzer und wären damit weniger aufwändig zu erkennen und zu notieren. Interessant erscheint der Sachverhalt, dass beim ersten Sprecher noch 12,1 % Wörter der *geschlossenen* Klasse angegeben wurden. Bei über die vier Stufen stetig fallendem Anteil waren es bei Sprecherin 4 nur noch 7,65 %. Das könnte ein Beleg dafür sein, dass mit zunehmendem Bekanntheitsgrad der vorgelesenen Allgemeinen Geschäftsbedingungen die Fokussierung auf deren Inhalte zunahm. Ebenfalls häufig wurden Floskeln vermerkt, die in einem ungünstigen SRV offensichtlich einen sehr hohen Wiedererkennungswert haben und daher auch in der Wahrnehmung sehr stabil sind. Viele Probanden haben statt einzelner Wörter Syntagmen angegeben, was ebenso dafür spricht, dass die menschliche Wahrnehmung nicht auf Einzelwörter ausgerichtet ist, sondern auf

die Sinn-Erschließung. Diese ist nicht zwangsläufig an Wörter gebunden, obwohl die Wörter der offenen Klasse dafür prädestiniert erscheinen. Darüber hinaus wird davon ausgegangen, dass Redewendungen, die auch sehr häufig im Bewertungsbogen vermerkt wurden, als Ganzheit rechtshemisphärisch verarbeitet werden. Redewendungen dieser Art sind z. B. auch bei aphasischen Patienten relativ gut stimulierbar und werden auch unter ungünstigen akustischen Bedingungen sicher wahrgenommen. In Übereinstimmung zur Hypothese kann davon ausgegangen werden, dass unter ungünstigen akustischen Bedingungen der Fokus der Wahrnehmung auf die sinntragenden Wörter der offenen Klasse ausgerichtet ist. Als Mittelwert aller vier Sprecher ergibt sich ein Anteil von 90,05 % an Wörtern dieser Klasse. Hypothese 4 gilt demnach als verifiziert.

Der *Fehlerquotient* erkannter Wörter im Verhältnis zur *Gewissheit des Textverstehens* ergab vor allem zwei interessante Ergebnisse. Während der *Fehlerquotient* zwischen den Geschlechtern nahezu identisch ausfiel, war die *Gewissheit des Verstehens* bei den weiblichen Probanden deutlich höher. In der bereits referierten Untersuchung Künzels zur Genauigkeit der Sprechererkennung zeigte sich, dass die männlichen Probanden häufiger Fehlurteile als die weiblichen Versuchspersonen produzierten (1989, 93 f.). Es ist jedoch für das hier referierte Experiment 3A anzumerken, dass die Geschlechtsverteilung keine Ausgewogenheit aufwies. Insgesamt standen den 60 Frauen nur 16 männliche Probanden gegenüber. Es ist durchaus denkbar, dass in diesem Experiment die männlichen Versuchspersonen tatsächlich weniger risikofreudig waren. Eindeutig hingegen lassen sich die Bewertungen der Gruppe des Weiterbildungsmasters interpretieren, die ihre Entsprechung in den Bewertungen der Altersgruppe *ab 26 Jahre* findet. Ältere Probanden scheinen bei der Beurteilung zurückhaltender und vorsichtiger zu sein. Dieser Befund deckt sich mit den Resultaten Künzels, wobei in seiner Untersuchung der Alterseffekt sogar überzufällige Beziehungen aufwies (1989, 93).

Nach **Hypothese 5** nimmt die Gewissheit, die Monologisierungen tatsächlich inhaltlich erfasst zu haben, mit zunehmendem Alter ab. Aufgrund der verschiedenen Gewissheitsgrade beim Textverstehen der *Altersgruppe bis 25 Jahre* (0,43) und *ab 26 Jahre* (0,28) lässt sich diese Vermutung tatsächlich stützen.

Auch eine indirekte Beweisführung kann hier angeführt werden: Die Subgruppe der Teilnehmer des Weiterbildungsmasters mit dem Gewissheitsgrad von 0,08 umfasste fast ausschließlich die Gruppe *ab 26 Jahre*. Insofern kann die Hypothese 5 angenommen werden.

Zusätzliche Befunde ergeben sich aus der Analyse der Qualität des Verhörens. Der hohe Prozentsatz an Hörfehlern ist ein Nachweis, dass die Wahrnehmung gesprochener Sprache gerade unter ungünstigen akustischen Bedingungen von Erwartung und Weltwissen geleitet wird. Bemerkenswert im Zusammenhang dieser Untersuchung sind die Hörfehler, die zum einen in Beziehung mit Ausbildung bzw. Studium der Probanden gesetzt werden können oder – wie im Fall der Sprecherin 4 – mit dem antizipierten Status der Sprecherin. Grundsätzlich kann jedoch die Annahme z. B. Hörmanns auch für dieses Experiment bestätigt werden (1981, 138): Bei der Aufnahme und Verarbeitung von Sprache kommt es nicht zu einem einfachen Informationsverarbeitungsprozess, sondern es werden auch stets neue Informationen geschaffen.

Die Forschungsfragen zum referierten Experiment lassen sich somit wie folgt beantworten:

1. Welche prosodischen Merkmale werden am sichersten erkannt?

Diese Frage kann nur im Hinblick auf die Auswahl der untersuchten prosodischen Merkmale beantwortet werden. Das am sichersten erkannte prosodische Merkmal ist das *Geschlecht* der sprechenden Person, wobei nicht davon ausgegangen wird, dass sich diese Sicherheit nur aus Tonhöhenverlaufsmerkmalen herleitet. Für die Eindrucksbildung können auch Aspekte der Klangqualität und der Art der Melodieführung entscheidend gewesen sein. Weiterhin konnten die *Stimmlage*, *Artikulationsrate* und *Modulationsbreite* sicher bestimmt werden. Von allen untersuchten Bestimmungsgrößen wurde jedoch die *Tatsache des Sprechens* mit Abstand am genauesten erkannt. Obwohl es sich nicht um ein eigentliches prosodisches Merkmal handelt müssen diesem Eindruck doch deren Erscheinungen zugrunde gelegen haben, da segmentale Merkmale erst sehr spät und mit großer Unsicherheit wahrgenommen werden.

2. Wie genau werden die verschiedenen prosodischen Merkmale erkannt?

Die *Tatsache des Sprechens* wurde über alle Maskierungsstufen mit 98,36 % erkannt. Für das *Geschlecht* der sprechenden Person ergab sich ein Mittelwert von 86,6 %, wobei der größte Zuwachs an Sicherheit zwischen der höchsten und zweithöchsten Maskierungsstufe erfolgte. Für die *Stimmlage* ergab sich ein Mittelwert von 76,56 %, für die *Artikulationsrate* von 75,41 %. Bei der *Modulationsbreite* beträgt der Mittelwert der Sicherheit 70,19 %. Für die letzten drei Merkmale wurden sowohl zwischen dem ersten und zweiten als auch zwischen der zweiten und dritten Maskierungsstufe große Zuwächse an Bewertungssicherheit errechnet.

3. Welche Wortart wird unter akustisch ungünstigen Hörbedingungen am besten erkannt?

Am häufigsten wurden Wörter der *offenen* Klasse mit einem Mittelwert von 90,05 % über alle vier Sprecher erkannt. Hier kann durchaus von einer gezielten Suche nach Schlüsselwörtern zur Erfassung der Aussage ausgegangen werden. Jedoch wurden auch gebräuchliche Redefloskeln erkannt, die im kommunikativen Geschehen häufig Verwendung finden.

4. Welche Arten von „Hörfehlern“ können unter akustisch ungünstigen Hörbedingungen festgestellt werden?

Die mit Abstand häufigsten Fehler des Verhörens waren *phonematischer* Natur, gefolgt von *semantisch* und *morphologisch* zu interpretierenden Hörfehlern. Die beiden letztgenannten traten mit annähernd gleicher Häufigkeit auf.

5. Gibt es einen Zusammenhang zwischen erkanntem verbalen Material und dem subjektiv empfundenem Verständlichkeitsgrad des Gesagten unter akustisch ungünstigen Hörbedingungen?

Lediglich ein Sachverhalt konnte statistisch abgesichert werden, wonach ältere Probanden trotz gut erkanntem verbalen Materials vorsichtiger bzgl. der Überzeugung sind, das perzipierte Material tatsächlich auch vollständig inhaltlich erfasst zu haben. Weiterhin hat sich gezeigt, dass die Fehlerquote des er-

kannten verbalen Materials bei beiden Geschlechtern vergleichbar ist, jedoch waren sich die Männer hinsichtlich der verstandenen Aussage unsicherer. Jedoch handelt es sich hier um eine Tendenz, die der Art der Stichproben geschuldet sein kann und nicht statistisch abgesichert ist.

11.2 Experiment 3B: Experiment zur relativen Wahrnehmung prosodischer Merkmale

11.2.1 Fragestellungen zum Experiment

Wie in den Kapiteln 3.2 und 6 dargestellt, werden prosodische Merkmale in Relation zu einer internalisierten Bezugsgröße wahrgenommen und interpretiert, die als *Hörmuster* angesehen werden können. Auch in den bisher beschriebenen Experimenten trat der *relative Charakter der Wahrnehmung* zumindest implizit immer wieder in Erscheinung. Relativität ist ein wesentlicher Aspekt der Konstruktion eines Perzepts, mit dem ein Abbild der Umwelt geschaffen wird und welches grundsätzlich alle Sinnesmodalitäten betrifft. Nimmt man die Sprechweise eines Menschen wahr, setzt man die Teilaspekte des wahrgenommenen Sprechausdrucksvorgangs in Beziehung zu Hörmustern, die sich im Laufe der Kommunikationsbiografie herausgebildet haben. Insofern hat jeder kommunikativ erfahrene Mensch ein internalisiertes Maß z. B. einer zu hohen Artikulationsrate, einer zu hohen Stimmlage, einer zu leisen Stimme oder einer auffälligen Klangqualität. Ähnliches gilt auch für das visuelle Wahrnehmungssystem, das uns beispielsweise vermittelt, inwieweit die Garderobe eines Menschen in einem konkreten Kontext passend oder unpassend erscheint. Die Relativität der Wahrnehmung als Teilfunktion etablierter Hörmuster ist ein wesentlicher Aspekt, welcher ermöglicht, die überaus große Informationsmenge, die einen wahrnehmenden Menschen in kürzester Zeit erreicht, auf die wesentlichen Inhalte zu reduzieren. Dieser aktive Filterprozess ist als Top-down-Prozess bei jedem Menschen verschieden und situationsbedingt instabil. Jedoch können interindividuelle Muster angenommen werden, die verschiedene Reize einer Sinnesmodalität vergleichbar machen. Ohne eine derartige Annahme wären phonetische Perzeptionsexperimente nicht auswertbar. Mit dem im Folgenden referierten Experiment soll der bisher impliziert angenommene Prozess der Relativität auditori-

scher Wahrnehmung für prosodische Merkmale in den Mittelpunkt der Betrachtung gerückt werden.

Wie dargestellt, wird bei der relativen Wahrnehmung ein eingehender Reiz in Beziehung zu einem schon bekannten, aktivierten Muster gesetzt. Eine grundsätzliche Problematik ist hierbei, inwieweit das als Bezugsgröße fungierende Hörmuster stabil ist oder Schwankungen unterliegt, seien sie situativer oder individueller Natur. Insofern sollen im Folgenden Wahrnehmungsbedingungen geschaffen werden, die ein inhärentes auditorisches Bezugsmuster durch einen akustischen Hintergrundreiz zu modifizieren versuchen. Die Stabilität auditorischer Wahrnehmung, wie sie beispielsweise im Experiment 2A nachgewiesen wurde, lässt diese Fragestellung unter kontrollierten Bedingungen auf der Ebene von Monologisierungen sinnvoll erscheinen.

Die Perzeption globaler und lokaler Tonhöhenverlaufsmuster hat sich in den vorangegangenen Experimenten als relativ stabil erwiesen. Unter der Fokussierung relativer Wahrnehmung ist beispielsweise zu erwarten, dass eine Sprechweise mit einer mittelhohen Stimmlage im Vergleichsexperiment anders wahrgenommen wird, wenn gleichzeitig eine hohe bzw. tiefe Stimme als Maskierungssignal, d. h. im Hintergrund in Erscheinung tritt. Ist die Hintergrundstimme *hoch*, könnte als Folge der relativen Wahrnehmung die zu bewertende Stimme im Vordergrund *tiefer* erscheinen. Umgekehrt wäre zu erwarten, dass bei einer *tiefen* Stimme im Hintergrund die Bewertung der Sprechweise als *höher* erfolgt. Entscheidende Untersuchungsbedingung ist hierbei, dass die hintergründigen Maskierungsreize zwar die Wahrnehmungsschwelle überschreiten, jedoch noch nicht im Sinne eines Vordergrundreizes bewusst werden. Was hier exemplarisch an der Stimmlage erörtert wurde, könnte in gleicher Weise für die Artikulationsrate und die melodische Variabilität zu beobachten sein. Insofern wird mit diesem Experiment in Bezug auf die im Theorieteil referierten Modelle auf charakteristische *Hörmuster* (s. 5.4) Bezug genommen. Da es um keine darüber hinaus gehenden Fragen zum Textverstehen geht, können spezifische linguistische Modelle hier nicht überprüft werden..

Mit diesem Experiment zur Relativität auditorischer Wahrnehmung sollen folgende Fragen beantwortet werden:

1. Ist der relative Charakter der Wahrnehmung für prosodische Merkmale im Maskierungsexperiment empirisch erfassbar?
2. Für welche prosodischen Merkmale kann eine Beeinflussung der Wahrnehmung durch das Maskierungssignal angenommen werden?
3. Welche unterschiedlichen Ausprägungsgrade der relativen Wahrnehmung sind für prosodische Merkmale erkennbar?
4. Können Hörergruppen abgegrenzt werden, die sich bei der Perzeption prosodischer Merkmale weniger vom Maskierungssignal beeinflussen lassen und in diesem Sinne genauer wahrnehmen?

11.2.2 Hypothesen

Für den anstehenden Untersuchungsschritt werden folgende Hypothesen postuliert:

Hypothese 1:

Bei der Perzeption prosodischer Merkmale beträgt die Abweichung bei Anwesenheit eines Maskierungssignals in Form gesprochener Sprache maximal 10%.

Hypothese 2:

Stabile Perzeptionsresultate ergeben sich unter einem Maskierungsreiz für die Parameter *Artikulationsrate*, *Sprechstimmlage*, *melodische Variabilität*, *Gespanntheit*, *Klangfülle* und *Helligkeit*.

Hypothese 3:

Die Stabilität der Perzeption ist unter einem Maskierungsreiz für *Artikulationsrate*, *Sprechstimmlage* und *melodischer Variabilität* höher als für *Gespanntheit*, *Klangfülle* und *Helligkeit*.

Hypothese 4:

Das Maskierungssignal beeinflusst die Perzeption der *Artikulationsrate*, *Sprechstimmlage* und *melodischen Variabilität* in gleicher Ausprägungsstärke.

Hypothese 5:

Das Maskierungssignal beeinflusst die Perzeption der *Artikulationsrate*, *Sprechstimmlage* und *melodischen Variabilität* in Richtung der gegenteiligen Ausprägung.

11.2.3 Versuchsaufbau und Versuchsdurchführung

Die Überprüfung der Hypothesen des Experimentes stellt eine große Herausforderung an die Konstruktion des Versuchsplans dar, da wiederum natürlich gesprochene Monologisierungen verwendet werden sollten. Die Schwierigkeit bestand darin, im Telekommunikationskorpus Sprecherinnen und Sprecher zu finden, die das zu untersuchende Merkmal in deutlicher Ausprägung realisieren, ohne dass Veränderungen am Schallsignal erfolgen. Aufgrund der großen Auswahl des Datenmaterials konnte diesem Anspruch Rechnung getragen werden, auch wenn dabei gelegentlich Kompromisse hinsichtlich der Kontrollierbarkeit der unabhängigen Variablen eingegangen werden mussten. Das sei am Beispiel der *Sprechstimmlage* erläutert: Wenn beim Testreiz eine Frauenstimme eine gemittelte Grundfrequenz von 225,2 Hz aufweist und der Maskierungsreiz 281,4 Hz als Mittelung ergab, so entspricht diese Differenz musikalisch ungefähr einer *großen Terz*. Für ein Item zur Untersuchung dieses Sachverhalts bei Männerstimmen betrug der Testreiz im Mittel 106,6 Hz, während das Maskierungssignal eine gemittelte Grundfrequenz von 167,2 Hz aufwies. Hier beträgt das musikalische Intervall ca. eine *kleine Sexte*. Damit ist der Kontrast für die zweite Bedingung größer, was auf Wahrnehmung einen Einfluss haben könnte. Diese Ungenauigkeiten wurden bewusst in Kauf genommen, um die Authentizität des Materials zu gewährleisten.

Als Untersuchungsdesigns war wiederum ein Wahrnehmungsexperiment unter der Bedingung eines Signal-Rausch-Verhältnisses (SRV) vorgesehen: Bei Telefonaten in Callcentern fällt häufig auf, dass aufgrund der Großraumbürosituation neben der Stimme des gerade anrufenden Callcenteragenten auch Sprecher im Hintergrund mehr oder weniger deutlich hörbar sind, wobei die Hintergrundstimmen als Maskierungssignal anzusehen sind. Insofern kann diese Konstellation als eine alltagstypische Bedingung angesehen werden. Für das hier

durchgeführte Experiment wurde dieses Setting kontrolliert eingesetzt: Es wurden für alle Versuchspersonen zwei Untersuchungstermine vorgegeben, wobei der Testreiz, d. h. die zu bewertende Sprechweise des jeweiligen Callcenteragenten, zweimal gleich war. Unterschiedlich war hingegen das Maskierungssignal, das bei einem der Testtermine auftrat, beim anderen hingegen nicht. Die Bedingungen hierfür waren randomisiert. Den Versuchspersonen wurde dieses Experiment als ein *Experiment zur Konstanz der Wahrnehmung hinsichtlich prosodischer Merkmale* angekündigt. Da zwischen beiden Untersuchungsterminen mindestens 14 Tage lagen und der Testreiz immer konstant war, wurde die Anwesenheit des Maskierungssignals zu den verschiedenen Untersuchungsterminen in keinem Fall bemerkt. Damit wurde der identische Testreiz zweimal bewertet, wobei einmal das Maskierungssignal auftrat und dessen Einfluss auf die unabhängige Variable überprüft werden konnte.

Nr.	Testkriterium	Sprecherge- schlecht	1. Versuchsdurchgang		2. Versuchsdurchgang	
			Testsignal	Maskierung	Testsignal	Maskierung
1	Sprechstimmlage	w	225,2 Hz (a)	–	225,2 Hz (a)	281,4 Hz (cis ¹)
2	melodische Variabilität	m	977 Cent	2795 Cent	977 Cent	–
3	Artikulationsrate	w	3,79 Silben/s	5,74 Silben/s	3,79 Silben/s	–
4	Sprechstimmlage	m	106,6 Hz (Gis)	–	106,6 Hz (Gis)	167,2 Hz (e)
5	melodische Variabilität	w	1833 Cent	867 Cent	1833 Cent	–
6	Sprechstimmlage	m	133,1 Hz (c)	–	133,1 Hz (c)	95,8 Hz (Fis)
7	melodische Variabilität	w	867 Cent	2409 Cent	867 Cent	–
8	Artikulationsrate	m	6,25 Silben/s	–	6,25 Silben/s	3,53 Silben/s
9	Sprechstimmlage	w	281,4 Hz (cis ¹)	w: 225,2 Hz (a)	281,4 Hz (cis ¹)	–
10	melodische Variabilität	m	2795 Cent	–	2795 Cent	977 Cent

Tab. 11-10: Übersicht über die Untersuchungsvariablen und Testbedingungen des Experimentes 3B (gemittelte Werte der Untersuchungsvariablen)

Das Testmaterial bestand wiederum aus dem Vorlesen der allgemeinen Geschäftsbedingungen bzw. aus den Ausführungen des Callcenteragenten beim Verkauf eines Festplattenreceivers. Für das Testsignal wurden Monologisierungen von fünf Sprecherinnen und fünf Sprecher verwendet. Die untersuchten Kri-

terien waren die *melodische Variabilität* (je 2 Sprecherinnen und 2 Sprecher), die *Sprechstimmlage* (je 2 Sprecherinnen und 2 Sprecher) und die *Artikulationsrate* (je eine Sprecherin und ein Sprecher). Die Abfolge des Experiments, die untersuchten Variablen, das Geschlecht der Sprecher und die jeweiligen Bedingungen für den 1. und 2. Versuchsdurchgang sind in Tabelle 10-10 dargestellt. In Tabelle 10-10 sind gleichfalls die Modifikationen der unabhängigen Variablen für den jeweiligen Testfall aufgeführt. Die Operationalisierung prosodischer Parameter ist zwangsläufig vom entsprechenden Merkmal abhängig, die im Kommunikationsverlauf natürlichen Schwankungen unterliegen. Insofern dürfen die angegebenen Bedingungen auch nur als Näherungswerte angesehen werden. Für die *Artikulationsrate* wurden die entsprechenden Silben in Beziehung zur Sprechzeit gesetzt, wobei die Sprechpausen nicht berücksichtigt wurden. Für die *Sprechstimmlage* wurde der Mittelwert der Frequenz bestimmt; der besseren Lesbarkeit wegen wurde die ungefähre musikalische Tonhöhe in Klammern angegeben. Am schwierigsten gestaltet sich die Quantifizierung der *melodischen Variabilität*: Für dieses Maß ist nicht nur das Intervall zwischen der höchsten und tiefsten Frequenz maßgebend, sondern auch die Geschwindigkeit und Größe der tonalen Veränderung über den Äußerungsverlauf hinweg. Für dieses Komplexphänomen wurde bisher in der Phonetik noch kein eigenes Maß postuliert. Tatsächlich ist ein derartiges Maß schwierig zu entwickeln, da für den Eindruck einer monotonen Sprechweise beispielsweise auch die Artikulationsgeschwindigkeit, Sprechspannung und Klangqualität in die Bewertung einfließt. Dass es dennoch möglich ist, eine *monotone* von einer *modulierten* Sprechweise zu unterscheiden, hängt wiederum mit erworbenen Hörmustern zusammen. Um dennoch ein Maß für die *melodische Variabilität* zu finden, wurde in Tabelle 10-10 das absolute Tonhöhenintervall in Cent angegeben. Da sich eine Tonhöhenkontur im Rahmen einer schriftlichen Ausarbeitung besser visuell erfassen lässt, wurden exemplarisch für vier Darbietungen jeweils eine Äußerung im Anhang C abgebildet. Abbildung 37 (Anhang C) stellt eine Sequenz aus dem Testitem 2 des Sprechers mit geringer melodischer Variabilität dar. Die entsprechende Äußerung lautet: „*dazu bekommen sie noch das kabel digital home paket für einen monat kostenlos dazu*“. Gegenübergestellt ist in Abbildung 38 (Anhang C) die Tonhöhenkontur mit hoher melodischer Variabilität für den Sprecher aus Testi-

tem 10 mit der Äußerung: „*leg ich ihnen mal eine ausgabe unserer zeitschrift tv digital kostenlos dazu*“. Die Parametereinstellungen im Programm war in beiden Fällen identisch. Deutlich sind die größeren Veränderungen und stärkeren Extreme der Tonhöhenkontur in der Abbildung 38 (Anhang C) zu erkennen.

In Abbildung 39 (Anhang C) wird aus dem Testitem 7 ein Ausschnitt der Melodiekontur der Sprecherin mit geringer melodischer Variabilität gezeigt. Die entsprechende Äußerung lautet: „*es läuft einfach so dass die drei monate für sie kostenlos sind*“. Kontrastiv zeigt die Abbildung 40 (Anhang C) die Sprecherin des Testitems 7 mit einer hohen melodischen Variabilität in der Äußerung: „*wie gesagt sie haben 3 monate um zuhause auszuprobieren. würde ihnen das gerät nicht zusagen*“. Da auch hier die Parametereinstellungen gleich sind, können auch hier die schnelleren Wechsel und größeren Extreme in der Tonhöhenkontur erkannt werden.

Um vergleichbare Bedingungen über alle zehn untersuchten Untersuchungstems zu schaffen, wurde der Intensitätspegel für alle Test- und Maskierungsvariablen normiert. Für die Untersuchungsbedingungen wurden unter dem SRV die Schalldruckpegel der Maskierungsvariablen um 15 dB abgesenkt.

Trotz der experimentellen Fragestellung nach prosodischen Merkmalen muss jedoch herausgestellt werden, dass die jeweils im Fokus stehende unabhängige Variable niemals isoliert zu einer Eindrucksbildung beiträgt. Sprechweisen werden immer als Komplexphänomene wahrgenommen, und so ist es auch unter den hier referierten Versuchsbedingungen zwangsläufig nicht auszuschließen, dass auch andere Merkmale das Wahrnehmungsergebnis beeinträchtigt haben. Als zu beurteilende Merkmale waren jedoch nicht nur die drei oben aufgeführten Untersuchungsvariablen aufgeführt, sondern zusätzlich drei Kriterien zur Klangqualität: *Gespanntheit*, *Klangfülle* und *Helligkeit*. Die Bedeutung und Relevanz dieser Merkmale wurde unter Punkt 2.2 referiert.

Zur Dokumentation der Wahrnehmungsergebnisse schien eine Intervallskalierung im Sinne einer Lickertskala zu ungenau zu sein. Auch bei Verwendung einer feinstufigen Skalierung wäre nicht ausgeschlossen, dass sich im Abstand von 14 Tagen die Probanden bei einer wiedererkannten Sprechweise noch an ihr

erstes Bewertungsurteil erinnern könnten. Es war vielmehr anzunehmen, dass die Einflüsse des Maskierungssignals sehr dezent ausfallen. Daher wurde eine Intervallskala für alle sechs zu beurteilenden Kriterien vorgegeben und als Referenzintervall der Bereich von 0-100 angenommen. Die Probanden sollten entsprechend ihres Wahrnehmungsergebnisses eine Markierung an entsprechender Position vornehmen. Je nach Untersuchungsvariable waren die Grenzbedingungen verbal vorgegeben (z. B. *minimal – maximal*). Im Anhang D ist ein Teil des Bewertungsbogens mit den entsprechenden Instruktionen wiedergegeben. Bei der Auswertung wurde die Markierung ausgemessen, wobei für die Skala das Intervall 0-100 zugrunde gelegt wurde. Der gemessene Wert galt als Rohwert für die nachfolgenden statistischen Berechnungen.

An der Untersuchung nahmen 67 Probanden (50 Frauen und 17 Männer) zwischen 20 und 58 Jahren (Medianwert: 24,0 Jahre) teil. Es handelte es sich um Studierende der Sprecherziehung aller Jahrgangsstufen. Um für die statistische Untersuchung einen naturgemäß zu erwartenden Fehlerquotienten als Inkonstanz der Wahrnehmung berücksichtigen zu können, wurden die Bewertungen von fünf phonetisch geschulten Kontrollhörern zugrunde gelegt, die an zwei verschiedenen Testterminen die Testvariablen bewerteten. In diesem Fall wurden die Maskierungsvariablen nicht eingespielt. Der gemittelte Abweichungswert dieser Expertengruppe wurde als tolerierter Fehler bei der Auswertung der Probanden zugrunde gelegt.

11.2.4 Untersuchungsergebnisse

11.2.4.1 Ergebnisse zur relativen Wahrnehmung melodischer und temporaler prosodischer Merkmale

Im Zentrum der Ergebnisdarstellung soll die Art und Weise der Beeinflussung der relativen Wahrnehmung durch das Maskierungssignal für die drei Untersuchungsvariablen stehen. Es war vermutet worden, dass beispielsweise die Artikulationsrate als *schneller* eingeschätzt wird, wenn die Sprechweise des Maskierungssignals *langsam* ist. Verallgemeinert müsste also die Wahrnehmung von *Artikulationsrate*, *Sprechstimmlage* und *melodischen Variabilität* durch das

Maskierungssignal in Richtung der gegenteiligen Ausprägung beeinflusst werden. Berechnet werden kann dies durch die Differenz der Mittel- bzw. Medianwerte der beiden Versuchsdurchgänge. Nach diesen Überlegungen wären folgende Differenzen zu erwarten:

- Testitem 1 und 4 (Sprechstimmlage): negativ
- Testitem 6 und 9 (Sprechstimmlage): positiv
- Testitem 2 und 7 (melodische Variabilität): negativ
- Testitem 5 und 10 (melodische Variabilität): positiv
- Testitem 3 (Artikulationsrate): negativ
- Testitem 8 (Artikulationsrate): positiv

Einen ersten graphisch vermittelten Eindruck gewähren die Abbildungen 11-7 bis 11-17. Die Mittelwerte und Streuungen für die jeweiligen Untersuchungsbedingungen aus dem 1. und 2. Untersuchungsdurchgang werden dabei als Boxplots nebeneinander in einer Abbildung dargestellt. Zum Vergleich sind die Ergebnisse der entsprechenden Bewertung der Experten (im Folgenden als *Kontrollgruppe* bezeichnet) gegenübergestellt. In Abbildung 11-7 und 11-8 sind die Ergebnisse für die *Sprechstimmlage* aufgeführt, wobei im jeweils 2. Untersuchungsdurchgang (jeweils 2. und 4. Kolumne) für die Versuchsgruppe eine *negative* Differenz erwartet wurde.

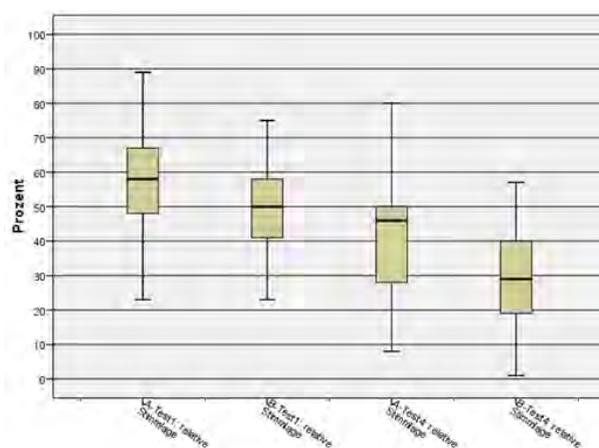


Abb. 11-7: Boxplots der Bewertungen für *Stimm-
lage* Item 1 und 4 der Versuchsgruppe

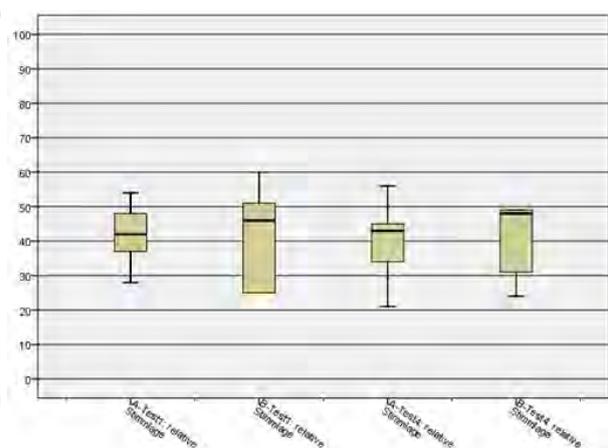


Abb. 11-8: Boxplots d. Bewertungen für *Stimm-
lage* Item 1 und 4 der Kontrollgruppe

Während die Mittelwerte für die Paarungen der jeweils zwei Untersuchungsvariablen in der Kontrollgruppe im 2. Untersuchungsdurchgang höhere Werte

annehmen (s. Abbildung 11-8), zeigen sich in Abbildung 11-7 für die Versuchsgruppe niedrigere Mittelwerte. In den Abbildungen 11-9 und 11-10 werden ebenfalls für die *Sprechstimmlage* die Bedingungen gepaart verglichen, bei denen im 2. Untersuchungsdurchgang *höhere* Mittelwerte erwartet wurden.

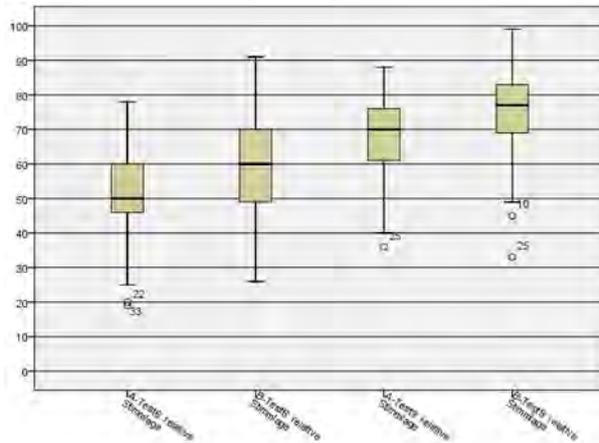


Abb. 11-9: Boxplots der Bewertungen für *Stimm-lage* Item 6 und 9 der Versuchsgruppe

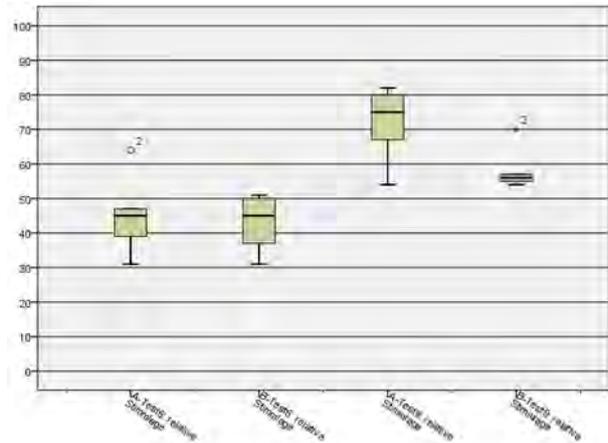


Abb. 11-10: Boxplots d. Bewertungen für *Stimm-lage* Item 6 und 9 der Kontrollgruppe

Während für die Kontrollgruppe bei Item 6 nahezu identische Mittelwerte erreicht werden, fällt beim Testitem 9 ein deutlich niedrigerer Wert im Wiederholungsdurchgang auf. Hingegen in der Versuchsgruppe (s. Abbildung 11-9) zeigen sich im 2. Untersuchungsdurchgang jeweils höhere Mittelwerte. In Abbildung 11-11 und 11-12 werden bezogen auf *die melodische Variabilität niedrigere* Werte im Wiederholungsdurchgang erwartet.

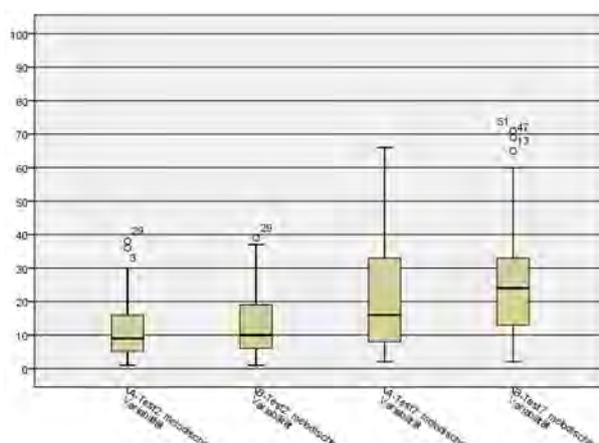


Abb. 11-11: Boxplots der Bewertungen für *melodische Variabilität* Item 2 und 7 der Versuchsgruppe

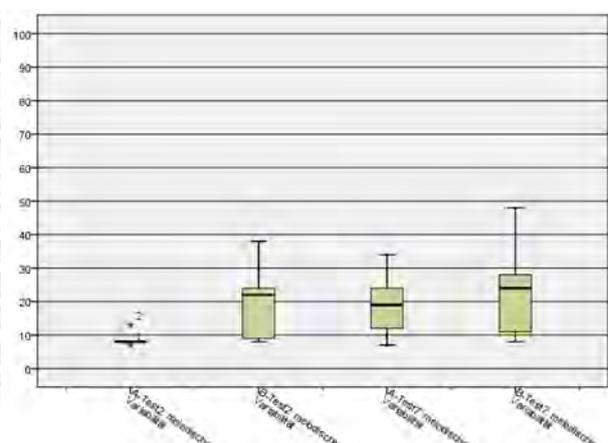


Abb. 11-12: Boxplots d. Bewertungen für *melodische Variabilität* Item 2 und 7 der Kontrollgruppe

Für die Versuchsgruppe in Abbildung 11-11 zeigen sich nahezu identische (Item 2) oder höhere Mittelwerte (Item 7). In der Kontrollgruppe fallen die Mittelwerte des Wiederholungsdurchganges prinzipiell höher aus. Die den Abbildungen 11-13 und 11-14 werden die Bedingungen der *melodischen Variabilität* gegenübergestellt, wobei im 2. Untersuchungsdurchgang *höhere* Mittelwerte erwartet werden.

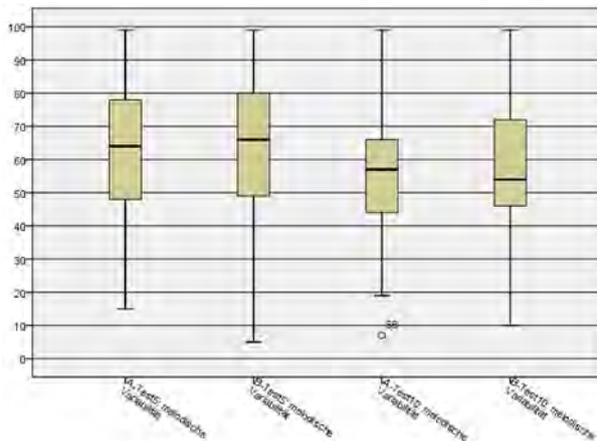


Abb. 11-13: Boxplots der Bewertungen für *melodische Variabilität* Item 5 und 10 der Versuchsgruppe

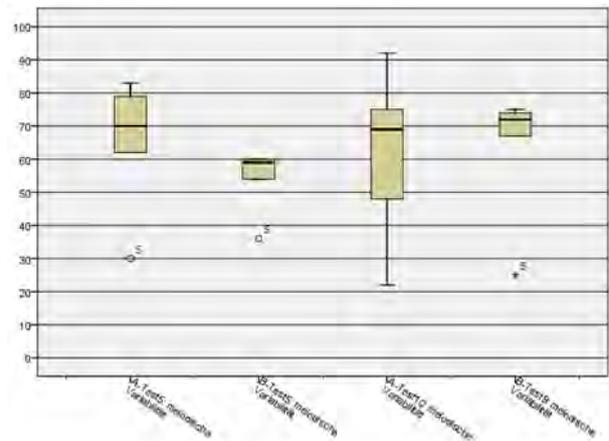


Abb. 11-13: Boxplots d. Bewertungen für *melodische Variabilität* Item 5 und 10 der Kontrollgruppe

Für die Versuchsgruppe zeigen sich nahezu identische Mittelwerte, die Ergebnisse der Kontrollgruppe hingegen fallen divergent aus. In Abbildung 11-14 und 11-15 wird die Bedingung erfasst, wo hinsichtlich der *Artikulationsrate* *niedrigere* Werte im 2. Versuchsdurchgang erwartet wurden.

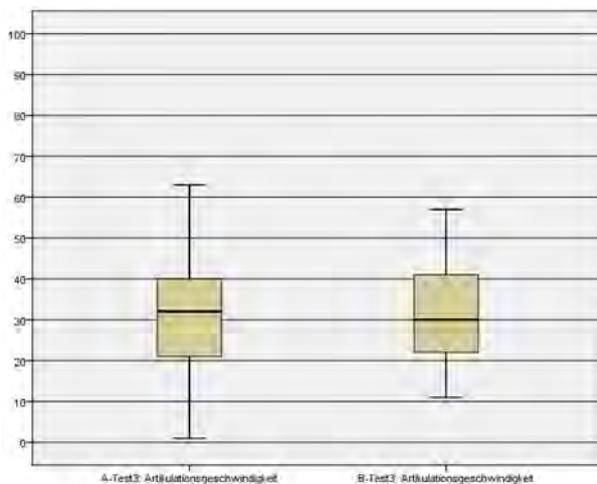


Abb. 11-14: Boxplots der Bewertungen für *Artikulationsrate* Item 3 der Versuchsgruppe

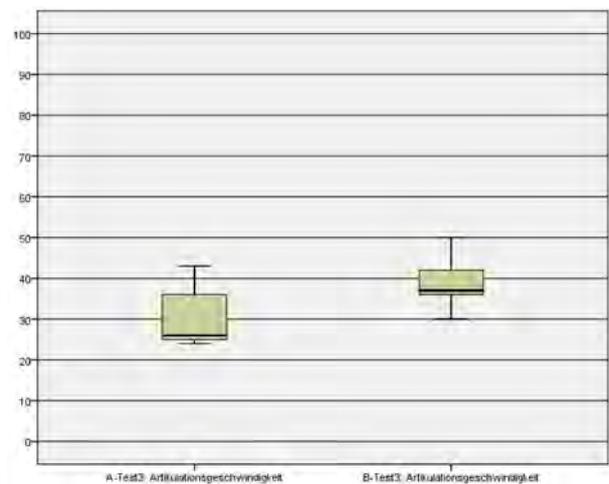


Abb. 11-15: Boxplots d. Bewertungen für *Artikulationsrate* Item 3 der Kontrollgruppe

Während sich für die Versuchsgruppe nahezu identische Mittelwerte ergeben (s. Abbildung 11-15), ist der Mittelwert der Kontrollgruppe in der Wiederholungsuntersuchung höher. Abbildungen 11-16 und 11-17 schließlich erfassen die Bedingung, wo bezüglich der *Artikulationsrate* im 2. Versuchsdurchgang *höhere* Mittelwerte erwartet werden.

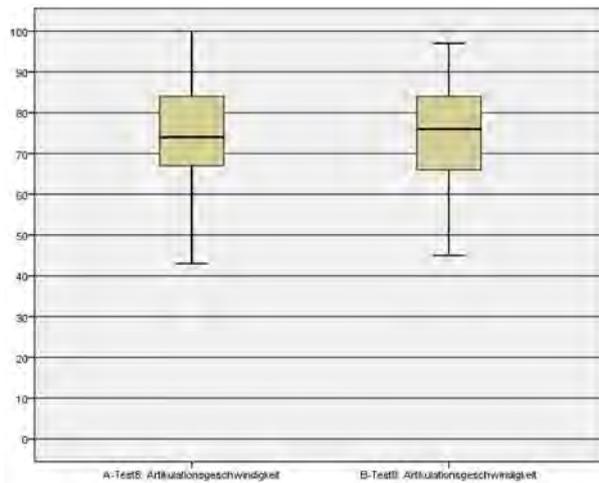


Abb. 11-16: Boxplots der Bewertungen für *Artikulationsrate* Item 8 der Versuchsgruppe

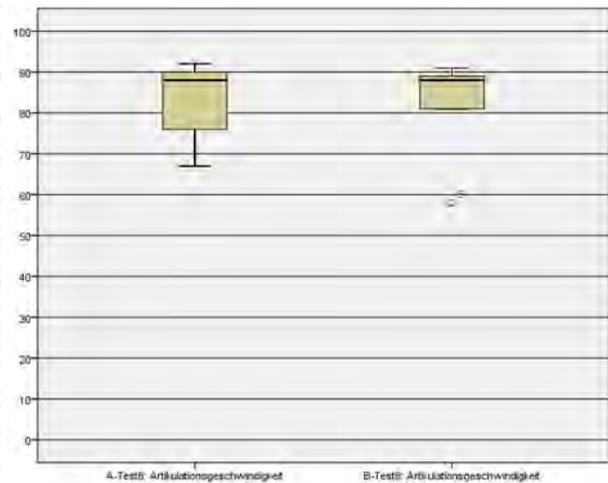


Abb. 11-17: Boxplots d. Bewertungen für *Artikulationsrate* Item 8 der Kontrollgruppe

Sowohl in Abbildung 11-16 für die Versuchsgruppe als auch in Abbildung 11-17 für die Kontrollgruppe ergeben sich nahezu identische Mittelwerte.

	Item	Mittelwert	Median	STABW
Stimmelage	1	-0,40	-3	8,961
	4	0,40	3	5,273
	6	-2,40	0	6,841
	9	-13,20	-12	10,521
Melodische Variabilität	2	-11,40	-9	12,137
	7	-4,60	-4	13,069
	5	11,00	10	13,928
	10	2,20	1	8,044
Artikulationsrate	3	-8,20	-6	9,985
	8	-1,20	-1	5,020

Tab. 10-11: Ergebnisse der deskriptiven Statistik der Kontrollgruppe sortiert nach den Untersuchungsgrößen

Der graphisch vermittelte Überblick zeigt für die Versuchsgruppe, dass hinsichtlich der *Sprechstimmlage* die Erwartungen offensichtlich erfüllt wurden, nicht hingegen für die *melodische Variabilität* und die *Artikulationsrate*. Außerdem ist ersichtlich, dass auch die Kontrollgruppe erhebliche Streuungen in den Beurteilungen aufweist. Einen genaueren Überblick über die Ergebnisse der er-

rechneten Differenzbeträge der drei untersuchten Bedingungen für die Versuchsgruppe ermöglichen die Tabellen 9-11 im Anhang C.

Vor der quantitativen Auswertung der Ergebnisse der Versuchsgruppe sollen zunächst die Resultate der Kontrollgruppe betrachtet werden. In Tabelle 10-11 sind deren Ergebnisse sortiert nach den drei Untersuchungsvariablen aufgeführt. Die Differenzen der Beurteiler über beide Versuchsdurchgänge sind hierbei als *Mittelwert*, *Medianwert* und *Standardabweichung* (STABW) aufgeführt. Aus Tabelle 10-11 ergibt sich, dass sowohl Mittelwerte als auch Mediane lediglich für Test-Item 9 und 5 (hinsichtlich des Mittelwertes auch Item 2) Werte im zweistelligen Bereich aufweisen. Bedenkt man, dass sich die Skalierung über 100 Stufen erstreckt, erscheinen Differenzwerte im einstelligen Bereich als bemerkenswert konstantes Wahrnehmungsergebnis. Bezogen auf die drei Testvariablen ergeben sich niedrigere Bewertungen im 2. Untersuchungsdurchgang für die *Artikulationsrate* und tendenziell die *Sprechstimmlage*. Bei der *melodischen Variabilität* sind die Verhältnisse hingegen ausgeglichen.

	Item	Erwartungswert	Mittelwert	Median	STABW
Stimmmlage	1	negativ	-7,51	-10	15,200
	4	negativ	-10,18	-12	17,330
	6	positiv	6,13	9,5	16,678
	9	positiv	7,07	8	13,673
Melodische Variabilität	2	negativ	-1,43	-1	10,382
	7	negativ	-2,77	-3	21,508
	5	positiv	-1,16	-3	22,952
	10	positiv	-3,00	-2	24,247
Artikulationsrate	3	negativ	0,41	1	13,747
	8	positiv	0,38	4	17,017

Tab. 10-12: Erwartungswerte und Ergebnisse der deskriptiven Statistik sortiert nach den Untersuchungsgrößen der Versuchsgruppe

In Tabelle 10-12 werden die entsprechenden Ergebnisse der errechneten Differenzen für die Versuchsgruppe dargestellt, wobei zusätzlich die Erwartung bezüglich des Median- bzw. Mittelwertes aufgeführt ist.

Sowohl durch die Mediane als auch Mittelwerte bestätigen sich die schon aus den Abbildungen ersichtlichen Tendenzen: Für die *Sprechstimmlage* zeigen sich bei allen vier Items die hypothetisch postulierten Differenzen. Bezüglich der *me-*

lodischen Variabilität fallen die Ergebnisse des 2. Versuchsdurchgangs immer etwas niedriger aus, während für die *Artikulationsrate* bezogen auf den Mittelwert nahezu identisch bewertet wurde. Insofern scheinen die Vorhersagen zumindest für die *Sprechstimmlage* erfüllt zu sein, wobei noch zu prüfen sein wird, inwieweit das in überzufälliger Weise der Fall ist.

Auf einen bemerkenswerten Nebentyp sei noch verwiesen: Bezogen auf den Mittel- bzw. Medianwert aller Items wurde im 2. Versuchsdurchgang etwas niedriger bewertet als das im Erstversuch der Fall war. In der Versuchspersonengruppe fällt diese Differenz etwas geringer aus (gemittelter Medianwert: -0,85; gemittelter Mittelwert: -1,21). In der Kontrollgruppe bei identischem Material zeigen sich mit einem gemittelten Medianwert von -2,1 und einem gemittelten Mittelwert von -2,78 noch deutlich negativere Differenzen.

11.2.4.2 Ergebnisse zur relativen Wahrnehmung der Klangqualitätsmerkmale

Mit dem vorgestellten Experiment sollte zusätzlich überprüft werden, inwieweit Probanden *Klangqualitätsmerkmale* einheitlich und stabil wahrnehmen bzw. wie stark die Abweichungen zu den hier primär untersuchten melodischen und temporalen prosodischen Merkmalen ausfallen. Zur Untersuchung gelangten dabei die Merkmale *Gespanntheit*, *Klangfülle* und *Helligkeit*, deren Relevanz für die Beschreibung von Klangqualitätseigenschaften an anderer Stelle hergeleitet wurde (Kranich 2014b, 22 ff.).

Die Ergebnisse für die einzelnen Testitems sind in den Tabellen 12-14 des Anhangs C aufgeführt. Ein Maß hinsichtlich der Stabilität der Wahrnehmung ist wiederum die Bestimmung der Differenzwerte zwischen den beiden Versuchsdurchgängen. In der nachfolgenden Tabelle 10-13 werden die gemittelten Differenzwerte für die Klangqualitätsmerkmale separiert nach Versuchspersonen- und Kontrollgruppe aufgeführt.

	Versuchspersonengruppe			Kontrollgruppe		
	Mittelwert	Median	STABW	Mittelwert	Median	STABW
Differenz Gespanntheit	-2,27	-2	25,638	-0,06	0,5	16,884
Differenz Klangfülle	-0,83	-2	21,748	-1,90	-2	12,838
Differenz Helligkeit	-1,46	-1	19,441	0,80	-0,5	14,645

Tab. 10-13: Gemittelte Differenzwerte der untersuchten Klangqualitätsmerkmale differenziert nach Versuchspersonen- und Kontrollgruppe

Es zeigt sich, dass für beide untersuchte Gruppen sowohl die Median- als auch Mittelwerte sehr niedrig ausfallen und – wie schon oben festgestellt – das tendenziell im 2. Untersuchungsdurchgang niedrigere Werte vergeben wurden. Insofern kann geschlussfolgert werden, dass auch die untersuchten Klangqualitätseigenschaften relativ stabil in der Wahrnehmung ausfallen. Zum Vergleich seien hierzu die Tabellen 9-11 des Anhangs C herangezogen, wo die gemittelten Differenzwerte der Parameter *Artikulationsrate*, *relative Sprechstimmlage* sowie *melodische Variabilität* aufgeführt sind, deren Beträge durchaus vergleichbar mit den untersuchten Klangqualitätseigenschaften ausfallen.

Item-Nr.	1. Versuchsdurchgang			2. Versuchsdurchgang		
	T-Wert	df	Signif. (2-seitig)	T-Wert	df	Signif. (2-seitig)
Item 1: relative Stimmlage-Helligkeit	-1,958	70	,054*	-2,534	68	,014
Item 2: relative Stimmlage-Helligkeit	3,688	70	,000	3,616	68	,001
Item 3: relative Stimmlage-Helligkeit	-,270	70	,788*	-,828	68	,411*
Item 4: relative Stimmlage-Helligkeit	3,901	70	,000	1,031	68	,306*
Item 5: relative Stimmlage-Helligkeit	-,552	70	,583*	1,222	68	,226*
Item 6: relative Stimmlage-Helligkeit	2,220	70	,030	4,053	68	,000
Item 7: relative Stimmlage-Helligkeit	-,601	70	,550*	-1,342	68	,184*
Item 8: relative Stimmlage-Helligkeit	1,346	70	,183*	,580	68	,564*
Item 9: relative Stimmlage-Helligkeit	-9,062	70	,000	-1,873	68	,065*
Item 10: relative Stimmlage-Helligkeit	1,967	70	,053*	2,131	68	,037

Tab. 10-14: Ergebnisse des T-Tests für verbundene Stichproben mit Darstellung der T-Werte, Freiheitsgrade und Korrelationskoeffizienten

Auf einen zusätzlichen Befund hinsichtlich der drei untersuchten Klangqualitätsparameter sei verwiesen: Ein in der Stimmfunktionsdiagnostik immer wieder beobachtbares Phänomen ist der Zusammenhang von *Sprechstimmlage* und *Helligkeit*. Oft werden Sprechweisen mit einer *hellen* Klangqualität generell als *hoch* eingeschätzt und ein *dunkler* Klang wird oft synonym für *tiefe* Stimme gebraucht. Tatsächlich bedarf es eines erheblichen Hörtrainings, um beide Parameter auditorisch voneinander trennen zu können. Da beide Parameter zum Unter-

suchungskatalog dieses Experimentes gehörten, erschien eine Gegenüberstellung der Bewertungen naheliegend. In Tabelle 10-14 werden die Ergebnisse des T-Tests für verbundene Stichproben der Merkmale *relative Stimmlage* und *Helligkeit* aufgeführt, wobei neben dem T-Wert auch die Freiheitsgrade und die Berechnung des Korrelationskoeffizienten vorgenommen wurden. In der Tabelle 10-14 werden die Korrelationskoeffizienten mit * markiert, die einen überzufälligen Zusammenhang zwischen der *relativen Stimmlage* und der *Helligkeit* aufzeigen. Für beide Versuchsdurchgänge zeigt sich, dass 6 von 10 Variablen diese Annahme bestätigen. Dieser Wert, der als Zusammenhangsmaß ansonsten für keine der gesamten untersuchten Parameter berechnet werden konnte. Insofern kann geschlussfolgert werden, dass ein derartiger Zusammenhang in der Wahrnehmung von *Stimmlage* und *Helligkeit* tatsächlich besteht, was jedoch einer eigenen Überprüfung speziell unter dieser Fragestellung bedarf.

Auch hinsichtlich der Bewertungen zur Klangqualität sei hervorgehoben, dass diese prosodischen Merkmale nicht unabhängig von anderen Eigenschaften wahrgenommen werden und – wie bei den Ausführungen unberücksichtigt gelassen wurde – die Untersuchungsbedingung des Maskierungsreizes nicht in die Auswertung eingeflossen ist. Dass es dennoch, wie in Tabelle 10-13 dargestellt, zu relativ niedrigen Differenzen zwischen beiden Untersuchungsgängen gekommen ist, kann als Indiz angesehen werden, dass die Wahrnehmungen von Eigenschaften der Klangqualität tatsächlich relativ stabil sind.

11.2.5 Ergebnisinterpretation und Hypothesenbeantwortung

Die Interpretation der Ergebnisse des hier vorgestellten Experiments bringt einige Probleme mit sich, was primär der Intention der Verwendung alltagstypischen Testmaterials geschuldet ist. Dadurch ergeben sich Schwierigkeiten, die unabhängigen Testvariablen zu isolieren. Bei einer phonetischen Veränderung einzelner Variablen hätte diesem Aspekt besser begegnet werden können, jedoch um den Preis der Authentizität des Untersuchungsmaterials. Außerdem ist nicht eindeutig abgrenzbar, welche Perzeptionsresultate der Instabilität der auditorischen Wahrnehmung geschuldet sind und welche tatsächlich dem Einfluss des Maskierungssignals zugeordnet werden können.

Nach **Hypothese 1** sollte die Wahrnehmung prosodischer Merkmale auch bei Anwesenheit eines Maskierungssignals in Form gesprochener Sprache mit 10 % Abweichung relativ stabil sein. Diese Fragestellung bezieht sich auf die Mittelung aller Ergebnisse hinsichtlich aller sechs untersuchten Merkmale. Hier bietet sich an, die Ergebnisse der Kontrollgruppe mit denen der Versuchspersonengruppe zu vergleichen, auch wenn nur die Versuchsgruppe mit dem in der Hypothese postulierten Maskierungsreiz als Versuchsbedingung konfrontiert wurde. Da der Untersuchung eine Intervallskala von 0 bis 100 zugrunde gelegt wurde, können direkt die Differenzwerte des 1. und 2. Untersuchungsdurchganges zur Überprüfung der Hypothese herangezogen werden. Jedoch ist unter dieser Fragestellung zu berücksichtigen, dass die Beträge der Differenzen in die Bewertung einfließen müssen, da sich andernfalls gegenläufige Differenzen aufheben würden. Danach ergibt sich für die Kontrollhörer ein Abweichungswert von 6,03 % bezogen auf die Medianwerte und von 5,7 % auf die Mittelwerte. Für die Versuchspersonengruppe zeigt sich hingegen eine Abweichung von 3,08 % hinsichtlich der Medianwerte und 3,15 % unter Berücksichtigung der Mittelwerte. Der höhere Abweichungsgrad der Kontrollgruppe resultiert aus der Gruppengröße. Bei der Kontrollgruppe standen fünf Hörer zur Verfügung, deren Ergebnisse sich natürlich nicht in gleichem Maße ausgleichen wie das bei den 67 Versuchspersonen der Fall ist. Da der Differenzwert zur postulierten Grenze für die Versuchspersonengruppe 6,85 % beträgt, kann Hypothese 1 angenommen werden: Der Abweichungsgrad ist kleiner als 10 %. Auch der Abweichungswert der Kontrollgruppe wäre noch 3,97 % von der postulierten 10%-Grenze entfernt. Es zeigt sich also auch bei dieser Untersuchung die immer wieder postulierte Notwendigkeit hinreichend großer Versuchspersonengruppen.

In **Hypothese 2** wurde behauptet, dass sich stabile Perzeptionsresultate unter einem Maskierungsreiz für die Parameter *Artikulationsrate*, *Sprechstimmlage*, *melodische Variabilität*, *Gespanntheit*, *Klangfülle* und *Helligkeit* ergeben. Die Bedingungen dieser Hypothese wären erfüllt, wenn sich für alle sechs prosodischen Merkmale ähnliche Differenzwerte aus den Wahrnehmungsergebnissen beider Untersuchungsdurchgänge abzeichnen würden. In Tabelle 10-15 sind die Ergebnisse der gemittelten Mittel- und Medianwerte für die Versuchspersonengruppe wiedergegeben.

	Mittelwert	Medianwert
Artikulationsrate	-0,47	-0,3
Relative Stimmlage	-0,73	-0,15
Melodische Variabilität	0,69	0,1
Gespanntheit	-2,27	-1,6
Klangfülle	-0,83	-0,7
Helligkeit	-1,05	-1,3

Tab. 10-15: Durchschnittliche Mittel- und Medianwerte der Differenzen der untersuchten prosodischen Merkmale für die Versuchspersonengruppe

Legt man die Medianwerte zugrunde, zeigt sich mit -1,6 für *Gespanntheit* der von Null am stärksten abweichende Wert. Berücksichtigt man hingegen die Mittelwertbildung, fällt bei dem entsprechenden Merkmal mit -2,27 das Ergebnis etwas höher aus. Für die *melodische Variabilität* ergeben sich hinsichtlich des Medianwertes mit 0,1 die besten Ergebnisse, für den Mittelwert errechnet sich mit -0,47 hinsichtlich der *Artikulationsrate* der der Null am nächsten liegende Wert.

Mit einer weiteren Maßzahl lässt sich die Gültigkeit der postulierten Hypothese überprüfen: Sollten unter der Bedingung des Maskierungsreizes die Beurteilungen des jeweiligen prosodischen Merkmals problematischer sein, müsste sich das in einer größeren Standardabweichung niederschlagen. In Tabelle 10-16 werden daher für alle untersuchten Merkmale die Standardabweichungen gegenübergestellt, wobei die Sortierung nach der Art des Untersuchungsreizes durchgeführt wurde (*unmask.* = keine Maskierung; *mask.* = Maskierung).

Nr.	Artikulationsrate		Relative Stimmlage		Melodische Variabilität		Gespanntheit		Klangfülle		Helligkeit	
	unmask.	mask.	unmask.	mask.	unmask.	mask.	unmask.	mask.	unmask.	mask.	unmask.	mask.
1	9,887	10,311	12,618	14,624	18,595	18,068	17,367	21,912	19,567	17,968	19,436	19,111
2	20,758	21,735	18,606	15,078	8,911	9,786	26,593	25,788	23,630	21,072	20,690	19,438
3	14,111	12,001	15,830	14,675	18,018	17,911	25,792	25,220	17,180	16,789	16,426	17,669
4	10,859	10,506	13,593	15,371	17,830	19,375	19,067	19,463	19,469	21,457	16,580	15,685
5	13,896	13,843	14,238	14,770	22,574	21,349	18,683	14,425	20,507	21,356	14,888	16,476
6	10,307	9,547	17,035	11,706	20,580	20,896	16,466	19,025	20,001	21,511	16,813	18,541
7	18,906	17,182	16,834	18,162	17,321	17,010	35,241	31,391	15,897	17,582	17,747	16,895
8	11,963	13,131	13,455	13,079	19,522	18,807	21,422	18,646	18,943	18,619	16,972	17,665
9	16,070	15,216	12,157	13,219	25,655	25,158	24,805	22,322	17,721	18,870	12,256	14,486
10	12,440	12,303	19,290	19,066	20,179	18,824	21,806	22,360	20,745	21,239	14,486	12,256

Tab. 10-16: Sortierte Standardabweichungen der Versuchspersonengruppe

Vergleicht man die entsprechenden Paare aller untersuchten Merkmale, zeigen sich keine durchgehend größeren Standardabweichungen unter Maskierungsbedingungen. Fallen die größeren Divergenzen auf, wie z. B. bei *Stimmlage* (Item 6) und *Gespanntheit* (Item 5 und 7), sind diese höheren Standardabweichungswerte jedoch immer unter der Bedingung der Nichtmaskierung auszumachen. Die insgesamt relativ niedrig ausfallenden Mittelwerte und Mediane der Differenzen, die sich in einem vergleichbaren Bereich bewegen (s. Tabelle 10-15) sowie die hohe Übereinstimmung der Standardabweichungen (s. Tabelle 10-16) legen nahe, dass sich für alle sechs untersuchten prosodischen Merkmale auch unter einem Maskierungsreiz relativ stabile Hörmuster einstellen. Insofern kann Hypothese 2 als verifiziert gelten.

Hypothese 3 postuliert, dass die Stabilität der Perzeption unter einem Maskierungsreiz für *Artikulationsrate*, *Sprechstimmlage* und *melodischen Variabilität* größer ausfällt als für *Gespanntheit*, *Klangfülle* und *Helligkeit*. Trifft diese Hypothese zu, sollten die Median- bzw. Mittelwerte für die drei Klangqualitätsmerkmale niedriger ausfallen als für *Artikulationsrate*, *Sprechstimmlage* und *melodische Variabilität*. Die Antwort darauf gibt wiederum Tabelle 10-15. Tatsächlich zeigen sich für die drei genannten Merkmale etwas niedrigere Werte. Leicht höher fällt der Differenzwert der *Klangfülle* aus während die *Gespanntheit* mit -1,6 die größte Abweichung aufweist. Bedenkt man jedoch, dass der maximal mögliche Differenzwert bei theoretisch ± 99 liegen könnte, erscheinen diese Beträge minimal.

Auch prüfstatisch lässt sich diese Vermutung untermauern: Unterzieht man alle Merkmale einem Mittelwertvergleich mittels des T-Tests für gepaarte Stichproben, ergeben sich für alle Bedingungen überzufällige Übereinstimmungen. Die einzige Ausnahme ist die Paarung *melodische Variabilität* zu *Gespanntheit* ($\alpha = ,020$). Aufgrund dieser Tatsache muss die Hypothese 3 zurückgewiesen werden, da die Stabilität der Hörmuster unter einem Maskierungsreiz für alle sechs geprüften prosodischen Merkmale annähernd gleich stabil ausfällt.

Mit **Hypothese 4** wird der Gedankengang aus der Ergebnisdarstellung (Punkt 11.2.4.1) wieder aufgegriffen. Danach sollte der Maskierungsreiz die Wahrnehmung der *Artikulationsrate*, *Sprechstimmlage* und *melodischen Variabilität* in

nahezu gleicher Stärke beeinflussen. Tatsächlich wird man hier mit der zu Beginn dieses Kapitels dargestellten Problematik konfrontiert, die sich aus dem Versuchsplan ergeben: Es lässt sich nicht zweifelsfrei bestimmen, ob die Wahrnehmungsveränderungen im 2. Untersuchungsdurchgang eine Folge der Instabilität der Hörwahrnehmungen sind oder ob diesen tatsächlich die Maskierungsbedingungen zugrunde liegen. Um diese Hypothese dennoch überprüfen zu können, wurden Untersuchungsbedingungen geschaffen, wonach je 2 Testitems der *relativen Sprechstimmlage* und *melodischen Variabilität* einen negativen bzw. positiven Differenzwert aufweisen sollten. Hinsichtlich der *Artikulationsrate* kam dabei jeweils nur eine Bedingung zur Untersuchung. Die entsprechenden Differenzwerte wurden in Tabelle 10-12 aufgeführt und zeigen, dass die Erwartungen lediglich für die *relative Sprechstimmlage* erfüllt wurden. Insofern kann im Sinne der Hypothese 4 lediglich überprüft werden, ob für das Merkmal *Sprechstimmlage* eine überzufällige Differenz errechnet werden kann, die wiederum ein Hinweis wäre, dass die differierenden Untersuchungsergebnisse tatsächlich auf das Maskierungssignal zurückzuführen sind. Um in diesem Fall das Merkmal der Instabilität der Hörwahrnehmung zu reduzieren, wird die Differenz der Kontrollgruppe als einkalkulierter Faktor berücksichtigt, gegen den die Ergebnisse der beiden Versuchsdurchgänge der Versuchspersonengruppe zu testen sind. Fasst man die Ergebnisse der Kontrollgruppe hinsichtlich der Merkmale *Sprechstimmlage*, *melodische Variabilität* und *Artikulationsrate* zusammen, ergeben sich die in Tabelle 10-17 aufgeführten Differenzwerte.

Kontrollgruppe	Mittelwerte	Medianwerte	Standardabweichung
Sprechstimmlage	-0,36	0,5	10,581
Melodische Variabilität	-0,94	-1	18,597
Artikulationsrate	-1,72	-1	10,955

Tab. 10-17: Differenz der beiden Untersuchungsdurchgänge

Unter den ausgeführten Bedingungen ist es sinnvoll, die vier Ergebnisse der Versuchspersonengruppe für das Merkmal *Sprechstimmlage* testen. Dafür bietet sich als statistisches Verfahren der Kontrasttest (ANOVA) gegen den Referenzwert $\pm 0,5$ (Medianwert der Kontrollgruppe für die *Sprechstimmlage*) an, da es sich zum einen um stetige Merkmale handelt und zum anderen von einer Normalverteilung ausgegangen werden kann. Die Nullhypothese gilt als

angenommen, wenn die Beurteilungen der beiden Versuchsdurchgänge sich nicht überzufällig unterscheiden. Es errechnen sich folgende Werte:

Item 1: $\alpha = ,161$	→	H_0 gilt
Item 4: $\alpha = ,006$	→	H_1 gilt
Item 6: $\alpha = ,049$	→	H_1 gilt
Item 9: $\alpha = ,009$	→	H_1 gilt

Auf der Basis der errechneten Irrtumswahrscheinlichkeiten ergibt sich, dass lediglich für den Test 1 keine überzufällige Verschiedenheit der Ergebnisse für die *Sprechstimmlage* beider Versuchsdurchgänge angenommen werden kann. Bei nichtparametrischer Überprüfung durch den Wilcoxon-Test kann auch für Item 1 die Nullhypothese angenommen werden. Die Hypothese 4 muss jedoch insgesamt zurückgewiesen werden, da sich für die *Artikulationsrate* und *melodische Variabilität* keine Veränderung des Wahrnehmungsergebnisses nachweisen lässt, die auf den Maskierungsschall zurückzuführen wäre.

Hypothese 5 zielt in eine ähnliche Richtung. Danach beeinflusst das Maskierungssignal die Wahrnehmung von *Artikulationsrate*, *Sprechstimmlage* und *melodischer Variabilität* in Richtung der gegenteiligen Ausprägung. Wie gerade referiert, kann ein entsprechender Einfluss nur für die *Sprechstimmlage* angenommen werden. Mit den Ergebnissen der deskriptiven Statistik, die durch prüfstatistische Verfahren weitgehend bestätigt wurden, ist zu vermuten, dass die relative Wahrnehmung unter der Bedingung eines Maskierungssignals die Perception der *Sprechstimmlage* verändert: Eine tiefere Stimme im Hintergrund lässt die vordergründige Stimme höher erscheinen und umgekehrt. In der postulierten Absolutheit muss jedoch auch die Hypothese 5 zurückgewiesen werden.

Abschließend sei noch ein Blick auf die Beurteilungsgenauigkeit der einzelnen Probanden gelenkt. Wie in den meisten Untersuchungen zeigt sich auch hier, dass einzelne Personen genauere Wahrnehmungsergebnisse erzielen als andere. Berücksichtigt man bei diesem Experiment die gemittelten Absolutwerte der Mediane, so weist der Differenzwert *Null* eine sehr hohe Konstanz der Wahrnehmung nach, insbesondere dann, wenn gleichzeitig die gemittelte Standardabweichung niedrig ausfällt. Tatsächlich errechnet sich bei sechs Versuchssper-

sonen ein Medianwert von Null, jedoch fiel hier die Standardabweichung relativ hoch aus. Auf der anderen Seite lag die höchste Differenz der gemittelten Mediane bei -7. Bezogen auf die Standardabweichung zeigte sich mit 9,271 der niedrigste Wert, der höchste lag hingegen bei 30,096. In Abbildung 11-18 wird die Konstanz der Wahrnehmung über die Probanden im Streudiagramm dargestellt. Zugrunde gelegt wurden dabei der *Absolutwert des Medianwertes* und die *Standardabweichung*.

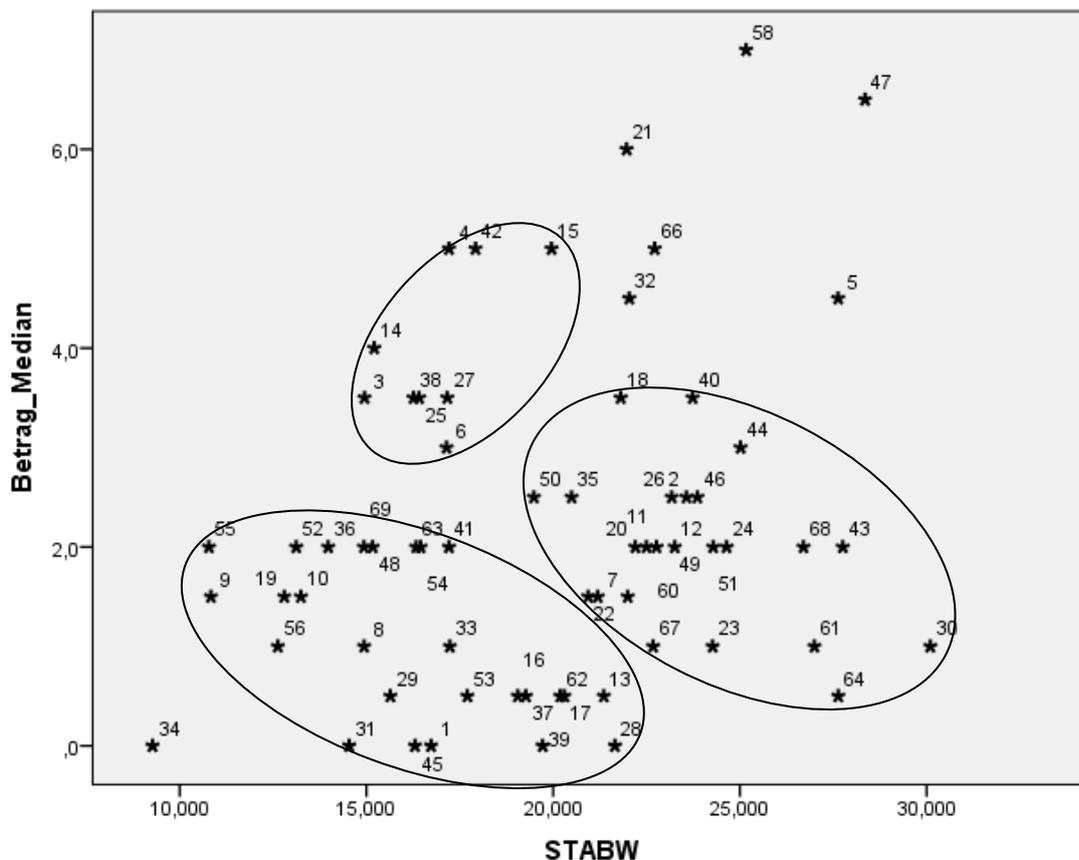


Abb. 11-18: Streudiagramm über den Betrag der Mediane und die Standardabweichungen der Probandengruppe

Die Darstellung weist ein breit gefächertes Feld hinsichtlich der Beurteilungskonstanz aus. Im Fall 34 zeigt sich bei einem Medianwert von *Null* eine sehr niedrige Standardabweichung, die Fälle 47 und 58 zeigen die gegenteiligen Ausprägungen. Insgesamt lassen sich bei hinreichender Trennschärfe drei Gruppen ausmachen, welche in Abbildung 11-18 entsprechend umrandet sind. Gruppe 1 weist niedrige Werte hinsichtlich beider Parameter auf, eine quantitativ annähernd gleichgroße zweite Gruppe zeigt durchschnittlich etwas höhere Medianwerte bei deutlich höherer Standardabweichung, während sich für eine klei-

neren dritte Gruppe ein höherer Absolutwert des Medians bei insgesamt moderaten Standardabweichungen errechnet. Es zeigt sich aber, dass zu keiner dieser statistisch abgrenzbaren Gruppen sich eine spezifische Gruppe – sei es die *Studiendauer*, das *Lebensalter* oder das *Geschlecht* – zuordnen lässt.

Abschließend sollen wiederum die Forschungsfragen dieses Experiments beantwortet werden:

1. Ist der relative Charakter der Wahrnehmung für prosodische Merkmale im Maskierungsexperiment empirisch erfassbar?

Um den relativen Charakter eines Wahrnehmungsprozesses zu bestimmen, müssen die Bezugs- und Bestimmungsgrößen separat erfasst werden. Da eine derartig formgeleitete Fokussierung den Perzeptionsvorgang zwangsläufig nachhaltig beeinflussen muss, werden hohe Anforderungen an das Untersuchungsdesign der empirischen Überprüfung gestellt, um Schlussfolgerungen auf alltagstypische Kommunikationen zu erlauben. Da sich im vorgestellten Experiment Ergebnisse gezeigt haben, die tatsächlich auf eine empirische Erfassung der Relativität der Wahrnehmung schließen lassen, erweist sich das Maskierungsexperiment im kontrollierten Wiederholungsversuch als probates Mittel der Bestimmung von prosodischen Merkmalen. Eine strengere Kontrolle der Variablen als im vorgestellten Experiment erscheint dafür jedoch unerlässlich.

2. Für welche prosodischen Merkmale kann eine Beeinflussung der Wahrnehmung durch das Maskierungssignal angenommen werden?

Das Untersuchungsdesign wurde auf die Konstanz der Wahrnehmung über einen ca. zweiwöchigen Zeitraum unter Berücksichtigung der relativen Wahrnehmung für prosodische Merkmale ausgerichtet. Insofern können hier Aussagen auf die Bezugsgrößen prosodischer Merkmale gemacht werden (siehe Forschungsfrage 1). Grundsätzlich hat sich für die Kontrollgruppe eine hohe Konstanz der Wahrnehmung gezeigt, deren Werte als Richtgröße für die Versuchsgruppe unter der Bedingung des Maskierungsreizes herangezogen wurden. Danach kann eine Beeinflussung der Wahrnehmung eindeutig für die *Sprechstimmlage* angenommen werden.

3. Welche unterschiedlichen Ausprägungsgrade der relativen Wahrnehmung sind für prosodische Merkmale erkennbar?

Für die untersuchten prosodischen Merkmale zeigte sich eine Beeinflussung der Wahrnehmung eindeutig für die *Sprechstimmlage*, was auch statistisch abgesichert werden konnte. Bei der *Artikulationsrate* kann nur die Hälfte der Bedingungen als erfüllt angesehen werden, was der Ratewahrscheinlichkeit entspricht. Kein Einfluss zeigte sich für die *melodische Variabilität* und die Klangqualitäts-eigenschaften *Gespanntheit*, *Klangfülle* und *Helligkeit*. Eine interessante Beziehung ergab sich jedoch dahingehend, dass ein Wahrnehmungszusammenhang zwischen der *Stimmlage* und der *Helligkeit* für die Mehrzahl der Fälle auch statistisch belegt werden konnte. Diese Beziehung wird oft im Zusammenhang stimmdiagnostischer Fragestellungen erkannt.

4. Können Hörergruppen abgegrenzt werden, die sich bei der Perzeption prosodischer Merkmale weniger vom Maskierungssignal beeinflussen lassen und in diesem Sinne genauer wahrnehmen?

Berücksichtigt man die Parameter *Medianwert* und *Standardabweichung* über beide Versuchsdurchgänge, so findet sich eine Gruppe von Probanden, die sehr konstant bewertet hat. Diese Personen lassen sich jedoch keiner natürlichen Gruppe zuordnen. Zwar waren alle diese Personen schon in einem frühen Lebensalter bzw. lange musikalisch aktiv, jedoch fallen in diese Gruppe auch nicht bzw. weniger musikalisch geschulte Probanden. Auch ergaben sich bei zahlreichen musikalisch geschulten Versuchspersonen höhere Abweichungen hinsichtlich der angelegten Parameter. Insofern muss diese Frage verneint werden.