

**Die Effekte kommunikationsunterstützender
Systeme auf natürliche Dialogsprache im
Auto
- eine phonetisch-linguistische Analyse**

Dissertation zur Erlangung des Doktorgrades
der Philosophischen Fakultät
der Christian-Albrechts-Universität
zu Kiel

vorgelegt von

Rabea Landgraf

Kiel

28. Februar 2018

Erstgutachter: Prof. Dr. Oliver Niebuhr

Zweitgutachter: Prof. Dr.-Ing. Gerhard Schmidt

Tag der mündlichen Prüfung: 26.06.2018

Durch den zweiten Prodekan, Prof. Dr. Ulrich Müller zum Druck
genehmigt: 10.10.2018

Danksagung

Mein Dank gilt meinem Doktorvater Prof. Dr. Oliver Niebuhr, der mich immer mit wissenschaftlichem und persönlichem Rat unterstützt und stets Zeit gefunden hat, mit mir Schwierigkeiten, Änderungen und Ergebnisse meiner Forschungsarbeit zu diskutieren. Vielen Dank für deine großartige Förderung, nicht nur während meiner Promotionsarbeit, sondern bereits seit den Jahren meines Bachelorstudiums.

Ich danke meinem Zweitgutachter Prof. Dr. Gerhard Schmidt, für die tatkräftige Hilfe bei der Generierung und Bearbeitung des Sprachkorpus SPID. Mein Dank gilt den Mitarbeiter*innen der Digitalen Signalverarbeitung und Systemtheorie (DSS) (Institut für Elektrotechnik und Informationstechnik) der CAU Kiel für ihre Unterstützung, insbesondere Christian Lüke, Dr. Jochen Withopf und Dr. Anne Theiß.

Zudem danke ich der Konrad-Adenauer-Stiftung, die mich als Stipendiatin im Rahmen ihrer Promotionsförderung während der gesamten Zeit meiner Promotion sowohl ideell als auch finanziell unterstützt und gefördert hat.

Ich danke Dr. Tina John für ihre vielen guten Ratschläge. Unsere gemeinsamen Praxistests zur Sprechaufgabe des SPID-Korpus und ihre fachliche Einschätzung meiner Ergebnisse haben mir sehr geholfen. Auch den weiteren Mitarbeiter*innen der Allgemeinen Sprachwissenschaft (Institut für Skandinavistik, Frisistik und Allgemeine Sprachwissenschaft) der CAU Kiel, wie Prof. Dr. Alena Witzlack-Makarevich und Dr. Benno Peters, danke ich für die vielfältige fachliche Unterstützung. Zudem danke ich den Praktikantinnen Jana Voße, Katrin Busse und Jana Thumm für ihre Hilfe bei der manuellen Transkription, Segmentierung und Annotation der Sprachsignale des SPID-Korpus, die viel Zeit in Anspruch nahm. Weiterhin möchte ich Jana Falk für ihre Unterstützung bei der Transkriptionsarbeit der SPID-Sprachsignale danken.

Herzlichen Dank an die Proband*innen, mit deren Hilfe das Sprachkorpus SPID überhaupt erst generiert werden konnte. Danke für eure Geduld und euren Einsatz. Dies gilt auch für die Teilnehmer*innen der Umfrage im Perzeptionsexperiment, die

Online durchgeführt wurde.

Ich möchte meiner Familie danken, meinen Eltern Gerd und Simone, meinem Bruder Boris und meiner Schwägerin Nicole, für ihre Liebe und ihre Unterstützung. Sie waren immer für mich da und haben mir stets bedingungslos geholfen, egal, welchen Weg ich eingeschlagen habe. Ich danke meinem Mann Johannes, der immer zu mir gehalten und mich besonders in schweren Phasen meiner Arbeit liebevoll ermutigt und in meinem Tun bestärkt hat.

Vielen Dank an meine Freundinnen und Freunde für die emotionale Unterstützung und eine Menge wertvoller Zeit. Ein besonderer Dank geht dabei an Britta Steinke, die meine gesamte Dissertation mit viel Geduld und großer Konzentration Korrektur gelesen hat.

Abstract

Lautsprachliche Kommunikation im Auto wird durch verschiedene Faktoren erschwert, zu denen vor allem die Geräuschbelastung während der Fahrt gehört. Dabei spielt der Lombard-Effekt, welcher die Veränderungen der Sprachproduktion bei lauten Umgebungsgeräuschen beschreibt, eine entscheidende Rolle. Die bei diesem Phänomen auftretenden sprachlichen Modifikationen werden als Ausdruck eines durch Lärm hervorgerufenen Stresses der Sprecher*innen interpretiert. Durch sogenannte Innenraumkommunikationssysteme (*In-Car Communication systems*, ICC-Systeme) wird versucht, diese Situation zu verbessern.

Die vorliegende Arbeit untersucht, wie sich die Kommunikation im Fahrzeug bei lauten Geräuschkulissen verändert und inwieweit diese Modifikationen durch den Einsatz eines ICC-Systems rückgängig gemacht werden können. Zu diesem Zweck wurden Sprachaufnahmen mit verschiedenen Dialogpaaren in einer akustischen und visuellen Fahrgeräuschsimulation bei Geräuschkulissen von 50 km/h, 130 km/h und einer stillen Referenzbedingung erhoben. Dadurch konnten Sprachsignale unter kontrollierten Laborbedingungen aufgenommen und die Geräusche nachträglich aus den Sprachaufnahmen entfernt werden, wodurch die Untersuchung unverfälschter Parameterwerte, wie z. B. der Grundfrequenz, möglich ist. Die Aufnahmen wurden sowohl ohne als auch mit Anwendung eines ICC-Systems durchgeführt. Die Umgebungs-simulation und das verwendete ICC-System wurden an der Christian-Albrechts-Universität zu Kiel (CAU) entwickelt. Auf diese Weise ist ein Sprachkorpus freigesprochener Dialoge (*SPontaneous In-car Dialogues*, SPID) entstanden, das der ICC-Forschung und weiterführenden linguistischen Fragestellungen zur Lombard-Sprache nachhaltig von Nutzen sein kann.

Die deskriptiven und inferenzstatistischen Analysen der aufgenommenen Sprachsignale und der aus einem Perzeptionsexperiment gewonnenen Daten ergeben, dass bei ansteigenden Fahrgeräuschen im Auto ein Lombard-Effekt auftritt. Die Sprache der

Passagier*innen wird „lauter“, „höher“ und „emotionaler“. Zudem benötigen die Personen weniger Zeit zum Lösen ihrer Sprecher*innenaufgabe im *Map-Task* Szenario (Signaldauer) und modifizieren ihre Sprachproduktionsleistung. Mit Aktivierung des ICC-Systems wird der Lombard-Effekt in den Parametern der durchschnittlichen Intensität und der durchschnittlichen Grundfrequenz abgeschwächt, während er hinsichtlich der Signaldauer und der in den Sprachaufnahmen wahrgenommenen Emotionalität nicht mehr nachgewiesen werden kann. Die Sprachproduktionsleistung der Sprecher*innen verschiebt sich auf den vorderen Sitzplatz.

Die Abschwächung bzw. Aufhebung des Lombard-Effekts wird als Reduktion des Stressniveaus der Passagier*innen interpretiert. Dabei entsteht durch das ICC-System eine neue Kommunikationssituation, welche der Umgebung eines stillen Autoinnenraums ähnelt. Aus diesen Ergebnissen lässt sich schließen, dass die Aktivierung des verwendeten ICC-Systems zu einer Verbesserung der lautsprachlichen Kommunikation im fahrenden Auto und dadurch gleichzeitig zu einer Erhöhung der Fahrsicherheit führt.

Inhaltsverzeichnis

Abbildungsverzeichnis	i
Tabellenverzeichnis	ix
Abkürzungsverzeichnis	xvii
1 Einleitung	1
1.1 Motivation des Forschungsvorhabens	1
1.2 Fragestellung und Hypothesen	3
1.3 Aufbau der Arbeit und Methoden	6
2 Sprachkommunikation im Auto	9
2.1 <i>In-Car Communication systems</i>	10
2.2 Evaluierung von ICC-Systemen	12
2.3 ICC an der CAU Kiel	14
3 Der Lombard-Effekt	17
3.1 Eigenschaften	19
3.2 Verschiedene Interpretationen des Phänomens	24
3.2.1 Der physiologische Reflex	25
3.2.2 Kommunikative Aspekte	26
3.2.3 Kombinierte Ursachen	28
3.3 Die Untersuchung des Phänomens	29
3.3.1 Im Labor	29
3.3.2 Ausmaße des Effekts	30
3.4 Arbeitsdefinition und Hypothesen	32

4	Die akustische und visuelle Umgebungssimulation	37
4.1	Die Geräuschbefreiung der Sprachsignale	39
4.2	Die Realitätsnähe der Simulation im Fahrzeug	41
4.3	Ergänzungen	43
5	<i>SPontaneous In-car Dialogues (SPID)</i>	45
5.1	Das <i>Map-Task</i> Szenario	45
5.2	Aufnahmebedingungen	50
5.3	Die Proband*innen	54
5.4	Instruktionen	55
5.5	Geräuschbefreiung	58
5.6	Dauer der Aufnahmen	62
5.7	Segmentierung, Transkription und Annotation	64
5.7.1	Die semi-automatische Segmentierung durch MAUS	66
5.7.2	Methoden zur verbesserten Arbeit mit MAUS	68
6	Die Analyse der Sprachaufnahmen	79
6.1	Der Sprecher*innen-Fragebogen	79
6.2	Die Zielwörter des <i>Map-Task</i> Szenarios	84
6.3	Regressionsanalyse der Sprachsignale	86
6.3.1	Intensität	89
6.3.2	Grundfrequenz	106
6.3.3	Signaldauer	121
6.3.4	Sprachproduktionsleistung	131
7	Perzeptionsexperiment	149
7.1	Das Semantische Differential	149
7.2	Hypothesen	153
7.3	Die Stimuli	156
7.4	Der <i>Online</i> -Fragebogen	157
7.5	Durchführung	159

7.6	Ergebnisanalyse	160
7.6.1	Bewertungen der Umfrage	162
7.6.2	Deskriptive Beobachtungen	166
7.6.3	Regressionsanalyse der Umfragedaten	173
8	Zusammenfassung	217
8.1	Diskussion der Ergebnisse	217
8.1.1	Teilfrage (1)	217
8.1.2	Teilfrage (2)	221
8.1.3	Teilfrage (3)	228
8.1.4	Zusammenfassung	235
8.2	Fazit	237
8.3	Methodenkritik	239
8.4	Ausblick	241
A	Anhang	253
A.1	Die Karten des <i>Map-Task</i> Szenarios	253
A.2	Die Instruktionen des <i>Map-Task</i> Szenarios	264
A.3	Der Fragebogen zum Produktionsexperiment	267
A.4	Der <i>Online</i> -Fragebogen des Perzeptionsexperiments	270
	Literaturverzeichnis	275

Abbildungsverzeichnis

2.1	Ein mit dem vorgestellten ICC-System ausgestattetes Testauto der CAU Kiel. Sichtbare Mikrophone des Systems sind rot eingekreist. (vgl. [48, S. 99])	14
4.1	Der künstliche Kopf zur Aufnahme von Geräuschen. ([25, S. 2] und [37, S. 100])	38
4.2	Die an den Innenseiten der rechten Autofenster angebrachten Lautsprecher der akustischen Umgebungssimulation. ([26, S. 196] und [38, S. 479])	39
4.3	Spektrogramm eines im Audiolabor aufgenommenen Sprachsignals während einer Fahrgeräuschsimulation von 100 km/h. Dargestellt sind das Originalsignal sowie verschiedene Stufen der Geräuschbefreiung. ([47, S. 7] und [49])	41
5.1	Die erste Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i>	47
5.2	Die erste Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i>	47
5.3	Die für die Aufnahmen während der echten Autofahrt verwendeten Routen in und um Kiel, jeweils rot gekennzeichnet. Links die Strecke für Sprachaufnahmen bei 50 km/h, rechts die Strecke für Sprachaufnahmen bei 130 km/h. (vgl. [87])	53
5.4	Eine kurze Unterbrechung der Sprechaufgabe während Aufnahmen im fahrenden Auto bei 50 km/h.	54

5.5	Verstärkung der Sprachsignale um 6 dB. Zeitgleiche Ausschnitte aus den Sprachsignalen zweier Dialogpartner bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Oben das Sprachsignal von Sprecher 3f (<okay>), unten das von Sprecher 3e (<ha: wie heißt das bei dir? wuss>), dessen Produktion der Laute [ha:] rosa markiert ist, ebenso wie der entsprechende Zeitabschnitt im oberen Sprachsignal von Sprecher 3f.	60
5.6	Verstärkung des Sprachsignals um 6 dB und <i>noise cancellation</i> (NC). Lautsprachliche Produktion von <okay> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System.	61
5.7	Verstärkung des Sprachsignals um 6 dB, <i>noise cancellation</i> (NC) und <i>noise suppression</i> (NS). Die Produktion von <okay> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System.	61
5.8	Verstärkung des Sprachsignals um 6 dB, <i>noise cancellation</i> (NC), <i>noise suppression</i> (NS) und Separierung. Die Produktion von <okay> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System.	62
5.9	Die Segmentierung und Annotation der mittels <ä>, <ö>, <ü> und <ß> in den orthographischen Transkriptionen gekennzeichneten Laute durch MAUS. Von links nach rechts sind folgende Wörter dargestellt: <fährt>, <nördlich>, <für> und <Straßen>. Die drei Ebenen in Praat entsprechen von oben nach unten „ORT“ (wortweise Segmentierung mit orthographischer Kennzeichnung), „KAN“ (wortweise Segmentierung mit kanonischer Annotation in SAMPA) und „MAU“ (phonemische Segmentierung mit SAMPA-Symbolen).	67
5.10	Die Produktion von <den Schulgang nördlich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System.	71

5.11	Die Produktion von <den Schulgang nördlich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Anwendung aller Stufen der Geräuschbefreiung nach Kapitel 4.1.	71
5.12	Die Produktion von <nördlich? Den find ich nich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Anwendung aller Stufen der Geräuschbefreiung nach Kapitel 4.1 sowie der oben beschriebenen Methoden zur Verbesserung der Segmentierung und Annotation durch MAUS.	75
5.13	Die Produktion von <den Schulgang nördlich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Anwendung aller Stufen der Geräuschbefreiung nach Kapitel 4.1 sowie der oben beschriebenen Methoden zur Verbesserung der Segmentierung und Annotation durch MAUS.	76
5.14	Die Produktion von <den Schulgang nördlich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Anwendung aller Stufen der Geräuschbefreiung nach Kapitel 4.1 sowie der oben beschriebenen Methoden zur Verbesserung der Segmentierung und Annotation durch MAUS. Zusätzliche manuelle Korrektur der Segmentierung und Annotation.	77
6.1	Die mittelwertbereinigte durchschnittliche Intensität in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren Daten einzelner Proband*innen. (vgl. [38, S. 482])	90
6.2	Die Mittelwerte zur mittelwertbereinigten durchschnittlichen Intensität über alle Sprecher*innen hinweg.	91
6.3	Die mittelwertbereinigte durchschnittliche Intensität in jeder ICC- und Sitzplatzbedingung.	96
6.4	Die mittelwertbereinigte durchschnittliche Intensität in jeder ICC- und Rollenbedingung.	97
6.5	Die mittelwertbereinigte durchschnittliche Intensität in jeder Geräusch- und Sitzplatzbedingung.	100

6.6	Die mittelwertbereinigte durchschnittliche Intensität in jeder Geräusch- und Rollenbedingung.	101
6.7	Die auf Basis des Regressionsmodells aus Tabelle 6.6 errechneten transformierten Schätzwerte zur durchschnittlichen Intensität. [39, S. 1486]	103
6.8	Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren Daten einzelner Proband*innen. [38, S. 481]	107
6.9	Die Mittelwerte zur mittelwertbereinigten durchschnittlichen Grundfrequenz über alle Sprecher*innen hinweg.	108
6.10	Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder ICC- und Sitzplatzbedingung.	112
6.11	Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder ICC- und Rollenbedingung.	113
6.12	Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder Geräusch- und Sitzplatzbedingung.	116
6.13	Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder Geräusch- und Rollenbedingung.	117
6.14	Die auf Basis des Regressionsmodells aus Tabelle 6.15 errechneten transformierten Schätzwerte zur durchschnittlichen Grundfrequenz.	119
6.15	Die mittelwertbereinigte Signaldauer in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren Daten einzelner Proband*innen.	122
6.16	Die Mittelwerte zur mittelwertbereinigten Signaldauer über alle Sprecher*innen hinweg.	123
6.17	Die mittelwertbereinigte Signaldauer in jeder ICC- und Geräuschbedingung.	124
6.18	Die auf Basis des Regressionsmodells aus Tabelle 6.21 errechneten transformierten Schätzwerte zur Signaldauer. ([39, S. 1487])	129
6.19	Der mittelwertbereinigte Anteil gesprochener Sprache in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren Daten einzelner Proband*innen.	134

6.20	Die Mittelwerte des mittelwertbereinigten Anteils gesprochener Sprache über alle Sprecher*innen hinweg auf dem hinteren Sitzplatz. . . .	135
6.21	Die Mittelwerte des mittelwertbereinigten Anteils gesprochener Sprache über alle Sprecher*innen hinweg auf dem vorderen Sitzplatz. . . .	136
6.22	Der mittelwertbereinigte Anteil gesprochener Sprache für unterschiedliche Sitzplätze, ohne und mit Verwendung des ICC-Systems.	138
6.23	Der mittelwertbereinigte Anteil gesprochener Sprache für unterschiedliche Rollen, ohne und mit Verwendung des ICC-Systems.	140
6.24	Der mittelwertbereinigte Anteil gesprochener Sprache für unterschiedliche Sitzplätze bei verschiedenen Geräuschbedingungen.	141
6.25	Der mittelwertbereinigte Anteil gesprochener Sprache für unterschiedliche Rollen bei verschiedenen Geräuschbedingungen.	141
6.26	Die auf Basis des Regressionsmodells aus Tabelle 6.30 errechneten transformierten Schätzwerte zur Sprachproduktionsleistung für den hinteren Sitzplatz. [39, S. 1487]	146
6.27	Die auf Basis des Regressionsmodells aus Tabelle 6.30 errechneten transformierten Schätzwerte zur Sprachproduktionsleistung für den vorderen Sitzplatz. [39, S. 1487]	147
7.1	Arithmetische Mittel der Skalen „unfreundlich“, „emotional“, „gehemmt“ und „steif“ in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren die Bewertungen von Sprachsignalen der drei Sprecher*innen 2d, 3f und 8o des Produktionsexperiments. .	169
7.2	Arithmetische Mittel der Skalen „angestrengt“, „stockend“, „laut“ und „verständlich“ in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren die Bewertungen von Sprachsignalen der drei Sprecher*innen 2d, 3f und 8o des Produktionsexperiments. .	170
7.3	<i>Boxplots</i> der Skalenbewertungen für „unfreundlich“, „emotional“, „gehemmt“ und „steif“ in jeder ICC- und Geräuschbedingung.	172
7.4	<i>Boxplots</i> der Skalenbewertungen für „angestrengt“, „stockend“, „laut“ und „verständlich“ in jeder ICC- und Geräuschbedingung.	173
7.5	Die auf Basis des Regressionsmodells aus Tabelle 7.8 errechneten transformierten Schätzwerte zur wahrgenommenen Lautstärke. . . .	179

Abbildungsverzeichnis

7.6	Die auf Basis des Regressionsmodells aus Tabelle 7.12 errechneten transformierten Schätzwerte zur wahrgenommenen Verständlichkeit.	184
7.7	Die auf Basis des Regressionsmodells aus Tabelle 7.15 errechneten transformierten Schätzwerte zur wahrgenommenen Unfreundlichkeit von Sprecherin 2d.	189
7.8	Die auf Basis des Regressionsmodells aus Tabelle 7.16 errechneten transformierten Schätzwerte zur wahrgenommenen Unfreundlichkeit von Sprecher 3f und 8o.	190
7.9	Die auf Basis des Regressionsmodells aus Tabelle 7.21 errechneten transformierten Schätzwerte zur wahrgenommenen Emotionalität.	195
7.10	Die auf Basis des Regressionsmodells aus Tabelle 7.25 errechneten transformierten Schätzwerte zur wahrgenommenen Gehemmtheit.	200
7.11	Die auf Basis des Regressionsmodells aus Tabelle 7.27 errechneten transformierten Schätzwerte zur wahrgenommenen Steifheit von Sprecher 3f.	205
7.12	Die auf Basis des Regressionsmodells aus Tabelle 7.29 errechneten transformierten Schätzwerte zur wahrgenommenen Steifheit von Sprecher*in 2d und 8o.	206
7.13	Die auf Basis des Regressionsmodells aus Tabelle 7.34 errechneten transformierten Schätzwerte zur wahrgenommenen Angestrengtheit.	211
7.14	Die auf Basis des Regressionsmodells aus Tabelle 7.38 errechneten transformierten Schätzwerte zur wahrgenommenen stockenden Sprechweise.	216
A.1	Die erste Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i>	254
A.2	Die erste Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i>	254
A.3	Die zweite Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i>	255
A.4	Die zweite Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i>	255
A.5	Die dritte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i>	256
A.6	Die dritte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i>	256
A.7	Die vierte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i>	257
A.8	Die vierte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i>	257
A.9	Die fünfte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i>	258

A.10 Die fünfte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i> . . .	258
A.11 Die sechste Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i> . . .	259
A.12 Die sechste Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i> . . .	259
A.13 Die siebte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i>	260
A.14 Die siebte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i> . . .	260
A.15 Die achte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i>	261
A.16 Die achte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i> . . .	261
A.17 Die neunte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i> . . .	262
A.18 Die neunte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i> . . .	262
A.19 Die zehnte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Giver</i>	263
A.20 Die zehnte Aufgabe des <i>Map-Task</i> Szenarios, Sprechrolle <i>Follower</i> . . .	263
A.21 Die erste Seite des <i>Online</i> -Fragebogens.	271
A.22 Beispielseite des <i>Online</i> -Fragebogens mit Semantischem Differential. .	271
A.23 Die vorletzte Seite des <i>Online</i> -Fragebogens.	272
A.24 Die letzte Seite des <i>Online</i> -Fragebogens.	273

Tabellenverzeichnis

1.1	Die Adjektive des Semantischen Differentials im Perzeptionsexperiment.	8
3.1	Die in Kapitel 3.1 beschriebenen Modifikationen der Sprachproduktion in Lombard-induzierenden Umgebungen. ↑ und ↓ geben eine Zu- oder Abnahme der jeweiligen Parameterwerte an.	23
5.1	Die Zielwörter des <i>Map-Task</i> Szenarios und die darin vorkommenden phonologischen Prozesse.	48
5.2	Die für die Bewältigung der <i>Map-Task</i> Aufgaben benötigte Zeit des Sprecherinnenpaars im Vortest.	50
5.3	Die Aufnahmebedingungen des SPID-Sprachkorpus.	51
5.4	Die Dauer der Sprachsignale des SPID-Korpus in der akustischen und visuellen Umgebungssimulation.	63
5.5	Die Dauer der Sprachsignale des SPID-Korpus während der echten Autofahrt.	63
5.6	Die Ersetzung der Grapheme <ü>, <ö>, <ä> und <ß> in den orthographischen Transkriptionen des SPID-Korpus.	67
5.7	Beispiel einer Erweiterung der orthographischen Transkriptionen um die zeitlichen Endpunkte der einzelnen Phrasen, die Pausen und die Start- und Endpunkte des gesamten Signals. Diese Zeitpunkte sind jeweils durch ein Semikolon vom Text bzw. den mit % gekennzeichneten Skript-Befehlen getrennt.	73
5.8	Die Transkriptionen und Annotationen der Sprachsignale des SPID-Korpus in der akustischen und visuellen Umgebungssimulation.	78
6.1	Die Produktion von Zielwörtern in den Sprachsignalen der akustischen und visuellen Umgebungssimulation.	85

6.2	Regressionsergebnisse für die abhängige Variable <i>intensity</i> und die erklärende Variable <i>ICC</i>	92
6.3	Regressionsergebnisse für die abhängige Variable <i>intensity</i> und die erklärende Variable <i>Geräusch</i>	92
6.4	Regressionsergebnisse für die abhängige Variable <i>intensity</i> und die Interaktion der erklärenden Variablen <i>ICC</i> und <i>Geräusch</i>	93
6.5	Regressionsergebnisse für die abhängige Variable <i>intensity</i> , die Interaktion der erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> und die erklärende Variable <i>Sitzplatz</i>	98
6.6	Regressionsergebnisse für die abhängige Variable <i>intensity</i> , die Interaktion der erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie die Interaktion der erklärenden Variablen <i>ICC</i> und <i>Sitzplatz</i>	99
6.7	Die auf Basis des Regressionsmodells aus Tabelle 6.6 errechneten transformierten Schätzwerte zur durchschnittlichen Intensität (dB). .	102
6.8	Die Effekte der unabhängigen Variablen <i>ICC</i> , <i>Geräusch</i> und <i>Sitzplatz</i> auf die durchschnittliche Intensität.	104
6.9	Ergebnisse des t-Tests zur durchschnittlichen Intensität für die abhängigen Stichproben der Experimentbedingungen <i>ohne ICC</i> , <i>Stille</i> , <i>hinten</i> und <i>mit ICC</i> , <i>Stille</i> , <i>hinten</i>	105
6.10	Ergebnisse des t-Tests zur durchschnittlichen Intensität für die abhängigen Stichproben der Experimentbedingungen <i>ohne ICC</i> , <i>Stille</i> , <i>vorne</i> und <i>mit ICC</i> , <i>Stille</i> , <i>vorne</i>	106
6.11	Regressionsergebnisse für die abhängige Variable <i>pitch</i> und die erklärende Variable <i>ICC</i>	109
6.12	Regressionsergebnisse für die abhängige Variable <i>pitch</i> und die erklärende Variable <i>Geräusch</i>	109
6.13	Regressionsergebnisse für die abhängige Variable <i>pitch</i> und die Interaktion der erklärenden Variablen <i>ICC</i> und <i>Geräusch</i>	111
6.14	Regressionsergebnisse für die abhängige Variable <i>pitch</i> , die Interaktion der erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> und die erklärende Variable <i>Sitzplatz</i>	114

6.15	Regressionsergebnisse für die abhängige Variable <i>pitch</i> , die Interaktion der erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie die Interaktion der erklärenden Variablen <i>ICC</i> und <i>Sitzplatz</i>	114
6.16	Die auf Basis des Regressionsmodells aus Tabelle 6.15 errechneten transformierten Schätzwerte zur durchschnittlichen Grundfrequenz (Hz).	118
6.17	Ergebnisse des t-Tests zur durchschnittlichen Grundfrequenz für die abhängigen Stichproben der Experimentbedingungen <i>ohne ICC, Stille, hinten</i> und <i>mit ICC, Stille, hinten</i>	120
6.18	Ergebnisse des t-Tests zur durchschnittlichen Grundfrequenz für die abhängigen Stichproben der Experimentbedingungen <i>ohne ICC, Stille, vorne</i> und <i>mit ICC, Stille, vorne</i>	120
6.19	Regressionsergebnisse für die abhängige Variable <i>dursignal</i> und die erklärende Variable <i>ICC</i>	125
6.20	Regressionsergebnisse für die abhängige Variable <i>dursignal</i> und die erklärende Variable <i>Geräusch</i>	126
6.21	Regressionsergebnisse für die abhängige Variable <i>dursignal</i> und die Interaktion der erklärenden Variablen <i>ICC</i> und <i>Geräusch</i>	127
6.22	Die auf Basis des Regressionsmodells aus Tabelle 6.21 errechneten transformierten Schätzwerte zur Signaldauer (Sek.).	128
6.23	Ergebnisse des t-Tests zur Signaldauer für die abhängigen Stichproben der Experimentbedingungen <i>mit ICC, Stille</i> und <i>mit ICC, 50 km/h</i>	130
6.24	Ergebnisse des t-Tests zur Signaldauer für die abhängigen Stichproben der Experimentbedingungen <i>mit ICC, 50 km/h</i> und <i>mit ICC, 130 km/h</i>	130
6.25	Ergebnisse des t-Tests zur Signaldauer für die abhängigen Stichproben der Experimentbedingungen <i>mit ICC, Stille</i> und <i>mit ICC, 130 km/h</i>	131
6.26	Regressionsergebnisse für die abhängige Variable <i>amountofspeech</i> und die erklärende Variable <i>ICC</i>	137
6.27	Regressionsergebnisse für die abhängige Variable <i>amountofspeech</i> und die erklärende Variable <i>Geräusch</i>	137

6.28	Regressionsergebnisse für die abhängige Variable <i>amountofspeech</i> und die erklärenden Variablen <i>ICC</i> und <i>Sitzplatz</i>	139
6.29	Regressionsergebnisse für die abhängige Variable <i>amountofspeech</i> und die Interaktionen der erklärenden Variablen <i>ICC</i> und <i>Sitzplatz</i> sowie <i>Geräusch</i> und <i>Sitzplatz</i>	142
6.30	Regressionsergebnisse für die abhängige Variable <i>amountofspeech</i> und die Interaktionen der erklärenden Variablen <i>ICC</i> und <i>Sitzplatz</i> , <i>Geräusch</i> und <i>Sitzplatz</i> sowie <i>Geräusch</i> und <i>Rolle</i>	143
6.31	Die auf Basis des Regressionsmodells aus Tabelle 6.30 errechneten transformierten Schätzwerte zur Sprachproduktionsleistung (%).	145
7.1	Faktoren und dazugehörige Skalen im Semantischen Differential des Perzeptionsexperiments.	152
7.2	Die durchschnittliche Intensität und die durchschnittliche Grundfrequenz der Stimuli des Perzeptionsexperiments.	157
7.3	Die vorläufige Kodierung der Skalenbewertungen im Semantischen Differential des Perzeptionsexperiments am Beispiel der Skala „laut-leise“.	161
7.4	Die finale Kodierung der Skalenbewertungen im Semantischen Differential des Perzeptionsexperiments.	161
7.5	Die Bewertung des <i>Online</i> -Fragebogens durch die Proband*innen.	162
7.6	Regressionsergebnisse der Skala „laut-leise“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>ICC</i>	175
7.7	Regressionsergebnisse der Skala „laut-leise“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>Geräusch</i>	176
7.8	Regressionsergebnisse der Skala „laut-leise“ für die abhängige Variable <i>evaluation</i> , die erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie deren Interaktionen.	176
7.9	Die auf Basis des Regressionsmodells aus Tabelle 7.8 errechneten transformierten Schätzwerte zur wahrgenommenen Lautstärke.	178
7.10	Regressionsergebnisse der Skala „verständlich-unverständlich“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>ICC</i>	180

7.11	Regressionsergebnisse der Skala „verständlich-unverständlich“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>Geräusch</i> .	181
7.12	Regressionsergebnisse der Skala „verständlich-unverständlich“ für die abhängige Variable <i>evaluation</i> , die erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie deren Interaktionen.	182
7.13	Die auf Basis des Regressionsmodells aus Tabelle 7.12 errechneten transformierten Schätzwerte zur wahrgenommenen Verständlichkeit. .	183
7.14	Regressionsergebnisse der Skala „unfreundlich-freundlich“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>ICC</i>	185
7.15	Regressionsergebnisse der Skala „unfreundlich-freundlich“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>Geräusch</i> . . .	186
7.16	Regressionsergebnisse der Skala „unfreundlich-freundlich“ für die abhängige Variable <i>evaluation</i> , die erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie deren Interaktionen.	187
7.17	Die auf Basis des Regressionsmodells aus Tabelle 7.15 errechneten transformierten Schätzwerte zur wahrgenommenen Unfreundlichkeit von Sprecherin 2d.	188
7.18	Die auf Basis des Regressionsmodells aus Tabelle 7.16 errechneten transformierten Schätzwerte zur wahrgenommenen Unfreundlichkeit von Sprecher 3f und 8o.	189
7.19	Regressionsergebnisse der Skala „emotional-sachlich“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>ICC</i>	191
7.20	Regressionsergebnisse der Skala „emotional-sachlich“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>Geräusch</i>	192
7.21	Regressionsergebnisse der Skala „emotional-sachlich“ für die abhängige Variable <i>evaluation</i> , die erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie deren Interaktionen.	193
7.22	Die auf Basis des Regressionsmodells aus Tabelle 7.21 errechneten transformierten Schätzwerte zur wahrgenommenen Emotionalität. . .	194
7.23	Regressionsergebnisse der Skala „gehemmt-ungehemmt“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>ICC</i>	196
7.24	Regressionsergebnisse der Skala „gehemmt-ungehemmt“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>Geräusch</i> . . .	197

7.25	Regressionsergebnisse der Skala „gehemmt-ungehemmt“ für die abhängige Variable <i>evaluation</i> , die erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie deren Interaktionen.	198
7.26	Die auf Basis des Regressionsmodells aus Tabelle 7.25 errechneten transformierten Schätzwerte zur wahrgenommenen Gehemtheit. . .	199
7.27	Regressionsergebnisse der Skala „steif-locker“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>ICC</i>	201
7.28	Regressionsergebnisse der Skala „steif-locker“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>Geräusch</i>	202
7.29	Regressionsergebnisse der Skala „steif-locker“ für die abhängige Variable <i>evaluation</i> , die erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie deren Interaktionen.	203
7.30	Die auf Basis des Regressionsmodells aus Tabelle 7.27 errechneten transformierten Schätzwerte zur wahrgenommenen Steifheit von Sprecher 3f.	204
7.31	Die auf Basis des Regressionsmodells aus Tabelle 7.29 errechneten transformierten Schätzwerte zur wahrgenommenen Steifheit von Sprecher*in 2d und 8o.	205
7.32	Regressionsergebnisse der Skala „angestrengt-entspannt“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>ICC</i>	207
7.33	Regressionsergebnisse der Skala „angestrengt-entspannt“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>Geräusch</i> . . .	208
7.34	Regressionsergebnisse der Skala „angestrengt-entspannt“ für die abhängige Variable <i>evaluation</i> , die erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie deren Interaktionen.	209
7.35	Die auf Basis des Regressionsmodells aus Tabelle 7.34 errechneten transformierten Schätzwerte zur wahrgenommenen Angestrengtheit. .	210
7.36	Regressionsergebnisse der Skala „stockend-flüssig“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>ICC</i>	212
7.37	Regressionsergebnisse der Skala „stockend-flüssig“ für die abhängige Variable <i>evaluation</i> und die erklärende Variable <i>Geräusch</i>	213

7.38	Regressionsergebnisse der Skala „stockend-flüssig“ für die abhängige Variable <i>evaluation</i> , die erklärenden Variablen <i>ICC</i> und <i>Geräusch</i> sowie deren Interaktionen.	214
7.39	Die auf Basis des Regressionsmodells aus Tabelle 7.38 errechneten transformierten Schätzwerte zur wahrgenommenen stockenden Sprechweise.	215
8.1	Die Veränderung der Messwerte aller abhängigen Variablen bei Zunahme der Geräuschbelastung. Grundlage der Berechnungen bilden die Tabellen 6.7, 6.16, 6.22 und 6.31. a) und b) geben die Differenzen der Messwerte zwischen zwei Geräuschbedingungen an, wobei jeweils der Messwert der niedrigeren Geräuschbedingung von dem der höheren subtrahiert wird.	218
8.2	Die Veränderung der Messwerte bei Aktivierung des ICC-Systems in allen Geräuschbedingungen. Grundlage der Berechnungen bilden die Tabellen 6.7, 6.16, 6.22 und 6.31. Die letzte Spalte beinhaltet die arithmetischen Mittel der vorangegangenen drei Spalten.	222
8.3	Die Ergebnisse der statistischen Analysen der Skalenbewertungen (vgl. Kapitel 7.6). Dargestellt werden die Veränderungen der Messwerte bei lauten Fahrgeräuschen ohne und mit ICC-System, differenziert nach den drei Sprecher*innen der verwendeten Sprachausschnitte.	229
8.4	Das Auftreten des Lombard-Effekts in den Parametern der Untersuchung sowie die Veränderung dieses Phänomens durch Aktivierung des ICC-Systems.	235
8.5	Die Verbesserung der Kommunikationsqualität durch mehrere ICC-Systeme, gemessen anhand des Lombard-Quotienten in einem Beispiel zur durchschnittlichen Intensität.	251

Abkürzungsverzeichnis

ANOVA	<i>analysis of variance</i> , Varianzanalyse
bzgl.	bezüglich
bzw.	beziehungsweise
CAU	Christian-Albrechts-Universität
CLARIN	<i>Common Language Resources and Technology Infrastructure</i>
d. h.	das heißt
dB	Dezibel
DSS	Digitale Signalverarbeitung und Systemtheorie
et al.	und andere
evtl.	eventuell
f.	folgende
ff.	fortfolgend
F0	Grundfrequenz
F1	Erster Formant
F2	Zweiter Formant
F3	Dritter Formant
Hz	Hertz
ICC	<i>In-Car Communication</i>
IPA	<i>International Phonetic Alphabet</i>
IPdS	Institut für Phonetik und digitale Sprachverarbeitung
KiRAT	Kiel <i>Real-Time Audio/Application Toolkit</i>
km/h	Kilometer pro Stunde
lm	lineares Modell
GLQ	General Lombard Quotient
LRM	Lineares Regressionsmodell

Abkürzungsverzeichnis

MAUS	<i>Munich AUtomatic Segmentation system</i>
ms	Millisekunden
NC	<i>noise cancellation</i>
NC+S	<i>noise cancellation</i> und <i>noise suppression</i>
NS	<i>noise suppression</i>
OLS	<i>Ordinary Least Squares</i>
RSS	<i>Residual Sum of Squares</i>
S.	Seite
SAMPA	<i>Speech Assessment Methods Phonetic Alphabet</i>
Sek.	Sekunden
SNR	<i>Signal-to-Noise Ratio</i> , Signal-Rausch-Verhältnis
sog.	sogenannter
SPID	<i>Spontaneous In-Car Dialogues</i>
SPL	<i>Sound-Pressure Level</i> , Schalldruckpegel
Sprachprod.	Sprachproduktion
u. a.	unter anderem
usw.	und so weiter
wahrgen.	wahrgenommene
vgl.	vergleiche
VOT	<i>Voice-Onset-Time</i>
vu	<i>volume units</i>
z. B.	zum Beispiel
z. T.	zum Teil

1 Einleitung

1.1 Motivation des Forschungsvorhabens

Die vorliegende Dissertation leistet einen wesentlichen Beitrag zur Verbesserung der lautsprachlichen Kommunikation im Fahrzeug. Dabei geht es um die Bewertung eines Innenraumkommunikationssystems im Auto (*In-Car Communication system*, ICC-System), das die Kommunikation der Passagier*innen¹ untereinander verbessern soll (vgl. [46, S. 1, 8], [24, S. 5], [26, S. 185, 197], [48, S. 97f., 116] und [71, S. 1]).

Diese Arbeit wird im Zuge eines interdisziplinären Forschungsprojekts zwischen der Allgemeinen und Vergleichenden Sprachwissenschaft und der Technischen Fakultät der Christian-Albrechts-Universität zu Kiel (CAU) durchgeführt, sodass wissenschaftliche und methodische Erkenntnisse und Erfahrungen ausgetauscht und kombiniert werden können. Zur Evaluierung eines Sprachverbesserungssystems ist es von Vorteil, verschiedene Qualitätskriterien unterschiedlicher wissenschaftlicher Fachgebiete zusammenzufassen, welche sich auf die Eigenschaften des Autos, die Charakteristika des Sprachverbesserungssystems sowie die Kommunikationsqualität beziehen (vgl. [26, S. 200ff.]). Kommunikationsunterstützende Systeme wurden bisher vorwiegend unter technischen Gesichtspunkten bewertet. Daneben fanden u. a. Hörtests zur Verständlichkeit einzelner Wörter statt (vgl. [65], [66] und Kapitel 2.2). Die lautsprachliche Kommunikation zwischen Sprecher*innen besteht jedoch nicht nur in der Wortverständlichkeit, sondern zeichnet sich darüber hinaus durch zahlrei-

¹Im Folgenden werden Personenbezeichnungen, bei denen das Geschlecht nicht bekannt ist, entweder durch Beidnennung der Geschlechter (Sprecherin bzw. Sprecher), genusunmarkierte Formen (Sprechende) oder durch die Verwendung eines Gender-Asterisk (Sprecher*innen) dargestellt. Durch diese Methoden soll die Verwendung gendergerechter Sprache ermöglicht werden, ohne den Textfluss für die Leser*innen wesentlich zu beeinflussen. Diese Vorgehensweise bezieht sich nur auf deutsche Wörter und nicht auf englische, auch wenn letztere in den deutschen Text integriert werden.

1 Einleitung

che Komponenten, wie Pausen, Wechsel der Sprecher*innen (*Turn-Taking*), Wortreduktion, Morphosyntax, Prosodie und deren Schnittstellen, aus. Daher ist weiterhin unklar, welche Kommunikationsprobleme im Auto entstehen und wie ICC-Systeme darauf abgestimmt werden müssen:

„Technische Qualitätsmaße, die bereits für die Evaluation [kommunikationsverbessernder Systeme] herangezogen werden, berücksichtigen zwar die Qualität des akustischen Signals, nicht aber die der eigentlichen Kommunikation im Sinne eines dynamischen Prozesses menschlich-sozialer Interaktion.“ ([26, S. 202])

Aus diesem Grund werden in der vorliegenden Studie akustisch-phonetische Eigenschaften von Dialogen im Auto, welche bei Fahrgeräuschen unterschiedlicher Geschwindigkeiten stattfinden, deren Wahrnehmung sowie die Effekte eines ICC-Systems auf diese Parameter untersucht. Durch die Analyse der zu diesem Zweck aufgenommenen Sprachsignale erfolgt eine Evaluierung des verwendeten ICC-Systems. Die Ergebnisse der Dissertation tragen dazu bei, neue interdisziplinäre Standards zur Bewertung von ICC-Systemen zu entwickeln und die Entstehung dieser Systeme dauerhaft und effektiv zu verbessern. Durch die Evaluierung eines ICC-Systems kann dieses zudem mit anderen Systemen verglichen und dadurch wiederum optimiert werden (vgl. [71, S. 1]).

Nach Jaschke (2012) werden durch ein ICC-System sowohl verschiedene Frequenzbänder als auch der gesamte Pegel des Gesprochenen an die vorhandene Geräuschkulisse angepasst. Um bandspezifisch gegenzusteuern, sollten die Sprachmodifikationen in unterschiedlichen Fahrsituationen bekannt sein. (vgl. [24, S. 5, 64]) Ein besseres Verständnis der im Fahrzeug ohne und mit ICC-System stattfindenden sprachlichen Veränderungen kann demnach einen großen Gewinn für die Weiterentwicklung und Verbesserung von Innenraumkommunikationssystemen bedeuten. Das für diese Untersuchung entstandene Sprachkorpus (*Spontaneous In-Car Dialogues*, SPID) soll zudem öffentlich zugänglich gemacht werden und dadurch der ICC-Forschung nachhaltig von Nutzen sein.

1.2 Fragestellung und Hypothesen

Die Ergebnisse der vorliegenden Dissertation sollen Aufschluss darüber geben, inwieweit ein ICC-System die lautsprachliche Kommunikation der Passagier*innen eines fahrenden Autos tatsächlich verbessert. Dafür wird zunächst festgestellt, welche Eigenschaften die Sprachproduktion aufweist, wenn das Auto still steht und der Motor ausgeschaltet ist. Da der Innenraum eines Fahrzeugs auch bei Stille eine eigene Kommunikationsbedingung darstellt, wird davon abgesehen, hierfür die Ergebnisse phonetischer Forschungen zu Dialogen im Labor als Basis heranzuziehen. Stattdessen wurden Sprachsignale im Auto aufgenommen, worauf im Folgenden näher eingegangen wird. Anschließend wird die Modifikation der Sprachsignale bei unterschiedlichen Geräuschszenarien untersucht, die verschiedene Fahrgeschwindigkeiten widerspiegeln. Die Veränderung der Sprachproduktion, welche bei lauten Hintergrundgeräuschen auftritt, wird als Lombard-Effekt beschrieben (vgl. [42, S. 1]).

Zu diesem Zweck wurden Sprachsignale unter verschiedenen Aufnahmebedingungen erhoben und als Korpus spontan gesprochener Dialoge im Auto (*SPontaneous In-car Dialogues*, SPID) zusammengefasst. Dieses enthält Sprachaufnahmen, welche in einem Fahrzeug ohne und mit ICC-System jeweils in Stille und mit den Fahrgeräuschen von 50 km/h und 130 km/h durchgeführt wurden.² Die oben beschriebene Fragestellung bzgl. der Verbesserung der lautsprachlichen Kommunikation im fahrenden Auto durch ein ICC-System ist weiterhin in drei verschiedene Teilfragen gegliedert. Die ersten zwei Teilfragen werden anhand akustischer und phonetischer Analysen der Sprachaufnahmen des SPID-Korpus beantwortet. Die Untersuchung der Sprachsignale findet dabei hinsichtlich der Parameter der durchschnittlichen Intensität, der durchschnittlichen Grundfrequenz, der Signaldauer und der Sprachproduktionsleistung statt:

1. Wie verändert sich die Kommunikation im Auto unter Lombard-Bedingungen, d. h., welche Modifikationen der Sprachproduktion finden unter dem Einfluss verschiedener Fahrgeräusche statt?

²Einem *online*-Artikel der „Auto Bild“ zufolge herrscht im Innenraum einer Mercedes S-Klasse bei 50 km/h ein Schallpegel von 54 dB und bei 130 km/h ein Schallpegel von 66 dB [81]. Durch diese Richtwerte wird deutlich, dass das Fahrzeug von Stille zu 50 km/h und von 50 km/h zu 130 km/h wesentlich lauter wird.

1 Einleitung

2. Wie verändert sich der Lombard-Effekt durch den Einsatz eines ICC-Systems?
 - 2.1 Kann die Verwendung eines ICC-Systems eine Rückverschiebung der sprachlichen Modifikationen hin zu einer Situation ohne Lombard-Umgebung bewirken?
 - 2.2 Führt der Einsatz eines ICC-Systems zu einer Verbesserung der Kommunikation? Kann sich ein ICC-System störend auf die Kommunikation der Dialogpartner*innen auswirken?

Ergänzend wurde mit Ausschnitten des SPID-Korpus ein Perzeptionsexperiment durchgeführt, um die Wahrnehmungen von Hörer*innen zu ermitteln, die nicht an der Erstellung des Sprachkorpus beteiligt waren. Anhand der Methode des Semantischen Differentials wurde die Wahrnehmung unterschiedlicher Eigenschaften der Sprechstimme abgefragt. Durch die subjektiven Bewertungen der Proband*innen sollen die Resultate der vorangegangenen akustischen Untersuchung verifiziert werden. Zudem wird überprüft, ob eine in der akustischen Analyse festgestellte Verbesserung der Kommunikationssituation durch das ICC-System von den Hörer*innen wahrgenommen wird oder ob die Hörer*innen auf emotionaler Ebene eine Verbesserung der Kommunikationssituation wahrnehmen, die in der akustischen Analyse zuvor nicht erkannt wurde. Es sollen dementsprechend Antworten auf eine dritte Teilfrage gefunden werden:

3. Wie hört sich die von Hintergrundgeräuschen befreite Lombard-Sprache für Hörer*innen an, die nicht an den Dialogen beteiligt waren?
 - 3.1 Ist in den geräuschbefreiten Sprachsignalen ein Lombard-Effekt wahrnehmbar, wenn diese bei erhöhter Lärmbelastung aufgenommen wurden?
 - 3.2 Werden aufgrund der unterschiedlichen ICC- und Geräuschbedingungen bestimmte emotionale Assoziationen bei den Hörer*innen hervorgerufen?

Im Vorwege der Untersuchung sind bereits verschiedene Hypothesen hinsichtlich der Ergebnisse der akustisch-phonetischen sowie der perzeptiven Signalanalysen formuliert worden. Vorangegangene Untersuchungen zum Lombard-Effekt stellten fest, dass bei lauten Hintergrundgeräuschen Steigerungen der Stimmintensität und der Grundfrequenz stattfinden. Des Weiteren lösen Proband*innen ihre Sprechaufgaben

schneller und steigern ihre Sprachproduktionsleistung. Diese vier Parameter der Intensität, der Grundfrequenz, der Signaldauer (Dauer eines Sprachsignals) und der Sprachproduktionsleistung, welche in Kapitel 3.1 ausführlich beschrieben sind, wurden in früheren Studien als Eigenschaften des Lombard-Effekts erkannt und werden daher auch in der vorliegenden Forschungsarbeit zur Untersuchung des Lombard-Effekts analysiert. Aufgrund dieser Ergebnisse vorangegangener Studien wird erwartet, dass durch die Präsenz bzw. Zunahme von Fahrgeräuschen ebenfalls Steigerungen der Stimmintensität und der Grundfrequenz stattfinden, die Sprechaufgaben der Proband*innen schneller gelöst bzw. die Dialoge kürzer sein werden und die Sprachproduktionsleistung eine Zunahme erfährt.

Die Verwendung eines ICC-Systems könnte Probleme der verbalen Kommunikation im Auto lösen. Wenn laute Umgebungsgeräusche bei den Sprecher*innen Stress hervorrufen (vgl. Kapitel 2), kann der Lombard-Effekt als Ausdruck dieses erhöhten Stresses angesehen werden. Eine Verbesserung der Kommunikationssituation im Fahrzeug würde daher mit einer Reduktion des Stresslevels einhergehen und damit zu einer Abnahme bzw. zu einem Ausbleiben des Lombard-Effekts führen. Es wird angenommen, dass die Modifikationen des Sprechverhaltens bei lauten Fahrgeräuschen durch die Aktivierung des ICC-Systems zurückgehen und die Dialogpartner*innen miteinander kommunizieren, als würden sie sich in einer stillen Fahrzeugumgebung befinden. D. h., das ICC-System trägt vermutlich dazu bei, dass im Auto trotz des umgebenden Fahrzeumlärms wie in einer stillen Situation gesprochen wird. Zwischen der Aktivierung und Deaktivierung des ICC-Systems in Stille wird hingegen keine bedeutende Veränderung der Sprachproduktion erwartet.

Weiterhin wird angenommen, dass sprachliche Veränderungen der Dialogpartner*innen, welche sich bei unterschiedlichen Fahrgeräuschen und der Aktivierung des ICC-Systems einstellen, von unabhängigen Zuhörer*innen wahrgenommen werden. Dabei liegt die Vermutung nahe, dass sich die Sprache der Proband*innen bei lauten Hintergrundgeräuschen ebenfalls lauter und dazu gestresster bzw. angestrenzter anhört. Bei einer Verbesserung der Kommunikation durch das ICC-System wird angenommen, dass die Sprecher*innen im Gegensatz dazu wieder leiser und entspannter klingen. Lombard-Sprache ist darüber hinaus verständlicher als im Stillen gesprochene Sprache (vgl. Kapitel 3.1). Daher wird vermutet, dass die Sprache mit ICC-System als Ausdruck eines reduzierten Lombard-Effekts wiederum weni-

ger verständlich ist und sich dies im Perzeptionsexperiment widerspiegelt. Wenn Lombard-Sprache ein Ausdruck physischen Stresses ist, wird eine geringere Sprachverständlichkeit mit ICC-System als Zeichen eines reduzierten Stressniveaus und einer verbesserten, d. h. entspannteren bzw. artikulatorisch weniger angestregten, Kommunikationssituation interpretiert.

1.3 Aufbau der Arbeit und Methoden

Einführend werden in Kapitel 2 zunächst verschiedene Faktoren beschrieben, welche die lautsprachliche Kommunikation im Autoinnenraum negativ beeinflussen. Anschließend wird auf die Funktionsweise von ICC-Systemen sowie auf unterschiedliche Ansätze der Evaluierung solcher Systeme eingegangen. Weiterhin erfolgt eine Beschreibung des ICC-Systems, welches in der vorliegenden Arbeit zur Generierung des Sprachkorpus SPID verwendet wurde. Dieses ICC-System wurde zuvor von der Arbeitsgruppe für Digitale Signalverarbeitung und Systemtheorie an der Technischen Fakultät der CAU Kiel entwickelt. In diesem und allen anderen Kapiteln der vorliegenden Arbeit finden sich Literaturhinweise und Querverweise innerhalb der Arbeit jeweils am Ende eines Satzes (vor dem Punkt) oder eines Sinnabschnitts (hinter dem Punkt), welcher aus mehreren Sätzen sowie einem vollständigen Absatz bestehen kann.

Kapitel 3 beinhaltet eine Definition des Lombard-Effekts sowie eine Beschreibung verschiedener Parameter dieses Phänomens. Dabei werden die Forschungsergebnisse unterschiedlicher Studien zusammengetragen und diskutiert. Klassische Eigenschaften des Lombard-Effekts, d. h. eindeutig nachzuweisende Parameter dieses Phänomens, sind u. a. die Steigerung von Intensität und Grundfrequenz bei lauten Hintergrundgeräuschen. Zudem wird die in der Forschungsliteratur wiederholt auftretende Diskussion zu den Ursachen des Lombard-Effekts aufgegriffen. Die Frage, ob es sich hierbei um eine bewusst gesteuerte Veränderung im Kommunikationsverhalten der Sprecher*innen oder um einen Reflex handelt, wird von verschiedenen Forscher*innen mit unterschiedlichen Argumenten zu beantworten versucht. Weiterhin wird darauf eingegangen, welche Methoden zur Untersuchung des Lombard-Effekts in der bisherigen Forschung angewendet wurden und welche Vor- und Nachteile sich dadurch ergaben. Durch unterschiedliche Vorgehensweisen bei der Aufnahme und

Untersuchung von Sprachsignalen, wie z. B. die Verwendung von Kopfhörern oder die Unterscheidung der Proband*innen nach ihrem Geschlecht, kann das Ausmaß des Lombard-Effekts deutlich beeinflusst werden.

Zur Untersuchung der Sprachkommunikation im fahrenden Auto ohne und mit ICC-System wurde eine akustische Umgebungssimulation verwendet, welche ebenfalls von der Arbeitsgruppe für Digitale Signalverarbeitung und Systemtheorie an der Technischen Fakultät der CAU Kiel entwickelt wurde (vgl. [26, S. 185]). Diese wurde im Vorwege der vorliegenden Arbeit um eine visuelle Umgebungssimulation erweitert (vgl. [76]). Kapitel 4 widmet sich der Beschreibung dieser Simulation, welche sich durch die Besonderheit auszeichnet, dass die in der Umgebungssimulation bei Lärm aufgenommenen Sprachsignale anschließend von den Hintergrundgeräuschen befreit werden konnten. Zudem wird auf die Realitätsnähe dieser Fahrzeuginnenraumsimulation eingegangen, welche in einer früheren Studie zum Lombard-Effekt bereits getestet wurde.

Die Aufnahmen des Sprachkorpus SPID wurden mit acht weiblichen und acht männlichen Proband*innen in der akustischen und visuellen Umgebungssimulation der CAU Kiel angefertigt. Als Sprechaufgabe diente das sogenannte *Map-Task* Szenario, welches in Kapitel 5.1 beschrieben wird. Die einzelnen Kartenaufgaben (*Maps*) sind im Anhang dieser Arbeit einsehbar (vgl. Anhang A.1). Kapitel 5 beinhaltet zudem eine ausführliche Darstellung der Experimentbedingungen, wie z. B. die verwendeten Fahrgeräusche, die Proband*innen, die Instruktionen vor und während der Aufnahmen, die Durchführung des Produktionsexperiments und die Geräuschbefreiung der Sprachsignale. Daran anschließend werden die Transkriptionen und Annotationen der Sprachaufnahmen in Kapitel 5.7 beschrieben. Dabei wird u. a. deutlich, welche Probleme sich bei der semi-automatischen Annotation mit der *Software* MAUS ergeben und welche Methoden entwickelt und angewandt wurden, um diese zu beheben.

In Kapitel 6 erfolgt zunächst eine Zusammenfassung der Antworten des Fragebogens, den die Sprecher*innen nach dem Produktionsexperiment ausfüllten, sowie eine Übersicht der im *Map-Task* Szenario von den Teilnehmer*innen artikulierten Zielwörter. Anschließend wird die deskriptive und inferenzstatistische Auswertung der Analyseparameter durchschnittliche Intensität, durchschnittliche Grundfrequenz, Signaldauer und Sprachproduktionsleistung anhand der aufgenommenen

1 Einleitung

Sprachsignale vorgenommen. Dazu wird eine Regressionsanalyse in der Programmiersprache R durchgeführt (vgl. [94]).

Methode, Vorbereitung, Durchführung und Analyse des Perzeptionsexperiments werden in Kapitel 7 vorgestellt. Durch dieses ergänzende Experiment soll eine umfassendere Beantwortung der Frage möglich sein, ob sich die lautsprachliche Kommunikation im Auto bei lauten Fahrgeräuschen durch die Anwendung des ICC-Systems verbessern lässt. Dazu wurden Ausschnitte der Sprachsignale des SPID-Korpus anhand eines *Online*-Fragebogens von Hörer*innen bewertet, die nicht an den Aufnahmen des Sprachkorpus beteiligt waren. Die Beurteilung des Gehörten erfolgte mithilfe eines Semantischen Differentials, indem verschiedene Adjektive auf siebenstufigen Skalen bewertet wurden (vgl. Tabelle 1.1). Die Gründe für die Wahl der Adjektive im Semantischen Differential sind in Kapitel 7.1 ausführlich dargestellt. Die Resultate des Perzeptionsexperiments dienen einer subjektiven Bestätigung der Ergebnisse der akustisch-phonetischen Analyse aus Kapitel 6.

Tabelle 1.1: Die Adjektive des Semantischen Differentials im Perzeptionsexperiment.

unfreundlich - freundlich
emotional - sachlich
gehemmt - ungehemmt
steif - locker
angestrengt - entspannt
stockend - flüssig
laut - leise
verständlich - unverständlich

In Kapitel 8 werden zunächst die Ergebnisse von Produktions- und Perzeptionsexperiment zusammengefasst und entsprechend der in Kapitel 1.2 beschriebenen drei Teilfragen interpretiert. Weiterhin erfolgt eine Auswertung der Forschungsarbeit und ihrer Resultate im Hinblick auf die übergeordnete Fragestellung (vgl. Kapitel 8.2). Methodenkritik und Ausblick sollen anschließend dazu beitragen, auf den vorliegenden Erkenntnissen aufzubauen. Weiterführende Forschungen können das für diese Arbeit erstellte Sprachkorpus SPID als Grundlage nutzen, um zahlreiche weitere linguistische Parameter in den Sprachsignalen der Sprecher*innen zu untersuchen.

2 Sprachkommunikation im Auto

Der moderne Kommunikationsort des Autos ist durch verschiedene Eigenschaften geprägt, die sich erschwerend auf die lautsprachliche Kommunikation der Passagier*innen auswirken. Dabei spielen vor allem die unterschiedlichen Geräuschkulissen während der Fahrt eine entscheidende Rolle. Lärm ist ein Stressfaktor ([28, S. 14]), der zu einer schlechteren Verständlichkeit von Sprache führt und die Kommunikation im Auto anstrengend werden lässt. Im Innenraum eines Fahrzeugs üben zudem weitere Faktoren einen nachteiligen Einfluss auf die Dialoge zwischen den Mitfahrenden aus. Schalldämpfende Materialien im Auto, welche das Lärmniveau im Fahrzeuginnenraum verringern sollen, führen dazu, dass die Verständlichkeit des Gesagten vermindert wird. Weiterhin findet während einer Unterhaltung häufig kein Blickkontakt statt. Dies ist z. B. der Fall, wenn eine am Dialog beteiligte Person vorne sitzt und die andere hinten oder wenn die Sprecher*innen ihre Blicke auf das Verkehrsgeschehen richten müssen. Dadurch kann die reduzierte Verständlichkeit nicht durch visuelle Informationen kompensiert werden, welche Hinweise auf artikulatorische Vorgänge liefern. Im Auto sprechen vorne sitzende Passagier*innen zudem in Richtung Windschutzscheibe, wodurch sie für die hinten Sitzenden nur schwer verständlich sind. Dabei sind besonders diejenigen Sprecher*innen benachteiligt, die unmittelbar hinter einer sprechenden Person sitzen. In größeren Fahrzeugen, wie z. B. Bussen, kann zudem eine größere Distanz zwischen den Passagier*innen liegen, welche die Verständlichkeit der Sprachsignale zusätzlich abschwächt. (vgl. [26, S. 186f.], [46, S. 1, 4], [48, S. 97f., 107], [65, S. 547, 549], [66, S. 73f., 80] und [71, S. 1])

Die während einer Autofahrt auftretenden Hintergrundgeräusche entstehen u. a. aufgrund verschiedener Geschwindigkeiten sowie durch Reifengeräusche, Lüftung, Wind, Motor und den umliegenden Straßenverkehr. Dieser Lärm führt zu einem niedrigen SNR (*Signal-to-Noise Ratio*, Signal-Rausch-Verhältnis). Die vom Motor

hervorgerufenen Geräusche sind in einem Frequenzbereich unter 500 Hz am energiereichsten, was besonders bei hohen Fahrgeschwindigkeiten der Fall ist. Als Reaktion auf diese ungünstigen Kommunikationsbedingungen verwenden die Passagier*innen Lombard-Sprache. Zudem neigen sich hinten sitzende Dialogpartner*innen häufig nach vorne, während sich vorne sitzende Personen meist nach hinten umdrehen, wodurch der SNR um bis zu 20 dB gesteigert werden kann. Bei Frequenzen im Bereich von 1 kHz können die Sprecher*innen dadurch den SNR um Werte bis zu 15 dB verbessern. Dies ist jedoch nicht nur unbequem und ermüdend, sondern kann darüber hinaus auch die Sicherheit der Passagier*innen gefährden, wenn sich die Fahrerin oder der Fahrer des Autos an der Unterhaltung beteiligt. (vgl. [24, S. 17f., 31, 33], [26, S. 188f., 201], [46, S. 1], [48, S. 97f.], [65, S. 547, 550], [66, S. 74] und [71, S. 1])

2.1 *In-Car Communication systems*

Zur Verbesserung der lautsprachlichen Kommunikation im Fahrzeug dienen sogenannte Innenraumkommunikationssysteme (*In-Car Communication systems*, ICC-Systeme), welche die Sprache der Passagier*innen mittels Mikrofonen und Lautsprechern innerhalb des Fahrzeugs übertragen.³ Durch die Verstärkung der Sprecher*innensignale wird der SNR am Ohr der Zuhörer*innen gesteigert. Diese Systeme funktionieren in einer geschlossenen elektroakustischen Schleife. Dabei nehmen die Mikrophone auch einen Teil der Lautsprechersignale auf. Ist dieser Anteil nicht ausreichend klein, kommt es zu anhaltenden Heul- oder Pfeifgeräuschen. Um die Stabilitätsgrenze zu verbessern, werden verschiedene Methoden der Signalverarbeitung angewandt. Die Gesamtverzögerung (*overall delay*) eines ICC-Systems darf nach Schmidt und Haulick (2006) (vgl. [65, S. 562]) 10 ms nicht überschreiten, da die Lautsprechersignale sonst als Echos wahrgenommen werden. Nach Schmidt et al. (2012) können Zuhörer*innen die Schallquelle, die direkt von den Sprecher*innen ausgeht, und ihre Wiedergabe durch das ICC-System einzeln wahrnehmen, wenn diese Gesamtverzögerung 15-20 ms überschreitet (vgl. [66, S. 82, 104]). Hinzu kommt, dass Sprecher*innen die eigene Stimme selbst bei einer geringen Verstärkung durch

³Diese Systeme werden auch als „*car interior communication system*“ (vgl. [65, S. 548]) oder „*intercom system*“ (vgl. [65, S. 547]) bezeichnet. Darüber hinaus existieren Begriffe wie „*in-vehicle communication system*“ (IVC) und „*digital voice enhancement system*“ (DVE) (vgl. [66, S. 74]).

das ICC-System als Echo wahrnehmen, wenn das *delay* zu hoch ist (vgl. [66, S. 83]). Allerdings können die Sprecher*innen ihre Stimmen ebenfalls als Echos wahrnehmen, wenn das *delay* ausreichend gering ist, aber die Verstärkung zu groß (vgl. [66, S. 104]). Weiterhin verfügen ICC-Systeme über eine Erkennung der Sprechaktivität. Nur, wenn eine Person spricht, werden ihre Signale über diejenigen Lautsprecher wiedergegeben, welche sich in der Nähe der anderen Passagier*innen befinden. (vgl. [49], [65, S. 547f., 560ff., 578f., 592] und [66, S. 73ff., 79f.]

Aufgrund von ICC-Systemen findet im Fahrzeug eine Steigerung der Nachhallzeit⁴ statt. Je größer *delay* und Verstärkung durch das System sind, desto größer ist die Zunahme des Nachhalls. Wenn die Nachhallzeit, abhängig vom Fahrzeugtyp, auf über 80-140 ms zunimmt, beginnen die Passagier*innen, sich über die halligen Eigenschaften des Autos zu beschweren. Je nach Art des Fahrzeugs sollte die Nachhallzeit eines Autos mit zwei Sitzreihen bei aktiviertem ICC-System nicht mehr als 80-150 ms betragen. Die Qualität eines ICC-Systems ist umso besser, je weniger die Nachhallzeit durch das System gesteigert wird. (vgl. [66, S. 83, 101ff.]

Im Fahrzeug stellen die Sprechrichtungen von vorne nach hinten und von hinten nach vorne jeweils erschwerende Kommunikationsbedingungen dar. Da Fahrer*in und Beifahrer*in eines Autos jedoch nicht in Richtung der anderen Passagier*innen sprechen (vgl. [65, S. 549]), ist die Richtung von vorne nach hinten besonders zu berücksichtigen (vgl. [71, S. 2]). Bei Autos mit mehr als zwei Sitzreihen, wie z. B. großen Limousinen und Kleinbussen, sollte ein ICC-System nach Schmidt und Haulick (2006) (vgl. [65, S. 549]) beide Richtungen unterstützen. Weiterhin beschreiben sie, dass die Sprachverständlichkeit bei Stillstand meistens ausreichend sei.⁵ Ein *intercom system* lasse das Auto halliger klingen, wodurch die Kommunikationsqualität herabgesetzt werde. Während mittlerer oder hoher Fahrgeschwindigkeiten könne ein solches System die Sprachverständlichkeit jedoch erheblich verbessern. (vgl. [65, S. 549f.]) Schmidt und Haulick (2006) schreiben:

⁴In der Nachhallzeit fällt ein Schallereignis um 60 dB ab (T_{60}). Im Fahrzeug beläuft sich die Nachhallzeit auf etwa 30-70 ms. (vgl. [26, S. 187], [46, S. 4], [48, S. 107] und [66, S. 102])

⁵An dieser Stelle soll angemerkt werden, dass die Kommunikationssituation zwischen vorne und hinten sitzenden Personen jedoch auch im Stillstand bei ausgeschaltetem Motor eine besondere ist, wie zu Beginn von Kapitel 2 beschrieben wurde.

„The output gain of a car intercom system needs to be adjusted continuously according to the current driving situation. While only a moderate gain is required whenever the car is in low noise conditions, a large gain is required and more artifacts will be tolerated at high speed.“ [65, S. 561f.]

2.2 Evaluierung von ICC-Systemen

Zur Evaluierung von ICC-Systemen, d. h., um die Verbesserung von Sprachqualität und Sprachverständlichkeit durch ICC-Systeme zu überprüfen, wurden bereits unterschiedliche Tests durchgeführt. Dazu gehören subjektive Untersuchungen, wie *diagnostic* bzw. *modified rhyme tests*⁶ zur Analyse der Sprachverständlichkeit und *comparison mean opinion score tests* (CMOS)⁷ zur Untersuchung der Sprachqualität. Des Weiteren existieren unterschiedliche objektive und nicht-kommunikative Methoden, wie z. B. Untersuchungen des SNR. Schmidt und Haulick (2006) fanden anhand ihrer Tests heraus, dass Sprachverständlichkeit und Sprachqualität durch das von ihnen verwendete ICC-System bei einer Geräuschbelastung von 130 km/h verbessert werden können. (vgl. [65, S. 585ff., 596] und [66, S. 90, 93ff., 99, 105f.])

Nach Schmidt et al. (2012) gibt es im Gegensatz zu beispielsweise Freisprechtelefonen bislang noch keine standardisierten Methoden zur Evaluierung der Qualität von ICC-Systemen. Standardisierte Verfahren zur Qualitätsbewertung könnten jedoch sowohl zur Unterstützung der Konstruktion solcher Systeme als auch für den Vergleich unterschiedlicher ICC-Systeme nützlich sein. Dabei sei die Entwicklung einer automatischen Systembewertung sinnvoll, welche auf objektiven Kriterien beruht, da subjektive Tests sehr zeitaufwendig seien und für ein repräsentatives Ergebnis eine hohe Anzahl an Testpersonen benötigt werde, was wiederum recht teuer sei. Objektive Testmethoden hingegen wären leichter zu reproduzieren, zuverlässiger und weniger kostspielig. (vgl. [65, S. 585f., 589] und [66, S. 78, 92, 104ff.]) Schmidt et al. (2012) zufolge müssen zur Evaluierung eines ICC-Systems zwei übergeordnete Fragestellungen beantwortet werden, welche sich auf die Passagier*innen des Fahrzeugs

⁶Dabei ist eine Untersuchung der Verständlichkeit einzelner Silben oder auch eine Untersuchung einzelner Phoneme durch Reimworte wie *game* und *name* möglich (vgl. [65, S. 586] und [66, S. 93]).

⁷Bei diesen Tests werden bekannte Phrasen verwendet, damit die Proband*innen sich beim Hören auf Faktoren konzentrieren können, welche die Sprachqualität beeinflussen. Dazu gehören der Umfang des Nachhalls, die Natürlichkeit der Sprache oder Artefakte. (vgl. [65, S. 586] und [66, S. 93])

beziehen, um die Qualität des Systems beurteilen zu können: (vgl. [66, S. 104])

1. Verbessert das ICC-System die Sprachqualität für die ZuhörerIn bzw. den Zuhörer? (vgl. [66, S. 104])
2. Wird durch das ICC-System die Kommunikationsqualität für die SprecherIn bzw. den Sprecher reduziert? (vgl. [66, S. 104])

Auf die erste Frage können zwei Teilantworten folgen, welche sich auf die Verbesserung oder die Beeinträchtigung der Kommunikationsqualität für die zuhörende Person beziehen. Dabei kann einerseits die Verbesserung der Sprachverständlichkeit z. B. durch die Analyse des SNR in den Ohren der bzw. des Zuhörenden bestimmt werden. Andererseits kann die Verschlechterung der Sprachqualität z. B. durch die Ermittlung der Nachhallzeit, der Lautstärke des reproduzierten Sprachsignals und des durch das ICC-System hervorgerufenen *delays* überprüft werden. Die zweite Frage kann ebenfalls anhand verschiedener objektiver Messungen beantwortet werden. Dabei wird z. B. untersucht, inwieweit eine SprecherIn bzw. ein Sprecher das *Feedback* des eigenen Sprachsignals oder ein Echo der eigenen Stimme hört. Die genannten Tests sind objektiv, doch es wird hierbei kein kommunikatives Verhalten untersucht. Zudem sind verschiedene weitere Faktoren noch immer schwer messbar. Dazu gehören subjektive Höreindrücke, wie die Natürlichkeit von Sprache. Daher wird vorgeschlagen, subjektive und objektive Evaluierungen von ICC-Systemen in Beziehung zu setzen. (vgl. [66, S. 101, 104f.])

Theiß et al. (2014) stellen eine Evaluierungsstrategie für ICC-Systeme vor, die sich aus den drei Evaluierungsgruppen Fahrzeugeigenschaften, ICC-System-Eigenschaften und Kommunikationsqualität zusammensetzt, welche jeweils mehrere unterschiedliche Messungen beinhalten (vgl. [71, S. 1ff., 4]). Die zuletzt genannte Evaluierungsgruppe der Kommunikationsqualität ist Gegenstand der vorliegenden Forschungsarbeit (vgl. Kapitel 1.2).

2.3 ICC an der CAU Kiel

Das für die vorliegende Untersuchung verwendete ICC-System wurde an der Technischen Fakultät der CAU Kiel von der Arbeitsgruppe für Digitale Signalverarbeitung und Systemtheorie (DSS) (vgl. [84]) entwickelt. Es nimmt das Sprachsignal einer sprechenden Person über im Sicherheitsgurt angebrachte Mikrophone auf und gibt dieses mittels Lautsprechern wieder, die in Kopfnähe der anderen Passagier*innen im Fahrzeug angebracht sind. Hierbei finden verschiedene Methoden der Signalverarbeitung Anwendung, deren einzelne Komponenten u. a. in [26], [46], [48] und [66] ausführlich beschrieben sind. Das vorgestellte ICC-System wurde in die *Software* KiRAT (Kiel *Real-Time Audio/Application Toolkit*) implementiert, die ebenfalls von der Arbeitsgruppe DSS entwickelt wurde. KiRAT wird weiterhin auch für andere Aspekte der Untersuchung von Sprachsignalen im Fahrzeug verwendet, wie z. B. zur Kalibrierung der Fahrzeuggeräusche für die akustische Umgebungssimulation (vgl. Kapitel 4) und zur Durchführung von Sprachaufnahmen in dieser Simulation (vgl. [34, S. 4, 37]). Abbildung 2.1 zeigt beispielhaft einige der Mikrophone, welche im Testauto mit dem System verbunden sind. Die *Hardware* zur Signalverarbeitung befindet sich im Kofferraum des entsprechenden Fahrzeugs. (vgl. [24, S. 5, 33], [26, S. 186, 197], [37, S. 100], [38, S. 479], [39, S. 1485], [46, S. 1, 8], [47, S. 4], [48, S. 97ff., 116], [49] und [71, S. 1])



Abbildung 2.1: Ein mit dem vorgestellten ICC-System ausgestattetes Testauto der CAU Kiel. Sichtbare Mikrophone des Systems sind rot eingekreist. (vgl. [48, S. 99])

Aufgrund der oben genannten verstärkten Kommunikationsprobleme im Auto bei einer Sprechrichtung von vorne nach hinten unterstützt das in dieser Forschungsarbeit verwendete ICC-System zum größten Teil diese und nur geringfügig die entgegengesetzte Richtung (vgl. [38, S. 480] und [39, S. 1485]). Weiterhin ist das verwendete System aufgrund der schlechten Kommunikationsbedingungen im Fahrzeug, welche z. B. durch fehlenden Blickkontakt und schalldämpfende Materialien (vgl. Kapitel 2) auch ohne zusätzliche Hintergrundgeräusche vorherrschen, bereits im stillen Innenraum des Fahrzeugs aktiv, steigert seine Funktion jedoch mit zunehmender Geräuschbelastung.

Eine Unterstützung der Kommunikationssituation durch das ICC-System findet in Abhängigkeit zur vorherrschenden Geräuschkulisse statt. Diese ist bei Stillstand des Fahrzeugs fast nicht vorhanden und steigert sich bei zunehmender Fahrgeschwindigkeit. Durch die Signalverarbeitung des ICC-Systems werden u. a. Umgebungsgeräusche, welche die lautsprachliche Kommunikation beeinträchtigen, entfernt. Des Weiteren wird ein Hochpassfilter angewandt, da die Sprachsignale in einer Fahrzeugumgebung in niedrigen Frequenzbereichen meist vom vorherrschenden Umgebungslärm überlagert werden. Abhängig von den Präferenzen der Anwender*innen und der Art des Fahrzeugs sollte die Grenzfrequenz des Hochpassfilters zwischen 80 Hz und 300 Hz gewählt werden. Die darunter liegenden Frequenzen schlechter Signalqualität werden entfernt. Zudem absorbieren die meisten Materialien zur Auskleidung von Innenräumen hochfrequente Schallenergien stärker als niederfrequente. Deshalb ist es zur Verbesserung der Sprachverständlichkeit wichtiger, mittlere und hohe Frequenzen zu verstärken als niedrige. (vgl. [26, S. 186, 197f.], [46, S. 1], [65, S. 553, 562, 578], [66, S. 84] und [71, S. 1])

3 Der Lombard-Effekt

Wenn Sprecher*innen lauten Umgebungsgeräuschen ausgesetzt sind, verändern sie ihre Sprachproduktion (vgl. [42, S. 1]). Der Name dieses Phänomens geht auf den französischen Hals-Nasen-Ohren-Arzt Étienne Lombard (1869-1920) zurück, welcher den Effekt im Jahre 1911 erstmals beschrieb (vgl. [6, S. 46], [26, S. 185], [36, S. 414], [37, S. 98], [42, S. 1] und [67, S. 191]). Er stellte fest, dass Sprecher*innen ihre Stimmintensität steigern, wenn sie während eines Gesprächs mit Lärm konfrontiert werden, und ihre vorherige Lautstärke wiederherstellen, wenn die Hintergrundgeräusche enden. Auch das eingeschränkte Hören der eigenen Stimme,⁸ z. B. beim Tragen von Ohrstöpseln, kann zu einer Erhöhung der Stimmintensität führen, da hierdurch die Selbstkontrolle der sprechenden Person eingeschränkt wird. (vgl. [40, S. 618f., 623], [41, S. 677], [37, S. 98] und [67, S. 191]) Sinnott (1975) betont die Verbindung zwischen menschlichem Hören und Sprechen, da Sprecher*innen ihre Lautstärke steigern, wenn Schwierigkeiten beim Verstehen der eigenen Stimme bestehen (vgl. [68, S. 412]). Der Lombard-Effekt bzw. Lombard-Reflex wird u. a. auch als „*hair dryer*“ *effect*“ bezeichnet (vgl. [68, S. 412]). Eine künstliche Steigerung des *Sidetones* führt hingegen dazu, dass Proband*innen leiser sprechen (vgl. [40, S. 618, 621ff.]).

Seit den Studien Étienne Lombards haben sich viele Forscher*innen diesem Phänomen gewidmet und erkannten, dass durch laute Umgebungsgeräusche nicht nur die Intensität des Sprachsignals modifiziert wird, sondern zahlreiche weitere Parameter verändert werden (vgl. [36, S. 414] und [37, S. 98f.]). Diese betreffen nicht nur akustisch-phonetische Eigenschaften des Sprachsignals, sondern auch andere Bereiche, wie eine gesteigerte Häufigkeit des Blickkontakts der Gesprächspartner*innen (vgl. [53, S. 1329, 1332]).

⁸In diesem Zusammenhang wird die eigene Stimme einer Sprecherin bzw. eines Sprechers häufig als *Sidetone* bezeichnet (vgl. [41, S. 677] und [67, S. 191]).

In Kapitel 3.1 erfolgt zunächst eine Übersicht zu verschiedenen Parametern des Lombard-Effekts, woraufhin in Kapitel 3.2 auf die Ursachen des Phänomens eingegangen wird. Dabei steht vor allem die Frage im Vordergrund, ob es sich um einen unbewussten Reflex oder eine von Sprecher*innen bewusst durchgeführte Modifikation der Sprachproduktion handelt. Dreher und O’Neill (1957) weisen darauf hin, dass eine Sprecherin bzw. ein Sprecher die eigene Stimmintensität unbewusst steigert, um den bestehenden Umgebungslärm zu kompensieren (vgl. [12, S. 1320]).

Um festzustellen, ob Personen beim Testen ihres Hörvermögens simulieren, wurde der sogenannte „*voice-reflex*-Test“ bzw. „Lombard-Test“ entwickelt (vgl. [40, S. 619] und [12, S. 1320f.]). Laut Stach (2003) wird einer sprechenden Patientin bzw. einem sprechenden Patienten dabei weißes Rauschen ins Ohr gespielt und festgestellt, ob sich die Intensität des Sprachsignals verändert (vgl. [69, S. 157]). Verschiedene medizinische Wörterbücher und Lexika beschreiben zudem eine als „Lombard-Versuch“ oder „Lombard-Test“ bezeichnete Untersuchung, anhand derer festgestellt wird, ob eine Person an Schwerhörigkeit leidet (vgl. [63, S. 1269] und [78, S. 1260]):

„**Lombard-Test** zur Überprüfung einer behaupteten beidseitigen Taubheit: während der Untersuchte einen Text laut vorliest, wird über Kopfhörer ein starkes Geräusch zugeschaltet. Wegen Wegfall der Eigenkontrolle verstärkt der normal Hörende seine Stimme; bei Ertaubung hingegen erfolgt keine Veränderung des Leseflusses.“ ([2, S. 205])

An anderer Stelle wird der „Lombard-Versuch“ als „Simulationsprüfung bei angebl. einseitiger Ertaubung“ beschrieben ([78, S. 1260]). Daneben gibt es weitere Tests zur Bestimmung einer einseitigen oder beidseitigen Taubheit, wie den „Lee-Test“ und den „Stenger-Versuch“ (vgl. [2, S. 205] und [78, S. 2008]). In diesem Zusammenhang wird häufig der Begriff des „Lombard-Zeichens“ genannt (vgl. [63, S. 1269]), welcher folgendermaßen beschrieben wird:

„**Lombard-Zeichen:** (engl.) *Lombard's sign*; Verstärkung der Sprechlautstärke bei einseitiger Schwerhörigkeit*, wenn die Eigenkontrolle durch das normal hörende Ohr mit Békány*-Lärmtrommel unterbrochen wird.“⁹ ([2, S. 1242])

In Kapitel 3.3 werden zunächst bisherige Vorgehensweisen bei der Untersuchung des Lombard-Effekts dargestellt und anschließend beschrieben, durch welche Faktoren das Ausmaß des Phänomens, d. h. die Stärke der bei Lärm stattfindenden Modifikationen im Sprachsignal, beeinflusst wird. Ein Überblick zu den unterschiedlichen Eigenschaften von Lombard-Sprache, den kontrovers diskutierten Ursachen des Phänomens sowie den Ergebnissen vorangegangener Untersuchungen ist u. a. auch in [26, S. 185f., 190], [34, S. 5ff.], [36, S. 414] und [37, S. 98ff.] zu finden. Kapitel 3.4 beschreibt die in der vorliegenden Arbeit verwendete Arbeitsdefinition des Lombard-Effekts.

3.1 Eigenschaften

Eine Untersuchung von Garnier et al. (2006) ergab, dass die Vokalintensität bei lauten Hintergrundgeräuschen stärker ansteigt als die Intensität von Konsonanten (vgl. [17, S. 2248]). Lane et al. (1970) zeigen auf, dass sowohl Veränderungen in der Verstärkung des *Sidetones* als auch im Pegel der Umgebungsgeräusche etwa denselben Effekt auf die Stimmproduktion der Sprecherin bzw. des Sprechers besitzen. Dabei sei die Veränderung in der Stimmintensität nur etwa halb so groß wie die Veränderung im Pegel des modifizierten Richtwerts. (vgl. [40, S. 618ff.]) Aus den Ergebnissen vorangegangener Studien schließen sie:

„[...] speakers compensate only halfway (in decibel units) for increases in the ambient noise during communication, just as they compensate only halfway for decreases in their signal level.“ ([40, S. 623])

⁹Die Békány-Lärmtrommel ist definiert als:

„(Robert B., Otol., Wien, Uppsala, 1876-1936): (engl.) *Békány's noise box*; historischer Apparat zur Erzeugung von Lärm i. R. von Hörprüfungen*; durch Ausschaltung der akust. Wahrnehmung eines Ohrs ist die isolierte Prüfung des Gegenohrs möglich. Vgl. Lombard-Zeichen.“ ([2, S. 234])

Die prozentuale Stimmintensitätserhöhung pro dB ansteigenden Hintergrundlärms fällt jedoch in verschiedenen Studien unterschiedlich aus (vgl. [12, S. 1322f.] und [24, S. 64]).

Sprecher*innen steigern bei lauten Hintergrundgeräuschen ihre Stimmintensität bzw. Lautstärke. Dieses Ergebnis wurde in zahlreichen Studien bestätigt (vgl. [7, S. 361f., 365], [8, S. 4], [14, S. 2701, 2704], [20, S. 364ff.], [42, S. 2f.], [53] und [70, S. 919f., 922]). Neben einer Steigerung der Sprachlautstärke beinhaltet der Lombard-Effekt eine Erhöhung der Grundfrequenz und eine Reduktion des *spectral tilt* (vgl. [3, S. 907, 909ff.], [13, S. 2611f.], [14, S. 2701, 2704], [17, S. 2246f.], [24, S. 42ff., 47ff., 51ff., 55ff., 58ff., 63], [45, S. 3261], [53] und [70, S. 920ff.]). Des Weiteren bewirkt der Lombard-Effekt eine Anhebung von F1 (vgl. [3, S. 910f.]) und F2 (vgl. [17, S. 2248] und [42, S. 4f.]). Eine dahingehende Tendenz wird auch bei Davis et al. (2006) beschrieben, wobei zudem eine starke Anhebung der Frequenzen von F3 festgestellt wird (vgl. [7, S. 362, 365]). Hinsichtlich des zweiten Formanten werden unterschiedliche Forschungsergebnisse beschrieben. Dieser nimmt, teilweise abhängig vom betroffenen Vokal, ab, steigt an oder bleibt unverändert (vgl. [3, S. 910f.] und [45, S. 3261]). Summers et al. (1988) vermuten aufgrund ihrer Ergebnisse, dass die bei Lombard-Sprache stattfindenden Veränderungen von F0 und F1 einander bedingen könnten, wobei eine Erhöhung der Grundfrequenz mit einer Anhebung des ersten Formanten einherginge (vgl. [70, S. 921f.]). Lau (2008) stellte zudem für das Amerikanische Englisch fest, dass die *Voice-Onset-Time* (VOT)¹⁰ stimmhafter Plosive bei lauten Hintergrundgeräuschen abnimmt (vgl. [42, S. 5f.]).

Bei Lombard-Sprache findet weiterhin eine Zunahme der Wortdauer statt (vgl. [42, S. 3f.] und [70, S. 920, 922]). Junqua (1993) beschreibt, dass Konsonanten in Lombard-Sprache gekürzt und Vokale verlängert werden (vgl. [15, S. 14] und [27, S. 515, 517]). Auch Drugman und Dutoit (2010) berichten, dass vorangegangene Studien bei Lärm einen Anstieg in der Dauer von Vokalen und Halbvokalen sowie eine Verkürzung der Konsonantendauern festgestellt haben (vgl. [13, S. 2610], [17, S. 2248] und [26, S. 185f.]). Fónagy und Fónagy (1966) führten hierzu eine Studie durch, in der

¹⁰Die VOT wird definiert als „[...] Verzögerung des Stimmtoneinsatzes gegenüber der Verschlusslösung [...]“ ([60, S. 125, 190]) bzw. als „Zeit zwischen oraler Verschlusslösung und dem Einsetzen der Stimmlippenschwingung als Merkmal der Stimmhaft-stimmlos-Opposition: negative VOT (Stimmlippen schwingen vor der Verschlusslösung): stimmhaft; 0 bis 20-30 ms VOT: stimmlos nichtaspiriert; VOT > 20-30 ms: stimmlos aspiriert.“ ([60, S. 323])

ungarische Proband*innen einen Text zunächst leise und anschließend laut vorlesen (vgl. [15, S. 14]). Sie beschreiben, dass im Falle des lauten Vorlesens Vokale deutlich gelängt und stimmlose Plosive sowie andere Konsonanten verkürzt werden, während Nasale und die Laute /l/, /j/ und /r/ gelängt werden (vgl. [15, S. 15ff., 21]). Dabei hänge die Veränderung der Dauer vom Grad der Sonorität einzelner Konsonanten ab (vgl. [15, S. 17f., 21]). Die sonorsten Konsonanten würden dabei jedoch nicht so stark gelängt werden wie die im Vergleich sonoreren Vokale, während die am wenigsten sonoren stimmlosen Plosive eine erhebliche Kürzung erfahren würden (vgl. [15, S. 17f.]).

In der Literatur finden sich unterschiedliche Angaben hinsichtlich der Sprechgeschwindigkeit. Während einige Forscher eine Reduktion dieses Parameters feststellen (vgl. [20, S. 364ff.] und [7, S. 361, 365]), konnte dieses Ergebnis in anderen Studien nicht bestätigt werden (vgl. [14, S. 2701, 2704]). Hanley und Steer (1949) stellten fest, dass mit der Steigerung der Geräuschbelastung eine Zunahme der Silbendauer einhergeht (vgl. [20, S. 364ff.]).

In einer Untersuchung von Mixdorff et al. (2007) tendierten die Proband*innen dazu, ihre Sprechaufgaben im Rahmen des *Map-Task* Szenarios¹¹ bei Lärm schneller zu lösen als im Stillen. Davon ausgenommen waren eine Aufwärmphase zu Beginn des Experiments sowie eine Geräuschbedingung, bei der Geflüster eingespielt wurde. Die Forscher*innen nehmen an, dass Sprecher*innen bei lauten Hintergrundgeräuschen nicht so entspannt seien wie in Stille und die Sprechaufgabe daher nicht so langsam bzw. gemächlich lösen würden. (vgl. [53]) Hazan et al. (2016) untersuchten Sprachaufnahmen von Erwachsenen sowie Kindern im Alter von neun bis vierzehn Jahren, welche Problemlösungsaufgaben bearbeiteten, bei denen jeweils zwei Gesprächspartner*innen über *Headsets* kommunizierten und den Unterschied zwischen verschiedenen Bildern finden sollten. Dabei wurde die Sprachverständlichkeit für die Zuhörer*innen bzw. den Zuhörer eines Aufnahmepaares künstlich vermindert, wodurch die Proband*innen mehr Zeit für die Lösung ihrer Sprechaufgabe benötigten. (vgl. [22, S. 1, 4ff., 9f.]) Aufgrund der unterschiedlichen Ergebnisse dieser beiden Studien wird vermutet, dass die zum Lösen einer Sprechaufgabe benötigte Dauer stark von der Art dieser Aufgabe sowie den Umständen der jeweiligen Kommunikationssituation abhängt.

¹¹Das *Map-Task* Paradigma ist in Kapitel 5.1 detailliert beschrieben.

Verschiedene Studien haben nachgewiesen, dass Sprachsignale für Zuhörer*innen verständlicher sind, wenn sie bei lauten Umgebungsgeräuschen produziert wurden, im Gegensatz zu einer stillen Umgebung (vgl. [12, S. 1320, 1322f.], [28, S. 16], [18, S. 255], [19, S. 589], [45, S. 3261f., 3268, 3270ff.], [70, S. 917, 924ff.] und [75, S. 48]). Eine reduzierte Verständlichkeit des akustischen Sprachsignals, die aufgrund lauter Umgebungsgeräusche auftritt, kann darüber hinaus durch die Verstärkung visueller Sprachinformationen, wie z. B. sprachbezogener Bewegungen des Mundes, der Lippen, des Kiefers sowie von Kopf und Gesicht, erhöht werden (vgl. [8, S. 1] und [19, S. 597]). Dabei ist die Verbesserung der Sprachverständlichkeit bei lauten Hintergrundgeräuschen durch visuelle Informationen größer, wenn diese visuellen Komponenten bei Lärm produziert worden sind, d. h., eine Art visueller Lombard-Sprache darstellen (vgl. [8, S. 7f.]).

Eine dieser Dissertation vorangehende Untersuchung (vgl. [34]) zeigte zudem, dass der im Auto bei unterschiedlichen Fahrgeschwindigkeiten hervorgerufene Lombard-Effekt die Anzahl der Redebeiträge im Dialog verringert. Dieses Ergebnis lässt darauf schließen, dass die Interaktion der Gesprächspartner*innen bei lauten Hintergrundgeräuschen abnimmt. Zudem wurde festgestellt, dass bei Lärm etwas mehr Hästitionen produziert werden, die Sprechgeschwindigkeit verringert ist und die Sprachproduktionsleistung¹² erhöht wird. (vgl. [34, S. 40ff.], [26, S. 190ff., 197], [37, S. 103ff.], [36, S. 415ff.] und Kapitel 4.2)

Viele Forscher*innen sprechen zudem davon, dass unter Lombard-Bedingungen eine Steigerung des *vocal effort* stattfindet (vgl. [22, S. 13], [27, S. 510, 517], [47, S. 1] und [70, S. 917]). Bottalico et al. (2016), die den Einfluss verschiedener Redestile, Raumakustiken und kurzfristiger Stimmermüdung auf den *vocal effort* untersuchten (vgl. [4, S. 2870]), beschrieben diesen Begriff dabei als:

„[...] physiological measure that accounts for changes in voice production as vocal loading increases. It has been quantified in terms of sound pressure level (SPL).“ ([4, S. 2870])

¹²Für die Berechnung dieses Parameters wurde „[...] die Sprachdauer in Sekunden auf die Gesamtdauer der Sprachaufnahme [...] in Sekunden normiert.“ ([34, S. 3])

Weiterhin sei die Anpassung von Sprecher*innen an abnehmende oder zunehmende Umgebungsgeräusche einer Anpassung an veränderte Kommunikationsdistanzen ähnlich, wobei jedoch nicht von einer vollständigen Entsprechung beider Phänomene ausgegangen werden kann (vgl. [74, S. 3438]). Tabelle 3.1 fasst die in diesem Kapitel beschriebenen Modifikationen der Sprachproduktion bei Lärm zusammen.

Tabelle 3.1: Die in Kapitel 3.1 beschriebenen Modifikationen der Sprachproduktion in Lombard-induzierenden Umgebungen. ↑ und ↓ geben eine Zu- oder Abnahme der jeweiligen Parameterwerte an.

Parameter	Modifikation	Parameter	Modifikation
Intensität	↑	Sprechgeschwindigkeit	(↑) (↓)
Grundfrequenz	↑	Dauer im Lösen von <i>Map-Task</i> Aufgaben	↓
<i>Spectral tilt</i>	↓	Sprachverständlichkeit	↑
F1	↑	visuelle Sprachinformationen	↑
F2	↑↓	Anzahl der Redebeiträge	↓
F3	↑	Anzahl der Häsitationen	(↑)
VOT	↓	Sprachproduktionsleistung	↑
Wortdauer	↑	<i>vocal effort</i>	↑
Dauer von weniger sonoren Konsonanten	↓		
Dauer von Vokalen und stark sonoren Konsonanten	↑		

Junqua (1996) beschreibt, dass laute Hintergrundgeräusche neben einem physiologischen ebenso einen psychologischen Einfluss auf Sprecher*innen besitzen und die dabei zustandekommenden Sprachmodifikationen u. a. davon abhängen, wie diese die vorhandene Geräuschbelastung bewältigen (vgl. [28, S. 20]). Dabei sei Lärm ein Umgebungsfaktor, durch den Stress¹³ ausgelöst werden kann, wobei das Ausmaß des Stresses erheblich von der einzelnen Person abhängt (vgl. [28, S. 14]). Aus diesen Erkenntnissen kann geschlossen werden, dass eine Reduktion des Lombard-Effekts eine Verringerung des Stresses für die Sprecherin bzw. den Sprecher bedeutet.

Aus den Ergebnissen der in diesem Kapitel vorgestellten Studien wird geschlossen,

¹³Eine kurze Diskussion zur Definition des englischen Begriffs *stress* ist in [28, S. 14] dargestellt.

dass die Parameter der Intensität, Grundfrequenz, Signaldauer (Dialogdauer beim Lösen einer Sprechaufgabe im *Map-Task* Szenario) und Sprachproduktionsleistung bewährte Eigenschaften des Lombard-Effekts sind. Daher wurden diese Parameter zur Untersuchung des Lombard-Effekts in den Sprachsignalen der vorliegenden Forschungsarbeit ausgewählt (vgl. Kapitel 1.2 und 3.4).

3.2 Verschiedene Interpretationen des Phänomens

Wie zu Beginn dieses Kapitels beschrieben, stellte Lombard fest, dass ein Anstieg der Stimmintensität nicht nur durch laute Hintergrundgeräusche, sondern auch durch eine Reduktion des Eigenechos bzw. *Sidetones* verursacht wird, wie es z. B. bei Gehörverlust der Fall ist (vgl. [41, S. 684f.]). Lane und Tranel (1971) nehmen an, dass die Steigerung der Lautstärke die Verständlichkeit der sprechenden Person aufrechterhalten soll (vgl. [41, S. 687]). Durch das Hören der eigenen Stimme kann eine Sprecherin bzw. ein Sprecher zudem ihre bzw. seine Lautstärke regulieren und Steigerungen des Umgebungslärms kompensieren (vgl. [16, S. 53f.] und [51, S. 177]). In der Forschungsliteratur wird zwischen zwei unterschiedlichen *Feedback-Loops* unterschieden. Die als intern bzw. privat bezeichnete *Feedback-Loop* ermöglicht es Sprecher*innen, Kontrolle über die eigene Stimme bzw. die eigene Stimmintensität ausüben zu können. Um diese Selbstkontrolle bei einem Anstieg der Hintergrundgeräusche weiterhin beizubehalten, muss die Lautstärke des eigenen Sprachsignals ebenfalls gesteigert werden. Durch die externe bzw. öffentliche *Feedback-Loop* können Sprecher*innen die eigene Stimme hingegen aufgrund der Reaktionen ihrer Zuhörer*innen verändern. (vgl. [41, S. 694f.], [58, S. 209] und [67, S. 191])

Die Ursachen des Lombard-Effekts werden im Rahmen unterschiedlicher Theorien besprochen. Dabei geht es vor allem um die Frage, ob und inwieweit eine Sprecherin bzw. ein Sprecher in der Lage ist, dieses Phänomen bewusst zu steuern. Es wird untersucht, ob sie oder er bewusst versucht, die eigene Verständlichkeit für Zuhörer*innen aufrechtzuerhalten, oder ob der Lombard-Effekt ein physiologischer Reflex ist, welcher aufgrund des abgeschwächten auditiven *Feedbacks* der eigenen Stimme die Lautstärke automatisch an die Umgebungsgeräusche anpasst (vgl. [18, S. 255] und [19, S. 588f.]). Diese Frage wird im Folgenden unter Bezugnahme auf die Ergebnisse unterschiedlicher Studien ausführlich diskutiert.

3.2.1 Der physiologische Reflex

Einige Forscher*innen charakterisieren den Lombard-Effekt als einen Reflex (vgl. [40, S. 619] und [27, S. 510f.]). Lane und Tranel (1971) beschreiben die Anpassung des eigenen Sprachsignals durch die öffentliche bzw. externe *Feedback-Loop* als eine unbewusste Handlungsweise (vgl. [41, S. 694f.]). Auch von anderen Autor*innen wird der Lombard-Effekt als eine unbewusste Tendenz bzw. eine unwillkürliche Anpassung an den Umgebungslärm charakterisiert (vgl. [15, S. 14, 21] und [69, S. 156]). Cowan (2016) definiert das Phänomen als

„[...] involuntary increase in voice level and change in voice characteristics (such as pitch, vocal rate, and enunciation) when background noise levels are elevated.“ ([6, S. 46])

In diesem Ansatz wird davon ausgegangen, dass Umgebungsgeräusche eine Störung der Kontrolle über die eigene Stimme verursachen (vgl. [1, S. 581]). Dafür spricht, dass der Lombard-Effekt nicht komplett unterbunden werden kann, selbst wenn diesbezüglich ein Training absolviert worden ist (vgl. [7, S. 361] und [18, S. 255]). Zudem weisen auch Tiere wie Vögel oder Affen vergleichbare Modifikationen, d. h. Veränderungen der Sprachproduktion und des Sprachverhaltens bei lauten Umgebungsgeräuschen, auf (vgl. [6, S. 46], [7, S. 361], [18, S. 255], [50, S. 1190, 1196], [61, S. 252f.] und [68]).

Manabe et al. (1997) schließen aus ihren Ergebnissen, dass Wellensittiche bei lauten Umgebungsgeräuschen ihre Stimmintensität reflexartig erhöhen. (vgl. [50, S. 1190ff., 1195]) Sinnott et al. (1975) erzielten ähnliche Untersuchungsergebnisse. Sie erforschten den Lombard-Effekt bei zwei Affen unterschiedlichen Geschlechts und verwendeten hierfür Umgebungsgeräusche mit unterschiedlichem SPL. Sie beschreiben, dass in fast allen Experimentbedingungen eine Steigerung der Stimmintensität von etwa 2 dB stattfand, wenn das Geräuschniveau sich um 10 dB erhöhte. Kein Affe musste lernen, die Stimmintensität bei Lärm zu steigern: (vgl. [68, S. 412f.]

„We believe we have found evidence of at least a rudimentary Lombard “reflex” or “response” in monkeys and that this behavior appears similar to that which takes place in isolated vocalizing humans. Monkeys appear to monitor their own voice and compensate for the masking effects of noise by changing vocal intensity in order to hear themselves.“ ([68, S. 414])

Auch Pick, Jr. et al. (1989) sehen aufgrund ihrer Untersuchungsergebnisse die Annahme bestätigt, dass der Lombard-Effekt ein automatischer Vorgang ist. Sprecher*innen könnten zwar ihre Stimme absichtlich heben und senken, aber eine Veränderung des auditiven *Feedbacks* bzw. den bestehenden Umgebungslärm nicht problemlos ignorieren. (vgl. [59, S. 894, 899f.])

3.2.2 Kommunikative Aspekte

Obwohl aufgrund der genannten Beobachtungen die Vermutung nahe liegt, dass der Lombard-Effekt ein Reflex ist, heben viele Forscher*innen die kommunikativen Aspekte dieses Phänomens hervor. Auch hierbei wird angenommen, dass die im Lombard-Effekt auftretenden Sprachmodifikationen der Aufrechterhaltung einer erfolgreichen Kommunikation mit Gesprächspartner*innen dienen (vgl. [27, S. 511], [1, S. 581] und [18, S. 255]):

„They generally implicitly make the assumption that the Lombard reflex is a physiological effect. However, it seems that in the real world, the magnitude of the response of the speakers is governed by the desire to obtain intelligible communication.“ ([28, S. 15])

Die Ergebnisse verschiedener Studien legen die Vermutung nahe, dass Sprecher*innen bei einer Veränderung ihres auditiven *Feedbacks* eine gewisse Kontrolle über die eigene Stimme ausüben können (vgl. [59, S. 894]). Sinnott et al. (1975) bestätigen zwar das Vorkommen von Lombard-Eigenschaften bei Affen (vgl. Kapitel 3.2.1), wodurch die Interpretation des Phänomens als Reflex bekräftigt wird, weisen jedoch zudem darauf hin, dass der Lombard-Effekt vor allem dann auftritt, wenn die kommunikative Effektivität eines Sprachsignals bei Lärm erhöht werden muss (vgl. [68, S. 412ff.]).

Eine Zunahme der Hintergrundgeräusche hat ungefähr die gleichen Auswirkungen auf die Stimmintensität wie ein damit vergleichbarer räumlicher Abstand der Gesprächspartner*innen (vgl. [41, S. 701]). In beiden Situationen findet ein Ausgleich der für die Hörerin bzw. den Hörer reduzierten Verständlichkeit statt, wobei die Selbstkontrolle der Sprecherin bzw. des Sprechers über die eigene Stimme lediglich durch die lauten Hintergrundgeräusche beeinträchtigt wird (vgl. [41, S. 701]). Diese Tatsache liefert Hinweise darauf, dass die von Silverman (2006) beschriebene „*social*

interpretation of the Lombard Effect“ ([67, S. 191]) Gültigkeit besitzt, bei der angenommen wird, dass Sprecher*innen von der öffentlichen *Feedback-Loop* Gebrauch machen, um die eigene Lautstärke anpassen zu können (vgl. [67, S. 191]). Auch die Verstärkung visueller Sprachinformationen bei lauten Hintergrundgeräuschen (vgl. Kapitel 3.1) weist darauf hin, dass die Sprecherin bzw. der Sprecher ihre bzw. seine Verständlichkeit für Gesprächspartner*innen erhöhen will, da ein visueller Lombard-Effekt vermutlich nicht zu einer verbesserten Eigenkontrolle der Sprechenden beiträgt.

Auch wenn der Lombard-Effekt nicht vollständig von Sprecher*innen kontrolliert werden kann, wie im vorangegangenen Kapitel 3.2.1 beschrieben wurde, versuchen diese unter nachteiligen Bedingungen dennoch, ihre Sprachproduktion den Bedürfnissen der Gesprächspartner*innen anzupassen (vgl. [1, S. 581] und [4, S. 2870]). Dies ist z. B. der Fall, wenn die Hörfähigkeit der Kommunikationspartnerin bzw. des Kommunikationspartners eingeschränkt ist (vgl. [4, S. 2870]), oder bei einem Gespräch, das in einer lauten Umgebung stattfindet. Diese Annahme wird u. a. durch die Vermutung von Lane und Tranel (1971) und weiteren Forscher*innen bestätigt, dass der Lombard-Effekt bei sehr kommunikativen bzw. interaktiven Sprechaufgaben stärker ausgeprägt ist als bei wenig oder nicht-kommunikativen bzw. nicht-interaktiven Aufgaben, wie z. B. dem Vorlesen einer Wortliste in ein Mikrofon ohne eine zuhörende Person (vgl. [1, S. 581] und [70, S. 925]).

Lane et al. (1970) schließen aus den Ergebnissen vorangegangener Studien, dass der Einfluss eines veränderten *Sidetones* auf die Stimmintensität vermutlich u. a. aufgrund einer fehlenden kommunikativen Aufgabe abgeschwächt wird (vgl. [40, S. 622]). Es gibt inzwischen zahlreiche Forschungsergebnisse, die darauf hinweisen, dass das Ausmaß des Lombard-Effekts mit den kommunikativen Ansprüchen einer Sprechaufgabe bzw. der Interaktivität mit einer anderen Person zusammenhängt. Die bei lauten Umgebungsgeräuschen stattfindenden Sprachmodifikationen sind in gelesener Sprache geringer als bei spontan gesprochener Sprache (vgl. [18, S. 255]). Des Weiteren gibt es Hinweise darauf, dass Inhaltswörter und weniger informative Funktionswörter unterschiedlich stark von den im Lombard-Effekt stattfindenden Veränderungen betroffen sein könnten (vgl. [18, S. 261] und [58, S. 213, 216ff.]). Die kommunikative Bedeutung des Phänomens wird durch die unterschiedliche Modifikation von Inhalts- und Funktionswörtern bei lauten Umgebungsgeräuschen zu-

sätzlich unterstrichen. Lau (2008) stellt zudem fest, dass auch dann ein geringer Lombard-Effekt hervorgerufen wird, wenn eine Sprecherin bzw. ein Sprecher keinem Umgebungslärm ausgesetzt ist, aber weiß, dass der Gesprächspartner bzw. die Gesprächspartnerin laute Geräusche hört (vgl. [42, S. 1, 6]).

3.2.3 Kombinierte Ursachen

Viele Forscher*innen vertreten die Ansicht, dass sich die bei lauten Hintergrundgeräuschen stattfindenden Sprachmodifikationen durch eine Kombination der beiden Interpretationsansätze aus Kapitel 3.2.1 und Kapitel 3.2.2 begründen lassen (vgl. [28, S. 14], [19, S. 599f.] und [58, S. 209]). Fraglich bleibt, ob das Auftreten einer gesteigerten Stimmintensität bei Lärm, wie es in Kapitel 3.2.1 für einige Tierarten beschrieben ist, tatsächlich einen Hinweis auf den reflexartigen Charakter des Lombard-Effekts darstellt. Theoretisch wäre es ebenso denkbar, dass dieses Phänomen den Tieren durchaus bewusst ist und auch der Mensch bei lauten Umgebungsgeräuschen seine Sprachproduktion wissentlich verändert. Zudem soll hier der Vorstellung Raum gegeben werden, dass sich die Kombination aus bewusster und unbewusster Sprachveränderung unter Lombard-Bedingungen eventuell in den einzelnen modifizierten Parametern widerspiegelt. Dabei könnte sich eine Sprecherin bzw. ein Sprecher z. B. durchaus der Steigerung von Intensität und Grundfrequenz bewusst sein, jedoch keine bewusste Kontrolle über die Veränderung von VOT und *spectral tilt* besitzen. Die wissenschaftliche Überprüfung dieser Annahme ist jedoch nicht Gegenstand der vorliegenden Arbeit.

Garnier et al. (2010) stellten fest, dass der Lombard-Effekt in Sprechaufgaben ohne und mit interaktiver Komponente auftrat, auch wenn der Effekt stärker ausgeprägt war, sobald eine Probandin bzw. ein Proband nicht mehr allein am Experiment teilnahm, sondern mit einer weiteren Person kommunizierte (vgl. [19, S. 595, 597ff.]). Des Weiteren wurde beobachtet, dass die Stärke des Lombard-Effekts u. a. durch eine Zunahme des Sprecher*innen-*Sidetones* und des visuellen *Feedbacks* verringert bzw. modifiziert werden kann. Das visuelle *Feedback* wurde in der Studie von Pick, Jr. et al. (1989) durch einen Bildschirm realisiert, auf dem die Proband*innen die Zu- und Abnahme ihrer Stimmintensität beobachten konnten. (vgl. [27, S. 511], [40, S. 618ff.] und [59, S. 894f., 899])

Diese unterschiedlichen Arten von Kompensierungsmethoden („*compensation methods*“ ([27, S. 511])) unterstützen jeweils die Interpretation des Lombard-Effekts als unbewussten Reflex oder als kommunikatives Phänomen. Hieraus ergibt sich eine Bekräftigung der Annahme, dass die Ursachen des Lombard-Effekts eine Kombination beider Erklärungsversuche bilden.

3.3 Die Untersuchung des Phänomens

Verschiedene Studien zum Lombard-Effekt haben ergeben, dass sich dieser künstlich hervorrufen lässt (vgl. [42, S. 1]). Das vorliegende Kapitel beschreibt, welche Faktoren bei der Untersuchung dieses Phänomens beachtet werden müssen. Dazu zählen einerseits verschiedene Umstände, die eine Erforschung im Labor erschweren. Andererseits wirken unterschiedliche Einflussgrößen auf den Umfang der Sprachmodifikationen bei lauten Hintergrundgeräuschen ein (vgl. [45, S. 3261] und [58, S. 210]). Diese sollten aus Gründen der besseren Vergleichbarkeit in jedem Experiment zum Lombard-Effekt besondere Aufmerksamkeit erhalten (vgl. [58, S. 210]).

3.3.1 Im Labor

Die Untersuchung des Lombard-Effekts im Labor hat den Vorteil, dass sie unter kontrollierten Bedingungen stattfindet. Diese künstlich erzeugten Gesprächssituationen spiegeln jedoch keine alltäglich auftretenden Kommunikationsbedingungen wider, zumal dazu häufig weißes Rauschen als Hintergrundgeräusch und Kopfhörer verwendet werden. Zudem wird dem Zusammenhang zwischen der Stärke des Lombard-Effekts und den kommunikativen Anforderungen einer Situation in vielen Studien, bei denen die Teilnehmer*innen ohne eine Gesprächspartnerin bzw. einen Gesprächspartner einen Text oder eine Liste von Wörtern vorlesen, nicht ausreichend Rechnung getragen. Die Herstellung einer Kommunikationssituation, welche möglichst wirklichkeitsgetreue Verhältnisse widerspiegelt, beinhaltet zahlreiche Hürden: (vgl. [14, S. 2701], [19, S. 589], [28, S. 14], [42, S. 1] und [58, S. 210])

„Even if the speakers were asked to speak in such a way as to be intelligible, there was no premium on intelligibility because the speakers were reading a list. Thus, most current databases do not place emphasis on communication, although this seems to be an important factor to consider.“ ([28, S. 15])

Bei der Untersuchung des Lombard-Effekts werden in den meisten Fällen Sprachaufnahmen benötigt, welche in lauten Umgebungen aufgenommen worden sind, die dabei auftretenden Hintergrundgeräusche jedoch nicht beinhalten. Der Grund hierfür ist, dass die so erhobenen Sprachsignale durch den Lärm beeinträchtigt sind, wodurch akustische Messungen verzerrt und Fehlinterpretationen formuliert werden können. Deshalb wird der Hintergrundlärm in Produktionsexperimenten zur Lombard-Sprache häufig durch Kopfhörer eingespielt, um geräuschfreie Sprachsignale zu erhalten. Doch Garnier et al. (2010) stellten fest, dass die Verwendung von Kopfhörern bereits einen Lombard-Effekt hervorruft, selbst wenn keine Geräusche durch die Kopfhörer abgespielt werden (vgl. [19, S. 595]). Durch das Tragen der Kopfhörer wird das eigene akustische *Feedback*, d. h. vom Mund hin zu den eigenen Ohren, verändert, wodurch die Sprachproduktion der jeweiligen Sprecher*innen beeinflusst wird. Daher wird manchmal auch eine direkte Kopplung vom Mikrofon zu den Kopfhörern (*Sidetone*) eingefügt. Doch obwohl diese Methode die Realitätsnähe der Umgebungssimulation mit Kopfhörern erhöht und die Sprecher*innen dadurch vermutlich ohne Lombard-Effekt auskommen, unterscheidet sich diese Situation dennoch erheblich von der Zielumgebung, in der die Sprachproduktion der Proband*innen untersucht werden soll. Jaschke (2012) beobachtete hingegen, dass seine Proband*innen bereits durch die Verwendung von Kopfhörern leiser sprachen und dass der Lombard-Effekt durch das Tragen von Kopfhörern weniger stark ausgeprägt ist. Die Ergebnisse von Jaschke (2012) sprechen daher gegen die Annahme, dass durch Kopfhörer ein zusätzlicher Lombard-Effekt hervorgerufen wird. Dennoch kann aufgrund der beschriebenen Forschungsergebnisse festgehalten werden, dass durch die Verwendung von Kopfhörern das Auftreten und die Stärke des Lombard-Effekts wesentlich beeinflusst wird. (vgl. [19, S. 589ff.], [24, S. 51, 53ff., 55ff., 63f.], [25, S. 1], [34, S. 36], [42, S. 1], [47, S. 1], [26, S. 192] und [37, S. 100])

3.3.2 Ausmaße des Effekts

Verschiedene Studien haben festgestellt, dass der Lombard-Effekt sprachübergreifend im Spanischen, Französischen, Japanischen und Amerikanischen Englisch vergleichbare Tendenzen aufweist. Dabei wurden jedoch auch Unterschiede in den akustischen Veränderungen zwischen den untersuchten Sprachen festgestellt. Die Grün-

de hierfür sind aufgrund verschiedener Ungleichheiten in den Vorgehensweisen der entsprechenden Studien schwer zu bestimmen. (vgl. [28, S. 18f.]) Es ist daher grundsätzlich möglich, dass der Lombard-Effekt von Sprache zu Sprache variiert. Um zu diesem Thema eine konkrete Aussage treffen zu können, sind jedoch weitere Untersuchungen nötig, welche sich speziell mit dieser Fragestellung befassen.

Die Art der Hintergrundgeräusche, welche zur Untersuchung des Lombard-Effekts verwendet werden, beeinflusst die Stärke der akustischen und artikulatorischen Veränderungen vieler unterschiedlicher Parameter (vgl. [17, S. 2247ff.]). Verschiedene Untersuchungen zur Modifikation sprachlicher Eigenschaften bei Lärm verwendeten dabei u. a. *babble noise* bzw. *cocktail party noise* oder auch *multitalker noise*, Geflüster, Flugzeuggeräusche sowie weißes Rauschen (vgl. [3, S. 907f.], [7, S. 365], [17, S. 2246], [20, S. 364], [53, S. 1329ff.] und [58, S. 209f.]). Nur wenige Studien untersuchten bislang die Veränderungen der Sprachproduktion bei Autogeräuschen, wie z. B. [13], [14], [24], [26], [34], [36], [37] und [53]. Garnier et al. (2010) beschreiben, dass die Verwendung von Kopfhörern statt Lautsprechern im Experiment einen größeren Einfluss auf die Veränderung von Lombard-Parametern hat, wenn *multitalker babble* anstelle von weißem Rauschen eingespielt wird (vgl. [19, S. 592f., 594f., 599]). Des Weiteren ist der Effekt auf verschiedene Aspekte der Sprachmodifikation geringer, wenn im Gegensatz zu stationären Umgebungsgeräuschen lediglich eine Sprecherin bzw. ein Sprecher im Hintergrund zu hören ist, wobei sich die Stärke des Lombard-Effekts mit der Anzahl der als Hintergrundgeräusch fungierenden Sprecher*innen erhöht (vgl. [45, S. 3268, 3273]). Darüber hinaus bewirkt eine Modifikation des Geräuschniveaus eine Veränderung im Ausmaß des Lombard-Effekts (vgl. [45, S. 3261, 3273] und [58, S. 210]). Summers et al. (1988) berichten, dass sich die Lautstärke eines Sprachsignals und die Wortdauer weiter erhöhen, je höher das Niveau der eingespielten Umgebungsgeräusche ist (vgl. [70, S. 919f.]). Zudem werden die bei Lombard-Sprache auftretenden akustischen Veränderungen u. a. durch die Frequenzverteilung bzw. den *spectral tilt* der Umgebungsgeräusche bestimmt (vgl. [28, S. 19f.] und [45, S. 3261]): „An increase of the speech energy in the frequency bands where the noise energy is most important was observed.“ ([28, S. 20])

Lane und Tranel (1971) gehen davon aus, dass die Stärke des Lombard-Effekts von dem dadurch erreichten Beitrag zu einer verständlichen Kommunikation abhängt (vgl. [41, S. 682]). In Kapitel 3.2.2 und Kapitel 3.2.3 wurde bereits beschrieben,

dass das Ausmaß des Lombard-Effekts in Spontansprache und kommunikativen bzw. interaktiven Sprechaufgaben größer ist, Inhalts- und Funktionswörter unterschiedlich stark modifiziert werden und eine Zunahme des visuellen *Feedbacks* bzw. des *Sidetones* die Stärke des Lombard-Effekts beeinflussen kann.

Der Lombard-Effekt variiert zudem im Hinblick auf das Geschlecht einer sprechenden Person und ist von Mensch zu Mensch unterschiedlich ausgeprägt (vgl. [6, S. 46], [12, S. 1321], [24, S. 44ff., 49ff., 63], [27, S. 510, 515, 518, 522f.], [28, S. 16f., 20] und [59, S. 896f.]). Die Stärke des Lombard-Effekts ist des Weiteren bei Kindern geringer als bei Erwachsenen (vgl. [18, S. 255]). Amazi und Garber (1982) fanden heraus, dass erwachsene Sprecher*innen bei einer interaktiven Sprechaufgabe stärker auf Umgebungsgeräusche reagieren als bei einer weniger interaktiven Aufgabe. Im Gegensatz dazu zeigten Kinder im Alter von fünf Jahren in beiden Aufgaben ähnliche Reaktionen. Als Begründung hierfür kann angenommen werden, dass das Bedürfnis nach Kontrolle der eigenen Stimme bei Kindern stärker ausgeprägt ist als das Bestreben, für eine ZuhörerIn bzw. einen Zuhörer verständlich zu sein. Daneben ist es zudem möglich, dass Kinder die Anforderungen einer GesprächspartnerIn bzw. eines Gesprächspartners nicht sehr gut einschätzen können. (vgl. [1, S. 581f., 584]) Auch Hazan et al. (2016) geben zu bedenken, dass Kinder größere Schwierigkeiten damit haben könnten, die Perspektive einer anderen Person anzunehmen, und ein gewisser Egozentrismus sei auch bei Jugendlichen noch immer offenkundig. (vgl. [22, S. 14])

3.4 Arbeitsdefinition und Hypothesen

Die Charakteristika des Lombard-Effekts sind nicht nur zahlreich, sondern auch sehr vielfältig (vgl. Kapitel 3.1). Es gibt Eigenschaften, die der Sprachproduktion bzw. der Artikulation von Lauten dienen, wie z. B. auf segmenteller Ebene die VOT und die Formantwerte, wobei das zuletzt genannte Merkmal die Vokalqualität bestimmt. Auf der anderen Seite existieren Charakteristika, welche vor allem die interaktive Kommunikation zwischen zwei oder mehreren Gesprächspartner*innen aufrecht erhalten sollen. Dazu gehören u. a. die Häufigkeit des Blickkontakts und die visuellen Sprachinformationen. Die Kapitel 3.2.2, 3.2.3 und 3.3.2 haben verdeutlicht, dass das Ausmaß des Lombard-Effekts von der Interaktivität und der Spontanität

der Sprechsituation abhängt. Eine umfassende Untersuchung des Phänomens sollte diese Faktoren daher berücksichtigen. Wenn Parameter aus unterschiedlichen Bereichen des Kommunikationsverhaltens analysiert werden sollen, ist demnach eine breite Arbeitsdefinition des Lombard-Effekts nötig.

Eine Beschränkung der Definition des Lombard-Effekts auf die Modifikationen der Sprachproduktion, ohne Berücksichtigung des interaktiven Verhaltens zwischen den Dialogpartner*innen, ist für die vorliegende Arbeit unzureichend, da hierbei auch die Verbesserung des Kommunikationsverhaltens zwischen den Passagier*innen durch ein ICC-System untersucht werden soll. D. h., neben den akustischen Parametern der Intensität und der Grundfrequenz werden dabei auch die Signaldauer (Dialogdauer) und die Sprachproduktionsleistung als interaktive Merkmale der Kommunikation analysiert (vgl. Kapitel 1.2, 3.1 und 6.3).

*Da die Effektrichtung für einzelne Parameter bei zunehmender Geräuschbelastung unterschiedlich sein kann und dabei Parameter wie die Sprachproduktionsleistung bisher noch wenig untersucht worden sind, wird jedwede Modifikation von Messwerten in diese Definition integriert. Im Hinblick auf die nachfolgenden Analysen und Ergebnisinterpretationen umfasst die Definition des Lombard-Effekts, welche dieser Studie zugrunde liegt, daher jegliche Veränderung des Kommunikationsverhaltens unter Geräuscheinwirkung. Darunter fällt alles, z. B. von der Modifikation der Sprecher*innenwechsel über den Grad der Wortreduktion bis zur Veränderung der Formantfrequenzen. Dies umfasst sowohl die Zu- als auch die Abnahme der entsprechenden Messwerte.*

In Kapitel 3.1 wurde gezeigt, dass neben der Veränderung der Intensität auch die Modifikation der Grundfrequenz zu den am häufigsten untersuchten Eigenschaften des Lombard-Effekts gehört. Da dieses Phänomen bisher in keiner der Autorin bekannten Studien für den Innenraum eines Autos unter Verwendung der im folgenden Kapitel 4 beschriebenen akustischen und visuellen Umgebungssimulation untersucht worden ist, wird die Analyse von Intensität und Grundfrequenz zunächst dazu verwendet, den Lombard-Effekt für den Kommunikationsort Autoinnenraum nachzuweisen. Um dem interaktiven Kommunikationsverhalten zwischen den Fahrzeugpassagier*innen Rechnung zu tragen, werden weiterhin die Parameter der Signaldauer

(Dialogdauer) und der Sprachproduktionsleistung in die Untersuchung aufgenommen, welche bis zu diesem Zeitpunkt nur selten untersucht worden sind. Aufgrund der in der bisherigen Forschungsliteratur angegebenen Resultate werden dabei folgende Hypothesen für den Kommunikationsort des Autoinnenraums aufgestellt:

Hypothesen zu Teilfrage (1)

1. Bei zunehmenden Fahrgeräuschen wird die Intensität gesteigert.
2. Bei zunehmenden Fahrgeräuschen wird die Grundfrequenz gesteigert.
3. Bei zunehmenden Fahrgeräuschen benötigen die Proband*innen weniger Zeit zum Lösen der Sprechaufgabe im *Map-Task* Szenario. Die Signaldauer (Dialogdauer) wird kürzer.
4. Bei zunehmenden Fahrgeräuschen wird die Sprachproduktionsleistung gesteigert.

Die Erforschung des Lombard-Effekts im Auto unter Verwendung eines ICC-Systems ist bisher ebenfalls in keiner der Autorin bekannten Studien durchgeführt worden. Die folgenden Hypothesen beschreiben, welche Veränderungen im Kommunikationsverhalten der Passagier*innen durch Aktivierung eines ICC-Systems erwartet werden:

Hypothesen zu Teilfrage (2)

5. Unter Verwendung eines ICC-Systems wird der Lombard-Effekt im Auto reduziert oder aufgehoben.
6. Im Auto wird mit ICC-System trotz des umgebenden Fahrzeuglärms wie in einer stillen Situation gesprochen. D. h., es findet weder ein Lombard-Effekt noch eine allgemeine Niveaushiftung von Messwerten statt.
7. In der stillen Referenzsituation hat das ICC-System keinen Einfluss auf das Kommunikationsverhalten der Passagier*innen.
8. Das ICC-System verbessert die Kommunikation zwischen den Sprecher*innen im Auto.

Des Weiteren wird erwartet, dass ein durch die akustisch-phonetische Analyse festgestellter Lombard-Effekt von Zuhörer*innen wahrgenommen werden kann, die nicht an den Dialogen beteiligt waren. Zudem liegt die Vermutung nahe, dass unabhängige Zuhörer*innen weitere Lombard-Eigenschaften feststellen, welche durch die zuvor durchgeführte Untersuchung nicht erkannt werden konnten, da nur ein Teil der spektralen und temporalen Variation in den Sprachsignalen im Rahmen der vorliegenden Arbeit parametrisch repräsentiert und analysiert werden kann und dabei die visuellen Aspekte der Kommunikation nicht berücksichtigt werden. Die über die akustisch-phonetische Analyse hinausgehenden Hypothesen sind in Kapitel 7.2 zusammengefasst.

4 Die akustische und visuelle Umgebungssimulation

Um Sprachkommunikation im Auto bei lauten Hintergrundgeräuschen unter kontrollierten Laborbedingungen untersuchen zu können, wurde am Institut für Digitale Signalverarbeitung und Systemtheorie (DSS) der CAU Kiel eine akustische Umgebungssimulation entwickelt, welche Lautsprecher zum Einspielen verschiedener Geräusche verwendet. Dadurch wird u. a. vermieden, dass aufgrund der Verwendung von Kopfhörern bei der Untersuchung des Lombard-Effekts Nachteile entstehen (vgl. Kapitel 3.3). In der Simulation aufgenommene Sprachsignale enthalten den eingespielten Umgebungslärm. Dieser kann durch unterschiedliche Techniken der Signalverarbeitung fast vollständig aus den Aufnahmen entfernt werden, wodurch nahezu dieselbe Sprachqualität wie bei der Verwendung von Kopfhörern erzielt wird. Zudem wird durch diese Methode die Realitätsnähe der Umgebungssimulation sowie die ökologische Validität der unter Lombard-Bedingungen erhobenen Sprachaufnahmen gesteigert. In der akustischen Umgebungssimulation können unterschiedliche Geräuschkategorien zur Erhebung von Lombard-Sprache verwendet und leicht reproduziert werden: (vgl. [24, S. 5, 63], [34, S. 36f.], [37, S. 99], [38, S. 479], [39, S. 1485], [47] und [49])

„The aim is to recreate any desired noise scenario within a sound-proof room (or within a vehicle) which can be reproduced at any time.“ ([47, S. 1])

Zur Vorbereitung der akustischen Umgebungssimulation werden die jeweils erforderlichen Geräuschkategorien zuvor mithilfe von kalibrierten Mikrofonen aufgenommen, welche sich an speziellen Positionen befinden. Um verschiedene Fahrsituationen akustisch nachzubilden zu können, werden Geräusche während einer echten Autofahrt bei verschiedenen Geschwindigkeiten aufgenommen. Dabei werden mehrere Mikro-

4 Die akustische und visuelle Umgebungssimulation

phone sowie ein künstlicher Kopf verwendet, in dessen Ohren weitere Mikrophone die Hintergrundgeräusche binaural aufnehmen, wobei ein Stereosignal entsteht. Diese Aufnahmesituation ist in Abbildung 4.1 dargestellt. Anschließend werden die erhobenen Signale zur Verwendung in einer Wiedergabeschleife vorbereitet sowie bearbeitet und dabei z. B. hinsichtlich vorhandener Artefakte überprüft. Innerhalb der akustischen Umgebungssimulation werden die so erhobenen und vorbereiteten Signale durch acht Lautsprecher und zwei *Subwoofer* wiedergegeben, welche zuvor kalibriert worden sind. Darüber hinaus wird die akustische Umgebungssimulation im Innenraum eines im Labor stehenden Autos realisiert. Zu diesem Zweck werden sechs Fensterlautsprecher an Front-, Heck- und Seitenscheiben angebracht sowie zwei *Subwoofer* auf dem Fahrersitz und dem linken Rücksitz platziert. Dadurch wird ein räumliches Geräuschempfinden erzeugt (vgl. Abbildung 4.2). Bei den hierfür verwendeten Testautos handelt es sich um eine Limousine (Mercedes S-Klasse) sowie ein Cabriolet (Mercedes E-Klasse). Durch die Übertragung der akustischen Fahrsimulation auf den Innenraum eines Autos soll eine erhöhte Realitätsnähe der Simulation erreicht und somit die Erhebung von realistischerer Lombard-Sprache ermöglicht werden. (vgl. [24, S. 5, 29f.], [25, S. 1ff.], [26, S. 192f., 196, 199f.], [34, S. 35ff.], [36, S. 414], [37, S. 99f.], [39, S. 1485], [47, S. 2, 5, 8] und [49])



Abbildung 4.1: Der künstliche Kopf zur Aufnahme von Geräuschen. ([25, S. 2] und [37, S. 100])



Abbildung 4.2: Die an den Innenseiten der rechten Autofenster angebrachten Lautsprecher der akustischen Umgebungssimulation. ([26, S. 196] und [38, S. 479])

Im Folgenden werden zunächst Methoden zur Geräuschbefreiung von in der akustischen Umgebungssimulation aufgenommenen Sprachsignalen dargestellt. Anschließend erfolgt eine Beschreibung verschiedener Evaluierungen und Ergänzungen, die in diesem Zusammenhang durchgeführt wurden, wie z. B. einer Erweiterung der akustischen Umgebungssimulation durch eine visuelle Komponente.

4.1 Die Geräuschbefreiung der Sprachsignale

Ein großer Vorteil der akustischen Umgebungssimulation ist, dass die bei Sprachaufnahmen eingespielten Hintergrundgeräusche im Nachhinein aus den erhobenen Signalen entfernt bzw. bedeutend reduziert werden können. Dadurch sollen phonetische Analysen durchführbar gemacht werden, denn Umgebungslärm kann die akustischen Untersuchungen beeinträchtigen (vgl. Kapitel 3.3.1). Dabei können z. B. die aus einem Sprachsignal gewonnenen Frequenzwerte der Formanten durch die Hintergrundgeräusche verfälscht sein. Des Weiteren sollen Hörer*innen die Aufnahmen als gut verständliche Sprachsignale einstufen können. Die Geräuschbefreiung geschieht durch unterschiedliche Methoden der Signalverarbeitung, der *noise cancellation* (NC) und der *noise suppression* (NS). (vgl. [25, S. 1ff.], [26, S. 192ff.], [36, S. 414], [37, S. 100], [47] und [49])

Mithilfe der oben beschriebenen akustischen Umgebungssimulation wurde im schallgeschützten Raum eine Lombard-Sprachdatenbank (vgl. [24]) erstellt. Bei Fahrgeräuschen von 0 km/h (Stillstand bei eingeschaltetem Motor), 50 km/h, 100 km/h und einer stillen Referenzsituation wurden dabei Sprachaufnahmen mit 5 weiblichen

4 Die akustische und visuelle Umgebungssimulation

und 5 männlichen Sprecher*innen durchgeführt, welche in jedem Geräuschszenario u. a. den Text „Nordwind und Sonne“ lasen (vgl. [25, S. 4]). Dabei wurden insgesamt $4 \times 10 = 40$ Sprachsignale erhoben. In Abbildung 4.3 ist die Anwendung der Methoden *noise cancellation* und *noise suppression* auf eines der Sprachsignale dargestellt, das während dieses Produktionsexperiments bei Hintergrundgeräuschen von 100 km/h aufgenommen wurde. In den abgebildeten Schmalbandspektrogrammen sind bis auf den Zeitabschnitt von etwa 0,2 bis 0,3 Sekunden Sprachlaute zu erkennen. Oben ist das nicht prozessierte Originalsignal zu sehen, welches das Sprachsignal mit den rot markierten Harmonischen beinhaltet. Darüber hinaus ist das gesamte Signal im tieffrequenten Bereich bis etwa 500 Hz von den ebenfalls rot gekennzeichneten Fahrgeräuschen durchzogen. Diese sind darüber bis in einen Frequenzbereich von ungefähr 1,5 kHz weniger energiereich und daher gelb gekennzeichnet. Das in der Mitte abgebildete Spektrogramm zeigt denselben Ausschnitt nach Anwendung der *noise cancellation*. Dabei werden die Hintergrundgeräusche im Signal stark abgeschwächt. Durch die zusätzliche Anwendung der *noise suppression* im unteren Spektrogramm wird der Lärm größtenteils entfernt, was besonders im nicht-sprachlichen Abschnitt zwischen 0,2 und 0,3 Sekunden zu sehen ist. An dieser Stelle ist keine gelb oder rot markierte Energie der Hintergrundgeräusche mehr zu erkennen. Auf diese Weise können durch die akustische Umgebungssimulation mit Lautsprechern der CAU Kiel Sprachaufnahmen unter Lombard-Bedingungen durchgeführt und anschließend fast vollständig von den Fahrgeräuschen befreit werden. (vgl. [24, S. 63], [47, S. 5ff.] und [49])

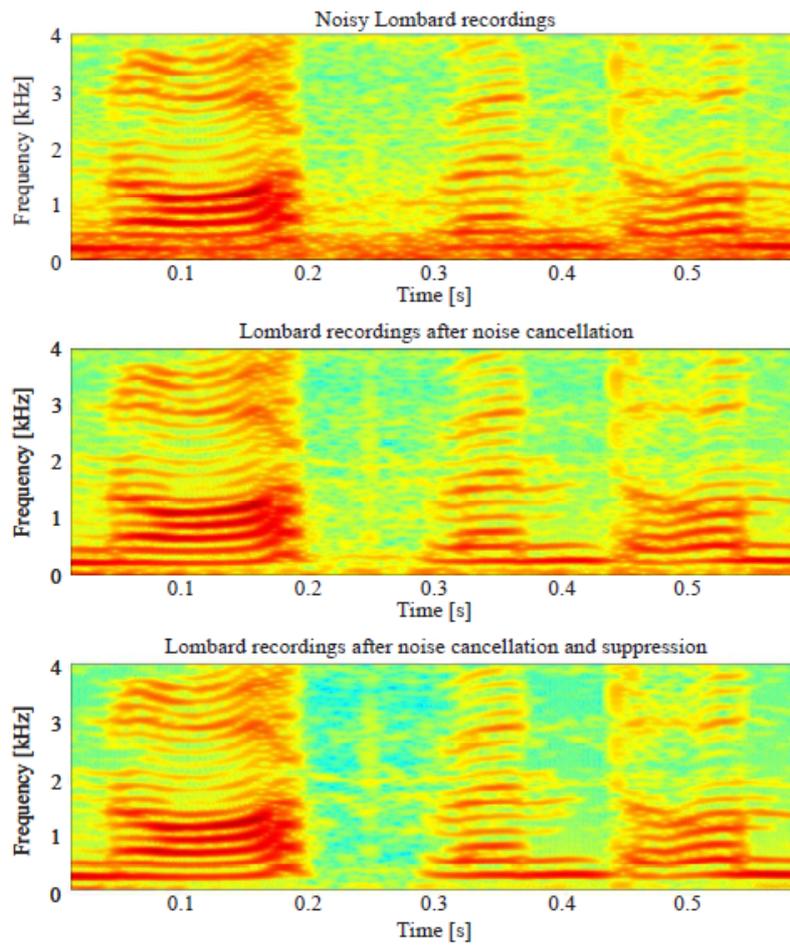


Abbildung 4.3: Spektrogramm eines im Audiolabor aufgenommenen Sprachsignals während einer Fahrgeräuschsimulation von 100 km/h. Dargestellt sind das Originalsignal sowie verschiedene Stufen der Geräuschbefreiung. ([47, S. 7] und [49])

4.2 Die Realitätsnähe der Simulation im Fahrzeug

In einer vorangegangenen linguistischen Untersuchung zum Lombard-Effekt wurde die Realitätsnähe der oben beschriebenen akustischen Fahrsimulation in der Limousine getestet. Dazu wurden spontansprachliche Dialoge in der Simulation und während einer echten Autofahrt im selben Auto bei Fahrgeräuschen von 0 km/h, 50 km/h, 100 km/h, 150 km/h und in einer stillen Referenzbedingung innerhalb des Autos aufgenommen, sodass sich insgesamt neun Aufnahmesituationen ergaben. Anschließend wurde untersucht, ob die Sprachproduktion bzw. das Sprachverhalten in beiden Kommunikationssituationen bei steigender Geräuschbelastung auf dieselbe Weise modifiziert wird. 6 Sprecher*innen, darunter 2 weibliche und 4 männliche,

4 Die akustische und visuelle Umgebungssimulation

nahmen am Produktionsexperiment dieser Untersuchung teil und bildeten insgesamt drei gleichgeschlechtliche Dialogpaare, um den Entrainment-Effekt¹⁴ zu minimieren. Die jeweiligen Sprecher*innen der Dialogpaare waren miteinander befreundet. Dabei saß eine Probandin bzw. ein Proband in allen Aufnahmebedingungen auf dem Beifahrersitz, während die andere Person dahinter Platz nahm, sodass kein Augenkontakt zwischen ihnen möglich war. Zur Elizitation der spontansprachlichen Dialoge wurde das *Video-Task* Paradigma verwendet, welches zuvor am Institut für Phonetik und digitale Sprachverarbeitung (IPdS) der CAU Kiel im sogenannten „*Daily Soap* Szenario“ angewendet worden war (vgl. [88]). Dabei sehen die Sprecher*innen eines Dialogpaares jeweils unterschiedliche Zusammenschnitte einer ihnen bekannten Fernsehserie, welche jedoch nicht vollständig identisch sind. In den Sprachaufnahmen sollen Unterschiede und Gemeinsamkeiten der gesehenen Videos erörtert werden. Anschließend beantworteten die Proband*innen zudem einen Fragebogen, der sich auf ihre subjektiven Eindrücke im Experiment bezog. (vgl. [26, S. 190, 197], [34, S. 26], [36], [37, S. 97f., 101ff.] und [38, S. 479])

Die Analyse der aufgenommenen Sprachsignale ergab, dass es quantitative Unterschiede zwischen der akustischen Umgebungssimulation und dem fahrenden Auto gibt, welche sich jedoch lediglich auf die Stärke des Lombard-Effekts beziehen. Im Labor wurden mehr *Turns* und eine etwas höhere Anzahl an Häsitationen produziert, während die Menge der produzierten Sprache und die Sprechgeschwindigkeit reduziert wurden. Es gibt jedoch an beiden Kommunikationsorten ähnliche Veränderungen des Sprachverhaltens und der Sprachproduktion. Aus den Ergebnissen dieser Untersuchung geht hervor, dass der im fahrenden Auto entstehende Lombard-Effekt durch Anwendung der akustischen Umgebungssimulation im Labor nachgebildet werden kann. Es wurden mit steigendem Umgebungslärm weniger *Turns*, mehr Sprache und etwas mehr Häsitationen produziert, und die Sprechgeschwindigkeit nahm ab. Aufgrund der oben beschriebenen Probleme bei akustischen Messungen in geräuschbehafteten Sprachsignalen und der Tatsache, dass lediglich die im Labor aufgenommenen Dialoge von den Hintergrundgeräuschen hätten befreit werden können, wurden Parameter wie Grundfrequenz und Intensität nicht analysiert. Die audi-

¹⁴Der Effekt des „phonetic entrainment“ wird bei Niebuhr und Michaud (2015) folgendermaßen definiert: „[...] when two speakers are engaged in a dialogue, they become phonetically more similar to each other, particularly when the interaction is cooperative and/or when the two dialogue partners are congenial with each other [...]“ ([56, S. 4])

tiven Eindrücke der Experimentleiterin und der Proband*innen bestätigten jedoch, dass Lautstärke und Grundfrequenz der aufgenommenen Sprache mit zunehmender Geräuschbelastung an beiden Kommunikationsorten anstiegen. Die akustische Umgebungssimulation kann demnach zur Untersuchung der lautsprachlichen Kommunikation im (fahrenden) Auto verwendet werden. Dies ermöglicht die ökologische und externe Validität der gewonnenen Ergebnisse. (vgl. [26, S. 191f., 197, 201], [34, S. 37ff.], [36], [37, S. 97f., 101f., 104ff.] und [38, S. 479f.]

4.3 Ergänzungen

Nachdem gezeigt wurde, dass der Lombard-Effekt in der akustischen Umgebungssimulation nachgebildet werden kann, wurde diese wiederum verbessert, da einige der Proband*innen berichteten, dass sie sich während der dort aufgenommenen Dialoge etwas unwohl fühlten. Sie hatten den Eindruck, dass die Lautstärke der Fahrgeräusche in der akustischen Umgebungssimulation höher sei als im fahrenden Auto. Dabei wurden die Hintergrundgeräusche in der Simulation als störender wahrgenommen als während der echten Autofahrt. Zudem wird vermutet, dass sich verschiedene Geräte während der Simulation, wie z. B. Mikrophone, Kabel und die den Innenraum verdunkelnden Fensterlautsprecher, negativ auf das Wohlbefinden der Proband*innen auswirkten. Dieses Unbehagen könnte einer der Gründe dafür sein, dass der Lombard-Effekt im fahrenden Auto und in der akustischen Umgebungssimulation quantitative Unterschiede aufweist. (vgl. [34, S. 39], [36, S. 417], [37, S. 107f.] und [38, S. 479])

Zur Steigerung der Realitätsnähe wurde die akustische Umgebungssimulation an der CAU Kiel durch eine visuelle Simulation ergänzt, die zusätzlich das Wohlbefinden der Proband*innen im Fahrzeug erhöhen könnte. Mithilfe einer großen Leinwand, welche sich vor dem Auto befindet, und einem Beamer werden Videos abgespielt. Das Testauto ist hierbei dasselbe, das bereits zuvor für die akustische Umgebungssimulation verwendet wurde. Die entsprechenden Videoaufnahmen wurden von Fahr-situationen bei 0 km/h, 50 km/h, 100 km/h, 130 km/h und 190 km/h aus Sicht der Beifahrerin bzw. des Beifahrers durchgeführt und mit den dazugehörigen Audioaufnahmen synchronisiert. Eine subjektive Bewertung dieser Simulation durch Proband*innen ergab, dass eine zusätzliche visuelle Simulation die Wirksamkeit der

4 Die akustische und visuelle Umgebungssimulation

akustischen Umgebungssimulation erhöht. D. h., die Realitätsnähe der Fahrsimulation wird gesteigert. Dies äußert sich u. a. darin, dass die Teilnehmer*innen dieser Studie bessere Schätzungen der Fahrgeschwindigkeit abgaben, wenn sowohl Audio- als auch Videosignale präsentiert wurden, als wenn nur Audio- bzw. nur Videosignale abgespielt wurden. Zusätzlich wurden den Proband*innen durch eine akustische und visuelle Simulation Situationen gezeigt, die eine Aktion von der Fahrerin bzw. vom Fahrer erfordern, wie z. B. steuern, beschleunigen und anhalten. Dabei hatten die meisten Teilnehmer*innen das Bedürfnis, zu bremsen oder zu lenken. (vgl. [36, S. 417], [38, S. 479], [39, S. 1485], [49] und [76])

5 *SPontaneous In-car Dialogues* (SPID)

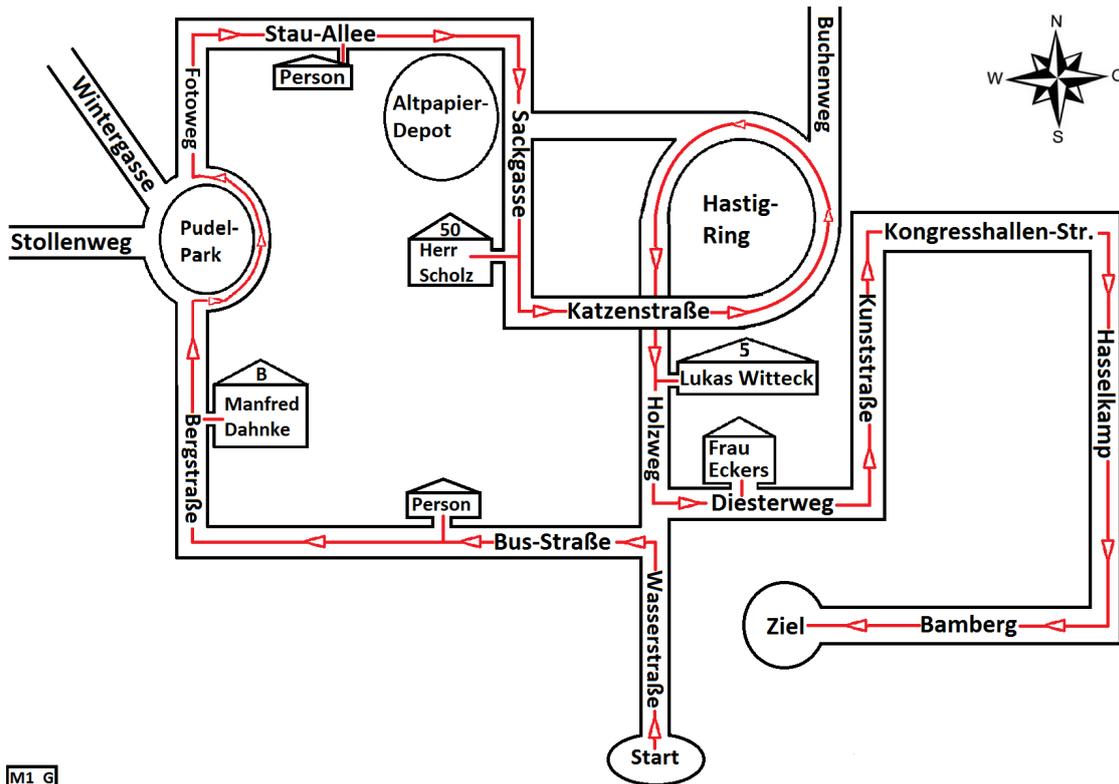
Im Folgenden werden die Vorbereitung sowie die Generierung des Sprachkorpus SPID (*SPontaneous In-car Dialogues*) ausführlich dargestellt. Dieses ist u. a. in [55, S. 1013f.], [38, S. 479ff.] sowie *online* auf einer Webseite der Arbeitsgruppe DSS beschrieben (vgl. [80]). Weiterhin erfolgt eine detaillierte Darstellung der zu den Sprachaufnahmen durchgeführten Segmentierungen, Transkriptionen und Annotationen. Dabei werden Probleme beschrieben, die während des Forschungsprozesses auftraten, verschiedene Lösungsansätze vorgestellt und deren Ergebnisse präsentiert. Zur Vorbereitung auf die Sprachaufnahmen des Sprachkorpus SPID wurde zunächst eine geeignete Sprechaufgabe im Rahmen des *Map-Task* Szenarios erstellt sowie eine Auswahl der am Produktionsexperiment beteiligten Proband*innen vorgenommen. Weiterhin werden im Folgenden die Umstände des Experiments beschrieben, welche u. a. die Instruktionen der Teilnehmer*innen, die Bedingungen der Sprachaufnahmen, die Durchführung der Aufnahmen sowie die Geräuschbefreiung der Sprachsignale beinhalten.

5.1 Das *Map-Task* Szenario

Da die Stärke des Lombard-Effekts durch die Interaktivität einer Sprechaufgabe und eine spontane Sprachproduktion erhöht wird (vgl. Kapitel 3.2.2 und 3.2.3), wurden für die Aufnahmen des SPID-Korpus spontansprachliche Dialoge elizitiert. Zu diesem Zweck wurden Sprechaufgaben im Rahmen des *Map-Task Szenarios* entwickelt, welches bereits in früheren Studien erfolgreich verwendet worden war (vgl. [72]). Ebenso wie das *Video-Task* Paradigma (vgl. Kapitel 4.2) bietet dieses Szenario eine für Proband*innen interessante Sprechaufgabe, welche die Teilnehmer*innen zu einem

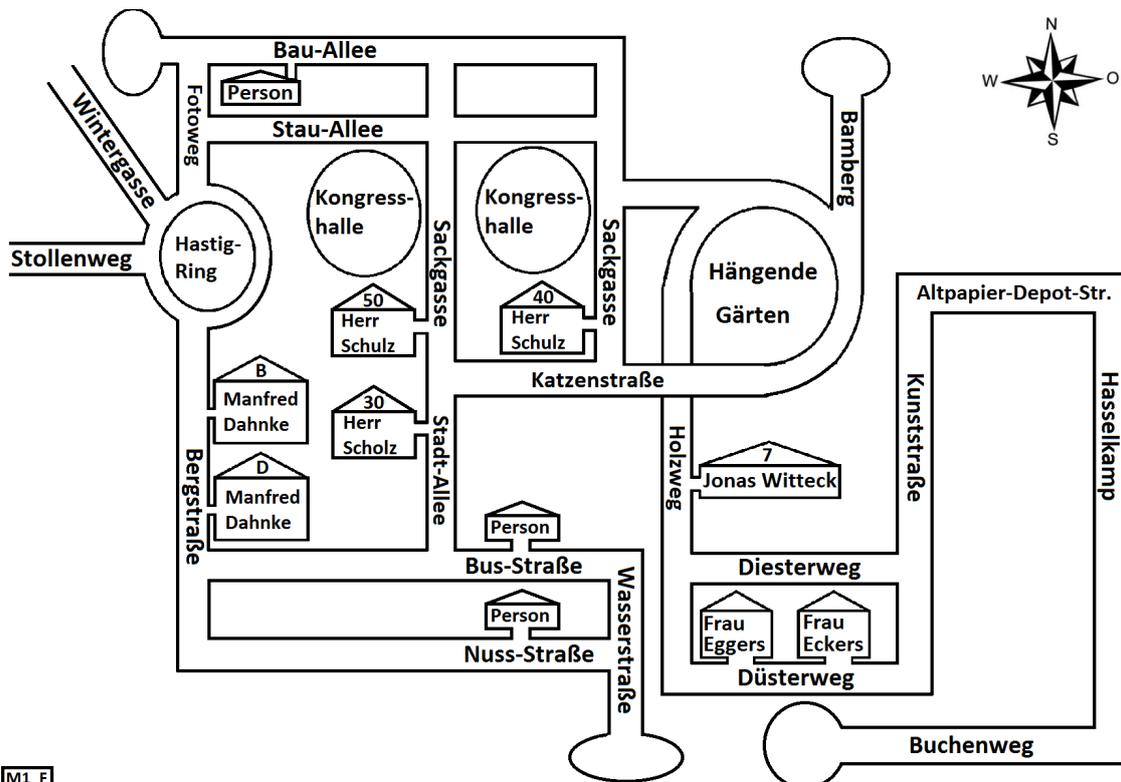
gewissen Grad von der Aufnahmesituation ablenkt (vgl. [37, S. 102]). Zur Erhebung der Sprachsignale wurden daher mehrere fast identische Paare von Straßenkarten erarbeitet, die jedes Dialogpaar in unterschiedlichen Geräusch- und ICC-Umgebungen besprechen sollte. Alle Kartenpaare des Produktionsexperiments sind im Anhang dieser Arbeit aufgeführt (vgl. Anhang A.1), während eines zudem beispielhaft in den Abbildungen 5.1 und 5.2 dargestellt ist. Die Kartenaufgaben sind frei erfunden und wurden von der Experimentleiterin erstellt. Eine bzw. einer der Sprecher*innen übernahm dabei jeweils die Rolle des *Instruction-Givers*, während die oder der andere die des *Instruction-Followers* innehatte. In der Karte des *Givers* ist eine Route dargestellt, die durch verschiedene Straßen vom Start- zum Zielpunkt führt und dabei die mit Nummern oder Buchstaben versehenen Häuser verschiedener Personen streift. Während der Aufnahmen war es Aufgabe des *Givers*, diese Route dem *Follower* mündlich zu erläutern, die bzw. der sie dann in der jeweils eigenen Karte nachzeichnete. Zudem sollten Personen, die in den Häusern der Karten verzeichnet sind, abgeholt werden. Für jede der 10 Aufnahmebedingungen des Produktionsexperiments wurden unterschiedliche Kartenpaare erstellt, um einem Lerneffekt der Proband*innen hinsichtlich der Karten vorzubeugen. Diese Paare bestanden jeweils aus einer *Giver*- und einer *Follower*-Karte. (vgl. [38, S. 480], [39, S. 1485] und [72, S. 26])

Durch das *Map-Task* Paradigma wird eine gute Kontrolle über die Verwendung bestimmter lexikalischer Einheiten im Dialog gewährleistet (vgl. [88]). Die Konzeption der Straßenkarten erlaubt zudem eine Untersuchung bestimmter sprachlicher Phänomene unter kontrollierten Bedingungen, wie z. B. den bei verschiedenen Lärmbedingungen stattfindenden phonologischen Prozessen. Vor diesem Hintergrund fand die Konzeption der Karten sowie der darauf verzeichneten Zahlen und Namen statt. Jedes Kartenpaar enthält unterschiedliche Zielwörter, welche bestimmte phonologische Prozesse ermöglichen. Diese sind in Tabelle 5.1 aufgeführt. Des Weiteren ist die Gestaltung der Straßenkarten darauf ausgerichtet, dass beide Dialogpartner*innen möglichst alle Zielwörter während der Aufnahmen mindestens einmal nennen. Die zwei Karten eines Paares stimmen daher in den meisten Eigenschaften miteinander überein, unterscheiden sich jedoch in einigen Punkten. D. h., in der Karte des *Givers* kommen Hausnummern, Straßen- und Personennamen vor, die in der Karte des *Followers* fehlen, und umgekehrt. Zudem weisen beide Karten z. T. leicht



M1_G

Abbildung 5.1: Die erste Aufgabe des *Map-Task* Szenarios, Sprechrolle *Giver*.



M1_F

Abbildung 5.2: Die erste Aufgabe des *Map-Task* Szenarios, Sprechrolle *Follower*.

unterschiedliche Strukturen auf. Die Länge der zu besprechenden Route sowie die zwischen zwei Karten eines Kartenpaares auftretenden Unstimmigkeiten wurden auf die Elizitation von etwa 6 Minuten (360 Sekunden) Dialogdauer abgestimmt. Die Kommunikation der Route zwischen *Giver* und *Follower* erforderte einen genauen mündlichen Abgleich der beiden Karten miteinander, in dem Unterschiede und Gemeinsamkeiten gründlich erörtert wurden. Dadurch kam es zwischen den Gesprächspartner*innen immer wieder zu kontroversen Diskussionen, in deren Folge häufig eine lautsprachliche Produktion der gewünschten Zielwörter erreicht werden konnte. Andere Hausnummern, Straßen- und Personennamen in den Karten dienten lediglich dem Zweck der vielfältigen Gestaltung der Karten und erfüllten somit eine *dummy*-Funktion. Durch sie sollte darüber hinaus der Wiedererkennungseffekt zwischen den einzelnen Kartenpaaren für die Sprecher*innen minimiert werden. (vgl. [38, S. 480], [39, S. 1485] und [72, S. 26, 28])

Tabelle 5.1: Die Zielwörter des *Map-Task* Szenarios und die darin vorkommenden phonologischen Prozesse.

Phonologische Prozesse	Zielwörter
Regressive Assimilation des Artikulationsortes $[nf] \rightarrow [mf]$ & $[nv] \rightarrow [mv]$	<u>fünf</u> Buchen <u>weg</u> ; Stollen <u>weg</u> ; Kuchen <u>weg</u> ; Gleisen <u>weg</u> ; Sternen <u>weg</u> ; Pfeifen <u>weg</u> ; Blüten <u>weg</u> ; Maisen <u>weg</u> ; Speisen <u>weg</u> ; Binsen <u>weg</u> ; Bären <u>weg</u> ; Reisen <u>weg</u> ; Lichten <u>weg</u> ; Wolken <u>weg</u> ; Campen <u>weg</u> ; Ranken <u>weg</u> ; Blumen <u>weg</u> ; Torten <u>weg</u>
Regressive Assimilation der Artikulationsart $[gʁ] \rightarrow [çʁ]$	Hastig-Ring; Billig-Ring; Honig-Ring; Günstig-Ring; Käfig-Ring; König-Ring; Fönig-Ring; Mutig-Ring; Butig-Ring; Dorfig-Ring; Torfig-Ring; Rosig-Ring; Hedwig-Ring; Hurtig-Ring
Regressive Assimilation der Stimmhaftigkeit $[kg] \rightarrow [g]$ & $[sv] \rightarrow [zv]$	Sackgasse Lukas <u>Witteck</u> ; Jonas <u>Witteck</u>
Frikativierung von $[g]$ $[g] \rightarrow [ç]$	zwanzig; dreißig; vierzig; fünfzig; sechzig; siebzig; achtzig; neunzig
Lenisierung $[t^h] \rightarrow [t]$	Altpapierdepot

Lenisierung und Fortisierung [eg] ↔ [ek ^h]	Frau <u>E</u> ggers ↔ Frau <u>E</u> ckers
Lenisierung & Geminat [t ^h d] → [d:]	Manfred <u>D</u> ahnke
Modifikation der Vokalqualität [i] ↔ [y] & [u] ↔ [ɔ]	<u>D</u> iesterweg ↔ <u>D</u> üsterweg Herr <u>S</u> chulz ↔ Herr <u>S</u> cholz
Velarisierung von [n] und Tilgung von [g] [ng] → [ŋg] [ŋg] → [ŋ] & [nk] → [ŋk]	<u>K</u> ongresshalle <u>H</u> ängende Gärten Manfred <u>D</u> ahnke
Elision des Glottallauts [auʔa] → [aua]	<u>B</u> auallee; <u>S</u> tauallee; <u>N</u> auallee; <u>M</u> auallee; <u>F</u> rau- <u>A</u> llée; <u>B</u> lau- <u>A</u> llée; <u>G</u> rau- <u>A</u> llée; <u>T</u> rau- <u>A</u> llée; <u>T</u> au- <u>A</u> llée; <u>K</u> lau- <u>A</u> llée; <u>H</u> au- <u>A</u> llée; <u>S</u> au- <u>A</u> llée <u>P</u> fau- <u>A</u> llée; <u>F</u> lau- <u>A</u> llée; <u>R</u> au- <u>A</u> llée; <u>K</u> au- <u>A</u> llée; <u>G</u> au- <u>A</u> llée; <u>D</u> au- <u>A</u> llée; <u>S</u> ch <u>au</u> - <u>A</u> llée; <u>W</u> au- <u>A</u> llée
Elision des Plosivs [mb] → [m]	<u>B</u> amberg; <u>R</u> ammberg; <u>F</u> lammberg; <u>L</u> ammberg <u>T</u> ammberg; <u>D</u> ammberg; <u>S</u> chrammberg; <u>S</u> chlammberg; <u>S</u> tammberg; <u>H</u> ammberg; <u>W</u> ammberg; <u>S</u> trammberg; <u>N</u> ammberg; <u>M</u> ammberg
Schwa-Elision [əl] → [l] & [ən] → [n]	<u>P</u> udelpark <u>H</u> ängende Gärten <u>B</u> uchenweg; <u>S</u> tollenweg; <u>K</u> uchenweg; <u>G</u> leisenweg; <u>S</u> ternenweg; <u>P</u> feifenweg; <u>B</u> lütenweg; <u>M</u> aisenweg; <u>S</u> peisenweg; <u>B</u> insenweg; <u>B</u> ärenweg; <u>R</u> eisenweg; <u>L</u> ichtenweg; <u>W</u> olkenweg; <u>C</u> ampenweg; <u>R</u> ankenweg; <u>B</u> lumenweg; <u>T</u> ortenweg

In Vorbereitung auf das Produktionsexperiment wurden die im Sinne des *Map-Task* Szenarios erstellten Straßenkarten mehrmals überprüft und korrigiert. Zudem fand ein Probedurchlauf mit einem Sprecherinnenpaar statt, das anschließend nicht an den Aufnahmen des SPID-Korpus beteiligt war. Die Probandinnen dieses Tests waren eine Studentin und eine Doktorandin der Allgemeinen und Vergleichenden Sprachwissenschaft an der CAU Kiel und zu diesem Zeitpunkt 28 und 35 Jahre

alt. Der Probedurchlauf fand in einem schallgeschützten Raum statt. Anhand der Ergebnisse dieses Testsprechens wurde überprüft, ob die Zielwörter für jedes Kartenpaar von den Dialogpartnerinnen genannt wurden und welche Schwierigkeiten sich bei der Bewältigung der Sprechaufgabe aufgrund der Gestaltung der Straßenkarten ergaben. Daraufhin fand eine erneute Überarbeitung der Karten statt. Neben dieser qualitativen Korrektur der Karten wurde auch der zeitliche Aufwand berücksichtigt, den die Sprecherinnen für die Bewältigung der jeweiligen Routenplanungen benötigten. Tabelle 5.2 zeigt auf, welche Routenplanung im Probedurchlauf wieviel Zeit erforderte. Dabei konnte bei einigen Kartenpaaren der zeitliche Rahmen von etwa 6 Minuten eingehalten werden, während andere Karten einer erneuten Überarbeitung zur Reduktion der benötigten Gesprächsdauer bedurften.

Tabelle 5.2: Die für die Bewältigung der *Map-Task* Aufgaben benötigte Zeit des Sprecherinnenpaars im Vortest.

Karte	Dauer (Sek.)
1	466,446
2	549,802
3	415,367
4	375,620
5	361,886
6	274,609
7	612,834
8	766,065
9	456,620
10	340,704

5.2 Aufnahmebedingungen

Die Sprachaufnahmen des SPID-Korpus wurden sowohl in der akustischen und visuellen Umgebungssimulation als auch während einer echten Autofahrt durchgeführt. Dabei wurde in beiden Aufnahmesituationen dasselbe Auto verwendet (Mercedes S-Klasse Limousine) (vgl. Kapitel 4). Die Aufnahmen in der Simulation fanden in einer stillen Referenzbedingung und bei Hintergrundgeräuschen von 50 km/h und

130 km/h statt sowie ohne und mit Verwendung des ICC-Systems (vgl. Kapitel 2.3). Daraus ergeben sich $3 \times 2 = 6$ Aufnahmebedingungen im Labor. In der ersten von zwei Aufnahmesitzungen wurden zunächst nur Dialoge ohne die Unterstützung des ICC-Systems aufgenommen. Anschließend erfolgten Dialoge bei aktiviertem ICC-System. In jeder dieser beiden Aufnahmesitzungen fanden zudem im Anschluss an die Aufnahmen im Labor weitere spontansprachliche Dialoge im fahrenden Auto statt. Dazu mussten die Lautsprecher der Umgebungssimulation aus dem Innenraum des Fahrzeugs entfernt werden, wodurch die Proband*innen Zeit für eine kurze Pause von etwa 15 bis 30 Minuten hatten. Während der echten Autofahrt fanden Aufnahmen bei Fahrgeschwindigkeiten von 50 km/h und 130 km/h statt, welche gemäß der jeweiligen Aufnahmesitzung ebenfalls zunächst ohne und später mit ICC-System aufgenommen wurden. Auf eine weitere stille Referenzsituation wurde dabei verzichtet, da sie bereits für die Aufnahmen im Labor im selben Auto erfolgt war. In Tabelle 5.3 sind alle zehn Aufnahmebedingungen des SPID-Korpus zusammengefasst. In jeder Aufnahmebedingung bewältigten die Dialogpaare jeweils eine Sprechaufgabe des *Map-Task* Szenarios, die in Kapitel 5.4 näher erläutert sind. Dafür wurden insgesamt zehn verschiedene Kartenpaare erstellt. Zwischen der ersten Aufnahmesitzung ohne ICC-System und der zweiten mit ICC-System fand eine Unterbrechung des Produktionsexperiments von drei bis vier Monaten statt. Diese Pause sollte einem erhöhten Lerneffekt der Proband*innen hinsichtlich der Sprechaufgabe sowie einer übermäßigen Gewöhnung an die Aufnahmesituation entgegenwirken. Zudem sollte durch die Zweiteilung der Aufnahmen die Länge der Sitzungen reduziert werden, damit Ermüdungserscheinungen die Konzentrationsfähigkeit der Proband*innen weniger stark beeinflussten. (vgl. [38, S. 480] und [39, S. 1486])

Tabelle 5.3: Die Aufnahmebedingungen des SPID-Sprachkorpus.

1. Aufnahmesitzung ohne ICC		2. Aufnahmesitzung mit ICC	
Simulation	Autofahrt	Simulation	Autofahrt
Stille	-	Stille	-
50 km/h	50 km/h	50 km/h	50 km/h
130 km/h	130 km/h	130 km/h	130 km/h

Da der Blickkontakt einen Parameter darstellt, der unter Lombard-Bedingungen gesteigert wird (vgl. Kapitel 3), sollte dieser während der Aufnahmen ausgeklammert werden. Deshalb saß eine Probandin bzw. ein Proband auf dem Beifahrersitz, während die Gesprächspartnerin bzw. der Gesprächspartner dahinter Platz nahm, sodass kein Blickkontakt zwischen ihnen möglich war. (vgl. [38, S. 480] und [39, S. 1486]) Die Instruktionen der Teilnehmer*innen sind in Kapitel 5.4 ausführlich beschrieben. Während der Aufnahmen im fahrenden Auto befanden sich zusätzlich zwei weitere Personen im Fahrzeug. Eine Mitarbeiterin bzw. ein Mitarbeiter der Technischen Fakultät fuhr das Auto. Die Experimentleiterin saß hinter dem Fahrersitz und steuerte mithilfe eines Laptops Start und Ende der Aufnahmen.

Die Reihenfolge der Geräuschbedingungen wurde innerhalb der Aufnahmen von Labor und Fahrt sowie innerhalb der einzelnen Aufnahmesitzungen durch ein *Python*-Skript (vgl. [93]) randomisiert (vgl. [39, S. 1486]). Auch die einzelnen Kartenaufgaben des *Map-Task* Szenarios wurden innerhalb dieser Situationen in eine Zufallsreihenfolge gebracht. Um in nachfolgenden Untersuchungen zusätzlich den Einfluss der verschiedenen Routenplanungen auf die Sprachaufnahmen in der Simulation untersuchen zu können, wurden sechs dieser Routenplanungen ausschließlich für die Sprachaufnahmen im Labor (Kartenpaar 2, 5, 6, 7, 8 und 10) und vier lediglich für die echte Autofahrt verwendet (Kartenpaar 1, 3, 4 und 9). Für diese Aufteilung wurden die Kartenpaare ebenfalls randomisiert. (vgl. [38, S. 480])

Um im fahrenden Auto bei gleichbleibender Geschwindigkeit Sprachaufnahmen von etwa sechs Minuten Länge durchführen zu können, wurden bestimmte Routen in und um Kiel gewählt, die möglichst wenige Unterbrechungen des fließenden Verkehrs beinhalten. Abbildung 5.3 stellt dar, welche Routen verwendet wurden. Auf der linken Seite der Darstellung ist die Route eingezeichnet, welche in und um Kiel für die Aufnahmen bei 50 km/h befahren wurde. Wenn die Länge dieser Strecke für die Teilnehmer*innen nicht ausreichte, um ihre Sprechaufgabe zu lösen, konnte am südlichen Ende der Route über die Edisonstraße und die Liebigstraße eine Schleife gefahren werden. Das nördliche Ende der Strecke bildet die Technische Fakultät der CAU Kiel in der Kaiserstraße 2, die als Ausgangspunkt der Fahrten diente. Rechts ist die Autobahnstrecke zu sehen, welche für die Aufnahmen bei 130 km/h befahren wurde. Die Strecke führt von Kiel über die A215 auf die A7 bis nach Neumünster. Waren die Proband*innen bis Neumünster nicht mit ihrer Sprechaufgabe fertig,

wurde auch die Rückfahrt nach Kiel dafür genutzt. Entsprechend der Randomisierung der Fahrgeschwindigkeiten unterlag auch die Reihenfolge beider Routen dieser Zufälligkeit. Bei Unterbrechungen im Verkehrsfluss, wie z. B. durch Ampeln, Bauustellen, usw., wurde die Sprechaufgabe kurz unterbrochen und wieder fortgesetzt, sobald die gewünschte Geschwindigkeit erneut erreicht war.

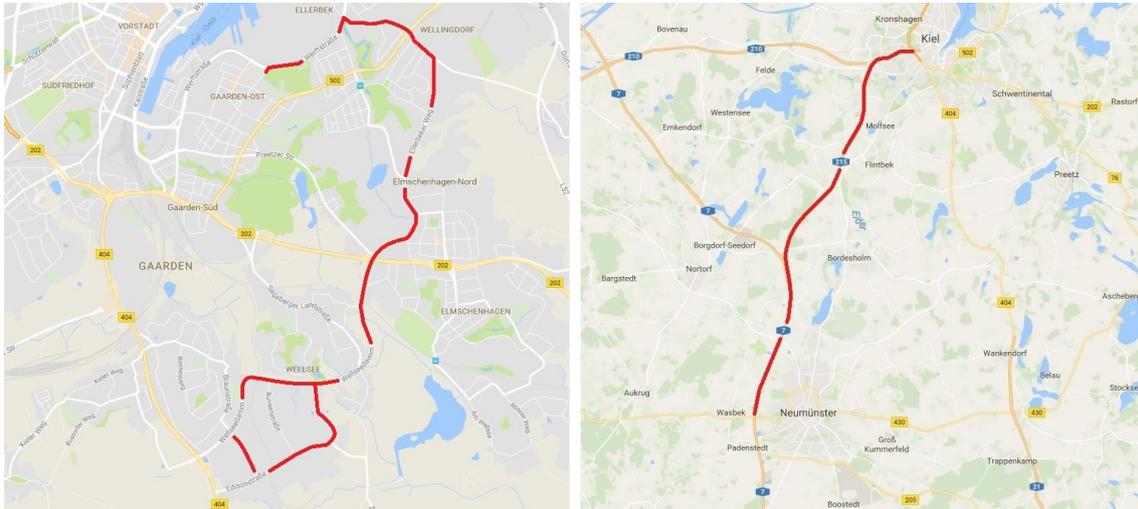


Abbildung 5.3: Die für die Aufnahmen während der echten Autofahrt verwendeten Routen in und um Kiel, jeweils rot gekennzeichnet. Links die Strecke für Sprachaufnahmen bei 50 km/h, rechts die Strecke für Sprachaufnahmen bei 130 km/h. (vgl. [87])

Die Sprachaufnahmen des SPID-Korpus, die im fahrenden Auto durchgeführt wurden, sind für die vorliegende Untersuchung nicht relevant. Ihre Erstellung dient dem Zweck, nachfolgenden Untersuchungen eingehendere Analysen der Sprachaufnahmen sowie einen weiterführenden Vergleich der Sprechsituation im Labor und während der Autofahrt zu ermöglichen. Für die Beantwortung der Fragestellungen der vorliegenden Arbeit (vgl. Kapitel 1.2) wurden lediglich die in der Simulation aufgenommenen spontansprachlichen Dialoge untersucht (vgl. [39, S. 1486]). Bei einer nachfolgenden Überprüfung der Sprachaufnahmen wurde festgestellt, dass die Aufnahmen von Paar 3 ohne ICC-System im fahrenden Auto bei 50 km/h und 130 km/h beschädigt sind und deshalb nicht für weitere Analysen verwendet werden können.

Die Sprache der einzelnen Sprecher*innen wurde durch zwei verschiedene Mikrophone aufgenommen. Eines davon war durch eine Schwanenhalskonstruktion zur Linken der Teilnehmerin bzw. des Teilnehmers stationär im Autoinnenraum befestigt, das

andere wurde durch ein kopfhörerloses Headset jeweils rechts nahe der Mundöffnung positioniert und bewegte sich mit dem Kopf der sprechenden Person im Raum. Die stationären Mikrophone waren mit einem Popschutz ausgestattet, welcher die Geräusche der Verschlusslösung von Plosiven dämpfte. (vgl. [26, S. 190]) Die Kanaltrennung der Aufnahmen wurde dadurch realisiert, dass beide Sprecher*innen jeweils in ein eigenes Headset und eine eigene Schwanenhalskonstruktion sprachen. Um weitergehenden Studien eine visuelle Untersuchung der Aufnahmen zu ermöglichen, wurde auf dem Amaturenbrett des Autos zudem eine *Webcam* angebracht, durch die u. a. das in Abbildung 5.4 dargestellte Bild entstand. Es zeigt den Ausschnitt eines während der echten Autofahrt aufgenommenen Dialogs bei 50 km/h. Das Bild veranschaulicht zudem die Anbringung der Kopfhörer. Zur Unterbrechung der Sprechaufgabe aufgrund der Verkehrssituation gaben die bzw. der Fahrer*in oder die Experimentleiterin ein kurzes Signal. (vgl. [38, S. 480])



Abbildung 5.4: Eine kurze Unterbrechung der Sprechaufgabe während Aufnahmen im fahrenden Auto bei 50 km/h.

5.3 Die Proband*innen

Am Produktionsexperiment der vorliegenden Studie partizipierten insgesamt 8 Männer und 8 Frauen. Die Teilnehmer*innen waren zwischen 22 und 31 Jahren alt. Während der ersten Aufnahmesitzung ohne die Verwendung des ICC-Systems betrug das Durchschnittsalter 26,4 Jahre, in der zweiten Aufnahmesitzung mit ICC-System 26,6 Jahre. Alle Proband*innen sind Muttersprachler*innen des Standarddeutschen. Zum

Zeitpunkt der Aufnahmen waren fast alle Teilnehmer*innen Studierende an der CAU Kiel, wobei eine Probandin Doktorandin und ein Proband berufstätig war. Keine*r der Sprecher*innen hat Phonetik studiert oder Informationen über den Zweck der Aufnahmen bzw. die Hintergründe des vorliegenden Forschungsvorhabens erhalten. (vgl. [38, S. 480] und [39, S. 1486])

Um die Möglichkeit einer anschließenden Analyse von Geschlechterunterschieden in den Aufnahmen zu gewährleisten, wurden die Proband*innen in Sprecher*innenpaare gleichen Geschlechts eingeteilt, sodass 4 männliche und 4 weibliche Dialogpaare entstanden. Darüber hinaus diente diese Vorgehensweise der Minimierung von Entrainment-Effekten (vgl. Kapitel 4.2). Zur Förderung eines authentischen, persönlichen und entspannten spontansprachlichen Dialogs während der Aufnahmen wurden zudem für jedes Dialogpaar Proband*innen ausgewählt, die gut miteinander befreundet oder zumindest gut miteinander bekannt waren. (vgl. [38, S. 480] und [39, S. 1486])

5.4 Instruktionen

Zu Beginn des Produktionsexperiments wurden die Proband*innen darüber informiert, dass es insgesamt zwei Aufnahmesitzungen im Abstand von einigen Monaten geben würde, die jeweils etwa drei bis vier Stunden lang seien. Zudem wurde ihnen mitgeteilt, dass Sprachaufnahmen bei verschiedenen Lärmbedingungen im stehenden Auto sowie bei unterschiedlichen Geschwindigkeiten im fahrenden Auto unter Verwendung desselben Fahrzeugs aufgenommen werden würden. Hierzu wurde das Prinzip der akustischen und visuellen Umgebungssimulation (vgl. Kapitel 4) kurz erläutert. (vgl. [38, S. 480])

Zu Beginn der ersten Aufnahmesitzung sollten die Sprecher*innen der jeweiligen Dialogpaare entscheiden, ob sie die Rolle des *Instruction-Givers* oder die des *Instruction-Followers* übernehmen wollten, ohne zu wissen, welche Bedeutung diese Rollen in der Sprechaufgabe haben würden. Dies diente dem Zweck einer zufälligen Auswahl der Sprechrollen. Anschließend erhielten beide Teilnehmer*innen eine speziell auf ihre Rolle zugeschnittene schriftliche Aufgabenbeschreibung. Um die spontansprachliche Produktion sowie die Realitätsnähe des *Map-Task* Szenarios weiterhin zu unterstützen, wurde diese Aufgabe in einen realistischen Kontext gesetzt, bei dem es sich

um eine kurze Geschichte handelte, mit der die Proband*innen sich identifizieren konnten. Dies sollte zudem die Anzahl der lexikalischen Einheiten im Gespräch steigern, welche gegenüber dem *Video-Task* Paradigma ansonsten vermutlich geringer ausgefallen wäre (vgl. [88]).

Die gesamte Aufgabenbeschreibung ist im Anhang dieser Arbeit aufgeführt (vgl. Anhang A.2). Die darin enthaltene Geschichte beschreibt, dass der *Giver* zur Hochzeit eines Cousins eingeladen ist und zugesagt hat, auf dem Weg dorthin weitere Gäste abzuholen. Besagter Cousin hat dem *Giver* eine selbst gezeichnete Straßenkarte übergeben, in der die Häuser der anderen Gäste eingetragen sind. Die Skizze enthält jedoch zahlreiche Fehler, da Personen, Straßennamen und Hausnummern teilweise falsch eingetragen worden sind. Aufgrund eines kurzfristigen Termins kann der *Giver* nun allerdings nicht mehr zur Hochzeit kommen. Daher bittet sie bzw. er den *Follower*, an ihrer bzw. seiner Stelle zur Hochzeit zu fahren und die eingeladenen Personen abzuholen. Während der Sprachaufnahmen war es die Aufgabe des *Givers*, dem *Follower* mündlich den Weg zu erklären und dabei alle abzuholenden Personen zu nennen. Der *Follower* zeichnete die Route infolgedessen auf einer fehlerfreien Straßenkarte nach, welche bereits die Häuser verschiedener Personen enthielt. Die Proband*innen wurden darüber informiert, dass beide Karten eines Dialogpaares sich in einigen Punkten voneinander unterscheiden. (vgl. [38, S. 480])

Zum Zwecke der Aufgabenbewältigung mussten die Karten des *Givers* und des *Followers* abgeglichen und Unstimmigkeiten zwischen ihnen aufgedeckt werden. Die Proband*innen wurden gebeten, eine möglichst genaue Beschreibung der eigenen Karte durchzuführen und alle Personen, Straßennamen und Hausnummern zu nennen, die darauf abgebildet sind. Bei Unstimmigkeiten in den Straßennamen sollte im Zweifelsfall immer der *Follower* Recht behalten, da sie bzw. er eine fehlerfreie Karte besaß. Bei Unstimmigkeiten in den Namen und Hausnummern der Personen sollte immer der *Giver* Recht behalten, da sie bzw. er die Informationen darüber von ihrem bzw. seinem Cousin erhalten hatte.

Nachdem beide Sprecher*innen ihre Lektüre der Aufgabenbeschreibung beendet hatten, fasste die Experimentleiterin diese noch einmal mündlich zusammen und beantwortete Fragen dazu. Für eine zufällige Verteilung der Sitzplätze (Beifahrer*innensitz und rechter Rücksitz) und der Sprechrollen (*Giver* und *Follower*) sollten die Proband*innen zu Beginn der ersten Aufnahmesitzung diese selbst wählen.

Sitzplatz und Sprechrolle mussten dann bis zum Ende des Produktionsexperiments beibehalten werden, um sie als Einflussfaktoren in der anschließenden Untersuchung der Sprachproduktion jeder einzelnen Teilnehmerin bzw. jedes einzelnen Teilnehmers ausschließen zu können. Am Ende der ersten Aufnahmesitzung wurden die Proband*innen darauf hingewiesen, dass in der zweiten Aufnahmesitzung ein ICC-System verwendet werden würde. Darauf folgte eine kurze Beschreibung der Funktionalität dieses Systems, welche zu Beginn der zweiten Aufnahmesitzung noch einmal wiederholt wurde. (vgl. [38, S. 480f.] und [39, S. 1486])

Zu Beginn der Aufnahmen in den jeweiligen Sitzungen wurden die Proband*innen aufgefordert, ihr Handy auszuschalten, um Störsignale in den Aufnahmen zu vermeiden. Zudem wurden sie gebeten, ein starkes Vorbeugen oder Umdrehen zur Kommunikationspartnerin bzw. zum Kommunikationspartner zu vermeiden, um eine starke Schwankung der Mikrofon- und Sprecher*innenposition zu verhindern und Blickkontakt zu vermeiden. Letzteres sollte die Realitätsnähe der Gesprächssituation steigern, da sich die Passagier*innen im fahrenden Auto beim Sprechen häufig nicht in die Augen sehen können (vgl. [26, S. 190]). Das Vorbeugen bzw. Umdrehen sollte darüber hinaus durch Anschnallen der Sprecher*innen reduziert werden. Die Proband*innen erhielten jeweils ein Klemmbrett und zwei Stifte, um sich während der Sprechaufgabe Notizen in ihren Karten machen zu können. Art und Menge des Geschriebenen wurde ihnen selbst überlassen. Weiterhin wurden die Teilnehmer*innen wiederholt darum gebeten, jeden Straßen- und Personennamen sowie jede Hausnummer, die auf der eigenen Karte oder auf der Karte ihrer Gesprächspartnerin bzw. ihres Gesprächspartners verzeichnet war, im Laufe der Diskussion mindestens einmal zu nennen. Die Gründe hierfür wurden offen gelassen. Des Weiteren wurden die Teilnehmer*innen darüber informiert, dass sich die Mikrophone des ICC-Systems in den Gurten befänden.

Da der Aufbau der sichtbaren *Hardware* für die akustische und visuelle Umgebungssimulation mehr Zeit in Anspruch nimmt als der Abbau, wurden in jeder der beiden Aufnahmesitzungen zunächst die Sprachaufnahmen in der Umgebungssimulation durchgeführt, anschließend die *Hardware* abgebaut und daraufhin die Aufnahmen im fahrenden Auto realisiert. Dadurch konnte die ohnehin sehr zeitaufwendige Aufnahmesitzung insgesamt etwas verkürzt werden, sodass der Ermüdungseffekt bei den Proband*innen weniger stark ausfallen würde. Vor Beginn der echten Auto-

fahrt wurden die Teilnehmer*innen zudem darauf hingewiesen, dass Ereignisse im Straßenverkehr, die eine starke Reduktion der Geschwindigkeit erfordern, eine kurze Unterbrechung der Sprechaufgabe nötig machen würden. Dazu wurde ein gemeinsames Signal vereinbart, bei dem die Fahrerin bzw. der Fahrer oder die Experimentleiterin ein kurzes „Stop“ rief. Am Ende der zweiten Aufnahmesitzung füllten die Teilnehmer*innen einen dreiseitigen Fragebogen mit zehn Teilfragen aus, der eine subjektive Bewertung der Dialogbedingungen ohne und mit ICC-System beinhaltete. Dieser Fragebogen ist im Anhang aufgeführt (vgl. Anhang A.3).

5.5 Geräuschbefreiung

Da die im Labor aufgenommenen Sprachsignale sehr leise sind, wurden diese um 6 dB verstärkt. Um die Segmentierung und Annotation sowie die phonetischen Analysen der Sprachsignale zu verbessern, wurden sie anschließend durch die zwei oben beschriebenen Methoden der Signalverarbeitung prozessiert (vgl. Kapitel 4.1). Trotz der vorgenommenen Kanaltrennung (vgl. Kapitel 5.2) sind die jeweiligen Dialogpartner*innen in den Sprachaufnahmen der Sprecher*innen z. T. leise zu hören. Da die Sprachsignale beider Gesprächspartner*innen durch die Kanaltrennung gesondert aufgenommen wurden, konnte jedoch eine automatische Separation der Sprachsignale jedes Dialogpaares durchgeführt werden (vgl. [26, S. 195f.]). Dadurch wurden die im Hintergrund teilweise wahrnehmbaren Sprachsignale der Gesprächspartner*innen aus den Signalen entfernt. Für jede Aufnahme wurden somit die folgenden Teilschritte durchgeführt, sodass zusätzlich zum Originalsignal jeweils vier unterschiedlich prozessierte Versionen des Signals vorhanden sind:

1. Verstärkung um 6 dB
2. Verstärkung um 6 dB + NC
3. Verstärkung um 6 dB + NC + NS
4. Verstärkung um 6 dB + NC + NS + Separierung

Um die Funktionalität dieser einzelnen Prozessierungen stärker zu verdeutlichen, sind diese im Folgenden bildhaft dargestellt. Abbildung 5.5 zeigt zeitgleiche Ausschnitte aus den Aufnahmen zweier männlicher Dialogpartner in der akustischen

und visuellen Umgebungssimulation bei einer Geräuschbelastung von 130 km/h ohne ICC-System. An dieser Stelle findet eine teilweise Überlagerung ihrer Sprachsignale statt. Das obere Signal wurde von Sprecher 3f produziert, wobei das Wort <okay> artikuliert wird. Auf das die Silbenkoda von [ok^he:] bildende [e:] folgt eine kurze Phase des Ausatmens. Das darunter abgebildete Signal zeigt die Sprachproduktion von Sprecher 3e, welcher die Phrase <ha: wie heißt das bei dir? wuss> artikuliert. Die Laute [ha:] sind rosa markiert und überlagern sich mit der Sprachproduktion von Sprecher 3f, was durch den ebenfalls rosa gekennzeichneten Zeitabschnitt im oberen Sprachsignal zu erkennen ist. Die hier dargestellten Sprachsignale wurden um 6 dB verstärkt.

Durch die zusätzliche Anwendung der Signalverarbeitungsmethoden *noise cancellation* (NC) und *noise suppression* (NS) fand eine wesentliche Reduktion der Umgebungsgeräusche in den Sprachsignalen statt (vgl. Abbildung 5.6 und 5.7). Durch eine weitere Separierung der Sprachsignale beider Dialogpartner verschwindet das Sprachsignal der im Hintergrund sprechenden Person fast vollständig aus den Aufnahmen (vgl. Abbildung 5.8).

5 SPontaneous In-car Dialogues (SPID)

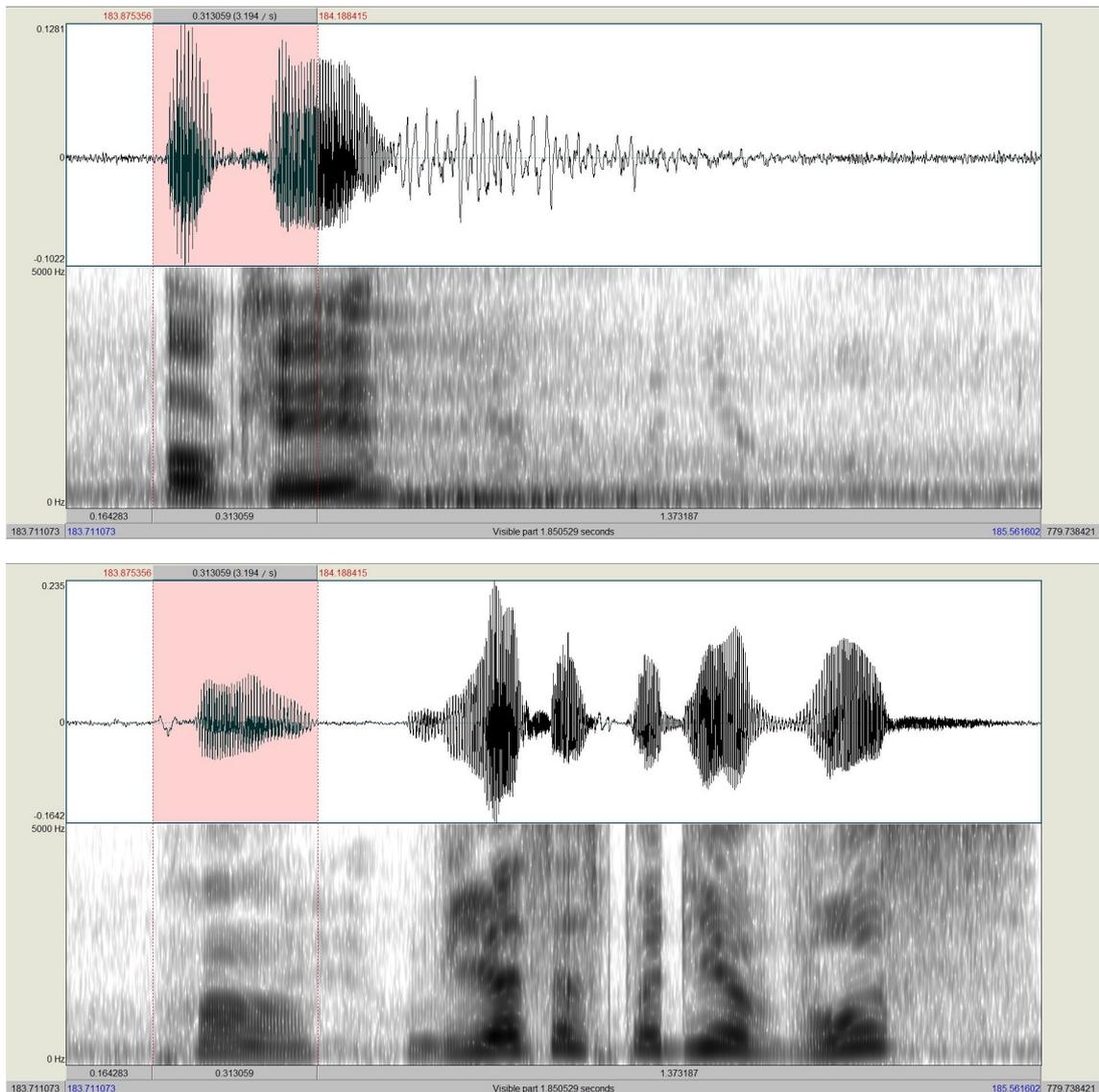


Abbildung 5.5: Verstärkung der Sprachsignale um 6 dB. Zeitgleiche Ausschnitte aus den Sprachsignalen zweier Dialogpartner bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Oben das Sprachsignal von Sprecher 3f (<okay>), unten das von Sprecher 3e (<ha: wie heißt das bei dir? wuss>), dessen Produktion der Laute [ha:] rosa markiert ist, ebenso wie der entsprechende Zeitabschnitt im oberen Sprachsignal von Sprecher 3f.

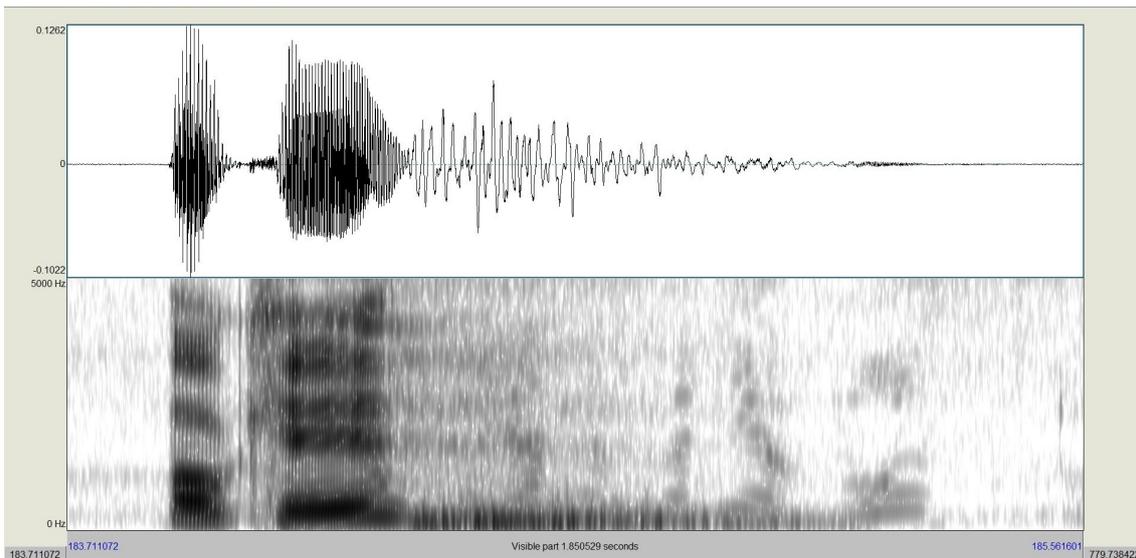


Abbildung 5.6: Verstärkung des Sprachsignals um 6 dB und *noise cancellation* (NC). Lautsprachliche Produktion von <okay> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System.

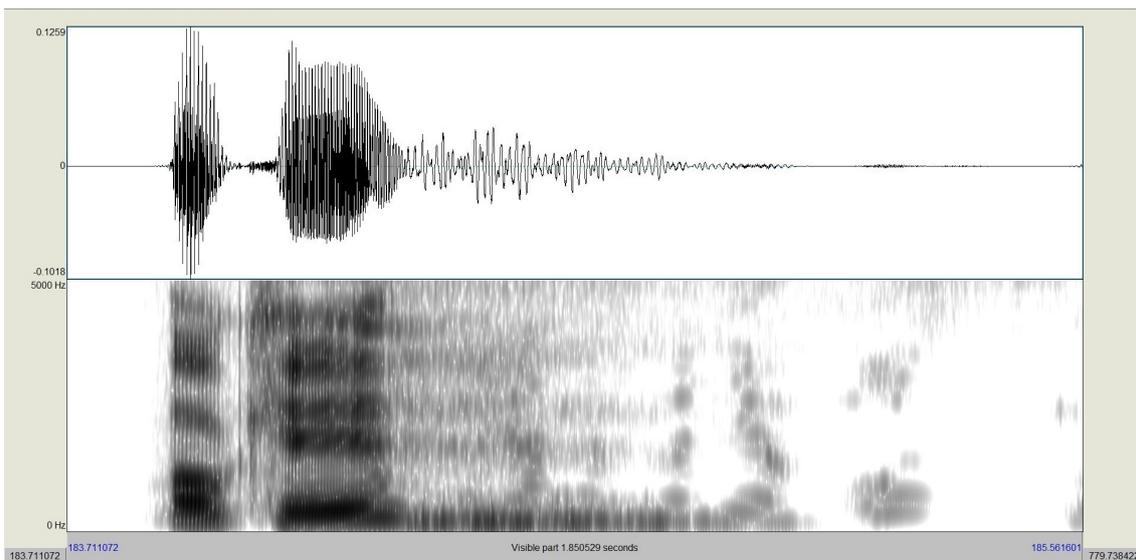


Abbildung 5.7: Verstärkung des Sprachsignals um 6 dB, *noise cancellation* (NC) und *noise suppression* (NS). Die Produktion von <okay> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System.

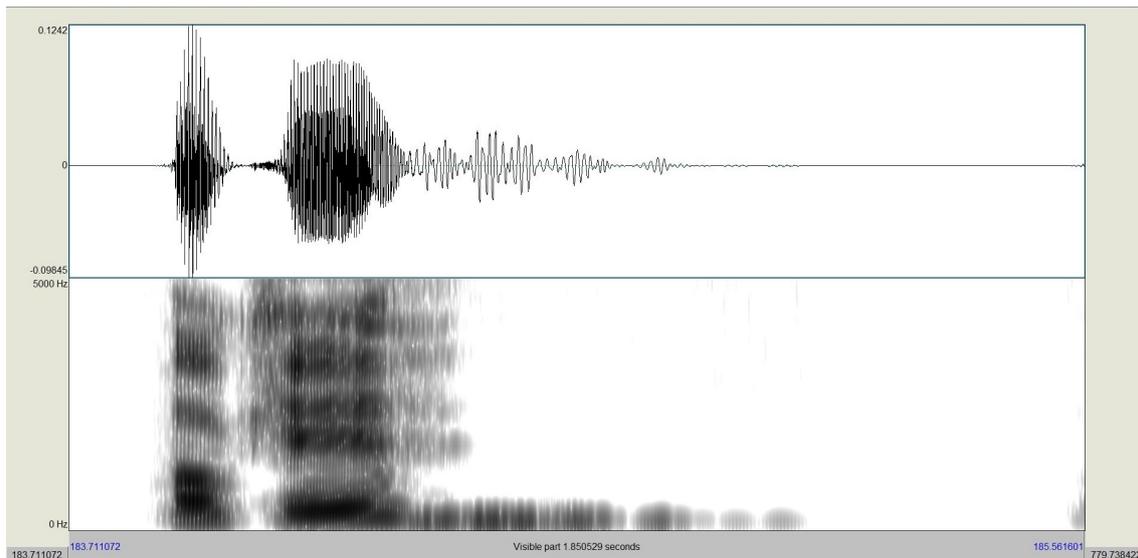


Abbildung 5.8: Verstärkung des Sprachsignals um 6 dB, *noise cancellation* (NC), *noise suppression* (NS) und Separierung. Die Produktion von <okay> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System.

5.6 Dauer der Aufnahmen

Die Aufnahmesitzungen ohne ICC-System waren zwischen $2 \frac{1}{4}$ und 3 Stunden lang, die Aufnahmesitzungen mit ICC-System zwischen $1 \frac{1}{4}$ und $3 \frac{1}{2}$ Stunden. Die Reaktionen der Proband*innen auf das Produktionsexperiment waren sehr unterschiedlich. Einige hatten großen Spaß am Lösen der Sprechaufgaben und freuten sich nach der ersten Aufnahmesitzung auf die Fortsetzung. Andere deuteten im Anschluss an das Experiment an, dass es nicht länger hätte dauern dürfen, da sie ansonsten aufgrund des zeitlichen Aufwands keine Freude mehr daran gehabt hätten.

Da in der akustischen und visuellen Umgebungssimulation häufig bereits einige Sekunden vor dem Start der eigentlichen Kartenaufgabe Sprache aufgenommen wurde und die Aufnahmen erst ein paar Sekunden nach dem Lösen der Aufgabe endeten, wurden alle vier Versionen eines Sprachsignals (vgl. Kapitel 4.1) gleichermaßen am Anfang und am Ende gekürzt. Das zu diesem Zweck verwendete *Matlab*-Skript (vgl. [79]) ist in Kapitel 5.7.2 beschrieben.

Die im Labor aufgenommenen spontansprachlichen Dialoge des SPID-Korpus umfassen dadurch insgesamt etwa 15 Stunden und 47 Minuten Sprachmaterial. Dabei wurden die Aufnahmen jeder Sprecherin bzw. jedes Sprechers innerhalb eines Dia-

logpaares einzeln aufsummiert. Der kürzeste Dialog ist etwa 4,6 Minuten lang, der längste etwa 23,6 Minuten. Tabelle 5.4 fasst die Dauern der Sprachaufnahmen in den sechs Aufnahmebedingungen der akustischen und visuellen Umgebungssimulation zusammen. (vgl. [38, S. 480] und [39, S. 1486])

Tabelle 5.4: Die Dauer der Sprachsignale des SPID-Korpus in der akustischen und visuellen Umgebungssimulation.

Simulation ohne ICC-System		
Stille	50 km/h	130 km/h
3 Std. 23 min.	3 Std. 7 min.	2 Std. 46 min.

Simulation mit ICC-System		
Stille	50 km/h	130 km/h
2 Std. 5 min.	2 Std. 10 min.	2 Std. 16 min.

Die während der echten Autofahrt durchgeführten Sprachaufnahmen haben eine Gesamtlänge von etwa 9 Stunden 55 Minuten. Hierbei muss jedoch berücksichtigt werden, dass diese Aufnahmen bisher nicht geschnitten worden sind. Sie enthalten außer den durch das *Map-Task* Szenario entstandenen Dialogen auch Gespräche der Autoinsass*innen, die aufgrund von Geschwindigkeitsbegrenzungen oder roten Ampeln während Unterbrechungen der Sprechaufgabe stattfanden. Diese sind z. T. mehrere Minuten lang und verzerren die Gesamtlänge der für das Korpus durchgeführten Aufnahmen (vgl. Tabelle 5.5). [39, S. 1486]

Tabelle 5.5: Die Dauer der Sprachsignale des SPID-Korpus während der echten Autofahrt.

Autofahrt ohne ICC-System	
50 km/h	130 km/h
3 Std. 0 Min.	2 Std. 19 Min.

Autofahrt mit ICC-System	
50 km/h	130 km/h
2 Std. 30 Min.	2 Std. 6 Min.

5.7 Segmentierung, Transkription und Annotation

Die orthographische Transkription von Sprachsignalen des SPID-Korpus wurde manuell durchgeführt. Eine anschließende Segmentierung und orthographische Annotation in Praat (vgl. [86]) erfolgte semi-automatisch mithilfe von MAUS (*Munich AUtomatic Segmentation system*) (vgl. [29], [30], [31] und [64]). MAUS kann in Windows und Linux ausgeführt werden. Das System wurde im Rahmen des CLARIN-Projekts (*Common Language Resources and Technology Infrastructure*) (vgl. [83]) als Web-Service frei zugänglich gemacht: (vgl. [31, S. 30f.] und [89])

„CLARIN-D stellt eine Forschungsinfrastruktur für die Geistes- Kultur- und Sozialwissenschaften dar, die Forschende beim Auffinden, Aufbewahren und Auswerten von Forschungsdaten unterstützt.“ ([82])

„Werkzeuge zur Analyse von Forschungsdaten sind über das Web zugänglich und können dadurch ohne zusätzlichen Aufwand verwendet werden. CLARIN-D, der deutsche Partner in der *Common Language Resources and Technology Infrastructure*, macht Werkzeuge für die Geistes- Kultur- und Sozialwissenschaften verfügbar, so dass unterschiedliche Werkzeuge für neue Forschungsfragestellungen zusammen verwendet werden können.“ ([82])

In MAUS wird ein sogenanntes „segmentation and labelling (S&L)“ ([31, S. 30]) durchgeführt, d. h. eine Anpassung der symbolischen Repräsentation der aufgenommenen Sprache an die dazugehörigen Abschnitte des Signals. Für die Segmentierung und Annotation eines Sprachsignals benötigt MAUS die folgenden zwei Eingaben: (vgl. [29, S. 106], [64, S. 607f.] und [31, S. 30])

1. Das Sprachsignal.
2. Eine orthographische oder phonologische Transkription.

Die MAUS-Resultate werden u. a. als TextGrid-Dateien für Praat ausgegeben und beinhalten neben einer in SAMPA (vgl. [32, S. 148ff.] und [92]) realisierten Transkription die dazugehörige Segmentierung (vgl. [29, S. 106] und [31, S. 30]). Zudem werden stille Intervalle des Signals von MAUS automatisch mit dem *Label* <p:> versehen. John et al. (2013) beschreiben in ihrer Untersuchung zur Funktionsweise

von MAUS, dass für eine erfolgreiche Segmentierung eine detaillierte kanonische Transkription erforderlich ist. Wenn ein Sprachsignal zusätzlichen Lärm enthält, führt dies zu Fehlinterpretationen im *alignment*-Vorgang von MAUS, d. h. in der Segmentierung und in der Anpassung der Annotation an das Sprachsignal. (vgl. [25, S. 6])

Die Verwendung eines automatischen Systems zur Transkription und Segmentierung von Sprachsignalen hat den Vorteil, dass hierfür wesentlich weniger Zeit benötigt wird als bei einer rein manuellen Vorgehensweise. Des Weiteren wird dadurch eine gewisse Konsistenz der Arbeit sichergestellt, die bei manueller Umsetzung nicht ausreichend vorhanden wäre, da kein Mensch denselben Satz auf genau dieselbe Art und Weise segmentiert wie ein anderer. Zudem kann es selbst bei einer Wiederholung des Transkriptions- und Segmentierungsvorgang durch dieselbe Person zu Abweichungen kommen. Um die Qualität der MAUS-Ausgaben überprüfen zu können, wurden diese in vorangegangenen Studien mit manuellen Segmentierungen verglichen. Da es jedoch keine eindeutig korrekte Segmentierung und Etikettierung eines Sprachsignals gibt, wurden drei unterschiedliche, von Hand durchgeführte Segmentierungen desselben Sprachsignals einander sowie einer automatisch erstellten Segmentierung gegenübergestellt. Dabei kommen zwischen den manuellen Segmentierungen Übereinstimmungen zwischen 93,1 % und 94,4 % zustande. Diese unterscheiden sich stärker von den automatischen Segmentierungen (Übereinstimmungen zwischen 87,5 % und 88,2 %). Der Unterschied beträgt jedoch weniger als 7 %. (vgl. [25, S. 3f.], [29, S. 106, 108f.] und [30, S. 1026])

Im Folgenden wird zunächst die semi-automatische Segmentierung der Laboraufnahmen des SPID-Korpus mithilfe von MAUS beschrieben. Als Grundlage hierfür dienten orthographische Verschriftlichungen, die im zeitlichen Rahmen der vorliegenden Arbeit nicht als phonologische Transkriptionen realisiert werden konnten. In den durch MAUS segmentierten Dateien traten unterschiedliche Fehler auf, welche durch verschiedene an der CAU Kiel entwickelte Methoden größtenteils behoben werden konnten. Diese sind in Kapitel 5.7.2 beschrieben. Im Anschluss daran wurden einzelne orthographische Segmentierungen und Annotationen von Zielwörtern des *Map-Task* Szenarios manuell durchgeführt, um nachfolgenden Untersuchungen als Grundlage zu dienen (vgl. Kapitel 6.2).

5.7.1 Die semi-automatische Segmentierung durch MAUS

Zur Durchführung der Annotationen wurde zunächst die Anwendung „*Basic MAUS*“ bzw. „*WebMAUS Basic*“ benutzt, die im Internet zur freien Verfügung steht. Dazu mussten eine einfache Textdatei sowie das Sprachsignal in den Web-Service hochgeladen werden. Diese Anwendung ist sinnvoll für einzelne Signal- und Textdateien, wobei keine zusätzlichen Optionen wählbar sind (vgl. [31, S. 32], [90] und [91]). Als Resultat gibt *WebMAUS Basic* eine TextGrid-Datei für Praat aus, welche drei verschiedene Ebenen beinhaltet. Die Ebene „ORT“ enthält eine wortweise Segmentierung und orthographische Kennzeichnung des Signals. „KAN“ beinhaltet ebenfalls eine Segmentierung nach Worten, jedoch mit einer kanonischen Annotation in SAMPA. Die dritte Ebene „MAU“ ist durch eine phonemische Segmentierung mit SAMPA-Symbolen gekennzeichnet (vgl. [91]).

In den orthographischen Transkriptionen der SPID-Sprachsignale wurden die deutschen Phoneme /y/ bzw. /ʏ/, /ø/ bzw. /œ/ und /ɛ/ zunächst durch die Grapheme <ü>, <ö> und <ä> repräsentiert. Weiterhin wurde der stimmlose, alveolare Frikativ /s/ nach den Regeln der deutschen Rechtschreibung (vgl. [85]) teilweise durch <ß> gekennzeichnet. Beides führte zu einer fehlerhaften Umsetzung der entsprechenden Laute durch MAUS, welche in Abbildung 5.9 dargestellt ist. Die in dieser Abbildung aufgezeigten Laute sind nicht vollständig in Oszillogramm und Spektrogramm zu sehen, da die Segmentierung durch MAUS teilweise fehlerhaft ist. Zudem werden die entsprechenden Laute häufig von dem Wort abgegrenzt, dem sie jeweils zugehörig sind. Die Ausgabedatei weist darüber hinaus in allen drei Ebenen „ORT“, „KAN“ und „MAU“ Zeichenfolgen auf, welche nur annähernd die artikulierten Lautfolgen repräsentieren.

Aufgrund dieser Fehler wurden die Grapheme <ü>, <ö> und <ä> in den orthographischen Transkriptionen durch Allographie ersetzt, welche gleichzeitig Symbole des *International Phonetic Alphabet* (IPA) darstellen, die sich im Vokaltrapez in der Umgebung der entsprechenden Laute befinden. Zum Beispiel repräsentiert das Graphem <ä> u. a. den Laut [ɛ]. Das Graphem <e> repräsentiert u. a. den halbgeschlossenen, vorderen, ungerundeten Vokal [e], der sich im Vokaltrapez in der Nähe des halboffenen, vorderen, ungerundeten Vokals [ɛ] befindet. Deshalb wurde in der orthographischen Transkription <ä> durch <e> ersetzt. Ebenfalls wurde das Graphem <ß> durch [ss] ersetzt. Diese Vorgehensweise sollte eine bessere Segmentierung

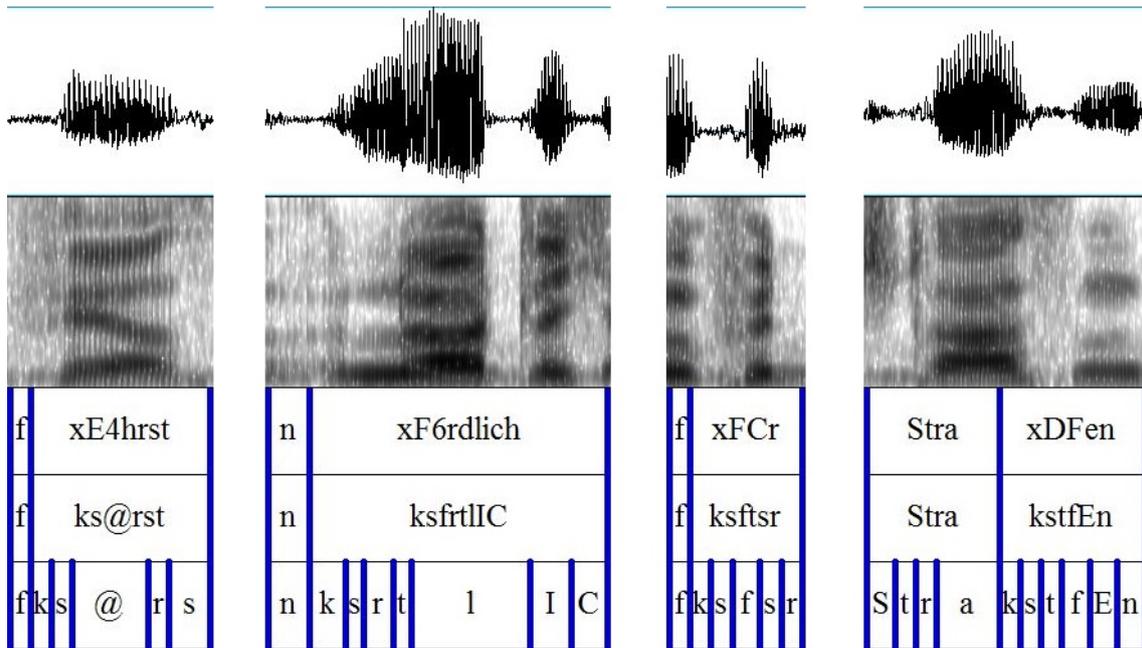


Abbildung 5.9: Die Segmentierung und Annotation der mittels <ä>, <ö>, <ü> und <ß> in den orthographischen Transkriptionen gekennzeichneten Laute durch MAUS. Von links nach rechts sind folgende Wörter dargestellt: <fährst>, <nördlich>, <für> und <Straßen>. Die drei Ebenen in Praat entsprechen von oben nach unten „ORT“ (wortweise Segmentierung mit orthographischer Kennzeichnung), „KAN“ (wortweise Segmentierung mit kanonischer Annotation in SAMPA) und „MAU“ (phonemische Segmentierung mit SAMPA-Symbolen).

und Annotation der Sprachsignale durch MAUS ermöglichen und die anschließend manuell durchgeführten Korrekturen der Ausgabedateien weniger zeitintensiv und damit effektiver werden lassen. Alle Ersetzungen dieser Art sind in Tabelle 5.6 aufgelistet.

Tabelle 5.6: Die Ersetzung der Grapheme <ü>, <ö>, <ä> und <ß> in den orthographischen Transkriptionen des SPID-Korpus.

<ä> → <e>
<ö> → <e>
<ü> → <i>
<ß> → <ss>

Durch die unerwartet langen Sprachsignale des SPID-Korpus (vgl. Kapitel 5.6) nahmen das Hochladen der Sprachdateien und die anschließende Prozessierung durch MAUS je nach Kapazität des verwendeten Computers jeweils mehrere Minuten in Anspruch. Ein großer Vorteil der Arbeit mit MAUS ist dagegen die erhebliche Zeiterparnis, die durch die automatische Segmentierung und Annotation des Sprachsignals entsteht (vgl. [31, S. 30]). Dadurch wird laut Schiel (1999) nicht die Genauigkeit einer manuellen Segmentierung und Annotation erreicht, obwohl sich die Transkripte innerhalb des Leistungsbereichs von erfahrenen Phonetiker*innen befinden (vgl. [64, S. 608]).

5.7.2 Methoden zur verbesserten Arbeit mit MAUS

In der vorliegenden Arbeit wurden, wie in Kapitel 5.7.1 bereits angesprochen, in den Ausgabedateien von MAUS verschiedene Unstimmigkeiten beobachtet, die anschließend eine manuelle Korrektur erforderlich machten. Dabei handelte es sich hauptsächlich um falsch bzw. ungenau gesetzte Wort- und Lautgrenzen im Signal und fehlerhafte Etikettierungen einzelner Laute. Um festzustellen, ob durch die an der CAU Kiel entwickelten Methoden der Geräuschbefeuerung eine Verbesserung der MAUS-Ausgaben erreicht werden kann, wurden zum einen die Originalsignale in die Web-Anwendung hochgeladen und zum anderen Signale, die durch die oben beschriebenen Methoden der Signalverarbeitung prozessiert worden sind (vgl. Kapitel 4.1).

John et al. (2013) erhielten bessere Ergebnisse, wenn lediglich *noise cancellation* (NC) ohne zusätzliche *noise suppression* (NS) auf die von Hintergrundlärm durchgezogenen Sprachsignale angewendet wurde. In den durch MAUS gewählten Wortvarianten existieren Unterschiede für die verbesserten Sprachsignale und die unprozessierten Signale. Es sind jedoch auch zwischen den Methoden NC und NC+S Unterschiede vorhanden. Ein Grund hierfür könnte sein, dass *noise suppression* einige phonetische Hinweise im Signal verdeckt, welche durch *noise cancellation* ohne *suppression* bestehen bleiben. *Noise suppression* unterdrückt dabei die akustische Energie im Signal stärker als *noise cancellation*. In ihrem Beispiel beschreiben die Forscher*innen eine durch NC+S hervorgerufene Unterdrückung des Frequenzbereichs zwischen 4000 Hz und 7000 Hz, durch welche sich die Intensität eines Nasals

nicht länger von der des vorangehenden Vokals unterscheidet. (vgl. [25, S. 5ff.]
 Aus zeitlichen Gründen war in der vorliegenden Arbeit keine phonologische, sondern lediglich eine orthographische Transkription der erhobenen Sprachsignale möglich. Denn obwohl die Ebenen „KAN“ und „MAU“ eine phonologische Annotation der Sprachsignale beinhalten, muss aufgrund der zahlreichen fehlerhaften Segmentierungen und Etikettierungen durch MAUS (vgl. Kapitel 5.7.1) eine umfangreiche manuelle Korrektur erfolgen, die im Rahmen der vorliegenden Forschungsarbeit nicht durchgeführt werden konnte. Ungeachtet der von John et al. (2013) beschriebenen Unterschiede zwischen den Prozessierungsstufen NC und NC+S wurden alle in Kapitel 5.5 aufgeführten Teilschritte zur Geräuschbefreiung bzw. Verbesserung auf die Sprachsignale des SPID-Korpus angewandt. Denn trotz einiger Fehlinterpretationen in der phonemischen Etikettierung der Sprachlaute sollte dadurch eine möglichst detaillierte Segmentierung von sprachlichen und nicht-sprachlichen Abschnitten des Signals durch MAUS erreicht werden.

Abbildung 5.10 verdeutlicht beispielhaft die von MAUS erstellte Segmentierung und Annotation der Laute in einem Sprachsignal des SPID-Korpus, welches nicht durch die Geräuschbefreiung der CAU Kiel prozessiert worden ist. Hierbei handelt es sich um das Signal von Sprecher 3f, welches in der akustischen und visuellen Umgebungssimulation in einer Geräuschumgebung von 130 km/h aufgenommen wurde. In diesem Ausschnitt des Sprachsignals wird die Phrase <den Schulgang nördlich> artikuliert. Anschließend erfolgt die im Hintergrund vom Gesprächspartner produzierte Ergänzung <ja>, welche im Spektrogramm rot eingekreist ist. MAUS erkennt hier weder die korrekten Wortgrenzen des vordergründigen Sprechers noch wird eine Unterscheidung zwischen den Sprachsignalen beider Probanden vorgenommen. Neben dem stark ausgeprägten Hintergrundgeräusch trägt das ergänzende <ja> des Dialogpartners vermutlich zusätzlich zu einer fehlerhaften Segmentierung und Annotation bei, da dieses Wort nicht Teil der in die *Software* eingegebenen orthographischen Transkription ist. Zudem werden Bereiche des Signals, in denen nicht gesprochen wird, von MAUS nur teilweise als Pausen erkannt und größtenteils mit sprachlichen Etiketten versehen. Des Weiteren lässt sich in Abbildung 5.10 erkennen, dass der initiale, stimmlose, postalveolare Frikativ [ʃ] des Wortes <Schulgang> von MAUS fälschlicherweise als Pause markiert wurde. Die korrekte Kennzeichnung dieses Lautes erfolgt erst im Anschluss an diese Pause.

Abbildung 5.11 stellt denselben Ausschnitt des Sprachsignals dar wie Abbildung 5.10, mit dem Unterschied, dass hierbei alle Stufen der Geräuschbefreiung auf das Signal angewendet worden sind (vgl. Kapitel 4.1). Durch den Vergleich beider Abbildungen wird deutlich, dass aufgrund der Geräuschbefreiung eine wesentlich deutlichere Darstellung des Sprachsignals in Oszillogramm und Spektrogramm gelingt. Zum einen sind fast keine Einflüsse des Hintergrundgeräuschs und des Gesprächspartners mehr zu erkennen, zum anderen können einzelne Komponenten des Sprachsignals besser voneinander unterschieden werden. Dies gilt vor allem für den Verlauf der Formanten. Des Weiteren ist mit der Separierung der Sprachsignale auch das hintergründige <ja> des Dialogpartners aus dem Signal verschwunden. In Abbildung 5.11 ist im Spektrogramm lediglich ein Nebengeräusch zu erkennen, welches durch einen roten Kreis markiert ist. Dabei handelt es sich um ein vom Sprecher produziertes Atemgeräusch. Auffällig ist weiterhin, dass durch die Geräuschbefreiung der Signale auch längere Pausen durch MAUS korrekt etikettiert wurden. So wird in Abbildung 5.11 eine akzeptable rechte Grenze des Wortes <nördlich> gesetzt, woraufhin eine längere Pause von 1,01 Sek. folgt. Positiv ist zudem, dass der zuvor als Pause annotierte stimmlose, postalveolare Frikativ [ʃ] nun als Sprachlaut erkannt wird, auch wenn dieser nicht korrekt segmentiert wird.

Die Geräuschbefreiung der Signale ermöglicht demnach eine Verbesserung der Segmentierung und Annotation des Sprachsignals durch MAUS. Dennoch gibt es weiterhin große Bereiche ohne Sprachproduktion und ohne die Merkmale von Hintergrundgeräuschen, die von MAUS ganz oder größtenteils als sprachliche Abschnitte etikettiert werden. Das Wort <den> wird in Abbildung 5.11 zwar als Sprache erkannt, aber fälschlicherweise in die Annotation von <Schulgang> integriert. Stattdessen bildet die Segmentierung und Kennzeichnung von <den> 5,38 Sek. des vorangehenden nichtsprachlichen Bereichs. Fast der gesamte Bereich der Pause wird dabei als sprachlicher Abschnitt gekennzeichnet. Lediglich eine vergleichsweise kurze Region von 0,29 Sekunden wird von MAUS als Pause erkannt.

Die Beobachtung der durch MAUS verarbeiteten Signale legt die Vermutung nahe, dass sich Fehler in der Segmentierung u. a. aufgrund der Länge der aufgenommenen Sprachsignale ergeben. Lange Pausen von mehreren Sekunden scheinen in vielen Fällen ein Problem für die *Software* darzustellen, was sich wiederholt in fehlerhaften Grenzziehungen widerspiegelt. John et al. (2013) beschreiben, dass MAUS bereits

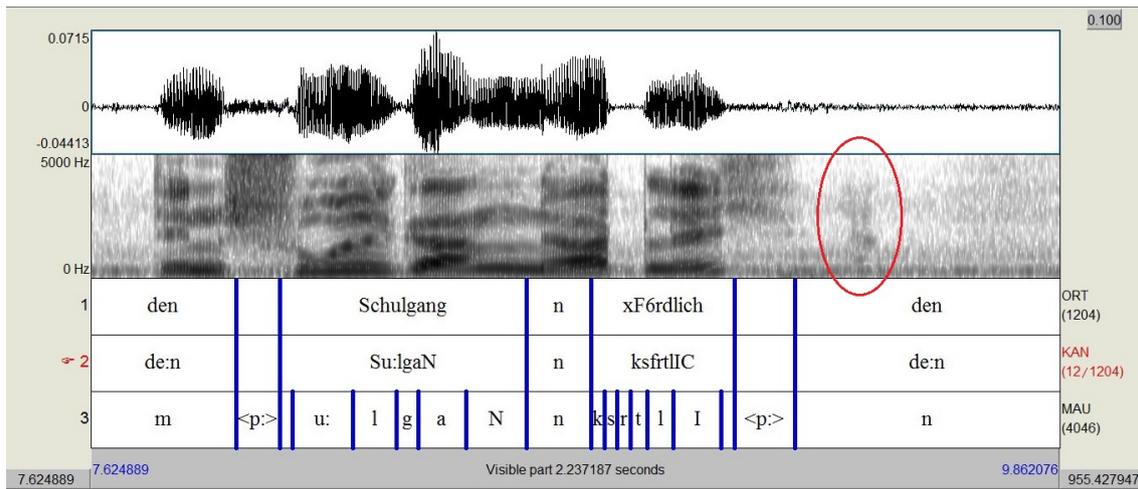


Abbildung 5.10: Die Produktion von <den Schulgang nördlich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System.

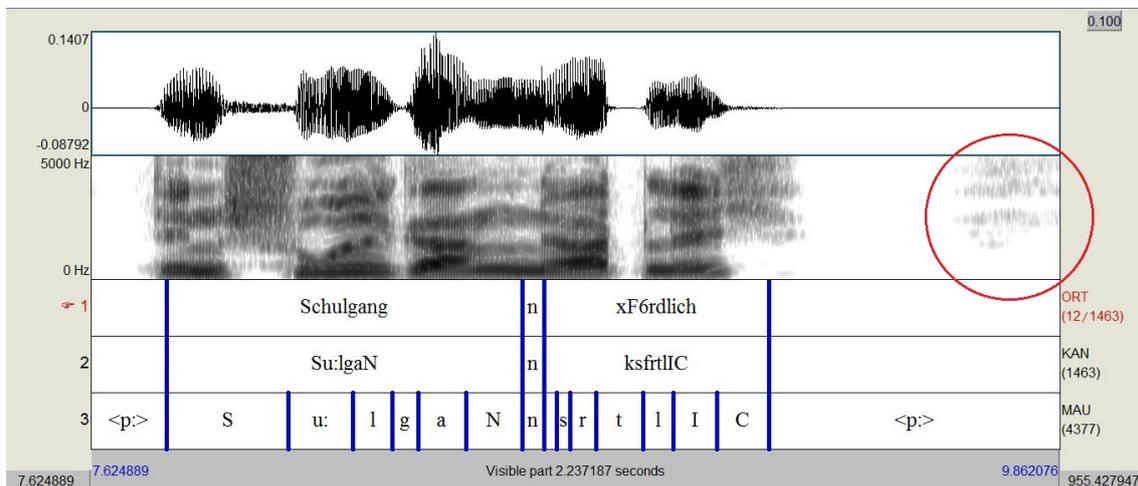


Abbildung 5.11: Die Produktion von <den Schulgang nördlich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Anwendung aller Stufen der Geräuschbefreiung nach Kapitel 4.1.

Phone am lärmdurchzogenen Signal ausrichtet, bevor das eigentliche Sprachsignal begonnen hat. Ein durch MAUS falsch segmentierter Anfang bewirkt Folgefehler, welche sich auf einen unbestimmten Zeitraum im Signal ausdehnen. (vgl. [25, S. 7]) Da es aufgrund der Sprecher*innenwechsel immer wieder zu langen Pausen in den Signalen des SPID-Korpus kommt, erscheint das Phänomen eines zu frühen Setzens einer Lautgrenze durch MAUS sehr häufig, und es sind fortlaufend viele weitere Fehler vorhanden. Deshalb wurde in Zusammenarbeit mit der Arbeitsgruppe DSS der CAU Kiel im Rahmen der vorliegenden Arbeit eine Methode entwickelt, mit welcher die von MAUS durchgeführte Segmentierung und Annotation bessere

Ergebnisse liefert. Diese Vorgehensweise beinhaltet die folgenden Schritte:

1. Durch ein *Matlab*-Skript werden einzelne Phrasen des Sprachsignals als eigene *Sounddateien* von wenigen Sekunden gespeichert. Zudem werden für jede Phrase entsprechende Textdateien erstellt, welche die orthographischen Transkriptionen der artikulierten Laute enthalten.
2. Dasselbe *Matlab*-Skript speichert Pausen zwischen diesen Abschnitten ebenfalls in eigene *Sounddateien*. Des Weiteren werden dazugehörige TextGrid-Dateien erstellt, welche die jeweilige Pause kennzeichnen.
3. Die in Schritt (1) erstellten Paare aus *Sound*- und Textdateien werden gemeinsam in MAUS hochgeladen.
4. Die einzelnen Ausgabedateien von MAUS und die in Schritt (2) erstellten TextGrid-Dateien für Pausen werden mittels eines zweiten *Matlab*-Skripts zu einer einzigen TextGrid-Datei zusammengefügt.

Durch die vier oben beschriebenen Schritte werden die Sprachsignale nicht in einzelne *Turns*, sondern in einzelne Phrasen von wenigen Sekunden eingeteilt, welche dann jeweils in MAUS verarbeitet werden. Dadurch wird eine zu frühe Segmentierung des Signals zuverlässig vermieden und potentiell auftretende Folgefehler werden unterbunden.

Die Schritte (1)-(2) werden durch ein speziell für diesen Zweck erstelltes *Matlab*-Skript durchgeführt. Hierfür werden die vier unterschiedlich prozessierten Versionen des jeweils zu bearbeitenden Sprachsignals benötigt (vgl. Kapitel 4.1), für welche die Originalsignale der *Headsets* verwendet worden sind. Diese sind weniger geeignet für die Methoden der Geräuschbefreiung, da sie nicht fest im Raum platziert sind, weisen jedoch einen wesentlich besseren SNR auf als die stationär befestigten Mikrophone. Des Weiteren muss eine erweiterte orthographische Transkription der im Signal produzierten Sprache vorliegen. Diese wurde durch Zeitangaben ergänzt, welche die Endpunkte der in Schritt (1) jeweils zu speichernden Phrasen beinhalten. Die Zeitangaben wurden dabei in Sekunden mit drei Nachkommastellen dargestellt, welche durch ein Semikolon von dem zum Abschnitt gehörenden orthographischen Text getrennt wurden. Durch dieses *Matlab*-Skript wurde gleichzeitig die in Kapitel 5.6 beschriebene Kürzung der Sprachaufnahmen vorgenommen, um die nicht zur

Sprechaufgabe gehörigen Zeiten am Anfang und am Ende einer Aufnahme zu tilgen. Die Anfangs- und Endzeiten des Signals wurden dem Skript dazu in der erweiterten Transkriptionsdatei durch zeitliche Angaben sowie die Befehle „%Start“ und „%Schluss“ signalisiert. Um dies zu veranschaulichen, ist in Tabelle 5.7 die gekürzte Version einer erweiterten orthographischen Transkriptionsdatei dargestellt.

Tabelle 5.7: Beispiel einer Erweiterung der orthographischen Transkriptionen um die zeitlichen Endpunkte der einzelnen Phrasen, die Pausen und die Start- und Endpunkte des gesamten Signals. Diese Zeitpunkte sind jeweils durch ein Semikolon vom Text bzw. den mit % gekennzeichneten Skript-Befehlen getrennt.

```

15.200;%Start
17.182;oke
22.805;%Pause
24.489;den Schulgang nerdlich
[...]
35.858;%Pause
37.041;em da
37.622;%Pause
38.814;ew w sag das nochmal
[...]
40.051;%Pause
980.381;ich hoffe auch
980.500;%Schluss

```

Aufgrund der zeitlichen Schnittpunkte in dieser erweiterten Transkriptionsdatei werden die jeweiligen Phrasen des Sprachsignals in einzelne *Sounddateien* und die dazugehörigen orthographischen Transkriptionen hinter dem Semikolon in einzelne *Textdateien* gespeichert. Die Grundlage für die einzelnen *Sounddateien* bildet dabei die vollständig prozessierte Version des Sprachsignals. Zusammengehörige Signal- und *Textdateien* erhalten dieselbe Nummer, um später durch MAUS als zusammenhängend identifiziert werden zu können. Zusätzlich werden Pausen zwischen den Sinnabschnitten ebenfalls in der erweiterten Transkriptionsdatei zeitlich markiert (vgl. Tabelle 5.7). Das Prozentzeichen vor dem Wort <Pause> signalisiert dem *Matlab*-Skript, dass hierzu keine einzelnen Signal- und *Textdateien* erstellt werden sollen. Stattdessen wird für den entsprechenden Zeitabschnitt eine *Textgrid-Datei* mit

dem Inhalt „<p:>“ und den dazugehörigen zeitlichen Informationen im Signal erstellt. Das Ergebnis sind, je nach Länge des ursprünglichen Signals, mehrere hundert Signal- und Textdateien sowie zahlreiche TextGrid-Dateien, die Pausen im Signal kennzeichnen. Pausen werden ebenfalls in einzelne *Sound*-dateien gespeichert, jedoch beim Hochladen aller Dateien in MAUS von der *Software* ignoriert, da es keine dazugehörigen Textdateien gibt.

Anschließend wird Schritt (3) der oben beschriebenen Methode durchgeführt, indem die einzelnen *Sound*- und Textdateien des jeweiligen Sprachsignals zusammen in MAUS hochgeladen und dort verarbeitet werden. Dies gelingt durch die Verwendung von „*Multiple MAUS*“ bzw. „*WebMAUS Multiple*“. Diese Anwendung ist für die Verarbeitung großer Datenmengen geeignet. D. h., es können mehrere Dateien von insgesamt bis zu 300 MB durch Ziehen und Ablegen hochgeladen und verarbeitet werden. Diese Anwendung wurde bereits für über 300 Paare aus Signal- und Textdateien erfolgreich getestet. Es können außerdem zahlreiche Einstellungen im Vorwege der Verarbeitung vorgenommen werden (vgl. [31, S. 32] und [90]). Bei der Verwendung von *WebMAUS Multiple* wurde in den angebotenen Optionen lediglich der Parameter „*ORT tier in TextGrid*“ auf *true* gesetzt.

Schritt (4) wurde durch ein zweites *Matlab*-Skript realisiert. Dieses fügt alle Ausgabedateien von MAUS sowie die durch das erste Skript erstellten TextGrid-Dateien für Pausen zu einer einzigen TextGrid-Datei zusammen. Zunächst stimmte jedoch die Länge dieser gesamten TextGrid-Datei nicht mit der Länge des Sprachsignals überein. Wenn in Praat eine einzelne Sounddatei gemeinsam mit der dazugehörigen, durch MAUS erstellten TextGrid-Datei geöffnet wird, verändert sich die Länge des Einzelsignals, z. B. von 1,148005 auf 1,13 Sekunden. Das Signal wird demnach um wenige Millisekunden kürzer. Bei mehreren hundert Einzelsignalen hat dies erhebliche Auswirkungen auf die gesamte Segmentierung des Sprachsignals. Deshalb passt das zweite Skript zunächst die in den einzelnen TextGrids beschriebenen Längen der Signalabschnitte an die Längen der entsprechenden Einzelsignale an. Anschließend werden alle TextGrid-Dateien zusammengefügt. Die von MAUS einzeln ausgegebenen TextGrid-Dateien enthalten jeweils eine Pause, die dem segmentierten und annotierten Signal vorausgeht, sowie eine Pause, die auf das Signal folgt. Werden zwei dieser TextGrid-Dateien zusammengefügt, existieren zwischen den segmentierten Phrasen des Signals zwei angrenzende Intervalle, die jeweils das Label

„<p:>“ enthalten. Dazwischen wird zusätzlich eine durch das erste *Matlab*-Skript erstellte Pausen-TextGrid-Datei eingefügt. Dies hat zur Folge, dass insgesamt drei Pausen-Intervalle zwischen den Phrasen aufeinanderfolgen, was die Annotationsdatei unnötig komplex und unübersichtlich werden lässt (vgl. Abbildung 5.12). Aus diesem Grund werden durch das zweite *Matlab*-Skript zusätzlich alle aufeinanderfolgenden Pausenintervalle zu einem einzigen Intervall zusammengefügt.

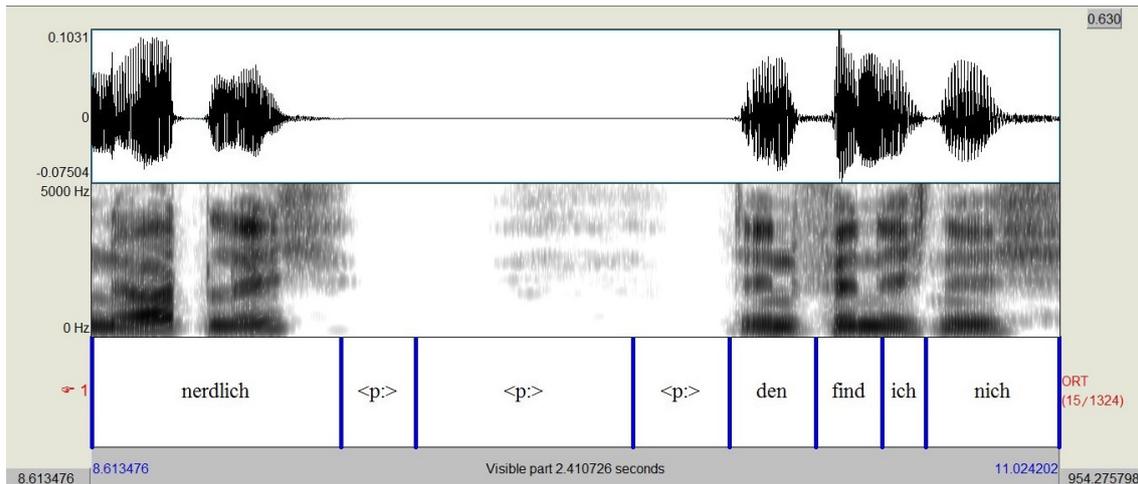


Abbildung 5.12: Die Produktion von <nördlich? Den find ich nich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Anwendung aller Stufen der Geräuschbefreiung nach Kapitel 4.1 sowie der oben beschriebenen Methoden zur Verbesserung der Segmentierung und Annotation durch MAUS.

Abbildung 5.13 zeigt denselben Ausschnitt des Sprachsignals wie die Abbildungen 5.10 und 5.11, mit dem Unterschied, dass hierbei zusätzlich zur Verarbeitung durch MAUS die oben beschriebenen Schritte (1)-(4) angewendet worden sind. Einzelne Segmentgrenzen werden durch MAUS jedoch immer noch etwas ungenau gesetzt. Um die anschließende manuelle Korrektur der Segmentierung möglichst effektiv und mit weniger Zeitaufwand durchführen zu können, sind die Ebenen „KAN“ und „MAU“ durch das zweite *Matlab*-Skript gelöscht worden. Dies birgt keinerlei Nachteile für die Analyse der Sprachsignale in der vorliegenden Studie.

Die Korrekturen der MAUS-Ausgaben durch die oben beschriebene Methode beziehen sich vor allem auf die Verbesserung der Segmentgrenzen zwischen Pausen (nicht-sprachlichen) und sprachlichen Intervallen. Dies ermöglicht z. B. eine genaue Unterscheidung zwischen der Dauer der gesprochenen Sprache und der Dauer der Pausen im Signal. Segmentgrenzen zwischen einzelnen Wörtern wurden vorerst nicht

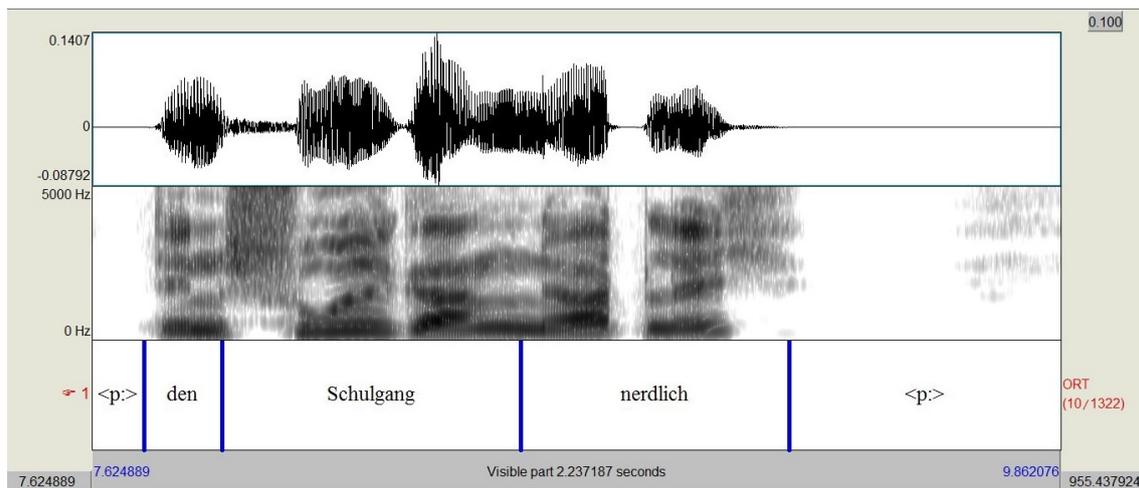


Abbildung 5.13: Die Produktion von <den Schulgang nördlich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Anwendung aller Stufen der Geräuschbefreiung nach Kapitel 4.1 sowie der oben beschriebenen Methoden zur Verbesserung der Segmentierung und Annotation durch MAUS.

korrigiert, da diese für die vorliegende Untersuchung kaum Relevanz besitzen. Lediglich zur Vorbereitung einer Untersuchung der exakten Länge von Zielwörtern des *Map-Task* Szenarios wurden diese Wörter teilweise bereits genau segmentiert. Neben der Korrektur der Segmentgrenzen wurden alle deutschen Umlaute zudem durch die Allographie <ae>, <oe> und <ue> gekennzeichnet (vgl. Abbildung 5.14), damit die orthographische Annotation der Sprachsignale auf internationaler Ebene im ASCII-Format interpretierbar ist. Zudem fand eine an den Regeln der deutschen Rechtschreibung orientierte orthographische Darstellung der Wörter statt. Dabei wurde z. B. das Wort <fahn> in den erweiterten orthographischen Transkriptionen absichtlich ohne den Schwa-Laut dargestellt, um MAUS eine leichtere Annotation zu ermöglichen. In der manuellen Korrektur wurden diese Wörter dann wieder der gültigen Rechtschreibung angeglichen und z. B. als <fahren> dargestellt. Dadurch sollen einzelne Wörter auf orthographischer Ebene schnell in den Daten gefunden werden können. Des Weiteren wurden Häsitationspartikel, wie z. B. <hä>, und Zögerungslängungen, wie z. B. <hm>, durch die Kennzeichnung HES ersetzt, damit nachfolgende Studien die Häufigkeit der in den Signalen auftretenden Häsitationen untersuchen können. Stöhnen, Lachen und andere kommunikative Produktionen von Sprecher*innen wurden ebenfalls als Sprache annotiert, auch wenn es sich hierbei lediglich um die Realisierung einzelner Laute handelte.

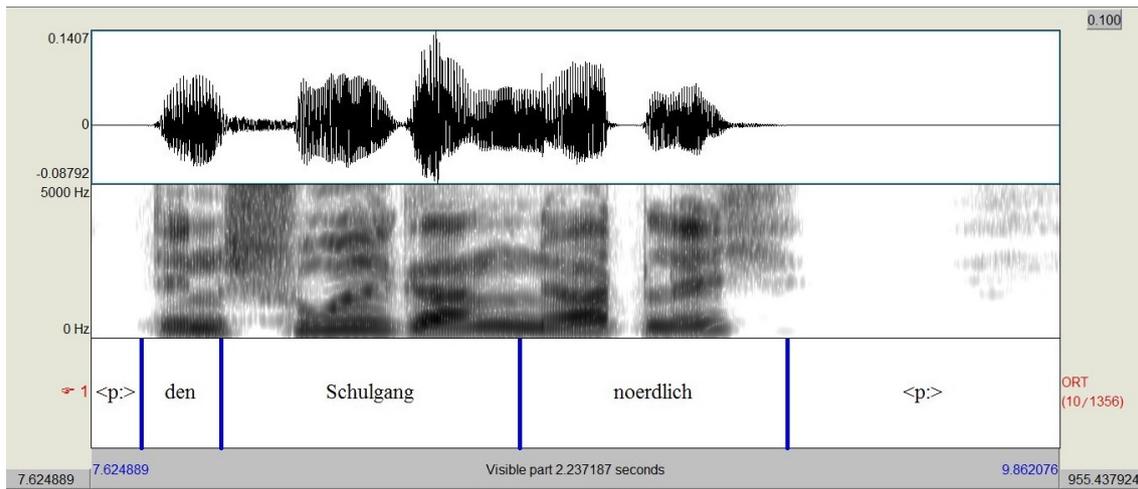


Abbildung 5.14: Die Produktion von <den Schulgang nördlich> durch Sprecher 3f bei 130 km/h in der akustischen und visuellen Umgebungssimulation ohne ICC-System. Anwendung aller Stufen der Geräuschbefreiung nach Kapitel 4.1 sowie der oben beschriebenen Methoden zur Verbesserung der Segmentierung und Annotation durch MAUS. Zusätzliche manuelle Korrektur der Segmentierung und Annotation.

Aus zeitlichen Gründen wurden lediglich die Sprachsignale von fünf Sprecher*innenpaaren (Paar 2, 3, 4, 7, 8) durch MAUS segmentiert und annotiert. Zudem erfolgte die manuelle Korrektur der Segmentgrenzen in diesen Sprachsignalen nur für die ersten fünf Minuten (300 Sekunden) in jedem Signal mit Annotationsdatei. Tabelle 5.8 fasst die durchgeführten Transkriptionen und Annotationen der Sprachsignale noch einmal zusammen:

Tabelle 5.8: Die Transkriptionen und Annotationen der Sprachsignale des SPID-Korpus in der akustischen und visuellen Umgebungssimulation.

Sprachsignal	orth. Transkription	orth. Annotation	Korrektur bis 300 Sek.
Paar 1 ohne ICC	✓	X	X
Paar 1 mit ICC	X	X	X
Paar 2 ohne ICC	✓	✓	✓
Paar 2 mit ICC	✓	✓	✓
Paar 3 ohne ICC	✓	✓	✓
Paar 3 mit ICC	✓	✓	✓
Paar 4 ohne ICC	✓	✓	✓
Paar 4 mit ICC	✓	✓	✓
Paar 5 ohne ICC	✓	X	X
Paar 5 mit ICC	X	X	X
Paar 6 ohne ICC	✓	X	X
Paar 6 mit ICC	X	X	X
Paar 7 ohne ICC	✓	✓	✓
Paar 7 mit ICC	✓	✓	✓
Paar 8 ohne ICC	✓	✓	✓
Paar 8 mit ICC	✓	✓	✓
Gesamtdauer:	13 Std. 1 Min.	8 Std. 59 Min.	0 Std. 50 Min.

6 Die Analyse der Sprachaufnahmen

In diesem Kapitel erfolgt zunächst eine Auswertung des Fragebogens, den die Sprecher*innen am Ende des Produktionsexperiments ausfüllten (vgl. Kapitel 5.4). Anschließend wird zusammengefasst, welche Zielwörter des *Map-Task* Szenarios (vgl. Kapitel 5.1) tatsächlich in jeder Sprachaufnahme von den Proband*innen artikuliert wurden. Weiterhin wird eine akustisch-phonetische Analyse der Sprachsignale durchgeführt, wobei die Auswirkungen der verschiedenen Fahrgeräusche, des ICC-Systems, des Sitzplatzes und der Sprechrolle auf die Parameter der durchschnittlichen Intensität, der durchschnittlichen Grundfrequenz, der Signaldauer (Gesamtdauer eines Dialogs) und der Sprachproduktionsleistung pro Sprecher*in untersucht werden.

6.1 Der Sprecher*innen-Fragebogen

Im Folgenden sind die einzelnen Punkte des Fragebogens aufgeführt, welche die Teilnehmer*innen des Produktionsexperiments am Ende der zweiten Aufnahmesitzung ausfüllten (vgl. Kapitel 5.4). Die Antworten der Proband*innen sind jeweils zusammengefasst und tabellarisch den entsprechenden Fragen zugeordnet, wobei aufgrund der verschiedenen Einstellungen des ICC-Systems für die Sitzplätze (vgl. Kapitel 2) zwischen vorne und hinten sitzenden Sprecher*innen unterschieden wird. Neben dem Ankreuzen der entsprechenden Antworten konnten zu Frage (9) und (10) eigene Kommentare hinzugefügt werden. Des Weiteren bot das Ende des Fragebogens die Möglichkeit, zusätzliche Anmerkungen zum gesamten Produktionsexperiment niederzuschreiben. Die in Frage (9) und (10) sowie in den sonstigen Anmerkungen ergänzten Kommentare der Proband*innen sind im Folgenden ebenfalls aufgeführt.

6 Die Analyse der Sprachaufnahmen

1. Bei der zweiten Aufnahmerunde, die soeben stattgefunden hat: Hast du die Sprache deiner/s Freundin/es häufig über Lautsprecher gehört, oder kam dies nur selten vor?

Antworten	immer	häufig	teilweise	selten	nie	n
vorne	1	5	1	0	1	8
hinten	0	6	2	0	0	8

2. Hast du deine/n Freund/in häufiger über Lautsprecher gehört, je lauter die Hintergrundgeräusche wurden?

Antworten	viel häufiger	etwas häufiger	kein Unterschied	etwas seltener	viel seltener	n
vorne	2	3	3	0	0	8
hinten	5	2	1	0	0	8

3. Wie laut hast du deine/n Freund/in über Lautsprecher gehört in Relation zu ihrer/seiner eigentlichen Stimme?

Antworten	viel lauter	etwas lauter	kein Unterschied	etwas leiser	viel leiser	n
vorne	1	6	1	0	0	8
hinten	1	6	1	0	0	8

4. Hast du deine/n Freund/in mit der Lautsprecher-Übertragung besser verstehen können als ohne?

Antworten	viel besser	etwas besser	kein Unterschied	etwas schlechter	viel schlechter	n
vorne	2	3	3	0	0	8
hinten	4	3	1	0	0	8

Die Fragen (1) bis (4) beziehen sich auf die Wahrnehmung der Stimme der Dialogpartnerin bzw. des Dialogpartners bei aktiviertem ICC-System. Hierbei soll noch keine subjektive Bewertung des Gehörten erfolgen, sondern lediglich eine Einschätzung der Stimmübertragung sowie einer damit einhergehenden Veränderung der Verständlichkeit vorgenommen werden. Es fällt auf, dass lediglich eine Person angegeben hat,

die Gesprächspartnerin bzw. den Gesprächspartner nie über Lautsprecher gehört zu haben. Alle weiteren Proband*innen haben die Lautsprecherübertragung wahrgenommen, wobei zwischen vorne und hinten sitzenden Personen kaum ein Unterschied besteht. Weiterhin hörte niemand die Gesprächspartnerin bzw. den Gesprächspartner seltener über Lautsprecher, je mehr die Fahrgeräusche anstiegen. Im Gegenteil hörte fast die Hälfte der Proband*innen (43,75 %) die andere Person viel häufiger über Lautsprecher, je lauter es wurde. Dies gilt verstärkt für den hinteren Sitzplatz. 31,25 % gaben an, die Gesprächspartnerin bzw. den Gesprächspartner etwas häufiger zu hören, 25 % nahmen keinen Unterschied in der Salienz der Lautsprecherübertragung mit ansteigender Lautstärke wahr. Frage (3) weist keine Differenz zwischen vorne und hinten sitzenden Sprecher*innen auf. Dabei nahm niemand die Dialogpartnerin bzw. den Dialogpartner mit Lautsprechern in Relation zu ihrer bzw. seiner gewöhnlichen Sprechstimme als leiser wahr. 75 % der Proband*innen empfanden die Stimme ihres Gegenübers über Lautsprecher als etwas lauter, während jeweils 12,5 % diese als viel lauter wahrnahmen oder keinen Unterschied hörten. Zudem konnte niemand die Dialogpartnerin bzw. den Dialogpartner aufgrund der Lautsprecherübertragung schlechter verstehen. 25 % nahmen keinen Unterschied wahr, während 37,5 % die andere Person etwas besser und 37,5 % die Gesprächspartnerin bzw. den Gesprächspartner viel besser verstanden, wobei letzteres vermehrt für den hinteren Sitzplatz gilt.

5. Konntest du deine eigene Stimme während der Dialoge über die Lautsprecher hören?

Antworten	Ja	nur selten	nein	n
vorne	1	2	5	8
hinten	4	3	1	8

Bei dieser Frage ist ein deutlicher Unterschied zwischen beiden Sitzplätzen feststellbar. Während 31,25 % der Proband*innen angaben, ihre Stimme nur selten über die Lautsprecher gehört zu haben, hörten sie 37,5 % überhaupt nicht. Im Gegensatz dazu antworteten 31,25 % der befragten Personen mit „ja“. Auffällig ist, dass die Mehrheit der Personen, die ihre eigene Stimme nicht hörten, auf dem vorderen Sitzplatz saßen, während diejenigen, die ihre Stimme hörten, hinten Platz genommen hatten.

6 Die Analyse der Sprachaufnahmen

6. Hat es dich gestört, wenn du deine/n Freund/in über Lautsprecher gehört hast?

Antworten	ja	etwas	kein Unterschied	es war eher angenehm	nein, ich mochte es	n
vorne	0	4	3	1	0	8
hinten	1	2	1	3	1	8

7. Hat dich die Übertragung von Sprache über die Lautsprecher generell gestört?

Antworten	Ja	etwas	kein Unterschied	es war eher angenehm	nein, ich mochte es	n
vorne	0	4	2	1	1	8
hinten	2	2	1	1	2	8

Frage (6) und (7) beziehen sich auf eine subjektive Bewertung der Proband*innen hinsichtlich der Lautsprecherübertragung. Bei Frage (6) bewerteten 25 % der Befragten das Hören der anderen Person über Lautsprecher weder als störend noch als angenehm. 43,75 % fühlten sich hingegen von der Lautsprecherübertragung der Dialogpartnerin bzw. des Dialogpartners gestört, während 31,25 % dies als positiv empfanden. Letzteres gilt für vier Personen auf dem Rücksitz, jedoch lediglich für eine Probandin bzw. einen Probanden des vorderen Sitzplatzes. Ein ähnliches Bild zeichnet sich ab, wenn nach der generellen Störung durch die Lautsprecherübertragung gefragt wird. 18,75 % der Proband*innen stellten keinen Unterschied fest, 50 % der Sprecher*innen fühlten sich gestört und 31,25 % empfanden dies eher als angenehm oder mochten es.

8. Falls du bei Frage 6 und/oder 7 Nummer 1 oder 2 eingekreist hast: Hat es dich gestört, deine eigene Stimme zu hören, oder die deiner/s Freundin/es? (Hierbei sind mehrere Antworten möglich).

Antworten	meine eigene	die meiner/s Freundin/es	beides	keine Antwort	n
vorne	1	2	1	4	8
hinten	1	1	2	4	8

9. Falls du die Sprach-Übertragung durch Lautsprecher als störend empfunden hast, was genau hat dich daran gestört? (Hierbei sind mehrere Antworten möglich).

Antworten	Die Sprache klang merkwürdig	Ich habe meine eigene Stimme durch die Lautsprecher gehört	Die Lautsprecher waren zu laut	Die Lautsprecher waren zu leise	n
vorne	2	1	1	0	4
hinten	4	2	0	0	6

vorne	<ul style="list-style-type: none"> • Es klang sehr mechanisch bzw. dumpf. • Hat etwas gestört, dass die Stimme der Freundin immer zwischendurch (schwankend) über die Lautsprecher kam.
hinten	<ul style="list-style-type: none"> • Bisschen Echo wie durch einen Wasserfall. • Blecherner Klang. • Irritierend. Komischer Schall.

10. Falls dir die Sprach-Übertragung durch Lautsprecher gefallen hat, was genau fandest du daran gut? (Hierbei sind mehrere Antworten möglich).

Antworten	Die Sprache klang natürlich.	Die Sprache war nicht zu laut oder zu leise.	Ich konnte meine/n Freund/in besser verstehen.	Die Kommunikation wurde leichter.	n
vorne	1	1	3	2	7
hinten	1	3	5	5	14

vorne	–
hinten	<ul style="list-style-type: none"> • Das einzige, was seltsam war, war, dass die Stimme wie „aus dem off“ kam und man sich beim 1. Mal hören gewundert hat, wo die Stimme herkam. • Ungewohnt, kein visuelles Feedback bei deutlicherer Stimme des Gesprächspartners zu haben, aber ich habe mich sehr schnell daran gewöhnt.

Durch Frage (8) bis (10) wird festgestellt, was die Proband*innen an der Lautsprecherübertragung der Sprachsignale gestört bzw. was ihnen daran gefallen hat. Bei Frage (8) enthielten sich acht Personen (50 %) einer Antwort. Zwei Sprecher*innen (12,5 %) hat es gestört, die eigene Stimme zu hören, drei Personen (18,75 %) empfanden es als störend, die Stimme der Dialogpartnerin bzw. des Dialogpartners zu hören und drei Teilnehmer*innen (18,75 %) empfanden beides als störend. Dabei sind keine eindeutigen Unterschiede zwischen vorne und hinten sitzenden Personen

zu erkennen. Von denjenigen, die die Lautsprecherübertragung als störend bewerteten, gaben sechs Personen (37,5 %) an, dass die Sprache merkwürdig klang, drei Proband*innen (18,75 %) hörten ihre eigene Stimme durch die Lautsprecher und eine Person (6,25 %) empfand die Übertragung als zu laut (vgl. Frage (9)). An der Sprachübertragung wurde vor allem kritisiert, dass diese sehr mechanisch bzw. dumpf klang, dass etwas Echo zu hören war und der Klang blechern war. Im Gegensatz dazu ging Frage (10) darauf ein, was den Proband*innen, denen die Sprachübertragung durch Lautsprecher gefallen hat, positiv daran aufgefallen ist. Dabei gab es mehr Bewertungen von Sprecher*innen des hinteren Sitzplatzes als von jenen des vorderen Sitzplatzes. Sie antworteten, dass die Sprache natürlich klang (12,5 %), dass die Sprache nicht zu laut oder zu leise war (25 %), dass sie ihre Dialogpartnerin bzw. ihren Dialogpartner besser verstehen konnten (50 %) und dass die Kommunikation leichter wurde (43,75 %). Als Kommentar wurde ergänzt, dass die Wahrnehmung einer deutlichen Stimme der anderen Person ohne visuelles *Feedback* zwar ungewohnt war, die Probandin bzw. der Proband sich jedoch schnell an die neue Situation gewöhnte.

11. Sonstige Anmerkungen

vorne	<ul style="list-style-type: none"> • Stimme klang etwas „metallisch“ / hallend, war aber nicht störend.
hinten	<ul style="list-style-type: none"> • Manchmal klang es etwas blechern, aber das war nicht wirklich störend. • Bei Kopfbewegungen des Partners schwankte die Tonqualität u. die Lautstärke.

Die unter (11) aufgeführten Kommentare bestätigen die bereits zuvor festgehaltenen Antworten der Proband*innen. Hinzu kommt, dass Kopfbewegungen der vorne sitzenden Personen anscheinend die subjektiv wahrgenommene Tonqualität und Lautstärke des übertragenen Sprachsignals beeinflussten. Die Ergebnisse dieser Fragebogenerhebung werden in Kapitel 8.1.4 im Rahmen der zusammenfassenden Ergebnisinterpretation diskutiert.

6.2 Die Zielwörter des *Map-Task* Szenarios

Für nachfolgende Analysen der einzelnen Zielwörter in den Sprachaufnahmen, die im Rahmen des *Map-Task* Szenarios produziert wurden (vgl. Kapitel 5.1), gibt Tabelle 6.1 Aufschluss darüber, welche in jedem der insgesamt 96 im Labor aufgenommenen

Signale produziert wurden. Eine vergleichende Untersuchung dieser Wortproduktionen über alle Experimentbedingungen hinweg ist nur dann sinnvoll, wenn diese in jeder Sprachaufnahme mindestens einmal artikuliert wurden. Es lässt sich jedoch feststellen, dass lediglich vier davon in jeder Aufnahme vorkommen. Dennoch treten auch die anderen Zielwörter in vielen Sprachsignalen auf, sodass Analysen dieser Wörter und damit einzelne Vergleiche unterschiedlicher ICC- und Geräuschbedingungen mit nur geringen Einschränkungen möglich sind. Aufgrund der zeitlichen Begrenzung der vorliegenden Forschungsarbeit konnte keine Analyse einzelner Wortproduktionen vorgenommen werden (vgl. Kapitel 8.3). Durch die Generierung dieses Sprachkorpus ist es jedoch nachfolgenden Studien möglich, Untersuchungen verschiedener Phänomene, wie z. B. unterschiedlicher phonologischer Prozesse, in den Zielwörtern des SPID-Korpus durchzuführen (vgl. Kapitel 8.4).

Tabelle 6.1: Die Produktion von Zielwörtern in den Sprachsignalen der akustischen und visuellen Umgebungssimulation.

Zielwort	Produktion in jeder Sprachaufnahme
Sackgasse	✓
Altpapierdepot	X
-auallee	X
Diesterweg	X
Düsterweg	X
Eckers	X
Eggers	X
-enweg	✓
fünf	X
-igring	✓
Kongresshalle	X
-mberg	X
Pudelpark	✓
Scholz	X
Schulz	X
Hängende Gärten	X
Lukas/Jonas Witteck	X
Manfred Dahnke	X

6.3 Regressionsanalyse der Sprachsignale

Die Analyse der im Folgenden beschriebenen Untersuchungsparameter wird gemäß der in Kapitel 1.2 dargestellten Fragestellung lediglich für jene Sprachaufnahmen durchgeführt, welche in der akustischen und visuellen Umgebungssimulation aufgenommen wurden (vgl. Kapitel 4). Dabei wird überprüft, wie die verschiedenen Untersuchungsparameter bzw. abhängigen Variablen (vgl. [39, S. 1486] und Kapitel 1.2 und 3.1)

- durchschnittliche Intensität,
- durchschnittliche Grundfrequenz,
- Signaldauer,
- Anteil der Sprachproduktion im Signal,

durch folgende unabhängige bzw. kategoriale Variablen beeinflusst werden, die jeweils zwei bis drei unterschiedliche Ausprägungen besitzen: (vgl. [38, S. 481f.] und [39, S. 1486])

- *ICC (ohne ICC, mit ICC)*,
- *Geräusch (Stille, 50 km/h, 130 km/h)*,
- *Rolle (Giver, Follower)*,
- *Sitzplatz (vorne, hinten)*.

Für jeden Untersuchungsparameter wird zudem ein Normalisierungsverfahren angewandt, um die Messwerte einzelner Sprecher*innen vergleichbar zu machen. Dabei wird für jedes Subjekt s über alle sechs Experimentbedingungen i im Labor (vgl. Kapitel 5.2) ein Mittelwert \bar{x}_s errechnet. Dieser wird anschließend von den jeweiligen Individualwerten $x_{s,i}$ der einzelnen Bedingungen subtrahiert, woraus sich die mittelwertbereinigten Werte $x'_{s,i}$ ergeben: (vgl. [38, S. 481f.] und [39, S. 1486])

$$x'_{s,i} = x_{s,i} - \bar{x}_s. \quad ([38, S. 481] \text{ und } [39, S. 1486]) \quad (6.1)$$

Anhand der Abweichungen der Einzelwerte von den individuellen Mittelwerten kann das durchschnittliche Verhalten der einzelnen Sprecher*innen in allen Experimentbedingungen miteinander verglichen werden (vgl. [38, S. 481f.]). Zur statistischen Analyse der Untersuchungsparameter wird eine multiple univariate lineare Regression (*linear least squares regression*), die eine abhängige und mehrere unabhängige Variablen beinhaltet, mit Hilfe der Programmiersprache R durchgeführt (vgl. [38, S. 481f.], [39, S. 1486], [21, S. 363ff.], [77, S. 186ff.], [23, S. 698ff.] und [94]). Dabei werden die erwarteten Werte der abhängigen Variablen durch folgende Regressionsgleichung berechnet: (vgl. [21, S. 363])

$$\hat{Y} = \alpha + \beta_1 X_1 + \dots + \beta_k X_k, \quad ([21, S. 363]) \quad (6.2)$$

wobei \hat{Y} die abhängige Variable bzw. *Responsevariable* (das Kriterium) ist, α und β_k die Regressionskoeffizienten und X_k die erklärenden bzw. unabhängigen Variablen (die Prädiktoren) sind. Die Regressionsparameter α und β werden dabei durch die Kleinst-Quadrat-Schätzung bzw. OLS-Schätzung (*Ordinary Least Squares*) bestimmt. Hierbei werden die Abweichungen der bestimmten von den tatsächlichen Werten (Residuen) quadriert und aufsummiert. Die kleinste Residuenquadratsumme (*Residual Sum of Squares*, RSS) wird zur Schätzung von α und β verwendet. Durch die Funktion „lm(Modellformel)“ wird in R die Berechnung sowohl einer einfachen als auch einer multiplen linearen Regression durchgeführt. Die darin enthaltene Modellformel wird dabei wie folgt aufgebaut: (vgl. [21, S. 363f., 353ff.] und [77, S. 180f., 186])

$$\text{Kriterium} \sim \text{Prädiktor}_1 + \dots + \text{Prädiktor}_p. \quad ([77, S. 186]) \quad (6.3)$$

Im Folgenden sind Analysen der einzelnen Untersuchungsparameter beschrieben. Zur Durchführung der multiplen univariaten linearen Regression kann eine unterschiedliche Anzahl der oben genannten Prädiktoren *ICC*, *Geräusch*, *Rolle* und *Sitzplatz* in die Modellformel 6.3 integriert werden. Dabei muss überprüft werden, welche Prädiktoren im Modell berücksichtigt werden müssen, um die Varianz der *Responsevariablen* am besten erklären zu können (vgl. [23, S. 708]). Hierzu wurde in verschiedenen Regressionsmodellen der Determinationskoeffizient R^2 betrachtet, welcher als „Maß für die Güte der Schätzung“ ([77, S. 184]) gilt. Dieses Bestimmt-

heitsmaß nähert sich dem höchsten Wert 1 an, je besser die Anpassung des Modells an die beobachteten Werte ist (vgl. [23, S. 697]). Das Bestimmtheitsmaß soll demnach möglichst hoch sein. Dabei muss berücksichtigt werden, dass R^2 mit zunehmender Anzahl an Prädiktoren in der Modellformel ebenfalls ansteigt. Durch eine Adjustierung bzw. Anpassung von R^2 ist es allerdings dennoch möglich, Modelle mit verschieden vielen Variablen zu vergleichen. (vgl. [23, S. 708]) Deshalb wird der Erklärungswert der in dieser Arbeit verwendeten Regressionsmodelle im Folgenden am Wert des *Adjusted R-squared* gemessen.

Im Folgenden wird jenes Modell als „bestes Modell“ bezeichnet, welches das *Adjusted R-squared* maximiert. Dieses zeichnet sich durch die bestmögliche Übereinstimmung zwischen der Erklärung der Ergebnisse und der Vermeidung zufälliger Resultate aus. Zur Bildung des besten Regressionsmodells können z. B. unterschiedliche „schrittweise Verfahren (*stepwise regression modeling*)“ ([23, S. 708]) angewendet werden. In der vorliegenden Untersuchung wird eine Methode zur Modellbildung verwendet, die einem sogenannten „Vorwärts-Einschluss“ ([23, S. 708f.]) ähnelt. Dabei werden einem Modell, welches beispielsweise lediglich die Konstante α , jedoch keine Prädiktoren enthält („Nullmodell“ ([23, S. 709])), schrittweise diejenigen Prädiktoren hinzugefügt, durch die das Regressionsmodell verbessert wird. Dies wird daran gemessen, ob das multiple Bestimmtheitsmaß eine Steigerung erfährt. (vgl. [23, S. 708f.]) Diese Methode wird in der folgenden Arbeit per Hand durchgeführt, mit dem Wert des *Adjusted R-squared* als Ausschlusskriterium. Zusätzlich werden diejenigen erklärenden Variablen in die Modellformel der Regression aufgenommen, welche in der jeweils vorangegangenen deskriptiven Untersuchung systematische Einflüsse auf die Messwerte des entsprechenden Untersuchungsparameters erkennen ließen.

In den von R ausgegebenen Regressions*outputs* werden Ergebnisse der Konstanten angezeigt, die mit *Intercept* beschriftet sind, während die darunter liegenden Zeilen Ergebnisse weiterer unabhängiger Variablen aufzeigen (vgl. [21, S. 356, 364]). Um die Zusammenhänge zwischen den verschiedenen Ausprägungen einer bzw. mehrerer erklärender Variablen besser darstellen zu können, wurde die Zeile *Intercept* in den Regressionsmodellen der folgenden Abschnitte entfernt. Dies geschieht, indem den erklärenden Variablen in der Regressionsformel ein „-1“ folgt. Durch diesen Schritt wird die Konstante der Regression auf 0 gesetzt und ermöglicht somit einen besseren Vergleich der Regressionskoeffizienten. Zusätzlich zum Einfluss einzelner

erklärender Variablen auf die *Responsevariable* werden Interaktionen zwischen verschiedenen Ausprägungen der erklärenden Variablen mit in die Analyse einbezogen. Diese Interaktionen werden in der Modellformel durch ein „*“ angezeigt, welches die unabhängigen Variablen verbindet.

Bei der Interpretation einzelner Signifikanzniveaus muss beachtet werden, dass beim Regressionsmodell kein signifikanter Unterschied zwischen einzelnen Variablenausprägungen aufgedeckt, sondern vielmehr die Unterscheidung der entsprechenden Koeffizienten von 0 berechnet wird. Für eine Überprüfung der Signifikanzniveaus zwischen zwei Einflussgrößen bietet sich hingegen die ANOVA (*analysis of variance*, Varianzanalyse) an. Die ANOVA ist ein Spezialfall der Regressionsanalyse, verarbeitet Einflussgrößen aber lediglich als kategoriale Variablen [62, S. 5]. Die Variablen des im Folgenden verwendeten Regressionsmodells sind ebenfalls kategorial modelliert. Der Vorteil der Regressionsanalyse liegt jedoch darin, dass sie auch die Verwendung stetiger Größen als unabhängige Variablen ermöglicht. Wenn weiterführende Analysen zu Kommunikationssystemen im Auto die Geschwindigkeit in der Regressionsanalyse als stetige Größe modellieren, z. B. um aufgrund ihrer Daten Vorhersagen für weitere, nicht gemessene Geschwindigkeiten zu berechnen, sind diese Untersuchungen mit den Methoden und Ergebnissen der vorliegenden Analyse besser vergleichbar.

6.3.1 Intensität

Die Analyse der Intensität in den Sprachaufnahmen der akustischen und visuellen Umgebungssimulation wird mit Hilfe von Praat durchgeführt (vgl. [86]). Das hierzu verwendete Praat-Skript liest nacheinander alle Sounddateien ein, die sich innerhalb eines Ordners befinden, und gibt für jedes Signal verschiedene Intensitätswerte wieder, wie z. B. die durchschnittliche Intensität, welche im Folgenden untersucht wird. Die Messwerte dieses Untersuchungsparameters werden anschließend durch die oben beschriebene Formel 6.1 mittelwertbereinigt. (vgl. [38, S. 481f.])

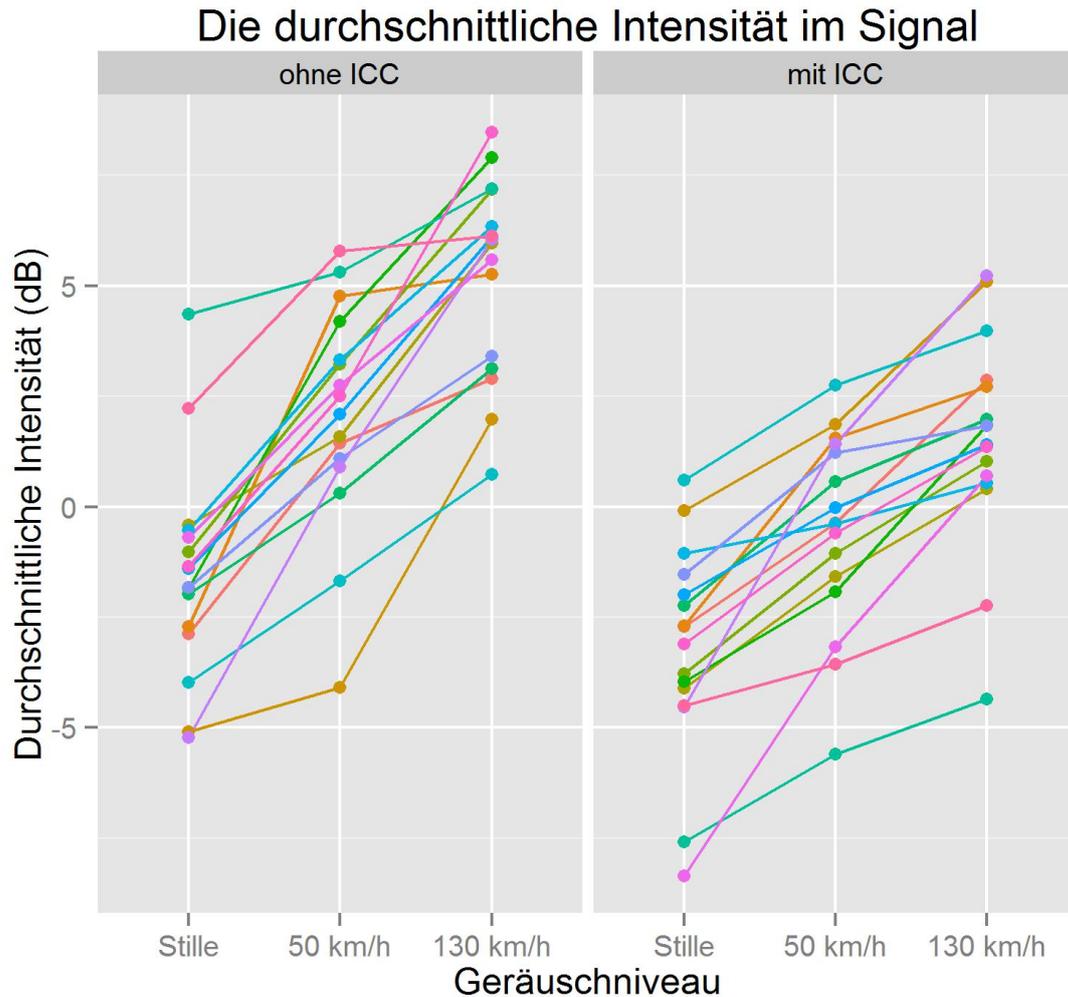


Abbildung 6.1: Die mittelwertbereinigte durchschnittliche Intensität in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren Daten einzelner Proband*innen. (vgl. [38, S. 482])

Das dargestellte Liniendiagramm 6.1 vermittelt einen ersten Eindruck davon, welchen Einfluss die verschiedenen ICC- und Geräuschbedingungen auf die durchschnittliche Intensität der untersuchten Sprachsignale haben. Dieser Messwert steigt ohne und mit ICC-System von Stille zu 50 km/h und von 50 km/h zu 130 km/h an. Darüber hinaus fallen die Messwerte im Mittel um etwa 1,7063 dB (Stille), 2,6519 dB (50 km/h) und 3,7469 dB (130 km/h) niedriger aus, wenn das ICC-System aktiviert ist. (vgl. [38, S. 481f.]) Dies wird durch die Illustrierung der Mittelwerte in Abbildung 6.2 bekräftigt. Dabei lässt sich zudem erkennen, dass ein Unterschied zwischen den Sitzplätzen der Proband*innen im Auto besteht und sich die durchschnittliche Intensität mit ICC-System bereits in der stillen Referenzbedingung verändert.

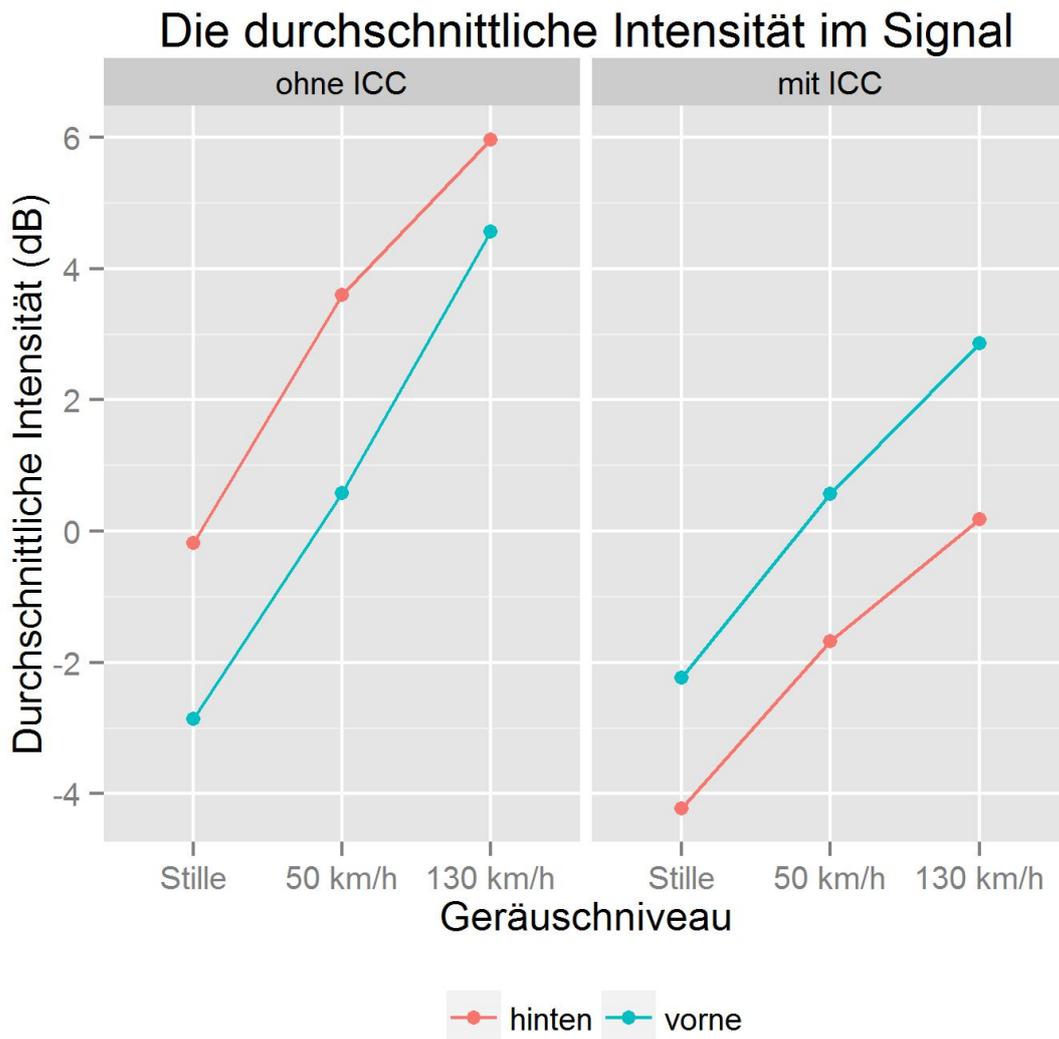


Abbildung 6.2: Die Mittelwerte zur mittelwertbereinigten durchschnittlichen Intensität über alle Sprecher*innen hinweg.

Gemäß den Erläuterungen zur Regressionsanalyse in 6.3 wird zunächst von einem Nullmodell ohne erklärende Variablen ausgegangen und anschließend durch Hinzunahme verschiedener erklärender Variablen überprüft, ob das *Adjusted R-squared* eine Steigerung erfährt. Die *Responsevariable* ist *intensity*. Die erste unabhängige Variable, welche zu diesem Zweck herangezogen wird, ist *ICC*. Die Ergebnisse dieses Regressionsmodells sind in Tabelle 6.2 dargestellt. Das *Adjusted R-squared* ist mit einem Wert von 0,1474 in diesem Modell noch sehr gering. Die Ausprägung *ohne ICC* unterscheidet sich jedoch signifikant von 0 ($p < 0,0005$). Anhand der durch das Modell errechneten Schätzwerte (*Estimate*) lässt sich erkennen, dass die durch-

schnittliche Intensität für Sprachaufnahmen mit ICC-System niedriger ist als für solche ohne ICC-System.

Tabelle 6.2: Regressionsergebnisse für die abhängige Variable *intensity* und die erklärende Variable *ICC*.

```
Call:
lm(formula = intensity ~ ICC - 1, data = dataicc)
```

Residuals:

Min	1Q	Median	3Q	Max
-7.5953	-2.5270	0.3289	2.4572	6.5280

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
ICCmit ICC	-0.7547	0.4841	-1.559	0.122396
ICCOhne ICC	1.9470	0.4841	4.022	0.000117 ***

—

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.354 on 94 degrees of freedom
Multiple R-squared: 0.1652, Adjusted R-squared: 0.1474
F-statistic: 9.301 on 2 and 94 DF, p-value: 0.0002062

In einem nächsten Schritt wird die unabhängige Variable *ICC* im Regressionsmodell durch *Geräusch* ersetzt. Die Ergebnisse dieser Analyse sind in Tabelle 6.3 dargestellt. Das *Adjusted R-squared* ist mit 0,4311 wesentlich höher als im vorangegangenen Modell. Diese Regression besitzt demnach einen höheren Erklärungswert. Zudem ist anhand der Signifikanzniveaus zu erkennen, dass sich die Ausprägungen *Stille* ($p < 0,0001$) und *130 km/h* ($p < 0,0001$) signifikant von 0 unterscheiden, während *50 km/h* einen p-Wert von 0,116 aufweist. Dementsprechend ist die zuletzt genannte Bedingung nicht signifikant unterschiedlich von 0, jedoch von *Stille* und *130 km/h*. Somit besitzen alle drei Geräuschbedingungen einen großen Einfluss auf die Werte der *Responsevariable*. Den dazugehörigen Schätzwerten ist zu entnehmen, dass die Sprachaufnahmen in *Stille* die niedrigste durchschnittliche Intensität aufweisen (-2,3759 dB), während bei *130 km/h* die Werte am höchsten liegen (3,3962 dB).

Tabelle 6.3: Regressionsergebnisse für die abhängige Variable *intensity* und die erklärende Variable *Geräusch*.

Call:

lm(formula = intensity ~ Geräusch - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-7.7562	-1.6303	0.1289	1.8404	6.7359

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
Geräusch130 km/h	3.3962	0.4844	7.012	3.70e-10	***
Geräusch50 km/h	0.7681	0.4844	1.586	0.116	
GeräuschStille	-2.3759	0.4844	-4.905	3.96e-06	***

—
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.74 on 93 degrees of freedom

Multiple R-squared: 0.4489, Adjusted R-squared: 0.4311

F-statistic: 25.25 on 3 and 93 DF, p-value: 4.9e-12

Des Weiteren wird überprüft, ob durch eine Interaktion der unabhängigen Variablen *ICC* und *Geräusch* ein zusätzlicher Erklärungswert des Regressionsmodells erreicht werden kann (vgl. Tabelle 6.4). Dabei steigt das *Adjusted R-squared* auf einen Wert von 0,5737 an. Dieses Regressionsmodell wird daher zunächst als das mit dem höchsten Informationsgehalt beibehalten.

Tabelle 6.4: Regressionsergebnisse für die abhängige Variable *intensity* und die Interaktion der erklärenden Variablen *ICC* und *Geräusch*.

Call:

lm(formula = intensity ~ ICC * Geräusch - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-6.1791	-1.1314	0.1828	1.2345	5.8828

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	1.5228	0.5930	2.568	0.0119	*
ICCoohne ICC	5.2697	0.5930	8.887	5.93e-14	***
Geräusch50 km/h	-2.0806	0.8386	-2.481	0.0150	*
GeräuschStille	-4.7519	0.8386	-5.666	1.73e-07	***
ICCoohne ICC:Geräusch50 km/h	-1.0950	1.1860	-0.923	0.3583	
ICCoohne ICC:GeräuschStille	-2.0406	1.1860	-1.721	0.0888	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.372 on 90 degrees of freedom

Multiple R-squared: 0.6003, Adjusted R-squared: 0.5737

F-statistic: 22.53 on 6 and 90 DF, p-value: 4.799e-16

Auffällig ist, dass in Tabelle 6.4 nicht alle Ausprägungen der erklärenden Variablen *Geräusch* bzw. der Interaktion zwischen *ICC* und *Geräusch* aufgeführt sind. Der Grund hierfür ist, dass jeweils eine der Ausprägungen von *R* „als Referenzkategorie gewählt“ ([21, S. 367]) wird. Durch die Anwendung der Regressionsgleichung 6.2 lassen sich für alle Kombinationen der Variablenausprägungen Werte berechnen. Dabei werden die vom Modell berechneten Schätzwerte in die Gleichung aufgenommen und diejenigen Ausprägungen bzw. Interaktionen auf 1 gesetzt, für die eine Berechnung durchgeführt werden soll. Alle anderen erhalten den Wert 0. Die entsprechende Formel für das Regressionsmodell aus Abbildung 6.4 ist im Folgenden dargestellt, mit \hat{Y} als erklärter Intensität:

$$\begin{aligned}
\hat{Y} = & 0 + 5,2697 \times (\text{ohne ICC}) + 1,5228 \times (\text{mit ICC}) \\
& + (-4,7519) \times (\text{Stille}) + (-2,0806) \times (50 \text{ km/h}) + 0 \times (130 \text{ km/h}) \\
& + (-2,0406) \times (\text{ohne ICC}) \times (\text{Stille}) + 0 \times (\text{mit ICC}) \times (\text{Stille}) \\
& + (-1,0950) \times (\text{ohne ICC}) \times (50 \text{ km/h}) + 0 \times (\text{mit ICC}) \times (50 \text{ km/h}) \\
& + 0 \times (\text{ohne ICC}) \times (130 \text{ km/h}) + 0 \times (\text{mit ICC}) \times (130 \text{ km/h}).
\end{aligned} \tag{6.4}$$

Um diese Methode näher zu erläutern, wird im Folgenden auf Basis der in Tabelle 6.4 aufgeführten Schätzwerte beispielhaft der erklärte Wert für die durchschnittliche Intensität in den Sprachsignalen *ohne ICC* bei *50 km/h* berechnet. In Formel 6.5 werden dementsprechend alle Bedingungen auf 0 oder 1 gesetzt und die *Responsevariable* errechnet, welche in diesem Fall einer Zunahme der durchschnittlichen Intensität um 2,0941 dB entspricht:

$$\begin{aligned}
2,0941 = & 0 + 5,2697 \times 1 + 1,5228 \times 0 \\
& + (-4,7519) \times 0 + (-2,0806) \times 1 + 0 \times 0 \\
& + (-2,0406) \times 1 \times 0 + 0 \times 0 \times 0 \\
& + (-1,0950) \times 1 \times 1 + 0 \times 0 \times 1 \\
& + 0 \times 1 \times 0 + 0 \times 0 \times 0.
\end{aligned} \tag{6.5}$$

Tabelle 6.4 zeigt, dass alle Ausprägungen beider erklärender Variablen mit einem Niveau von $p=0,0119$ und $p=0,0150$ signifikant sind. Die durchschnittliche Intensität ist weiterhin niedriger bei der Verwendung eines ICC-Systems und steigt mit zunehmender Geräuschbelastung an. Im Hinblick auf die Interaktion von *ICC* und *Geräusch* wird lediglich ein geringer signifikanter Trend für die Interaktion zwischen den Ausprägungen *ohne ICC* und *Stille* angezeigt ($p=0,0888$).

Darüber hinaus wird untersucht, welchen Einfluss die unabhängigen Variablen *Sitzplatz* und *Rolle* auf die durchschnittliche Intensität der Sprachsignale besitzen. Dabei wird in Abbildung 6.3 zunächst deskriptiv verdeutlicht, dass eine Wechselwirkung zwischen den unabhängigen Variablen *ICC* und *Sitzplatz* existiert. Während sich auf dem vorderen Sitzplatz mit Aktivierung des ICC-Systems kaum eine Verände-

6 Die Analyse der Sprachaufnahmen

lung einstellt (die Messwerte liegen im Mittel nur geringfügig tiefer), lässt sich auf dem hinteren Sitzplatz eine Abnahme der Messwerte feststellen. Auffällig ist dabei, dass die durchschnittliche Intensität *ohne ICC* hinten höher ist und mit ICC-System vorne höher.

Abbildung 6.4 weist kaum systematische Unterschiede zwischen den Sprechrollen des *Givers* und des *Followers* auf. Mit ICC-System kann im Falle des *Followers* eine niedrigere Varianz vermutet werden als ohne ICC-System. Des Weiteren fallen die Messwerte mit ICC-System für beide Sitzrollen wiederum ab. Eine Interaktion zwischen den unabhängigen Variablen *ICC* und *Rolle* wird hieraus jedoch nicht ersichtlich.

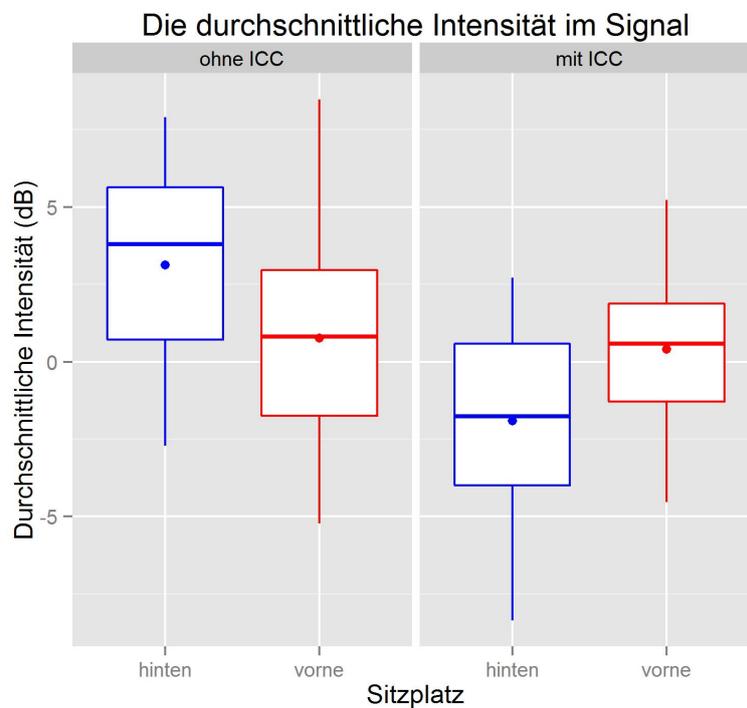


Abbildung 6.3: Die mittelwertbereinigte durchschnittliche Intensität in jeder ICC- und Sitzplatzbedingung.

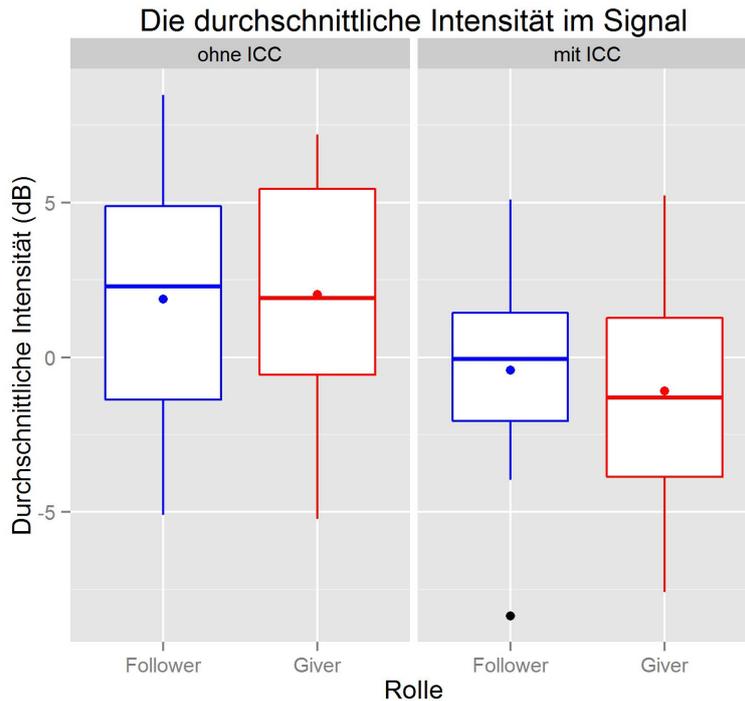


Abbildung 6.4: Die mittelwertbereinigte durchschnittliche Intensität in jeder ICC- und Rollenbedingung.

Angesichts dieser deskriptiven Ergebnisse wird dem bisher besten Modell aus Tabelle 6.4 die unabhängige Variable *Sitzplatz* hinzugefügt (vgl. Tabelle 6.5). Das *Adjusted R-squared* weist dabei einen Wert von 0,5689 auf und ist somit niedriger als im vorangegangenen Regressionsmodell. Damit besitzt dieses Modell einen geringeren Erklärungswert als das vorherige. Die Messwerte des vorderen Sitzplatzes weisen darüber hinaus einen p-Wert von 0,95 auf. Die unabhängige Variable *Sitzplatz* wird nicht in der Modellformel beibehalten, da durch sie kein erhöhter Erklärungswert des Regressionsmodells erreicht werden kann.

Tabelle 6.5: Regressionsergebnisse für die abhängige Variable *intensity*, die Interaktion der erklärenden Variablen *ICC* und *Geräusch* und die erklärende Variable *Sitzplatz*.

Call:

lm(formula = intensity ~ ICC * Geräusch + Sitzplatz - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-6.1638	-1.1433	0.1931	1.2278	5.8675

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	1.53813	0.64408	2.388	0.0190	*
ICCoohne ICC	5.28500	0.64408	8.206	1.64e-12	***
Geräusch50 km/h	-2.08063	0.84330	-2.467	0.0155	*
GeräuschStille	-4.75188	0.84330	-5.635	2.03e-07	***
Sitzplatzvorne	-0.03062	0.48688	-0.063	0.9500	
ICCoohne ICC:Geräusch50 km/h	-1.09500	1.19260	-0.918	0.3610	
ICCoohne ICC:GeräuschStille	-2.04062	1.19260	-1.711	0.0906	.

—
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.385 on 89 degrees of freedom

Multiple R-squared: 0.6003, Adjusted R-squared: 0.5689

F-statistic: 19.1 on 7 and 89 DF, p-value: 2.395e-15

Aufgrund der in Abbildung 6.3 erkennbaren Wechselwirkung zwischen den erklärenden Variablen *ICC* und *Sitzplatz* wird das Regressionsmodell aus Tabelle 6.4 anschließend durch die Interaktion dieser beiden unabhängigen Variablen erweitert. Dadurch erhöht sich das *Adjusted R-squared* auf den Wert 0,6767 (vgl. Tabelle 6.6). Die dargestellte Ausprägung *vorne* ($p < 0,0005$) sowie die Interaktion *ohne ICC, vorne* ($p < 0,0001$) weisen zudem hoch signifikante Unterscheidungen zu 0 auf (vgl. [39, S. 1486]). Die durchschnittliche Intensität ist auf dem vorderen Sitzplatz um 2,3052 dB höher. Auffällig ist, dass die Ausprägung *vorne* in Tabelle 6.5 keinen signifikanten Effekt aufweist, während in Tabelle 6.6 mit einem p-Wert von $p < 0,0005$ das Gegenteil der Fall ist. Der Grund hierfür ist, dass in Tabelle 6.6 in den Werten der Ausprägung

vorne bereits die Interaktion mit der erklärenden Variablen *ICC* einbezogen wird.

Tabelle 6.6: Regressionsergebnisse für die abhängige Variable *intensity*, die Interaktion der erklärenden Variablen *ICC* und *Geräusch* sowie die Interaktion der erklärenden Variablen *ICC* und *Sitzplatz*.

```
Call:
lm(formula = intensity ~ ICC * Geräusch + ICC * Sitzplatz - 1, data = dataicc)
```

Residuals:

Min	1Q	Median	3Q	Max
-4.9958	-1.2829	-0.0526	1.4685	4.6996

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	0.3702	0.5962	0.621	0.536257	
ICCoohne ICC	6.4529	0.5962	10.823	< 2e-16	***
Geräusch50 km/h	-2.0806	0.7302	-2.849	0.005456	**
GeräuschStille	-4.7519	0.7302	-6.507	4.53e-09	***
Sitzplatzvorne	2.3052	0.5962	3.866	0.000211	***
ICCoohne ICC:Geräusch50 km/h	-1.0950	1.0327	-1.060	0.291893	
ICCoohne ICC:GeräuschStille	-2.0406	1.0327	-1.976	0.051284	.
ICCoohne ICC:Sitzplatzvorne	-4.6717	0.8432	-5.540	3.10e-07	***

—

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.065 on 88 degrees of freedom
Multiple R-squared: 0.7037, Adjusted R-squared: 0.6767
F-statistic: 26.12 on 8 and 88 DF, p-value: < 2.2e-16

Des Weiteren wird untersucht, ob Interaktionen zwischen *Geräusch* und *Sitzplatz* bzw. *Geräusch* und *Rolle* einen Einfluss auf die durchschnittliche Intensität besitzen. Die Abbildungen 6.5 und 6.6 bestätigen, dass die Messwerte mit steigender Geräuschbelastung ansteigen. Dabei scheinen die Messwerte der hinten und vorne sitzenden Sprecher*innen bzw. der *Giver* und *Follower* in jeder Geräuschbedingung ähnlich hoch zu sein. Es soll die Frage beantwortet werden, ob durch das Hinzufügen dieser deskriptiv beobachteten Wechselwirkungen der Erklärungswert des Regressionsmodells aus Tabelle 6.6 erhöht werden kann. Durch die Hinzunahme einer

Interaktion zwischen *Geräusch* und *Sitzplatz* fällt das *Adjusted R-squared* auf einen Wert von 0,674 ab, bei einer Hinzunahme der Wechselwirkung von *Geräusch* und *Rolle* auf 0,6698. Zudem sind die hinzugefügten Interaktionen in keinem der beiden Modelle signifikant.

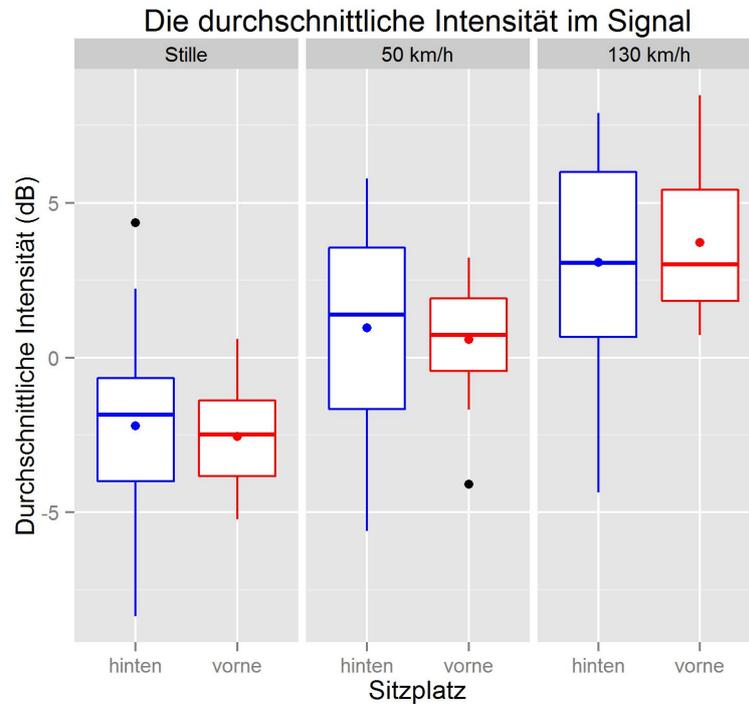


Abbildung 6.5: Die mittelwertbereinigte durchschnittliche Intensität in jeder Geräusch- und Sitzplatzbedingung.

Das Regressionsmodell aus Tabelle 6.6 bietet somit unter allen beschriebenen Modellen den besten Erklärungswert zur durchschnittlichen Intensität. Auf Basis der durch das Modell ausgegebenen Regressionskoeffizienten und durch Anwendung der oben beschriebenen Formel 6.2 werden zudem Schätzwerte für die Ausprägungen der abhängigen Variable unter den verschiedenen Bedingungen berechnet. Das LRM (Lineares Regressionsmodell) verwendet dazu nur die hier ermittelten relevanten Einflussfaktoren, d. h. die erklärenden Variablen. Die Darstellung ist demnach besser auf die vorliegende Analyse abgestimmt als eine einfache Untersuchung der Mittelwerte. Um eine bessere Vergleichbarkeit der Schätzwerte zu gewährleisten, werden diese durch eine lineare Transformation an der Experimentbedingung *ohne ICC, Stille, hinten* ausgerichtet. D. h., dieser Wert wird von allen aufgeführten Bedingungen subtrahiert, sodass *ohne ICC, Stille, hinten* als Vergleichsbasis den Wert 0

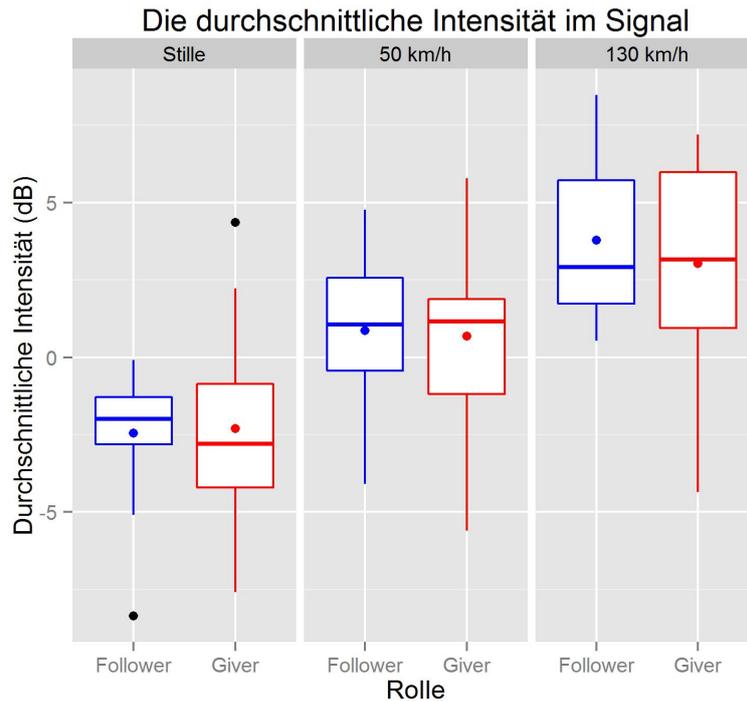


Abbildung 6.6: Die mittelwertbereinigte durchschnittliche Intensität in jeder Geräusch- und Rollenbedingung.

annimmt. Die auf diese Weise angepassten Schätzwerte sind in Tabelle 6.7 aufgelistet und zur Veranschaulichung in Abbildung 6.7 dargestellt. Dadurch wird die absolute Veränderung der durchschnittlichen Intensität in allen Experimentbedingungen im Hinblick auf die Basis *ohne ICC, Stille, hinten* beschrieben. (vgl. [39, S. 1486])

Es wird deutlich, dass die durchschnittliche Intensität auf beiden Sitzplätzen *ohne ICC* und *mit ICC* mit zunehmender Geräuschbelastung ansteigt. Dieser Effekt wird jedoch durch die Aktivierung des ICC-Systems abgeschwächt, da die Differenzen der Messwerte zwischen den einzelnen Geräuschbedingungen abnehmen. Weiterhin sind die Werte des vorderen Sitzplatzes ohne ICC-System niedriger als auf dem hinteren Sitzplatz. Auf der anderen Seite wird *mit ICC vorne* eine höhere durchschnittliche Intensität produziert als *hinten*. Daraus lässt sich schließen, dass sich die Proband*innen entsprechend ihrer Sitzposition unterschiedlich an das ICC-System anpassen. Zudem fallen die Werte in der stillen Bedingung *mit ICC hinten* ab, während sie *vorne* ansteigen. Die Differenz zwischen den Werten *ohne ICC* und *mit ICC* steigt bei zunehmendem Geräuschniveau an. Dies gilt jedoch nicht bei einer Steigerung der Geräuschbedingung von *Stille* zu *50 km/h vorne*. Es kann jedoch

6 Die Analyse der Sprachaufnahmen

festgestellt werden, dass die Unterschiede zwischen *ohne ICC* und *mit ICC* in den einzelnen Geräuschbedingungen *hinten* größer sind als *vorne*. Bis auf die Kombination *Stille, vorne* sind zudem die Werte mit ICC-System in allen Bedingungen niedriger als ohne. (vgl. [38, S. 481f.] und [39, S. 1486])

Tabelle 6.7: Die auf Basis des Regressionsmodells aus Tabelle 6.6 errechneten transformierten Schätzwerte zur durchschnittlichen Intensität (dB).

hinten	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
	0	3,6169	6,7925	-4,0421	-1,3708	0,7098
vorne	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
	-2,3665	1,2504	4,426	-1,7369	0,9344	3,015

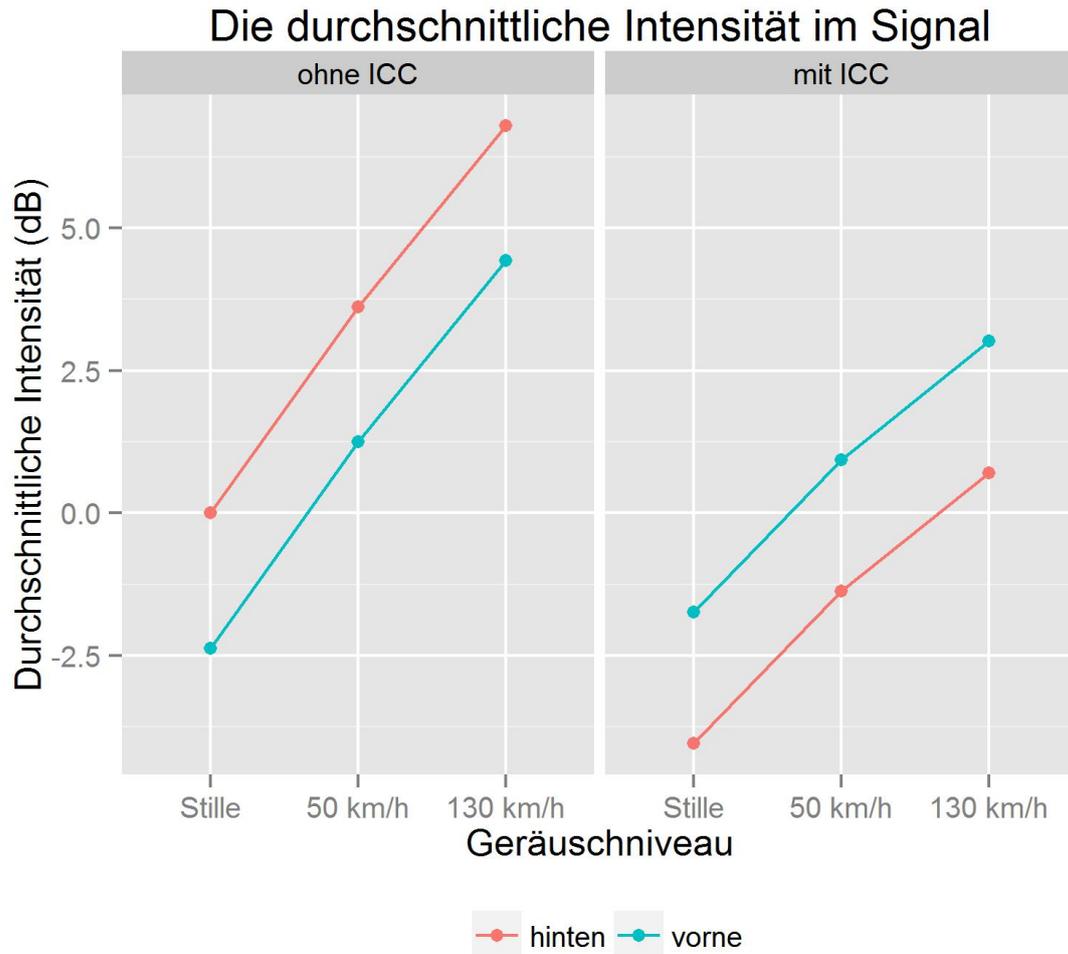


Abbildung 6.7: Die auf Basis des Regressionsmodells aus Tabelle 6.6 errechneten transformierten Schätzwerte zur durchschnittlichen Intensität. [39, S. 1486]

Aus der statistisch bereinigten Darstellung in Abbildung 6.7 wird deutlich, dass es sich hierbei um eine vom Regressionsmodell stilisierte Illustration der signifikanten Zusammenhänge handelt. Dabei scheinen die Veränderungen der Messwerte aufgrund unterschiedlicher Geräuschbedingungen in den einzelnen ICC-Bedingungen für beide Sitzplätze genau parallel zu verlaufen. In Tabelle 6.8 werden die signifikanten Ergebnisse noch einmal dargestellt:

Tabelle 6.8: Die Effekte der unabhängigen Variablen *ICC*, *Geräusch* und *Sitzplatz* auf die durchschnittliche Intensität.

unabhängige Variable	Effekt
Geräusch	In beiden ICC-Bedingungen findet ein Lombard-Effekt statt.
ICC	Das ICC-System reduziert die Stärke des Lombard-Effekts.
ICC	Es gibt einen Niveauunterschied zwischen den ICC-Bedingungen.
Sitzplatz	Es gibt einen Niveauunterschied zwischen den Sitzplätzen.
ICC	Es gibt keine Veränderung des Lombard-Effekts je nach Sitzplatz.

Eine Kombination dieser Ergebnisse ergibt im vorliegenden Regressionsmodell die in Abbildung 6.7 aufgeführte Darstellung. Hier ist besonders auffällig, dass sich die in Stille gemessenen Werte ohne ICC-System und mit ICC-System hinten stark und vorne leicht unterscheiden. Aufgrund der in Kapitel 1.2 beschriebenen Annahmen wurde jedoch ein ähnliches Niveau in der stillen Umgebung beider ICC-Bedingungen erwartet. Wie in den Abbildungen 6.2 und 6.7 zu erkennen ist, besteht diese Diskrepanz zwischen den einzelnen ICC-Bedingungen in Stille auch nach der statistischen Bereinigung fort.

Der Unterschied zwischen den Darstellungen der Mittelwerte und der Schätzwerte liegt darin, dass die Schätzwerte des besten Regressionsmodells in Tabelle 6.6 (vgl. Abbildung 6.7) lediglich die Effekte der Variablenkombinationen $ICC * Geräusch + ICC * Sitzplatz$ darstellen, während die Mittelwerte in Abbildung 6.2 zudem die Interaktion von $Geräusch * Sitzplatz$ miteinbeziehen. Es lässt sich feststellen, dass bei Aktivierung des ICC-Systems auch in der stillen Referenzbedingung eine deutliche Veränderung der durchschnittlichen Intensität auftritt. Dabei findet auf dem hinteren Sitzplatz eine stärkere Modifikation der Intensität statt, als auf dem vorderen Sitzplatz. Diese Differenzen der Messwerte in stiller Umgebung werden zusätzlich zur oben beschriebenen Regressionsanalyse auf Signifikanz geprüft. Dazu werden innerhalb der Programmiersprache R und unter Anwendung der Modellformel 6.6 t-Tests durchgeführt, welche sich auf jeweils zwei abhängige Stichproben beziehen: (vgl. [77, S. 215ff.])

```
t.test(x =< Vektor >, y =< Vektor >, paired = TRUE, conf.level = 0.95,
       alternative = "two.sided", var.equal = FALSE) (6.6)
```

Aus Tabelle 6.9 wird ersichtlich, dass es hinsichtlich der durchschnittlichen Intensität einen signifikanten Unterschied zwischen der stillen Bedingung ohne und mit ICC-System auf dem hinteren Sitzplatz gibt ($p=0,0351$). Ein vergleichsweise nicht sehr hoher t-Wert von 2,6072 liegt darin begründet, dass es sich hierbei um abhängige Stichproben handelt und die Wahrscheinlichkeit einer signifikanten Differenz damit geringer wird. Tabelle 6.10 zeigt hingegen auf, dass die Differenz in *Stille* zwischen *ohne ICC* und *mit ICC* auf dem vorderen Sitzplatz nicht signifikant ist ($p=0,5441$).

Tabelle 6.9: Ergebnisse des t-Tests zur durchschnittlichen Intensität für die abhängigen Stichproben der Experimentbedingungen *ohne ICC*, *Stille*, *hinten* und *mit ICC*, *Stille*, *hinten*.

Paired t-test

data: ohneICC_stille_hinten\$intensity and mitICC_stille_hinten\$intensity

t = 2.6072, df = 7, p-value = 0.03506

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

0.3758705 7.7041295

sample estimates:

mean of the differences

4.04

Tabelle 6.10: Ergebnisse des t-Tests zur durchschnittlichen Intensität für die abhängigen Stichproben der Experimentbedingungen *ohne ICC, Stille, vorne* und *mit ICC, Stille, vorne*.

```
Paired t-test

data: ohneICC_stille_vorne$intensity and mitICC_stille_vorne$intensity
t = -0.63754, df = 7, p-value = 0.5441
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-2.9549 1.6999
sample estimates:
mean of the differences
-0.6275
```

6.3.2 Grundfrequenz

Die Untersuchung der durchschnittlichen Grundfrequenz findet ebenfalls mit Hilfe der *Software* Praat statt. Auch hierfür wird ein Praat-Skript verwendet, welches für alle *Sounddateien* eines Zielordners unterschiedliche Werte zur *pitch* ausgibt, wie die durchschnittliche Grundfrequenz des jeweiligen Signals. Dabei werden nach einer Empfehlung in [52, S. 94] verschiedene Tonhöhenbereiche für Männer (50 Hz - 300 Hz) und Frauen (100 Hz - 500 Hz) festgelegt. Anschließend findet wiederum eine Mittelwertbereinigung der Messwerte durch Formel 6.1 statt. (vgl. [38, S. 481f.]

Ähnlich wie bei der durchschnittlichen Intensität lässt sich in Abbildung 6.8 beobachten, dass die durchschnittliche Grundfrequenz mit zunehmender Geräuschbelastung ansteigt, sowohl ohne als auch mit ICC-System. Zudem scheinen die Messwerte mit ICC-System eine geringere Varianz aufzuweisen und liegen im Mittel um etwa 2,498 Hz (Stille), 7,77 Hz (50 km/h) und 13,708 Hz (130 km/h) niedriger. Obwohl die vorliegenden Sprachsignale Männer- und Frauenstimmen beinhalten, kann die Grundfrequenz in Hz analysiert und muss nicht in Halbtönen (*semitones*) bzw. Halbtonschritten untersucht werden, da eine Mittelwertbereinigung der Messwerte durchgeführt wurde. (vgl. [38, S. 481f.]) Bei den Mittelwerten (vgl. Abbildung 6.9) wird eine Differenz zwischen dem vorderen und hinteren Sitzplatz und eine Veränderung der Messwerte mit ICC-System deutlich, die bereits in Stille auftritt.

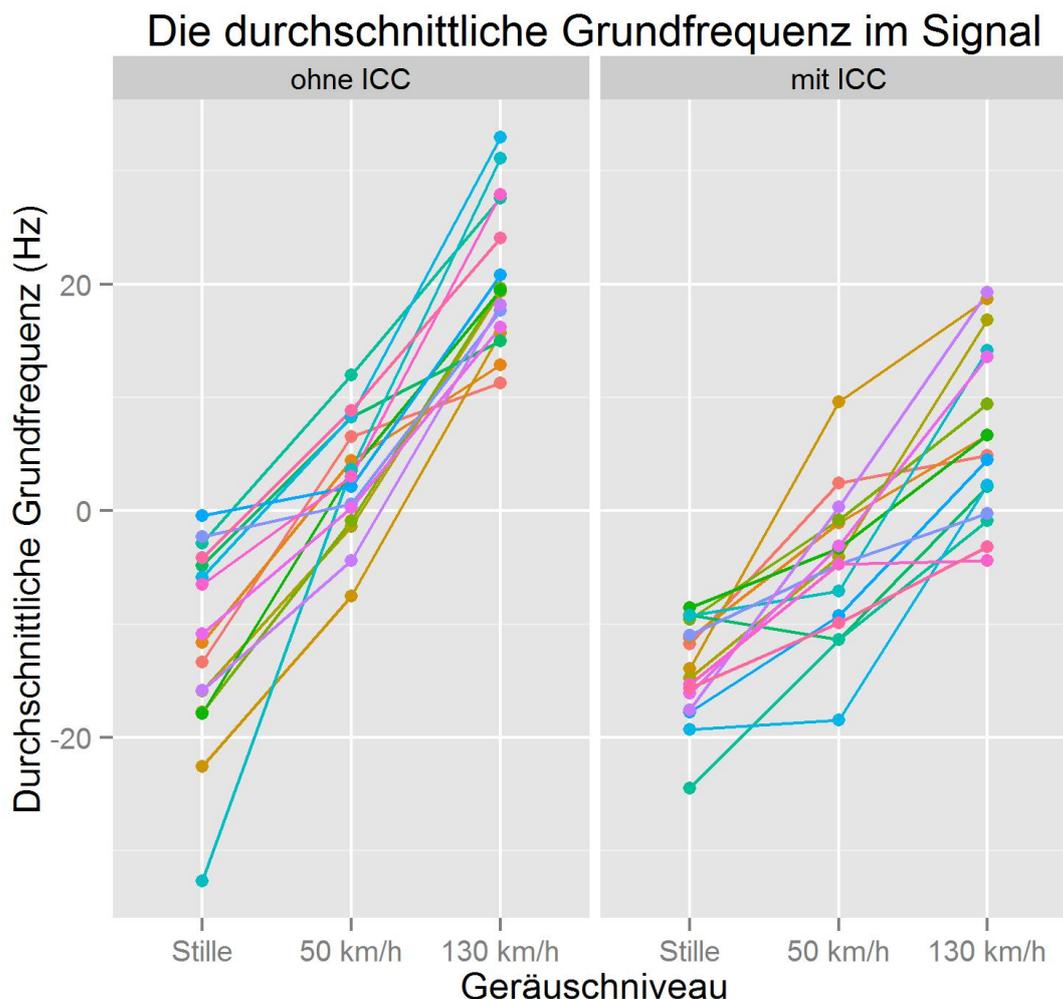


Abbildung 6.8: Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren Daten einzelner Proband*innen. [38, S. 481]

Zur statistischen Überprüfung der Fragestellungen aus Kapitel 1.2 wird wiederum zunächst von einem Nullmodell ohne erklärende Variablen ausgegangen. Aufgrund der Übereinstimmungen mit den deskriptiven Ergebnissen der durchschnittlichen Intensität erfolgt das Hinzufügen von unabhängigen Variablen in die Modellformel der Regression zunächst nach einem ähnlichen Muster wie in Kapitel 6.3.1. Die *Responsevariable* ist *pitch*, die erste hinzugefügte erklärende Variable *ICC*. Tabelle 6.11 stellt die R-Ausgabe für dieses Modell dar. Das *Adjusted R-squared* ist mit einem Wert von 0,06891 noch sehr niedrig und sogar geringer als für die durchschnittliche Intensität in Kapitel 6.3.1. Allerdings kann festgestellt werden, dass sich die Ausprägungen der unabhängigen Variablen *ICC* mit $p < 0,05$ signifikant von 0 unter-

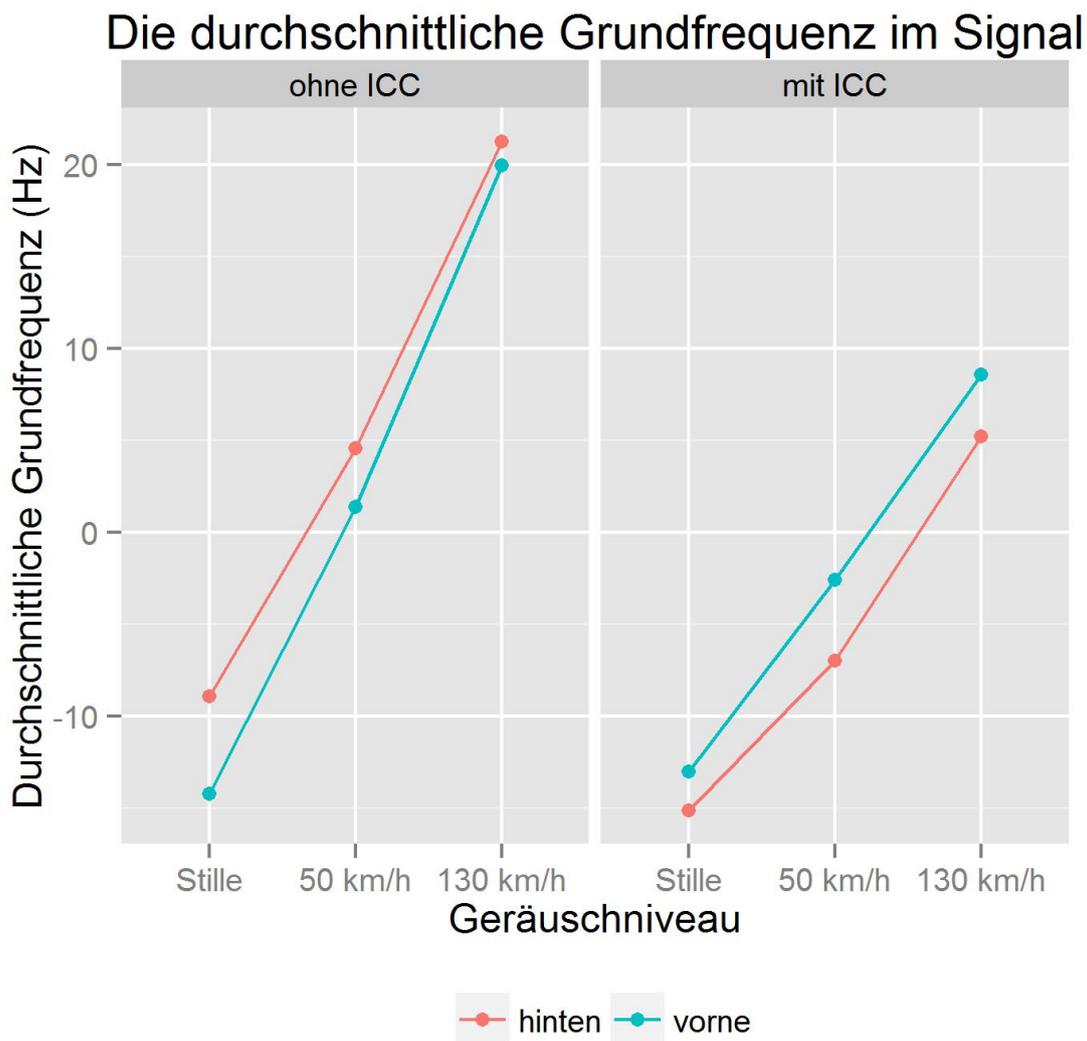


Abbildung 6.9: Die Mittelwerte zur mittelwertbereinigten durchschnittlichen Grundfrequenz über alle Sprecher*innen hinweg.

scheiden. Wie in der Analyse der durchschnittlichen Intensität sind auch die Werte der durchschnittlichen Grundfrequenz geringer für Aufnahmen mit ICC-System.

Tabelle 6.11: Regressionsergebnisse für die abhängige Variable *pitch* und die erklärende Variable *ICC*.

```
Call:
lm(formula = pitch ~ ICC - 1, data = dataicc)
```

```
Residuals:
```

Min	1Q	Median	3Q	Max
-36.634	-8.212	-0.588	8.851	28.969

```
Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	-3.996	1.873	-2.134	0.0355	*
ICCoohne ICC	3.996	1.873	2.134	0.0355	*

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
Residual standard error: 12.98 on 94 degrees of freedom
Multiple R-squared: 0.08831, Adjusted R-squared: 0.06891
F-statistic: 4.552 on 2 and 94 DF, p-value: 0.01297
```

Des Weiteren wird der Einfluss der erklärenden Variable *Geräusch* auf die durchschnittliche Grundfrequenz in den Sprachaufnahmen untersucht, indem *ICC* in der Modellformel ersetzt wird. Die Ergebnisse dieses Regressionsmodells sind in Tabelle 6.12 aufgeführt. Hierbei ist zu erkennen, dass der Wert des *Adjusted R-squared* mit 0,6427 deutlich höher liegt als für die erklärende Variable *ICC*. Zudem sind die Signifikanzniveaus der verschiedenen Ausprägungen von *Geräusch* erheblich höher. Die Bedingung *Stille* unterscheidet sich mit $p < 0,0001$ signifikant von 0. Dabei zeigt der dazugehörige Regressionskoeffizient (-12,8322) an, dass die durchschnittliche Grundfrequenz im Vergleich zu allen anderen Bedingungen niedriger liegt. Die Ausprägung *50 km/h* unterscheidet sich mit einem Schätzwert von -0,9216 und $p = 0,518$ nicht signifikant von 0. Die Bedingung *130 km/h* erhält einen p-Wert von weniger als 0,0001 und weist mit einem Regressionskoeffizienten von 13,7538 die höchsten Werte der durchschnittlichen Grundfrequenz auf. Aufgrund dieser Ergebnisse kann festgehalten werden, dass sich die Ausprägungen der erklärenden Variable *Geräusch* signifikant voneinander unterscheiden.

Tabelle 6.12: Regressionsergebnisse für die abhängige Variable *pitch* und die erklärende Variable *Geräusch*.

```
Call:
lm(formula = pitch ~ Geräusch - 1, data = dataicc)
```

Residuals:

Min	1Q	Median	3Q	Max
-19.8061	-4.7616	0.7517	5.0567	19.2112

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
Geräusch130 km/h	13.7538	1.4209	9.679	9.75e-16	***
Geräusch50 km/h	-0.9216	1.4209	-0.649	0.518	
GeräuschStille	-12.8322	1.4209	-9.031	2.30e-14	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8.038 on 93 degrees of freedom
Multiple R-squared: 0.6538, Adjusted R-squared: 0.6427
F-statistic: 58.56 on 3 and 93 DF, p-value: < 2.2e-16

Nachdem bewiesen ist, dass die Einflussgrößen *ICC* und *Geräusch* signifikante Effekte auf die durchschnittliche Grundfrequenz der Sprachsignale haben, können zusätzliche Informationen durch ein Regressionsmodell gewonnen werden, welches eine Interaktion zwischen den Ausprägungen beider unabhängiger Variablen berücksichtigt. Die Ergebnisse dieses Modells sind in Tabelle 6.13 dargestellt. Dieses Modell liefert wiederum einen höheren Erklärungswert als die vorangegangenen Regressionsmodelle (*Adjusted R-squared*=0,7559). Durch die Signifikanzniveaus der einzelnen Ausprägungen werden die Ergebnisse der zuvor berechneten Modelle bestätigt. Zudem sind in Zeile 5 und 6 des Abschnitts *Coefficients* Interaktionen zwischen den Ausprägungen beider erklärender Variablen aufgeführt. Dabei wird deutlich, dass es signifikante Wechselwirkungen zwischen den unabhängigen Variablen gibt ($p < 0,005$).

Tabelle 6.13: Regressionsergebnisse für die abhängige Variable *pitch* und die Interaktion der erklärenden Variablen *ICC* und *Geräusch*.

Call:

lm(formula = pitch ~ ICC * Geräusch - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-21.0549	-4.3797	-0.1365	4.9386	14.4116

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	6.900	1.661	4.154	7.42e-05	***
ICCoohne ICC	20.608	1.661	12.407	< 2e-16	***
Geräusch50 km/h	-11.706	2.349	-4.984	3.00e-06	***
GeräuschStille	-20.981	2.349	-8.932	4.77e-14	***
ICCoohne ICC:Geräusch50 km/h	-5.938	3.322	-1.788	0.07721	.
ICCoohne ICC:GeräuschStille	-11.211	3.322	-3.375	0.00109	**

—
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.644 on 90 degrees of freedom

Multiple R-squared: 0.7711, Adjusted R-squared: 0.7559

F-statistic: 50.54 on 6 and 90 DF, p-value: < 2.2e-16

Des Weiteren werden die Effekte der unabhängigen Variablen *Sitzplatz* und *Rolle* auf die *Responsevariable pitch* untersucht. Die Abbildungen 6.10 und 6.11 bestätigen die oben beschriebene Beobachtung, dass die durchschnittliche Grundfrequenz mit ICC-System im Mittel niedriger ist. Dies gilt für beide Sitzplätze und beide Sprechrollen gleichermaßen. Darüber hinaus weisen die Messwerte mit ICC-System eine geringere Varianz auf. In Abbildung 6.10 lässt sich zudem eine geringe Interaktion zwischen den unabhängigen Variablen *ICC* und *Sitzplatz* erkennen. Mit Blick auf die jeweiligen Mittelwerte ist die durchschnittliche Grundfrequenz ohne ICC-System auf dem hinteren Sitzplatz mit 5,637 Hz höher als vorne mit 2,355 Hz. Mit ICC-System kehrt sich das Verhältnis um (-5,637 Hz hinten und -2,355 Hz vorne). Die Werte ohne und mit ICC-System ergeben jeweils für den hinteren und vorderen Sitzplatz im Mittel 0, da die Messwerte mittelwertbereinigt sind und die

6 Die Analyse der Sprachaufnahmen

Proband*innen in allen Sprachaufnahmen ihre Sitzplätze beibehalten haben. Diese Beobachtung stimmt mit Ergebnissen zur durchschnittlichen Intensität (vgl. Kapitel 6.3.1) überein, wobei die Aktivierung des ICC-Systems einen stärkeren Einfluss auf hinten produzierte Sprachsignale hat (vgl. Abbildung 6.3). Die unabhängige Variable *Rolle* scheint hingegen keinen systematischen Einfluss auf die Messwerte zu besitzen.

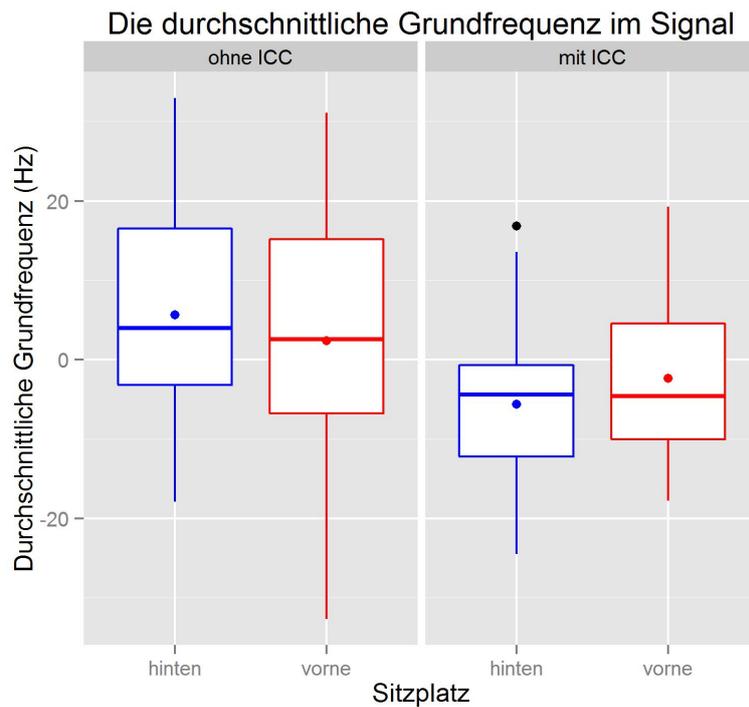


Abbildung 6.10: Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder ICC- und Sitzplatzbedingung.

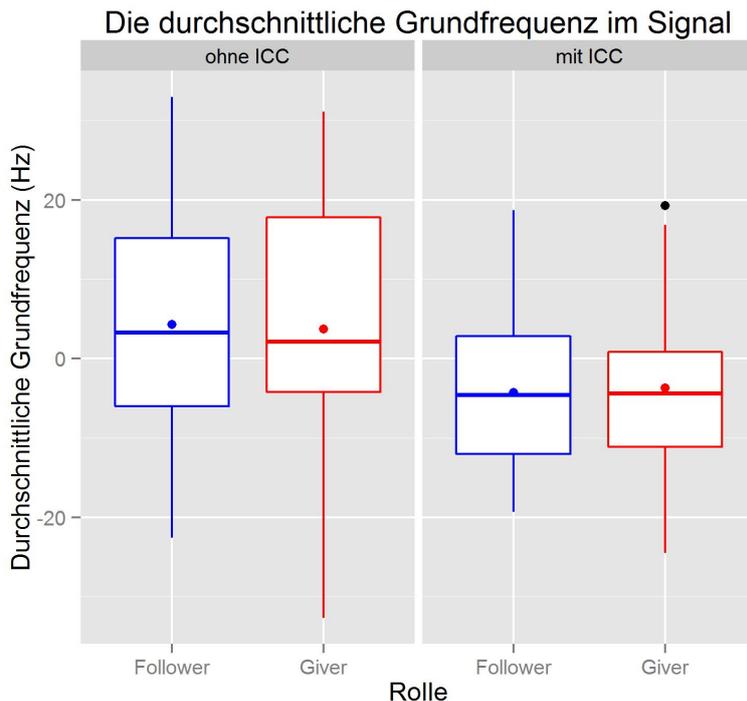


Abbildung 6.11: Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder ICC- und Rollenbedingung.

Ausgehend von den oben beschriebenen deskriptiven Beobachtungen wird das Regressionsmodell in Tabelle 6.13 zunächst um die unabhängige Variable *Sitzplatz* erweitert. Dabei wird deutlich, dass das *Adjusted R-squared* eine geringe Reduktion erfährt, da dieser Wert statt 0,7559 nur noch 0,7531 beträgt (vgl. Tabelle 6.14). Zudem wird deutlich, dass sich die Ausprägung *vorne* der hinzugefügten unabhängigen Variable nicht signifikant von 0 unterscheidet ($p=1$). Aus diesen Gründen wird die Variable *Sitzplatz* wieder aus der Modellformel entfernt.

Tabelle 6.14: Regressionsergebnisse für die abhängige Variable *pitch*, die Interaktion der erklärenden Variablen *ICC* und *Geräusch* und die erklärende Variable *Sitzplatz*.

Call:

lm(formula = pitch ~ ICC * Geräusch + Sitzplatz - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-21.0549	-4.3797	-0.1365	4.9386	14.4116

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	6.900e+00	1.804e+00	3.825	0.000243	***
ICCoohne ICC	2.061e+01	1.804e+00	11.423	< 2e-16	***
Geräusch50 km/h	-1.171e+01	2.362e+00	-4.956	3.41e-06	***
GeräuschStille	-2.098e+01	2.362e+00	-8.882	6.58e-14	***
Sitzplatzvorne	-1.146e-09	1.364e+00	0.000	1.000000	
ICCoohne ICC:Geräusch50 km/h	-5.938e+00	3.340e+00	-1.778	0.078881	.
ICCoohne ICC:GeräuschStille	-1.121e+01	3.340e+00	-3.356	0.001164	**

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.681 on 89 degrees of freedom

Multiple R-squared: 0.7711, Adjusted R-squared: 0.7531

F-statistic: 42.84 on 7 and 89 DF, p-value: < 2.2e-16

Um die in Abbildung 6.10 erkennbare Wechselwirkung zwischen *ICC* und *Sitzplatz* näher zu untersuchen, wird diese Interaktion in die Modellformel aus Tabelle 6.13 aufgenommen. Tabelle 6.15 zeigt auf, dass sich der Wert des *Adjusted R-squared* auf 0,7666 erhöht. Dabei wird deutlich, dass sowohl die Ausprägung *vorne* ($p=0,0836$) als auch die Interaktion *ohne ICC, vorne* ($p=0,0152$) eine schwach signifikante bzw. signifikante Unterscheidung von 0 aufweisen. Die hier angegebenen Signifikanzniveaus sind demnach niedriger als für den Untersuchungsparameter der durchschnittlichen Intensität (vgl. Kapitel 6.3.1), zu dem ähnliche Ergebnisse erzielt wurden. (vgl. [39, S. 1486])

Tabelle 6.15: Regressionsergebnisse für die abhängige Variable *pitch*, die Interaktion der erklärenden Variablen *ICC* und *Geräusch* sowie die Interaktion der erklärenden Variablen *ICC* und *Sitzplatz*.

Call:
lm(formula = pitch ~ ICC * Geräusch + ICC * Sitzplatz - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-19.4140	-4.3922	-0.0202	4.3001	12.7794

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	5.259	1.875	2.804	0.006208	**
ICCoohne ICC	22.249	1.875	11.864	< 2e-16	***
Geräusch50 km/h	-11.706	2.297	-5.097	1.96e-06	***
GeräuschStille	-20.981	2.297	-9.134	2.17e-14	***
Sitzplatzvorne	3.282	1.875	1.750	0.083609	.
ICCoohne ICC:Geräusch50 km/h	-5.938	3.248	-1.828	0.070924	.
ICCoohne ICC:GeräuschStille	-11.211	3.248	-3.451	0.000859	***
ICCoohne ICC:Sitzplatzvorne	-6.564	2.652	-2.475	0.015247	*

—

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.497 on 88 degrees of freedom
Multiple R-squared: 0.786, Adjusted R-squared: 0.7666
F-statistic: 40.41 on 8 and 88 DF, p-value: < 2.2e-16

Darüber hinaus wird untersucht, welche Effekte *Sitzplatz* und *Rolle* in Verbindung mit ansteigenden Fahrgeräuschen haben. Aus Abbildung 6.12 und 6.13 geht hervor, dass die durchschnittliche Grundfrequenz mit zunehmender Geräuschbelastung ansteigt. In Stille besitzen hinten sitzende Sprecher*innen mit 1,617 Hz und *Follower* mit 1,678 Hz im Mittel ein höheres F0 als ihre Gesprächspartner*innen. Mit dem Auftreten der Fahrgeräusche kehrt sich dieses Verhältnis um. Bei 50 km/h sprechen vorne sitzende Personen um 0,587 Hz und *Giver* um 0,546 Hz höher als ihre Dialogpartner*innen, während sich diese Differenz bei 130 km/h auf 1,031 Hz für den Sitzplatz und 1,133 Hz für die Sprechrolle erhöht. Durch das Hinzufügen dieser Interaktionen zu der Regression in Tabelle 6.15 wird kein höherer Erklärungswert des Modells erreicht. Für die zusätzliche Wechselwirkung von *Geräusch* und *Sitzplatz*

6 Die Analyse der Sprachaufnahmen

beträgt das *Adjusted R-squared* 0,7632, während dieser Wert bei einer zusätzlichen Interaktion von *Geräusch* und *Rolle* auf 0,7606 abfällt. Zudem werden keine signifikanten p-Werte für die jeweils hinzugefügte Interaktion erreicht.

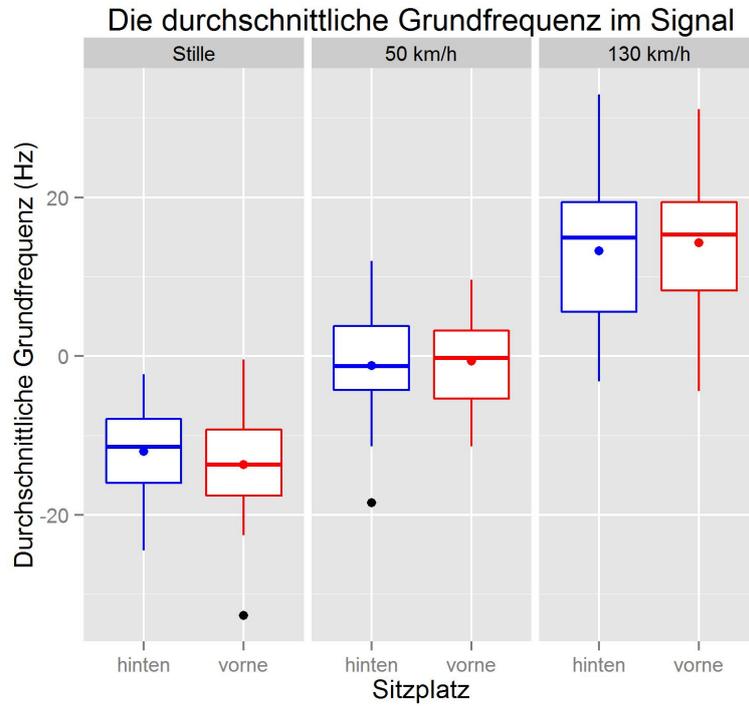


Abbildung 6.12: Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder Geräusch- und Sitzplatzbedingung.

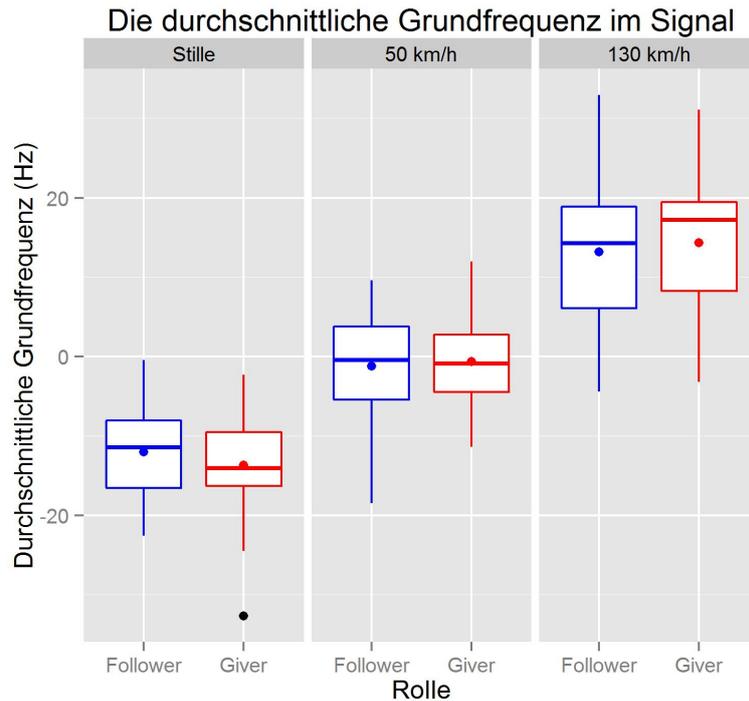


Abbildung 6.13: Die mittelwertbereinigte durchschnittliche Grundfrequenz in jeder Geräusch- und Rollenbedingung.

Die in Tabelle 6.15 dargestellte Regression wird somit als Modell mit dem besten Erklärungswert zur durchschnittlichen Grundfrequenz beibehalten. Wie bereits in Kapitel 6.3.1 für die durchschnittliche Intensität, werden anhand der Regressionskoeffizienten des Modells 6.15 alle Schätzwerte der Variablenausprägungen und ihrer Interaktionen für die durchschnittliche Grundfrequenz anhand der Formel 6.2 berechnet. Zur besseren Vergleichbarkeit dieser Werte fungiert wiederum die Kombination *ohne ICC*, *Stille*, *hinten* als Basis, deren Schätzwert von allen Experimentbedingungen subtrahiert worden ist und somit 0 beträgt. Die dadurch erhaltenen Schätzwerte sind in Tabelle 6.16 beschrieben und zeigen die absolute Veränderung der gesamten durchschnittlichen Grundfrequenz hinsichtlich der Basis an. Abbildung 6.14 stellt zusätzlich eine graphische Übersicht dieser Schätzwerte dar. Aus Tabelle 6.16 lässt sich schließen, dass *vorne* und *hinten* bei steigender Geräuschbelastung von *Stille* zu *50 km/h* und von *50 km/h* zu *130 km/h* eine Erhöhung der durchschnittlichen Grundfrequenz in den Sprachsignalen stattfindet, sowohl *ohne ICC* als auch *mit ICC*. Auch hier führt die Verwendung des ICC-Systems *hinten* und *vorne* zu einer Verminderung dieses Anstiegs. Zudem ist der festgestellte Lombard-Effekt

6 Die Analyse der Sprachaufnahmen

von 50 km/h zu 130 km/h stärker als von *Stille* zu 50 km/h. Dabei kann kaum ein Unterschied zwischen den beiden Sitzplätzen festgestellt werden. D. h., die Abstände der Werte zwischen den einzelnen Geräuschbedingungen sowohl *ohne ICC* als auch *mit ICC* sind *hinten* und *vorne* gleich. Ebenso findet auch hier je nach Sitzposition eine andersartige Anpassung an das ICC-System statt, denn wie für den Parameter der durchschnittlichen Intensität sind die Werte der durchschnittlichen Grundfrequenz auf dem vorderen Sitzplatz *ohne ICC* niedriger und *mit ICC* höher als *hinten*. Ebenso steigen die Werte der stillen Referenzbedingung auf dem vorderen Sitzplatz *mit ICC*, während sie *hinten* mit Verwendung des ICC-Systems abfallen. Darüber hinaus vergrößern sich die Abstände zwischen den Werten *ohne ICC* und *mit ICC* mit steigender Geräuschbelastung. Dieser Effekt ist etwas stärker für den hinteren Sitzplatz in der Geräuschzunahme von *Stille* zu 50 km/h. Die absoluten Unterschiede zwischen *ohne ICC* und *mit ICC* sind *hinten* größer als *vorne*. Des Weiteren sind die Werte *mit ICC* in fast allen Bedingungen geringer als *ohne ICC*. Davon ausgenommen ist lediglich die Bedingung *Stille, vorne*. (vgl. [38, S. 481f.] und [39, S. 1486])

Tabelle 6.16: Die auf Basis des Regressionsmodells aus Tabelle 6.15 errechneten transformierten Schätzwerte zur durchschnittlichen Grundfrequenz (Hz).

hinten	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
	0	14,548	32,192	-5,779	3,496	15,202
vorne	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
	-3,282	11,266	28,910	-2,497	6,778	18,484

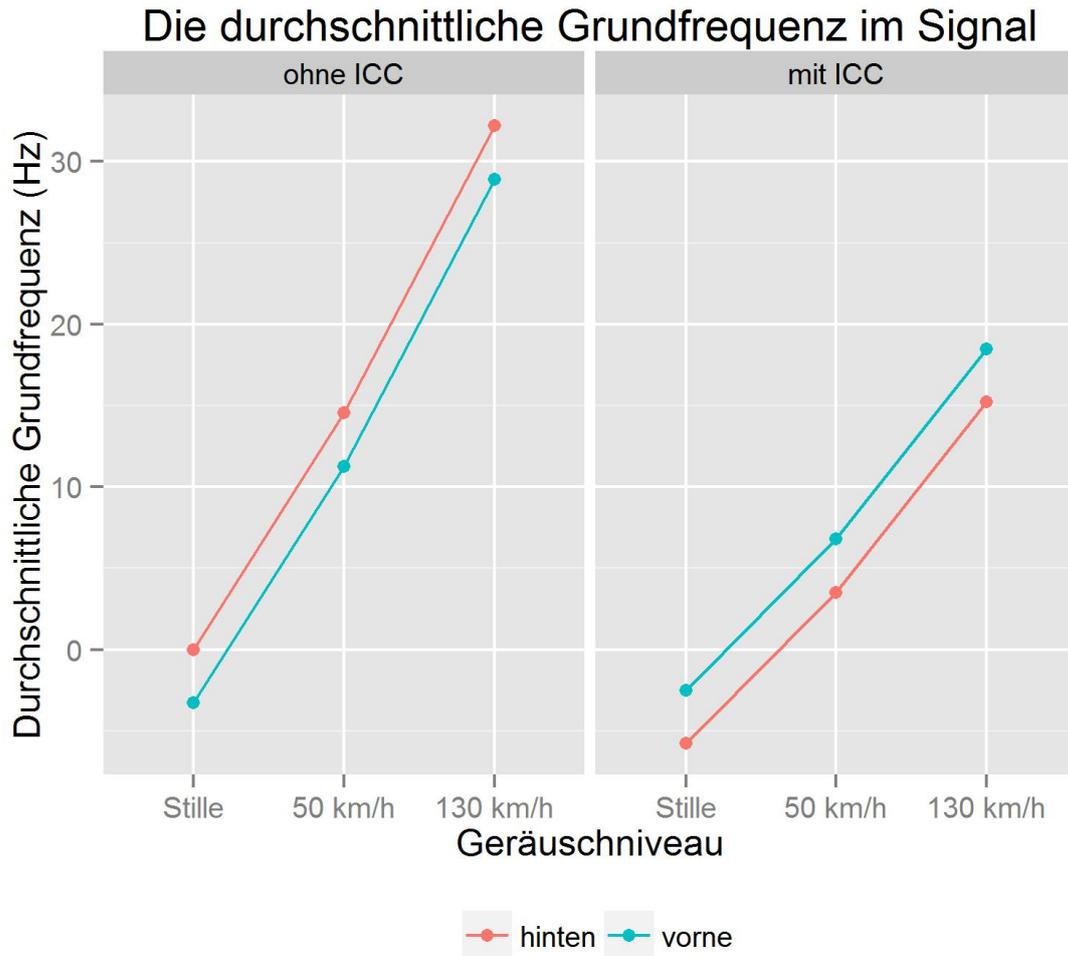


Abbildung 6.14: Die auf Basis des Regressionsmodells aus Tabelle 6.15 errechneten transformierten Schätzwerte zur durchschnittlichen Grundfrequenz.

Die signifikanten Zusammenhänge zur durchschnittlichen Grundfrequenz werden in Abbildung 6.14 vom Regressionsmodell stilisiert dargestellt. Dabei wird ebenfalls eine Kombination der in Tabelle 6.8 aufgeführten Ergebnisse aufgezeigt. Die Darstellung der Regressionsergebnisse bestätigt die in den Mittelwerten (vgl. Abbildung 6.9) in Stille beobachteten Unterschiede zwischen den Sitzplätzen und den ICC-Bedingungen. Wie in Kapitel 6.3.1 werden auch in diesen Mittelwerten nicht nur die Variablenkombinationen des besten Regressionsmodells 6.15 ($ICC * Geräusch + ICC * Sitzplatz$) aufgezeigt, sondern zudem die Interaktion von $Geräusch * Sitzplatz$.

Wie in Kapitel 6.3.1 zur durchschnittlichen Intensität ist auch hier die Differenz zwischen den Werten der stillen Referenzbedingung ohne und mit ICC-System beson-

ders auffällig (vgl. Abbildung 6.14). Dabei findet eine starke Abnahme der Messwerte *hinten* und ein leichter Anstieg der Messwerte *vorne* statt. Die durch Aktivierung des ICC-Systems stattfindenden Veränderungen der durchschnittlichen Grundfrequenz sind auf dem hinteren Sitzplatz stärker ausgebildet als auf dem vorderen Sitzplatz. Deshalb werden für die durchschnittliche Grundfrequenz ebenso t-Tests in R nach der oben beschriebenen Formel 6.6 durchgeführt. Aus Tabelle 6.17 und 6.18 wird ersichtlich, dass die Differenz zwischen inaktivem und aktivem ICC-System in stiller Referenzbedingung weder auf dem hinteren Sitzplatz ($p=0,1144$) noch auf dem vorderen Sitzplatz ($p=0,7906$) signifikant ist.

Tabelle 6.17: Ergebnisse des t-Tests zur durchschnittlichen Grundfrequenz für die abhängigen Stichproben der Experimentbedingungen *ohne ICC, Stille, hinten* und *mit ICC, Stille, hinten*.

```

Paired t-test

data:  ohneICC_stille_hinten$pitch and mitICC_stille_hinten$pitch
t = 1.8028, df = 7, p-value = 0.1144
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-1.933549 14.343549
sample estimates:
mean of the differences
      6.205
    
```

Tabelle 6.18: Ergebnisse des t-Tests zur durchschnittlichen Grundfrequenz für die abhängigen Stichproben der Experimentbedingungen *ohne ICC, Stille, vorne* und *mit ICC, Stille, vorne*.

```

Paired t-test

data:  ohneICC_stille_vorne$pitch and mitICC_stille_vorne$pitch
t = -0.27583, df = 7, p-value = 0.7906
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-11.583151 9.163151
sample estimates:
mean of the differences
     -1.21
    
```

6.3.3 Signaldauer

Hinsichtlich der Gesamtdauer der im Labor aufgenommenen Dialoge wird untersucht, ob ein Zusammenhang zwischen der für die Lösung einer *Map-Task* Aufgabe benötigten Zeit und den unterschiedlichen ICC- und Geräuschbedingungen des Produktionsexperiments existiert. Dazu werden die Dauern der Sprachsignale in Sekunden wiederum für jede Sprecherin bzw. jeden Sprecher mittelwertbereinigt (vgl. Kapitel 6.3).

Das hierzu erstellte Liniendiagramm in Abbildung 6.15 weist im Gegensatz zu den Diagrammen der vorangegangenen Parameter lediglich acht verschiedenfarbige Linien auf. Der Grund hierfür ist, dass die Dauer einer Aufnahme für jede Sprecherin bzw. jeden Sprecher eines Dialogpaares dieselbe ist und daher paarweise Signaldauern abgebildet werden. Das Liniendiagramm zeigt sowohl für die Geräuschbedingungen ohne ICC-System als auch für jene mit ICC-System keine systematischen Veränderungen in der Dauer der aufgenommenen Dialoge. Es ist jedoch zu erkennen, dass der *range* der Messwerte über die Geräuschbedingungen hinweg mit ICC-System wesentlich geringer ist. Dabei liegen fast alle Messwerte unterhalb des sprecher*innenspezifischen Durchschnitts.

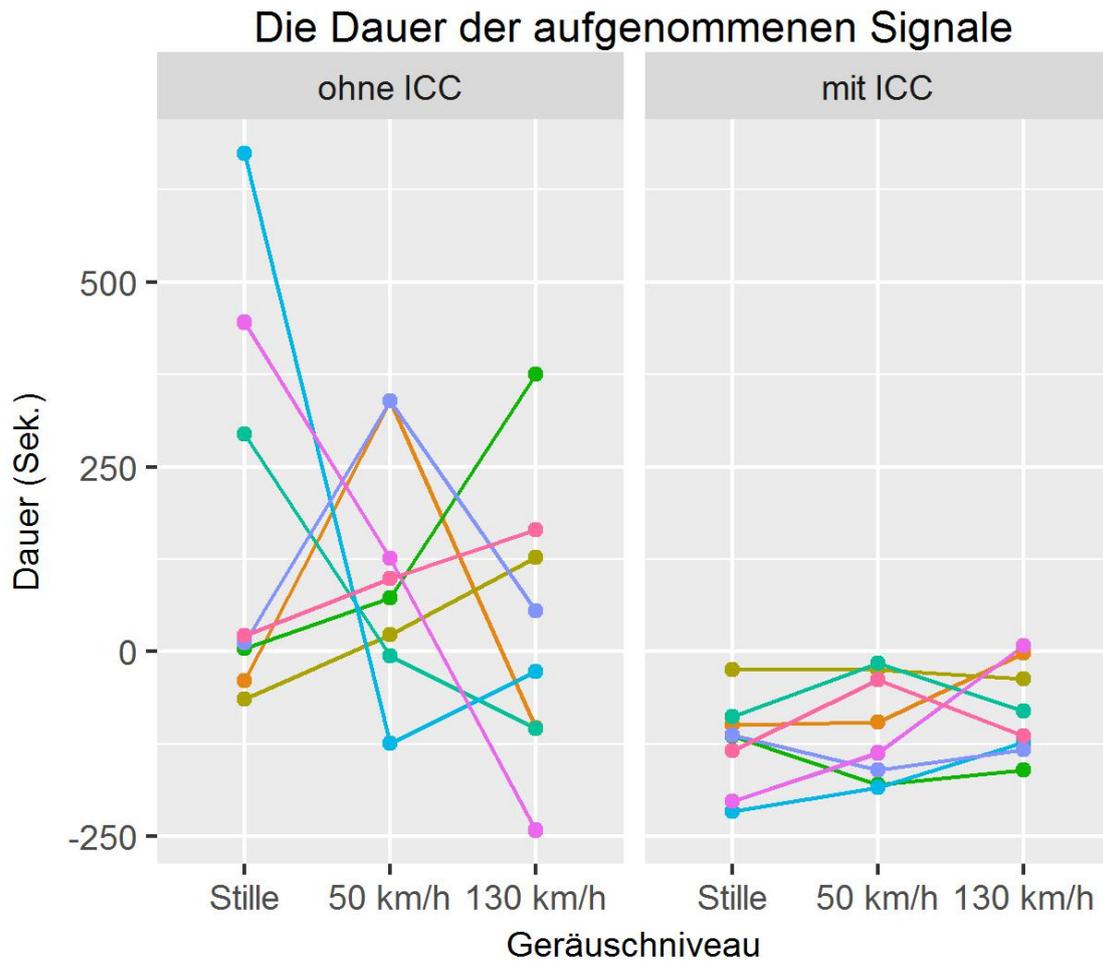


Abbildung 6.15: Die mittelwertbereinigte Signaldauer in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren Daten einzelner Proband*innen.

Die in Abbildung 6.16 dargestellten Mittelwerte zeigen, dass die Signaldauer mit zunehmenden Fahrgeräuschen ohne ICC-System abfällt, während sie bei Anwendung des ICC-Systems leicht zunimmt. Dabei sind die Messwerte mit ICC-System in jeder Geräuschbedingung niedriger als ohne.

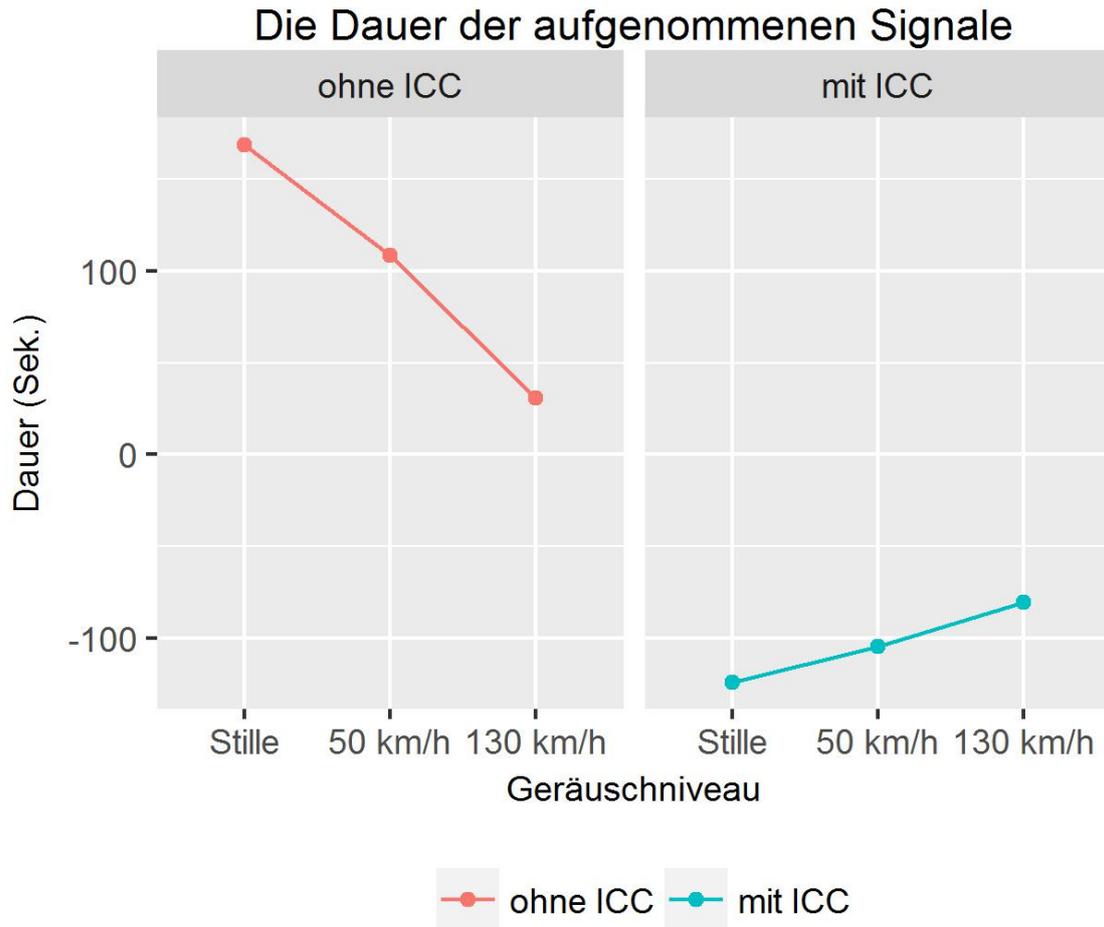


Abbildung 6.16: Die Mittelwerte zur mittelwertbereinigten Signaldauer über alle Sprecher*innen hinweg.

Die oben beschriebene Beobachtung wird weiterhin durch das in Abbildung 6.17 dargestellte *Boxplot*-Diagramm bestätigt. Dabei ist zu erkennen, dass die Signaldauer mit steigenden Umgebungsgeräuschen immer weiter abnimmt, wenn ICC nicht aktiviert ist. Mit ICC-System zeichnet sich indessen ein geringerer, jedoch entgegengesetzter Effekt ab. Die deskriptive Analyse des vorliegenden Parameters weist darauf hin, dass mit ICC-System weniger Zeit zum Lösen einer Sprechaufgabe benötigt wird. Mit steigender Geräuschbelastung scheinen sich die Unterschiede zwischen den ICC-Bedingungen einander anzunähern. Eine weitergehende Untersuchung der Messwerte hinsichtlich der unabhängigen Variablen *Rolle* und *Sitzplatz* ist für den vorliegenden Parameter nicht relevant, da sowohl *Giver* und *Follower* als auch *vorne* und *hinten* sitzende Proband*innen eines Dialogpaares jeweils dieselbe Signaldauer aufweisen (vgl. [39, S. 1487]).

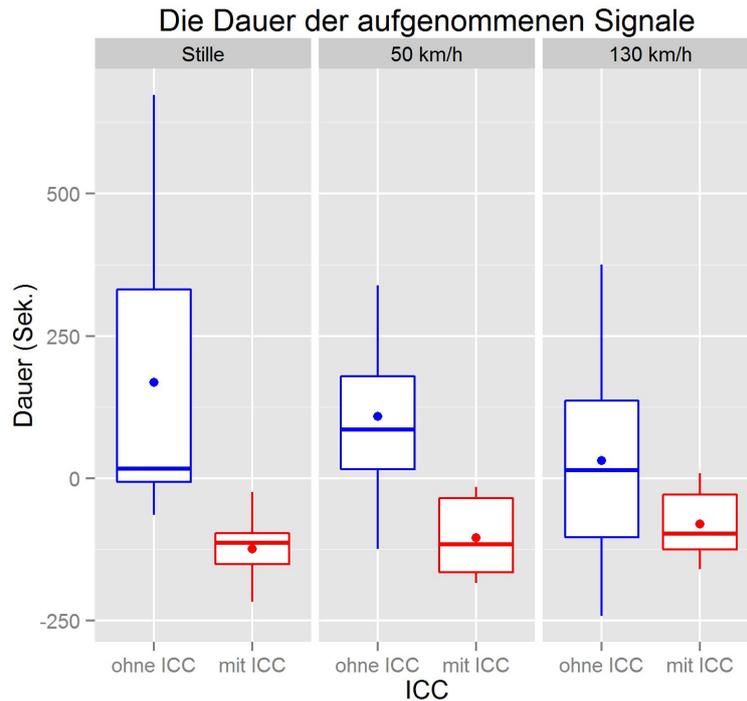


Abbildung 6.17: Die mittelwertbereinigte Signaldauer in jeder ICC- und Geräuschbedingung.

Ausgehend von einem Nullmodell ohne erklärende Variablen wird anschließend eine Regressionsanalyse unter Verwendung der unabhängigen Variable *ICC* durchgeführt (vgl. Tabelle 6.19). Dabei wird die oben beschriebene Beobachtung bestätigt, dass die Aktivierung des ICC-Systems einen signifikanten Effekt ($p < 0,0001$) auf die zur Bewältigung der Sprechaufgabe benötigte Zeit besitzt. Weiterhin zeigt Tabelle 6.20, dass die Steigerung der Umgebungsgeräusche allein keine signifikanten Auswirkungen auf die Dauer der aufgenommenen Sprachsignale besitzt.

Tabelle 6.19: Regressionsergebnisse für die abhängige Variable *dursignal* und die erklärende Variable *ICC*.

```

Call:
lm(formula = dursignal ~ ICC - 1, data = dataicc)

```

```

Residuals:
Min          1Q      Median        3Q         Max
-344.19    -83.23    -15.59     64.51    570.89

```

```

Coefficients:
              Estimate Std. Error  t value Pr(>|t|)
ICCmit ICC   -102.68     22.41   -4.581  1.42e-05 ***
ICCoohne ICC  102.68     22.41    4.581  1.42e-05 ***
—
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```

Residual standard error: 155.3 on 94 degrees of freedom
Multiple R-squared: 0.3087, Adjusted R-squared: 0.294
F-statistic: 20.99 on 2 and 94 DF, p-value: 2.915e-08

```

Tabelle 6.20: Regressionsergebnisse für die abhängige Variable *dursignal* und die erklärende Variable *Geräusch*.

Call:
lm(formula = dursignal ~ Geraeusch - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-238.93	-122.68	-43.45	42.55	651.10

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
Geraeusch130 km/h	-24.560	33.011	-0.744	0.459
Geraeusch50 km/h	2.098	33.011	0.064	0.949
GeraeuschStille	22.462	33.011	0.680	0.498

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 186.7 on 93 degrees of freedom
Multiple R-squared: 0.01085, Adjusted R-squared: -0.02105
F-statistic: 0.3402 on 3 and 93 DF, p-value: 0.7963

Das in Tabelle 6.20 dargestellte Resultat entspricht den durch die oben beschriebene deskriptive Untersuchung gewonnenen Ergebnissen (vgl. Abbildung 6.17). Die Effekte der steigenden Umgebungsgeräusche auf die Dauer der Sprachaufnahmen verlaufen für beide ICC-Bedingungen in unterschiedliche Richtungen. Ohne die Unterscheidung zwischen den ICC-Bedingungen fallen die Messwerte der Variablenausprägungen *ohne ICC* und *mit ICC* zusammen und zeigen keine systematischen Veränderungen über die einzelnen Geräuschbedingungen hinweg. Wenn beide erklärenden Variablen *ICC* und *Geräusch* sowie deren Interaktion in das Regressionsmodell integriert werden (vgl. Tabelle 6.21), lässt sich feststellen, dass *ICC* weiterhin einen signifikanten Effekt auf die Dauer der Sprachaufnahmen besitzt ($p < 0,05$). Des Weiteren wird deutlich, dass es eine signifikante Interaktion zwischen den unabhängigen Variablen gibt ($p < 0,05$). Dies bestätigt wiederum die Ergebnisse der zuvor durchgeführten deskriptiven Untersuchung. (vgl. [39, S. 1487]) Anhand von Abbildung 6.15 ist zu erkennen, dass die Messwerte der Signaldauer eine hohe Varianz

aufweisen, weshalb eine größere Stichprobe benötigt wird, um eine stärkere statistische Sicherheit der Ergebnisse zu erreichen.

Tabelle 6.21: Regressionsergebnisse für die abhängige Variable *dursignal* und die Interaktion der erklärenden Variablen *ICC* und *Geräusch*.

Call:
lm(formula = dursignal ~ ICC * Geraeusch - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-272.45	-81.21	-9.84	78.32	505.02

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	-80.06	38.18	-2.097	0.0388	*
ICCOhne ICC	30.94	38.18	0.810	0.4199	
Geraeusch50 km/h	-24.29	53.99	-0.450	0.6539	
GeraeuschStille	-43.56	53.99	-0.807	0.4219	
ICCOhne ICC:Geraeusch50 km/h	101.90	76.36	1.335	0.1854	
ICCOhne ICC:GeraeuschStille	181.17	76.36	2.373	0.0198	*

—

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 152.7 on 90 degrees of freedom
Multiple R-squared: 0.3598, Adjusted R-squared: 0.3171
F-statistic: 8.43 on 6 and 90 DF, p-value: 2.909e-07

Auf Basis der im Regressionsmodell 6.21 dargestellten erklärenden Variablen werden nach der Formel 6.2 für alle ICC- und Geräuschbedingungen des Produktions-experiments Schätzwerte errechnet. Anschließend wird eine lineare Transformation durchgeführt, im Zuge derer der Schätzwert der Bedingung *ohne ICC, Stille* von allen anderen Schätzwerten subtrahiert wird und somit als Basis den Wert 0 erhält. An diesen Schätzwerten, welche in Tabelle 6.22 aufgelistet sind, ist zu erkennen, dass die Dauer der Sprachaufnahmen ohne ICC-System mit steigender Geräuschbelastung immer weiter abnimmt. D. h., ohne ICC-System und bei Fahrgeräuschen von 50 km/h verkürzt sich die Dialogdauer gegenüber der stillen Referenzbedingung um durchschnittlich 59,998 Sekunden, während die Signaldauer bei einem weiteren Anstieg der Geräuschbelastung auf 130 km/h um weitere 77,607 Sekunden reduziert

6 Die Analyse der Sprachaufnahmen

wird. Durch Aktivierung des ICC-Systems steigt die Signaldauer bei einer Zunahme der Umgebungsgeräusche hingegen kontinuierlich an, wobei dieser Effekt jedoch deutlich geringer ist. Die Proband*innen sprechen mit ICC-System von Stille auf 50 km/h 19,271 Sekunden und bei einer Steigerung der Fahrgeräusche auf 130 km/h weitere 24,291 Sekunden länger miteinander. Auffällig ist, dass alle Schätzwerte mit ICC-System deutlich niedriger sind als ohne, wobei die größte Differenz zwischen den Werten der stillen Bedingung besteht. Mit zunehmender Geräuschbelastung nähern sich die Schätzwerte beider ICC-Bedingungen einander an. (vgl. [39, S. 1486f.])

Tabelle 6.22: Die auf Basis des Regressionsmodells aus Tabelle 6.21 errechneten transformierten Schätzwerte zur Signaldauer (Sek.).

ohne ICC			mit ICC		
Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
0	-59,998	-137,605	-292,163	-272,892	-248,601

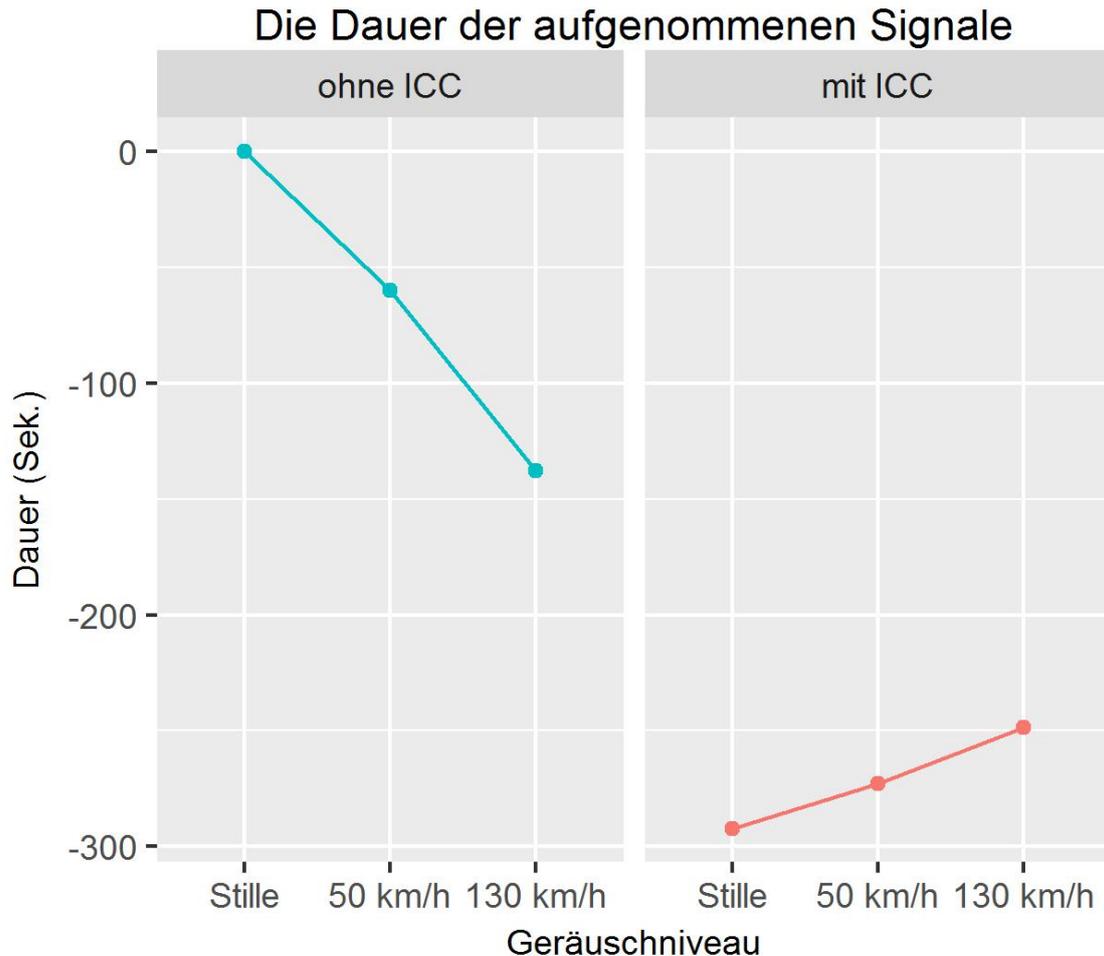


Abbildung 6.18: Die auf Basis des Regressionsmodells aus Tabelle 6.21 errechneten transformierten Schätzwerte zur Signaldauer. ([39, S. 1487])

Die absoluten Mittelwerte der mittelwertbereinigten Messwerte in Abbildung 6.16 zeigen hinsichtlich der Differenzen zwischen den einzelnen Geräuschbedingungen bei aktiviertem ICC-System dieselben Tendenzen auf wie die in Abbildung 6.18 dargestellten Schätzwerte. Anschließend durchgeführte t-Tests in R bestätigen jedoch, dass mit aktiviertem ICC-System weder ein signifikanter Unterschied zwischen der stillen Referenzbedingung und 50 km/h ($p=0,1869$) noch zwischen 50 km/h und 130 km/h ($p=0,2028$) nachgewiesen werden kann. Ein Vergleich der Bedingungen *Stille* und 130 km/h weist hingegen mit $p=0,05055$ einen schwach signifikanten Trend auf. Die Ergebnisse deuten darauf hin, dass ein Zusammenhang zwischen einem hohen Anstieg der Fahrgeräusche und der Veränderung der Signaldauer besteht. Solch ein Effekt müsste jedoch durch eine größere Stichprobe bzw. eine auf diese Systematik spezialisierte Studie zusätzlich bestätigt werden. Somit kann an dieser Stelle keine

6 Die Analyse der Sprachaufnahmen

systematische Veränderung der Dialogdauer bei unterschiedlichen Fahrgeräuschen mit ICC-System nachgewiesen werden (vgl. [39, S. 1487]).

Tabelle 6.23: Ergebnisse des t-Tests zur Signaldauer für die abhängigen Stichproben der Experimentbedingungen *mit ICC, Stille* und *mit ICC, 50 km/h*.

```
Paired t-test
data:  mitICC_stille$dursignal and mitICC_50kmh$dursignal
t = -1.3829, df = 15, p-value = 0.1869
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-48.97289 10.43164
sample estimates:
mean of the differences
-19.27063
```

Tabelle 6.24: Ergebnisse des t-Tests zur Signaldauer für die abhängigen Stichproben der Experimentbedingungen *mit ICC, 50 km/h* und *mit ICC, 130 km/h*.

```
Paired t-test
data:  mitICC_50kmh$dursignal and mitICC_130kmh$dursignal
t = -1.3319, df = 15, p-value = 0.2028
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-63.16232 14.58107
sample estimates:
mean of the differences
-24.29062
```

Tabelle 6.25: Ergebnisse des t-Tests zur Signaldauer für die abhängigen Stichproben der Experimentbedingungen *mit ICC, Stille* und *mit ICC, 130 km/h*.

```

Paired t-test

data:  mitICC_stille$dursignal and mitICC_130kmh$dursignal
t = -2.1257, df = 15, p-value = 0.05055
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-87.240356 0.117856
sample estimates:
mean of the differences
-43.56125

```

6.3.4 Sprachproduktionsleistung

Auf Basis der oben beschriebenen Segmentierungen und orthographischen Annotationen der Sprachsignale (vgl. Kapitel 5.7) wird untersucht, wie hoch der Anteil gesprochener Sprache gegenüber den Pausen in jedem Signal ist. Dazu wird die Dauer der Sprachproduktion in den Signalen jeder Probandin bzw. jedes Probanden s in jeder Experimentbedingung i ($p_{s,i}$) auf die Gesamtdauer des entsprechenden Signals $d_{s,i}$ normiert:

$$p'_{s,i} = \frac{p_{s,i}}{d_{s,i}}. \quad (6.7)$$

Zur Bestimmung der Sprachproduktionsdauer im Signal ist eine Separierung zwischen sprachlichen und nicht-sprachlichen Abschnitten nötig. Deshalb bezieht sich diese Analyse lediglich auf Sprachaufnahmen, deren Segmentierungen und Annotationen mithilfe von MAUS semi-automatisch durchgeführt und anschließend manuell korrigiert wurden. Dabei handelt es sich um die ersten 300 Sekunden der Sprachsignale von Dialogpaar 2, 3, 4, 7 und 8. Für diese 10 Sprecher*innen \times 6 Experimentbedingungen ergeben sich insgesamt 60 Sprachsignale. Die zuvor erstellten TextGrid-Dateien werden zur Durchführung der vorliegenden Analyse derart gekürzt, dass sie jeweils nur die ersten 300 Sekunden umfassen. Daraufhin werden mithilfe eines Praat-Skripts in jedem Signal die Dauern der Intervalle zusammengefasst, die nicht mit $\langle p \rangle$ gekennzeichnet sind. Die Ergebnisse werden in ms ausgegeben. Deshalb erfolgt zunächst eine Division durch 1000, um die entsprechenden Werte in Sekun-

den zu erhalten. Weiterhin wird die oben beschriebene Normierung an der Dauer des gesamten Signals (300 Sekunden) durchgeführt sowie eine Multiplikation mit 100, um Prozentwerte zu erhalten. Gleichung 6.7 kann demnach mit

$$d_{s,i} = 300 \quad (6.8)$$

folgendermaßen erweitert

$$p''_{s,i} = \frac{p'_{s,i}}{1000} \times 100 = \frac{p_{s,i}/300}{1000} \times 100 \quad (6.9)$$

und entsprechend gekürzt werden:

$$p''_{s,i} = \frac{p_{s,i}}{3000}. \quad (6.10)$$

Während der Sprachaufnahmen behielten alle Proband*innen den ausgewählten Sitzplatz (*vorne, hinten*) und die Sprechrolle (*Giver, Follower*) in allen Experimentbedingungen bei (vgl. Kapitel 5.4). Aufgrund der Paarung je zweier Dialogpartner*innen wird vermutet, dass sich die Sprachproduktionsleistungen von *Giver* und *Follower* bzw. von *vorne* und *hinten* sitzenden Sprecher*innen gegenseitig bedingen. D. h., dass eine Probandin bzw. ein Proband immer weniger Sprache produziert, je mehr die Dialogpartnerin bzw. der Dialogpartner spricht und umgekehrt. Sollte diese Annahme Gültigkeit besitzen, ist davon auszugehen, dass die unabhängigen Variablen *ICC* und *Geräusch* allein keinen Einfluss auf die Sprachproduktion im Signal haben. Demnach würden sich die Messwerte in unterschiedlichen ICC- und Geräuschbedingungen nicht signifikant verändern, solange die Dialogpartner*innen nicht anhand des Sitzplatzes oder der Sprechrolle voneinander getrennt werden. Aufgrund dieser Überlegungen wird das Verhältnis der Sprachproduktionsleistungen *innerhalb* jedes einzelnen Sprecher*innenpaares und dessen Veränderungen durch die Anwendung des ICC-Systems bzw. durch wechselnde Hintergrundgeräusche in den Vordergrund gerückt.

Um den Anteil der Sprachproduktion einzelner Sprecher*innen miteinander vergleichen zu können, werden die Ergebnisse aus Formel 6.10 durch die in Formel 6.1 dargestellte Berechnung wiederum mittelwertbereinigt. Bei der deskriptiven Analyse der Daten wird, ähnlich wie in den Kapiteln 6.3.1, 6.3.2 und 6.3.3, zunächst ein Linien-

diagramm betrachtet (vgl. Abbildung 6.19). Anders als bei den Untersuchungsparametern der durchschnittlichen Intensität und der durchschnittlichen Grundfrequenz sind hier keine systematischen Veränderungen der Messwerte zu erkennen. Dennoch scheint die Sprachproduktionsleistung mit ICC-System eine geringere Varianz aufzuweisen als ohne ICC-System. Die Mittelwerte in Abbildung 6.20 und 6.21 weisen darauf hin, dass sowohl zwischen den Sitzplätzen als auch zwischen den Sprechrollen der Proband*innen ein Unterschied besteht. Je nach Sitzplatz und Rolle findet ein anders gearteter Lombard-Effekt bei steigender Geräuschbelastung statt, wobei sich die Messwerte der Dialogpartner*innen gegenteilig verschieben. Z. B. sprechen *hinten* sitzende *Follower* ohne ICC-System immer weniger, während *vorne* sitzende *Giver* immer mehr sprechen. Mit ICC-System sind diese Effekte weniger eindeutig.

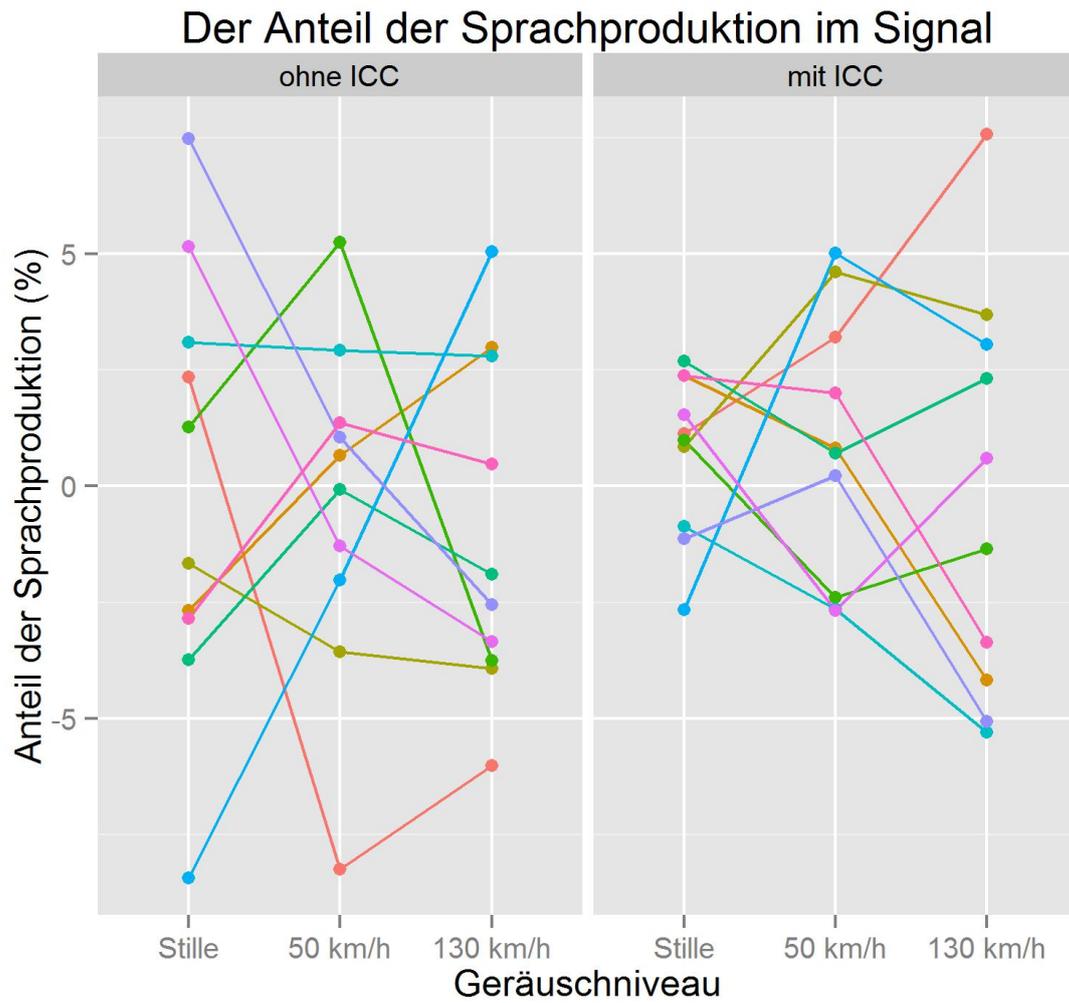


Abbildung 6.19: Der mittelwertbereinigte Anteil gesprochener Sprache in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren Daten einzelner Proband*innen.

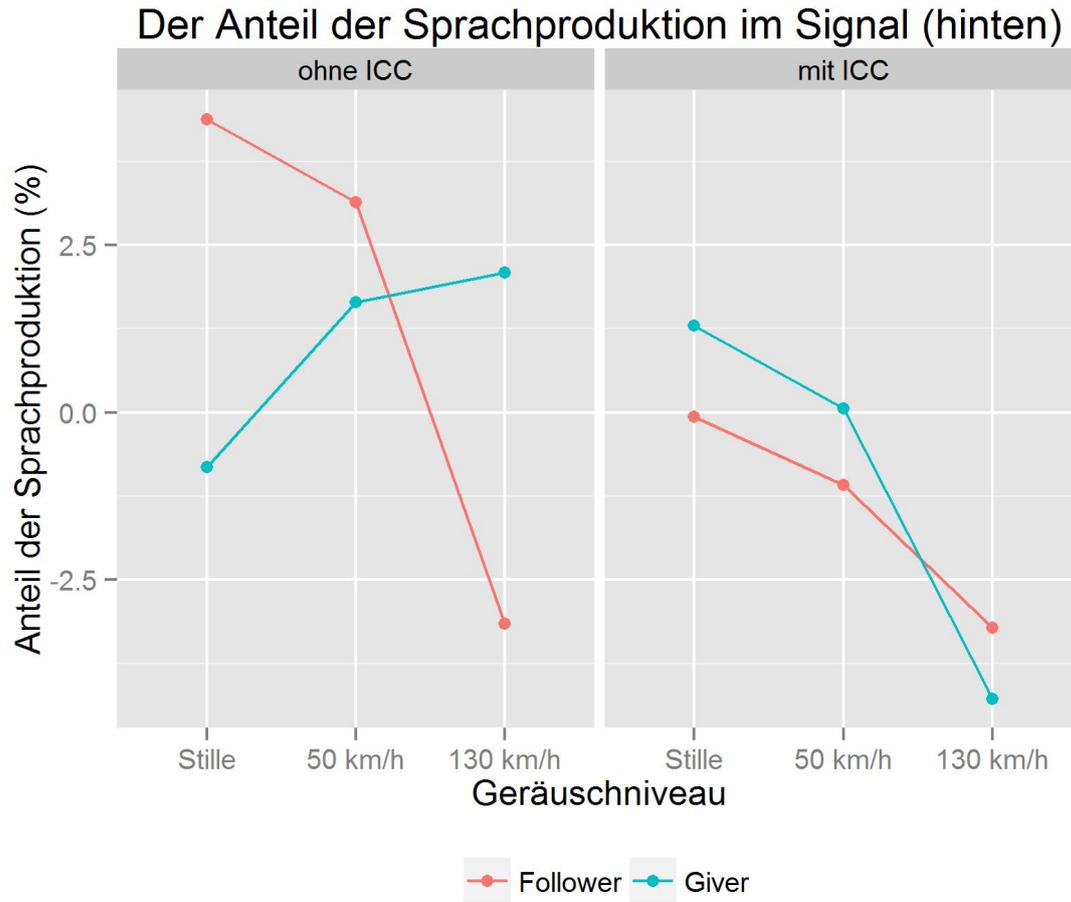


Abbildung 6.20: Die Mittelwerte des mittelwertbereinigten Anteils gesprochener Sprache über alle Sprecher*innen hinweg auf dem hinteren Sitzplatz.

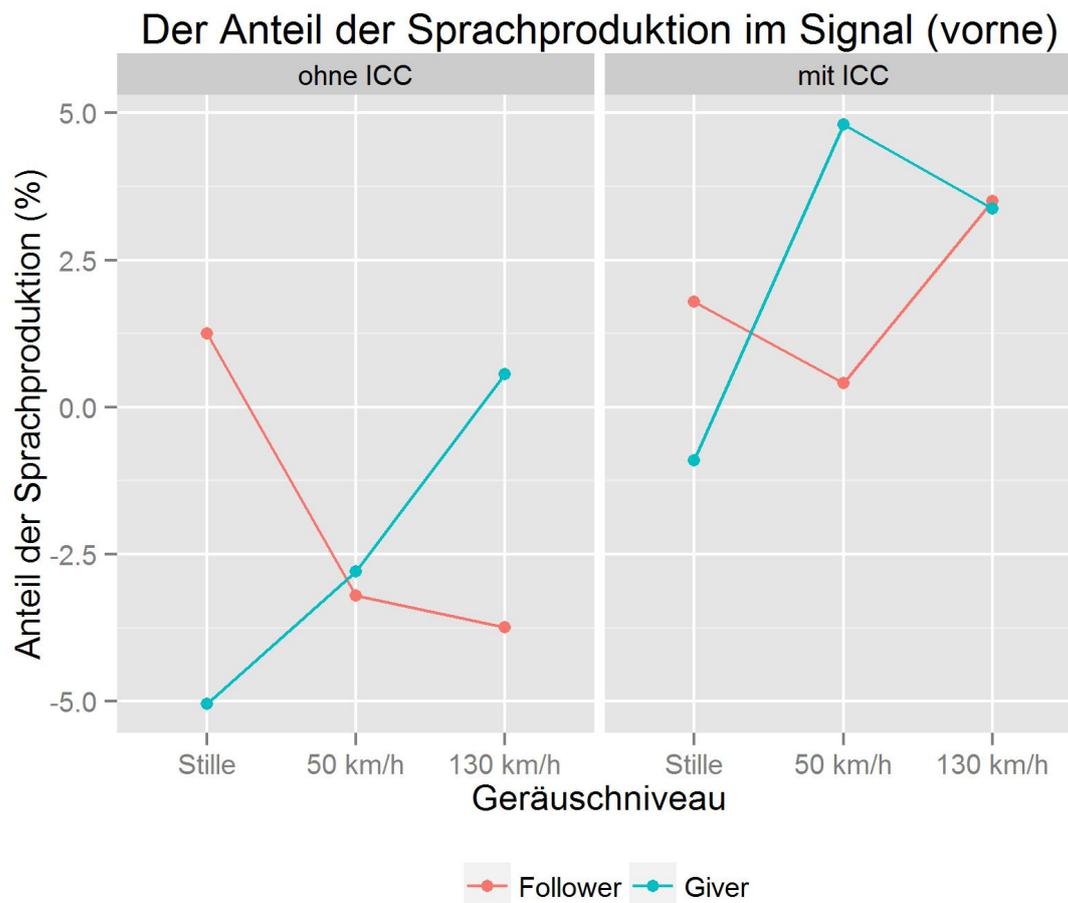


Abbildung 6.21: Die Mittelwerte des mittelwertbereinigten Anteils gesprochener Sprache über alle Sprecher*innen hinweg auf dem vorderen Sitzplatz.

In der Regressionsanalyse wird wiederum ausgehend von einem Nullmodell mit der abhängigen Variable *amountofspeech* zunächst die erklärende Variable *ICC* hinzugefügt (vgl. Tabelle 6.26). Die Resultate zeigen entsprechend dem Liniendiagramm in Abbildung 6.19 keinen signifikanten Einfluss des ICC-Systems auf die Menge der Sprachproduktion im Dialog auf. Ebenso verhält es sich mit der unabhängigen Variable *Geräusch*, die wiederum allein keinen Einfluss auf den untersuchten Parameter hat (vgl. Tabelle 6.27). Weiterhin fällt auf, dass sich der Wert des *Adjusted R-squared* für beide Modelle 6.26 und 6.27 im negativen Bereich befindet und diese somit keinerlei Aussagekraft besitzen. Diese Ergebnisse bestätigen die oben beschriebene Annahme, dass die unabhängigen Variablen *ICC* und *Geräusch* allein keinen Einfluss auf die Messwerte haben. Wird dasselbe Modell nur mit *Sitzplatz* oder *Rolle* berechnet, ergibt sich jeweils ein *Adjusted R-squared* von -0,03448 und $p=1$.

Tabelle 6.26: Regressionsergebnisse für die abhängige Variable *amountofspeech* und die erklärende Variable *ICC*.

```
Call:
lm(formula = amountofspeech ~ ICC - 1, data = dataicc)
```

Residuals:

Min	1Q	Median	3Q	Max
-7.9583	-2.8688	0.3617	2.3093	7.9598

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
ICCmit ICC	0.4713	0.6476	0.728	0.47
ICCoohne ICC	-0.4713	0.6476	-0.728	0.47

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.547 on 58 degrees of freedom
Multiple R-squared: 0.01793, Adjusted R-squared: -0.01593
F-statistic: 0.5296 on 2 and 58 DF, p-value: 0.5917

Tabelle 6.27: Regressionsergebnisse für die abhängige Variable *amountofspeech* und die erklärende Variable *Geräusch*.

```
Call:
lm(formula = amountofspeech ~ Geraeuschk130 + Geraeuschk50 + GeraeuschkStille - 1, data = dataicc)
```

Residuals:

Min	1Q	Median	3Q	Max
-8.7949	-2.7926	0.4764	2.4117	8.1883

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
Geraeuschk130 km/h	-0.6099	0.8011	-0.761	0.450
Geraeuschk50 km/h	0.2446	0.8011	0.305	0.761
GeraeuschkStille	0.3653	0.8011	0.456	0.650

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.583 on 57 degrees of freedom
Multiple R-squared: 0.01522, Adjusted R-squared: -0.03661
F-statistic: 0.2936 on 3 and 57 DF, p-value: 0.8299

6 Die Analyse der Sprachaufnahmen

Die Verwendung des ICC-Systems im Zusammenhang mit den unabhängigen Variablen *Sitzplatz* und *Rolle* ergibt indes ein anderes Bild. Abbildung 6.22 zeigt auf, dass das ICC-System abhängig vom *Sitzplatz* einen Einfluss auf die untersuchten Daten hat. Ohne ICC-System wird *hinten* wesentlich mehr Sprache produziert als *vorne*, wobei sich dieses Verhältnis mit Verwendung des ICC-Systems umkehrt (vgl. [39, S. 1487]). Das dazugehörige Regressionsmodell in Tabelle 6.28 weist eine hohe Signifikanz für die Interaktion von *ICC* und *Sitzplatz* ($p < 0,0005$) sowie ein *Adjusted R-squared* von 0,1802 auf. Die p-Werte der Ausprägungen von *ICC* sind zudem deutlich niedriger als im vorangegangenen Modell 6.26. Der *Sitzplatz* besitzt darüber hinaus einen signifikanten Einfluss auf die abhängige Variable ($p < 0,01$). Aufgrund der obigen Beobachtungen kann jedoch festgehalten werden, dass diese Einflüsse nur durch die Interaktion beider unabhängiger Variablen entstehen.

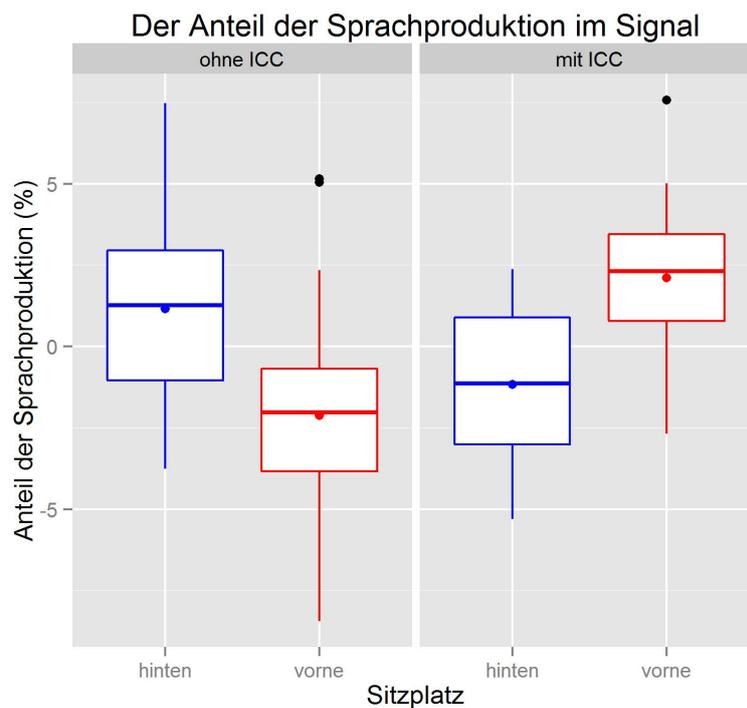


Abbildung 6.22: Der mittelwertbereinigte Anteil gesprochener Sprache für unterschiedliche Sitzplätze, ohne und mit Verwendung des ICC-Systems.

Tabelle 6.28: Regressionsergebnisse für die abhängige Variable *amountofspeech* und die erklärenden Variablen *ICC* und *Sitzplatz*.

Call:
lm(formula = amountofspeech ~ ICC * Sitzplatz - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-6.3194	-1.5353	0.0962	1.8482	7.2685

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	-1.1676	0.8227	-1.419	0.161335	
ICCoohne ICC	1.1676	0.8227	1.419	0.161335	
Sitzplatzvorne	3.2778	1.1634	2.817	0.006676	**
ICCoohne ICC:Sitzplatzvorne	-6.5556	1.6453	-3.984	0.000197	***

—

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.186 on 56 degrees of freedom
Multiple R-squared: 0.2348, Adjusted R-squared: 0.1802
F-statistic: 4.297 on 4 and 56 DF, p-value: 0.004211

Die Sprechrolle einer Probandin bzw. eines Probanden scheint zudem kaum Einfluss auf die Sprachproduktionsleistung zu haben. Lediglich ohne ICC-System zeichnet sich eine sehr geringe Tendenz dahingehend ab, dass *Follower* generell etwas weniger Sprache produzieren als *Giver* (vgl. Abbildung 6.23). Mit Anwendung des ICC-Systems ist das Verhältnis der Sprachproduktionsleistung zwischen *Giver* und *Follower* hingegen relativ ausgeglichen. Eine Erweiterung des Regressionsmodells aus Tabelle 6.28 um die Interaktion der unabhängigen Variablen *ICC* und *Rolle* führt zu einem geringeren *Adjusted R-squared* von 0,1553 und keinen signifikanten p-Werten für die unabhängige Variable *Rolle* oder deren Interaktion mit *ICC*.

Weiterhin wird der Einfluss der unabhängigen Variablen *Geräusch* in Wechselwirkung mit *Sitzplatz* und *Rolle* untersucht. Abbildung 6.24 verdeutlicht, dass die Sprachproduktionsleistung der *vorne* und *hinten* sitzenden Proband*innen in *Stille* relativ ausgeglichen zu sein scheint, wobei die Durchschnittswerte *vorne* jedoch

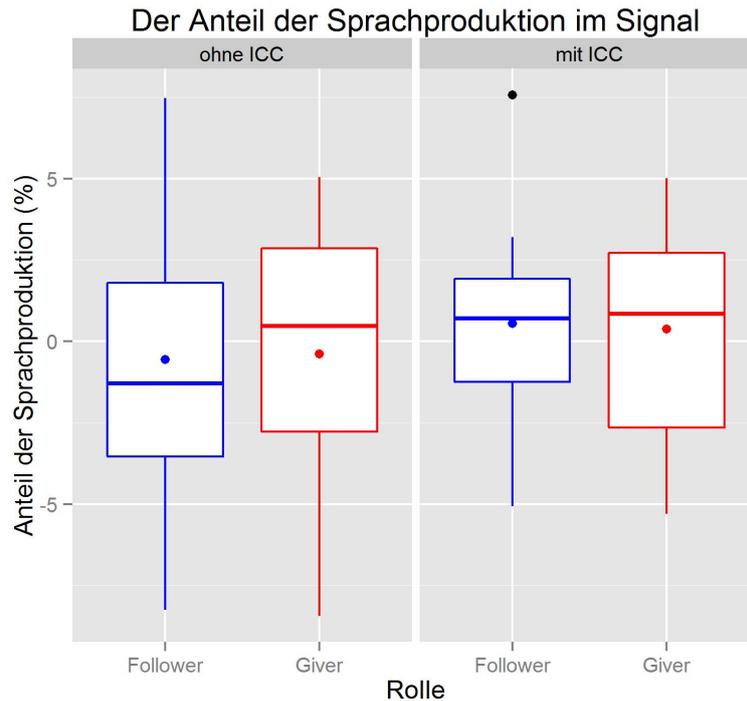


Abbildung 6.23: Der mittelwertbereinigte Anteil gesprochener Sprache für unterschiedliche Rollen, ohne und mit Verwendung des ICC-Systems.

niedriger liegen. Bei 50 km/h spricht die *hinten* sitzende Person im Mittel mehr als die *vorne* sitzende. Mit einer Geräuschbelastung von 130 km/h kehrt sich dieses Verhältnis um. Weitere systematische Veränderungen in den Messwerten sind in Abbildung 6.25 erkennbar. Während der *Follower* in *Stille* mehr spricht als der *Giver*, ist mit steigenden Hintergrundgeräuschen das Gegenteil der Fall. Dabei wird die Sprachproduktionsleistung des *Followers* immer geringer. Die Sprachproduktionsleistung des *Givers* steigt von *Stille* zu 50 km/h an und nimmt von 50 km/h zu 130 km/h im Mittel wieder ab. Aufgrund dieser Beobachtungen wird dem in Tabelle 6.28 dargestellten Regressionsmodell zunächst die Interaktion von *Geräusch* und *Sitzplatz* hinzugefügt. Dabei steigert sich das *Adjusted R-squared* des Regressionsmodells auf einen Wert von 0,216 (vgl. Tabelle 6.29). Weiterhin wird dieses Modell um die Interaktion der unabhängigen Variablen *Geräusch* und *Rolle* erweitert. In diesem Regressionsmodell steigert sich das *Adjusted R-squared* wiederum auf einen Wert von 0,2936 (vgl. Tabelle 6.30). Somit ist das für diesen Parameter beste Regressionsmodell erreicht.

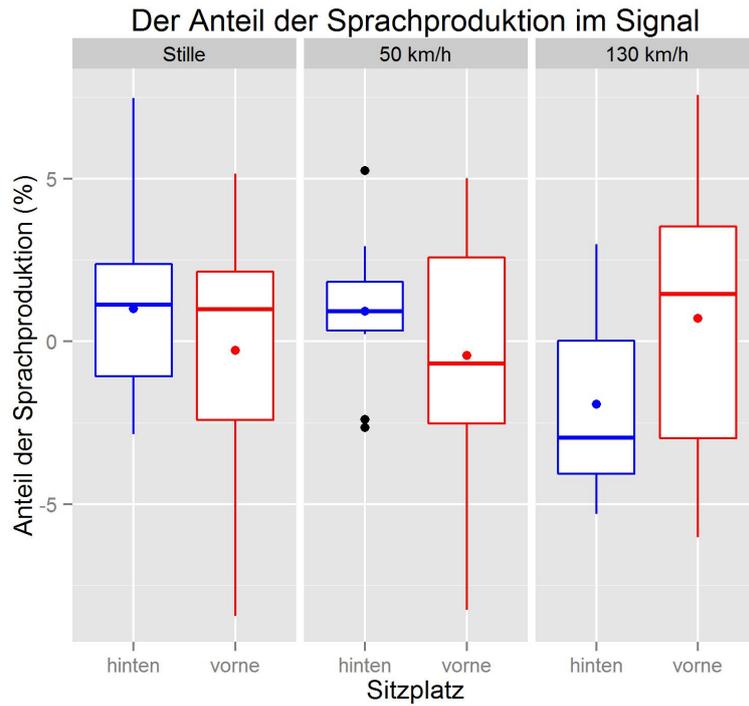


Abbildung 6.24: Der mittelwertbereinigte Anteil gesprochener Sprache für unterschiedliche Sitzplätze bei verschiedenen Geräuschbedingungen.

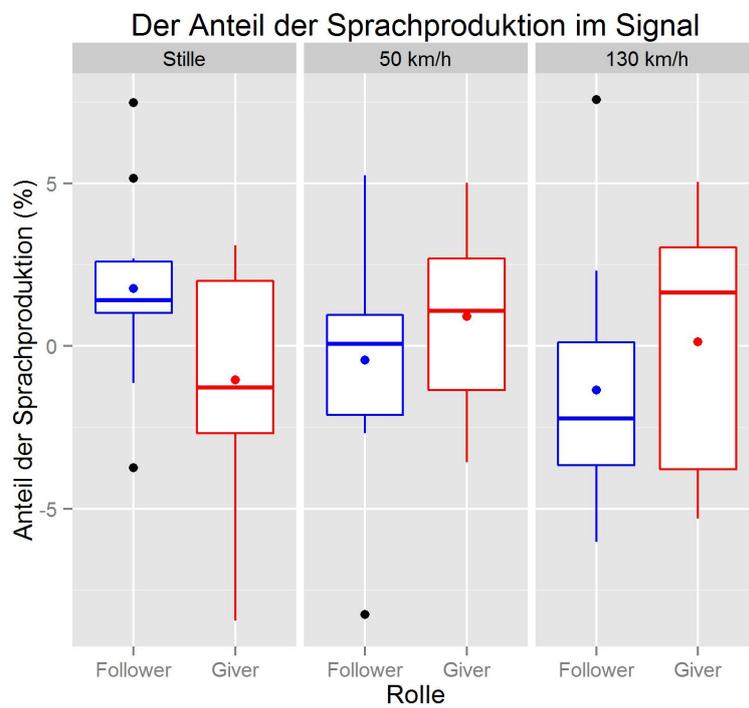


Abbildung 6.25: Der mittelwertbereinigte Anteil gesprochener Sprache für unterschiedliche Rollen bei verschiedenen Geräuschbedingungen.

Tabelle 6.29: Regressionsergebnisse für die abhängige Variable *amountofspeech* und die Interaktionen der erklärenden Variablen *ICC* und *Sitzplatz* sowie *Geräusch* und *Sitzplatz*.

Call:

lm(formula = amountofspeech ~ ICC * Sitzplatz + Geraeusch * Sitzplatz - 1, data = dataicc)

Residuals:

Min	1Q	Median	3Q	Max
-6.0416	-1.8318	-0.3956	1.5817	7.5463

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	-3.0986	1.1377	-2.724	0.008772	**
ICCoohne ICC	-0.7633	1.1377	-0.671	0.505226	
Sitzplatzvorne	5.9199	1.6090	3.679	0.000556	***
Geraeusch50 km/h	2.8535	1.3934	2.048	0.045638	*
GeraeuschStille	2.9394	1.3934	2.110	0.039729	*
ICCoohne ICC:Sitzplatzvorne	-6.5556	1.6090	-4.074	0.000158	***
Sitzplatzvorne:Geraeusch50 km/h	-3.9980	1.9706	-2.029	0.047608	*
Sitzplatzvorne:GeraeuschStille	-3.9284	1.9706	-1.994	0.051461	.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.116 on 52 degrees of freedom

Multiple R-squared: 0.3206, Adjusted R-squared: 0.216

F-statistic: 3.067 on 8 and 52 DF, p-value: 0.006632

Tabelle 6.30: Regressionsergebnisse für die abhängige Variable *amountofspeech* und die Interaktionen der erklärenden Variablen *ICC* und *Sitzplatz*, *Geräusch* und *Sitzplatz* sowie *Geräusch* und *Rolle*.

Call:

```
lm(formula = amountofspeech ~ ICC * Sitzplatz + Geraeusch * Sitzplatz + Geraeusch * Rolle
- 1, data = dataicc)
```

Residuals:

Min	1Q	Median	3Q	Max
-5.2548	-1.8829	-0.4041	2.2110	6.2651

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
ICCmit ICC	-4.3529	1.3500	-3.224	0.002249	**
ICCoohne ICC	-2.0176	1.3500	-1.495	0.141440	
Sitzplatzvorne	6.3380	1.5510	4.086	0.000162	***
Geraeusch50 km/h	3.4402	1.7498	1.966	0.054968	.
GeraeuschStille	6.1156	1.7498	3.495	0.001016	**
RolleGiver	2.0905	1.3500	1.549	0.127927	
ICCoohne ICC:Sitzplatzvorne	-6.5556	1.5273	-4.292	8.32e-05	***
Sitzplatzvorne:Geraeusch50 km/h	-4.1935	1.9091	-2.197	0.032814	*
Sitzplatzvorne:GeraeuschStille	-4.9871	1.9091	-2.612	0.011909	*
Geraeusch50 km/h:RolleGiver	-0.9778	1.9091	-0.512	0.610834	
GeraeuschStille:RolleGiver	-5.2937	1.9091	-2.773	0.007838	**

—
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.958 on 49 degrees of freedom

Multiple R-squared: 0.4231, Adjusted R-squared: 0.2936

F-statistic: 3.267 on 11 and 49 DF, p-value: 0.002066

Die in Tabelle 6.30 dargestellten Ergebnisse bestätigen, dass es eine hoch signifikante Interaktion zwischen *ICC* und *Sitzplatz* ($p < 0,0001$), signifikante Wechselwirkungen zwischen *Geräusch* und *Rolle* ($p < 0,01$) sowie eine signifikante Tendenz für die Interaktion von *Geräusch* und *Sitzplatz* ($p < 0,5$) gibt. Die *Rolle* allein hat keinen signifikanten Einfluss auf die Messwerte. Für die Ausprägungen der unabhängigen Variablen *ICC* ($p < 0,005$), *Geräusch* ($p < 0,005$) und *Sitzplatz* ($p < 0,0005$)

werden dagegen signifikante Einflüsse auf die Messwerte angezeigt. Dies widerspricht den Ergebnissen der oben beschriebenen Regressionsmodelle 6.26 und 6.27, dass die unabhängigen Variablen *ICC* und *Geräusch* allein keinen Einfluss auf die Sprachproduktionsleistung der Proband*innen besitzen. Die in Tabelle 6.30 dargestellten Signifikanzniveaus ergeben sich aus den Interaktionen der unabhängigen Variablen. Dies wird auch aus den Abbildungen der deskriptiven Analyse in 6.22, 6.24 und 6.25 ersichtlich. (vgl. [39, S. 1487]).

Die Schätzwerte der Variablenausprägungen aus Tabelle 6.30 werden, ebenso wie in den vorangegangenen Unterkapiteln, auch für die Sprachproduktionsleistung anhand der Regressionskoeffizienten aus Tabelle 6.30 und der Formel 6.2 berechnet, wobei die Experimentbedingung *ohne ICC, Stille, hinten, Follower* als Basis von allen Schätzwerten subtrahiert wird und somit 0 ergibt. Aus den dadurch erhaltenen Schätzwerten in Tabelle 6.31 und Abbildung 6.26 bzw. 6.27 ist zu erkennen, dass sich die Messwerte für gepaarte Sprechrollen gegenteilig verhalten, wenn sich die Geräuschbedingungen ändern. An dieser Stelle werden unter gepaarten Sprechrollen *Giver* und *Follower* verstanden, die jeweils unterschiedliche Sitzplätze einnehmen. D. h., während die Sprachproduktionsleistung z. B. bei *hinten* sitzenden *Followern ohne ICC* bei zunehmender Geräuschbelastung immer geringer wird, steigt sie für *vorne* sitzende *Giver* in den entsprechenden ICC- und Geräuschbedingungen an. *Hinten* sitzende *Giver* sprechen bei *50 km/h ohne ICC* zunächst mehr, bevor die Messwerte bei *130 km/h* abfallen. Entsprechend reden *vorne* sitzende *Follower ohne ICC* bei *50 km/h* weniger und bei *130 km/h* wieder etwas mehr. Insgesamt haben die ansteigenden Umgebungsgeräusche einen stärkeren Effekt auf die Sprachproduktionsleistung, wenn der *Follower hinten* sitzt und der *Giver vorne*. Dies gilt für beide ICC-Bedingungen gleichermaßen. Weiterhin wird aus Tabelle 6.31 deutlich, dass ohne ICC-System die Werte in jeder Geräuschbedingung *hinten* höher sind als *vorne*. Dies gilt sowohl für *Giver* als auch für *Follower*. Mit Anwendung des ICC-Systems kehrt sich dieser Effekt um, denn nun sind die Messwerte beider Sprechrollen *hinten* niedriger als *vorne*. Zudem wird aus den Schätzwerten ersichtlich, dass *Giver* in der stillen Referenzbedingung immer weniger sprechen als *Follower*, sowohl *ohne ICC* als auch *mit ICC*. Sobald Hintergrundgeräusche auftreten, kehrt sich dieser Effekt um. (vgl. [39, S. 1486ff.])

Tabelle 6.31: Die auf Basis des Regressionsmodells aus Tabelle 6.30 errechneten transformierten Schätzwerte zur Sprachproduktionsleistung (%).

Hinten:

Follower	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
	0	-2,68	-6,12	-2,34	-5,01	-8,45
Giver	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
	-3,20	-1,56	-4,03	-5,54	-3,9	-6,36

Vorne:

Follower	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
	-5,21	-7,09	-6,33	-0,98	-2,87	-2,11
Giver	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
	-8,41	-5,97	-4,24	-4,19	-1,75	-0,02

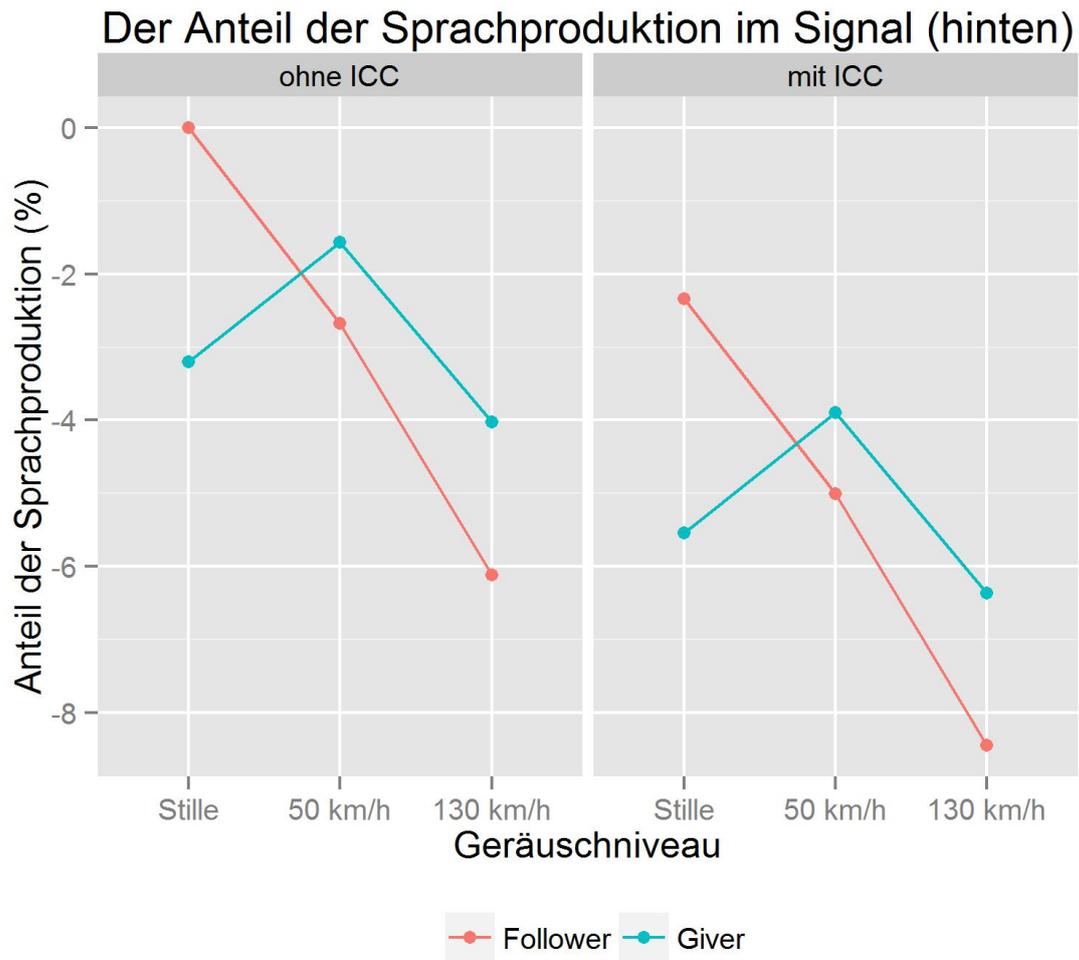


Abbildung 6.26: Die auf Basis des Regressionsmodells aus Tabelle 6.30 errechneten transformierten Schätzwerte zur Sprachproduktionsleistung für den hinteren Sitzplatz. [39, S. 1487]

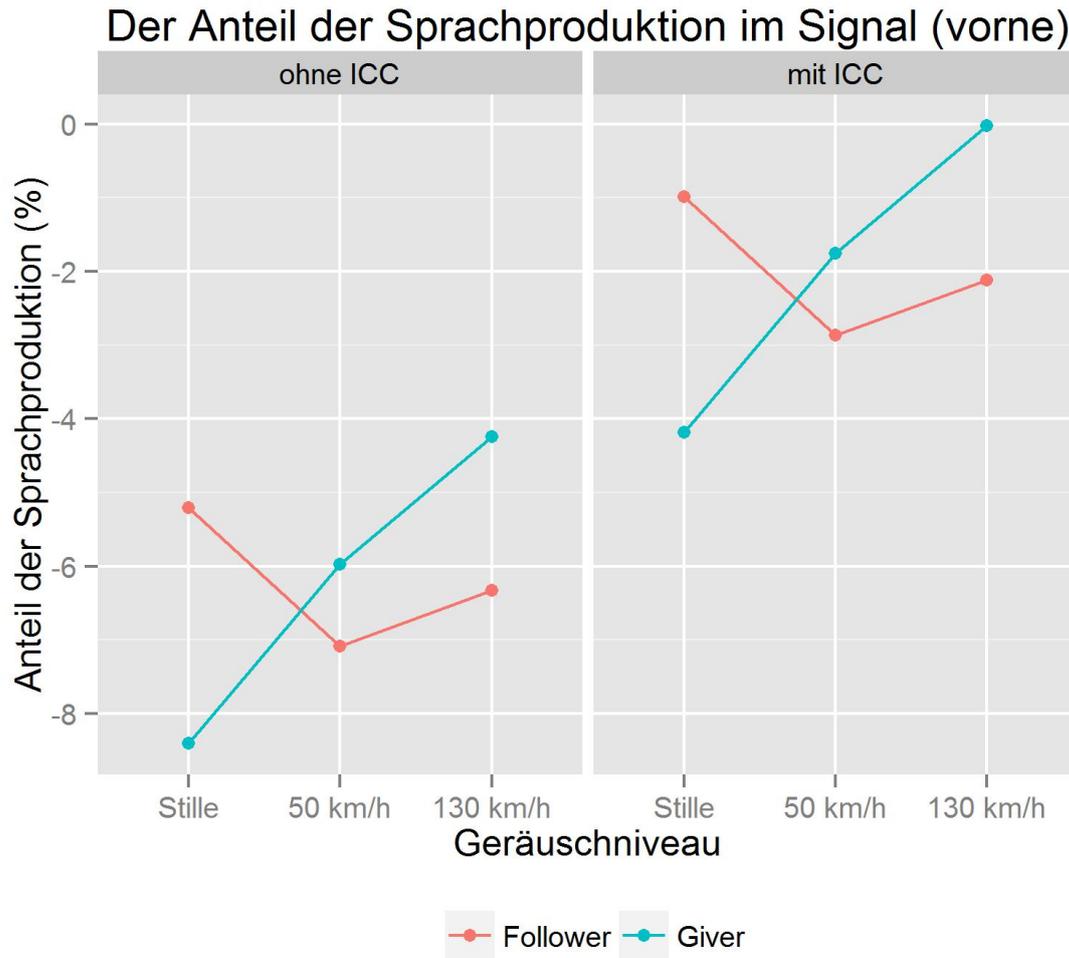


Abbildung 6.27: Die auf Basis des Regressionsmodells aus Tabelle 6.30 errechneten transformierten Schätzwerte zur Sprachproduktionsleistung für den vorderen Sitzplatz. [39, S. 1487]

6 Die Analyse der Sprachaufnahmen

Die Abbildungen 6.20 und 6.21 zeigen die Mittelwerte der Sprachproduktionsleistung für den hinteren und vorderen Sitzplatz. Im Vergleich zu den in Abbildung 6.26 und 6.27 dargestellten Schätzwerten der Regressionsanalyse werden in den Mittelwerten auch die Variablenkombinationen *ICC * Geräusch*, *ICC * Rolle* und *Sitzplatz * Rolle* in die Darstellung der Messwerte integriert. Diese Interaktionen besitzen jedoch keine signifikanten Einflüsse auf die Messwerte der Sprachproduktionsleistung, wie in der vorliegenden statistischen Analyse bewiesen wurde. Weiterhin werden die in den Mittelwerten aus Abbildung 6.20 und 6.21 beobachteten Effekte der Fahrgeräusche auf die Messwerte nur teilweise durch die Regressionsanalyse als signifikante Einflüsse auf die Sprachproduktionsleistung bestätigt. So zeigt sich in den Mittelwerten (Abbildung 6.20) für *ohne ICC, hinten, Giver* ein Anstieg der Messwerte von *50 km/h* zu *130 km/h*, wohingegen die Ergebnisse der Regressionsanalyse (Abbildung 6.26) für diese Experimentbedingungen eine Abnahme der Messwerte von *50 km/h* zu *130 km/h* aufdecken. In derselben Bedingung *ohne ICC, hinten, Giver* zeigen die Messwerte jedoch von *Stille* zu *50 km/h* sowohl in den Mittelwerten als auch in den Schätzwerten der Regression einen Anstieg der Messwerte, d. h., dieselbe Effektrichtung, an.

7 Perzeptionsexperiment

Das folgende Kapitel beschreibt die Methode, Durchführung und Analyse des Wahrnehmungsexperiments. Dabei wird ein Semantisches Differential verwendet, um sowohl wahrgenommene Lombard-Eigenschaften als auch emotionale Assoziationen in den Sprachaufnahmen des SPID-Korpus feststellen zu können. Anhand eines *Online*-Fragebogens werden Ausschnitte der zuvor generierten Sprachsignale von Proband*innen bewertet, welche nicht am vorangegangenen Produktionsexperiment teilgenommen haben. Anschließend erfolgt eine deskriptive und inferenzstatistische Analyse der erhobenen Daten mit Hilfe der Programmiersprache R.

7.1 Das Semantische Differential

Die Methode des Semantischen Differentials (Polaritätsprofil) ist ein „Von OSGOOD ET AL. [1957] entwickeltes Messverfahren zur Bestimmung der konnotativen emotionalen Bedeutungskomponenten von sprachlichen Ausdrücken“ ([5, S. 617]). Dabei bewerten Proband*innen ein Wort anhand von Skalen, an deren Enden antonyme Adjektive stehen, wie z. B. „glücklich – traurig“ (vgl. [5, S. 617]). Diese Messtechnik besitzt keine festen Einheiten, sondern wird an die Anforderungen der jeweiligen Forschungsfrage angepasst. Es existieren jedoch die drei festen Dimensionen der Valenz (*evaluation*), Potenz (*potency*) und Aktivität (*activity*), auf die im Folgenden näher eingegangen wird. Osgood et al. (1975) beziehen sich mit dem Begriff *concept* auf den Stimulus, der von einem Subjekt bewertet wird. Ein solches *concept* kann z. B. verbal sein, gemalt oder ein Sonarsignal. Aufgrund der zeitlichen Begrenzungen und der beschränkten Anzahl an Proband*innen muss zudem eine Auswahl an Stimuli getroffen werden, da nicht alle relevanten Konzepte eines Bereichs komplett abgedeckt werden können. Dabei sollten die zu beurteilenden Objekte relevant für den betreffenden Forschungsbereich sein und diesen repräsentieren. (vgl. [57, S. 76f.]

7 Perzeptionsexperiment

Die Stimuli des vorliegenden Perzeptionsexperiments sind Ausschnitte derjenigen im Labor aufgenommenen Sprachsignale des SPID-Korpus, welche zuvor nach akustisch-prosodischen Parametern analysiert worden sind (vgl. Kapitel 6.3). Um der zeitlichen Begrenzung bei der Durchführung des Semantischen Differentials Rechnung zu tragen, d. h., Ermüdungseffekte bei den Teilnehmer*innen zu vermeiden, sollte das Wahrnehmungsexperiment nicht viel mehr als 10 Minuten in Anspruch nehmen. Als Stimuli werden deshalb lediglich die Sprachaufnahmen von drei verschiedenen Proband*innen verwendet. Bei zwei ICC- und drei Geräuschbedingungen pro Sprecher*in ergeben sich daraus 3 (Sprecher*innen) \times 6 (Bedingungen) = 18 Stimuli. Diese werden in Kapitel 7.3 näher beschrieben.

Der semantische Raum ist multidimensional. Nach Osgood et al. (1975) sind dabei drei Faktoren dominant: *evaluation*, *potency* und *activity*. Wenn Proband*innen die Bedeutungen von Konzepten ausdifferenzieren, kann die Varianz auf bestimmten Skalen, wie z. B. Aktivitätsskalen, relativ unabhängig von der Variation auf anderen Skalen, wie z. B. Evaluierungsskalen, sein. Aus ihren Untersuchungen schließen sie, dass der bewertende Faktor (*evaluative*) der erste Faktor des semantischen Raums ist (z. B. „schön-hässlich“) und bei einer Beurteilung annähernd $\frac{1}{2}$ bis $\frac{3}{4}$ der extrahierbaren Varianz erklärt. Er bezieht sich auf die Einstellung bzw. Haltung der Probandin bzw. des Probanden. Dieser wird gefolgt vom Stärkefaktor (*potency*) (z. B. „stark-schwach“), welcher im Gegensatz zum ersten Faktor etwa halb so viel Varianz ausmacht bzw. erklärt. Er beinhaltet Kraft und damit zusammenhängende Aspekte, wie Gewicht und Größe. Als drittes folgt der Aktivitätsfaktor (*activity*) (z. B. „aktiv-passiv“), welcher in seinem explikativen Varianzeffekt der zweiten Dimension entspricht oder leicht darunter liegt. Er bezieht sich z. B. auf Unruhe, Geschwindigkeit und Aufregung. Darüber hinaus können weitere Faktoren erschlossen werden, die jedoch höchstens halb so viel Varianz erklären wie der zweite und dritte Faktor. Wenn die Auswahl der zu beurteilenden Konzepte eingeschränkt ist, kann sich der Erklärungsumfang dieser Faktoren verändern. Dabei können z. B. der zweite und der dritte Faktor zu einem Dynamikfaktor (*dynamism*) verschmelzen, wenn die Beurteilung sich auf gesellschaftspolitische Konzepte beschränkt. Somit scheinen Dinge, die „stark“ sind, auch „aktiv“ zu sein und umgekehrt. Durch das zu beurteilende Konzept kann sich die faktorielle Zusammensetzung einzelner Skalen verändern. Während bei einem Sonarsignal „niedrig (pitch)“ angenehmer zu sein

scheint, entspricht der Gegensatz „hoch-niedrig“ bei der Beurteilung sozialer Konzepte dem Gegensatz von „gut-schlecht“. (vgl. [5, S. 617f.], [10, S. 259] und [57, S. 71ff.])

Um einen spezifischen Faktor zu repräsentieren, verwenden Osgood et al. (1975) üblicherweise je drei Skalen, welche linear sind und zwischen zwei polaren Gegensätzen liegen (vgl. [57, S. 78f.]). Sie bezeichnen die Paarung eines Konzepts mit einer bestimmten Skala als *item*, wobei jede Beurteilung eines solchen *items* durch eine Probandin bzw. einen Probanden einen *bit* an Information liefert (vgl. [57, S. 80]). Zur Beurteilung der Stimuli im vorliegenden Perzeptionsexperiment werden zum einen die drei genannten Hauptfaktoren nach Osgood et al. (1975) gewählt und zum anderen der Faktor „Lombard“, welcher im Vorfeld des Wahrnehmungsexperiments als bedeutungsvoll eingestuft und in Kapitel 7.2 näher beschrieben wird. Wenn jedem dieser vier Faktoren drei Skalen zugeordnet werden würden, ergäbe sich eine Menge von $18 \text{ (Stimuli)} \times 12 \text{ (Skalen)} = 216 \text{ items}$, wodurch auf $18 \text{ (Stimuli)} \times 12 \text{ (Skalen)} \times X \text{ (Proband*innen)} = Y \text{ bits}$ geschlossen werden kann. Nach Osgood et al. (1975) kann sogar die bzw. der langsamste Hochschulstudierende mindestens 10 *items* pro Minute beurteilen, wobei sich die meisten einer Urteilsrate von 20 *items* pro Minute annähern, sobald sie sich eingewöhnt haben (vgl. [57, S. 80]). Im Perzeptionsexperiment würde eine Probandin bzw. ein Proband demnach zwischen 10,8 und 21,6 Minuten brauchen, um alle Skalen des Semantischen Differentials zu beurteilen. Um einem erhöhten Ermüdungseffekt und damit einer geringer werdenden Konzentrationsfähigkeit bei den Hörer*innen vorzubeugen, werden jedem Faktor des Semantischen Differentials deshalb lediglich zwei Skalen zugeordnet, sodass eine Gesamtmenge von $18 \text{ (Stimuli)} \times 8 \text{ (Skalen)} = 144 \text{ items}$ entsteht und eine Teilnehmerin bzw. ein Teilnehmer zwischen 7,2 und 14,4 Minuten zur Durchführung des Experiments benötigen sollte. Die im Perzeptionsexperiment verwendeten Faktoren und Skalen zur Durchführung des Semantischen Differentials sind in Tabelle 7.1 dargestellt:

Tabelle 7.1: Faktoren und dazugehörige Skalen im Semantischen Differential des Perzeptionsexperiments.

Faktor	Skala
evaluation	unfreundlich - freundlich emotional - sachlich
potency	gehemmt - ungehemmt steif - locker
activity	angestrengt - entspannt stockend - flüssig
lombard	laut - leise verständlich - unverständlich

Die Methode des Semantischen Differentials ist bereits in verschiedenen Studien erfolgreich angewendet worden. Hierzu zählen u. a. prosodische Untersuchungen zur wahrgenommenen Bedeutung unterschiedlicher F0-Gipfel und -Täler bzw. Intonationskonturen (vgl. [9], [10], [33] und [54]) sowie zu den phonetischen Profilen negativ und positiv intensivierender Emphase (vgl. [35]). Niebuhr und Dombrowski (2010) untersuchten Bedeutungsunterschiede bei konkav und konvex ansteigenden Intonationskonturen am Phrasenende. Dazu wählten die Autoren Skalen, welche das kommunikative Konzept von *restriction* und *activation* repräsentieren, wobei diese Termini zwei kommunikative Grundsätze bzw. Gesten zur Kontrolle der zwischenmenschlichen Interaktion bezeichnen (vgl. [11, S. 193f.] und [54, S. 1f.]):

„The first pattern is seen as a *restrictive* gesture, e.g. preventing the dialogue partner from turn taking. The second one is viewed as an *activating* gesture, inviting the coparticipant to contribute.“ ([11, S. 176])

Auf ähnliche Weise wird in der vorliegenden Arbeit den Skalen der Faktoren *evaluation*, *potency* und *activity* in Tabelle 7.1 das kommunikative Konzept *Anstrengung - Entspannung* zugrunde gelegt, wobei ebenfalls eine der Skalen mit diesem Konzept übereinstimmt (vgl. [54, S. 2]). Dabei werden die jeweils links aufgeführten Adjektive der ersten sechs Skalen der „Anstrengung“ und die rechts dargestellten Adjektive der „Entspannung“ zugeordnet (vgl. [54, S. 2]). Der vierte Faktor des für diese Forschungsarbeit verwendeten Semantischen Differentials bezieht sich auf das Ausmaß

des Lombard-Effekts in den Stimuli. Dabei wird dem Umstand Rechnung getragen, dass Sprache mit der Steigerung des Umgebungslärms lauter und verständlicher wird (vgl. Kapitel 3). Die Veränderung der Lautstärke ist in Kapitel 6.3.1 auch für die im Labor aufgenommenen Sprachsignale des SPID-Korpus nachgewiesen worden. Die Lombard-Parameter der Intensität und Verständlichkeit werden als Skalen des Semantischen Differentials verwendet, weil sie für ungeschulte Hörer*innen leichter zu identifizieren sind als z. B. die Grundfrequenz.

7.2 Hypothesen

Das vorliegende Perzeptionsexperiment dient als Gegenprobe zu den vorangegangenen akustisch-phonetischen Untersuchungen der Sprachsignale des SPID-Korpus. In Kapitel 6.3 wurde festgestellt, welche lautsprachlichen Veränderungen sich im Auto für die jeweilige Sprecherin bzw. den jeweiligen Sprecher eines Dialogpaares ergeben. Die vorliegende Forschungsarbeit untersucht den Einfluss von Fahrgeräuschen und ICC-System auf die Kommunikationssituation im Auto. Da Kommunikation jedoch nicht nur aus der Sprecher*innenebene besteht, sondern auch die Hörer*innenebene umfasst, muss auch aus Sicht der Hörer*innen festgestellt werden, ob bei lauten Umgebungsgeräuschen im Auto ein Lombard-Effekt wahrgenommen wird und ob durch die Verwendung des ICC-Systems eine Verbesserung der Kommunikation entsteht. Dabei wird zum einen überprüft, ob der in Kapitel 6.3 festgestellte Lombard-Effekt bzgl. der Lautstärke von den Zuhörer*innen der Perzeptionsstudie wahrgenommen wird und ob eine gesteigerte Verständlichkeit als zusätzlicher Lombard-Parameter in den Sprachsignalen festgestellt werden kann. Zum anderen wird durch die zusätzlichen Skalen des Semantischen Differentials untersucht, ob die Zuhörer*innen weitere Lombard-Eigenschaften in den Sprachsignalen feststellen, die durch die akustische Analyse nicht erkannt werden konnten. (vgl. Kapitel 1.2)

Das Semantische Differential ist ein für die Untersuchung von emotionalen Bedeutungen entwickelter Test (vgl. Kapitel 7.1). Die Verknüpfung des Faktors „Lombard“ mit den drei klassischen Faktoren dieser Methode dient der Aufdeckung von Korrelationen zwischen den Ausmaßen des Lombard-Effekts und emotionalen Assoziationen der Proband*innen. Hierdurch soll über die Ergebnisse der vorangegangenen Sprachsignalanalyse hinaus eine umfassendere Beantwortung der Frage ermöglicht werden,

ob sich die lautsprachliche Kommunikation im Fahrzeug bei lauten Hintergrundgeräuschen durch Anwendung des ICC-Systems verbessern lässt. Es soll überprüft werden, mit welchen konnotativen Bedeutungskomponenten die in Lombard-Sprache stattfindenden Modifikationen von Hörer*innen assoziiert werden.

Bei lauten Hintergrundgeräuschen wird die Sprachkommunikation für die Sprecher*innen anstrengender (vgl. [43, S. 117f.]). Vermutlich wird dies auch durch unabhängige Zuhörer*innen wahrgenommen und die entsprechenden Sprachsignale klingen emotionaler, weil die erhöhte Anstrengung der Sprecherin bzw. dem Sprecher eine sachliche Auseinandersetzung mit Dialogpartner*innen erschwert. Weiterhin ist es wahrscheinlich, dass laute Hintergrundgeräusche kognitiv von der Sprachplanung ablenken und die Proband*innen dadurch stockender und auch steifer klingen. Zudem wird vermutet, dass die Belastung durch laute Hintergrundgeräusche und die dabei zu erfüllende Sprechaufgabe es den Proband*innen erschweren, sich darüber hinaus auf die Aufrechterhaltung einer freundlichen Sprechweise zu konzentrieren. Zusätzlich wird angenommen, dass sich die Dialogpartner*innen gehemmter anhören, weil sie sich der Veränderung ihrer Sprachproduktion zumindest teilweise bewusst sind und sie dies im Hinblick auf ihre Sprachplanung verunsichert. Für die Skalen des Faktors „Lombard“ wird angenommen, dass die Sprachsignale entsprechend der Erkenntnisse aus der Literatur und den vorangegangenen akustisch-phonetischen Analysen (vgl. Kapitel 3.1 und 6.3) mit zunehmender Geräuschbelastung lauter und verständlicher klingen, während dieser wahrgenommene Lombard-Effekt mit Anwendung des ICC-Systems abgeschwächt wird. Die hier aufgestellten Hypothesen werden folgendermaßen zusammengefasst:

Hypothesen zu Teilfrage (3)

9. Die in der akustischen Analyse festgestellten Veränderungen der Intensität bei unterschiedlichen Fahrgeräusch- und ICC-Bedingungen werden von unabhängigen Zuhörer*innen wahrgenommen.
10. Bei zunehmenden Fahrgeräuschen klingt die Sprache der Passagier*innen verständlicher. Unter Verwendung eines ICC-Systems im Auto wird dieser Effekt reduziert oder aufgehoben.

11. Bei zunehmenden Fahrgeräuschen klingt die Sprache der Passagier*innen unfreundlicher. Unter Verwendung eines ICC-Systems im Auto wird dieser Effekt reduziert oder aufgehoben.
12. Bei zunehmenden Fahrgeräuschen klingt die Sprache der Passagier*innen emotionaler. Unter Verwendung eines ICC-Systems im Auto wird dieser Effekt reduziert oder aufgehoben.
13. Bei zunehmenden Fahrgeräuschen klingt die Sprache der Passagier*innen gehemmter. Unter Verwendung eines ICC-Systems im Auto wird dieser Effekt reduziert oder aufgehoben.
14. Bei zunehmenden Fahrgeräuschen klingt die Sprache der Passagier*innen steifer. Unter Verwendung eines ICC-Systems im Auto wird dieser Effekt reduziert oder aufgehoben.
15. Bei zunehmenden Fahrgeräuschen klingt die Sprache der Passagier*innen angestrengter. Unter Verwendung eines ICC-Systems im Auto wird dieser Effekt reduziert oder aufgehoben.
16. Bei zunehmenden Fahrgeräuschen klingt die Sprache der Passagier*innen stockender. Unter Verwendung eines ICC-Systems im Auto wird dieser Effekt reduziert oder aufgehoben.
17. Die Skalen der Faktoren *evaluation*, *potency* und *activity* repräsentieren das kommunikative Konzept *Anstrengung - Entspannung*.

7.3 Die Stimuli

Die Stimuli des Perzeptionsexperiments wurden aus den Sprachsignalen der Sprecher*innen 2d (*weiblich, hinten, Giver*), 3f (*männlich, hinten, Follower*) und 8o (*männlich, vorne, Follower*) zusammengestellt. Sie sind jeweils zwischen 6 und 6,5 Sekunden lang und ähneln sich inhaltlich dahingehend, dass die jeweils sprechende Person die Problematik um das Abholen von Herrn Schulz bzw. Herrn Scholz in den Karten des *Map-Task* Szenarios diskutiert. Weiterhin sind die Werte der durchschnittlichen Intensität und der durchschnittlichen Grundfrequenz in den Sprachausschnitten repräsentativ, da sie den Ergebnissen der Sprachsignalanalyse in Kapitel 6.3.1 und 6.3.2 entsprechen. Das bedeutet, dass die durchschnittliche Intensität und die durchschnittliche Grundfrequenz bei erhöhtem Geräuschniveau ohne und mit ICC-System ansteigen, wobei die Werte mit ICC jeweils niedriger sind als ohne ICC. Die einzige Ausnahme hierbei bildet der vorne sitzende Sprecher 8o, bei dem die Werte in der stillen Referenzbedingung mit ICC etwas höher sind als ohne ICC. Die Messung der Grundfrequenz erfolgte wiederum anhand der in Kapitel 6.3.2 beschriebenen Tonhöhenbereiche. Eine Auflistung der Messwerte für die vorliegenden Stimuli ist in Tabelle 7.2 abgebildet. Da für das Perzeptionsexperiment lediglich Ausschnitte des aufgenommenen Sprachmaterials verwendet wurden, welche für einzelne Sprecher*innen dieselbe Dauer besitzen, sind die Messwerte der Parameter Signaldauer und Sprachproduktionsleistung für die ausgewählten Stimuli nicht relevant. Zur Beurteilung im Perzeptionsexperiment wurden lediglich die Stimuli verwendet, die zuvor durch alle vier Methoden der Geräuschbefreiung prozessiert worden waren (vgl. Kapitel 4.1).

Tabelle 7.2: Die durchschnittliche Intensität und die durchschnittliche Grundfrequenz der Stimuli des Perzeptionsexperiments.

Stimuli			∅ Grundfrequenz (Hz)	∅ Intensität (dB)
Sprecher*in	ICC	Geräusch		
2d	ohne	Stille	233,70	52,96
2d	ohne	50	244,09	53,37
2d	ohne	130	264,09	58,18
2d	mit	Stille	214,88	47,14
2d	mit	50	237,80	51,24
2d	mit	130	259,45	53,12
3f	ohne	Stille	116,28	50,99
3f	ohne	50	145,35	57,38
3f	ohne	130	158,66	62,37
3f	mit	Stille	108,75	45,91
3f	mit	50	141,66	52,29
3f	mit	130	156,52	56,00
8o	ohne	Stille	118,25	46,99
8o	ohne	50	132,08	53,76
8o	ohne	130	175,50	60,00
8o	mit	Stille	120,64	49,00
8o	mit	50	127,86	49,35
8o	mit	130	137,76	51,64

7.4 Der *Online*-Fragebogen

Die subjektiven Bewertungen der lautsprachlichen Stimuli anhand des Semantischen Differentials wurden mit Hilfe eines *Online*-Fragebogens durchgeführt, welcher mit der *Software SoSci Survey* erstellt worden ist (vgl. [95]). Einzelne Seiten dieses *Online*-Fragebogens sind im Anhang dieser Arbeit aufgeführt (vgl. Anhang A.4). Auf der ersten Seite des Fragebogens erfolgten zunächst eine Begrüßung der Teilnehmer*innen, eine kurze Erläuterung der Vorgehensweise während der Befragung und das Abspielen eines Teststimulus durch Klick auf den *play button*. Diese Beispiel-datei beinhaltete den Ausschnitt einer Sprachaufnahme von Sprecherin 4g (*weiblich, vorne, Follower*), welche ohne ICC-System in der stillen Referenzbedingung

7 Perzeptionsexperiment

im Labor aufgenommen worden ist. Weiterhin wurden die Proband*innen gebeten, während der Umfrage nicht zwischen Lautsprechern und Kopfhörern zu wechseln und die Lautstärke nicht zu verändern, um eine gute Vergleichbarkeit der einzelnen Sprachausschnitte zu gewährleisten. Durch die Anpassung der Lautstärke anhand des Teststimulus, einer im Stillen aufgenommenen Sprachproduktion, sollte verhindert werden, dass nachfolgende Stimuli zu leise abgespielt und dadurch zu unverständlich werden, um sie mit den anderen Stimuli vergleichen zu können. Das als Beispiel aufgeführte Sprachsignal hat ebenfalls den Wohnort von Herrn Schulz bzw. Herrn Scholz zum Inhalt (vgl. Kapitel 7.3). Die Teilnehmer*innen wurden darüber hinaus gebeten, nur dann an der Umfrage teilzunehmen, wenn sie Muttersprachler*innen des Deutschen sind, damit Variationen in den Skalenbewertungen nicht auf ein ungenügendes Verständnis der deutschen Sprache zurückgeführt werden können. Weiterhin wurde den Proband*innen mitgeteilt, dass die Umfrage etwa 10 bis 15 Minuten in Anspruch nehmen würde. Durch Klicken auf *weiter* gelangten die Teilnehmer*innen zur nächsten Seite des Fragebogens, durch Klicken auf *zurück* kehrten sie zur vorangehenden Seite zurück. In den Instruktionen der ersten Fragebogenseite wurden diese beiden Funktionen erläutert, ohne darauf hinzuweisen, dass durch Klicken auf *zurück* frühere Bewertungen korrigiert werden könnten. Hierauf wurde bewusst verzichtet, um die Teilnehmer*innen nicht unter Druck zu setzen und ihnen eine unbefangene Beurteilung der Stimuli zu ermöglichen. In der Umfrage wurde zudem nicht nach dem musikalischen Interesse der Proband*innen gefragt, weil es hierbei nicht um eine reine Prosodiewahrnehmung geht.

Der Hauptteil des Fragebogens umfasste 18 einzelne Fragebogenseiten. Diese bestanden aus jeweils einem Stimuli, welcher durch Klicken eines oben abgebildeten *play buttons* wiederholt abgespielt werden konnte. Darunter waren die acht Skalen des Semantischen Differentials zur Bewertung dargestellt (vgl. Tabelle 7.1). Dabei wurden sowohl die Reihenfolge der einzelnen Skalen auf jeder Seite als auch die Abfolge der zu bewertenden Stimuli, d. h. die einzelnen Seiten des Fragebogens, zufällig rotiert. Um einer Bevorzugung von linken oder rechten Positionen in den Skalen vorzubeugen, wurden zwei annähernd identische Fragebögen erstellt. Dabei enthielt Fragebogen *SemDiff_NP* Skalen, deren Polarität denen in Tabelle 7.1 entspricht, während Fragebogen *SemDiff_PN* Skalen mit entgegengesetzter Polarität beinhaltete. Jede Probandin bzw. jeder Proband wurde automatisch und zufällig

einem der beiden Fragebögen zugeordnet, wobei diese Verteilung nicht gleichmäßig erfolgte.

Zur Beurteilung der Audiosignale wurden siebenstufige Skalen verwendet, wie es bei Osgood et al. (1975) empfohlen ist. Ihre Erkenntnisse beruhen auf Untersuchungen zur Verwendung verschiedener Skalen mit unterschiedlichen Intervallen durch Proband*innen. (vgl. [57, S. 81f., 85]) Auch in den Untersuchungen [9], [33] [54] und [35] wurden siebenstufige Skalen erfolgreich zur Durchführung der Methode des Semantischen Differentials angewendet, wobei der linke und rechte Rand der Skala *3-2-1* und *1-2-3* jeweils einer schwachen, mäßigen oder starken Zuordnung des jeweiligen Skalenpols zum gehörten Stimulus entspricht und die in der Mitte stehende *0* der Antwort *weder-noch* zugeordnet wird (vgl. [9, S. 1219], [33, S. 92f.], [54, S. 2] und [35, S. 92]). Beide Adjektive eines *items* stehen jeweils am Ende der entsprechenden Skala (vgl. [35, S. 92]). In Anhang A.4 wird beispielhaft eine Seite eines der beiden Fragebögen, welche zur Durchführung des Semantischen Differentials erstellt wurden, aufgeführt. Dabei handelt es sich um den Fragebogen, in dem die Polarität der Skalen der in Tabelle 7.1 entspricht.

Die vorletzte Seite des Fragebogens fragte nach Alter und Geschlecht der Teilnehmenden. Anschließend konnten die Proband*innen ankreuzen, wie einfach oder schwer die Umfrage für sie war und eine Bewertung der einzelnen Skalen durchführen. Ein Feld zur freien Texteingabe bot zudem die Möglichkeit einer individuellen Beurteilung. Am Ende der Befragung erfolgte nochmals ein kurzer Dankestext und die Bitte, den Fragebogen an andere potentielle Teilnehmer*innen weiterzuleiten.

7.5 Durchführung

Vor dem Start der Umfrage ist durch die Experimentleiterin und zwei weitere Personen mit Hilfe der *Software Sosci Survey* (vgl. [95]) mehrmals ein *Pretest* durchgeführt worden. Nach jedem *Pretest* wurden der Fragebogen und dessen Konditionen modifiziert und schließlich der Befragungszeitraum vom 22.11.2016 bis zum 11.01.2017 festgelegt, innerhalb dessen die Umfrage für potentielle Teilnehmer*innen freigeschaltet wurde. Die Proband*innen der Studie nahmen anhand eines Internet-*Links* an der Befragung teil (<https://www.soscisurvey.de/sprachsignale/>). Sie wurden gebeten, den *Online*-Fragebogen an andere Personen weiterzuleiten, sodass eine mög-

lichst hohe Anzahl an Teilnehmer*innen und eine absolute Zufallsstichprobe zustande kommen konnte. Als Grundgesamtheit wurde bei dieser Erhebung die Menge aller deutschen Muttersprachler*innen angenommen, welche nicht an den Aufnahmen des SPID-Korpus beteiligt waren und Internetzugang besaßen.

Der Nachteil an der Teilnahme per Internet-*Link* ist, dass nicht kontrolliert wird, ob Proband*innen die Umfrage mehrfach durchführen. Dies könnte z. B. mit Serien*mails* oder Passwörtern verhindert werden. Der Vorteil ist jedoch, dass durch die einfache Verbreitung eines *Links* viele Internetnutzer*innen den Fragebogen öffnen können und die Hemmschwelle zur Teilnahme vermutlich geringer ist, wenn die Proband*innen z. B. ihre *Mail*adresse nicht angeben müssen. Der *Link* zum Fragebogen wurde u. a. über persönliche und berufliche Kontakte, soziale Netzwerke im Internet sowie die Internetplattform *SurveyCircle* (vgl. [96]) verbreitet. Insgesamt nahmen 132 Personen am Perzeptionsexperiment teil, die alle in der vorliegenden Analyse der Erhebung berücksichtigt wurden.

7.6 Ergebnisanalyse

Für die Auswertung der erhobenen Daten werden lediglich diejenigen Fragebögen berücksichtigt, welche bis zur letzten Seite ausgefüllt wurden. D. h., abgebrochene Umfragen werden bereits beim Herunterladen der Daten ignoriert (vgl. [95]). Ebenso werden beim Herunterladen Daten aussortiert, welche sich aus der Fragebogenentwicklung oder den *Pretests* ergeben haben. Die Rückgabewerte in den von *SoSci Survey* heruntergeladenen Daten entsprechen nicht den auf den Fragebogenen dargestellten Skalierungen des Semantischen Differentials. In der *Software* werden nur positive Werte verwendet, um die Antwort einer Teilnehmerin bzw. eines Teilnehmers zu speichern (vgl. [95]). Die Kodierung der Bewertungen erfolgt daher nach dem in Tabelle 7.3 dargestellten Muster:

Tabelle 7.3: Die vorläufige Kodierung der Skalenbewertungen im Semantischen Differential des Perzeptionsexperiments am Beispiel der Skala „laut-leise“.

Skalierung während der Umfrage

<i>laut</i>	-3	-2	-1	0	1	2	3	<i>leise</i>
-------------	----	----	----	---	---	---	---	--------------

Kodierung während der Analyse

<i>laut</i>	1	2	3	4	5	6	7	<i>leise</i>
-------------	---	---	---	---	---	---	---	--------------

Bei der in Tabelle 7.3 unten dargestellten Kodierung ist die Polarität der Skalen jedoch kontraintuitiv, da der niedrigste Wert die höchste Lautstärke beschreibt. Aus diesem Grund werden in einem R-Skript zur Datenaufbereitung alle Rückgabewerte von der Zahl 8 subtrahiert, um die Polarität der Skalen für die nachfolgende Analyse umzukehren. Die Kodierung der Skalierungen des vorliegenden Semantischen Differentials ist in Tabelle 7.4 dargestellt.

Tabelle 7.4: Die finale Kodierung der Skalenbewertungen im Semantischen Differential des Perzeptionsexperiments.

Skalierung während der Umfrage

<i>X</i>	-3	-2	-1	0	1	2	3	<i>Y</i>
----------	----	----	----	---	---	---	---	----------

Kodierung während der Analyse

<i>freundlich</i>	1	2	3	4	5	6	7	<i>unfreundlich</i>
<i>sachlich</i>	1	2	3	4	5	6	7	<i>emotional</i>
<i>ungehemmt</i>	1	2	3	4	5	6	7	<i>gehemmt</i>
<i>locker</i>	1	2	3	4	5	6	7	<i>steif</i>
<i>entspannt</i>	1	2	3	4	5	6	7	<i>angestrengt</i>
<i>flüssig</i>	1	2	3	4	5	6	7	<i>stockend</i>
<i>leise</i>	1	2	3	4	5	6	7	<i>laut</i>
<i>unverständlich</i>	1	2	3	4	5	6	7	<i>verständlich</i>

Im Folgenden werden zunächst die von den Teilnehmer*innen am Ende der Umfrage in das offene Textfeld eingetragenen Kommentare thematisch zusammengefasst und diskutiert. Anschließend erfolgt die statistische Untersuchung der Skalenbewertungen. Wie in Kapitel 6 wird die Umformung der Daten als Vorbereitung der Ergeb-

nisuntersuchung sowie die deskriptive und inferenzstatistische Analyse wiederum in R durchgeführt (vgl. [94]).

7.6.1 **Bewertungen der Umfrage**

Von den 132 Proband*innen des Perzeptionsexperiments haben insgesamt 67 Personen die offene Texteingabe genutzt, um die Fragebogen-Erhebung mit eigenen Worten zu bewerten. Ihre Kommentare beziehen sich neben den einzelnen Skalen des Semantischen Differentials auf die Audiosignale, die Instruktionen und weitere Eigenschaften der Umfrage. Die Hauptaussagen dieser Evaluation sind in Tabelle 7.5 zusammengefasst und thematisch sortiert. Aufgrund von inhaltlichen Wiederholungen sowie syntaktischer und kontextueller Unklarheiten in den Texteingaben der Teilnehmer*innen handelt es sich hierbei größtenteils um eine durch die Experimentleiterin vorgenommene Paraphrasierung und nicht um die Originalkommentare der Proband*innen.

Tabelle 7.5: Die Bewertung des *Online*-Fragebogens durch die Proband*innen.

Instruktionen

1.	Das Konzept der Erhebung ist nicht klar. Unklare Zielsetzung.
2.	Definitionen bzw. Begriffserklärungen fehlen.
3.	Die Einstufungen der Stimuli anhand der Skalen sind schwierig, da die Teilnehmer*innen auf keine Erfahrungen im Umgang mit diesen Skalen zurückgreifen können. Dies gilt besonders für den Beginn der Befragung. Hier hätte ein Beispiel am Anfang mit Referenzwerten eventuell helfen können.

Anonymität

4.	Informationen über die Hintergründe der Sprecher*innen, ihre normale Sprechstimme und ihr Verhältnis zueinander fehlen.
5.	Es ist schwer, die Skalen zu bewerten, wenn ein reines Audiosignal vorliegt (fehlender Gesichtsausdruck bzw. Gestik und Mimik und fehlende Körperhaltung), z. B. bei „gehemmt“, „locker“, „angestrengt“ und „freundlich“.

Audiosignale

6.	Die Dauer der Sprachsignale ist zu kurz für eine sichere Bewertung (z. B. bei „freundlich“, „emotional“, „gehemmt“, „locker“ und „angestrengt“).
7.	Die Sprachsignale sind sich sehr ähnlich bzw. es gibt nur geringe Unterschiede zwischen ihnen.
8.	Die Adjektive der Skalen sind z. T. nicht für die Bewertung der Sprachausschnitte geeignet, da sie nur schwer aus dem gegebenen Sprachmaterial herausgehört werden können (z. B. bei „gehemmt“, „freundlich“, „emotional“ oder „locker“). Der sachliche Inhalt des Gesagten erschwert die Bewertung einiger Skalen (z. B. bei „angestrengt“ und „emotional“).
9.	Die Einschätzung von „emotional“ wird für männliche Sprecher z. T. als schwer empfunden.
10.	Die Sprachsignale brechen z. T. ab bzw. Sätze werden z. T. nicht beendet. Dazu gibt es unverständlich Gesprochenes vor einem Satz.
11.	Bei der Evaluierung wird z. T. ein Abgleich mit vorherigen Bewertungen vorgenommen bzw. die Bewertung der Sprachausschnitte nicht in Relation zu vorherigen Signalen vorzunehmen, ist schwierig.
12.	Es gibt Umgebungsgeräusche, Wiederholungen der Inhalte und Stimmlagenwiederholungen.
13.	Es gibt z. T. auch innerhalb eines Signals deutliche und undeutliche Abschnitte, was die Bewertungen des Stimulus als Ganzes erschwert.

Skalen - Allgemein

14.	Die Kategorien „verständlich“ und „laut“ sind selbsterklärend.
15.	Die Adjektive der Skalen beziehen sich teilweise auf dasselbe semantische Konzept oder sind mehrdeutig.
16.	Die einzelnen Stufen der Skalen sind nicht ausreichend selbsterklärend. Dies gilt z. B. für den Wert 0 der Skalen „verständlich-unverständlich“ oder auch „emotional-sachlich“, da „sachlich“ z. T. bereits mit „neutral“ assoziiert wird und die Bedeutung des Wertes 0 hierbei nicht klar ist. Die Skalierung ist schwierig. Die Proband*innen neigen z. T. dazu, nie einen Extremwert zu nutzen.
17.	Die Skalen scheinen teilweise zu korrelieren. Einige sind schwer voneinander abzugrenzen: <ul style="list-style-type: none"> ● „emotional-angestrengt“ ● „emotional-gehemmt“ ● „emotional-laut“ ● „angestrengt-laut“ ● „angestrengt-unfreundlich“ ● „angestrengt-steif“ ● „gehemmt-steif“

Skalen - Details

„emotional“	<p>18. Das Gesagte klingt eher sachlich, weil dessen Inhalt auf Fakten beruht.</p> <p>19. Die Bewertung ist schwierig, wenn der Inhalt sachlich ist, aber die Stimmlage emotional.</p> <p>20. Es fehlt eine allgemein gültige Definition der Begrifflichkeiten.</p>
„gehemmt“	<p>21. Auch hier ist die Bewertung eines sachlichen Inhalts mit der gegebenen Skala schwer.</p> <p>22. Es ist z. T. unklar, was unter gehemmtem bzw. ungehemmtem Sprechen verstanden werden soll; es wird z. T. eine Ähnlichkeit zu „stockend“ vermutet.</p>
„steif“	<p>23. Es ist z. T. unklar, was unter steifem Sprechen verstanden werden soll.</p>
„flüssig“	<p>24. Es ist z. T. unklar, was unter stockendem bzw. flüssigem Sprechen verstanden werden soll; es wird z. T. als „langsam-schnell“ definiert und ein Zusammenhang zu „gehemmt-ungehemmt“ hergestellt.</p>
„verständlich“	<p>25. Es ist unklar, ob Störungen bzw. eine veränderte <i>Soundqualität</i> in den Aufnahmen bewertet werden sollen, selbst wenn das eigentliche Sprachsignal gut zu verstehen ist.</p> <p>26. Es stellt sich die Frage, ob der Inhalt oder die Artikulation bzw. Lautstärke des Gesagten bewertet werden soll.</p>
„laut“	<p>27. Frauen sind leiser als Männer.</p> <p>28. Alle Sprachsignale klingen ähnlich laut.</p> <p>29. Für einen gelungenen Vergleich der Lautstärke müssen zunächst mehrere Sprachsignale angehört werden.</p> <p>30. Diese Skala wird von Tonhöhe und Gehemtheit der Sprecherin bzw. des Sprechers beeinflusst, wodurch die Bewertung erschwert wird.</p> <p>31. Es ist nicht klar, ob eine tatsächliche Zunahme oder Abnahme der Signallautstärke gegenüber anderen Signalen bewertet werden soll oder eine als hoch oder niedrig wahrgenommene Sprechlautstärke, auch wenn das Signal dabei nicht lauter oder leiser ist als andere Signale.</p>

Hinsichtlich der Punkte (1), (2), (4) und (5) in Tabelle 7.5 ist zu sagen, dass die vorliegende Forschungsarbeit und ihre Ziele zu Beginn der Befragung absichtlich nicht beschrieben worden sind, um die Teilnehmer*innen in ihrer Bewertung nicht zu beeinflussen. Der Begriff der „normalen Sprechstimme“ bezieht sich wahrscheinlich auf eine stille Kommunikationssituation, in der die Gesprächspartner*innen Blickkontakt halten können. Diese Situation, das Verhältnis der Sprecher*innen zueinander, ihre persönlichen Hintergründe, Gestik, Mimik und Körperhaltung sind nicht Teil des

Forschungsobjekts und mussten für den Vergleich der Sprachsignale in den unterschiedlichen ICC- und Geräuschbedingungen nicht bekannt sein.

Die Punkte (8), (9), (14), (15) und (17) sind subjektive Einzelbewertungen der Proband*innen hinsichtlich der Eignung verschiedener Adjektive zur Bewertung der Stimuli und der Auswahl der Sprachausschnitte. Dasselbe gilt für die unter der Überschrift „Skalen - Details“ in Tabelle 7.5 zusammengefassten Kommentare. Diese zeigen, dass unter den Teilnehmer*innen z. T. Unsicherheiten bzw. Unklarheiten im Umgang mit den Adjektiven der Umfrage herrschten. Die Experimentleitung geht jedoch davon aus, dass verschiedene Hörer*innen stets leicht unterschiedliche Interpretationen der skalierten Adjektive vornehmen und dass sich dies auch durch eine vorangehende Begriffserklärung niemals vollständig wird aufheben lassen. Anfangsschwierigkeiten im Umgang mit den Skalenbewertungen (Punkt (3)) wird weiterhin durch die Randomisierung der Skalen, der Skalenpolarität und der einzelnen Fragebogenseiten entgegengewirkt (vgl. Kapitel 7.4). Auch der Umstand, dass die Evaluierung eines Stimulus mit vorangegangenen Evaluierungen anderer Stimuli abgeglichen wird (Punkt (11)), ist aufgrund dieser Randomisierungen kein Einflussfaktor für die Bewertungen der Sprachausschnitte.

Die kurze Dauer der Sprachsignale (Punkt (6)) ist dem Umstand geschuldet, dass die Umfrage nicht zu lang sein sollte (vgl. Kapitel 7.1). Die inhaltliche und akustische Ähnlichkeit der Stimuli (Punkt (7) und (12)) verhindert, dass weitere Faktoren außer den ICC- und Geräuschbedingungen des SPID-Korpus einen Einfluss auf die Bewertungen der Proband*innen nehmen. Dennoch ließen sich gewisse Abweichungen in manchen Sprachausschnitten, wie nicht beendete Sätze oder undeutlich gesprochene Abschnitte (Punkt (10) und (13)), nicht vollständig vermeiden. Die Wahrnehmung von Umgebungsgeräuschen (Punkt (12)) bezieht sich wahrscheinlich auf Schmatz- oder Atemgeräusche der Sprecher*innen, da die Stimuli zuvor von jeglichen Hintergrundgeräuschen befreit worden sind (vgl. Kapitel 7.3).

Bei Punkt (16) wird kritisiert, dass die Skalenwerte nicht ausreichend erklärt wurden. Kohler (2005) weist darauf hin, dass die Bewertung eines Stimulus mit dem Skalenpunkt 0 schwierig ist, wenn ein Skalenpol anhand der Verneinung des gegenüberliegenden Pols benannt wurde (vgl. [33, S. 93]). In der vorliegenden Studie betrifft dies die Skalen „unfreundlich-freundlich“, „gehemmt-ungehemmt“ und „verständlich-unverständlich“. Dabei wird in allen Skalen für den Wert 0 die Zu-

7 Perzeptionsexperiment

ordnung des Kommentars *weder-noch* angenommen, was einer Enthaltung bei der Bewertung entspricht. Um die Beeinflussung der Teilnehmer*innen im Vorfeld der Umfrage möglichst gering zu halten, wurden, wie oben beschrieben, nur wenige Angaben zu Zweck und Inhalt der Befragung gemacht. Aufgrund der Unsicherheiten der Proband*innen sollten die Stufen der Skalen in nachfolgenden Untersuchungen jedoch zu Beginn der Erhebung thematisiert werden.

Probandin Nr. 693 gibt an, Muttersprachlerin des österreichischen Deutschen zu sein. Da während der Umfrage jedoch nicht nach Akzenten oder Dialekten des Deutschen gefragt wurde und dementsprechend keine dazugehörigen Daten anderer Proband*innen vorliegen, werden die Daten dieser Teilnehmerin mit in die Analyse der Umfrage einbezogen. Zudem liegt trotz eventuellen Akzenten oder Dialekten der Hörer*innen ein hinreichendes Verständnis des Standarddeutschen vor. Probandin Nr. 229 schreibt in der offenen Texteingabe den folgenden Kommentar:

„Man hat keine Ahnung, ab wann etwas als gehemmt/steif/emotional gilt und klickt halt irgendetwas an. Mit den Kategorien verständlich und laut kann hingegen vermutlich jeder etwas anfangen.“

Obwohl diese Teilnehmerin zugibt, bei der Bewertung der Stimuli teilweise willkürlich gehandelt zu haben, werden ihre Evaluationen nicht aus der Analyse des Semantischen Differentials ausgeschlossen. Zum einen ist damit zu rechnen, dass auch andere Proband*innen mitunter willkürlich geantwortet haben, zum anderen ist es darüber hinaus möglich, dass es sich bei diesem Kommentar um eine Übertreibung handelt, welche einem vermeintlichen Problem in der Erhebungsmethode besonderes Gewicht verleihen soll. Des Weiteren besteht die Möglichkeit, dass die Proband*innen sich selbst unterschätzen und systematischer urteilen, als es ihnen bewusst ist.

7.6.2 Deskriptive Beobachtungen

Die Abbildungen 7.1 und 7.2 bilden die arithmetischen Mittel der Skalenbewertungen ab, welche sich aus dem Perzeptionsexperiment für jede der acht Skalen des Semantischen Differentials ergeben. Zunächst werden die Skalen betrachtet, welche sich auf das Ausmaß des Lombard-Effekts im Sinne einer Modifikation von Lautstärke und Verständlichkeit der Sprachsignale beziehen. Anschließend werden die

deskriptiven Ergebnisse der weiteren sechs Skalen des Semantischen Differentials diskutiert.

Hinsichtlich der Skala „laut“ lässt sich feststellen, dass die Stimuli aller drei Sprecher*innen mit zunehmender Geräuschbelastung als lauter empfunden werden, sowohl ohne als auch mit ICC-System (vgl. Abbildung 7.2). Mit ICC-System wirken die Sprachsignale der einzelnen Experimentbedingungen etwas leiser auf die Proband*innen. Der Effekt der ansteigenden Hintergrundgeräusche auf die wahrgenommene Lautstärke der Stimuli scheint hierdurch jedoch kaum beeinflusst zu werden. Insgesamt machen die Stimuli einen eher lauten Eindruck auf die Teilnehmer*innen der Umfrage. Weiterhin werden Sprecher 3f und 8o ohne ICC von Stille zu 130 km/h verständlicher, während Sprecherin 2d diesen Trend nicht aufweist (vgl. Abbildung 7.2). Mit aktiviertem ICC-System lässt sich feststellen, dass Sprecher*in 2d und 3f mit steigender Geräuschbelastung immer verständlicher werden. Dies ist bei Sprecher 8o nicht der Fall. Im Hinblick auf die arithmetischen Mittel der Skala kann beobachtet werden, dass diese einen Wert von 4,0 nicht unterschreiten und somit alle Sprachsignale tendenziell als verständlich bzw. als nicht unverständlich eingestuft werden.

Ohne ICC-System hören sich alle drei Sprecher*innen mit zunehmender Geräuschbelastung weniger sachlich bzw. emotionaler sowie weniger gehemmt an (vgl. Abbildung 7.1). Mit ICC lassen sich hierbei keine Systematiken mehr erkennen, außer dass Sprecherin 2d emotional klingt, wohingegen Sprecher 3f und 8o sachlich wirken. Den Mittelwerten in Abbildung 7.1 zufolge werden die Sprecher*innen in allen Experimentbedingungen von den Hörer*innen als ungehemmt wahrgenommen. Die Skala „steif-locker“ weist keine systematischen Effekte der ICC- und Geräuschbedingungen auf (vgl. Abbildung 7.1). Es ist jedoch zu erkennen, dass die Sprachsignale in den meisten Experimentbedingungen als locker klingend wahrgenommen werden. Die Ergebnisse der Skalen „unfreundlich-freundlich“ (vgl. Abbildung 7.1) und „angestrengt-entspannt“ (vgl. Abbildung 7.2) ähneln sich hinsichtlich ihrer Ergebnisse stark. Die Sprachsignale der Umfrage hören sich mit steigenden Umgebungsgläuschen sowohl ohne als auch mit ICC-System angestrenzter und unfreundlicher an. Zwei Ausnahmen bilden Sprecher 8o, welcher ohne ICC bei steigenden Geräuschen entspannter und freundlicher wirkt, sowie Sprecher 3f, der mit ICC-System kaum eine Systematik über die Geräuschveränderungen aufzeigt und von Stille zu 130

7 *Perzeptionsexperiment*

km/h etwas weniger entspannt klingt. Wie in Abbildung 7.2 zu erkennen ist, klingt Sprecherin 2d in allen Experimentbedingungen im Durchschnitt angestrengt. Sprecher 3f und 8o klingen hingegen in den meisten Experimentbedingungen entspannt. Die in Abbildung 7.1 dargestellten Mittelwerte zur wahrgenommenen Freundlichkeit lassen darauf schließen, dass kaum ein Sprachsignal des Perzeptionsexperiments als unfreundlich bewertet wird. Es finden hauptsächlich Unterscheidungen zwischen der Freundlichkeit einzelner Sprecher*innen in verschiedenen Geräuschbedingungen statt. Die Skala „stockend-flüssig“ weist keinerlei systematische Veränderungen über die Ausprägungen der ICC- und Geräuschbedingungen hinweg auf. Anhand der in Abbildung 7.2 dargestellten Mittelwerte kann lediglich festgestellt werden, dass ohne ICC-System keine Sprecherin bzw. kein Sprecher als stockend eingestuft wird und mit ICC-System nur wenige Sprecher*innen- und Geräuschkombinationen.

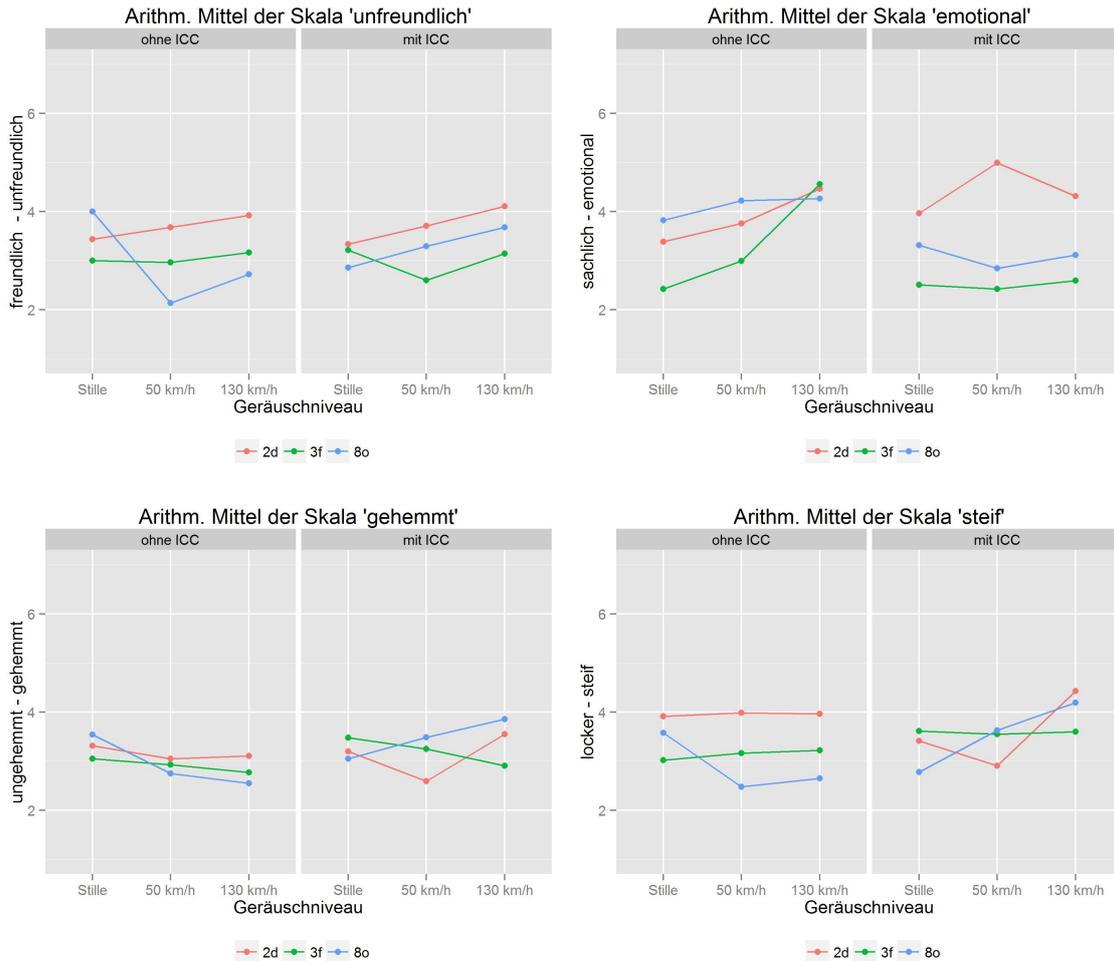


Abbildung 7.1: Arithmetische Mittel der Skalen „unfreundlich“, „emotional“, „gehemmt“ und „steif“ in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren die Bewertungen von Sprachsignalen der drei Sprecher*innen 2d, 3f und 8o des Produktionsexperiments.

Die Abbildungen 7.3 und 7.4 zeigen *Boxplots* über die Skalenbewertungen aller drei Sprecher*innen hinweg, in jeder ICC- und Geräuschbedingung. Dabei bestätigt sich für die Skala „laut“ das zuvor gewonnene Bild (vgl. Abbildung 7.4). Hierbei wird sowohl ohne als auch mit ICC-System ein Lombard-Effekt wahrgenommen, d. h., die Sprachsignale hören sich bei steigender Geräuschbelastung lauter an. Mit ICC-System klingen diese in jeder Geräuschbedingung im Mittel etwas leiser. Weiterhin lässt sich feststellen, dass die Varianz der Evaluierungen mit dem Auftreten der Fahrgeräusche abnimmt. Die Skala der Verständlichkeit zeigt weniger eindeutige Ergebnisse (vgl. Abbildung 7.4). Der Grund hierfür ist vermutlich die fehlende Differenzierung der Daten anhand der unterschiedlichen Sprecher*innenstimuli. Mit ICC-System weisen die Bewertungen der Proband*innen eine höhere Varianz auf,

7 Perceptionsexperiment

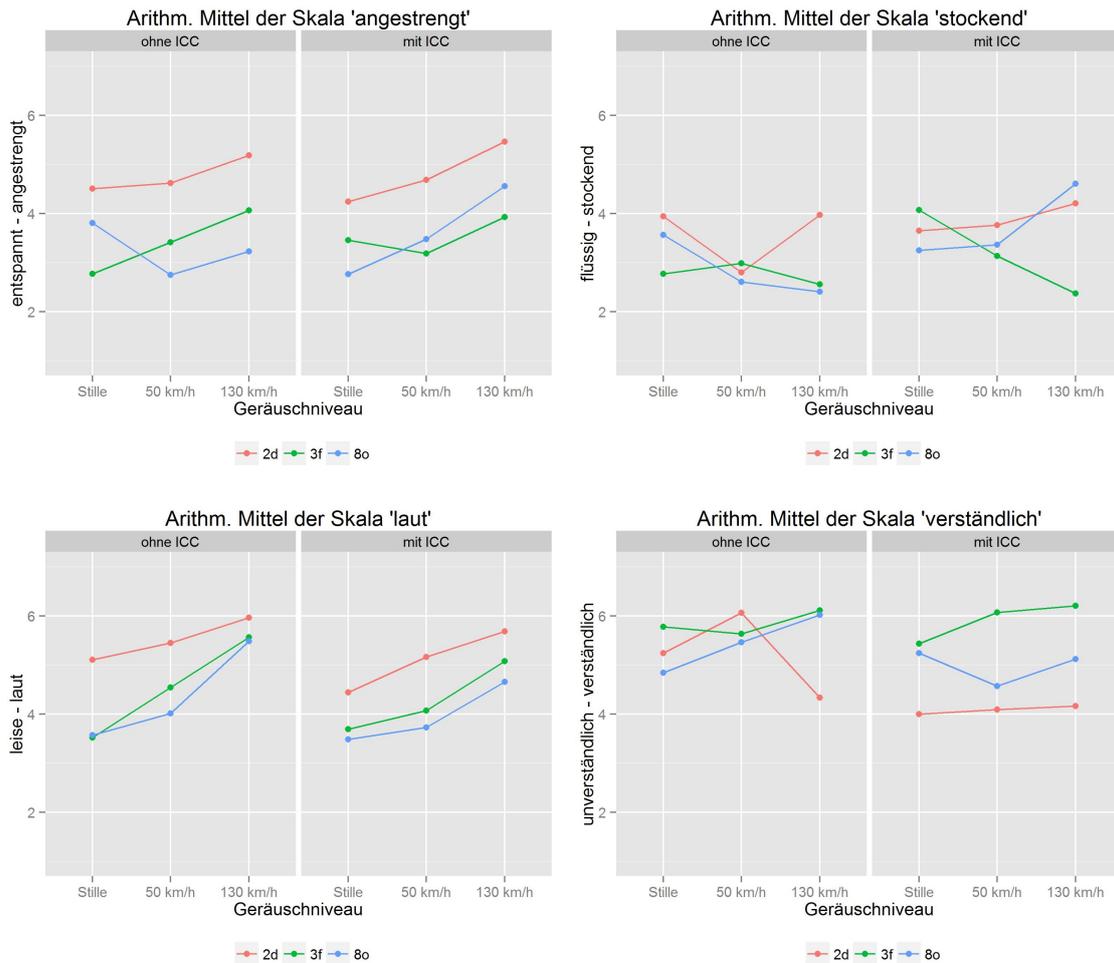


Abbildung 7.2: Arithmetische Mittel der Skalen „angestrengt“, „stockend“, „laut“ und „verständlich“ in jeder ICC- und Geräuschbedingung. Verschiedenfarbige Linien repräsentieren die Bewertungen von Sprachsignalen der drei Sprecher*innen 2d, 3f und 8o des Produktionsexperiments.

während die Sprachsignale der einzelnen Geräuschbedingungen im Mittel unverständlicher klingen.

Die Skalen „emotional“ und „gehemmt“ weisen in Abbildung 7.3 ähnliche Tendenzen auf wie in Abbildung 7.1. Dabei werden die Sprachsignale ohne ICC-System mit zunehmender Geräuschbelastung im Mittel als weniger sachlich bzw. emotionaler und als ungehemmter wahrgenommen. Die Varianz der Evaluierungen sinkt hin zu 130 km/h. Mit ICC-System können hierbei keine systematischen Effekte der unterschiedlichen Hintergrundgeräusche auf die Bewertungen dieser beiden Skalen mehr festgestellt werden. Die Skala „steif“ (vgl. Abbildung 7.3) lässt ohne ICC keine eindeutigen Effekte erkennen, während die Sprachsignale mit ICC-System bei zunehmenden Fahrgeräuschen weniger locker bzw. steifer klingen, wobei auch die Varianz

der Bewertungen abnimmt. Weiterhin bestätigen die vorliegenden *Boxplots*, dass die Stimuli mit ansteigender Geräuschbelastung hin zu 130 km/h ohne und mit ICC-System als angestrenzter bewertet werden, während die Varianz der Evaluierungen abnimmt (vgl. Abbildung 7.4). Die Skala „unfreundlich“ zeigt weniger eindeutige Effekte (vgl. Abbildung 7.3). Ohne ICC-System lassen sich keine systematischen Einflüsse der Fahrgeräusche auf die Wahrnehmung der Stimuli erkennen, während die Sprachsignale mit ICC-System bei abnehmender Varianz der Bewertungen hin zu 130 km/h tendenziell weniger freundlich bzw. unfreundlicher klingen. Die Skala „stockend“ weist in Abbildung 7.4 ebenfalls keine systematischen Effekte der unterschiedlichen Fahrgeräusche auf die Wahrnehmung der Stimuli auf. Die Varianz der Messwerte scheint in fast allen Geräuschbedingungen mit ICC-System anzusteigen.

7 Perzeptionsexperiment

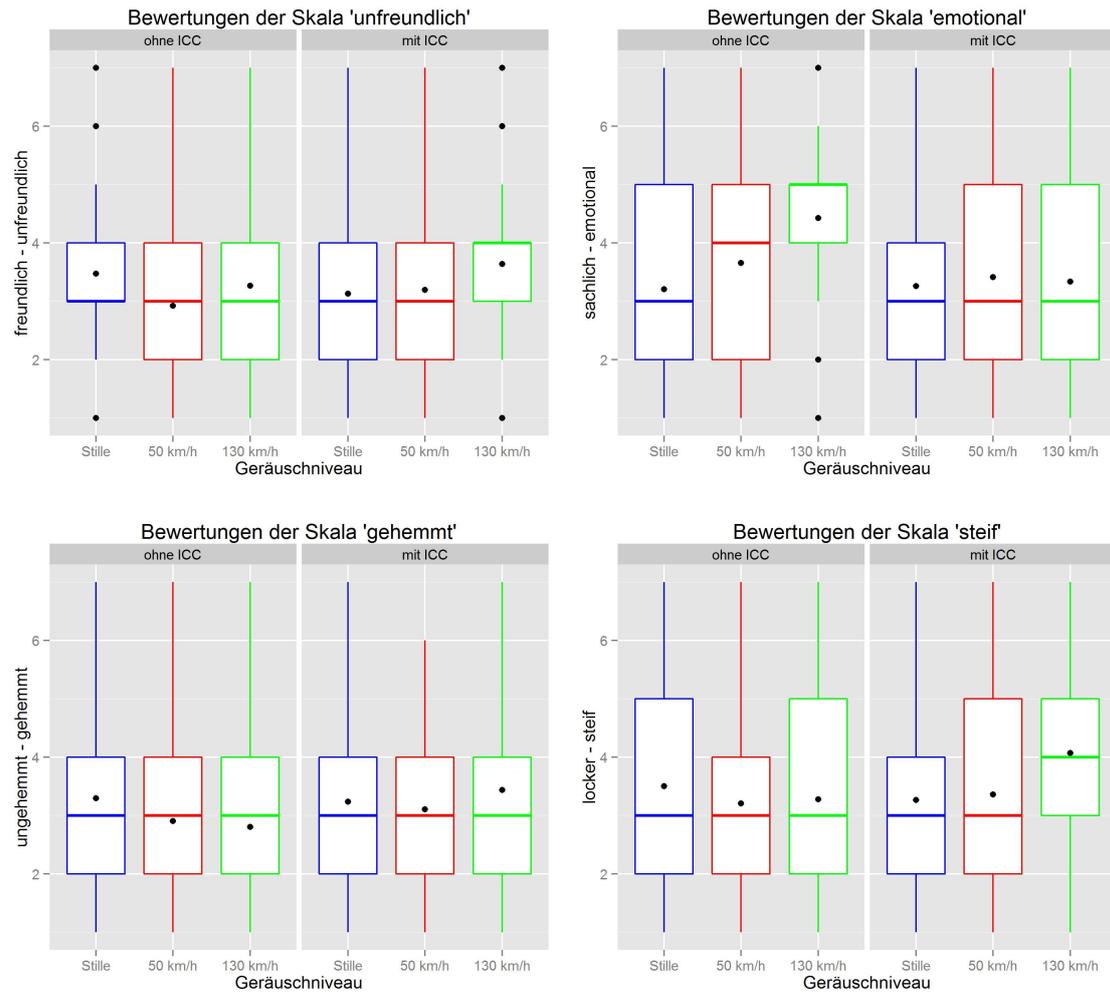


Abbildung 7.3: *Boxplots* der Skalenbewertungen für „unfreundlich“, „emotional“, „gehemmt“ und „steif“ in jeder ICC- und Geräuschbedingung.

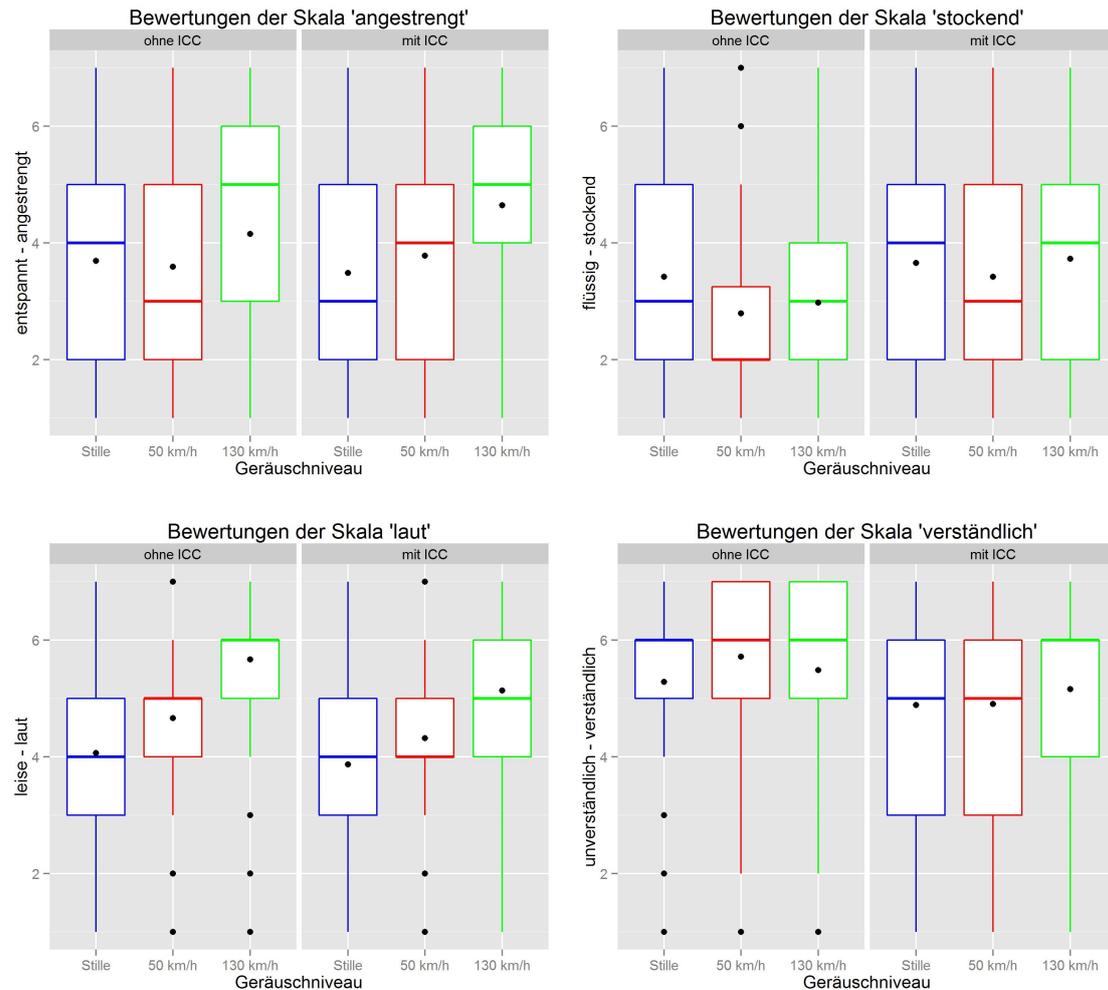


Abbildung 7.4: *Boxplots* der Skalenbewertungen für „angestrengt“, „stockend“, „laut“ und „verständlich“ in jeder ICC- und Geräuschbedingung.

7.6.3 Regressionsanalyse der Umfragedaten

Die inferenzstatistische Analyse der durch das Semantische Differential erhobenen Daten wird in R (vgl. [94]) anhand einer Regressionsanalyse durchgeführt, wie sie in Kapitel 6.3 beschrieben ist. Im Gegensatz zur vorangegangenen Untersuchung der Sprachsignale wird die Zeile *Intercept* im Folgenden jedoch nicht entfernt (vgl. Kapitel 6.3), da zur Untersuchung der Messwerte keine Mittelwertbereinigung der Daten notwendig ist, deren Mitte in jedem Fall 4 (*weder-noch*) beträgt (vgl. Kapitel 7.4 und Tabelle 7.4). Die Zeile *Intercept* beinhaltet dabei die Experimentbedingungen, deren Ausprägungen in den darunter liegenden Zeilen nicht aufgeführt sind, und liefert hierfür einen Schätzwert. Anders als in Kapitel 6.3 müssen die unter *Estimate*

7 Perzeptionsexperiment

aufgeführten Werte der nachfolgenden Zeilen mit dem Wert des *Intercept* addiert werden, um den Schätzwert der jeweiligen Variablenausprägung zu erhalten.

Aufgrund der deskriptiven Ergebnisse des Perzeptionsexperiments (vgl. Kapitel 7.6.2) wird angenommen, dass es z. T. individuelle Unterschiede in den Bewertungen der Stimuli von Sprecher*in 2d, 3f und 8o gibt. Die Abbildungen 7.1 und 7.2 zeigen auf, dass sich die unterschiedlichen Bewertungen der Sprecher*innenstimuli sowohl in den Effekten der ICC- und Geräuschbedingungen als auch in den Niveauunterschieden zwischen den Sprecher*innen widerspiegeln. Weiterhin sind manche der dabei aufgedeckten Effekte in den *Boxplots* der Abbildungen 7.3 und 7.4 nicht mehr feststellbar, wenn die Bewertungen über alle drei Sprecher*innen hinweg betrachtet werden.

Im Folgenden werden zunächst die Skalen analysiert, welche die Lombard-Eigenschaften der Lautstärke und der Verständlichkeit widerspiegeln. Anschließend erfolgt die Analyse der anderen sechs Skalen des Semantischen Differentials. Durch die Separierung der Daten anhand der einzelnen Sprecher*innen entstehen drei verschiedene Teildatensätze, in denen jeweils die Effekte der Ausprägungen von *ICC* und *Geräusch* sowie deren Interaktionen auf die acht Skalen des Semantischen Differentials untersucht werden. Die Ergebnisse der Regressionsanalyse werden dabei verkürzt dargestellt, wobei die für die Analyse wesentlichen Resultate zusammengefasst und in tabellarischer Form abgebildet werden.

laut/leise

Innerhalb der Regressionsanalyse wird entsprechend den Ausführungen in Kapitel 6.3 einem Nullmodell ohne erklärende Variablen vorerst lediglich eine unabhängige Variable hinzugefügt, bei der es sich um *ICC* handelt (vgl. Tabelle 7.6). Dabei wird deutlich, dass das ICC-System bei allen drei Sprecher*innen einen signifikanten Einfluss auf die Bewertungen der Stimuli ausübt (Sprecher*in 2d und 8o: $p < 0,0001$; Sprecher 3f: $p = 0,0058$). Die Bewertungen *mit ICC* werden durch das *Intercept* repräsentiert. Ohne ICC-System werden zu den Schätzwerten des *Intercept* je nach Sprecher*in Werte zwischen 0,26263 und 0,4116 hinzuaddiert. Die Bewertungen ohne ICC-System sind daher höher. D. h., die wahrgenommene Lautstärke der Sprachsignale nimmt mit Aktivierung des ICC-Systems um 0,4116 (2d), 0,26263 (3f) und 0,40152 (8o) Skalenpunkte ab.

Tabelle 7.6: Regressionsergebnisse der Skala „laut-leise“ für die abhängige Variable *evaluation* und die erklärende Variable *ICC*.

Call:

`lm(formula = evaluation ~ ICC, data = dataicc_laut)`

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	5.0934	0.0553	92.113	< 2e-16	***	0.03266
	ICCoohne ICC	0.4116	0.0782	5.264	1.82e-07	***	
3f	(Intercept)	4.27778	0.06713	63.724	<2e-16	***	0.00834
	ICCoohne ICC	0.26263	0.09494	2.766	0.0058	**	
8o	(Intercept)	3.95455	0.06718	58.863	< 2e-16	***	0.02087
	ICCoohne ICC	0.40152	0.09501	4.226	2.66e-05	***	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Anschließend wird die unabhängige Variable *ICC* in der Modellformel aus Tabelle 7.6 durch die unabhängige Variable *Geräusch* ersetzt (vgl. Tabelle 7.7). Dabei unterscheiden sich die Ausprägungen dieser Variablen bei jeder Sprecherin bzw. jedem Sprecher hoch signifikant voneinander ($p < 0,0001$). Zudem steigt das *Adjusted R-squared* im Gegensatz zum vorangegangenen Modell erheblich. Anhand der Schätzwerte lässt sich feststellen, dass die wahrgenommene Lautstärke der Stimuli unabhängig vom ICC-System mit steigender Geräuschbelastung von *Stille* zu *50 km/h*

7 Perzeptionsexperiment

um 0,5303 (2d), 0,69697 (3f) und 0,3447 (8o) Skalenpunkte und von *50 km/h* zu *130 km/h* um 0,51894 (2d), 1,01515 (3f) und 1,19697 (8o) Skalenpunkte zunimmt.

Tabelle 7.7: Regressionsergebnisse der Skala „laut-leise“ für die abhängige Variable *evaluation* und die erklärende Variable *Geräusch*.

Call:

lm(formula = evaluation ~ Geräusch, data = dataicc_laut)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)	Adjusted R ²	
2d	(Intercept)	5.82197	0.06368	91.421	< 2e-16	***	
	Geraeusch50 km/h	-0.51894	0.09006	-5.762	1.19e-08	***	0.1446
	GeraeuschStille	-1.04924	0.09006	-11.650	< 2e-16	***	
3f	(Intercept)	5.31818	0.07039	75.55	<2e-16	***	
	Geraeusch50 km/h	-1.01515	0.09955	-10.20	<2e-16	***	0.2731
	GeraeuschStille	-1.71212	0.09955	-17.20	<2e-16	***	
8o	(Intercept)	5.06818	0.07261	69.80	<2e-16	***	
	Geraeusch50 km/h	-1.19697	0.10269	-11.66	<2e-16	***	0.2375
	GeraeuschStille	-1.54167	0.10269	-15.01	<2e-16	***	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Aufgrund der Ergebnisse des Produktionsexperiments in Kapitel 6.3.1 wird für diesen Parameter ebenfalls eine Interaktion zwischen den unabhängigen Variablen *ICC* und *Geräusch* vermutet. Tatsächlich steigt das *Adjusted R-squared* in Tabelle 7.8 gegenüber dem vorangegangenen Regressionsmodell für jede Sprecherin bzw. jeden Sprecher an. Dabei ist die Ausprägung *ohne ICC* bei Sprecherin 2d nur schwach signifikant unterschiedlich vom *Intercept* ($p=0,0247$), wohingegen Sprecher 3f für diese Variablenausprägung eine höhere Signifikanz aufweist ($p=0,0005$). Weiterhin sind wesentliche Interaktionen zwischen *ICC* und *Geräusch* feststellbar. Diese sind jedoch bei Sprecher*in 2d und 3f nur in der stillen Referenzbedingung (schwach) signifikant (Sprecherin 2d: $p=0,0286$; Sprecher 3f: $p=0,0009$). Für Sprecher 8o weisen beide dargestellten Interaktionen signifikante Effekte auf ($p=0,0068$ und $p=0,0002$). Generell fallen die Werte des *Adjusted R-squared* in diesen Regressionsmodellen eher niedrig aus, aufgrund der hohen Stichprobengröße von $n=132$ sind die Signifikanzniveaus einzelner Variablenausprägungen dennoch sehr hoch.

Tabelle 7.8: Regressionsergebnisse der Skala „laut-leise“ für die abhängige Variable *evaluation*, die erklärenden Variablen *ICC* und *Geräusch* sowie deren Interaktionen.

Call:

lm(formula = evaluation ~ ICC * Geräusch, data = dataicc_laut)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	5.681818	0.088072	64.514	< 2e-16	***	0.182
	ICCOhne ICC	0.280303	0.124552	2.250	0.0247	*	
	Geraeusch50 km/h	-0.522727	0.124552	-4.197	3.02e-05	***	
	GeraeuschStille	-1.242424	0.124552	-9.975	< 2e-16	***	
	ICCOhne ICC:	0.007576	0.176143	0.043	0.9657		
	Geraeusch50 km/h						
	ICCOhne ICC:	0.386364	0.176143	2.193	0.0286	*	
GeraeuschStille							
3f	(Intercept)	5.07576	0.09818	51.697	< 2e-16	***	0.2929
	ICCOhne ICC	0.48485	0.13885	3.492	0.000506	***	
	Geraeusch50 km/h	-1.00758	0.13885	-7.257	9.54e-13	***	
	GeraeuschStille	-1.38636	0.13885	-9.985	< 2e-16	***	
	ICCOhne ICC:	-0.01515	0.19637	-0.077	0.938516		
	Geraeusch50 km/h						
	ICCOhne ICC	-0.65152	0.19637	-3.318	0.000949	***	
:GeraeuschStille							
8o	(Intercept)	4.6515	0.1004	46.316	< 2e-16	***	0.2707
	ICCOhne ICC	0.8333	0.1420	5.867	6.52e-09	***	
	Geraeusch50 km/h	-0.9242	0.1420	-6.507	1.36e-10	***	
	GeraeuschStille	-1.1667	0.1420	-8.214	8.75e-16	***	
	ICCOhne ICC	-0.5455	0.2009	-2.716	0.006761	**	
	:Geraeusch50 km/h						
	ICCOhne ICC	-0.7500	0.2009	-3.734	0.000202	***	
:GeraeuschStille							

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

7 Perzeptionsexperiment

Die Mittelwerte dieses Skalenparameters (vgl. Abbildung 7.2) zeigen bereits deutliche Effektrichtungen auf, welche durch die Ergebnisse der Regressionsanalyse bestätigt werden. Dabei wird zudem deutlich, dass Sprecherin 2d als einzige Person in jeder ICC- und Geräuschbedingung als eher *laut* bewertet wird, wohingegen Sprecher 3f und 8o in den Bedingungen *ohne ICC, Stille* und *mit ICC, Stille* noch eher leise klingen. Für Sprecher 8o gilt dies darüber hinaus auch *mit ICC* bei *50 km/h*. Tabelle 7.9 und Abbildung 7.5 zeigen die transformierten Schätzwerte des besten Regressionsmodells aus Tabelle 7.8 an. Diese *Estimates* sind für jede Sprecherin bzw. jeden Sprecher an der Experimentbedingung *ohne ICC, Stille* ausgerichtet, indem der Wert dieser Bedingung jeweils von allen Schätzwerten subtrahiert worden ist. Es wird deutlich, dass sowohl ohne als auch mit ICC-System eine Steigerung der wahrgenommenen Lautstärke stattfindet, je höher die Geräuschbelastung der Sprecherin bzw. des Sprechers ist. Zudem sind alle Werte mit ICC-System niedriger als ohne, mit Ausnahme der Bedingung *Stille* bei Sprecher 3f. Für Sprecher 3f und 8o gilt zudem, dass die Zunahme der wahrgenommenen Lautstärke bei steigenden Hintergrundgeräuschen mit ICC-System weniger ausgeprägt ist als ohne ICC-System, da die Differenzen zwischen den einzelnen Geräuschbedingungen *mit ICC* geringer sind. Für Sprecherin 2d ist das Gegenteil der Fall.

Tabelle 7.9: Die auf Basis des Regressionsmodells aus Tabelle 7.8 errechneten transformierten Schätzwerte zur wahrgenommenen Lautstärke.

	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
2d	0	0.3409	0.8561	-0.6667	0.0530	0.5758
3f	0	1.0152	2.0379	0.1667	0.5455	1.5530
8o	0	0.4470	1.9167	-0.0833	0.1591	1.0833

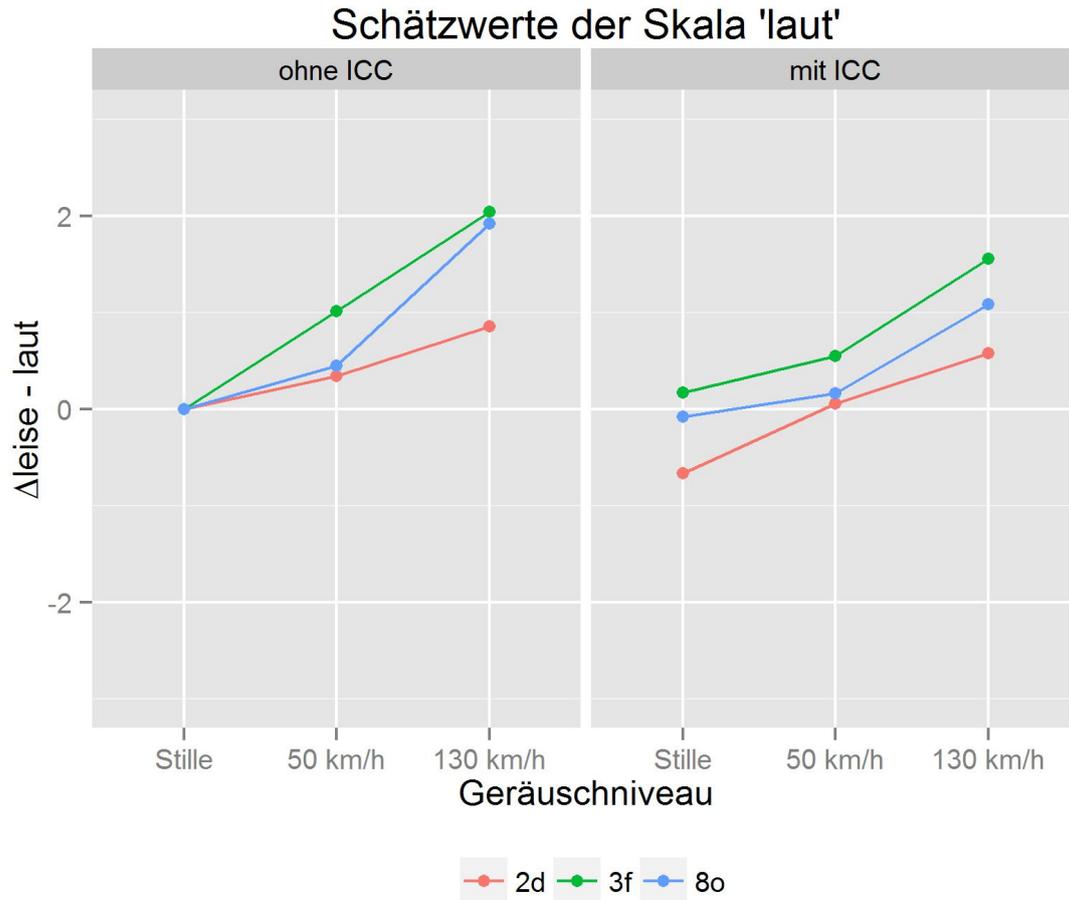


Abbildung 7.5: Die auf Basis des Regressionsmodells aus Tabelle 7.8 errechneten transformierten Schätzwerte zur wahrgenommenen Lautstärke.

verständlich/unverständlich

Bei der Verständlichkeitsskala bildet ein Regressionsmodell mit der einzigen unabhängigen Variablen *ICC* hohe Signifikanzniveaus für die Stimuli der Sprecher*innen 2d und 8o ($p < 0,0001$) ab (vgl. Tabelle 7.10). Den Schätzwerten zufolge sind die Sprachsignale ohne ICC-System um 1,12879 (2d) und 0,46212 (8o) Skaleneinheiten verständlicher. Für Sprecher 3f hingegen ist kein signifikanter Unterschied zwischen den Messwerten ohne und mit ICC-System feststellbar, wobei der Erklärungswert des Regressionsmodells für diesen Sprecher sehr gering ist (*Adjusted R-squared* = -0,0006761).

Tabelle 7.10: Regressionsergebnisse der Skala „verständlich-unverständlich“ für die abhängige Variable *evaluation* und die erklärende Variable *ICC*.

Call:

lm(formula = evaluation ~ ICC, data = dataicc_verstaendlich)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.08333	0.08504	48.016	<2e-16	***	0.09918
	ICCoohne ICC	1.12879	0.12027	9.386	<2e-16	***	
3f	(Intercept)	5.90152	0.06281	93.960	<2e-16	***	-0.0006761
	ICCoohne ICC	-0.06061	0.08883	-0.682	0.495		
8o	(Intercept)	4.97727	0.07777	63.996	< 2e-16	***	0.02062
	ICCoohne ICC	0.46212	0.10999	4.201	2.95e-05	***	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Nach Ersetzen der unabhängigen Variable *ICC* durch *Geräusch* können ebenfalls signifikante Effekte festgestellt werden (vgl. Tabelle 7.11). Den *Estimates* ist zu entnehmen, dass die Verständlichkeit der Sprachsignale bei Sprecherin 2d von *Stille* zu *50 km/h* um 0,4545 Skaleneinheiten zunimmt und zu *130 km/h* hin wiederum am niedrigsten ist. Dabei ist der Erklärungswert dieses Modells für Sprecherin 2d niedriger als in der vorangegangenen Tabelle 7.10 (*Adjusted R-squared* = 0,03379). Die Sprachsignale von Sprecher 3f werden immer verständlicher, je lauter es wird (0,25 Skaleneinheiten von *Stille* zu *50 km/h* und 0,3068 Skaleneinheiten von *50 km/h* zu *130 km/h*). Für Sprecher 8o hingegen fallen die Werte von *Stille* zu *50 km/h* hin ab (0,02651 Skaleneinheiten), wobei jedoch die Geräuschbelastung mit *130 km/h* die höchste Verständlichkeit birgt. Zudem erfährt das *Adjusted R-squared* bei 3f und 8o

im Vergleich zum vorangegangenen Regressionsmodell in Tabelle 7.10 eine Steigerung. Zusammenfassend lässt sich bei allen drei Sprecher*innen eine Steigerung der Verständlichkeit von der stillen Referenzbedingung zu *50 km/h* oder *130 km/h* hin feststellen. Durch diese Ergebnisse lässt sich jedoch kein ähnlich klarer Lombard-Effekt wie bei der Skala „laut-leise“ nachweisen.

Tabelle 7.11: Regressionsergebnisse der Skala „verständlich-unverständlich“ für die abhängige Variable *evaluation* und die erklärende Variable *Geräusch*.

Call:

lm(formula = evaluation ~ Geräusch, data = dataicc_verstaendlich)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.2462	0.1079	39.365	< 2e-16	***	0.03379
	Geraeusch50 km/h	0.8295	0.1525	5.438	7.19e-08	***	
	GeraeuschStille	0.3750	0.1525	2.458	0.0142	*	
3f	(Intercept)	6.1591	0.0757	81.356	< 2e-16	***	0.03081
	Geraeusch50 km/h	-0.3068	0.1071	-2.866	0.00427	**	
	GeraeuschStille	-0.5568	0.1071	-5.201	2.53e-07	***	
8o	(Intercept)	5.56818	0.09509	58.560	< 2e-16	***	0.02409
	Geraeusch50 km/h	-0.55303	0.13447	-4.113	4.32e-05	***	
	GeraeuschStille	-0.52652	0.13447	-3.915	9.80e-05	***	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

In Tabelle 7.12 werden zudem Interaktionen der unabhängigen Variablen *ICC* und *Geräusch* untersucht. Dabei werden nicht alle ICC- und Geräuschbedingungen der drei Sprecher*innen als signifikant unterschiedlich angezeigt. Der Grund hierfür ist, dass sich die signifikanten Effekte in die Interaktionen von *ICC* und *Geräusch* verlagern, weil die Interaktionen diese Effekte besser erklären. Das *Adjusted R-squared* aller drei Sprecher*innen ist höher als in den vorangegangenen Tabellen 7.10 und 7.11, weshalb dieses Regressionsmodell als das mit dem besten Erklärungswert beibehalten wird.

7 Perzeptionsexperiment

Tabelle 7.12: Regressionsergebnisse der Skala „verständlich-unverständlich“ für die abhängige Variable *evaluation*, die erklärenden Variablen *ICC* und *Geräusch* sowie deren Interaktionen.

Call:

lm(formula = evaluation ~ ICC * Geräusch, data = dataicc_verstaendlich)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.15909	0.14103	29.490	< 2e-16	***	0.1741
	ICCoohne ICC	0.17424	0.19945	0.874	0.382597		
	Geraeusch50 km/h	-0.06818	0.19945	-0.342	0.732557		
	GeraeuschStille	-0.15909	0.19945	-0.798	0.425317		
	ICCoohne ICC :Geraeusch50 km/h	1.79545	0.28207	6.365	3.31e-10	***	
	ICCoohne ICC :GeraeuschStille	1.06818	0.28207	3.787	0.000164	***	
3f	(Intercept)	6.20455	0.10634	58.346	< 2e-16	***	0.04383
	ICCoohne ICC	-0.09091	0.15039	-0.604	0.5457		
	Geraeusch50 km/h	-0.13636	0.15039	-0.907	0.3648		
	GeraeuschStille	-0.77273	0.15039	-5.138	3.5e-07	***	
	ICCoohne ICC :Geraeusch50 km/h	-0.34091	0.21268	-1.603	0.1094		
	ICCoohne ICC :GeraeuschStille	0.43182	0.21268	2.030	0.0427	*	
8o	(Intercept)	5.121e+00	1.305e-01	39.241	< 2e-16	***	0.08077
	ICCoohne ICC	8.939e-01	1.846e-01	4.843	1.54e-06	***	
	Geraeusch50 km/h	-5.530e-01	1.846e-01	-2.996	0.00282	**	
	GeraeuschStille	1.212e-01	1.846e-01	0.657	0.51154		
	ICCoohne ICC :Geraeusch50 km/h	-1.486e-14	2.610e-01	0.000	1.00000		
	ICCoohne ICC :GeraeuschStille	-1.295e+00	2.610e-01	-4.963	8.51e-07	***	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Tabelle 7.13 und Abbildung 7.6 zeigen die transformierten Schätzwerte des Regressionsmodells mit dem besten Erklärungswert aus Tabelle 7.12. Dabei wird wiederum der Wert der Ausprägung *ohne ICC, Stille* für alle Sprecher*innen von jedem Schätzwert subtrahiert. Hinsichtlich der Geräuschbedingungen zeigen alle drei Sprecher*innen verschiedene Tendenzen auf. Für Sprecherin 2d existieren ohne und mit ICC-System unterschiedliche Effekte hinsichtlich der zunehmenden Fahrgeräusche. Ohne ICC-System nimmt die Verständlichkeit von Sprecherin 2d von *Stille* zu *50 km/h* zu und fällt dann zu *130 km/h* wieder ab. Das Gegenteil ist für Sprecher 3f der Fall. Sprecher 8o zeigt einen stetigen Anstieg der Verständlichkeit mit jeder Zunahme der Hintergrundgeräusche. Demnach lässt sich für Sprecher 3f und 8o eine Erhöhung der Verständlichkeit von *Stille* zu *130 km/h* feststellen sowie für Sprecherin 2d von *Stille* zu *50 km/h*. Mit ICC-System werden die Sprachsignale der Sprecher*innen 2d und 3f verständlicher, je lauter es wird. Bei Sprecher 8o nimmt die Verständlichkeit von *Stille* zu *50 km/h* zunächst ab, während sie von *50 km/h* zu *130 km/h* wieder etwas zunimmt. Weiterhin ist das Gesprochene in fast jeder Experimentbedingung mit ICC-System unverständlicher. Ausnahmen hierzu bilden Sprecher 3f bei *50 km/h* und *130 km/h* sowie Sprecher 8o in *Stille*. Die Ergebnisse deuten darauf hin, dass die Auswahl der Sprachausschnitte einen entscheidenden Beitrag zu den Ergebnissen der vorliegenden Analyse leistet.

Tabelle 7.13: Die auf Basis des Regressionsmodells aus Tabelle 7.12 errechneten transformierten Schätzwerte zur wahrgenommenen Verständlichkeit.

	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
2d	0	0.8182	-0.9091	-1.2424	-1.1515	-1.0833
3f	0	-0.1364	0.3409	-0.3409	0.2955	0.4318
8o	0	0.6212	1.1742	0.4015	-0.2727	0.2803

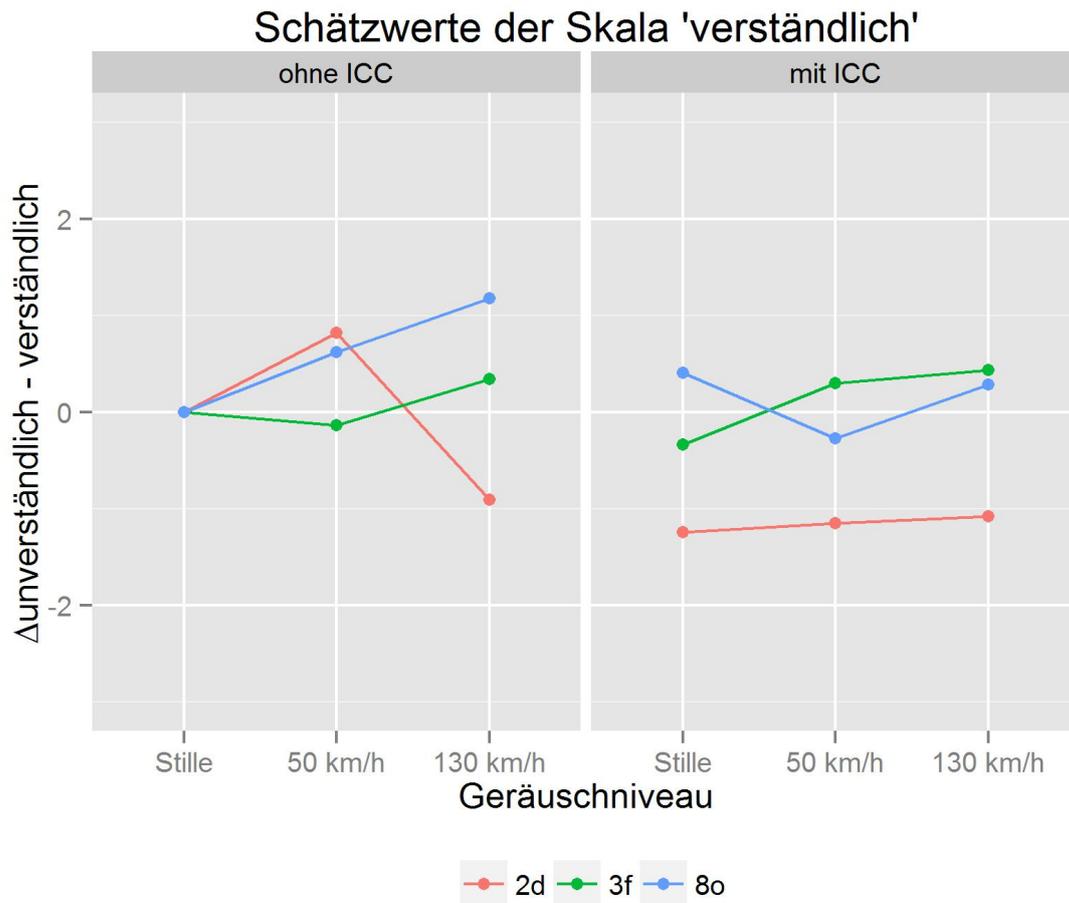


Abbildung 7.6: Die auf Basis des Regressionsmodells aus Tabelle 7.12 errechneten transformierten Schätzwerte zur wahrgenommenen Verständlichkeit.

unfreundlich/freundlich

Hinsichtlich der Skala „unfreundlich-freundlich“ werden für die unabhängige Variable *ICC* im Regressionsmodell aus Tabelle 7.14 bei zwei von drei Sprecher*innen keine signifikanten Effekte erkannt. Sprecher 8o hingegen weist einen signifikanten Einfluss des ICC-Systems auf ($p=0,0004$), bei dessen Aktivierung unfreundlicher gesprochen wird.

Tabelle 7.14: Regressionsergebnisse der Skala „unfreundlich-freundlich“ für die abhängige Variable *evaluation* und die erklärende Variable *ICC*.

Call:

lm(formula = evaluation ~ ICC, data = dataicc_unfreundlich)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	3.71465	0.06475	57.366	<2e-16	***	-0.001019
	ICCoohne ICC	-0.04040	0.09157	-0.441	0.659		
3f	(Intercept)	2.98485	0.05572	53.569	<2e-16	***	-0.0006362
	ICCoohne ICC	0.05556	0.07880	0.705	0.481		
8o	(Intercept)	3.27273	0.06343	51.597	< 2e-16	***	0.01468
	ICCoohne ICC	-0.32071	0.08970	-3.575	0.000371	***	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Die erklärende Variable *Geräusch* erzielt für alle Sprecher*innen ein höheres *Adjusted R-squared* als im vorangegangenen Regressionsmodell (vgl. Tabelle 7.15). Sprecherin 2d weist hohe Signifikanzen für die Unterscheidung der drei Geräuschszenarien auf ($p=0,0035$ und $p<0,0001$). Dabei hören sich die Sprachsignale dieser Probandin immer unfreundlicher an, je lauter es wird. Für Sprecher 3f und 8o sind hoch signifikante Effekte zwischen den Bedingungen *50 km/h* und *130 km/h* zu verzeichnen (Sprecher 3f: $p=0,0001$; Sprecher 8o: $p<0,0001$). Da sich die stille Referenzbedingung nicht bzw. nur schwach signifikant von *130 km/h* unterscheidet (Sprecher 8o: $p=0,0323$), weist diese eine große Differenz zu *50 km/h* auf. Dabei wird das Gesagte bei Sprecher 3f mit steigender Geräuschbelastung von *Stille* zu *130 km/h* unfreundlicher, während es sich bei Sprecher 8o andersherum verhält.

Tabelle 7.15: Regressionsergebnisse der Skala „unfreundlich-freundlich“ für die abhängige Variable *evaluation* und die erklärende Variable *Geräusch*.

Call:

lm(formula = evaluation ~ Geraeusch, data = dataicc_unfreundlich)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.01136	0.07777	51.579	< 2e-16	***	0.03735
	Geraeusch50 km/h	-0.32197	0.10999	-2.927	0.00352	**	
	GeraeuschStille	-0.62879	0.10999	-5.717	1.54e-08	***	
3f	(Intercept)	3.15152	0.06754	46.660	< 2e-16	***	0.01979
	Geraeusch50 km/h	-0.37121	0.09552	-3.886	0.00011	***	
	GeraeuschStille	-0.04545	0.09552	-0.476	0.63430		
8o	(Intercept)	3.19697	0.07617	41.971	< 2e-16	***	0.05272
	Geraeusch50 km/h	-0.48485	0.10772	-4.501	7.78e-06	***	
	GeraeuschStille	0.23106	0.10772	2.145	0.0323	*	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Für Sprecher 3f und 8o ist das Regressionsmodell in Tabelle 7.16, welches Interaktionen der unabhängigen Variablen *ICC* und *Geräusch* berücksichtigt, das mit dem höchsten *Adjusted R-squared* und somit dem besten Erklärungswert. In Bezug auf Sprecherin 2d ist das *Adjusted R-squared* in diesem Modell etwas geringer als im vorangegangenen. Deshalb wird für diese Sprecherin das Regressionsmodell aus Tabelle 7.15 als das mit dem besten Erklärungswert beibehalten.

Tabelle 7.16: Regressionsergebnisse der Skala „unfreundlich-freundlich“ für die abhängige Variable *evaluation*, die erklärenden Variablen *ICC* und *Geräusch* sowie deren Interaktionen.

Call:

lm(formula = evaluation ~ ICC * Geräusch, data = dataicc_unfreundlich)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.1061	0.1101	37.307	< 2e-16	***	0.03603
	ICCOhne ICC	-0.1894	0.1556	-1.217	0.2240		
	Geraeusch50 km/h	-0.4015	0.1556	-2.580	0.0101	*	
	GeraeuschStille	-0.7727	0.1556	-4.965	8.45e-07	***	
	ICCOhne ICC :Geraeusch50 km/h	0.1591	0.2201	0.723	0.4701		
	ICCOhne ICC :GeraeuschStille	0.2879	0.2201	1.308	0.1913		
3f	(Intercept)	3.14394	0.09511	33.056	< 2e-16	***	0.02818
	ICCOhne ICC	0.01515	0.13450	0.113	0.9103		
	Geraeusch50 km/h	-0.54545	0.13450	-4.055	5.51e-05	***	
	GeraeuschStille	0.06818	0.13450	0.507	0.6124		
	ICCOhne ICC :Geraeusch50 km/h	0.34848	0.19022	1.832	0.0673	.	
	ICCOhne ICC :GeraeuschStille	-0.22727	0.19022	-1.195	0.2325		
8o	(Intercept)	3.67424	0.09691	37.912	< 2e-16	***	0.2332
	ICCOhne ICC	-0.95455	0.13706	-6.965	6.97e-12	***	
	Geraeusch50 km/h	-0.38636	0.13706	-2.819	0.00494	**	
	GeraeuschStille	-0.81818	0.13706	-5.970	3.60e-09	***	
	ICCOhne ICC :Geraeusch50 km/h	-0.19697	0.19383	-1.016	0.30985		
	ICCOhne ICC :GeraeuschStille	2.09848	0.19383	10.826	< 2e-16	***	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

7 Perzeptionsexperiment

Die folgende Tabelle 7.17 sowie Abbildung 7.7 fassen die Schätzwerte von Sprecherin 2d zusammen, für welche das Regressionsmodell in Tabelle 7.15 den besten Erklärungswert hat (*Adjusted R-squared*=0.03735). Diese Werte werden derart transformiert, dass die Bedingung *Stille* als Basis den Wert 0 annimmt. Das ICC-System übt keinen signifikanten Einfluss auf die wahrgenommene Freundlichkeit dieser Sprecherin aus. Es lässt sich jedoch feststellen, dass Sprecherin 2d immer unfreundlicher klingt, je stärker die Umgebungsgeräusche ansteigen. Hierbei ist ein Anstieg der Skalenpunkte um 0,3068 von *Stille* auf *50 km/h* und um 0,322 von *50 km/h* auf *130 km/h* zu verzeichnen.

Weiterhin zeigen Tabelle 7.18 und Abbildung 7.8 die Schätzwerte von Sprecher 3f und 8o an, welche auf dem Regressionsmodell in Tabelle 7.16 beruhen und ebenfalls transformiert werden, sodass die Bedingung *ohne ICC, Stille* zu 0 wird. Dabei ist erkennbar, dass ohne ICC-System zunächst ein Abfall der Messwerte von *Stille* zu *50 km/h* und anschließend eine Zunahme der Werte von *50 km/h* zu *130 km/h* stattfindet. Ein ähnliches Muster ergibt sich für Sprecher 3f bei aktiviertem ICC-System. Sprecher 8o klingt *mit ICC* immer unfreundlicher, je lauter es wird. Es lässt sich wiederum feststellen, dass die Vorauswahl der Sprachausschnitte die Ergebnisse der Untersuchung beeinflusst.

Tabelle 7.17: Die auf Basis des Regressionsmodells aus Tabelle 7.15 errechneten transformierten Schätzwerte zur wahrgenommenen Unfreundlichkeit von Sprecherin 2d.

	Stille	50 km/h	130 km/h
2d	0	0.3068	0.6288

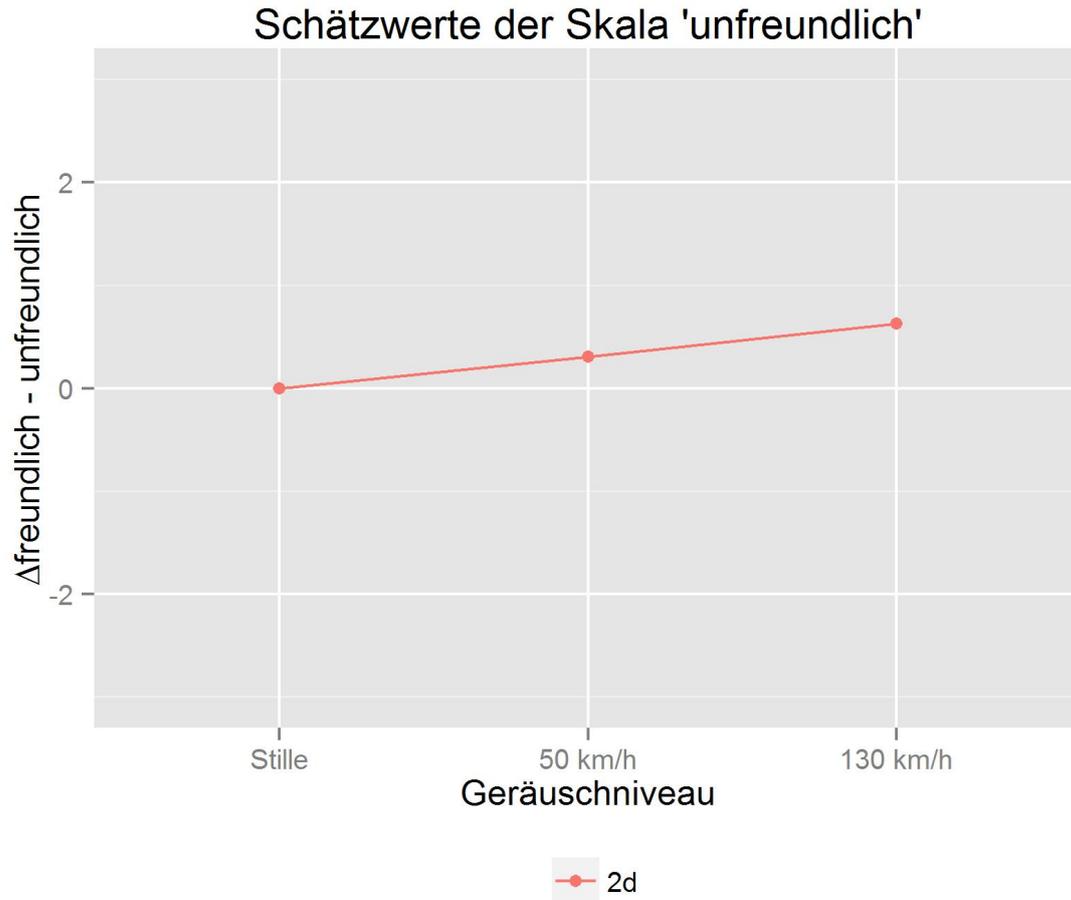


Abbildung 7.7: Die auf Basis des Regressionsmodells aus Tabelle 7.15 errechneten transformierten Schätzwerte zur wahrgenommenen Unfreundlichkeit von Sprecherin 2d.

Tabelle 7.18: Die auf Basis des Regressionsmodells aus Tabelle 7.16 errechneten transformierten Schätzwerte zur wahrgenommenen Unfreundlichkeit von Sprecher 3f und 8o.

	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
3f	0	-0.0379	0.1591	0.2121	-0.4015	0.1439
8o	0	-1.8636	-1.2803	-1.1439	-0.7121	-0.3258

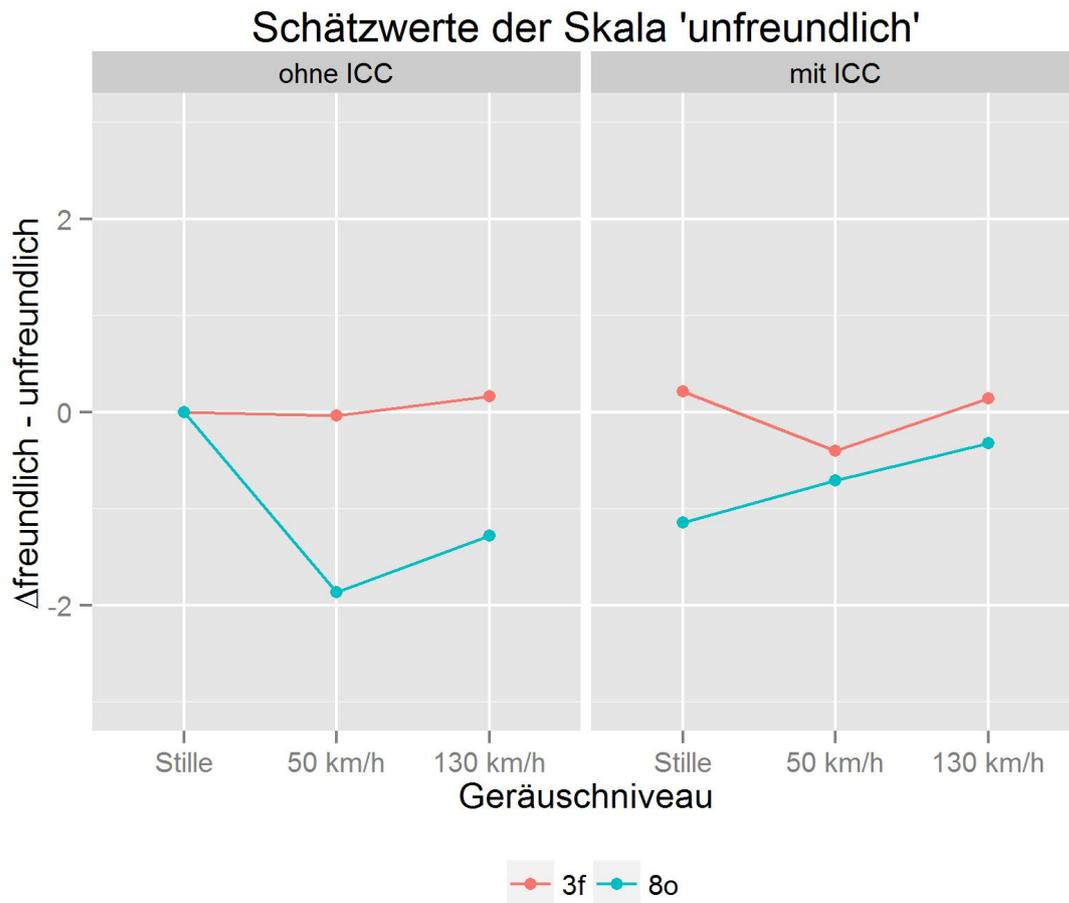


Abbildung 7.8: Die auf Basis des Regressionsmodells aus Tabelle 7.16 errechneten transformierten Schätzwerte zur wahrgenommenen Unfreundlichkeit von Sprecher 3f und 8o.

emotional/sachlich

Die Effekte des ICC-Systems sind in dieser Skala für alle drei Sprecher*innen hoch signifikant ($p < 0,0001$) (vgl. Tabelle 7.19). Dabei werden die Sprachsignale von Sprecherin 2d mit ICC emotionaler, während sich die von Sprecher 3f und 8o sachlicher anhören.

Tabelle 7.19: Regressionsergebnisse der Skala „emotional-sachlich“ für die abhängige Variable *evaluation* und die erklärende Variable *ICC*.

Call:

`lm(formula = evaluation ~ ICC, data = dataicc_emotional)`

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.42172	0.07288	60.669	< 2e-16	***	0.03394
	ICCoohne ICC	-0.55303	0.10307	-5.366	1.06e-07	***	
3f	(Intercept)	2.50505	0.07194	34.819	< 2e-16	***	0.07406
	ICCoohne ICC	0.81566	0.10174	8.017	3.9e-15	***	
8o	(Intercept)	3.08838	0.07030	43.94	<2e-16	***	0.115
	ICCoohne ICC	1.01263	0.09941	10.19	<2e-16	***	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Für Sprecher*in 2d und 3f ergibt sich durch das Ersetzen von *ICC* durch *Geräusch* in der Regressionsformel ein erhöhter Erklärungswert (vgl. Tabelle 7.20). Dabei ist eine signifikante Zunahme der Emotionalität von *Stille* zu *130 km/h* zu verzeichnen ($p < 0,0001$). Hinsichtlich Sprecher 8o sind keine bedeutsamen Einflüsse des Hintergrundgeräuschs auf die Messwerte festzustellen und das *Adjusted R-squared* fällt hierbei in den negativen Bereich.

7 Perzeptionsexperiment

Tabelle 7.20: Regressionsergebnisse der Skala „emotional-sachlich“ für die abhängige Variable *evaluation* und die erklärende Variable *Geräusch*.

Call:

```
lm(formula = evaluation ~ Geräusch, data = dataicc_emotional)
```

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)	Adjusted R ²
2d	(Intercept)	4.38636	0.08858	49.518	< 2e-16	***
	Geraeusch50 km/h	-0.01136	0.12527	-0.091	0.928	0.0486
	GeraeuschStille	-0.71212	0.12527	-5.685	1.85e-08	***
3f	(Intercept)	3.57197	0.08685	41.127	< 2e-16	***
	Geraeusch50 km/h	-0.86742	0.12283	-7.062	3.6e-12	***
	GeraeuschStille	-1.10985	0.12283	-9.036	< 2e-16	***
8o	(Intercept)	3.68939	0.09153	40.306	<2e-16	***
	Geraeusch50 km/h	-0.15909	0.12945	-1.229	0.219	-0.0004113
	GeraeuschStille	-0.12500	0.12945	-0.966	0.335	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Das in Tabelle 7.21 dargestellte Regressionsmodell zeigt für alle drei Sprecher*innen den höchsten Erklärungswert an, da das *Adjusted R-squared* jeweils höher ist als in den vorangegangenen Regressionsmodellen.

Tabelle 7.21: Regressionsergebnisse der Skala „emotional-sachlich“ für die abhängige Variable *evaluation*, die erklärenden Variablen *ICC* und *Geräusch* sowie deren Interaktionen.

Call:

lm(formula = evaluation ~ ICC * Geraeusch, data = dataicc_emotional)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.3106	0.1207	35.726	< 2e-16	***	0.1174
	ICCOhne ICC	0.1515	0.1706	0.888	0.37484		
	Geraeusch50 km/h	0.6818	0.1706	3.996	7.05e-05	***	
	GeraeuschStille	-0.3485	0.1706	-2.042	0.04146	*	
	ICCOhne ICC :Geraeusch50 km/h	-1.3864	0.2413	-5.745	1.31e-08	***	
	ICCOhne ICC :GeraeuschStille	-0.7273	0.2413	-3.014	0.00266	**	
3f	(Intercept)	2.59091	0.11171	23.194	< 2e-16	***	0.2559
	ICCOhne ICC	1.96212	0.15798	12.420	< 2e-16	***	
	Geraeusch50 km/h	-0.17424	0.15798	-1.103	0.270		
	GeraeuschStille	-0.08333	0.15798	-0.527	0.598		
	ICCOhne ICC :Geraeusch50 km/h	-1.38636	0.22342	-6.205	8.83e-10	***	
	ICCOhne ICC :GeraeuschStille	-2.05303	0.22342	-9.189	< 2e-16	***	
8o	(Intercept)	3.1136	0.1208	25.765	< 2e-16	***	0.1281
	ICCOhne ICC	1.1515	0.1709	6.738	3.11e-11	***	
	Geraeusch50 km/h	-0.2727	0.1709	-1.596	0.11094		
	GeraeuschStille	0.1970	0.1709	1.153	0.24946		
	ICCOhne ICC :Geraeusch50 km/h	0.2273	0.2417	0.940	0.34734		
	ICCOhne ICC :GeraeuschStille	-0.6439	0.2417	-2.664	0.00787	**	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

7 Perzeptionsexperiment

Tabelle 7.22 und Abbildung 7.9 stellen die dem Regressionsmodell in Tabelle 7.21 entsprechenden Schätzwerte dar, welche wiederum derart transformiert werden, dass die Bedingung *ohne ICC, Stille* den Wert 0 erhält. Hinsichtlich dieser Schätzwerte wird deutlich, dass die Sprecher*innen ohne ICC-System mit ansteigenden Hintergrundgeräuschen weniger sachlich bzw. emotionaler sprechen. Weiterhin scheint Sprecherin 2d mit ICC-System bei Fahrgeräuschen emotionaler zu sprechen als in *Stille*, während 8o sachlicher wird. Sprecher 3f klingt *mit ICC* von *Stille* zu *50 km/h* sachlicher und bei *130 km/h* wieder etwas emotionaler als in *Stille*. Demnach gibt es ohne ICC-System von *Stille* zu *50 km/h* und von *50 km/h* zu *130 km/h* einen generellen Anstieg der Emotionalität, der mit Aktivierung des ICC-Systems nicht mehr vorhanden ist bzw. stark abgeschwächt wird.

Tabelle 7.22: Die auf Basis des Regressionsmodells aus Tabelle 7.21 errechneten transformierten Schätzwerte zur wahrgenommenen Emotionalität.

	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
2d	0	0.3712	1.0758	0.5758	1.6061	0.9242
3f	0	0.5758	2.1364	0.0909	1.065814e-14	0.1742
8o	0	0.4015	0.4470	-0.5076	-0.9773	-0.7046

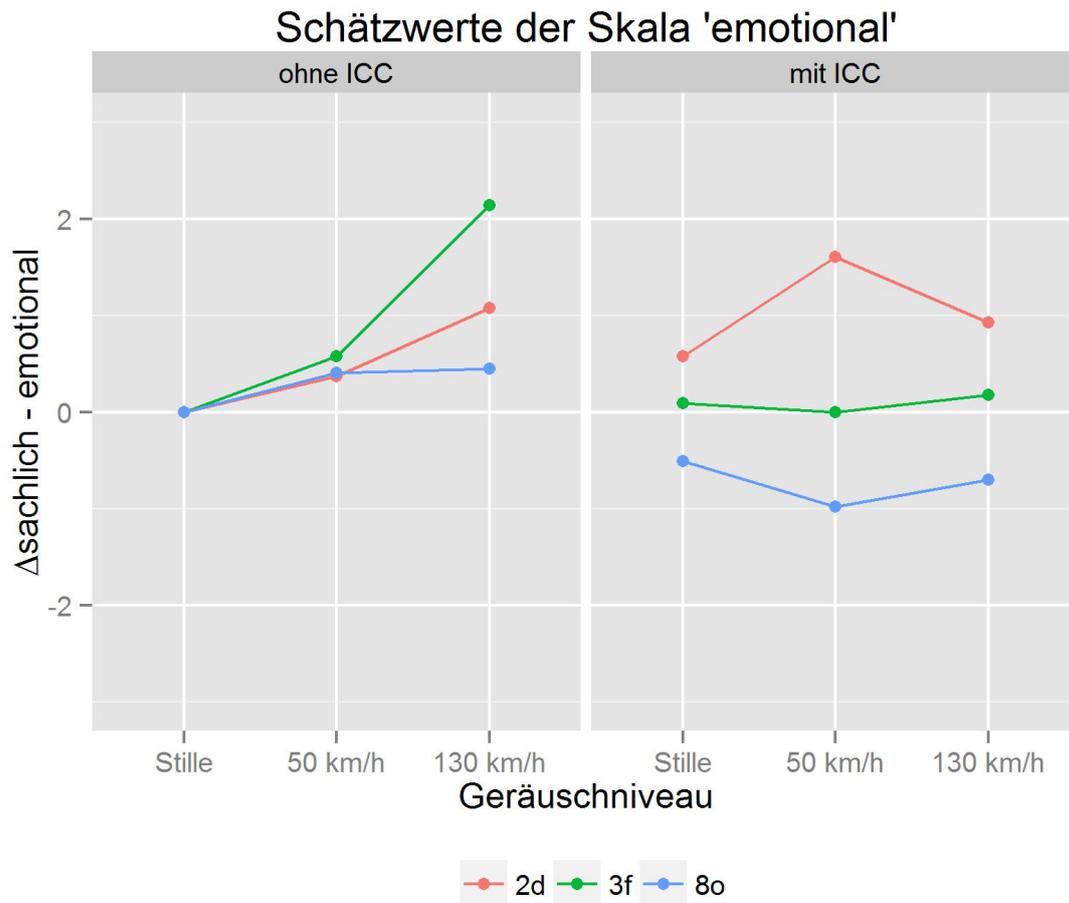


Abbildung 7.9: Die auf Basis des Regressionsmodells aus Tabelle 7.21 errechneten transformierten Schätzwerte zur wahrgenommenen Emotionalität.

gehemmt/ungehemmt

Das Regressionsmodell mit der einzigen erklärenden Variablen *ICC* in Tabelle 7.23 weist nur für die Sprachsignale von Sprecher 3f und 8o signifikante Effekte auf (Sprecher 3f: $p=0,0021$; Sprecher 8o: $p<0,0001$). Dabei wirkt das Gesprochene mit ICC-System gehemmter.

Tabelle 7.23: Regressionsergebnisse der Skala „gehemmt-ungehemmt“ für die abhängige Variable *evaluation* und die erklärende Variable *ICC*.

Call:

```
lm(formula = evaluation ~ ICC, data = dataicc_gehemmt)
```

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)	Adjusted R ²
2d	(Intercept)	3.11111	0.06725	46.264	<2e-16	***
	ICCoohne ICC	0.04293	0.09510	0.451	0.652	-0.001008
3f	(Intercept)	3.21212	0.06825	47.063	< 2e-16	***
	ICCoohne ICC	-0.29798	0.09652	-3.087	0.00209	**
8o	(Intercept)	3.46212	0.06913	50.079	< 2e-16	***
	ICCoohne ICC	-0.51768	0.09777	-5.295	1.55e-07	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Das Regressionsmodell in Tabelle 7.24, welches *Geräusch* als einzige erklärende Variable beinhaltet, weist keinen höheren Erklärungswert für Sprecher 8o auf, jedoch für die beiden anderen Sprecher*innen. Dabei werden die Sprachsignale der Sprecherin 2d von *Stille* zu *50 km/h* zunächst ungehemmter und zu *130 km/h* dann gehemmter, während die Stimuli von Sprecher 3f mit zunehmenden Hintergrundgeräuschen als immer ungehemmter empfunden werden.

Tabelle 7.24: Regressionsergebnisse der Skala „gehemmt-ungehemmt“ für die abhängige Variable *evaluation* und die erklärende Variable *Geräusch*.

Call:

lm(formula = evaluation ~ Geraeusch, data = dataicc_gehemmt)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)	Adjusted R ²	
2d	(Intercept)	3.32576	0.08126	40.930	< 2e-16	***	
	Geraeusch50 km/h	-0.50758	0.11491	-4.417	1.14e-05	***	0.02568
	GeraeuschStille	-0.07197	0.11491	-0.626	0.531		
3f	(Intercept)	2.84091	0.08347	34.035	< 2e-16	***	
	Geraeusch50 km/h	0.24621	0.11805	2.086	0.03732	*	0.01348
	GeraeuschStille	0.42045	0.11805	3.562	0.00039	***	
8o	(Intercept)	3.20076	0.08610	37.173	<2e-16	***	
	Geraeusch50 km/h	-0.08333	0.12177	-0.684	0.494		6.176e-05
	GeraeuschStille	0.09091	0.12177	0.747	0.456		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Tabelle 7.25 stellt für alle Sprecher*innen das Regressionsmodell mit dem höchsten *Adjusted R-squared* und damit dem besten Erklärungswert dar.

7 Perzeptionsexperiment

Tabelle 7.25: Regressionsergebnisse der Skala „gehemmt-ungehemmt“ für die abhängige Variable *evaluation*, die erklärenden Variablen *ICC* und *Geräusch* sowie deren Interaktionen.

Call:

lm(formula = evaluation ~ ICC * Geräusch, data = dataicc_gehemmt)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	3.5455	0.1140	31.104	< 2e-16	***	0.04133
	ICCOhne ICC	-0.4394	0.1612	-2.726	0.00656	**	
	Geraeusch50 km/h	-0.9545	0.1612	-5.921	4.77e-09	***	
	GeraeuschStille	-0.3485	0.1612	-2.162	0.03093	*	
	ICCOhne ICC :Geraeusch50 km/h	0.8939	0.2280	3.921	9.58e-05	***	
	ICCOhne ICC :GeraeuschStille	0.5530	0.2280	2.426	0.01550	*	
3f	(Intercept)	2.9091	0.1174	24.773	< 2e-16	***	0.02373
	ICCOhne ICC	-0.1364	0.1661	-0.821	0.411833		
	Geraeusch50 km/h	0.3409	0.1661	2.053	0.040425	*	
	GeraeuschStille	0.5682	0.1661	3.421	0.000655	***	
	ICCOhne ICC :Geraeusch50 km/h	-0.1894	0.2349	-0.806	0.420253		
	ICCOhne ICC :GeraeuschStille	-0.2955	0.2349	-1.258	0.208768		
8o	(Intercept)	3.8561	0.1153	33.444	< 2e-16	***	0.1035
	ICCOhne ICC	-1.3106	0.1631	-8.038	3.35e-15	***	
	Geraeusch50 km/h	-0.3712	0.1631	-2.277	0.0231	*	
	GeraeuschStille	-0.8106	0.1631	-4.971	8.17e-07	***	
	ICCOhne ICC :Geraeusch50 km/h	0.5758	0.2306	2.497	0.0127	*	
	ICCOhne ICC :GeraeuschStille	1.8030	0.2306	7.819	1.71e-14	***	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Tabelle 7.26 und Abbildung 7.10 zeigen wiederum die transformierten Schätzwerte, wobei der Wert der Bedingung *ohne ICC*, *Stille* bei jeder Sprecherin bzw. jedem Sprecher 0 wird. Ohne ICC-System zeigt sich, dass alle Sprecher*innen mit Hintergrundgeräuschen ungehemmter sprechen als in *Stille*. Bei Sprecher 3f und 8o fallen die Werte stetig ab. Sprecherin 2d zeigt wiederum einen leichten Anstieg der Werte von *50 km/h* zu *130 km/h*. Mit ICC-System wird deutlich, dass der Effekt der ansteigenden Fahrgeräusche für Sprecher*in 2d und 3f gleich bleibt, bei 2d nun jedoch ausgeprägter ist und bei 3f eine Niveauverschiebung nach oben erlebt. Weiterhin spricht 2d bei *130 km/h* nun gehemmt als in *Stille*. Sprecher 8o zeigt hingegen im Vergleich zur Bedingung *ohne ICC* einen gegenteiligen Effekt.

Tabelle 7.26: Die auf Basis des Regressionsmodells aus Tabelle 7.25 errechneten transformierten Schätzwerte zur wahrgenommenen Gehemtheit.

	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
2d	0	-0.2652	-0.2046	-0.1136	-0.7197	0.2349
3f	0	-0.1212	-0.2727	0.4318	0.2046	-0.1364
8o	0	-0.7879	-0.9924	-0.4924	-0.0530	0.3182

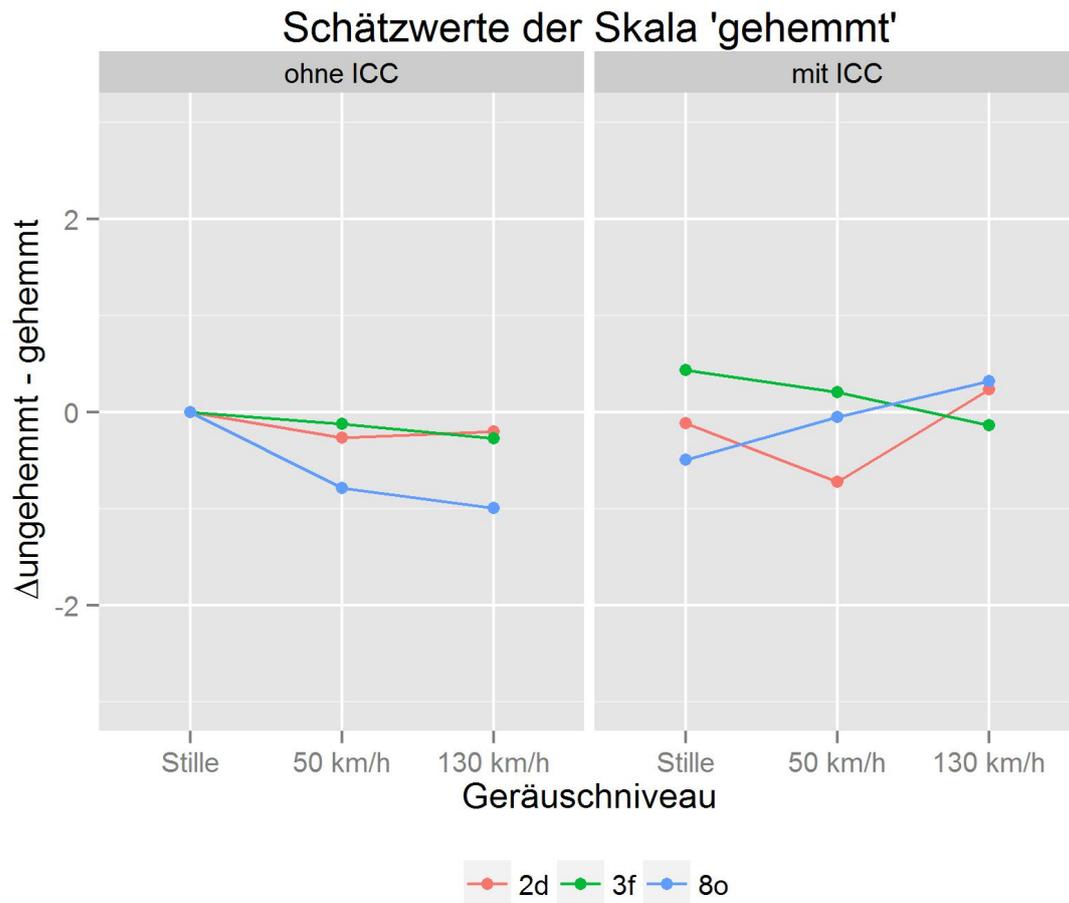


Abbildung 7.10: Die auf Basis des Regressionsmodells aus Tabelle 7.25 errechneten transformierten Schätzwerte zur wahrgenommenen Gehemtheit.

steif/locker

Das Regressionsmodell in Tabelle 7.27 zeigt hinsichtlich der Skala „steif-locker“ für alle drei Sprecher*innen signifikante Effekte bei der Anwendung des ICC-Systems (Sprecherin 2d: $p=0,0005$; Sprecher 3f und 8o: $p<0,0001$). Dabei klingt Sprecherin 2d *mit ICC* lockerer, während sich Sprecher 3f und 8o steifer anhören.

Tabelle 7.27: Regressionsergebnisse der Skala „steif-locker“ für die abhängige Variable *evaluation* und die erklärende Variable *ICC*.

Call:

`lm(formula = evaluation ~ ICC, data = dataicc_steif)`

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)	Adjusted R ²
2d	(Intercept)	3.58081	0.07553	47.411	< 2e-16	***
	ICCoohne ICC	0.37121	0.10681	3.475	0.000538	***
3f	(Intercept)	3.58586	0.07363	48.702	< 2e-16	***
	ICCoohne ICC	-0.45202	0.10413	-4.341	1.6e-05	***
8o	(Intercept)	3.53283	0.07559	46.734	< 2e-16	***
	ICCoohne ICC	-0.63131	0.10691	-5.905	5.23e-09	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Für die erklärende Variable *Geräusch* weist lediglich Sprecherin 2d ein höheres *Adjusted R-squared* auf als im vorangegangenen Modell mit der unabhängigen Variablen *ICC* (vgl. Tabelle 7.28). Die Sprachsignale dieser Sprecherin werden von *Stille* zu *50 km/h* lockerer und von *Stille* zu *130 km/h* steifer. Ebenso verhält es sich mit dem Gesagten des Sprechers 8o, wobei hier jedoch z. T. nur schwach signifikante Effekte vorliegen ($p=0,0692$ und $p=0,0059$). Für Sprecher 3f können keine bedeutsamen Einflüsse der Geräuschbelastung auf die Messwerte nachgewiesen werden.

7 Perzeptionsexperiment

Tabelle 7.28: Regressionsergebnisse der Skala „steif-locker“ für die abhängige Variable *evaluation* und die erklärende Variable *Geräusch*.

Call:
lm(formula = evaluation ~ Geräusch, data = dataicc_steif)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)	Adjusted R ²	
2d	(Intercept)	4.19318	0.09123	45.962	< 2e-16	***	
	Geraeusch50 km/h	-0.74621	0.12902	-5.784	1.05e-08	***	0.04066
	GeraeuschStille	-0.53409	0.12902	-4.140	3.85e-05	***	
3f	(Intercept)	3.40909	0.09127	37.350	<2e-16	***	
	Geraeusch50 km/h	-0.05682	0.12908	-0.440	0.660		-0.001892
	GeraeuschStille	-0.09091	0.12908	-0.704	0.481		
8o	(Intercept)	3.4205	0.0942	36.312	< 2e-16	***	
	Geraeusch50 km/h	-0.3674	0.1332	-2.758	0.00595	**	0.007362
	GeraeuschStille	-0.2424	0.1332	-1.820	0.06917	.	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Das in Tabelle 7.29 dargestellte Regressionsmodell berücksichtigt Interaktionen zwischen *ICC* und *Geräusch* und stellt für Sprecher*in 2d und 8o das Modell mit dem höchsten *Adjusted R-squared* dar. Bei Sprecher 3f birgt das in Tabelle 7.27 beschriebene Modell den besten Erklärungswert. Bezüglich der Hintergrundgeräusche sind für Sprecher 3f keine signifikanten Effekte nachweisbar.

Tabelle 7.29: Regressionsergebnisse der Skala „steif-locker“ für die abhängige Variable *evaluation*, die erklärenden Variablen *ICC* und *Geräusch* sowie deren Interaktionen.

Call:

lm(formula = evaluation ~ ICC * Geräusch, data = dataicc_steif)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.4242	0.1252	35.333	< 2e-16	***	0.09642
	ICCoohne ICC	-0.4621	0.1771	-2.610	0.009236	**	
	Geraeusch50 km/h	-1.5152	0.1771	-8.556	< 2e-16	***	
	GeraeuschStille	-1.0152	0.1771	-5.733	1.41e-08	***	
	ICCoohne ICC :Geraeusch50 km/h	1.5379	0.2504	6.141	1.30e-09	***	
	ICCoohne ICC :GeraeuschStille	0.9621	0.2504	3.842	0.000132	***	
3f	(Intercept)	3.598485	0.127740	28.170	<2e-16	***	0.01884
	ICCoohne ICC	-0.378788	0.180651	-2.097	0.0363	*	
	Geraeusch50 km/h	-0.053030	0.180651	-0.294	0.7692		
	GeraeuschStille	0.015152	0.180651	0.084	0.9332		
	ICCoohne ICC :Geraeusch50 km/h	-0.007576	0.255479	-0.030	0.9764		
	ICCoohne ICC :GeraeuschStille	-0.212121	0.255479	-0.830	0.4066		
8o	(Intercept)	4.1894	0.1227	34.134	< 2e-16	***	0.1574
	ICCoohne ICC	-1.5379	0.1736	-8.860	< 2e-16	***	
	Geraeusch50 km/h	-0.5606	0.1736	-3.230	0.00129	**	
	GeraeuschStille	-1.4091	0.1736	-8.118	1.82e-15	***	
	ICCoohne ICC :Geraeusch50 km/h	0.3864	0.2455	1.574	0.11589		
	ICCoohne ICC :GeraeuschStille	2.3333	0.2455	9.506	< 2e-16	***	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

7 Perzeptionsexperiment

Für Sprecher*in 2d und 8o ist das Regressionsmodell aus Tabelle 7.29 das mit dem besten Erklärungswert bzw. dem höchsten *Adjusted R-squared*. Für Sprecher 3f ist das Modell aus Tabelle 7.27 das am besten geeignete, da die Hintergrundgeräusche keinen signifikanten Einfluss auf die wahrgenommene Steifheit in den Stimuli dieses Sprechers ausüben. Die Tabellen 7.30 und 7.31 sowie die Abbildungen 7.11 und 7.12 fassen die transformierten Schätzwerte der besten Regressionsmodelle jeder Sprecherin bzw. jedes Sprechers zusammen, wobei die Bedingungen *ohne ICC* bzw. *ohne ICC, Stille* als jeweilige Basen den Wert 0 annehmen. Die Aktivierung des ICC-Systems führt dazu, dass Sprecher 3f für die Teilnehmer*innen des Perzeptionsexperiments steifer klingt. Ohne ICC-System bilden Sprecher*in 2d und 8o unterschiedliche Effekte hinsichtlich der Geräuschbelastung ab. 2d klingt steifer von *Stille* zu *50 km/h* und wieder etwas lockerer von *50 km/h* zu *130 km/h*. 8o zeigt jedoch einen gegenteiligen Effekt. *Mit ICC* klingen die Sprachsignale von *Stille* zu *50 km/h* bei Sprecherin 2d lockerer. Hin zu *130 km/h* werden diese jedoch als steifer wahrgenommen als in *Stille*. Sprecher 8o klingt immer steifer, je lauter es wird. Bei *130 km/h* klingen die Sprachsignale der Sprecher*innen 2d und 8o mit ICC-System steif (vgl. Abbildung 7.1).

Tabelle 7.30: Die auf Basis des Regressionsmodells aus Tabelle 7.27 errechneten transformierten Schätzwerte zur wahrgenommenen Steifheit von Sprecher 3f.

	ohne ICC	mit ICC
3f	0	0.4520

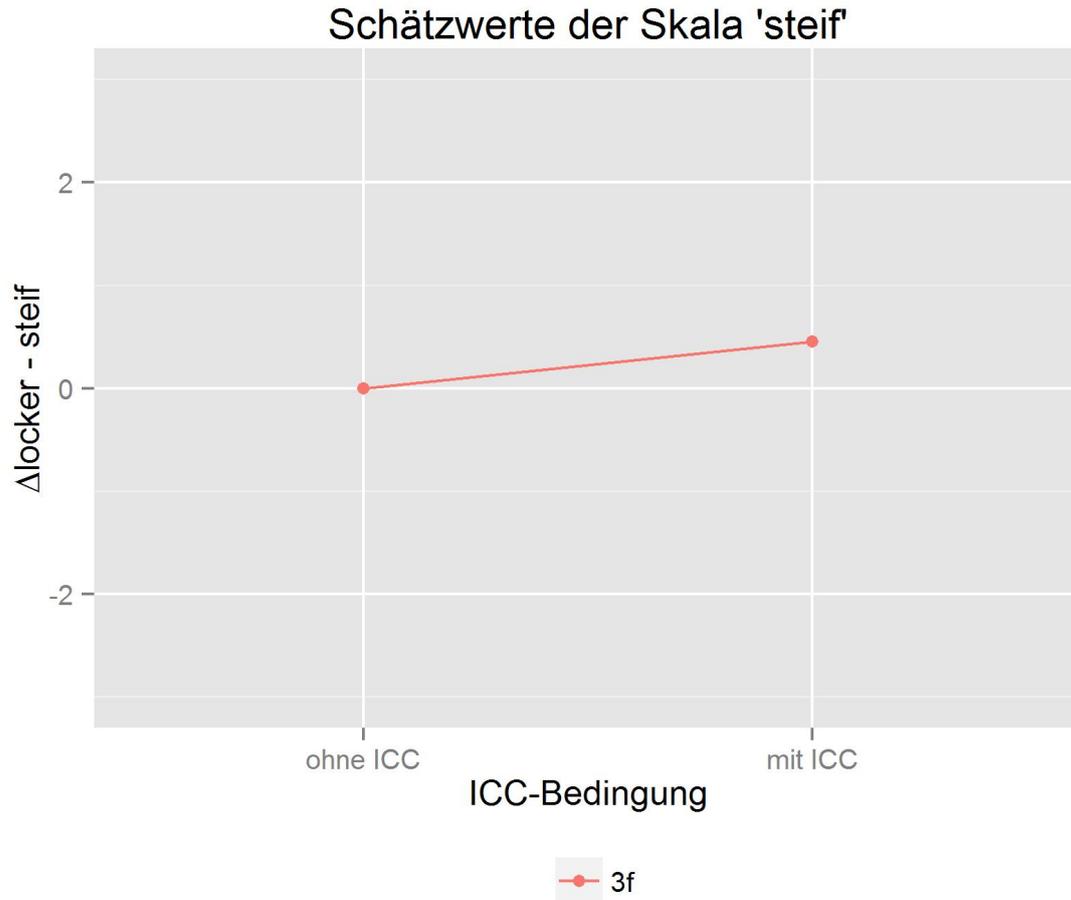


Abbildung 7.11: Die auf Basis des Regressionsmodells aus Tabelle 7.27 errechneten transformierten Schätzwerte zur wahrgenommenen Steifheit von Sprecher 3f.

Tabelle 7.31: Die auf Basis des Regressionsmodells aus Tabelle 7.29 errechneten transformierten Schätzwerte zur wahrgenommenen Steifheit von Sprecher*in 2d und 8o.

	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
2d	0	0.0758	0.0530	-0.5000	-1.0000	0.5152
8o	0	-1.0985	-0.9242	-0.7955	0.0530	0.6136

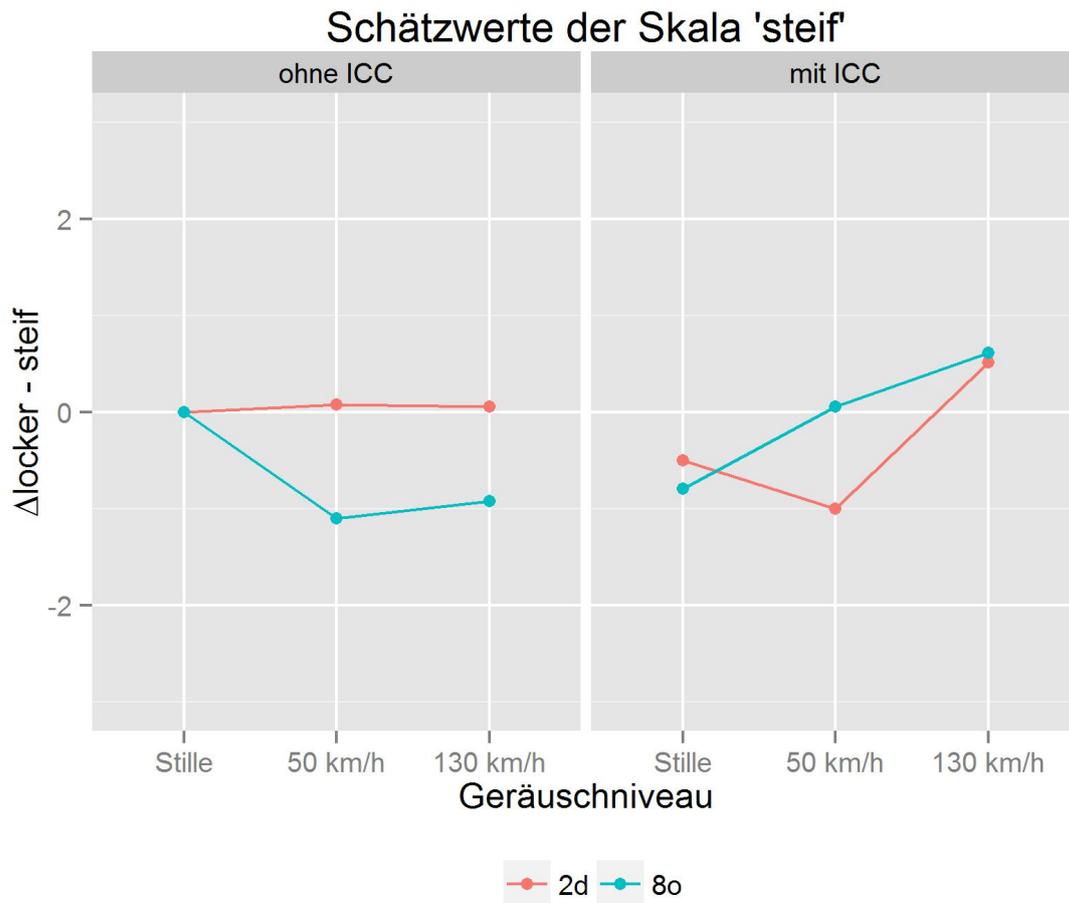


Abbildung 7.12: Die auf Basis des Regressionsmodells aus Tabelle 7.29 errechneten transformierten Schätzwerte zur wahrgenommenen Steifheit von Sprecher*in 2d und 8o.

angestrengt/entspannt

Die Aktivierung des ICC-Systems hat nach dem Regressionsmodell in Tabelle 7.32 lediglich bei Sprecher 8o einen Einfluss auf die Wahrnehmung der Sprachsignale als angestrengt oder entspannt ($p=0,0037$). Dabei klingt das Gesagte des Probanden *mit ICC* angestrengter.

Tabelle 7.32: Regressionsergebnisse der Skala „angestrengt-entspannt“ für die abhängige Variable *evaluation* und die erklärende Variable *ICC*.

Call:

```
lm(formula = evaluation ~ ICC, data = dataicc_angestrengt)
```

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)	Adjusted R ²	
2d	(Intercept)	4.79545	0.07099	67.552	<2e-16	***	-0.001186
	ICCoohne ICC	-0.02525	0.10039	-0.252	0.801		
3f	(Intercept)	3.5202	0.0784	44.900	<2e-16	***	-0.0001074
	ICCoohne ICC	-0.1061	0.1109	-0.957	0.339		
8o	(Intercept)	3.59848	0.08215	43.802	< 2e-16	***	0.009371
	ICCoohne ICC	-0.33838	0.11618	-2.913	0.00369	**	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Wird die unabhängige Variable *Geräusch* als einzige erklärende Variable in das Regressionsmodell aufgenommen, ergeben sich für alle drei Sprecher*innen hoch signifikante Effekte in allen Ausprägungen dieser Variable ($p<0,0001$) (vgl. Tabelle 7.33). Dabei klingen die Sprachsignale immer angestrengter, je lauter die Hintergrundgeräusche werden. Lediglich bei Sprecher 8o gibt es zunächst einen geringen Abfall der Messwerte von *Stille* zu *50 km/h*, bevor diese eine bedeutende Steigerung hin zu *130 km/h* erfahren.

7 Perzeptionsexperiment

Tabelle 7.33: Regressionsergebnisse der Skala „angestrengt-entspannt“ für die abhängige Variable *evaluation* und die erklärende Variable *Geräusch*.

Call:

```
lm(formula = evaluation ~ Geraeusch, data = dataicc_angestrengt)
```

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)	Adjusted R ²	
2d	(Intercept)	5.32197	0.08348	63.755	< 2e-16	***	
	Geraeusch50 km/h	-0.67045	0.11805	-5.679	1.90e-08	***	0.07708
	GeraeuschStille	-0.94697	0.11805	-8.022	3.77e-15	***	
3f	(Intercept)	3.99242	0.09326	42.811	< 2e-16	***	
	Geraeusch50 km/h	-0.69697	0.13189	-5.285	1.63e-07	***	0.05663
	GeraeuschStille	-0.87879	0.13189	-6.663	5.02e-11	***	
8o	(Intercept)	3.89015	0.09911	39.250	< 2e-16	***	
	Geraeusch50 km/h	-0.77652	0.14016	-5.540	4.12e-08	***	0.03878
	GeraeuschStille	-0.60606	0.14016	-4.324	1.73e-05	***	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Für die Untersuchung der Interaktionen zwischen *ICC* und *Geräusch* beschreibt das Regressionsmodell in Tabelle 7.34 für alle drei Sprecher*innen das höchste *Adjusted R-squared*.

Tabelle 7.34: Regressionsergebnisse der Skala „angestrengt-entspannt“ für die abhängige Variable *evaluation*, die erklärenden Variablen *ICC* und *Geräusch* sowie deren Interaktionen.

Call:

lm(formula = evaluation ~ ICC * Geräusch, data = dataicc_angestrengt)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	5.4621	0.1179	46.342	< 2e-16	***	0.07998
	ICCOhne ICC	-0.2803	0.1667	-1.682	0.0930	.	
	Geraeusch50 km/h	-0.7803	0.1667	-4.681	3.36e-06	***	
	GeraeuschStille	-1.2197	0.1667	-7.317	6.26e-13	***	
	ICCOhne ICC	0.2197	0.2357	0.932	0.3516		
	:Geraeusch50 km/h						
	ICCOhne ICC	0.5455	0.2357	2.314	0.0209	*	
3f	(Intercept)	3.92424	0.13084	29.992	< 2e-16	***	0.07149
	ICCOhne ICC	0.13636	0.18504	0.737	0.46138		
	Geraeusch50 km/h	-0.74242	0.18504	-4.012	6.59e-05	***	
	GeraeuschStille	-0.46970	0.18504	-2.538	0.01133	*	
	ICCOhne ICC	0.09091	0.26169	0.347	0.72838		
	:Geraeusch50 km/h						
	ICCOhne ICC	-0.81818	0.26169	-3.127	0.00183	**	
8o	(Intercept)	4.5530	0.1326	34.338	< 2e-16	***	0.1398
	ICCOhne ICC	-1.3258	0.1875	-7.070	3.42e-12	***	
	Geraeusch50 km/h	-1.0758	0.1875	-5.737	1.38e-08	***	
	GeraeuschStille	-1.7879	0.1875	-9.535	< 2e-16	***	
	ICCOhne ICC	0.5985	0.2652	2.257	0.0243	*	
	:Geraeusch50 km/h						
	ICCOhne ICC	2.3636	0.2652	8.913	< 2e-16	***	
:GeraeuschStille							

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

7 Perzeptionsexperiment

Tabelle 7.35 und Abbildung 7.13 zeigen die Schätzwerte des obigen Regressionsmodells (vgl. Tabelle 7.34) mit der Bedingung *ohne ICC*, *Stille* als Basis des Wertes 0. Es wird deutlich, dass die Sprachsignale von Sprecher*in 2d und 3f ohne ICC-System angestregter klingen, je lauter es wird. Sprecher 8o hört sich von *Stille* zu *50 km/h* zunächst entspannter und von *50 km/h* zu *130 km/h* wieder etwas weniger entspannt an. Bei Aktivierung des ICC-Systems werden alle Sprecher*innen als angestregter wahrgenommen, je lauter es wird. Die einzige Ausnahme dabei ist, dass Sprecher 3f von *Stille* zu *50 km/h* zunächst entspannter klingt.

Tabelle 7.35: Die auf Basis des Regressionsmodells aus Tabelle 7.34 errechneten transformierten Schätzwerte zur wahrgenommenen Angestregtheit.

	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
2d	0	0.1136	0.6742	-0.2652	0.1742	0.9546
3f	0	0.6364	1.2879	0.6818	0.4091	1.1515
8o	0	-1.0530	-0.5758	-1.0379	-0.3258	0.7500

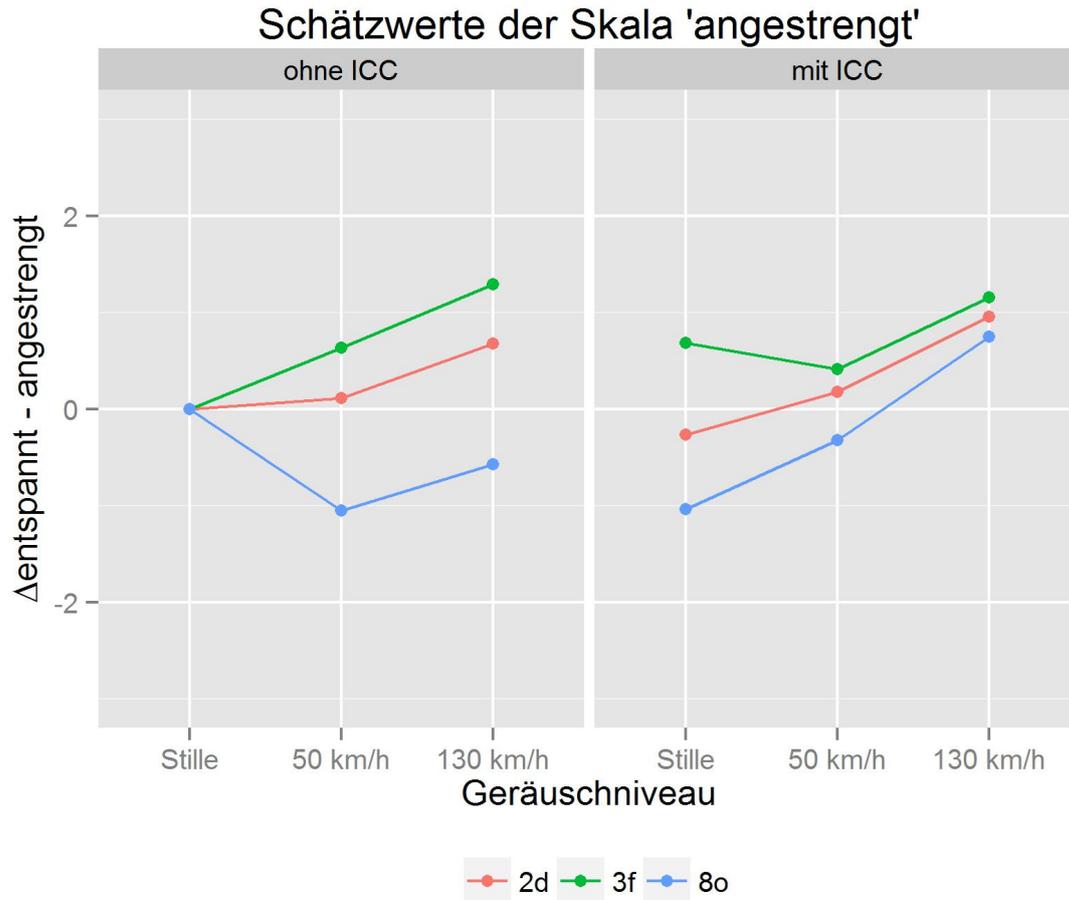


Abbildung 7.13: Die auf Basis des Regressionsmodells aus Tabelle 7.34 errechneten transformierten Schätzwerte zur wahrgenommenen Angestrengtheit.

stockend/flüssig

Die Ergebnisse in Tabelle 7.36 machen deutlich, dass die Sprachsignale aller drei Sprecher*innen mit ICC-System stockender klingen (Sprecherin 2d: $p=0,0102$; Sprecher 3f: $p=0,0001$; Sprecher 8o: $p<0,0001$).

Tabelle 7.36: Regressionsergebnisse der Skala „stockend-flüssig“ für die abhängige Variable *evaluation* und die erklärende Variable *ICC*.

Call:

lm(formula = evaluation ~ ICC, data = dataicc_stockend)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	3.87374	0.08391	46.164	<2e-16	***	0.007067
	ICCoohne ICC	-0.30556	0.11867	-2.575	0.0102	*	
3f	(Intercept)	3.19192	0.07792	40.963	< 2e-16	***	0.01696
	ICCoohne ICC	-0.42172	0.11020	-3.827	0.00014	***	
8o	(Intercept)	3.73990	0.07777	48.087	< 2e-16	***	0.07399
	ICCoohne ICC	-0.88131	0.10999	-8.013	4.02e-15	***	

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Nach Ersetzen der erklärenden Variablen *ICC* durch *Geräusch* zeigt sich ein unbeständigeres Bild. Dabei ergibt sich lediglich für Sprecher*in 2d und 3f ein höheres *Adjusted R-squared*. Die Sprachsignale von 3f werden mit zunehmender Geräuschbelastung immer flüssiger ($p<0,0001$), während das Gesprochene bei 2d von *Stille* zu *50 km/h* flüssiger und von *50 km/h* zu *130 km/h* am stockendsten klingt ($p<0,0001$ und $p=0,0419$). Sprecher 8o zeigt für den Unterschied zwischen *Stille* und *130 km/h* keinen signifikanten Effekt auf, jedoch für die Differenz zwischen *50 km/h* und *130 km/h* ($p=0,0002$), wobei das Gesprochene von *50 km/h* zu *130 km/h* stockender klingt.

Tabelle 7.37: Regressionsergebnisse der Skala „stockend-flüssig“ für die abhängige Variable *evaluation* und die erklärende Variable *Geräusch*.

Call:

lm(formula = evaluation ~ Geraeus, data = dataicc_stockend)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.0871	0.1012	40.387	< 2e-16	***	0.03724
	Geraeus50 km/h	-0.8068	0.1431	-5.637	2.4e-08	***	
	GeraeusStille	-0.2917	0.1431	-2.038	0.0419	*	
3f	(Intercept)	2.46212	0.09324	26.405	< 2e-16	***	0.06158
	Geraeus50 km/h	0.59848	0.13187	4.539	6.55e-06	***	
	GeraeusStille	0.95833	0.13187	7.268	8.82e-13	***	
8o	(Intercept)	3.50758	0.09812	35.746	< 2e-16	***	0.01732
	Geraeus50 km/h	-0.52273	0.13877	-3.767	0.000178	***	
	GeraeusStille	-0.10227	0.13877	-0.737	0.461341		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Auch bei dieser Bewertungsskala weist das Regressionsmodell mit den Interaktionen zwischen *ICC* und *Geräusch* wiederum für alle drei Sprecher*innen das höchste *Adjusted R-squared* auf (vgl. Tabelle 7.38).

7 Perzeptionsexperiment

Tabelle 7.38: Regressionsergebnisse der Skala „stockend-flüssig“ für die abhängige Variable *evaluation*, die erklärenden Variablen *ICC* und *Geräusch* sowie deren Interaktionen.

Call:

lm(formula = evaluation ~ ICC * Geräusch, data = dataicc_stockend)

Sprecher	Ausprägungen	Estimate	Std. Error	t value	Pr(> t)		Adjusted R ²
2d	(Intercept)	4.2045	0.1410	29.824	< 2e-16	***	0.06581
	ICCoohne ICC	-0.2348	0.1994	-1.178	0.23918		
	Geraeusch50 km/h	-0.4394	0.1994	-2.204	0.02782	*	
	GeraeuschStille	-0.5530	0.1994	-2.774	0.00567	**	
	ICCoohne ICC :Geraeusch50 km/h	-0.7348	0.2820	-2.606	0.00933	**	
	ICCoohne ICC :GeraeuschStille	0.5227	0.2820	1.854	0.06412	.	
3f	(Intercept)	2.3712	0.1279	18.543	< 2e-16	***	0.1175
	ICCoohne ICC	0.1818	0.1808	1.005	0.315		
	Geraeusch50 km/h	0.7652	0.1808	4.231	2.6e-05	***	
	GeraeuschStille	1.6970	0.1808	9.384	< 2e-16	***	
	ICCoohne ICC :Geraeusch50 km/h	-0.3333	0.2557	-1.303	0.193		
	ICCoohne ICC :GeraeuschStille	-1.4773	0.2557	-5.776	1.1e-08	***	
8o	(Intercept)	4.6061	0.1258	36.605	< 2e-16	***	0.192
	ICCoohne ICC	-2.1970	0.1780	-12.346	< 2e-16	***	
	Geraeusch50 km/h	-1.2424	0.1780	-6.982	6.20e-12	***	
	GeraeuschStille	-1.3561	0.1780	-7.620	7.28e-14	***	
	ICCoohne ICC :Geraeusch50 km/h	1.4394	0.2517	5.720	1.52e-08	***	
	ICCoohne ICC :GeraeuschStille	2.5076	0.2517	9.964	< 2e-16	***	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Tabelle 7.39 und Abbildung 7.14 stellen die aus dem besten Regressionsmodell 7.38 errechneten Schätzwerte dar, wobei die Bedingung *ohne ICC*, *Stille* wiederum von allen Werten subtrahiert wird und als Basis den Wert 0 annimmt. Es lässt sich feststellen, dass die verschiedenen Hintergrundgeräusche ohne ICC-System andersartige Effekte auf alle drei Sprecher*innen haben. Sprecherin 2d klingt von *Stille* zu *50 km/h* flüssiger und von *50 km/h* zu *130 km/h* etwas stockend, während Sprecher 3f einen gegenteiligen Effekt aufweist. Allein die Sprachsignale von Sprecher 8o werden mit zunehmender Geräuschbelastung als flüssiger wahrgenommen. Mit ICC-System sprechen 2d und 8o mit zunehmenden Fahrgeräuschen immer weniger flüssig, wobei sie bei *130 km/h* sogar stockend klingen. Sprecher 3f hingegen wird mit steigender Geräuschbelastung als immer flüssiger sprechend wahrgenommen.

Tabelle 7.39: Die auf Basis des Regressionsmodells aus Tabelle 7.38 errechneten transformierten Schätzwerte zur wahrgenommenen stockenden Sprechweise.

	ohne ICC			mit ICC		
	Stille	50 km/h	130 km/h	Stille	50 km/h	130 km/h
2d	0	-1.1439	0.0303	-0.2879	-0.1742	0.2652
3f	0	0.2121	-0.2197	1.2955	0.3636	-0.4015
8o	0	-0.9546	-1.1515	-0.3106	-0.1970	1.0455

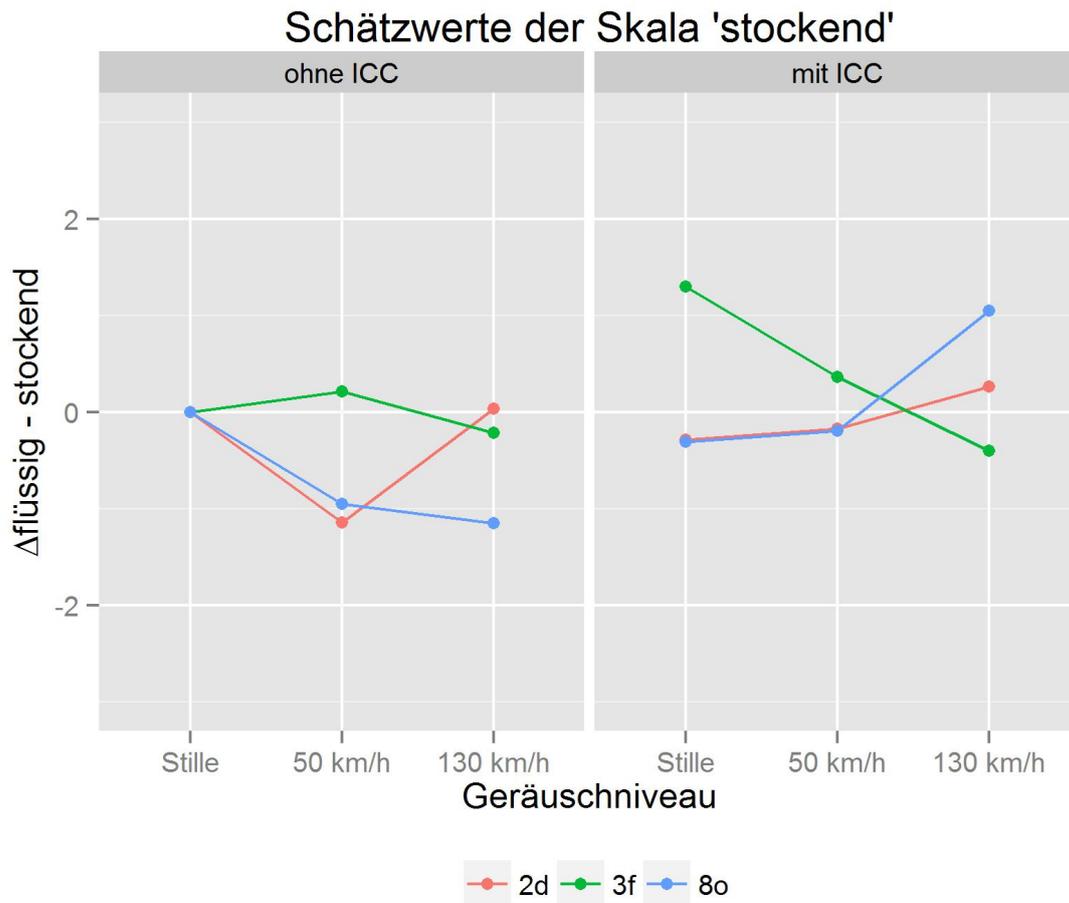


Abbildung 7.14: Die auf Basis des Regressionsmodells aus Tabelle 7.38 errechneten transformierten Schätzwerte zur wahrgenommenen stockenden Sprechweise.

8 Zusammenfassung

Im Folgenden werden die Ergebnisse der vorliegenden Dissertationsarbeit diskutiert. Das Fazit stellt anschließend eine Übersicht der Resultate dar. Auf dieser Grundlage sind Aussagen zur qualitativen Veränderung der Sprachkommunikation im Fahrzeug durch Anwendung des untersuchten ICC-Systems möglich. Weiterhin erfolgen eine Methodenkritik sowie ein Ausblick auf weiterführende Untersuchungen.

8.1 Diskussion der Ergebnisse

Die deskriptiven und inferenzstatistischen Analysen des Produktions- und Perzeptionsexperiments dieser Forschungsarbeit (vgl. Kapitel 6.3 und 7.6) werden im Folgenden zusammengefasst und interpretiert. Dies geschieht durch die Beantwortung der in Kapitel 1.2 beschriebenen Teilfragen sowie durch die Überprüfung der in Kapitel 3.4 und 7.2 aufgestellten Hypothesen. Teilfrage (1) und (2) beziehen sich auf die Analyse der Sprachsignale des SPID-Korpus, während Teilfrage (3) auf die Ergebnisse der Umfragedaten eingeht.

8.1.1 Teilfrage (1)

1. Wie verändert sich die Kommunikation im Auto unter Lombard-Bedingungen, d. h., welche Modifikationen der Sprachproduktion finden unter dem Einfluss verschiedener Fahrgeräusche statt?

Diese Teilfrage bezieht sich auf die Analyse der ohne ICC-System aufgenommenen Sprachsignale des SPID-Korpus. Die Auswirkungen der unterschiedlichen Fahrgeräusche auf die vier abhängigen Variablen durchschnittliche Intensität, durchschnittliche Grundfrequenz, Signaldauer und Anteil der Sprachproduktion (vgl. Kapitel 6.3) ohne und mit ICC-System werden anhand der Differenzen zwischen den Messwerten

der einzelnen Geräuschbedingungen deutlich. Diese werden aus den Schätzwerten der Untersuchungsparameter in Tabelle 6.7, 6.16, 6.22 und 6.31 berechnet und sind in Tabelle 8.1 aufgeführt. Positive Differenzen zeigen eine Steigerung, negative Differenzen eine Abnahme der Messwerte bei zunehmender Geräuschbelastung auf. Die Inhalte der Tabelle 8.1 werden nachfolgend für jeden der vier Untersuchungsparameter noch einmal beschrieben. Da die Differenzen der Messwerte von hinten und vorne aufgenommenen Sprachsignalen für die durchschnittliche Intensität und die durchschnittliche Grundfrequenz jeweils gleich sind, werden diese in Tabelle 8.1 zusammengefasst.

Tabelle 8.1: Die Veränderung der Messwerte aller abhängigen Variablen bei Zunahme der Geräuschbelastung. Grundlage der Berechnungen bilden die Tabellen 6.7, 6.16, 6.22 und 6.31. a) und b) geben die Differenzen der Messwerte zwischen zwei Geräuschbedingungen an, wobei jeweils der Messwert der niedrigeren Geräuschbedingung von dem der höheren subtrahiert wird.

Parameter	Sitzplatz	ICC	Rolle	a) Differenz 50km/h-Stille	b) Differenz 130km/h-50km/h
Intensität	beide	ohne ICC	beide	3,6169 dB	3,1756 dB
Intensität	beide	mit ICC	beide	2,6713 dB	2,0806 dB
Grundfrequenz	beide	ohne ICC	beide	14,548 Hz	17,644 Hz
Grundfrequenz	beide	mit ICC	beide	9,275 Hz	11,706 Hz
Signaldauer	beide	ohne ICC	beide	-59,998 Sek.	-77,607 Sek.
Signaldauer	beide	mit ICC	beide	19,271 Sek.	24,291 Sek.
Sprachprod.	hinten	beide	Follower	-2,68 %	-3,44 %
Sprachprod.	hinten	beide	Giver	1,64 %	-2,47 %
Sprachprod.	vorne	beide	Follower	-1,89 %	0,76 %
Sprachprod.	vorne	beide	Giver	2,44 %	1,73 %

Intensität

Hinsichtlich der durchschnittlichen Intensität der untersuchten Sprachsignale wird aus Tabelle 8.1 deutlich, dass ohne ICC-System ein Lombard-Effekt stattfindet. Die Messwerte dieses Parameters steigen mit Zunahme der Hintergrundgeräusche an,

was sich in Tabelle 8.1 als positive Differenz zwischen den einzelnen Geräuschbedingungen widerspiegelt. Aufgrund dieser Ergebnisse kann Hypothese (1.) bestätigt werden (vgl. Kapitel 3.4). Dabei ist die Differenz zwischen Stille und 50 km/h etwas größer als zwischen 50 km/h und 130 km/h. Das Ausmaß des Lombard-Effekts fällt demnach bei steigender Geräuschbelastung um 0,4413 dB (3,6169 dB - 3,1756 dB) ab bzw. bleibt annähernd konstant.

Grundfrequenz

Die durchschnittliche Grundfrequenz der Sprachsignale weist ohne ICC-System ebenfalls einen Lombard-Effekt, d. h. eine Zunahme der Messwerte bei steigender Geräuschbelastung, auf. Deshalb kann Hypothese (2.) ebenfalls bestätigt werden (vgl. Kapitel 3.4). Im Gegensatz zur durchschnittlichen Intensität nimmt das Ausmaß des Phänomens bei anwachsender Geräuschbelastung jedoch zu, da die Differenz zwischen 50 km/h und 130 km/h gegenüber der zwischen Stille und 50 km/h um 3,096 Hz (17,644 Hz - 14,548 Hz) größer ist (vgl. Tabelle 8.1).

Signaldauer

Die Dauer der ohne ICC-System aufgenommenen Dialoge bestätigt ebenfalls das Auftreten des Lombard-Effekts, da die Signaldauer mit steigender Geräuschbelastung abfällt. Hypothese (3.) wird ebenfalls bestätigt (vgl. Kapitel 3.4). Dabei ist die Differenz zwischen Stille und 50 km/h um 17,609 Sek. geringer (-77,607 Sek. - (-59,998 Sek.)) als zwischen 50 km/h und 130 km/h (vgl. Tabelle 8.1). Der Lombard-Effekt steigt demnach, wie im Falle der durchschnittlichen Grundfrequenz, mit wachsendem Geräuschniveau an.

Sprachproduktionsleistung

Aufgrund der Tatsache, dass die Veränderung der Geräusch- und ICC-Bedingungen allein keinen Einfluss auf die abhängige Variable der Sprachproduktionsleistung hat, bestätigt sich die in Kapitel 6.3.4 angestellte Vermutung, dass sich der Anteil der produzierten Sprache zweier Dialogpartner*innen größtenteils gegenseitig bedingt. Lediglich eine Trennung der Sprecher*innen anhand von Sitzplatz oder Sprechrolle zeigt einen Einfluss auf die Messwerte des untersuchten Parameters auf, denn

vorne und hinten sitzende Sprecher*innen bzw. *Follower* und *Giver* beeinflussen sich gegenseitig in ihrer Sprachproduktionsleistung. Es lässt sich jedoch feststellen, dass sich das Verhältnis der Sprachproduktionsleistung zweier Dialogpartner*innen ohne ICC-System durch ansteigende Hintergrundgeräusche gegenteilig verschiebt. Dazu werden in Tabelle 8.1 jeweils die hinten sitzenden *Follower* mit den vorne sitzenden *Givern* sowie die vorne sitzenden *Follower* mit den hinten sitzenden *Givern* verglichen. Hypothese (4.) aus Kapitel 3.4 kann nicht bestätigt werden, da die Sprachproduktionsleistung der Dialogpartner*innen bei zunehmenden Fahrgeräuschen nicht nur zu-, sondern auch abnimmt und dabei nicht für jede Person eine gleichbleibende Effektrichtung aufzeigt. Aufgrund der breiten Arbeitsdefinition des Lombard-Effekts (vgl. Kapitel 3.4) kann das Auftreten dieses Phänomens jedoch für den Parameter der Sprachproduktionsleistung bestätigt werden.

Da die Fahrgeräusche allein keinen signifikanten Einfluss auf diesen Parameter haben, zeigen die Messwerte in Tabelle 8.1 lediglich die Niveauunterschiede der Sprachproduktionsleistung zwischen den beiden Sitzplätzen bzw. Sprechrollen auf. Ohne und mit ICC-System sprechen *Giver* in Stille weniger als *Follower*. Mit Auftreten der Hintergrundgeräusche ab 50 km/h kehrt sich dieses Verhältnis um. (vgl. Abbildung 6.26 und 6.27) Die Sprachproduktionsleistung auf dem hinteren Sitzplatz wird weiterhin von 50 km/h zu 130 km/h stärker beeinflusst als von der stillen Referenzbedingung zu 50 km/h. Die Differenz zwischen Stille und 50 km/h ist dabei für den *Follower* um 0,76 % (-3,44 % - (-2,68 %)) und für den *Giver* um 4,11 % (-2,47 % - 1,64 %) niedriger als die Differenz zwischen 50 km/h und 130 km/h. Auf dem vorderen Sitzplatz ergibt sich ein gegenteiliges Bild. Hierbei wird der Lombard-Effekt mit steigender Geräuschbelastung abgeschwächt, da die Differenz zwischen Stille und 50 km/h für den *Follower* um 2,65 % (-1,89 % - 0,76 %) und für den *Giver* um 0,71 % (2,44 % - 1,73 %) größer ist als die Differenz zwischen 50 km/h und 130 km/h. (vgl. Tabelle 8.1)

Ergebnisinterpretation

Aus den in Kapitel 6.3 gewonnenen Ergebnissen geht hervor, dass mit dem Auftreten und der Steigerung von Fahrgeräuschen ein Lombard-Effekt in allen Untersuchungsparametern hervorgerufen wird. Dabei findet eine Zunahme der durchschnittlichen Intensität und der durchschnittlichen Grundfrequenz sowie eine Reduktion der Signaldauer statt. Diese Resultate bestätigen die in der Forschungsliteratur beschriebenen Eigenschaften des Lombard-Effekts (vgl. Kapitel 3). Auch hinsichtlich der Sprachproduktionsleistung lässt sich feststellen, dass mit der Steigerung der Umgebungsgeräusche eine Veränderung der Messwerte einhergeht. Die prozentuale Sprachproduktion verschiebt sich dabei gegenteilig für die jeweiligen Dialogpartner*innen, wie z. B. hinten sitzende *Giver* und vorne sitzende *Follower*.

8.1.2 Teilfrage (2)

2. Wie verändert sich der Lombard-Effekt durch den Einsatz eines ICC-Systems?

Die zweite Teilfrage nimmt die Veränderungen der Sprachproduktion bei lauten Fahrgeräuschen und der Aktivierung des ICC-Systems in den Blick. Neben der Analyse von Messwerten erfolgt eine Interpretation der Ergebnisse dahingehend, ob der Einsatz des ICC-Systems eine Verbesserung der Kommunikationssituation im fahrenden Auto bewirkt.

In Tabelle 8.2 werden Unterschiede in den Messwerten zwischen den ICC-Bedingungen in jeder Geräuschbedingung aufgezeigt, welche ebenfalls auf Grundlage der Tabellen 6.7, 6.16, 6.22 und 6.31 berechnet werden. Durch das arithmetische Mittel der letzten Spalte werden positive oder negative Niveauverschiebungen der Messwerte deutlich.

Tabelle 8.2: Die Veränderung der Messwerte bei Aktivierung des ICC-Systems in allen Geräuschbedingungen. Grundlage der Berechnungen bilden die Tabellen 6.7, 6.16, 6.22 und 6.31. Die letzte Spalte beinhaltet die arithmetischen Mittel der vorangegangenen drei Spalten.

Parameter	Sitzpl.	Rolle	Differenz mit ICC - ohne ICC			Ø
			Stille	50 km/h	130 km/h	
Intensität	hinten	beide	-4,0421 dB	-4,9877 dB	-6,0827 dB	-5,0375 dB
Intensität	vorne	beide	0,6296 dB	-0,316 dB	-1,411 dB	-0,3658 dB
Grundfr.	hinten	beide	-5,779 Hz	-11,052 Hz	-16,99 Hz	-11,2737 Hz
Grundfr.	vorne	beide	0,785 Hz	-4,488 Hz	-10,426 Hz	-4,7097 Hz
Signaldauer	beide	beide	-292,163 Sek.	-212,894 Sek.	-110,996 Sek.	-205,351 Sek.
Sprachprod.	hinten	beide	-2,34 %	-2,34 %	-2,33 %	-2,3367 %
Sprachprod.	vorne	beide	4,23 %	4,22 %	4,22 %	4,2233 %

2.1 Kann die Verwendung eines ICC-Systems eine Rückverschiebung der sprachlichen Modifikationen hin zu einer Situation ohne Lombard-Umgebung bewirken?

Die Anpassung an eine Sprechsituation mit aktiviertem ICC-System fällt z. T. je nach Sitzplatz der Proband*innen unterschiedlich aus. Die Messwerte der Intensität sowie der Grundfrequenz sind ohne ICC-System auf dem hinteren Sitzplatz höher und mit ICC-System auf dem vorderen Sitzplatz höher. Zudem hat die Anwendung des ICC-Systems einen stärkeren Einfluss auf den hinteren Sitzplatz. Bei den Parametern der Intensität und der Grundfrequenz sind die Unterschiede zwischen den beiden ICC-Bedingungen hinten größer (vgl. Kapitel 6.3.1 und 6.3.2). Diese Ergebnisse spiegeln die Tatsache wider, dass das ICC-System während der Aufnahmen für die Vermittlung der Sprachsignale von vorne nach hinten stärker eingestellt war als andersherum. (vgl. Kapitel 2.1 und [39, S. 1488])

Intensität

Durch die Aktivierung des ICC-Systems (vgl. Kapitel 2) wird der Lombard-Effekt für den Parameter der durchschnittlichen Intensität nicht aufgehoben. Aus Tabelle

8.1 lässt sich jedoch erkennen, dass die Differenzen zwischen den einzelnen Geräuschbedingungen mit ICC-System wesentlich geringer ausfallen als ohne. Der Lombard-Effekt wird demnach durch Anwendung des ICC-Systems abgeschwächt. Dies wird durch die Werte in Tabelle 8.2 zusätzlich bestätigt, da die Differenzen zwischen den beiden ICC-Bedingungen mit steigendem Hintergrundlärm hinten und vorne in fast jeder Geräuschbedingung größer werden.

Bei Subtraktion der Werte ohne ICC von den Werten mit ICC zeigen negative Differenzen an, dass die durchschnittliche Intensität der Sprachsignale mit ICC-System niedriger ist als ohne (vgl. Tabelle 8.2). Dies gilt für alle Experimentbedingungen bis auf die Kombination *Stille, vorne*. D. h., die Proband*innen sprechen mit ICC-System leiser, da sie den vorhandenen Umgebungslärm stimmlich weniger stark übertönen müssen. Dabei besteht zwischen dem hinteren und dem vorderen Sitzplatz eine Differenz von 4,6717 dB ($-5,0375 \text{ dB} - (-0,3658 \text{ dB})$). Demnach ist die Niveauverschiebung der durchschnittlichen Intensität mit ICC-System über alle Geräuschbedingungen hinweg auf dem vorderen Sitzplatz geringer.

Grundfrequenz

Für die durchschnittliche Grundfrequenz wird mit ICC-System ebenfalls eine Reduktion des Lombard-Effekts erreicht, da auch hier die Differenzen der Messwerte zwischen den Geräuschbedingungen mit ICC-System deutlich geringer ausfallen (vgl. Tabelle 8.1). Zudem wird dieses Ergebnis in Tabelle 8.2 dadurch bestätigt, dass die Differenzen zwischen den Bedingungen ohne und mit ICC-System bei zunehmenden Fahrgeräuschen auf beiden Sitzplätzen ansteigen.

Tabelle 8.2 zeigt durch negative Differenzen auf, dass die Messwerte mit ICC-System niedriger sind als ohne. Wie für die durchschnittliche Intensität bildet auch für diesen Parameter die Bedingung *Stille, vorne* eine Ausnahme. Die Ergebnisse zeigen, dass die Dialogpartner*innen bei Aktivierung des ICC-Systems mit einer im Vergleich zur Situation ohne ICC-System tieferen Stimme gesprochen haben. Weiterhin besteht zwischen dem hinteren und dem vorderen Sitzplatz eine Differenz des arithmetischen Mittels von 6,564 Hz ($-11,2737 \text{ Hz} - (-4,7097 \text{ Hz})$). Wie im Falle der durchschnittlichen Intensität ist der Niveauunterschied der Messwerte zwischen den ICC-Bedingungen auch für diesen Parameter über alle Geräuschbedingungen hinweg vorne geringer.

Signaldauer

Hinsichtlich der Signaldauer ist mit ICC-System kein Lombard-Effekt wie in den Analysen ohne ICC-System (vgl. Kapitel 6.3) und den Ergebnissen von Mixdorff et al. 2007 (vgl. [53]) mehr nachweisbar, da die Messwerte mit zunehmenden Fahrgeräuschen nicht mehr abfallen, sondern im Gegenteil sogar ansteigen (vgl. Tabelle 8.1).

Tabelle 8.2 zeigt durch negative Differenzen auf, dass die Signaldauer mit ICC-System wesentlich niedriger ist als ohne. Demnach benötigen die Sprecher*innen mit ICC-System weniger Zeit zum Lösen der Kartenaufgaben. Bei steigendem Geräuschniveau nähern sich die Werte der ICC-Bedingungen für die Signaldauer einander an, d. h., die Differenzen werden geringer. Die in Kapitel 6.3.3 durchgeführten t-Tests zeigen jedoch, dass die Unterschiede in der Signaldauer mit ICC-System zwischen Stille und 50 km/h bzw. 50 km/h und 130 km/h nicht signifikant sind. Lediglich für die Differenz zwischen den Geräuschbedingungen Stille und 130 km/h gibt es einen signifikanten Trend ($p=0,0506$). D. h., es findet mit Aktivierung des ICC-Systems kein eindeutiger Effekt statt, denn die Dialoge der Proband*innen werden mit zunehmender Geräuschbelastung nicht signifikant länger oder kürzer.

Sprachproduktionsleistung

Die Sprachproduktionsleistung der einzelnen Dialogpartner*innen wird, wie oben beschrieben, durch die Anwendung des ICC-Systems allein nicht signifikant beeinflusst. In Interaktion mit den Sitzplätzen der Sprecher*innen finden jedoch deutliche Veränderungen statt. Dabei werden die Messwerte des hinteren Sitzplatzes mit ICC-System reduziert, während die des vorderen Sitzplatzes ansteigen. Die Werte verschieben sich demnach gegenteilig zueinander (vgl. Tabelle 8.2).

Der ohne ICC beobachtete Lombard-Effekt wird mit Aktivierung des ICC-Systems nicht aufgehoben, da sich die Differenzen der Messwerte ohne und mit ICC in Tabelle 8.2 bei steigender Geräuschbelastung nicht bzw. kaum verändern. Es finden lediglich Niveauverschiebungen der Messwerte mit ICC-System über alle Geräuschbedingungen hinweg statt. Diese verlaufen auf jedem Sitzplatz in unterschiedliche Richtungen. Dabei wird mit Anwendung des ICC-Systems vorne mehr Sprache produziert (4,2233 %) und hinten weniger (-2,3367 %).

Ergebnisinterpretation

Hinsichtlich der durchschnittlichen Intensität und der durchschnittlichen Grundfrequenz wird eine Verringerung im Ausmaß des Lombard-Effekts beobachtet, im Falle der Signaldauer sogar die vollständige Aufhebung des Phänomens. Daher kann Hypothese (5.) für diese drei Parameter bestätigt werden (vgl. Kapitel 3.4). Diese Ergebnisse zeigen, dass durch die Aktivierung des ICC-Systems tatsächlich eine Rückverschiebung der sprachlichen Modifikationen hin zu einer Situation ohne Lombard-Umgebung erzielt werden kann. Hinzu kommt jedoch, dass mit ICC-System eine generelle Niveaushiftung der Messwerte stattfindet, wobei Intensität und Grundfrequenz stark abfallen, besonders auf dem hinteren Sitzplatz. Auch die Dauer der produzierten Sprachsignale erfährt mit ICC-System eine starke Reduktion in allen drei Geräuschbedingungen. Für die Sprachproduktionsleistung wird der zuvor beobachtete Lombard-Effekt durch das ICC-System nicht aufgehoben, wodurch Hypothese (5.) für diesen Parameter nicht bestätigt werden kann (vgl. Kapitel 3.4). Dabei finden jedoch wiederum Niveaushiftungen der Messwerte statt. Mit ICC wird vorne mehr Sprache produziert, während auf dem hinteren Sitzplatz weniger gesprochen wird.

Die vorliegenden Ergebnisse legen die Interpretation nahe, dass durch Anwendung des ICC-Systems nicht die Gesprächssituation des stillen Autoinnenraums wiederhergestellt wird, sondern eine neue Kommunikationsumgebung entsteht, welche ebenfalls eine Reduktion bzw. Aufhebung des Lombard-Effekts beinhaltet, der stillen Referenzbedingung jedoch nicht in allen Parameterausprägungen gleicht. Hypothese (6.) wird daher verworfen (vgl. Kapitel 3.4).

- 2.2 Führt der Einsatz eines ICC-Systems zu einer Verbesserung der Kommunikation? Kann sich ein ICC-System störend auf die Kommunikation der Dialogpartner*innen auswirken?

Intensität, Grundfrequenz und Signaldauer

Die Abschwächung des Lombard-Effekts in den Parametern der durchschnittlichen Intensität und der durchschnittlichen Grundfrequenz sowie die Aufhebung dieses Phänomens hinsichtlich der Signaldauer lassen darauf schließen, dass das ICC-System eine Strategie zur Kompensierung des Lombard-Effekts darstellt (vgl. Kapi-

tel 3, 3.2 und 3.2.3).

Durch t-Tests wird gezeigt, dass die Unterschiede in der durchschnittlichen Intensität zwischen den Bedingungen ohne und mit ICC-System in der stillen Referenzbedingung hinten signifikant sind, jedoch nicht vorne (vgl. Kapitel 6.3.1). Ein ähnliches Verhältnis lässt sich für die durchschnittliche Grundfrequenz feststellen (vgl. Kapitel 6.3.2), wobei für den Unterschied zwischen ohne und mit ICC in Stille hinten lediglich ein schwach signifikanter Trend ($p=0,1144$) nachgewiesen werden kann. Dennoch unterstützen diese Ergebnisse grundsätzlich die der Intensität und zeigen dieselben Effektrichtungen auf. Diese Resultate zeigen, dass das ICC-System bereits in einer stillen Situation die ungünstige Kommunikationsumgebung im Autoinnenraum kompensiert, auch wenn noch keine Hintergrundgeräusche vorhanden sind. Dies gilt jedoch nur für den hinteren Sitzplatz, weil das ICC-System hierfür stärker eingestellt war als für den vorderen Sitzplatz. Ein weiterer Grund für diese Ergebnisse könnte sein, dass eine nachteiligere Abstrahlrichtung des Sprachschalls bereits ohne erhöhte Umgebungsgeräusche auf dem hinteren Sitzplatz eine Zunahme des *vocal effort* bewirkt (vgl. Kapitel 3.1). Dieser gesteigerte Stimm Aufwand könnte dann durch die Verwendung des ICC-Systems kompensiert werden. (vgl. [38, S. 481f.] und [39, S. 1488]) Das ICC-System würde demnach nicht nur der Kompensation von Kommunikationsnachteilen dienen, die durch laute Umgebungsgeräusche entstehen, sondern auch solchen, die sich aufgrund einer ungünstigen akustischen und visuellen Kommunikationssituation im Auto ergeben (vgl. Kapitel 2). Hypothese (7.) muss demnach für die Parameter der Intensität und der Grundfrequenz verworfen werden (vgl. Kapitel 3.4). Dasselbe gilt für den Parameter der Signaldauer, wobei die Dialoge aufgrund der Niveauverschiebungen der Messwerte mit ICC-System in Stille wesentlich kürzer sind als ohne ICC-System.

Sprachproduktionsleistung

Da die ungünstige Abstrahlrichtung des Sprachschalls für die hinteren Sitzplätze größere Schwierigkeiten bei der lautsprachlichen Kommunikation bedeutet und das ICC-System deshalb die Sprache der Proband*innen von vorne nach hinten stärker weitergeleitet hat (vgl. Kapitel 2), sind die Veränderungen auf dem hinteren Sitzplatz von größerer Bedeutung für die Effekte des ICC-Systems. Unabhängig von ihrer Rolle sprechen hinten sitzende Proband*innen mehr als vorne sitzende. Die-

ses Phänomen kehrt sich durch das ICC-System um, die Sprachproduktionsleistung wird auf den vorderen Sitzplatz verschoben (vgl. Kapitel 6.3.4). Die Untersuchung der vorangegangenen drei Parameter ergab eine Verbesserung der lautsprachlichen Kommunikation durch das ICC-System. Dieses Resultat lässt sich auf die Ergebnisse der Sprachproduktionsleistung übertragen, auch wenn hinsichtlich dieses Untersuchungsparameters keine Reduktion oder Aufhebung des Lombard-Effekts festgestellt werden kann. Durch eine Senkung des Stressniveaus der im Auto sitzenden Dialogpartner*innen können hinten sitzende *Giver* präzisere und damit kürzere Aussagen treffen, während hinten sitzende *Follower* mehr verstehen und daher weniger nachfragen müssen. Da sich für diesen Untersuchungsparameter die Messwerte der Dialogpartner*innen jeweils gegenseitig bedingen, steigt dementsprechend die Sprachproduktion auf dem vorderen Sitzplatz an.

Unter Berücksichtigung der einzelnen Sitzplätze und Sprechrollen hat das ICC-System auch in der stillen Referenzbedingung einen Einfluss auf die Sprachproduktionsleistung. Daher wird Hypothese (7.) auch für den vierten Parameter der akustisch-phonetischen Analyse verworfen (vgl. Kapitel 3.4).

Der Lerneffekt

Für den Parameter der Signaldauer wird mit ICC-System eine vollständige Aufhebung des zuvor beobachteten Lombard-Effekts festgestellt. Es bleibt jedoch die Frage bestehen, weshalb die Signaldauer durch Anwendung des ICC-Systems in allen drei Geräuschbedingungen wesentlich kürzer ist als ohne ICC. Der Grund hierfür könnte sein, dass nach der ersten Aufnahmesitzung ohne ICC-System ein Lerneffekt bei den Proband*innen eintrat, weil Diskussionen über Art und Durchführung der Sprechaufgabe, über zuvor aufgetretene Missverständnisse und sonstige Verständnisfragen in der ersten Aufnahmesitzung bereits ausreichend geklärt worden sind und daher in einem späteren Durchlauf nicht erneut aufgegriffen werden mussten. Daher tritt trotz der zwischen den Aufnahmesitzungen liegenden Pause eine vom Umgebungslärm und vom ICC-System unabhängige Verkürzung der Dialogdauer auf. Es gilt jedoch zu beachten, dass dieser Lerneffekt nicht innerhalb der ersten oder zweiten Aufnahmesitzung vorkommen konnte, da die Reihenfolge der Geräuschbedingungen innerhalb jeder Aufnahmesitzung für alle Dialogpaare randomisiert wurde (vgl. Kapitel 5.2). Dennoch kann auch die Anwendung des ICC-Systems der Grund für dieses

Phänomen sein, ebenso wie ein Placebo-Effekt, welcher durch die Kenntnis der Proband*innen über die Aktivierung des ICC-Systems zustande gekommen sein könnte. (vgl. [39, S. 1488]) Die oben beschriebene Verschiebung der Sprachproduktionsleistung von hinten sitzenden Sprecher*innen auf vorne sitzende Personen kann nicht durch einen Lerneffekt entstehen, da diese Niveauverschiebung je nach Sitzplatz der Proband*innen unterschiedlich ausfällt. Der Grund hierfür ist die Aktivierung des ICC-Systems.

Ergebnisinterpretation

Da Lärm ein Stressfaktor ist (vgl. Kapitel 1.2, 2 und 3.1), können die in Lombard-Sprache stattfindenden sprachlichen Veränderungen als Ausdruck dieses Stresses angesehen werden. Demnach würde eine Abnahme dieser Sprachmodifikationen in lauten Umgebungen eine Verringerung des Stresses bedeuten. Aufgrund dieser Annahme, welche ansatzweise bereits in Kapitel 3.1 beschrieben wurde, wird gefolgert, dass die Abschwächung bzw. Aufhebung des Lombard-Effekts durch das ICC-System in den Parametern der durchschnittlichen Intensität, der durchschnittlichen Grundfrequenz und der Signaldauer eine Reduktion des Stresses und somit eine Zunahme der Kommunikationsqualität bzw. eine Verbesserung der lautsprachlichen Kommunikation im Fahrzeug bewirkt. Hypothese (8.) kann daher bestätigt werden (vgl. Kapitel 3.4). Mit ICC-System sprechen die Proband*innen zudem leiser, mit einer niedrigeren Grundfrequenz und werden schneller im Lösen ihrer Kartenaufgaben.

8.1.3 Teilfrage (3)

3. Wie hört sich die von Hintergrundgeräuschen befreite Lombard-Sprache für Hörer*innen an, die nicht an den Dialogen beteiligt waren?

Im Folgenden werden die Ergebnisse des Perzeptionsexperiments, welche durch den *Online*-Fragebogen gewonnen wurden, ausgewertet und interpretiert. Dabei spiegeln die zwei Skalen „laut-leise“ und „verständlich-unverständlich“ die Lombard-Parameter der Intensität und der Verständlichkeit wider (vgl. Kapitel 7). Die Resultate der weiteren sechs Skalenbewertungen werden im Hinblick auf Teilfrage (3.2) dargestellt. Tabelle 8.3 fasst die Ergebnisse der in Kapitel 7.6 durchgeführten statistischen Analysen zusammen.

Tabelle 8.3: Die Ergebnisse der statistischen Analysen der Skalenbewertungen (vgl. Kapitel 7.6). Dargestellt werden die Veränderungen der Messwerte bei lauten Fahrgeräuschen ohne und mit ICC-System, differenziert nach den drei Sprecher*innen der verwendeten Sprachausschnitte.

Skala	2d		3f		8o	
	ohne ICC	mit ICC	ohne ICC	mit ICC	ohne ICC	mit ICC
laut	[+]	+	+	[+]	+	[+]
verständlich	~	+	~	+	+	~
unfreundlich	+	+	~	~	~	+
emotional	+	~	+	~	+	~
gehemmt	~	~	-	-	-	+
steif	~	~	~	~	~	+
angestrengt	+	+	+	~	~	+
stockend	~	+	~	-	-	+

Kennzeichnungen:

- + = Lombard-Effekt mit steigenden Messwerten.
- = Lombard-Effekt mit fallenden Messwerten.
- ~ = keine systematischen Effekte.
- [] = Effekt ist weniger ausgeprägt als in der anderen ICC-Bedingung.

3.1 Ist in den geräuschbefreiten Sprachsignalen ein Lombard-Effekt wahrnehmbar, wenn diese bei erhöhter Lärmbelastung aufgenommen wurden?

laut/leise

Teilfrage 3.1 kann für diesen Parameter eindeutig mit „Ja“ beantwortet werden. Im Hinblick auf die Lautstärke der geräuschbefreiten Sprachsignale werden die Auswirkungen des Lombard-Effekts von den Teilnehmer*innen der Perzeptionsstudie wahrgenommen. Sowohl ohne als auch mit ICC-System klingen die Stimuli der Umfrage lauter, wenn diese bei zunehmender Geräuschbelastung aufgenommen wurden. Es wird jedoch mit ICC-System nur für zwei der drei Sprecher*innen eine Abschwächung des Phänomens erkannt. Zudem klingen die Sprachsignale mit ICC-System fast alle leiser als ohne, d. h., es findet in fast allen Experimentbedingungen eine Niveauverschiebung nach unten statt. (vgl. Tabelle 8.3 und Kapitel 7.6.3) Die Skala der wahrgenommenen Lautstärke bestätigt größtenteils die im Produktionsexperiment gewonnenen Ergebnisse der durchschnittlichen Intensität. Hypothese (9.) kann da-

her bestätigt werden, mit der Einschränkung, dass bei einer Person der Produktionsstudie keine Abschwächung des Lombard-Effekts bei Aktivierung des ICC-Systems auditiv wahrgenommen wurde. (vgl. Kapitel 7.2)

verständlich/unverständlich

Für die zweite Lombard-Skala des Semantischen Differentials kann kein systematisch auftretender Lombard-Effekt über die drei Sprecher*innen des Perzeptionsexperiments hinweg in den einzelnen ICC-Bedingungen festgestellt werden. Die Messwerte zeigen, dass ohne ICC-System lediglich bei Sprecher 8o ein Lombard-Effekt, d. h. eine Zunahme der Verständlichkeit bei ansteigender Geräuschbelastung, wahrgenommen wird. Mit ICC-System findet hingegen ein derartiger Lombard-Effekt nur bei den Sprecher*innen 2d und 3f statt. Weiterhin klingen die Sprachsignale der einzelnen Geräuschbedingungen bei Sprecherin 2d und größtenteils auch bei Sprecher 8o mit ICC-System im Durchschnitt unverständlicher, d. h., es findet bei diesem Parameter ebenfalls eine Niveaushiftung der Messwerte nach unten statt. (vgl. Tabelle 8.3 und Kapitel 7.6.3) Hypothese (10.) muss daher verworfen werden. (vgl. Kapitel 7.2)

Ergebnisinterpretation

Hinsichtlich der Lautstärke bestätigt die Wahrnehmung eines ohne ICC-System stattfindenden Lombard-Effekts in den Sprachsignalen, dass das Perzeptionsexperiment korrekt durchgeführt und für die Teilnehmer*innen nachvollziehbar gestaltet worden ist. Dabei werden die zuvor gewonnenen Ergebnisse der akustischen Analyse (vgl. Kapitel 6.3.1) mit dem Unterschied bekräftigt, dass eine Abschwächung des Lombard-Effekts durch das ICC-System nur teilweise nachgewiesen werden kann. Die Skala der Verständlichkeit zeigt als zweite Lombard-Skala dieses Experiments ohne und mit ICC-System hingegen lediglich einen unregelmäßig auftretenden Lombard-Effekt im Sinne einer gesteigerten Verständlichkeit auf, womit insgesamt nicht auf eindeutige Veränderungen der Messwerte durch das ICC-System geschlossen werden kann. Weiterhin weisen beide Lombard-Skalen bis auf einige Ausnahmen mit ICC-System eine Niveaushiftung der Messwerte nach unten auf. Dies bestätigt die oben beschriebene Vermutung, dass durch die Anwendung des ICC-Systems nicht die Situation des stillen Autoinnenraums wiederhergestellt, sondern eine neue Kommunikationssituation geschaffen wird (vgl. Kapitel 8.1.2).

- 3.2 Werden aufgrund der unterschiedlichen ICC- und Geräuschbedingungen bestimmte emotionale Assoziationen bei den Hörer*innen hervorgerufen?

unfreundlich/freundlich

Für diesen Parameter wird ohne ICC-System größtenteils kein Lombard-Effekt in den Sprachsignalen des Perzeptionsexperiments wahrgenommen. Einzeln betrachtet wurden die drei Sprecher*innen der Studie von den Hörer*innen unterschiedlich bewertet. Bei Sprecherin 2d tritt ein Lombard-Effekt auf, welcher durch die Aktivierung des ICC-Systems nicht signifikant beeinflusst wird. Hierbei klingen die Sprachsignale dieser Probandin mit steigenden Umgebungsgeräuschen immer unfreundlicher. Ein ähnliches Phänomen wird lediglich für Sprecher 8o mit ICC-System beobachtet. Somit können jedoch keine allgemeingültigen Aussagen über die Effekte von ICC-System und Hintergrundgeräuschen auf die wahrgenommene Freundlichkeit in den Sprachsignalen getroffen werden. (vgl. Tabelle 8.3 und Kapitel 7.6.3) Hypothese (11.) wird daher verworfen. (vgl. Kapitel 7.2)

emotional/sachlich

Ohne ICC-System wird in den Sprachsignalen der Umfrage ein Lombard-Effekt wahrgenommen. Dabei klingen die Sprachsignale mit zunehmender Geräuschbelastung weniger sachlich bzw. emotionaler. Mit Aktivierung des ICC-Systems sind für die drei Sprecher*innen vollkommen unterschiedliche Effekte festzustellen. D. h., der ohne ICC auftretende Lombard-Effekt kann mit ICC nicht mehr nachgewiesen werden. (vgl. Tabelle 8.3 und Kapitel 7.6.3) Daher wird Hypothese (12.) bestätigt. (vgl. Kapitel 7.2)

gehemmt/ungehemmt

Bei diesem Skalenparameter zeichnet sich ohne ICC-System größtenteils die Wahrnehmung eines Lombard-Effekts ab. Dabei reden die Sprecher*innen mit zunehmenden Hintergrundgeräuschen ungehemmter. Lediglich bei Sprecherin 2d ist von 50 km/h zu 130 km/h ein leichter Anstieg der Messwerte zu verzeichnen. Mit Anwendung des ICC-Systems ist für Sprecher 3f ein ähnlicher Effekt der Fahrgeräusche auf die wahrgenommene Gehemtheit festzustellen wie ohne ICC. Dabei wird eine

leichte Niveauverschiebung der Messwerte nach oben festgestellt. Für Sprecher 8o lässt sich mit ICC-System beobachten, dass die Sprachsignale bei ansteigender Geräuschbelastung nicht mehr ungehemmter, sondern im Gegensatz zur Situation ohne ICC-System immer gehemmter klingen. Sprecherin 2d weist mit ICC-System keinen Lombard-Effekt auf. (vgl. Tabelle 8.3 und Kapitel 7.6.3) Aufgrund dieser Ergebnisse wird Hypothese (13.) verworfen. (vgl. Kapitel 7.2)

steif/locker

Bis auf eine Ausnahme wird für diesen Parameter kein Lombard-Effekt in den Sprachsignalen der Perzeptionsstudie erkannt. Sprecher 3f klingt mit ICC steifer als ohne ICC. Die unterschiedlichen Geräuschbedingungen haben hierauf keinen signifikanten Einfluss. Sprecherin 2d klingt ohne ICC von Stille zu 50 km/h steifer und von 50 km/h zu 130 km/h wieder ein wenig lockerer. Sprecher 8o hingegen weist gegenteilige Effekte auf. Mit ICC-System werden die Sprachsignale von Sprecherin 2d von Stille zu 50 km/h als lockerer und von Stille zu 130 km/h als steifer wahrgenommen. Sprecher 8o weist mit ICC-System den einzigen Lombard-Effekt dieses Parameters auf, indem er bei steigenden Fahrgeräuschen immer steifer klingt. Somit lassen sich über alle drei Sprecher*innen hinweg keine systematischen Effekte der einzelnen ICC- und Geräuschbedingungen auf die wahrgenommene Steifheit des Gesagten feststellen. (vgl. Tabelle 8.3 und Kapitel 7.6.3) Hypothese (14.) kann ebenfalls nicht bestätigt werden. (vgl. Kapitel 7.2)

angestrengt/entspannt

Sowohl ohne als auch mit ICC-System wird für die wahrgenommene Angestrenztheit größtenteils ein Lombard-Effekt erkannt. Dabei klingen die Sprachsignale mit zunehmender Geräuschbelastung angestrenzter. Für Sprecher 3f mit ICC-System und Sprecher 8o ohne ICC-System wird jedoch kein Lombard-Effekt wahrgenommen. Es können hierdurch ebenfalls über alle Sprecher*innen hinweg keine systematischen Effekte der ICC- und Geräuschbedingungen auf den vorliegenden Skalenparameter festgestellt werden. (vgl. Tabelle 8.3 und Kapitel 7.6.3) Hypothese (15.) kann nur für Sprecher 3f bestätigt werden, hält den Ergebnissen der anderen beiden Sprecher*innen jedoch nicht stand und wird daher verworfen. (vgl. Kapitel 7.2)

stockend/flüssig

Bei diesem Untersuchungsparameter wird für die drei Sprecher*innen größtenteils ein Lombard-Effekt beobachtet, der allerdings je nach Sprecher*in und ICC-Bedingung in unterschiedliche Richtungen verläuft. Bei Sprecher 8o ohne ICC und Sprecher 3f mit ICC-System klingt das Gesagte mit ansteigenden Fahrgeräuschen immer flüssiger, während sich Sprecher*in 2d und 8o mit ICC-System dabei immer weniger flüssig bzw. stockender anhören. Daher kann auch bei diesem Parameter nicht von einem systematisch auftretenden Lombard-Effekt durch die ICC- und Geräuschbedingungen gesprochen werden. (vgl. Tabelle 8.3 und Kapitel 7.6.3) Hypothese (16.) wird nicht bestätigt. (vgl. Kapitel 7.2)

Ergebnisinterpretation

Die Ergebnisse der Skala „laut-leise“ zeigen, dass die Sprachsignale des Perzeptionsexperiments mit zunehmenden Fahrgeräuschen lauter klingen. Dies ist in beiden ICC-Bedingungen der Fall, wobei das Ausmaß des wahrgenommenen Lombard-Effekts jedoch nur bei zwei von drei Sprecher*innen mit ICC-System abgeschwächt wird. Für die Skala der Emotionalität lässt sich ebenfalls die Wahrnehmung eines Lombard-Effekts feststellen, bei dem die Sprecher*innen mit ansteigender Geräuschbelastung emotionaler klingen. Dieses Phänomen kann mit Aktivierung des ICC-Systems nicht mehr nachgewiesen werden. Die Resultate bestätigen die in Kapitel 8.1.2 vorgenommene Interpretation der Ergebnisse des Produktionsexperiments, nach der das ICC-System eine Reduktion bzw. Aufhebung des Lombard-Effekts bewirkt und damit eine Verbesserung der Kommunikationssituation herstellt.

Weiterhin bestärkt die in den Skalen „laut-leise“, „verständlich-unverständlich“, „gehemmt-ungehemmt“ und „steif-locker“ teilweise auftretende Niveauverschiebung die in Kapitel 8.1.2 aufgestellte Vermutung, dass durch Anwendung des ICC-Systems keine Rückverschiebung zu einer Situation ohne Lombard-Umgebung bewirkt wird, sondern sich hierdurch eine neue Kommunikationssituation ergibt.

Für die Skalen „verständlich-unverständlich“, „unfreundlich-freundlich“, „gehemmt-ungehemmt“, „steif-locker“, „angestrengt-entspannt“ und „stockend-flüssig“ gilt, dass trotz vereinzelt festgestellter Lombard-Effekte und anderer Auffälligkeiten in den Daten keine systematischen Einflüsse von ICC-System und Fahrgeräuschen auf

die Skalenbewertungen über alle drei Sprecher*innen hinweg nachgewiesen werden können.

Hypothese (17.) wird verworfen (vgl. Kapitel 7.2), da ein kommunikatives Konzept *Anstrengung - Entspannung* (vgl. Kapitel 7.1) anhand der Skalenbewertungen der Faktoren *evaluation*, *potency* und *activity* nicht bestätigt werden kann. Denn obwohl unter Lombard-Bedingungen im fahrenden Auto ohne ICC-System die wahrgenommene Emotionalität in den Sprachaufnahmen ansteigt, wird bei den weiteren fünf Skalen kein systematischer Effekt der Umgebungsgeräusche auf die Wahrnehmung der Sprachsignale über alle drei Sprecher*innen hinweg festgestellt.

8.1.4 Zusammenfassung

Tabelle 8.4 bietet eine Übersicht aller Parameter, welche ohne ICC-System den Lombard-Effekt aufweisen, sowie die Veränderungen dieser Effekte durch die Aktivierung des ICC-Systems.

Tabelle 8.4: Das Auftreten des Lombard-Effekts in den Parametern der Untersuchung sowie die Veränderung dieses Phänomens durch Aktivierung des ICC-Systems.

Parameter	ohne ICC-System		mit ICC-System
	Lombard-Effekt	Auswirkungen	Lombard-Effekt
∅ Intensität	✓	Steigerung	[✓]
∅ Grundfrequenz	✓	Steigerung	[✓]
Signaldauer	✓	Abnahme	~
Sprachproduktionsleistung	✓	Modifikation	✓
wahrgen. Lautstärke	✓	Steigerung	(✓)
wahrgen. Emotionalität	✓	Steigerung	~

Kennzeichnungen:

~ = keine systematischen Effekte.

[] = Effekt ist weniger ausgeprägt als in der anderen ICC-Bedingung.

() = Effekt ist teilweise weniger ausgeprägt als in der anderen ICC-Bedingung.

Die Zunahme der Fahrgeräusche im Innenraum des Testautos ohne ICC-System ruft einen Lombard-Effekt in allen Parametern der akustisch-phonetischen Analyse hervor. Dabei findet eine Steigerung der durchschnittlichen Intensität und der durchschnittlichen Grundfrequenz sowie eine Abnahme der Signaldauer statt, wodurch die in Kapitel 1.2 formulierten Hypothesen zu diesen Untersuchungsparametern bestätigt werden können. Hinsichtlich der Sprachproduktionsleistung einzelner Proband*innen wird ebenfalls ein Lombard-Effekt im Sinne einer Modifikation der Messwerte festgestellt, es kann jedoch keine systematische Zunahme der Sprachproduktionsleistung beobachtet werden. Der im Hinblick auf die durchschnittliche Intensität ohne ICC-System festgestellte Lombard-Effekt wird zudem auf perceptiver Ebene durch die Steigerung der wahrgenommenen Lautstärke bestätigt, was die Korrektheit von Vorbereitung und Durchführung des Perceptionsexperiments nachweist. Neben dieser Lombard-Skala der Perzeptionsstudie wird eine weitere Skala

des Semantischen Differentials durch die ICC- und Geräuschbedingungen des Produktionsexperiments signifikant beeinflusst. Dabei klingen die Sprecher*innen ohne ICC-System bei zunehmenden Fahrgeräuschen emotionaler.

Mit ICC-System nahmen die meisten Sprecher*innen des Produktionsexperiments die Stimme der Dialogpartnerin bzw. des Dialogpartners über Lautsprecher als lauter wahr (vgl. Kapitel 6.1). Durch Aktivierung des ICC-Systems wurde der Lombard-Effekt der Signaldauer sowie der wahrgenommenen Emotionalität aufgehoben. In den akustischen Parametern der durchschnittlichen Intensität und der durchschnittlichen Grundfrequenz wird das Phänomen abgeschwächt. Diese Reduktion des Lombard-Effekts durch das ICC-System kann hinsichtlich des Parameters der wahrgenommenen Lautstärke nur teilweise bestätigt werden. Der für die Sprachproduktionsleistung nachgewiesene Lombard-Effekt wird mit Anwendung des ICC-Systems beibehalten.

Das ICC-System zur Verbesserung der Sprachkommunikation im Auto wird aufgrund dieser Ergebnisse als eine Strategie zur Kompensierung des Lombard-Effekts interpretiert. Es dient dabei nicht nur der Kompensation von Kommunikationsnachteilen, die durch laute Umgebungsgeräusche entstehen, sondern auch von solchen, die sich durch eine ungünstige akustische und visuelle Kommunikationssituation im Auto ergeben. Der Lombard-Effekt kann als Ausdruck des Stresses, welcher durch laute Umgebungsgeräusche hervorgerufen wird, durch das ICC-System reduziert bzw. aufgehoben werden, wodurch eine Verbesserung der Kommunikationssituation entsteht. Darüber hinaus wird hierdurch die Fahrsicherheit der Passagier*innen erhöht. (vgl. [39, S. 1488])

Diese Interpretation der Ergebnisse wird teilweise durch die Fragebogenerhebung am Ende des Produktionsexperiments unterstrichen, welche in Kapitel 6.1 beschrieben wurde. Keine Sprecherin bzw. kein Sprecher des SPID-Korpus verstand die jeweils andere Person durch die Lautsprecherübertragung schlechter und die Mehrheit der Proband*innen gab dabei sogar eine Verbesserung der Verständlichkeit an. Daneben gab es geteilte Meinungen darüber, ob die Lautsprecherübertragung durch das ICC-System als störend empfunden wurde oder nicht. Störende Komponenten waren dabei z. T. das Hören der eigenen Stimme oder der Stimme der Dialogpartnerin bzw. des Dialogpartners. Zudem wurde teilweise eine merkwürdig bzw. mechanisch, dumpf oder blechern klingende Sprache, ein wahrnehmbares Echo oder eine

zu laute Sprachübertragung kritisiert. Weiterhin schienen Kopfbewegungen der vor-
sitzenden Person vereinzelt die wahrgenommene Lautstärke und Tonqualität des
übertragenen Signals zu beeinflussen. Positiv wurde u. a. bewertet, dass die durch
Lautsprecher übertragene Sprache natürlich klang, nicht zu laut oder zu leise war
und die Kommunikation leichter wurde. Zudem wurde eine schnelle Gewöhnung an
die ungewohnte Situation mit ICC-System beschrieben. (vgl. Kapitel 6.1)

Bis auf eine Person nahmen alle Proband*innen des Produktionsexperiments die
Übertragung von Sprachsignalen durch das ICC-System wahr. Gleichzeitig gab der
Großteil aller Sprecher*innen an, die Gesprächspartnerin bzw. den Gesprächspartner
mit zunehmenden Fahrgeräuschen häufiger über Lautsprecher zu hören. (vgl. Ka-
pitel 6.1) Letzteres lässt darauf schließen, dass sich die meisten Teilnehmer*innen
der geräuschabhängigen Einstellung des ICC-Systems bewusst waren. Diese subjek-
tiven Einschätzungen bestätigen die in Kapitel 8.1.2 beschriebene Annahme, dass
mit ICC-System nicht die ursprünglich stille Situation des Autoinnenraums wieder-
hergestellt, sondern eine neue, veränderte Kommunikationssituation erreicht wird.
Diese Interpretation wird weiterhin durch die Niveauverschiebungen der Messwerte
mit ICC-System in allen Parametern der akustisch-phonetischen Analyse (vgl. Ka-
pitel 6.3) sowie in den perzeptiven Parametern der wahrgenommenen Lautstärke,
Verständlichkeit, Gehemmtheit und Steifheit bestätigt. Durch die Aufhebung bzw.
Abschwächung des Lombard-Effekts trägt die neuartige Kommunikationssituation
jedoch ebenso wie die Situation des stillen Autoinnenraums zur Verbesserung der
lautsprachlichen Kommunikation zwischen den Passagier*innen bei.

8.2 Fazit

Im Zuge der vorliegenden Studie ist ein Korpus spontan gesprochener Dialoge im
Auto (SPID, *SPontaneous In-car Dialogues*) entstanden, welche ohne und mit ICC-
System bei unterschiedlichen Geräuschkulissen von Stille, 50 km/h und 130 km/h
produziert wurden. Diese sind sowohl in der akustischen und visuellen Umgebungssi-
mulation der CAU Kiel als auch im fahrenden Auto aufgenommen worden. Die Um-
gebungssimulation stellt eine neuartige Methode zur Untersuchung lautsprachlicher
Kommunikation im Fahrzeug dar. Hierdurch werden realitätsnahe Aufnahmen von
Lombard-Sprache im Auto möglich, welche in vorangegangenen Studien durch den

Einsatz von Kopfhörern, durch unrealistische Laborumgebungen oder Geräusche im Sprachsignal, welche die akustischen Analysen stark beeinflussen können, deutlich eingeschränkt waren.

An den zwei Aufnahmesitzungen des Produktionsexperiments nahmen insgesamt 16 Muttersprachler*innen des Standarddeutschen teil. Die Sprachsignale wurden weitgehend orthographisch transkribiert und annotiert, wobei semi-automatische Verfahren zur Anwendung kamen. Nach einer akustisch-phonetischen Analyse der Signale wurde zudem ein Wahrnehmungsexperiment durchgeführt, innerhalb dessen 132 Proband*innen Ausschnitte der Sprachsignale des SPID-Korpus mit Hilfe eines *Online*-Fragebogens subjektiv evaluierten. Durch Produktions- und Perzeptionsexperiment dieser Studie und die jeweiligen Ergebnisanalysen sollte festgestellt werden, welche lautsprachlichen Modifikationen durch laute Fahrgeräusche im Innenraum eines Autos entstehen bzw. wahrgenommen werden und ob die Kommunikationssituation im fahrenden Auto durch Aktivierung eines ICC-Systems verbessert werden kann.

Die deskriptiven und inferenzstatistischen Resultate beider Experimente ergeben, dass die Geräusche eines fahrenden Autos im Innenraum einen Lombard-Effekt hervorrufen, welcher sowohl durch akustisch-phonetische als auch perzeptive Untersuchungen nachgewiesen werden kann. Dabei sprechen die Fahrgäste lauter, höher und emotionaler. Zudem nimmt bei steigenden Hintergrundgeräuschen die Signaldauer ab, d. h., die Dialoge werden kürzer und die Proband*innen benötigen weniger Zeit zum Lösen der Kartenaufgaben im *Map-Task* Szenario, während sich ihre Sprachproduktionsleistung je nach Sitzplatz und Sprechrolle verändert. Die Modifikationen der Intensität, Grundfrequenz und Signaldauer in Lombard-Sprache sind bereits aus der Literatur bekannte Phänomene. In dieser Arbeit wurden sie für den Kommunikationsort des Autoinnenraums nachgewiesen. Zudem wurde erkannt, dass die in den Sprachsignalen wahrgenommene Emotionalität einen neuen Lombard-Parameter darstellt. Durch die Aktivierung des ICC-Systems wird der Lombard-Effekt der durchschnittlichen Intensität bzw. der Lautstärke und der durchschnittlichen Grundfrequenz abgeschwächt und hinsichtlich der Signaldauer und der wahrgenommenen Emotionalität sogar gänzlich aufgehoben. Die Sprachproduktionsleistung der Fahrgäste verschiebt sich mit ICC-System vom hinteren auf den vorderen Sitzplatz.

Aufgrund dieser Ergebnisse wird das ICC-System als neue Strategie zur Kompensierung des Lombard-Effekts betrachtet. Es trägt somit zur Verbesserung der Kommunikationssituation und damit ebenfalls zur Steigerung der Fahrsicherheit bei, da sich auch die Fahrerin bzw. der Fahrer weniger zu den Gesprächspartner*innen umdrehen muss, um sie verstehen zu können und verstanden zu werden (vgl. Kapitel 2). Zum einen werden Kommunikationsnachteile kompensiert, die sich aufgrund lauter Fahrgeräusche ergeben, und zum anderen solche, die bereits im stillen Fahrzeug durch die akustisch und visuell nachteilige Situation des Autoinnenraums entstehen. Dabei wird durch das ICC-System nicht, wie anfangs vermutet, die stille Situation des Fahrzeuginnenraums wiederhergestellt, was durch Niveauverschiebungen der unterschiedlichen Messwerte deutlich wird. Vielmehr entsteht eine neue Kommunikationssituation, welche jedoch in Übereinstimmung mit den in Kapitel 1.2 formulierten Hypothesen eine Aufhebung bzw. Reduktion des Lombard-Effekts bewirkt und dadurch die lautsprachliche Kommunikation der Passagier*innen untereinander verbessert.

8.3 Methodenkritik

Die Anwendung der akustischen und visuellen Umgebungssimulation birgt zahlreiche Vorteile, welche in Kapitel 3.3 und 4 eingehend diskutiert wurden. Dazu gehören, dass geräuschfreie Lombard-Aufnahmen ohne die Verwendung von Kopfhörern gelingen, welche einen zusätzlichen Lombard-Effekt hervorrufen. Aufgrund der verschiedenen Methoden der Geräuschbefreiung sind akustische Analysen der Sprachsignale möglich, ohne dass diese durch störende Hintergrundgeräusche verzerrt werden. Zudem bietet die Simulation realitätsnahe Bedingungen im Labor und ist bereits durch verschiedene Untersuchungen und Ergänzungen weiterentwickelt und verbessert worden.

Bei einer Untersuchung des Lombard-Effekts müssen immer dessen Ursachen sowie mögliche Einflussfaktoren auf die Ausmaße des Phänomens berücksichtigt werden (vgl. Kapitel 3). Dabei ist die Stärke des Lombard-Effekts bei spontan gesprochener Sprache größer als bei gelesener Sprache. Auch durch eine interaktive bzw. kommunikative Sprechaufgabe kann das Ausmaß des Phänomens angehoben werden. Diese Voraussetzungen werden durch das *Map-Task* Szenario (vgl. Kapitel 5) erfüllt, bei

zeitgleicher Aufrechterhaltung einer hohen Kontrolle über die Wortproduktionen im Dialog. Weiterhin erlaubt diese Sprecheraufgabe den Proband*innen aufgrund der klaren Aufgabenstellung eine relativ ungezwungene Kommunikationssituation, welche durch die emotionale Nähe zur Dialogpartnerin bzw. zum Dialogpartner zusätzlich unterstützt wird. Trotz der zahlreichen Tests und Überarbeitungen der Kartenaufgaben sowie der wiederholten Instruktionen für die Sprecher*innen wurden nicht alle Zielwörter in allen Sprachaufnahmen des SPID-Korpus produziert, was einen Vergleich der unterschiedlichen ICC- und Geräuschbedingungen auf Basis der vorgegebenen Zielwörter erschwert. Obschon diese Wortproduktionen bei gelesener Sprache in allen Signalen gewährleistet sind, überwiegt dies nicht die zahlreichen Vorteile spontan produzierter Dialoge für die Untersuchung von Lombard-Sprache, wie sie oben beschrieben sind.

Die Dauer der Dialoge hat den vorgesehenen Rahmen des Produktionsexperiments teilweise deutlich überschritten. Mögliche Ursachen hierfür sind im Folgenden beschrieben. Die Proband*innen wurden wiederholt darum gebeten, alle Personen, Hausnummern und Straßennamen, welche auf der eigenen Karte oder der Karte der anderen Person abgebildet sind, bei jeder Aufnahme mindestens einmal zu nennen (vgl. Kapitel 5.4). In dem Bestreben, die Sprecheraufgabe möglichst korrekt durchzuführen, und um ihrer Dialogpartnerin bzw. ihrem Dialogpartner die eigene Karte detailliert zu schildern, nannten viele Sprecher*innen die dargestellten Wörter mehr als einmal. Darüber hinaus erschien die Lösung der Aufgabe manchen Teilnehmer*innen schwerer als anderen, zumal die Karten relativ viele Anregungen zur Diskussion boten.

Aufgrund des zeitlichen Rahmens dieser Arbeit fand keine Untersuchung der Wortreduktion, der Dauer und der phonologischen Prozesse in den Zielwörtern des *Map-Task* Szenarios statt. Auch zusätzliche akustisch-phonetische Analysen wie die der Formantfrequenzen bzw. der Vokalqualitäten wurden deshalb nicht vorgenommen. Im Hinblick auf die in Kapitel 3.1 beschriebenen Eigenschaften des Lombard-Effekts wird deutlich, dass in der vorliegenden Arbeit nur ein kleiner Ausschnitt aller bekannten Lombard-Parameter untersucht werden konnte. Die Segmentierung und Annotation der Sprachsignale konnte jedoch trotz der zeitlichen Begrenzung in großem Umfang durchgeführt werden und kann somit nachfolgenden Untersuchungen als Basis dienen.

Die semi-automatische Segmentierung und Annotation der Sprachaufnahmen mithilfe der *Software* MAUS (vgl. Kapitel 5.7) barg zahlreiche Schwierigkeiten. Dabei entstanden durch das Programm unterschiedliche Fehler, welche u. a. auf die Länge der einzelnen Signale zurückzuführen waren. Durch verschiedene, speziell zu diesem Zweck entwickelte Algorithmen konnte die Arbeit mit MAUS jedoch verbessert werden. Dies wirkte sich positiv auf die ohnehin umfangreiche manuelle Korrektur der Segmentierungen und Annotationen in den Ausgabedateien von MAUS aus.

Um den Lerneffekt zwischen den beiden Aufnahmesitzungen des Produktionsexperiments zu isolieren, hätte eine Kontrollgruppe dieselbe Sprechaufgabe in derselben Aufnahmesituation ausführen müssen, wobei in der zweiten Aufnahmesitzung keine Aktivierung des ICC-Systems stattgefunden hätte. Diese Vorgehensweise war aufgrund des begrenzten zeitlichen Rahmens der Forschungsarbeit jedoch nicht durchführbar.

Die Ergebnisse der Perzeptionsstudie weisen z. T. darauf hin, dass die Vorauswahl der Stimuli im Semantischen Differential entscheidend für die Bewertung durch die Teilnehmer*innen war. Dies wird anhand der teilweise unterschiedlichen Beurteilungen der Sprachausschnitte von Sprecher*in 2d, 3f und 8o deutlich. Um diesem Umstand Rechnung zu tragen, könnten z. B. mehrere Sprachausschnitte derselben Experimentbedingungen in das Wahrnehmungsexperiment integriert werden, was die Umfrage jedoch wesentlich in die Länge ziehen und bei den Teilnehmer*innen einen Ermüdungseffekt hervorrufen würde, durch den die Daten verzerrt werden könnten. Diese Vorgehensweise wäre daher lediglich für einen Teil der untersuchten Skalen im Semantischen Differential anwendbar gewesen.

8.4 Ausblick

Das Sprachkorpus SPID wird im Anschluss an diese Untersuchung öffentlich zugänglich gemacht und soll damit nachfolgenden Studien von Nutzen sein. Dabei kann das Korpus dazu dienen, den Lombard-Effekt im Auto hinsichtlich zahlreicher weiterer Parameter zu untersuchen. Hierzu zählen linguistische, akustische und visuelle Spracheigenschaften, wie z. B. unterschiedliche Modifikationen von Inhalts- und Funktionswörtern, Veränderungen von Formantfrequenzen und Modifikationen der Mundöffnung (vgl. Kapitel 3.1 und 3.2.2). Die Untersuchung der visuellen Sprachei-

genschaften ist aufgrund der Anbringung einer Kamera auf dem Amaturenbrett des Fahrzeugs während der Aufnahmen möglich. Zudem können jene Lombard-Eigenschaften studiert werden, welche in der bisherigen Erforschung dieses Phänomens nicht oder nur wenig Berücksichtigung fanden, wie die veränderte Häufigkeit von Häsitationen und Zögerungslängungen (vgl. Kapitel 3.1). Die Sprechaufgabe des *Map-Task* Szenarios ermöglicht eine hohe Kontrolle über die Produktion einzelner Zielwörter in den Sprachsignalen (vgl. Kapitel 5.1). Dadurch können zudem verschiedene phonologische Prozesse, Wortreduktionen und die Veränderung der Wort- und Silbendauer analysiert werden.

Auch die anhand des *Online*-Fragebogens gewonnenen Daten zur Sprachwahrnehmung bieten weiterführende Möglichkeiten der Untersuchung. Dabei könnten zusätzliche statistische Methoden wie die der Faktorenanalyse dazu beitragen, verschiedene Skalen aufgrund ihrer Korrelationen zu einem gemeinsamen Lombard-Faktor zu kombinieren. Zudem können die Segmentierungen, Transkriptionen und Annotationen der im Labor und im fahrenden Auto aufgenommenen SPID-Sprachsignale auf orthographischer und phonologischer Ebene vervollständigt werden. Dadurch ergeben sich neue Möglichkeiten für weiterführende Erkenntnisse zur Sprachmodifikation bei unterschiedlichen ICC- und Fahrgeräuschbedingungen. Die weiterführende Untersuchung von Lombard-Eigenschaften kann zum einen die Ergebnisse der vorliegenden Arbeit untermauern und zum anderen weitere Effekte des verwendeten ICC-Systems auf Sprachproduktion und Sprachwahrnehmung in den Aufnahmen des SPID-Korpus aufdecken. Darüber hinaus können zusätzliche Vergleiche zwischen der Sprechsituation in der akustischen und visuellen Umgebungssimulation und der im fahrenden Auto angestellt werden.

Aus den Ergebnissen vorangegangener Studien geht hervor, dass der Lombard-Effekt je nach Geschlecht einer Person unterschiedlich stark ausgeprägt ist (vgl. Kapitel 3.3.2). Während der akustisch-phonetischen Analyse der vorliegenden Arbeit wurden keine Hinweise auf geschlechtsspezifische Unterschiede in den Daten des SPID-Korpus gefunden (vgl. Kapitel 6.3). Die Ergebnisse des Perzeptionsexperiments legen hingegen nahe, dass es derartige Einflüsse auf die *Wahrnehmung* des Lombard-Effekts geben könnte (vgl. Kapitel 7.6). Dabei wird z. B. ohne ICC-System für Sprecherin 2d kein eindeutiger Lombard-Effekt in der Skala „gehemmt-ungehemmt“ wahrgenommen, während dies für die männlichen Probanden 3f und 8o der Fall ist.

In der Skala „verständlich-unverständlich“ wird ohne ICC-System für Sprecher 3f und 8o ein annähernd konsequent vorhandener Lombard-Effekt wahrgenommen, jedoch nicht für Sprecherin 2d. Das ICC-System hat zudem keinen Einfluss auf die perzipierte Freundlichkeit dieser Probandin. Auch hinsichtlich der Stärke der Hörerurteile zeichnet sich eine Tendenz zu geschlechtsspezifischen Unterschieden ab. Dabei klingt Probandin 2d leiser als die männlichen Sprecher 3f und 8o und hört sich mit ICC-System zudem wesentlich emotionaler an. Zur Bestätigung dieser geschlechtsabhängigen Auswirkungen von Fahrgeräuschen und ICC-System bedarf es jedoch einer weiteren Perzeptionsstudie mit größerer Stichprobe. Das Ziel einer weiterführenden Untersuchung könnte u. a. die Beantwortung der Frage sein, ob sich ein Geschlecht besser zur Evaluierung von ICC-Systemen eignet, z. B. aufgrund einer stärkeren Ausprägung des Lombard-Effekts, wie im Falle der Skala „gehemmt-ungehemmt“ für die männlichen Probanden 3f und 8o. Zudem wäre es vorteilhaft zu wissen, ob eine geschlechts- bzw. frequenzspezifische Einstellung von ICC-Systemen sinnvoll ist. Bei einer entsprechenden Untersuchung würden die verschiedenen Stimmlagen der Sprecher*innen und ihre Modifikationen durch Fahrgeräusche und ICC-System in Beziehung gesetzt werden. Dabei könnte erkannt werden, für welche Sprecher*inneneigenschaften das ICC-System Schwächen aufweist und wie dieses an die Anforderungen der einzelnen Fahrgäste angepasst werden kann.

Die Ergebnisse der vorliegenden Studie haben gezeigt, dass sich das ICC-System auf verschiedene Parameter des Lombard-Effekts unterschiedlich stark auswirkt. Bei einigen Parametern wird der Effekt vollständig aufgehoben, bei anderen lediglich reduziert. Davon ausgehend, dass es verschiedene Ursachen für den Lombard-Effekt geben kann (vgl. Kapitel 3.2), lässt dies auf einen kompositionalen Charakter des Phänomens schließen. Dabei könnte die Veränderung akustischer Parameter wie der Intensität und der Grundfrequenz ihren Ursprung neben der externen bzw. öffentlichen *Feedback-Loop* maßgeblich in der internen bzw. privaten *Feedback-Loop* haben, da durch laute Hintergrundgeräusche die Kontrolle über die eigene Stimme beeinträchtigt wird. Lombard-Eigenschaften wie die Dialogdauer und die Sprachproduktionsleistung hängen eventuell stärker von der Reaktion der Gesprächspartnerin bzw. des Gesprächspartners ab, da diese Parameter vor allem den interaktiven Charakter des Gesprächs widerspiegeln. Auch die räumliche Distanz zwischen den Personen und der auf sie einwirkende Stress könnten unterschiedliche Merk-

male des Kommunikationsverhaltens auf verschiedene Weisen beeinflussen. Daraus würde sich ergeben, dass das ICC-System nicht alle nachteiligen Faktoren der Kommunikationssituation in gleichem Maße kompensiert. Beispielsweise könnte dadurch die Interaktion der Gesprächspartner*innen stärker verbessert werden als die lautsprachliche Eigenkontrolle der Passagier*innen. Um diese Annahme zu überprüfen, müssten verschiedene Ursachen des Lombard-Effekts isoliert werden. Eine darauf abzielende Studie könnte z. B. testen, inwieweit sich der Lombard-Effekt im Auto ohne und mit ICC-System verändert, wenn die Sprecher*innen hintereinander oder nebeneinander sitzen, um die Distanz zwischen ihnen zu verringern. Zudem würde eine weniger angestregte und damit stressfreiere Kommunikationssituation entstehen, wenn beide Sprecher*innen Blickkontakt halten und in die Richtung der jeweils anderen Person sprechen würden, sodass keiner von beiden Richtung Windschutzscheibe spricht. Somit ließe sich der Effekt der verminderten Selbstkontrolle im fahrenden Auto isolieren. Andererseits wäre es möglich, das Hauptaugenmerk der Untersuchung durch das Zuschalten eines (lauteren) *Sidetones* in den Ohren der sprechenden Person auf den Faktor der Sprecher*inneninteraktion zu legen. Hierbei besteht jedoch die Gefahr, dass die Sprecherin bzw. der Sprecher einen zu lauten *Sidetone* als Echo wahrnimmt. Darüber hinaus wäre es möglich, dass sich bei den Passagier*innen in Bezug auf das ICC-System ein Gewöhnungseffekt einstellt, wodurch der Lombard-Effekt für einige Parameter erst nach längerer Anwendung des ICC-Systems reduziert bzw. vollständig aufgehoben wird. Um dies zu überprüfen, ist eine Langzeitstudie nötig, welche z. B. durch sich wöchentlich wiederholende Aufnahmen mit denselben Proband*innen realisiert werden könnte.

Eine andere mögliche Erklärung für die unterschiedlichen Effekte des ICC-Systems ist, dass den Sprecher*innen einige Eigenschaften des Lombard-Effekts bewusst sind und andere wiederum nicht (vgl. Kapitel 3.2.3). Aufgrund der Tatsache, dass die Proband*innen des Produktionsexperiments von der Anwendung des ICC-Systems wussten, könnten sie eine Verbesserung der Kommunikation erwartet und einige der bei Lärm auftretenden Modifikationen aktiv reduziert bzw. aufgehoben haben. Unbewusste Veränderungen des Kommunikationsverhaltens würden dadurch nicht beeinflusst werden. Diese Annahme ließe sich überprüfen, indem bei Aufnahmen im Fahrzeug die Aktivierung des ICC-Systems nicht angekündigt wird. Alternativ könnte das ICC-System während der Sprachaufnahmen sukzessive heruntergefahren

und die Effekte auf den Lombard-Effekt analysiert werden. Weiterhin könnten die Sprecher*innen nach den Aufnahmen dahingehend befragt werden, wie sich ihre Sprachproduktion und ihr kommunikatives Verhalten durch die Geräuschbelastung im Auto und die Aktivierung eines ICC-Systems verändert haben. Dadurch wird getestet, welche Parameter den Sprecher*innen zumindest annähernd bewusst oder unbewusst sind.

Weiterhin stellt sich die Frage, ob die im Perzeptionsexperiment (vgl. Kapitel 7) getestete wahrgenommene Emotionalität einen weiteren Lombard-Parameter darstellt oder ob sich diese Beurteilung der Sprecher*innen aus der geräuschbefreiten Präsentation der Sprachsignale ergibt. Würden Zuhörer*innen die SPID-Stimuli als ebenso sachlich oder emotional einstufen, wenn sie wüssten, dass die Sprecher*innen z. T. lauten Fahrgeräuschen ausgesetzt waren oder wenn die Zuhörer*innen diese Geräusche sogar selbst hören?¹⁵ Um diese Frage zu beantworten, wäre ein weiteres Wahrnehmungsexperiment nötig, in dem Proband*innen dieselben Stimuli mit und ohne Hintergrundgeräusche hören und jeweils nach dem Grad ihrer Emotionalität beurteilen. Falls es sich bei diesem Parameter tatsächlich um eine Eigenschaft des Lombard-Effekts handelt, ergibt sich daraus, dass akustisch-phonetische Eigenschaften der Sprachproduktion, die dem Ausdruck verschiedener Gefühlslagen dienen, bei lauten Fahrgeräuschen im Autoinnenraum verändert werden. Diese Modifikationen würden dann durch Aktivierung des ICC-Systems aufgehoben werden.

Die Ergebnisse der vorliegenden Studie haben gezeigt, dass Intensität und Grundfrequenz bei zunehmenden Fahrgeräuschen im Autoinnenraum einen Lombard-Effekt aufweisen, welcher durch das ICC-System abgeschwächt wird. Aus der bisherigen Forschungsliteratur ist bekannt, dass laute Hintergrundgeräusche bei weiteren Eigenschaften der Sprachproduktion ebenfalls einen Lombard-Effekt hervorrufen. Dazu gehören u. a. die Formantwerte, die VOT und der *spectral tilt*. Daher wird angenommen, dass der Lombard-Effekt für diese Parameter auch im Innenraum eines Fahrzeugs festgestellt werden kann und, wie es für die Intensität und die Grundfrequenz der Fall ist, durch Aktivierung eines ICC-Systems abgeschwächt wird. Bei

¹⁵An dieser Stelle soll darauf hingewiesen werden, dass die Interpretation phonetischer *Acoustic Cues*, sog. „akustischer Hinweise“ (vgl. [5, S. 5]) kontextsensitiv ist. Beispielsweise ist nicht nur die Position des ersten Formanten, sondern auch der Abstand zwischen F1 und F0 entscheidend für den wahrgenommenen Öffnungsgrad von Vokalen [73]. Hieraus ergibt sich die Vermutung, dass auch die Wahrnehmung von Emotionalität vom Kontext der beurteilten Sprachsignale abhängt.

Parametern, die der Interaktion zwischen zwei oder mehreren Personen geschuldet sind, wie z. B. die Häufigkeit des Blickkontakts, die visuellen Kommunikationssignale und die Anzahl der *Turns* im Gespräch, wird ebenfalls erwartet, dass sie im stillen Autoinnenraum ohne ICC-System einen mit den Ergebnissen vorangegangener Studien vergleichbaren Lombard-Effekt aufweisen. Es wird vermutet, dass diese Eigenschaften durch das ICC-System eine vollständige Aufhebung des Lombard-Effekts erfahren, da dies auch für den von der Interaktion der Sprecher*innen abhängigen Parameter der Dialogdauer der Fall ist. Der Grad der Wortreduktion wird dabei ebenfalls durch die externe bzw. öffentliche *Feedback-Loop* beeinflusst, d. h., die sprechende Person passt die artikulatorische Genauigkeit den situationsbedingten Anforderungen der ZuhörerIn bzw. des Zuhörers im Sinne der H&H-Theorie (Lindblom 1990) an:

„Speech production is adaptive. Speakers can, and typically do, tune their performance according to communicative and situational demands, controlling the interplay between production-oriented factors on the one hand, and output-oriented constraints on the other. [...] Hence speakers are expected to vary their output along a continuum of **hyper-** and **hypospeech**.“ ([44, S. 403])

Aufgrund der Abschwächung bzw. Aufhebung anderer (interaktiver) Lombard-Eigenschaften im Fahrzeug durch Aktivierung des ICC-Systems wird dies auch für den Lombard-Parameter der Wortreduktion erwartet. D. h., die situationsbedingte artikulatorische Anpassung der sprechenden Person wird mit Anwendung des ICC-Systems vermindert.

Die Untersuchung verschiedener Lombard-Parameter aus unterschiedlichen Bereichen der Kommunikation, d. h. mehr und weniger interaktiv geprägter Eigenschaften, Merkmalen der Artikulation oder des Kommunikationsverhaltens, gewährleistet eine umfangreiche Kenntnis dieses Phänomens. Es ist daher von Vorteil, möglichst viele Modifikationen der Sprachproduktion und des Kommunikationsverhaltens bei Lärm zu kennen und in Relation setzen zu können. Durch die Analyse weiterer prosodischer Spracheigenschaften in spontansprachlicher Autoinnenraumkommunikation können Zusammenhänge zwischen diesen und anderen interaktiven Parametern festgestellt werden. Die bei Summers et al. (1988) beschriebene Vermutung, dass

in Lombard-Sprache Modifikationen von F0 und F1 zusammenhängen (vgl. Kapitel 3.1), würde z. B. für die Evaluation von ICC-Systemen bedeuten, dass nur einer dieser Parameter untersucht werden muss. Eine umfangreiche Bewertung kommunikationsunterstützender Systeme würde daher durch die Kenntnis von Korrelationen zwischen einzelnen Lombard-Parametern vereinfacht werden. Auch unabhängig von der Bewertung solcher ICC-Systeme würde dies für die Analyse des Lombard-Effekts bedeuten, dass weniger Parameter untersucht werden müssen, um eine Veränderung des Phänomens unter verschiedenen Bedingungen festzustellen.

Zudem könnten sich weiterführende Studien eingehender mit der Frage nach dem reflexartigen Charakter des Lombard-Effekts im Fahrzeug befassen. Die Ergebnisse der vorliegenden Untersuchung geben keine eindeutigen Hinweise darauf, ob der im Fahrzeug produzierte Lombard-Effekt reflexartig entsteht oder von den Sprecher*innen bewusst wahrgenommen wird (vgl. Kapitel 3.2). Es wäre u. a. möglich, dass der Lombard-Effekt den Sprecher*innen abhängig von ihrem Kommunikationsort bewusst oder nicht bewusst ist, da an verschiedenen Orten unterschiedliche Bedingungen zur Kommunikation herrschen. Im Vergleich zu einer Kantine Umgebung, in der die verminderte Verständlichkeit des Gesagten durch Blickkontakt zwischen den Sprecher*innen vermutlich bis zu einem gewissen Grad kompensiert werden kann (vgl. Kapitel 3.1 zu visueller Lombard-Sprache), können sich die Passagier*innen eines Autos während einer Unterhaltung häufig nicht ins Gesicht sehen. Dadurch sind sie sich der ungünstigen Kommunikationssituation und der Notwendigkeit ihrer Kompensation wahrscheinlich weitaus stärker bewusst. Subjektive Tests zur Wahrnehmung des Lombard-Effekts aus Sicht der Sprecher*innen könnten einen Vergleich zwischen dem Innenraum eines Autos und einer anderen lauten Umgebung, wie z. B. einer Kantine, anstellen.

Die in den Aufnahmen des SPID-Korpus durch das ICC-System neu entstandene Kommunikationssituation zeichnet sich vor allem dadurch aus, dass die am Gespräch beteiligten Personen über ein technisches System lautsprachlich kommunizieren, obwohl sie sich in unmittelbarer Nähe zueinander befinden. Dabei können sich die Sprecher*innen trotz ungünstiger Geräuschbedingungen gut miteinander unterhalten. Weiterhin hebt sie sich dadurch von vielen anderen Situationen ab, dass durch das ICC-System mitunter ein Eigenecho entsteht (vgl. Kapitel 6.1), und dass die eigene Stimme automatisch verstärkt wird, was an die Verwendung ei-

nes Megaphons erinnert. Diese Kommunikationssituation ähnelt einer Unterhaltung zwischen Einsatzkräften, z. B. denen der Feuerwehr, die Atemschutzmasken tragen, während sie ihr weiteres Vorgehen besprechen. Auch hierbei kommt ein kommunikationsunterstützendes System zum Einsatz. Zudem ist ein unabhängig von den lauten Geräuschen erhöhtes Stresslevel vorhanden, das auch im Auto bei ungünstigen Verkehrssituationen vorkommen kann. In beiden Situationen ist der Blickkontakt zwischen den beteiligten Personen häufig eingeschränkt oder nicht vorhanden. D. h., auch die Feuerwehrleute unterhalten sich durch das Kommunikationssystem in den Atemschutzmasken nicht miteinander, als würden sie sich in einer ruhigen, entspannten Situation befinden, da sie einem erhöhten Stresslevel ausgesetzt sind. Dennoch wird durch das technische System eine Verbesserung der Kommunikation erreicht und damit, wie auch im Auto, u. a. eine erhöhte Sicherheit der beteiligten Personen ermöglicht.

Die Passagier*innen eines fahrenden Autos sind trotz kommunikationsunterstützendem System weiterhin den sie umgebenden lauten Geräuschen und anderen anstrengenden Faktoren, wie z. B. dem Verkehrsgeschehen, ausgesetzt (vgl. Kapitel 2). Ein ICC-System kann die dadurch entstehenden negativen Auswirkungen auf die lautsprachliche Kommunikation der Passagier*innen kompensieren, jedoch sind der Hintergrundlärm und der Straßenverkehr weiterhin vorhanden. Eine Verbesserung der lautsprachlichen Kommunikation durch ein solches System führt daher zu einer neuen Kommunikationssituation, anstatt die relativ stressfreie Situation eines stillen Autoinnenraums wiederherzustellen. Weiterführende Studien könnten sich jedoch mit der Frage befassen, inwieweit sich diese neue Situation dem stillen Autoinnenraum immer mehr annähern lässt. Evtl. könnte dies durch eine unterschiedliche Anzahl oder unterschiedliche Positionen der Mikrophone und Lautsprecher im Auto bewerkstelligt werden. Zudem kann eine anschließende Erforschung der durch das ICC-System neu entstandenen Kommunikationssituation (vgl. Kapitel 8.1) z. B. durch eine Perzeptionsstudie aus Sicht der Sprecher*innen erfolgen. Diese würde auf individuelle Bedürfnisse der Proband*innen, wie z. B. eine zu laute Signalübertragung oder ein wahrnehmbares Eigenecho, eingehen.

Aufgrund der Optimierung der Kommunikation zwischen den Passagier*innen ist diese neue Kommunikationssituation wesentlich besser als die Situation im Auto ohne ICC-System. Denn viele Sprecher*innen des Produktionsexperiments verstan-

den ihre Gesprächspartnerin bzw. ihren Gesprächspartner im Auto besser, wenn das ICC-System aktiviert war. Eine Verschlechterung der Verständlichkeit wurde nicht angegeben. Die einzigen Kritikpunkte schienen aus Sicht der Passagier*innen zu sein, dass sie z. T. ihre eigene Stimme hörten und sich die Stimme der anderen Person mitunter etwas merkwürdig anhörte. Jedoch gaben einige der Sprecher*innen an, sich schnell an die neue Situation gewöhnt zu haben. (vgl. Kapitel 6.1) Eine notwendige Anpassung der Passagier*innen an die neue Kommunikationssituation erscheint dabei vollkommen sinnvoll, da es sich hierbei um eine ungewohnte Situation handelt. Die Ergebnisse dieser Dissertation sowie die Resultate nachfolgender Studien können dazu beitragen, neue Standards zur Evaluierung von ICC-Systemen zu generieren. Dabei geht es vor allem um die Erkenntnis, dass zur Bewertung der Kommunikationsqualität im Auto (vgl. Kapitel 2.2) die dynamische Interaktion zwischen den Passagier*innen in die Untersuchung einbezogen werden muss. Dies gelingt durch Anwendung einer spontansprachlichen Aufgabe. Durch die von Schmidt und Haulick (2006) und Schmidt et al. (2012) beschriebenen subjektiven und objektiven Methoden zur Evaluierung von ICC-Systemen (vgl. Kapitel 2.2) werden beide Teilnehmer*innen eines Dialogs im Autoinnenraum berücksichtigt, d. h. sowohl Sprecher*in als auch Zuhörer*in. Dennoch wird mit diesen Tests von Reimwörtern, Nachhall, Sprachnatürlichkeit, SNR, usw. kein kommunikatives Verhalten zwischen den Passagier*innen untersucht. Die Verbesserung der interaktiven lautsprachlichen Kommunikation ist jedoch vorrangiges Ziel eines ICC-Systems.

Die in dieser Arbeit beschriebene Methode zur Evaluierung von ICC-Systemen hat gezeigt, dass sich der Lombard-Effekt, welcher in spontansprachlicher Interaktion aussagekräftiger ist als bei gelesenen Monologen, als Maß der Sprachqualität eignet. D. h., die vorliegenden Ergebnisse schaffen die Grundlage dafür, mit dem Lombard-Effekt ein Maß für die Verbesserung der Sprach- bzw. Kommunikationsqualität durch ICC-Systeme zu entwickeln. Die Qualität eines ICC-Systems wird dabei durch die Reduktion des Lombard-Effekts bei Anwendung dieses Systems messbar gemacht. Als konkretes Maß wird die Quote der Lombard-Reduktion vorgeschlagen und entsprechend als „Allgemeiner Lombard-Quotient“ (*General Lombard Quotient*, GLQ) benannt. Er wird deshalb als allgemein bezeichnet, weil sich dieser Quotient nicht nur auf die Veränderung des Lombard-Effekts durch ICC-, sondern darüber hinaus auch durch andere kommunikationsunterstützende Systeme anwen-

den lässt. Der GLQ ist für unterschiedliche Parameter anwendbar und wird definiert als:

Grad der Reduktion des Lombard-Effekts als Ausdruck der Qualität eines kommunikationsunterstützenden Systems.

Die Formel des Allgemeinen Lombard-Quotienten kann dabei folgendermaßen dargestellt werden:

$$GLQ = 1 - \frac{\Delta g(x_s)}{\Delta g(x)}, \quad (8.1)$$

wobei x eine Geräuschbedingung ist, z. B. 130 km/h, und $g(x)$ die Ausprägung eines bestimmten Lombard-Parameters unter dieser Geräuschbedingung, z. B. 51 dB. Der Index s bezeichnet das untersuchte kommunikationsunterstützende System. Das folgende Beispiel zeigt auf, wie Formel 8.1 durch konkrete Werte ergänzt und umgeformt werden kann:

$$\begin{aligned} GLQ &= 1 - \frac{\Delta g(x_s)}{\Delta g(x)} \\ &= 1 - \frac{\Delta g(130_{ICC})}{\Delta g(130)} \\ &= 1 - \frac{g(130_{ICC}) - g(Stille_{ICC})}{g(130) - g(Stille)} \\ &= 1 - \frac{(53 - 50)}{(60 - 50)} \\ &= 1 - \frac{3}{10} \\ &= 0,7 \hat{=} 70 \% \text{ Reduktion des Lombard-Effekts.} \end{aligned} \quad (8.2)$$

In der folgenden Tabelle 8.5 werden zwei fiktive ICC-Systeme anhand des GLQ gegenübergestellt. Dabei wird in den Kommunikationssituationen ohne ICC-System, mit ICC_1 und mit ICC_2 jeweils das Ausmaß des Lombard-Effekts von Stille zu 130 km/h im Parameter der akustischen Intensität gemessen und verglichen.

Tabelle 8.5: Die Verbesserung der Kommunikationsqualität durch mehrere ICC-Systeme, gemessen anhand des Lombard-Quotienten in einem Beispiel zur durchschnittlichen Intensität.

System	Stille	130 km/h	$\Delta_{130km/h}$	Lombard-Reduktion	GLQ
-	50 dB	60 dB	10 dB	-	-
ICC_1	50 dB	53 dB	3 dB	70 %	0,7
ICC_2	50 dB	51 dB	1 dB	90 %	0,9

Tabelle 8.5 zeigt beispielhaft auf, wie die Qualitäten zweier ICC-Systeme anhand des GLQ miteinander verglichen werden können. Im Gegensatz zur Situation ohne ICC-System zeigt ICC_1 bei 130 km/h eine Abschwächung des Lombard-Effekts um 7 dB (70 %), ICC_2 sogar um 9 dB (90 %). ICC_1 erhält dafür einen GLQ von 0,7 und ICC_2 einen von 0,9. Je mehr sich der GLQ demnach einem Wert von 1 annähert, desto größer ist die kommunikationsverbessernde Qualität des Systems. Wird der niedrigste Wert von 0 erreicht, führt die Anwendung des Systems zu keiner Verbesserung im untersuchten Parameter. Durch diese Methode wird weiterhin berücksichtigt, dass sich aufgrund der Entstehung einer neuen Kommunikationssituation im Fahrzeug mit ICC-System die Messwerte der stillen Referenzbedingung ändern können. Bei der Messung in mehreren Geräuschenarien kann zudem eine Verlaufskurve mit mehreren Lombard-Quotienten erstellt werden, anhand derer festgestellt werden kann, wie linear die Verbesserung der Kommunikationsqualität durch das ICC-System verläuft. Dabei ist der GLQ auf verschiedene Parameter anwendbar. Aufgrund der Ergebnisse der vorliegenden Untersuchung wird die Anwendung auf die Intensität und die Grundfrequenz des Sprachsignals sowie die Dialogdauer empfohlen. Auch die Wahrnehmung des Lombard-Effekts im Sinne der Lautstärke bzw. die Perzeption von Emotionalität in den Sprachsignalen kann auf diese Weise analysiert werden.

Durch die Anwendung des Lombard-Effekts als Maß der Kommunikationsqualität kann die Evaluierung eines ICC-Systems und dadurch der Vergleich mehrerer ICC-Systeme standardisiert werden. Die Erhebung spontansprachlicher, interaktiver Dialoge und die Anwendung einer realitätsnahen Umgebungssimulation, durch welche eine Geräuschbefreiung der Sprachsignale zur Analyse unverzerrter akustischer Parameter möglich wird, sind dabei vergleichsweise zeitaufwendig. Deshalb könnten

die bisherigen, nicht-kommunikativen Evaluierungsstrategien (vgl. Kapitel 2.2) als Vortests eingesetzt werden und die hier vorgeschlagene Methode der finalen Analyse der Kommunikationsverbesserung dienen. Zudem kann eine Sprechaufgabe verwendet werden, die kürzer ist als die in dieser Arbeit beschriebene, und die nach ihrer Erstellung wiederholt angewendet werden kann. Dabei ist es auch möglich, eine andere Sprechaufgabe zu benutzen, wie z. B. die des *Video-Task* Paradigmas (vgl. Kapitel 4.2). Da die Proband*innen sich über die Gemeinsamkeiten und Unterschiede von Videos unterhalten, welche sie sich vor den Sprachaufnahmen angesehen haben, wird hierbei kein Kartenmaterial benötigt. Dadurch ist eine stärker inhaltsgezielte und direktere Kommunikation zwischen den Dialogpartner*innen möglich. Durch die Erhebung spontansprachlicher, interaktiver Dialoge und die Anwendung des Lombard-Effekts wird bei der Evaluierung von ICC-Systemen das kommunikative Verhalten der Passagier*innen untersucht. Erst dadurch kann getestet werden, ob ICC-Systeme tatsächlich einen positiven Effekt auf die Kommunikationsqualität der Fahrgäste haben.

A Anhang

A.1 Die Karten des *Map-Task* Szenarios

Die im Folgenden dargestellten Kartenaufgaben des *Map-Task* Szenarios sind jeweils mit einem *Label* versehen, welches auf die Nummerierung des Kartenpaares sowie die Sprechrolle der teilnehmenden Person verweist, die diese Karte erhielt. Dabei bezieht sich z. B. das *Label* „M1_G“ auf Kartenpaar Nummer eins und die Sprechrolle des *Givers*.

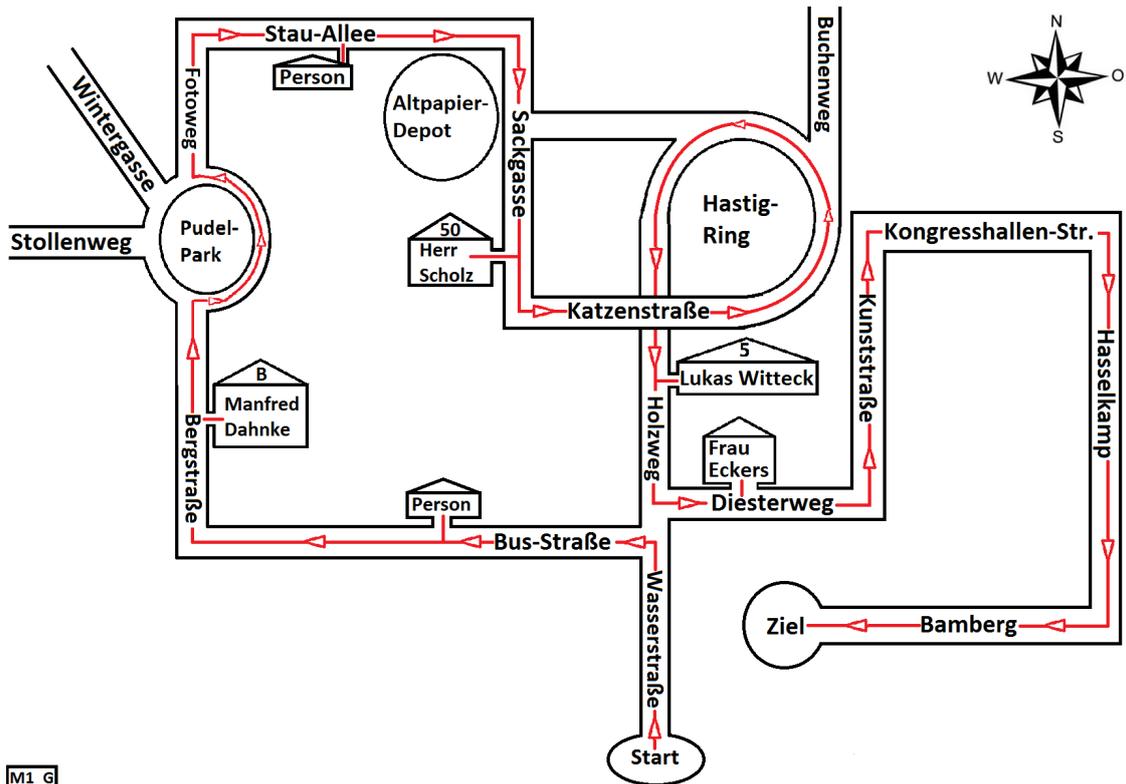


Abbildung A.1: Die erste Aufgabe des *Map-Task* Szenarios, Sprechrolle *Giver*.

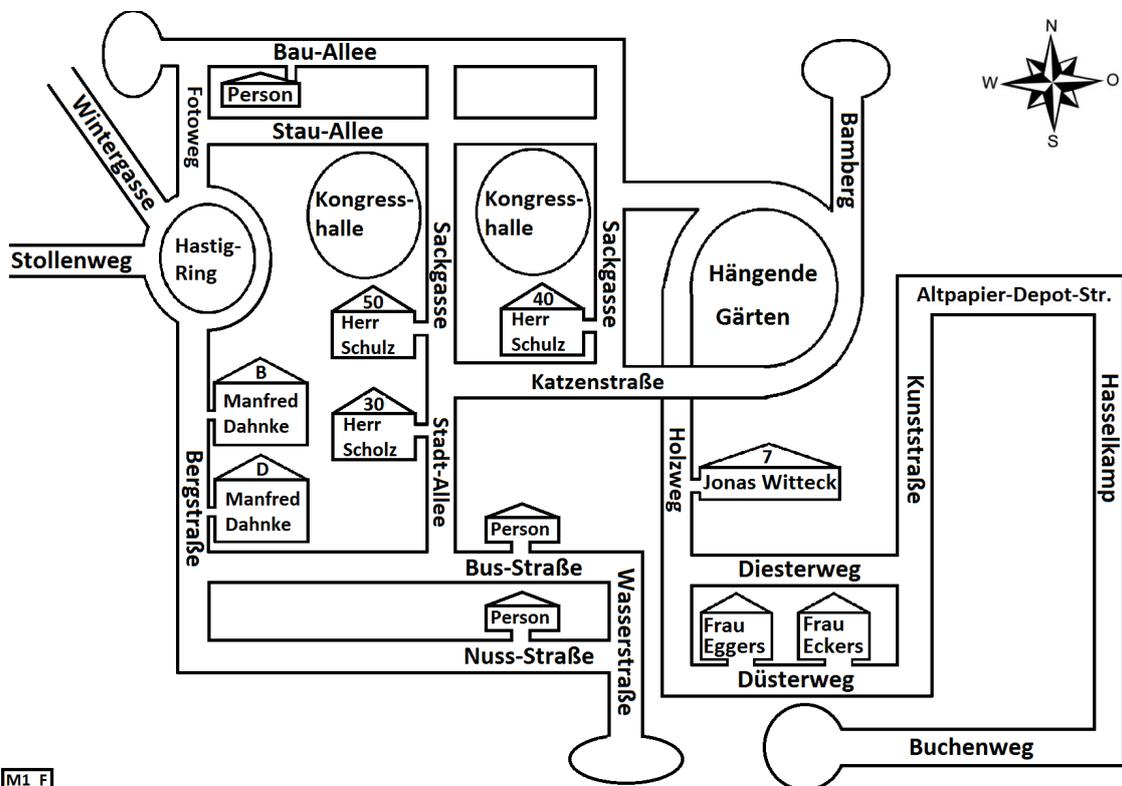
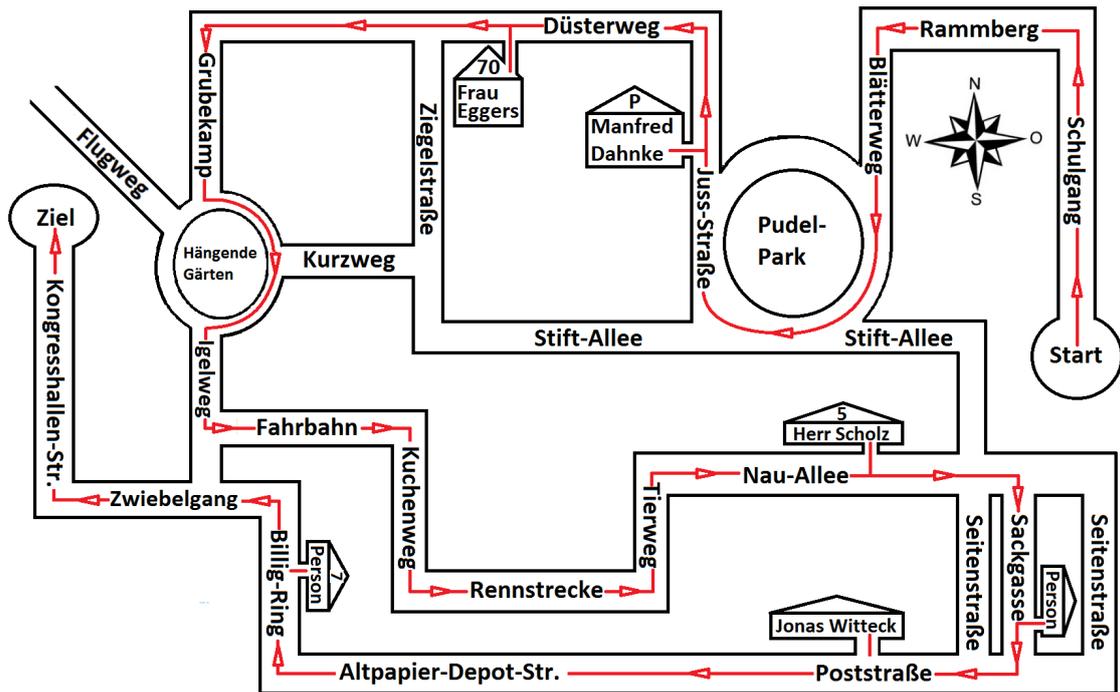
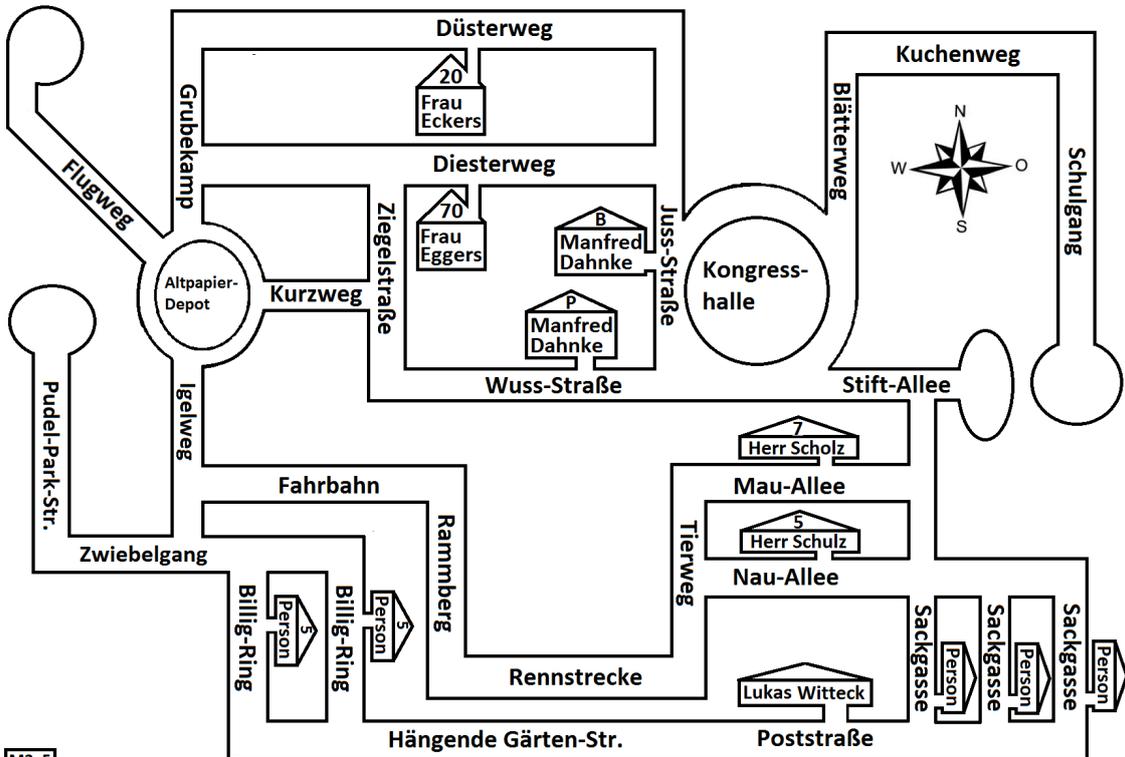


Abbildung A.2: Die erste Aufgabe des *Map-Task* Szenarios, Sprechrolle *Follower*.



M2_G

Abbildung A.3: Die zweite Aufgabe des Map-Task Szenarios, Sprechrolle *Giver*.



M2_F

Abbildung A.4: Die zweite Aufgabe des Map-Task Szenarios, Sprechrolle *Follower*.

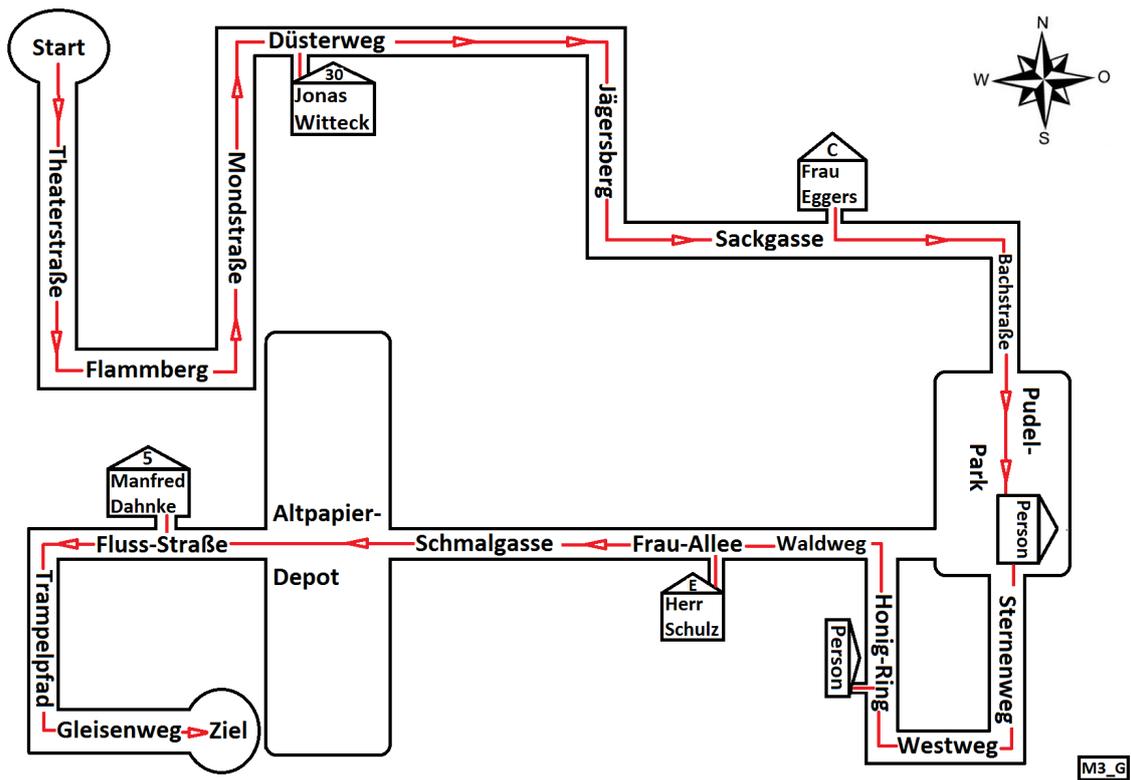


Abbildung A.5: Die dritte Aufgabe des *Map-Task* Szenarios, Sprechrolle *Giver*.

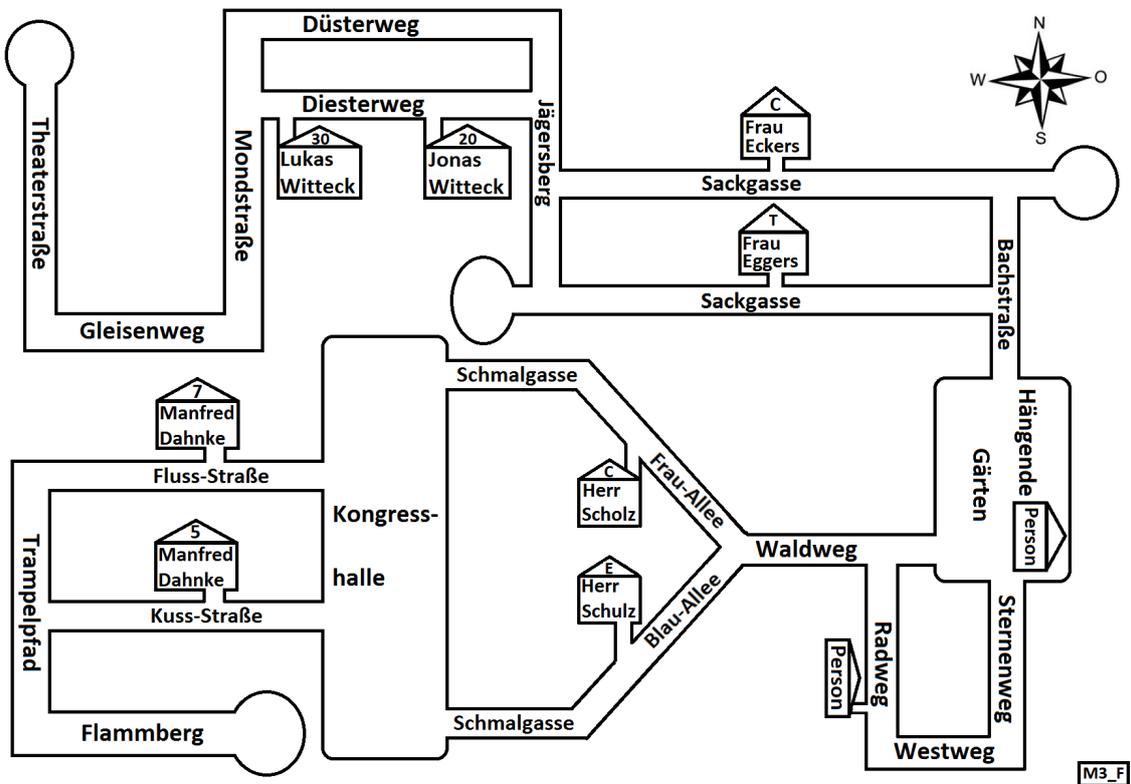


Abbildung A.6: Die dritte Aufgabe des *Map-Task* Szenarios, Sprechrolle *Follower*.

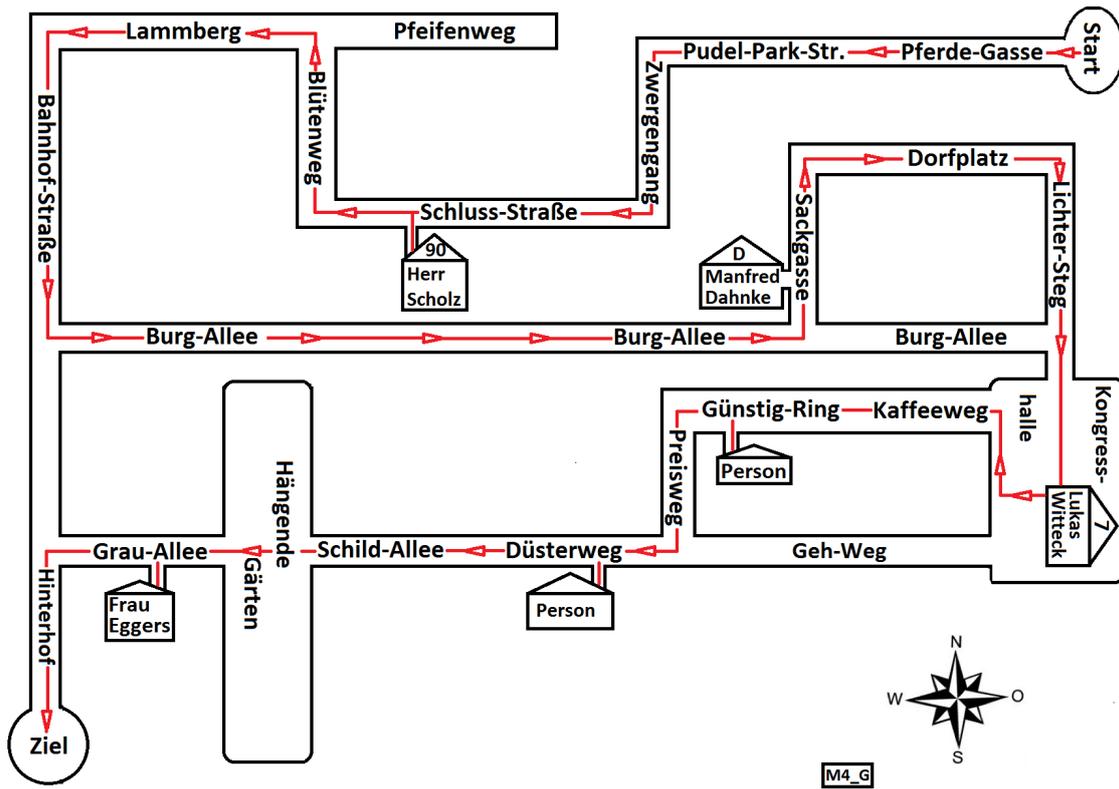


Abbildung A.7: Die vierte Aufgabe des Map-Task Szenarios, Sprechrolle *Giver*.

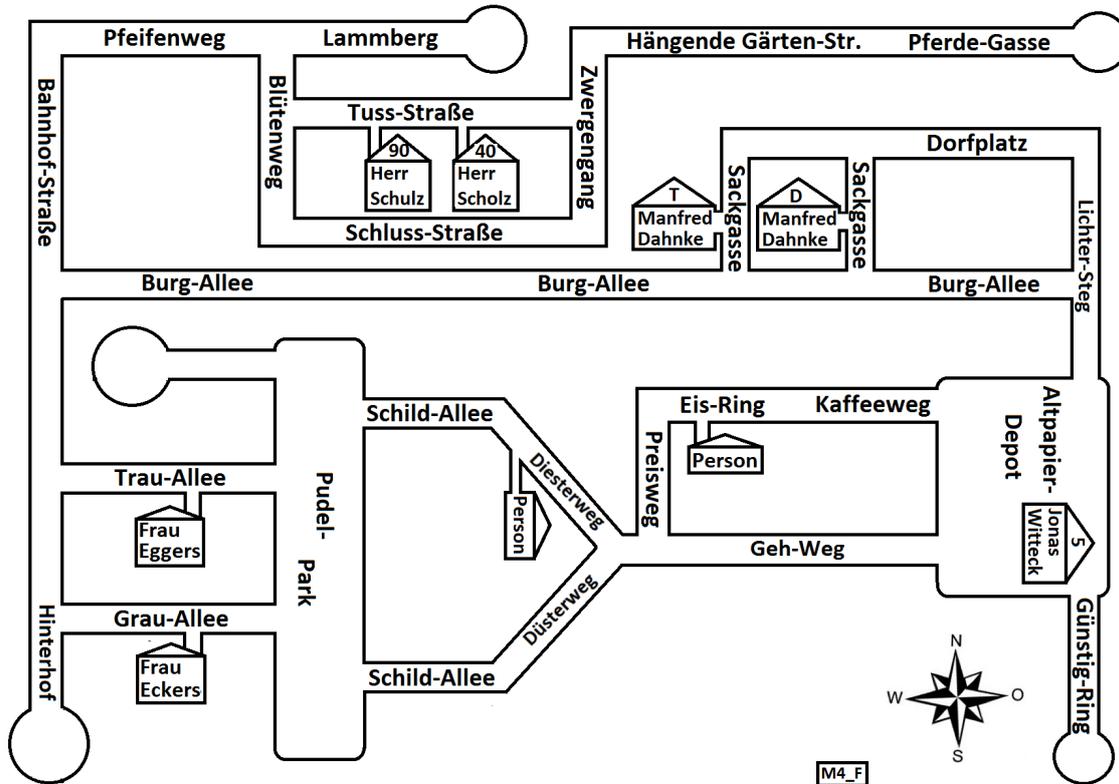


Abbildung A.8: Die vierte Aufgabe des Map-Task Szenarios, Sprechrolle *Follower*.

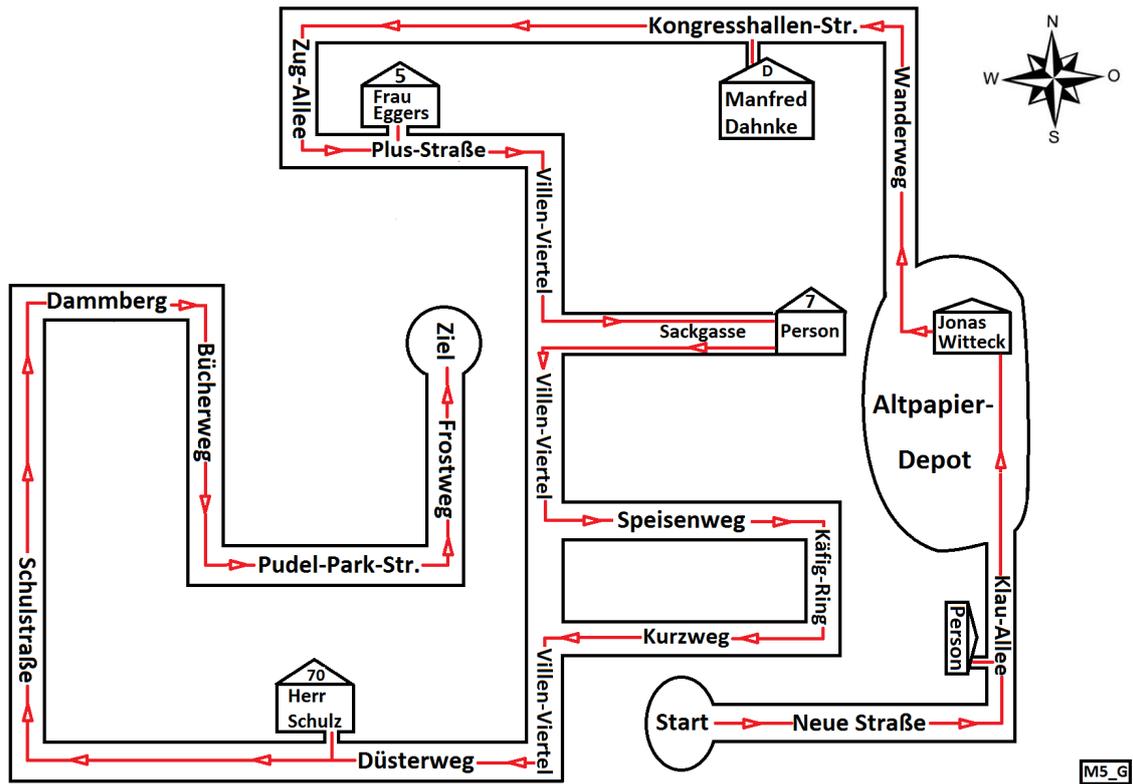


Abbildung A.9: Die fünfte Aufgabe des *Map-Task* Szenarios, Sprechrolle *Giver*.

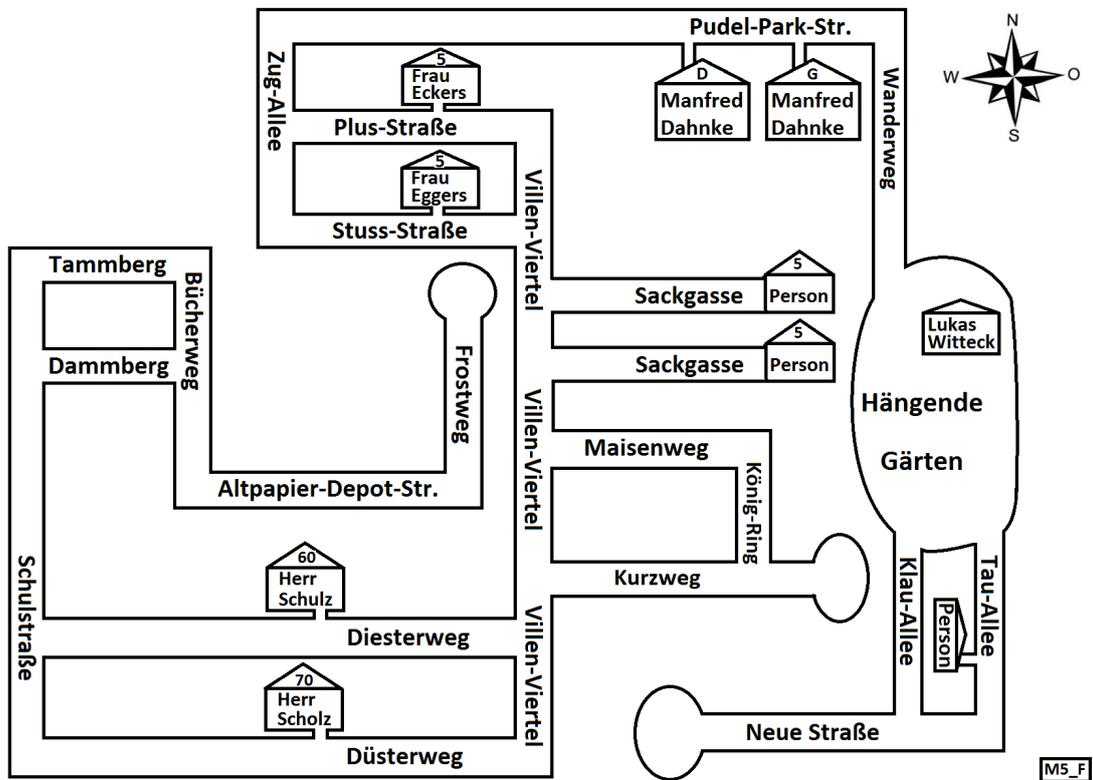


Abbildung A.10: Die fünfte Aufgabe des *Map-Task* Szenarios, Sprechrolle *Follower*.

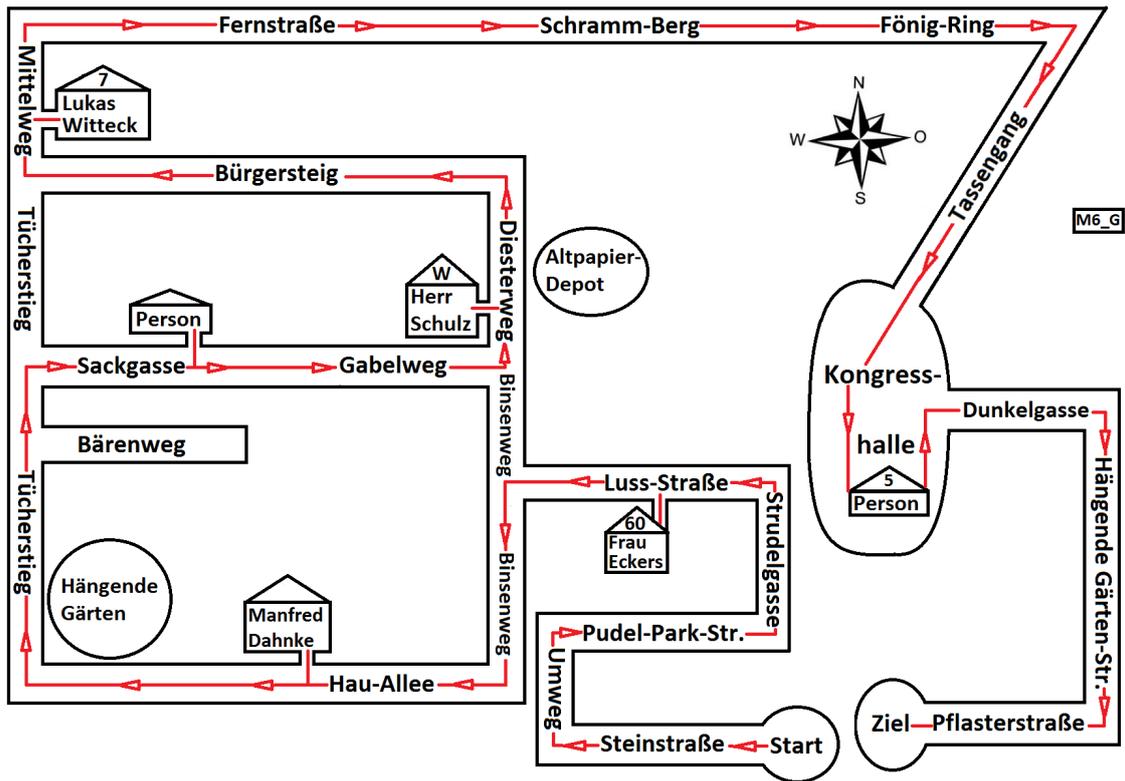


Abbildung A.11: Die sechste Aufgabe des Map-Task Szenarios, Sprechrolle *Giver*.

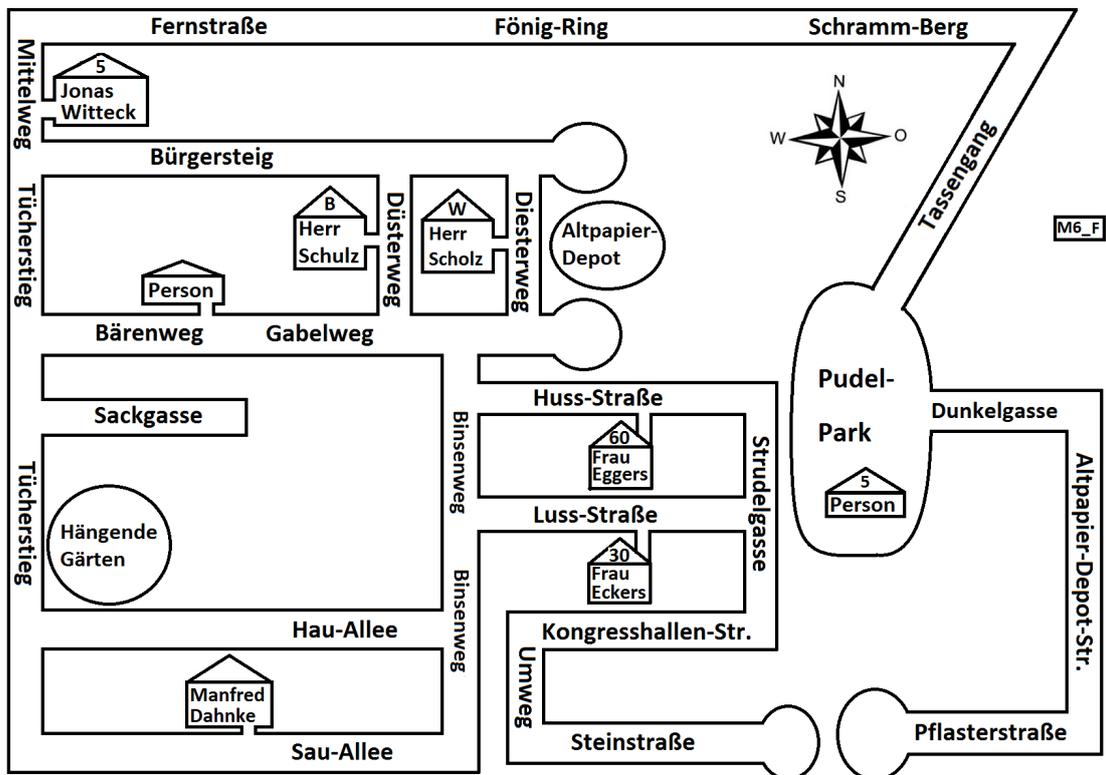


Abbildung A.12: Die sechste Aufgabe des Map-Task Szenarios, Sprechrolle *Follower*.

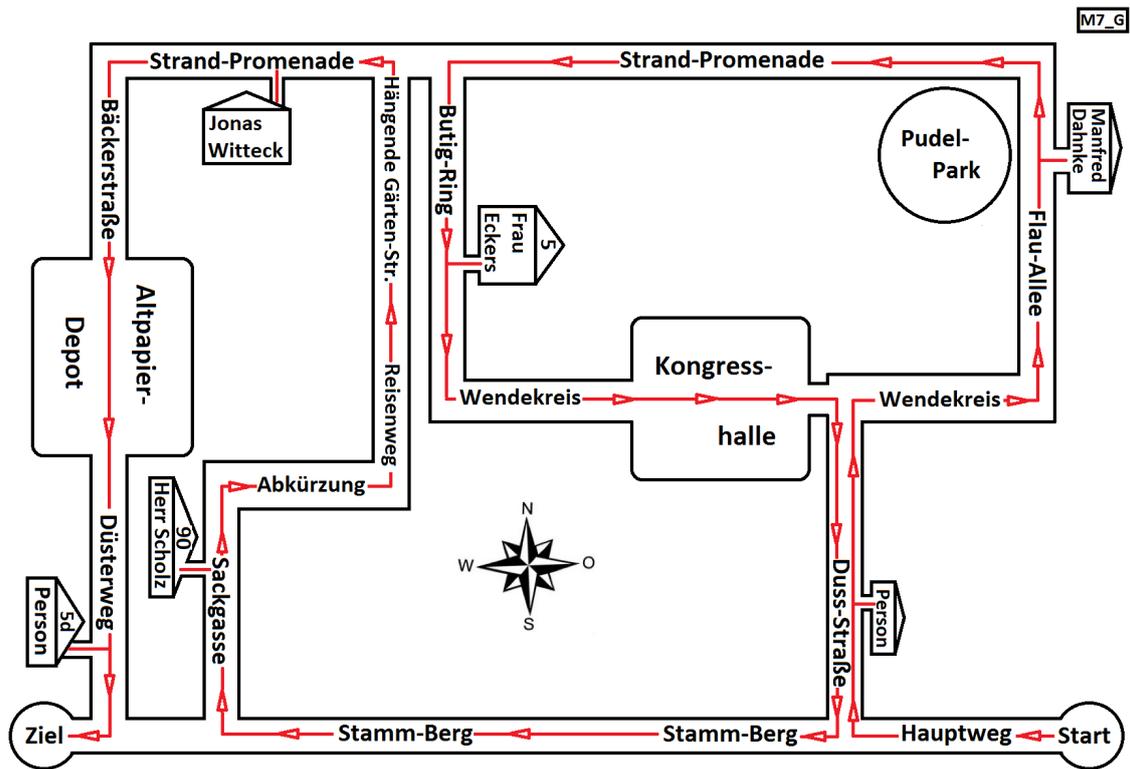


Abbildung A.13: Die siebte Aufgabe des *Map-Task* Szenarios, Sprechrolle *Giver*.

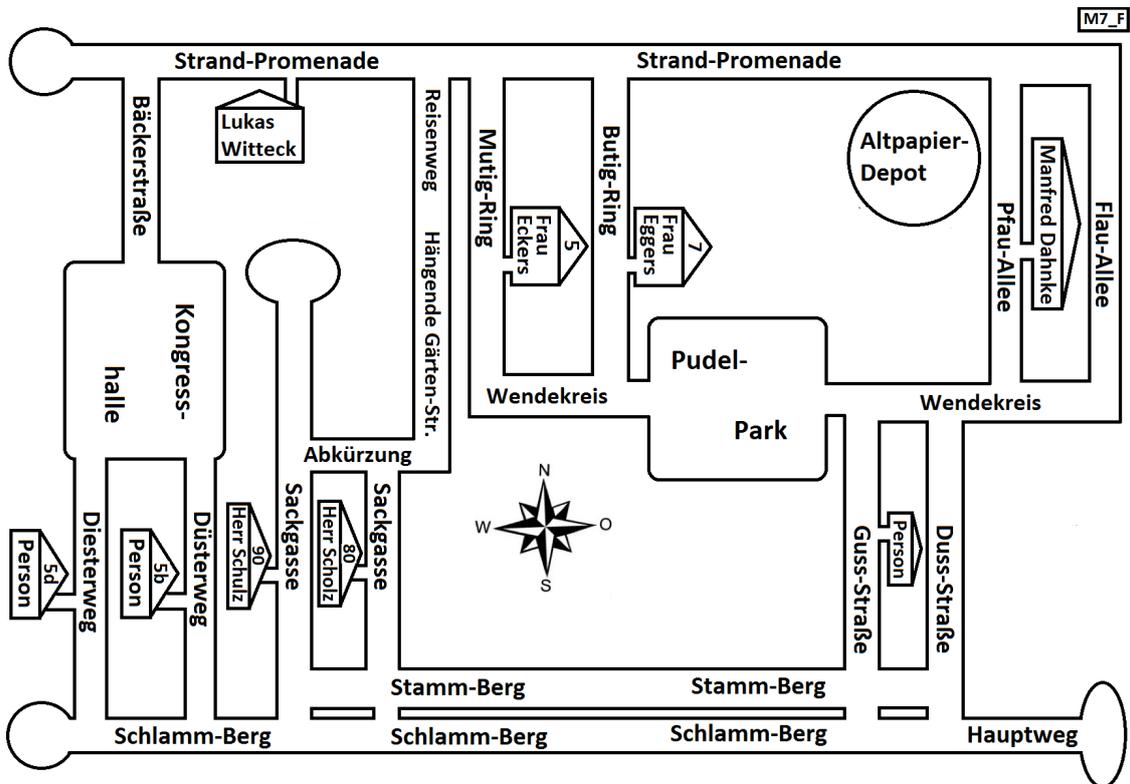


Abbildung A.14: Die siebte Aufgabe des *Map-Task* Szenarios, Sprechrolle *Follower*.

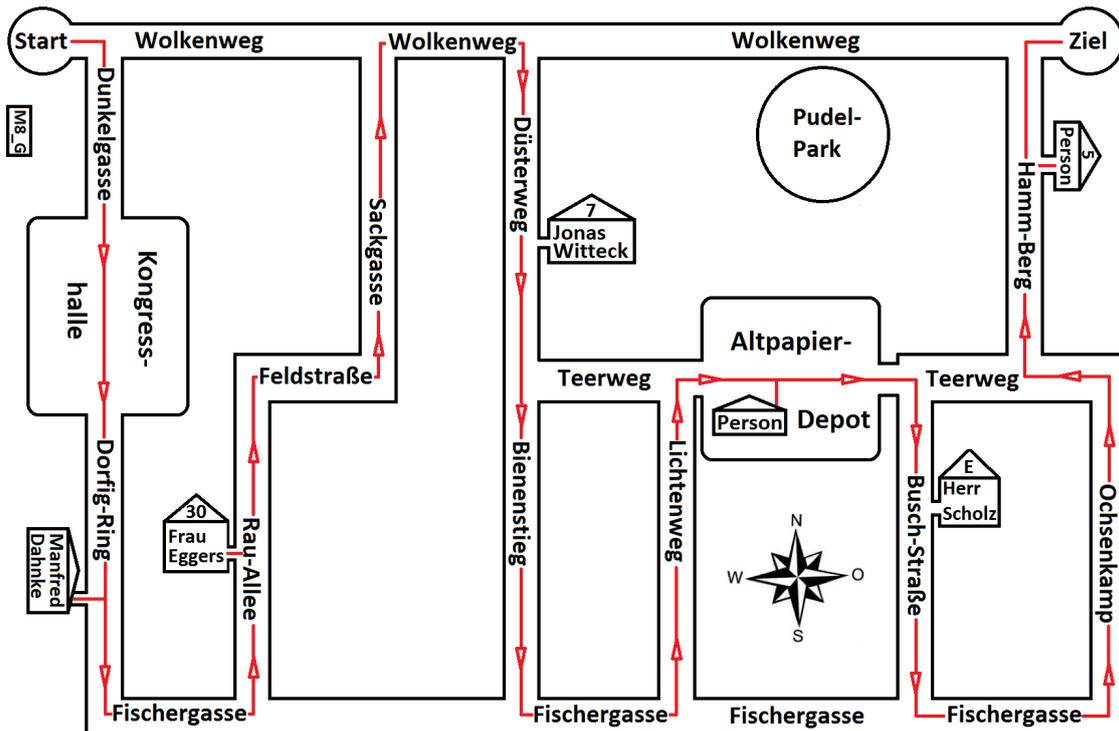


Abbildung A.15: Die achte Aufgabe des Map-Task Szenarios, Sprechrolle *Giver*.

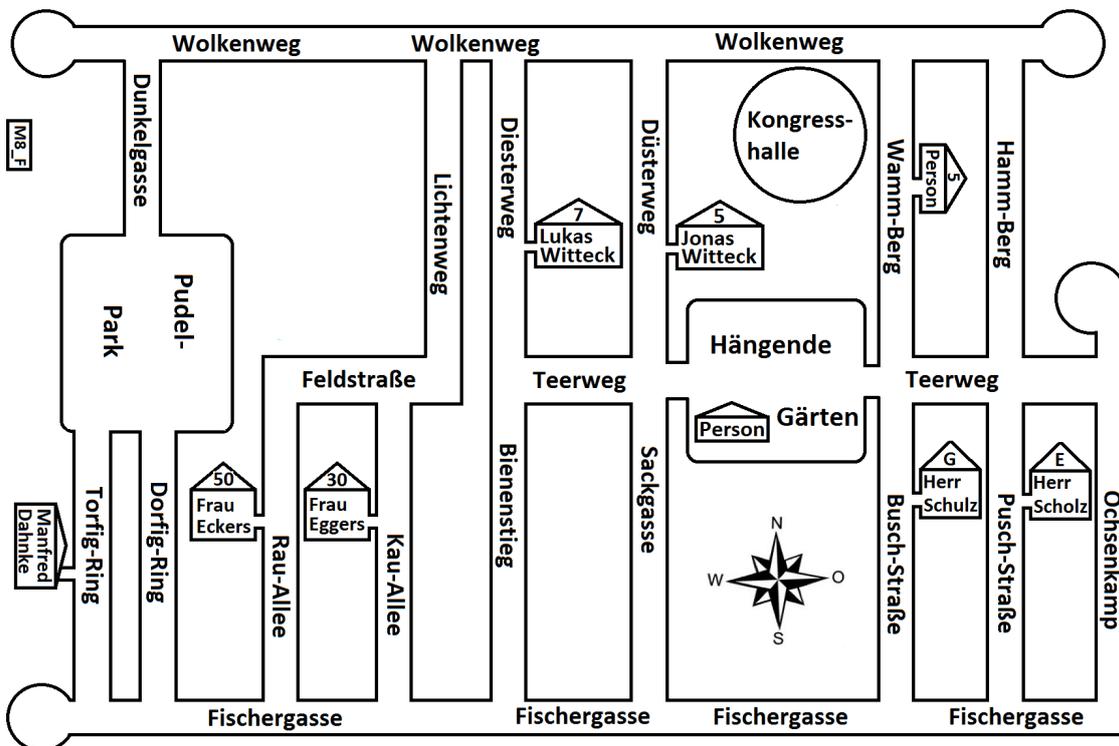


Abbildung A.16: Die achte Aufgabe des Map-Task Szenarios, Sprechrolle *Follower*.

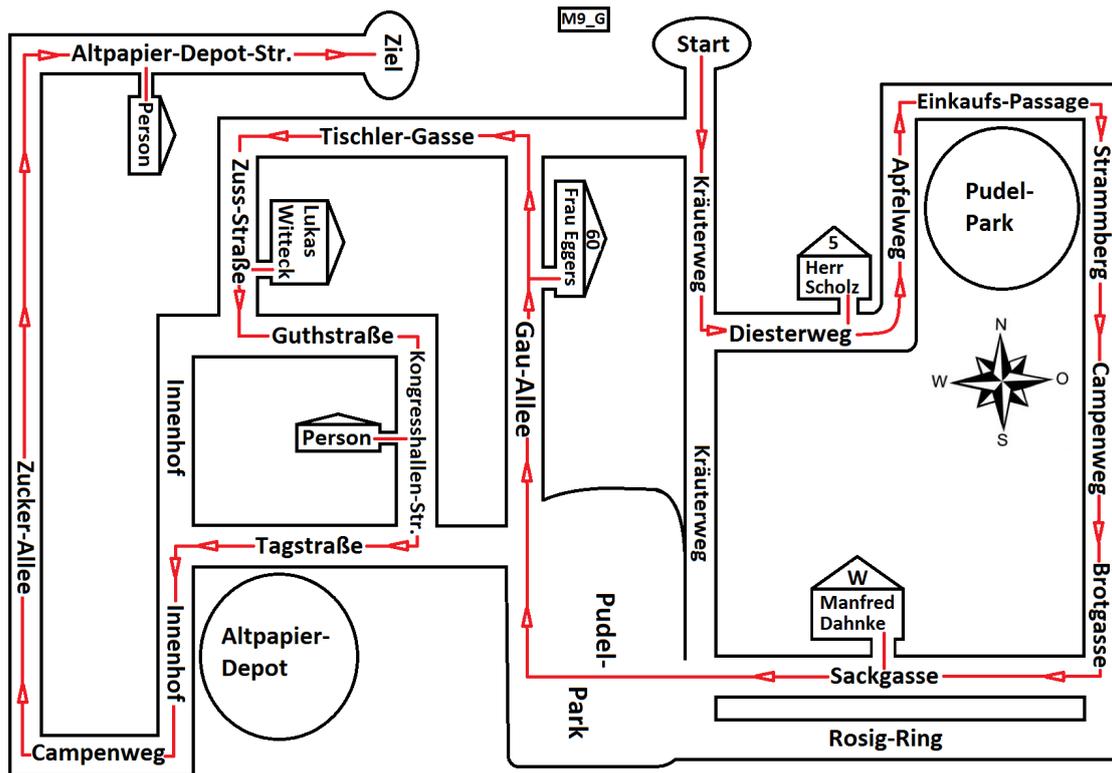


Abbildung A.17: Die neunte Aufgabe des *Map-Task* Szenarios, Sprechrolle *Giver*.

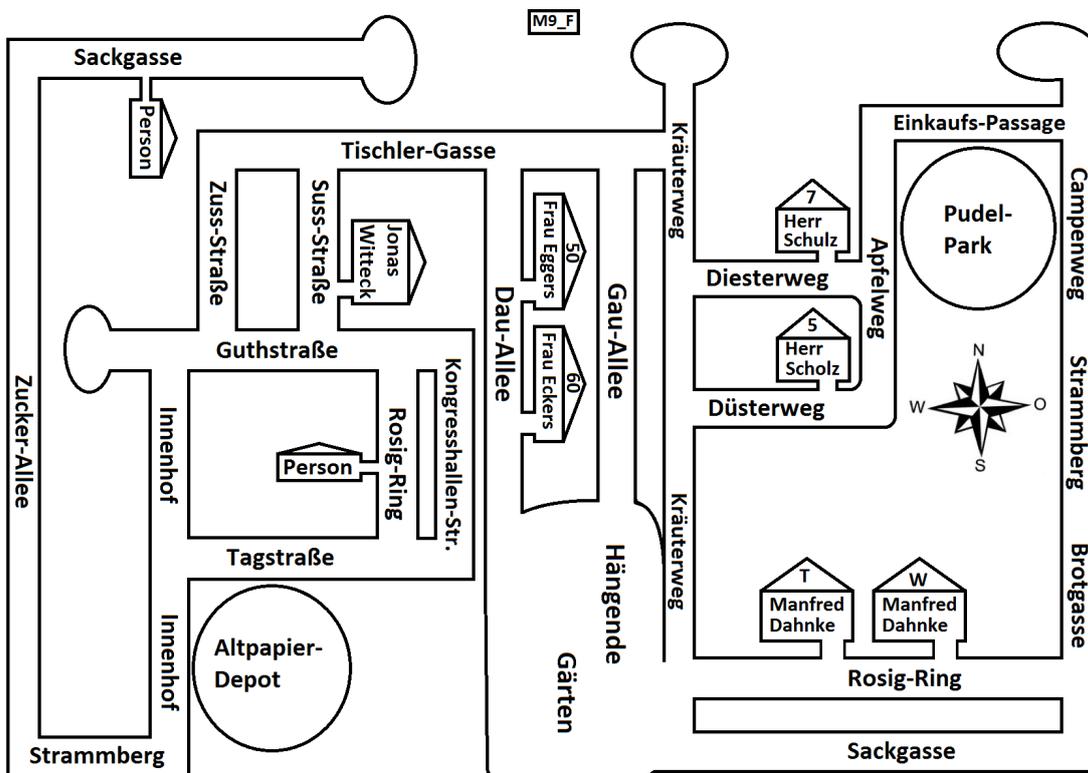


Abbildung A.18: Die neunte Aufgabe des *Map-Task* Szenarios, Sprechrolle *Follower*.

M10_G

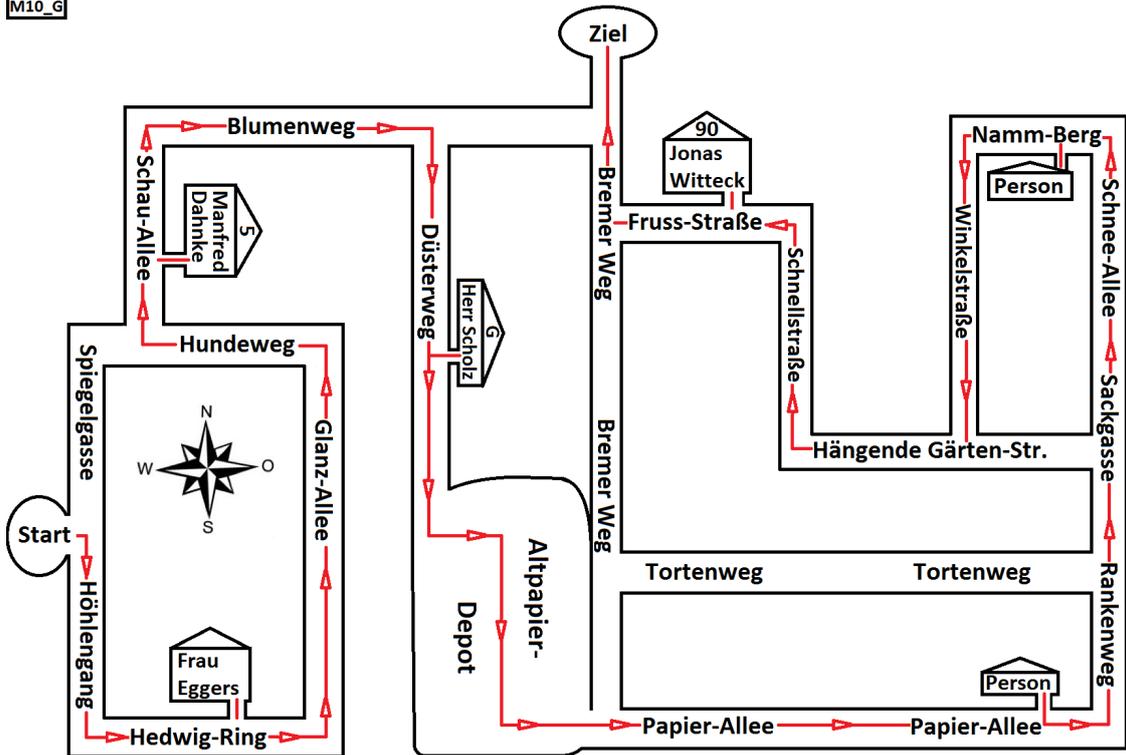


Abbildung A.19: Die zehnte Aufgabe des Map-Task Szenarios, Sprechrolle *Giver*.

M10_F

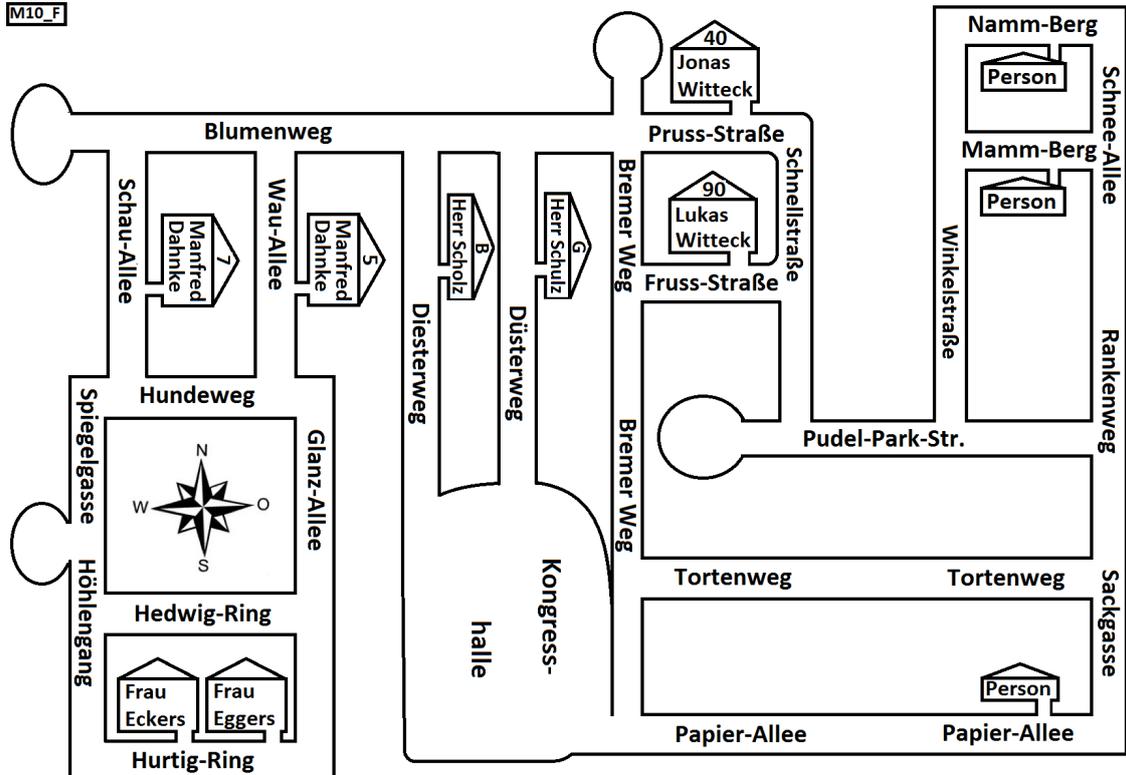


Abbildung A.20: Die zehnte Aufgabe des Map-Task Szenarios, Sprechrolle *Follower*.

A.2 Die Instruktionen des *Map-Task* Szenarios

Im Folgenden sind Instruktionen zum Produktionsexperiment abgebildet, welche sich an weibliche Teilnehmerinnen richten. Diese Texte sind für männliche Probanden der Studie entsprechend ihres Geschlechts umgeschrieben worden. Letztere werden aufgrund der geringfügigen Unterschiede, welche den Inhalt der Experimentanweisungen nicht verändern, hier nicht mit aufgeführt.

Instruction-Giver:

Du bist zur Hochzeit deines Cousins eingeladen und sollst auf dem Weg dorthin noch weitere Gäste abholen. Dein Cousin hat dir eine Skizze von der Route gezeichnet, die du nehmen musst. Darin sind auch die Häuser der anderen Personen eingetragen.

Leider kannst du nun doch nicht zur Hochzeit gehen, weil dir ein wichtiger Termin dazwischen gekommen ist. Deshalb bittest du eine Freundin, für dich als Fahrer einzuspringen.

Deine Aufgabe ist es nun, deiner Freundin mündlich den Weg zu erklären und die Personen zu nennen, die abgeholt werden müssen. Dazu benutzt du die Skizze, die dein Cousin angefertigt hat. Deine Freundin zeichnet die Route auf einer echten Stadtkarte nach. Aufgrund von Informationen aus dem Telefonbuch sind darauf bereits die Häuser verschiedener Leute abgebildet.

Das Problem ist, dass dein Cousin viele Fehler in seiner selbst gezeichneten Karte gemacht hat, die du in den Händen hältst. So können z.B. Personen, Straßennamen und Hausnummern von deinem Cousin falsch notiert worden sein. Während du deiner Freundin mitteilst, wie die Route verläuft und welche Personen abzuholen sind, müsst ihr diese Fehler entdecken und ausräumen.

Es ist deshalb wichtig, dass deine Beschreibung möglichst genau ist und du deiner Freundin alle Straßen, Hausnummern und Personen nennst, die auf deiner Karte abgebildet sind!
Nutze zur Beschreibung der Route auch die jeweiligen Himmelsrichtungen Nord, Süd, Ost und West.

Bei Unstimmigkeiten in den Straßennamen hat im Zweifelsfall immer deine Freundin recht, da sie eine echte Stadtkarte besitzt! Fehler, die in deiner Skizze auftauchen, musst du entsprechend mit einem Stift korrigieren! Bei Unstimmigkeiten in den Namen und Hausnummern der Personen hast immer du recht!

Instruction-Follower:

Deine Freundin ist zur Hochzeit ihres Cousins eingeladen und muss auf dem Weg dorthin noch andere Gäste abholen. Leider ist ihr nun ein wichtiger Termin dazwischen gekommen, sodass sie nicht zur Hochzeit fahren kann. Deshalb hast du dich bereit erklärt, für sie als Fahrer einzuspringen.

Der Cousin deiner Freundin hat ihr eine selbst gezeichnete Karte gegeben, auf der die Route und die abzuholenden Personen eingezeichnet sind. Deine Freundin erklärt dir nun mündlich, was auf ihrer Skizze abgebildet ist.

Du hingegen hältst eine echte Stadtkarte in den Händen. Aufgrund von Informationen aus dem Telefonbuch sind darauf bereits die Häuser verschiedener Leute abgebildet. Deine Aufgabe ist es, die beschriebene Route und alle dazu gehörigen Informationen auf dieser Karte nachzuzeichnen.

Das Problem ist jedoch, dass die Skizze deiner Freundin viele Fehler enthält. So können z.B. Personen, Straßennamen und Hausnummern falsch notiert sein. Während deine Freundin dir mitteilt, wie die Route verläuft und welche Personen abzuholen sind, müsst ihr diese Fehler entdecken und ausräumen.

Die Beschreibung deiner Freundin muss deshalb sehr genau sein. Auch du solltest ihr während des Karten-Vergleichs alle Straßen, Hausnummern und Personen nennen, die auf deiner Karte abgebildet sind.

Nutzt zum richtigen Abgleich der Karten auch die jeweiligen Himmelsrichtungen Nord, Süd, Ost und West.

Bei Unstimmigkeiten in den Straßennamen hast im Zweifelsfall immer du recht, da du eine echte Stadtkarte besitzt! Deine Freundin muss ihre Skizze daher entsprechend korrigieren! Bei Unstimmigkeiten in den Namen und Hausnummern der Personen hat immer deine Freundin recht!

A.3 Der Fragebogen zum Produktionsexperiment

Fragebogen zum Experiment

Die folgenden Fragen beziehen sich auf die Sprachaufnahmen, an denen du vor etwa 2 Monaten und gerade eben teilgenommen hast. Als Antworten auf die jeweiligen Fragen kreise eine Zahl bzw. ein Kreuz auf der dazugehörigen Skala ein. Falls nicht anders beschrieben, gib jeweils nur eine Antwort an. Bitte lies alle Fragen erst einmal komplett durch, bevor du deine Antworten einträgst.

1. Bei der zweiten Aufnahmerunde, die soeben stattgefunden hat:
Hast du die Sprache deiner/s Freundin/es häufig über Lautsprecher gehört, oder kam dies nur selten vor?

1	2	3	4	5
immer	häufig	teilweise	selten	nie

2. Hast du deine/n Freund/in häufiger über Lautsprecher gehört, je lauter die Hintergrundgeräusche wurden?

1	2	3	4	5
viel häufiger	etwas häufiger	kein Unterschied	etwas seltener	viel seltener

3. Wie laut hast du deine/n Freund/in über Lautsprecher gehört in Relation zu ihrer/seiner eigentlichen Stimme?

1	2	3	4	5
viel lauter	etwas lauter	kein Unterschied	etwas leiser	viel leiser

4. Hast du deine/n Freund/in mit der Lautsprecher-Übertragung besser verstehen können als ohne?

1	2	3	4	5
viel besser	etwas besser	kein Unterschied	etwas schlechter	viel schlechter

5. Konntest du deine eigene Stimme während der Dialoge über die Lautsprecher hören?

1	2	3
Ja	nur selten	nein

A Anhang

6. Hat es dich gestört, wenn du deine/n Freund/in über Lautsprecher gehört hast?

1	2	3	4	5
ja	etwas	kein Unterschied	es war eher angenehm	nein, ich mochte es

7. Hat dich die Übertragung von Sprache über die Lautsprecher generell gestört?

1	2	3	4	5
Ja	etwas	kein Unterschied	es war eher angenehm	nein, ich mochte es

8. Falls du bei Frage 6 und/oder 7 Nummer 1 oder 2 eingekreist hast:

Hat es dich gestört, deine eigene Stimme zu hören, oder die deiner/s Freundin/es?
(Hierbei sind mehrere Antworten möglich).

1	2	3
meine eigene	die meiner/s Freundin/es	beides

A.3 Der Fragebogen zum Produktionsexperiment

9. Falls du die Sprach-Übertragung durch Lautsprecher als störend empfunden hast, was genau hat dich daran gestört? (Hierbei sind mehrere Antworten möglich).

X	X	X	X
Die Sprache klang merkwürdig.	Ich habe meine eigene Stimme durch die Lautsprecher gehört.	Die Lautsprecher waren zu laut.	Die Lautsprecher waren zu leise.

Sonstige Antworten: _____

10. Falls dir die Sprach-Übertragung durch Lautsprecher gefallen hat, was genau fandest du daran gut? (Hierbei sind mehrere Antworten möglich).

X	X	X	X
Die Sprache klang natürlich.	Die Sprache war nicht zu laut oder zu leise.	Ich konnte meine/n Freund/in besser verstehen.	Die Kommunikation wurde leichter.

Sonstige Antworten: _____

Sonstige Anmerkungen: _____

Vielen Dank für deine Mitarbeit!

A.4 Der *Online*-Fragebogen des Perzeptionsexperiments

In diesem Kapitel sind zunächst die erste Seite des *Online*-Fragebogens sowie ein Beispiel des Semantischen Differentials abgebildet, das die Proband*innen der Studie bearbeitet haben und welches mit insgesamt 18 Seiten den Mittelteil der Erhebung darstellte. Anschließend sind die vorletzte und die letzte Seite des Fragebogens aufgeführt.

A.4 Der Online-Fragebogen des Perzeptionsexperiments

soSci
oFb - der onlineFragebogen

0% ausgefüllt

Herzlich willkommen bei dieser Umfrage!

Der folgende Fragebogen ist Teil einer Doktorarbeit an der Christian-Albrechts-Universität zu Kiel. Um Ihnen eine möglichst unbefangene Teilnahme zu ermöglichen, werden nähere Details zu dieser Arbeit erst am Ende der Umfrage preisgegeben.

Eines ist aber sehr wichtig: Bitte nehmen Sie nur dann an der Umfrage teil, wenn Sie deutscher Muttersprachler sind.

Auf den folgenden Seiten hören Sie jeweils einen Sprachausschnitt, den Sie durch wiederholtes Klicken des *Play-Buttons* auch mehrmals abspielen können. Unter dem Play-Button finden Sie 8 Skalen. Bitte bewerten Sie das Gehörte anhand der dort angeführten Kriterien.

Bitte wechseln Sie während der Umfrage nicht zwischen Lautsprechern und Kopfhörern und verändern Sie die Lautstärke nicht mehr. Anhand des folgenden Beispiels können Sie die Lautstärke jetzt einstellen:



Es geht los!

Die Umfrage wird etwa 10-15 Minuten in Anspruch nehmen.

Fahren Sie fort, indem Sie auf *weiter* klicken. Kehren Sie zu einer vorangegangenen Seite zurück, indem Sie auf *zurück* klicken.

[Weiter](#)

R. Landgraf, Kiel – landgraf@isfas.uni-kiel.de

Abbildung A.21: Die erste Seite des *Online*-Fragebogens.

soSci
oFb - der onlineFragebogen

5% ausgefüllt



1. Wie klingt das Gesprochene?

	3	2	1	0	1	2	3	
angestrengt	<input type="radio"/>	entspannt						
steif	<input type="radio"/>	locker						
gehemmt	<input type="radio"/>	ungehemmt						
verständlich	<input type="radio"/>	unverständlich						
unfreundlich	<input type="radio"/>	freundlich						
stockend	<input type="radio"/>	flüssig						
laut	<input type="radio"/>	leise						
emotional	<input type="radio"/>	sachlich						

[Zurück](#) [Weiter](#)

R. Landgraf, Kiel – landgraf@isfas.uni-kiel.de

Abbildung A.22: Beispielseite des *Online*-Fragebogens mit Semantischem Differential.



95% ausgefüllt

Vielen Dank für Ihre Teilnahme!

Zum Schluss bitte ich Sie darum, noch zwei Angaben zu Ihrer Person zu machen und eine kurze Bewertung dieser Umfrage vorzunehmen:

Alter:

Geschlecht:

1. Wie schwierig fanden Sie die Bewertung der Sprachausschnitte?

sehr einfach
einfach
weder noch
schwierig
sehr schwierig

2. Wie schwierig fanden Sie die Bewertung der einzelnen Skalen?

	sehr einfach	einfach	weder noch	schwierig	sehr schwierig
unfreundlich – freundlich	<input type="radio"/>				
emotional – sachlich	<input type="radio"/>				
gehemmt – ungehemmt	<input type="radio"/>				
steif – locker	<input type="radio"/>				
angestrengt – entspannt	<input type="radio"/>				
stockend – flüssig	<input type="radio"/>				
laut – leise	<input type="radio"/>				
verständlich – unverständlich	<input type="radio"/>				

3. Wenn Sie die Bewertung der Sprachausschnitte bzw. der einzelnen Skalen schwierig fanden, was sind die Gründe dafür?

Zurück
Weiter

R. Landgraf, Kiel – landgraf@isfas.uni-kiel.de

Abbildung A.23: Die vorletzte Seite des *Online*-Fragebogens.

soSci
oFb - der onlineFragebogen

Ich möchte mich ganz herzlich für Ihre Mithilfe bedanken!

Bitte leiten Sie den Link dieser Umfrage an andere Teilnehmer weiter.

Ihre Antworten wurden gespeichert, Sie können das Browser-Fenster nun schließen.

Zur ihrer Information:

Diese Studie widmet sich der Erforschung von lautsprachlicher Kommunikation im fahrenden Auto. Dabei wird untersucht, ob die Kommunikation zwischen den Passagieren bei lauten Fahrgeräuschen verbessert werden kann, wenn ein sogenanntes Innenraum-Kommunikationssystem (ICC, In-Car Communication) angewendet wird. Die von Ihnen soeben beurteilten Sprachaufnahmen wurden im Auto bei unterschiedlichen Fahrgeräuschen, jeweils mit und ohne ICC-System aufgenommen. Durch Ihre Bewertung und die Bewertungen weiterer Teilnehmer können wir feststellen, wie sich die Sprache der Passagiere bei lauten Fahrgeräuschen verändert und ob die Sprecher z. B. (un)verständlicher oder angespannter klingen. Zudem wird überprüft, ob die bei Fahrgeräuschen auftretenden Veränderungen in der Sprache durch Aktivierung des ICC-Systems wieder reduziert werden können.

Für Nutzer von SurveyCircle (www.surveycircle.com): Der Survey Code dieser Studie lautet:
M1YS-G58L-BZ5Y-PLJ4

Einladung zum SoSci Panel

Guten Tag,

das nicht-kommerzielle **SoSci Panel** würde Sie gerne zu interessanten wissenschaftlichen Onlinebefragungen einladen. Wir würden uns sehr freuen, wenn Sie die Forschung durch Ihre Teilnahme unterstützen.

E-Mail:

Das SoSci Panel sendet Ihnen pro Jahr maximal 4 Einladungen per E-Mail. Sie gehen keine Verpflichtung ein und können die Teilnahme jederzeit mit drei Klicks beenden. Wenn Sie oben Ihre E-Mail-Adresse angeben, erhalten Sie zunächst eine Bestätigungsmail mit welcher Sie Ihre Teilnahme bestätigen oder widerrufen können.

Das SoSci Panel unterliegt dem strengen deutschen Datenschutzrecht. Wir senden Ihnen keine Werbung und geben Ihre E-Mail-Adresse selbstverständlich nicht an Dritte weiter.

Der Fragebogen, den Sie gerade ausgefüllt haben, wurde gespeichert. Sie können das Browserfenster selbstverständlich auch schließen, ohne am SoSci Panel teilzunehmen.

R. Landgraf, Kiel – landgraf@isfas.uni-kiel.de

Abbildung A.24: Die letzte Seite des *Online*-Fragebogens.

Literaturverzeichnis

- [1] **Amazi, Deborah K., Sharon R. Garber (1982)**
The Lombard sign as a function of age and task. In: Journal of Speech and Hearing Research 25(4), S. 581-585.
- [2] **Arnold, Ulrike, Alexandra Hinz, Verena Mahlstedt, Miriam Mailahn, Aydan Wilck (2014)**
Psyhyrembel Klinisches Wörterbuch. 266. Auflage, Berlin, Boston: Walter de Gruyter.
- [3] **Bond, Z. S., Thomas J. Moore, Beverley Gable (1989)**
Acoustic-phonetic characteristics of speech produced in noise and while wearing an oxygen mask. In: The Journal of the Acoustical Society of America 85(2), S. 907-912.
- [4] **Bottalico, Pasquale, Simone Graetzer, Eric J. Hunter (2016)**
Effects of speech style, room acoustics, and vocal fatigue on vocal effort. In: The Journal of the Acoustical Society of America 139(5), S. 2870-2879.
- [5] **Bußmann, Hadumod (2008)**
Lexikon der Sprachwissenschaft. 4. Auflage, Stuttgart: Alfred Kröner Verlag.
- [6] **Cowan, James P. (2016)**
The effects of sound on people. Chichester: John Wiley & Sons.
- [7] **Davis, Chris, Jeesun Kim, Katja Grauwinkel, Hansjörg Mixdorff (2006)**
Lombard speech: auditory (A), visual (V) and AV effects. In: Proc. of the 3rd International Conference of Speech Prosody, S. 361-365. Dresden, Deutschland.

- [8] **Davis, Chris, Jeesun Kim, Amanda Sironic (2006)**
In-quiet and Lombard audiovisual speech. In: Proc. of the 7th International Seminar on Speech Production. Ubatuba, Brasilien.
- [9] **Dombrowski, Ernst (2003)**
Semantic features of accent contours: effects of F₀ peak position and F₀ time shape. In: Proc. of the 15th International Congress of Phonetic Sciences, S. 1217-1220. Barcelona, Spanien.
- [10] **Dombrowski, Ernst (2013)**
Semantic features of 'stepped' versus 'continuous' contours in german intonation. In: *Phonetica* 70(4), S. 247-273.
- [11] **Dombrowski, Ernst, Oliver Niebuhr (2005)**
Acoustic patterns and communicative functions of phrase-final F₀ rises in german: activating and restricting contours. In: *Phonetica* 62(2-4), S. 176-195.
- [12] **Dreher, John J., John J. O'Neill (1957)**
Effects of ambient noise on speaker intelligibility for words and phrases. In: *The Journal of the Acoustical Society of America* 29(12), S. 1320-1323.
- [13] **Drugman, Thomas, Thierry Dutoit (2010)**
Glottal-based analysis of the Lombard effect. In: Proc. of the 11th Interspeech Conference, S. 2610-2613. Makuhari, Japan.
- [14] **Folk, Laura, Florian Schiel (2011)**
The Lombard effect in spontaneous dialog speech. In: Proc. of the 12th Interspeech Conference, S. 2701-2704. Florenz, Italien.
- [15] **Fónagy, Ivan, Judith Fónagy (1966)**
Sound pressure level and duration. In: *Phonetica* 15, S. 14-21.
- [16] **Fricke, J. E. (1970)**
Syllabic duration and the Lombard effect. In: *International Audiology* 9, S. 53-57.

- [17] **Garnier, Maëva, Lucie Bailly, Marion Dohen, Pauline Welby, H el ene L evenbruck (2006)**
An acoustic and articulatory study of Lombard speech: global effects on the utterance. In: Proc. of the Interspeech and 9th International Conference on Spoken Language Processing (ICSLP 5), S. 2246-2249. Pittsburgh, USA.
- [18] **Garnier, Ma eva, Marion Dohen, H el ene L evenbruck, Pauline Welby, Lucie Bailly (2006)**
The Lombard Effect: a physiological reflex or a controlled intelligibility enhancement? In: Proc. of the 7th International Seminar on Speech Production, S. 255-262. Ubatuba, Brasilien.
- [19] **Garnier, Ma eva, Nathalie Henrich, Dani ele Dubois (2010)**
Influence of sound immersion and communicative interaction on the Lombard Effect. In: Journal of Speech, Language, and Hearing Research 53(3), S. 588-608.
- [20] **Hanley, T. D., M. D. Steer (1949)**
Effect of level of distracting noise upon speaking rate, duration and intensity. In: Journal of Speech and Hearing Disorders 14, S. 363-368.
- [21] **Hatzinger, Reinhold, Kurt Hornik, Herbert Nagel, Marco J. Maier (2014)**
R: Einf uhrung durch angewandte Statistik. 2. Auflage, Hallbergmoos: Pearson.
- [22] **Hazan, Valerie, Outi Tuomainen, Mich ele Pettinato (2016)**
Suprasegmental characteristics of spontaneous speech produced in good and challenging communicative conditions by talkers aged 9 to 14 years old. In: Journal of Speech, Language, and Hearing Research 59(6), S. 1-19.
- [23] **Hedderich, J urgen, Lothar Sachs (2012)**
Angewandte Statistik: Methodensammlung mit R. 14. Auflage, Berlin, Heidelberg: Springer.
- [24] **Jaschke, Sven (2012)**
Generierung und Analyse einer Sprachdatenbank unter Ber ucksichtigung des Lombard-Effekts. Bachelorarbeit an der Technischen Fakult at der Christian-Albrechts-Universit at zu Kiel.

- [25] **John, Tina, Oliver Niebuhr, Gerhard Schmidt, Anne Theiß (2013)**
Phonetic analyses vs. dirty signals: fixing the paradox. In: Proc. der 24. Konferenz zur Elektronischen Sprachsignalverarbeitung (ESSV), S. 1-8.
- [26] **John, Tina, Rabea Landgraf, Christian Lüke, Sebastian Rohde, Gerhard Schmidt, Anne Theiß, Jochen Withopf (2014)**
Über die Verbesserung der Sprachkommunikation in geräuschbehafteten Umgebungen. In: Niebuhr, Oliver (Hrsg.): Formen des Nicht-Verstehens, S. 185-203. Frankfurt am Main: Peter Lang.
- [27] **Junqua, Jean-Claude (1993)**
The Lombard reflex and its role on human listeners and automatic speech recognizers. In: The Journal of the Acoustical Society of America 93(1), S. 510-524.
- [28] **Junqua, Jean-Claude (1996)**
The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex. In: Speech Communication 20(1-2), S. 13-22.
- [29] **Kipp, Andreas, Maria-Barbara Wesenick, Florian Schiel (1996)**
Automatic detection and segmentation of pronunciation variants in german speech corpora. In: Proc. of the 4th International Conference on Spoken Language Processing (ICSLP), S. 106-109. Philadelphia, USA.
- [30] **Kipp, Andreas, Maria-Barbara Wesenick, Florian Schiel (1997)**
Pronunciation modeling applied to automatic segmentation of spontaneous speech. In: Proc. of the 5th European Conference on Speech Communication and Technology (EUROSPEECH), S. 1023-1026. Rhodos, Griechenland.
- [31] **Kisler, Thomas, Florian Schiel, Han Sloetjes (2012)**
Signal processing via web services: the use case WebMAUS. In: Proc. Digital Humanities 2012, S. 30-34. Hamburg, Deutschland.
- [32] **Kohler, Klaus J. (1995)**
Einführung in die Phonetik des Deutschen. 2. Auflage, Berlin, Münster, Regensburg: Erich Schmidt Verlag.

- [33] **Kohler, Klaus J. (2005)**
Timing and communicative functions of pitch contours. In: *Phonetica* 62(2-4), S. 88-105.
- [34] **Landgraf, Rabea (2013)**
Linguistische Charakteristika von Dialogen im fahrenden Auto und wie sie sich im Labor simulieren lassen. Masterarbeit an der Philosophischen Fakultät der Christian-Albrechts-Universität zu Kiel.
- [35] **Landgraf, Rabea (2014)**
Are you serious? Irony and the perception of emphatic intensification. In: Proc. of the 4th International Symposium on Tonal Aspects of Languages (TAL), S. 91-94. Nijmegen, Niederlande.
- [36] **Landgraf, Rabea, Oliver Niebuhr, Gerhard Schmidt, Tina John, Christian Lüke, Anne Theiß (2015)**
Von der Straße ins Labor: Die Modifikation der Sprachproduktion bei lauten Fahrgeräuschen. In: 41. Jahrestagung für Akustik (DAGA), S. 414-417. Nürnberg, Deutschland.
- [37] **Landgraf, Rabea (2015)**
Simulating complex speech-production environments. In: Niebuhr, Oliver, Radek Skarnitzl (Hrsg.): *Tackling the complexity in speech*, S. 97-110. Prag: Prague University Press.
- [38] **Landgraf, Rabea, Johannes Köhler-Kaeß, Christian Lüke, Oliver Niebuhr, Gerhard Schmidt (2016)**
Can you hear me now? Reducing the Lombard effect in a driving car using an In-Car Communication system. In: Proc. of the 8th International Conference on Speech Prosody, S. 479-483. Boston, USA.
- [39] **Landgraf, Rabea, Gerhard Schmidt, Johannes Köhler-Kaeß, Oliver Niebuhr, Tina John (2017)**
More noise, less talk - the impact of driving noise and In-Car Communication systems on acoustic-prosodic parameters in dialogue. In: 43. Jahrestagung für Akustik (DAGA), S. 1485-1488. Kiel, Deutschland.

- [40] **Lane, Harlan, Bernard Tranel, Cyrus Sisson (1970)**
Regulation of voice communication by sensory dynamics. In: The Journal of the Acoustical Society of America 47(2), S. 618-624.
- [41] **Lane, Harlan, Bernard Tranel (1971)**
The Lombard sign and the role of hearing in speech. In: Journal of Speech, Language, and Hearing Research 14, S. 677-709.
- [42] **Lau, Priscilla (2008)**
The Lombard effect as a communicative phenomenon. In: UC Berkeley Phonology Lab Annual Report, S. 1-9.
- [43] **Lazarus, Hans, Charlotte A. Sust, Rita Steckel, Marko Kulka, Patrick Kurtz (2007)**
Akustische Grundlagen sprachlicher Kommunikation. Berlin, Heidelberg, New York: Springer.
- [44] **Lindblom, Björn (1990)**
Explaining phonetic variation: A sketch of the H&H theory. In: Hardcastle, William J., Alain Marchal (Hrsg.): Speech Production and Speech Modelling, S. 403-439. Dordrecht, Boston, London: Kluwer Academic Publishers.
- [45] **Lu, Youyi, Martin Cooke (2008)**
Speech production modifications produced by competing talkers, babble, and stationary noise. In: The Journal of the Acoustical Society of America 124(5), S. 3261-3275.
- [46] **Lüke, Christian, Halil Özer, Gerhard Schmidt, Anne Theiß, Jochen Withopf (2011)**
Signal processing for In-Car Communication systems. In: 5th biennial workshop on DSP for In-Vehicle Systems, S. 1-8. Kiel, Deutschland.
- [47] **Lüke, Christian, Anne Theiß, Gerhard Schmidt, Oliver Niebuhr, Tina John (2013)**
Creation of a Lombard speech database using an acoustic ambiance simulation with loudspeakers. In: 6th biennial workshop on DSP for In-Vehicle Systems, S. 1-8. Seoul, Südkorea.

- [48] **Lüke, Christian, Gerhard Schmidt, Anne Theiß, Jochen Withopf (2014)**
In-Car Communication. In: Schmidt, Gerhard, Huseyin Abut, Kazuya Takeda, John H. L. Hansen (Hrsg.): Smart Mobile In-Vehicle Systems: Next Generation Advancements, S. 97-118. New York, Heidelberg, Dordrecht, London: Springer.
- [49] **Lüke, Christian, Anne Theiß, Gerhard Schmidt, Rabea Landgraf, Tina John, Oliver Niebuhr (2017)**
How to create a clean Lombard speech database using loudspeakers. In: Abut, Hüseyin, John H. L. Hansen, Kazuya Takeda, Gerhard Schmidt, Hanseok Ko (Hrsg.): Vehicle Systems and Driver Modelling: DSP, human-to-vehicle interfaces, driver behavior, and safety, S. 209-226. Berlin, Boston: Walter de Gruyter.
- [50] **Manabe, Kazuchika, Ehsanee I. Sadr, Robert J. Dooling (1998)**
*Control of vocal intensity in budgerigars (*Melopsittacus undulatus*): Differential reinforcement of vocal intensity and the Lombard effect*. In: The Journal of the Acoustical Society of America 103(2), S. 1190-1198.
- [51] **Mandel, Steven, Robert Thayer Sataloff, Sarita R. Schapiro (1993)**
Minor head trauma: Assessment, management, and rehabilitation. New York: Springer.
- [52] **Mayer, Jörg (2017)**
Phonetische Analysen mit Praat. Ein Handbuch für Ein- und Umsteiger.
<http://praatpfanne.lingphon.net/das-praat-handbuch>
- [53] **Mixdorff, Hansjörg, Ulrich Pech, Chris Davis, Jeeseun Kim (2007)**
Map Task dialogs in noise – a paradigm for examining Lombard speech. In: Proc. of the 16th International Congress of Phonetic Sciences (ICPhS), S. 1329-1332. Saarbrücken, Deutschland.
- [54] **Niebuhr, Oliver, Ernst Dombrowski (2010)**
Shaping phrase-final rising intonation in German. In: Proc. of the 5th International Conference on Speech Prosody, S. 1-4. Chicago, USA.

- [55] **Niebuhr, Oliver, Benno Peters, Rabea Landgraf, Gerhard Schmidt (2015)**
The kiel corpora of “speech & emotion” - a summary. In: 41. Jahrestagung für Akustik (DAGA), S. 1011-1014. Nürnberg, Deutschland.
- [56] **Niebuhr, Oliver, Alexis Michaud (2015)**
Speech Data Acquisition: The Underestimated Challenge. In: Niebuhr, Oliver, Sarah Buchberger (Hrsg.): Kieler Arbeiten zur Linguistik und Phonetik (KALIPHO) (Band 3), S. 1-42. Institut für Skandinavistik, Frisistik und Allgemeine Sprachwissenschaft (ISFAS), Christian-Albrechts-Universität zu Kiel.
- [57] **Osgood, Charles E., George J. Suci, Percy H. Tannenbaum (1975)**
The measurement of meaning. 9. Auflage, Illinois: University of Illinois Press.
- [58] **Patel, Rupal, Kevin W. Schell (2008)**
The influence of linguistic content on the Lombard effect. In: Journal of Speech, Language, and Hearing Research 51(1), S. 209-220.
- [59] **Pick, Herbert L. Jr., Gerald M. Siegel, Paul W. Fox, Sharon R. Garber, Joseph K. Kearney (1989)**
Inhibiting the Lombard effect. In: The Journal of the Acoustical Society of America 85(2), S. 894-900.
- [60] **Pompino-Marschall, Bernd (2009)**
Einführung in die Phonetik. 3. Auflage, Berlin, New York: Walter de Gruyter.
- [61] **Potash, L. M. (1972)**
Noise-induced changes in calls of the Japanese quail. In: Psychonomic Science 26(5), S. 252-254.
- [62] **Rutherford, Andrew (2001)**
Introducing ANOVA and ANCOVA: A GLM approach. London, Thousand Oaks, New Delhi: Sage Publications.
- [63] **Reuter, Peter (2004)**
Springer Lexikon Medizin. Berlin, Heidelberg, New York: Springer.

- [64] **Schiel, Florian (1999)**
Automatic phonetic transcription of non-prompted speech. In: Proc. of the 14th International Congress of Phonetic Sciences (ICPhS), S. 607-610. San Francisco, USA.
- [65] **Schmidt, Gerhard, Tim Haulick (2006)**
Signal processing for In-Car Communication systems. In: Hänslér, Eberhard, Gerhard Schmidt (Hrsg.): Topics in Acoustic Echo and Noise Control: Selected Methods for the Cancellation of Acoustical Echoes, the Reduction of Background Noise, and Speech Processing, S. 547-598. Berlin, Heidelberg, New York: Springer.
- [66] **Schmidt, Gerhard, Anne Theiß, Jochen Withopf, Arthur Wolf (2012)**
Evaluation of In-Car Communication systems. In: Hansen, John H. L., Pinar Boyraz, Kazuya Takeda, Hüseyin Abut (Hrsg.): Digital Signal Processing for In-Vehicle Systems and Safety, S. 73-107. New York: Springer.
- [67] **Silverman, Daniel (2006)**
A critical introduction to phonology: Of sound, mind, and body. London, New York: Continuum.
- [68] **Sinnott, Joan M., William C. Stebbins, David B. Moody (1975)**
Regulation of voice amplitude by the monkey. In: The Journal of the Acoustical Society of America 58(2), S. 412-414.
- [69] **Stach, Brad A. (2003)**
Comprehensive dictionary of audiology: Illustrated. 2. Auflage, New York: Thomson Delmar Learning.
- [70] **Summers, W. Van, David B. Pisoni, Robert H. Bernacki, Robert I. Pedlow, Michael A. Stokes (1988)**
Effects of noise on speech production: Acoustic and perceptual analyses. In: The Journal of the Acoustical Society of America 84(3), S. 917-928.

- [71] **Theiß, Anne, Gerhard Schmidt, Jochen Withopf, Christian Lüke (2014)**
Instrumental evaluation of In-Car Communication systems. In: Proc. of Speech Communication, 11. ITG Symposium, S. 1-4. Erlangen, Deutschland.
- [72] **Thompson, Henry S., Anne Anderson, Ellen Gurman Bard, Gwyneth Doherty-Sneddon, Alison Newlands, Cathy Sotillo (1993)**
The HCRC map task corpus: natural dialogue for speech recognition. In: Proc. of the Workshop on Human Language Technology, S. 25-30. Princeton, USA.
- [73] **Trautmüller, Hartmut (1981)**
Perceptual dimension of openness in vowels. In: The Journal of the Acoustical Society of America 69, S. 1465-1475.
- [74] **Trautmüller, Hartmut, Anders Eriksson (2000)**
Acoustic effects of variation in vocal effort by men, women, and children. In: The Journal of the Acoustical Society of America 107(6), S. 3438-3451.
- [75] **Vatikiotis-Bateson, Eric, Adriano V. Barbosa, Cheuk Yi Chow, Martin Oberg, Johanna Tan, Hani C. Yehia (2007)**
Audiovisual Lombard speech: Reconciling production and perception. In: Proc. of the International Conference on Auditory-Visual Speech Processing (AVSP), S. 45-50. Hilvarenbeek, Niederlande.
- [76] **Warhadpande, Alankar, Christian Lüke, Anne Theiß, Gerhard Schmidt (2015)**
Improvement by adding video feature in an acoustic ambiance simulation for automobiles. In: 7th biennial workshop on DSP for In-Vehicle Systems and Safety, S. 1-6. Berkeley, USA.
- [77] **Wollschläger, Daniel (2014)**
Grundlagen der Datenanalyse mit R: Eine anwendungsorientierte Einführung. 3. Auflage, Berlin, Heidelberg: Springer.
- [78] **Zetkin, Maxim, Herbert Schaldach (1992)**
Wörterbuch der Medizin. 15. Auflage, Berlin: Ullstein Mosby.

- [79] de.mathworks.com/products/matlab.html
(Zugriff: 11.09.2017)
- [80] dss.kirat-online.de/index.php/media-center/data-bases/spid-corpus
(Zugriff: 11.09.2017)
- [81] <http://www.autobild.de/bilder/top-30-die-leisesten-autos-3432065.html#bild22>
(Zugriff: 22.02.2018)
- [82] www.clarin-d.de/de/
(Zugriff: 11.09.2017)
- [83] www.clarin.eu
(Zugriff: 11.09.2017)
- [84] www.dss.tf.uni-kiel.de
(Zugriff: 11.09.2017)
- [85] www.duden.de/sprachwissen/rechtschreibregeln/doppel-s-und-scharfes-s
(Zugriff: 11.09.2017)
- [86] www.fon.hum.uva.nl/praat/
(Zugriff: 11.09.2017)
- [87] www.google.de/maps/
(Zugriff: 19.09.2017)
- [88] www.ipds.uni-kiel.de/pub_exx/bp2001_1/Linda21.html
(Zugriff: 11.09.2017)
- [89] www.bas.uni-muenchen.de/Bas/BasMAUS.html
(Zugriff: 11.09.2017)
- [90] clarin.phonetik.uni-muenchen.de/BASWebServices/#!/services
(Zugriff: 11.09.2017)
- [91] <http://clarin.phonetik.uni-muenchen.de/BASWebServices/services/help>
(Zugriff: 11.09.2017)

LITERATURVERZEICHNIS

- [92] www.phon.ucl.ac.uk/home/sampa/index.html
(Zugriff: 11.09.2017)
- [93] www.python.org
(Zugriff: 11.09.2017)
- [94] www.r-project.org
(Zugriff: 11.09.2017)
- [95] www.soscisurvey.de
(Zugriff: 11.09.2017)
- [96] www.surveycircle.com/de/
(Zugriff: 11.09.2017)

Curriculum Vitae

Rabea Landgraf

Akademischer Werdegang

- 04/2014 - 06/2018** Christian-Albrechts-Universität (CAU) zu Kiel, Deutschland. Promotionsstudium (Dr. Phil.) im Fach: Allgemeine und Vergleichende Sprachwissenschaft.
- 10/2011 - 01/2014** CAU Kiel. Master of Arts (M.A.) in den Fächern: Sprachdokumentation und Korpuslinguistik und Europäische Ethnologie/Volkskunde. Titel der Masterarbeit: „Linguistische Charakteristika von Dialogen im fahrenden Auto und wie sie sich im Labor simulieren lassen“.
- 10/2007 - 10/2011** CAU Kiel. Bachelor of Arts (B.A.) in den Fächern: Empirische Sprachwissenschaft und Europäische Ethnologie/Volkskunde. Titel der Bachelorarbeit: „Reduplikation im Teop - Eine phonetische Perspektive“.
- 08/1997 - 06/2006** Gymnasium (Auguste-Viktoria-Schule) in Itzehoe, Deutschland. Abitur. Leistungskurse: Englisch und Deutsch.

Stipendium

- 12/2014 - 06/2018** Stipendium in der Promotionsförderung der Konrad-Adenauer-Stiftung.

Universitäre Tätigkeiten

- 10/2014 - 03/2015** CAU Kiel. Lehrauftrag für ein Sprachwissenschaftliches Kolloquium.
- 05/2014 - 11/2014** CAU Kiel. Wissenschaftliche Hilfskraft am Institut für Skandinavistik, Frisistik und Allgemeine Sprachwissenschaft (ISFAS). PerLe Fonds für Lehrinnovation.
- 10/2011 - 02/2012** CAU Kiel. Durchführung eines begleitenden Tutoriums zum Proseminar „Einführung in die Europäische Ethnologie/Volkskunde“.
- 04/2010 - 04/2011** CAU Kiel. Wissenschaftliche Hilfskraft am interdisziplinären Forschungszentrum „Stimme & Emotion“.