

Von der Straße ins Labor: Die Modifikation der Sprachproduktion bei lauten Fahrgeräuschen

Rabea Landgraf¹, Oliver Niebuhr², Gerhard Schmidt³, Tina John¹, Christian Lüke³, Anne Theiß³

¹ *Abt. für Allgemeine Sprachwissenschaft, ISFAS, Universität Kiel, E-Mail: landgraf@isfas.uni-kiel.de; tjohn@isfas.uni-kiel.de*

² *Dept. of Design & Communication, IRCA, University of Southern Denmark, E-Mail: olni@sdu.dk*

³ *Digitale Signalverarbeitung und Systemtheorie, Universität Kiel, E-Mail: gus@tf.uni-kiel.de; ath@tf.uni-kiel.de; cl@tf.uni-kiel.de*

Abstract

Lautsprachliche Kommunikation im Auto wird durch die Geräuschbelastung unterschiedlicher Geschwindigkeiten stark beeinflusst. Dabei spielt vor allem der Lombard-Effekt eine entscheidende Rolle, der die Veränderungen der Sprachproduktion bei Lärm beschreibt. Um festzustellen, ob die im Fahrzeug entstehende Lombard-Sprache unter kontrollierten Laborbedingungen nachgebildet und somit im Labor möglichst realitätsnah untersucht werden kann, wurde eine an der Universität Kiel entwickelte akustische Umgebungssimulation verwendet. In dieser Laborsituation sowie während einer echten Fahrsituation wurden bei jeweils gleichen Geräuschbelastungen Sprachaufnahmen durchgeführt. Während frühere Arbeiten zur Lombard-Sprache überwiegend prosodische und segmentelle Aspekte betrachteten, wurden in dieser Arbeit vor allem kommunikative Parameter der Gesprächssteuerung untersucht. Die Ergebnisse zeigen, dass der im Labor und im fahrenden Auto auftretende Lombard-Effekt qualitativ gleich ist.

1. Einleitung

Das Phänomen der Lombard-Sprache wurde zum ersten Mal 1911 vom französischen Ohrenarzt Étienne Lombard (1869-1920) untersucht und beschreibt die Modifikationen der Sprachproduktion bei lauten Hintergrundgeräuschen [8]. Seitdem gab es zahlreiche weitere Studien, die herausfanden, dass der Lombard-Effekt nicht nur die Stimmintensität, sondern viele weitere Parameter der Sprachproduktion beeinflusst. Dabei ergibt sich z. B. eine Steigerung der Grundfrequenz [10,13] und der Wortdauer [2,8,13] sowie eine Veränderung der Sprechgeschwindigkeit [2,10] und der unteren Formantfrequenzen [2,8,13]. Des Weiteren wurde durch Perzeptionstests festgestellt, dass Sprache, die bei Lärm produziert worden ist, für Zuhörer bis zu einem gewissen Grad immer verständlicher wird, je stärker die Geräuschbelastung für den Sprecher wurde [3].

Das Ausmaß des Lombard-Effekts wird von verschiedenen Faktoren beeinflusst, wie z. B. der Art der Sprechaufgabe und der Lärmbedingung in einem Experiment [10]. Dabei ist der Effekt stärker in Spontansprache und interaktiven Sprechaufgaben, als bei gelesener Sprache und nicht-interaktiven Aufgaben [4,7,13]. Dies sollte in dieser Untersuchung zum Lombard-Effekt Berücksichtigung finden.

2. Die akustische Umgebungssimulation

Für die Sprachaufnahmen wurde eine akustische Umgebungssimulation verwendet, die am Institut für Digitale Signalverarbeitung und Systemtheorie (DSS) der Universität Kiel entwickelt worden ist [9]. Diese nutzt Lautsprecher im Inneren des Fahrzeugs, um dort eine realistische Reproduktion der Fahrgeräusche zu erzielen. Dazu wird zunächst für jeden Lautsprecher ein weißes Rauschsignal erzeugt. Aus der Aufnahme einer realen Fahrt werden die Betragsspektren errechnet und den Rauschsignalen zeitvariant aufgeprägt. Eine Kalibrierung stellt sicher, dass die Autoleistungsdichtespektren an den Ohren der Probanden korrekt wiederhergestellt werden. Die derart zeitvariant spektral gefärbten Rauschsignale werden durch die Lautsprecher im Fahrzeuginnenraum wiedergegeben.

In dieser simulierten akustischen Umgebung wird die Sprache der Probanden über Mikrofone aufgezeichnet. Um ein möglichst reines Sprachsignal zu erhalten, kann das simulierte Fahrgeräusch aus den Mikrofonsignalen entfernt werden. Da sich die Probanden während der Aufnahme bewegen, werden dazu adaptive Filter eingesetzt, welche fortwährend die Übertragungswege von den Lautsprechern zu den Mikrofonen schätzen und mit den Lautsprecher-signalen als Referenz das Fahrgeräusch weitgehend kompensieren können. Die Rauschgeneratoren als Anregungssignale stellen dabei sicher, dass die Lautsprecher-signalen nicht gekreuzkorreliert sind, was den Einsatz adaptiver Filter ermöglicht.

3. Erhebung der Sprachdaten

Die Sprachsignale wurden in der akustischen Umgebungssimulation und während einer echten Fahrsituation aufgenommen, bei jeweils gleichen Geräuschbelastungen von 0 km/h (Motor an), 50 km/h, 100 km/h und 150 km/h, sowie in einer stillen Referenzsituation. Drei Sprecherpaare jeweils gleichen Geschlechts nahmen an dem Experiment teil, die sich während der Aufnahmen in jeder der 9 Bedingungen im Dialog unterhielten. Während der Aufnahmen saß einer der Probanden auf dem Beifahrersitz und der andere dahinter. Beide sprachen jeweils in ein eigenes Mikrofon, das am Vordersitz befestigt war. Abbildung 1 und 2 zeigen beispielhaft die Position des Mikrofons für den Beifahrersitz und die Anbringung der Lautsprecher auf der rechten Fahrzeugseite während der Sprachaufnahmen. Nach den Aufnahmen füllte jeder Dialogpartner einen Fragebogen zu seinen subjektiven Eindrücken der Experimentsituation aus.



Abbildung 1: Die Position des Mikrofons am Beifahrersitz während der Sprachaufnahmen.



Abbildung 2: Die Positionen der Lautsprecher an den Fenstern der rechten Fahrzeugseite während der Sprachaufnahmen.

Zur Elizitation von Spontansprache und damit zur Steigerung des Lombard-Effekts wurde die interaktive Sprechaufgabe des *Video Task* gewählt [6], die bereits am Institut für Phonetik und Digitale Sprachverarbeitung der Universität Kiel als sogenanntes ‘*Daily Soap Szenario*’ angewendet wurde [11,12]. Dabei werden verschiedene Sequenzen einer bekannten Fernsehserie derart zusammengeschnitten, dass zwei ähnliche, aber nicht identische Videos entstehen. Diese unterscheiden sich in der Auswahl, der Reihenfolge und der Vollständigkeit der Szenen und sind etwa 15 Minuten lang. Im Experiment der vorliegenden Studie sah sich jeder Dialogpartner eines der Videos zweimal an. Während der darauffolgenden Sprachaufnahmen sollten die Probanden das Gesehene vergleichen und dabei Ähnlichkeiten und Unterschiede diskutieren. Um möglichst ungezwungene spontansprachliche Dialoge zu erhalten, wurden nur Dialogpartner gewählt, die sich bereits gut kannten.

4. Akustische und auditive Analysen

Die Untersuchung der Sprachsignale wurde mittels Praat [1] durchgeführt. Für die anschließende Analyse hätten nur die in der akustischen Umgebungssimulation aufgenommenen Sprachsignale von den eingespielten Hintergrundgeräuschen befreit werden können. Da die Aufnahmen der echten Fahrsituation vom Hintergrundlärm beeinflusst sind, können akustische Messungen wie die Bestimmung von spektralen Eigenschaften und Formanten verfälscht sein und leicht zu Fehlinterpretationen führen [5]. Aus diesem Grund wurden die Intensität, Grundfrequenz und weitere spektrale Eigenschaften nicht analysiert, sondern aufgrund der auditiven

Eindrücke von Experimentleiterin und Probanden beschrieben.

Die phonetischen Analysen der Sprachsignale wurden auf zwei verschiedenen Ebenen durchgeführt. Auf der einen Seite wurden Parameter untersucht, die sich auf einen gesamten Dialog beziehen (*Ges.*). Auf der anderen Seite wurden Parameter analysiert, die sich auf Veränderungen innerhalb eines Dialogs beziehen, d. h. Eigenschaften von einzelnen Redebeiträgen am Anfang, in der Mitte und am Ende eines Dialogs (*Turn*). Letztere wurden untersucht, um festzustellen, ob sich die Probanden während eines Dialogs an die laute Umgebung gewöhnen. Die Parameter der Untersuchung sind im Folgenden aufgelistet:

- Anzahl der Turns. (Turns.Ges.)
- Anzahl der Häsitationen. (Häs.Ges.)
- Anteil der produzierten Sprache im Gesamtgespräch, d. h. Normierung der Sprachdauer auf die Gesamtdauer des Dialogs. (Spr.Ges.)
- Sprechgeschwindigkeit, d. h. Division der Silbenanzahl durch die Dauer der produzierten Sprache ohne Pausen in Sekunden. (Silb.Turn)
- Anteil der produzierten Sprache im Turn, d. h. Normierung der Sprachdauer im Turn auf die Gesamtdauer des Turns. (Spr.Turn)

5. Ergebnisse

5.1. Inferenzstatistik

Zur statistischen Analyse der Daten wurden fünf zweifaktorielle Varianzanalysen mit Messwiederholung durchgeführt, eine für jeden der fünf Untersuchungsparameter. Sie basieren auf den zwei festen Innersubjektfaktoren Ort (Labor und Fahrt) und Tempo (0 km/h, 50 km/h, 100 km/h und 150 km/h). Zur Berechnung der fünf ANOVAs wurde die statistische Software SPSS verwendet. Die Ergebnisse sind in Tabelle 1 und 2 zusammengefasst.

Tabelle 1: Deskriptive und inferenzstatistische Auswertung der fünf ANOVAs für den Faktor Ort. Statistisch unbedeutende Vergleiche sind durch “-“ markiert.

Abhängige Variablen	Ort der Untersuchung					
	Mittelwerte		Ergebnisse der ANOVA			
	Labor	Fahrt	df1/df2	F	p	η_p^2
Turns.Ges.	9,792	8,375	1/5	3,658	0,016	0,423
Häs.Ges.	10,375	8,792	-	-	-	-
Spr.Ges.	0,353	0,425	1/5	7,254	0,002	0,592
Silb.Turn	5,519	5,741	1/5	4,493	0,009	0,473
Spr.Turn	0,798	0,826	1/5	7,83	0,002	0,61

Tabelle 2: Deskriptive und inferenzstatistische Auswertung der fünf ANOVAs für den Faktor Tempo. Statistisch unbedeutende Vergleiche sind durch “-“ markiert. Ein signifikanter Trend ist in Klammern gesetzt.

Abhängige Variablen	Lärmbedingung			
	Mittelwerte			
	0 km/h	50 km/h	100 km/h	150 km/h
Turns.Ges.	11,5	9,833	8,083	6,917
Häs.Ges.	8,833	10,583	9,583	9,333
Spr.Ges.	0,385	0,39	0,382	0,4
Silb.Turn	5,767	5,707	5,422	5,623
Spr.Turn	0,796	0,825	0,791	0,836
	Ergebnisse der ANOVA			
	df1/df2	F	p	η_p^2
Turns.Ges.	3/15	3,144	0,012	0,386
Häs.Ges.	-	-	-	-
Spr.Ges.	-	-	-	-
Silb.Turn	3/15	2,401	0,023	0,324
Spr.Turn	3/15	1,711	(0,062)	0,255

Aus Tabelle 1 wird deutlich, dass es einen signifikanten Haupteffekt des Faktors Ort für jeden der fünf Parameter gibt, außer für Häs.Ges. In der Laborbedingung gibt es mehr Turns, die Sprechgeschwindigkeit ist geringer, und weniger Sprache wird produziert, sowohl im ganzen Dialog als auch in den einzelnen untersuchten Turns. Weiterhin kann den Mittelwerten aus Häs.Ges. entnommen werden, dass im Labor etwas mehr Häsitationen produziert werden.

Des Weiteren zeigen die Ergebnisse der Varianzanalysen, dass die Anzahl der Turns bei steigender Geschwindigkeit abnimmt, d. h. die Interaktion beider Sprecher geht zurück, und die Redebeiträge werden länger. Tabelle 2 macht zudem deutlich, dass die Sprechgeschwindigkeit abnimmt, insbesondere bei 100 km/h. Für die anderen Parameter gab es keinen signifikanten Effekt. Dennoch wird aus den Mittelwerten deutlich, dass die Anzahl der Häsitationen bei zunehmender Lärmbedingung leicht ansteigt, besonders von 0 km/h zu 50 km/h. Zusätzlich wird, außer bei 100 km/h, mehr Sprache im ganzen Dialog und in den jeweiligen Turns produziert. Da der Anstieg der Häsitationen eher gering ist, wird davon ausgegangen, dass dies nicht der Grund für die vermehrte Sprachproduktion ist. Abbildung 3 und 4 verdeutlichen beispielhaft, dass bei steigendem Umgebungslärm in beiden Versuchsorten dieselben Effekte auftreten, auch wenn die Tendenzen in Labor und Fahrt unterschiedlich stark ausgeprägt sind.

Sprachproduktion im Turn

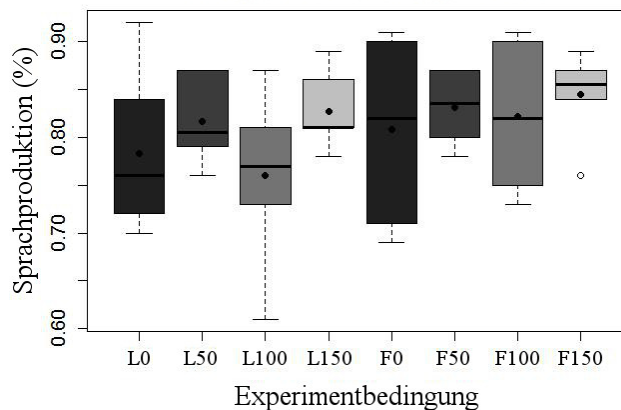


Abbildung 3: Der Anteil der produzierten Sprache im Turn. Die Versuchsorte L (Labor) und F (Fahrt) sind mit den Lärmkonditionen 0 km/h, 50 km/h, 100 km/h und 150 km/h kombiniert.

Zudem gibt es eine signifikante Interaktion zwischen den beiden Faktoren Ort und Tempo für den Parameter Turns.Ges. ($F[3,15]=3,441$; $p=0,015$; $\eta_p^2=0,408$). Die Mittelwerte der unterschiedlichen Lärmbedingungen nehmen im Labor stärker ab ($\bar{x}_{L0}=13,167$; $\bar{x}_{L50}=12,000$; $\bar{x}_{L100}=7,500$; $\bar{x}_{L150}=6,500$) als im fahrenden Auto ($\bar{x}_{F0}=9,833$; $\bar{x}_{F50}=7,667$; $\bar{x}_{F100}=8,667$; $\bar{x}_{F150}=7,333$). Da jedoch die Effektrichtung in beiden Versuchsorten dieselbe ist, handelt es sich hierbei nicht um eine qualitative, sondern um eine quantitative Interaktion.

Anzahl der Turns im Dialog

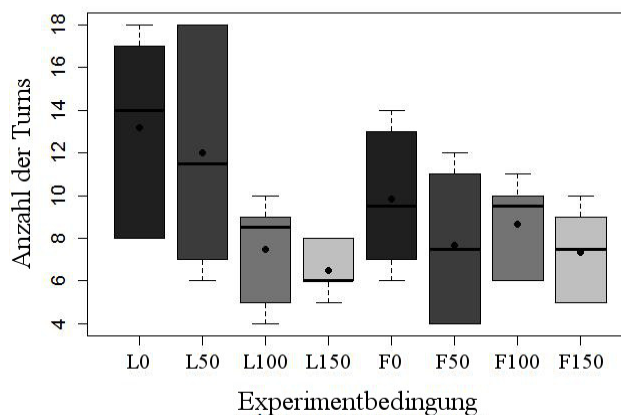


Abbildung 4: Die Anzahl der Turns in einem Dialog. Die Versuchsorte L (Labor) und F (Fahrt) sind mit den Lärmkonditionen 0 km/h, 50 km/h, 100 km/h und 150 km/h kombiniert.

5.2. Deskriptive Analysen

Für die deskriptive Untersuchung der Sprachsignale wurden zusätzliche Detailvergleiche der Messwertkurven zwischen Labor und Fahrt durchgeführt, sowie zwischen den unterschiedlichen Lärmbedingungen beider Versuchsorte.

Für keinen der fünf Untersuchungsparameter wurden systematische Unterschiede oder Gemeinsamkeiten in den Kurvenverläufen der Messwerte zwischen den unterschiedlichen Experimentbedingungen festgestellt. Selten auf-

trete Ähnlichkeiten zwischen Messwertkurven sind unbeständig und werden daher als zufällig betrachtet. Deshalb ist davon auszugehen, dass keine Anpassung der Sprecher an die laute Umgebung über die drei Turns am Anfang, in der Mitte und am Ende eines jeden Dialogs stattfindet. Des Weiteren konnten keine geschlechtsspezifischen Variationen in den Daten nachgewiesen werden.

Gemäß den auditiven Eindrücken von Sprechern und Experimentleiterin stieg die Lautheit der Dialoge in beiden Versuchsorten mit steigender Lärmbelastung an. Zudem wurde die Grundfrequenz als höher empfunden. In der Stärke der Anstiege wurden keine Unterschiede von Lautheit und F0 zwischen den Labor- und den Fahrtbedingungen festgestellt.

Allerdings fühlten sich einige Sprecher während der Umgebungssimulation im Labor leicht unwohl. Dies liegt vermutlich an der ungewöhnlichen Experimentsituation und den zahlreichen sichtbaren Gerätschaften und Kabeln im Auto. Zusätzlich ist das Innere des Wagens aufgrund der großen Lautsprecher an den Scheiben stark verdunkelt.

6. Fazit

Während einige der untersuchten Sprachmodifikationen wie die Veränderung der Sprechgeschwindigkeit bereits aus der Literatur bekannt sind, wurden zudem neue Lombard-Parameter aufgezeigt, wie z. B. die abnehmende Anzahl von Redebeiträgen pro Sprecher bei gleichzeitig steigender Sprachproduktion und sinkender Sprechgeschwindigkeit. Basierend auf den statistischen und deskriptiven Ergebnissen konnten über die Sprecherstichprobe hinweg Auswirkungen von sowohl Geschwindigkeits- als auch Umgebungsbedingung auf das Kommunikationsverhalten nachgewiesen werden. Es gibt Unterschiede zwischen Labor und Fahrt, da der Versuchsort Einfluss auf die Messwerte der fünf untersuchten Parameter hat. Es treten jedoch in beiden Versuchsorten ähnliche Modifikationen der Sprachproduktion und des Kommunikationsverhaltens auf. Quantitative Unterschiede im Ausmaß des Lombard-Effekts zwischen Labor und Fahrt wurden vermutlich durch die ungewohnte Situation im Labor hervorgerufen.

Die Ergebnisse der vorliegenden Studie bestätigen, dass der Lombard-Effekt, welcher bei unterschiedlichen Geschwindigkeiten in einem fahrenden Auto auftritt, im Labor simuliert werden kann. Zusammenfassend stellt die hier vorgestellte akustische Umgebungssimulation einen innovativen und bedeutenden Schritt hin zu einer besseren Untersuchung der Kommunikation im fahrenden Auto dar, sowie des darin auftretenden Lombard-Effekts. Im Anschluss an diese Untersuchung wurde die Wiedergabe von Geräuschen in der akustischen Umgebungssimulation, z. T. angeregt durch die subjektiven Einschätzungen der Probanden, zusätzlich überarbeitet und verbessert. Zudem ist eine visuelle Simulation der Fahrsituation in Form einer Leinwand hinzugefügt worden, die sich direkt vor der Motorhaube des Autos befindet. Die auf diese Weise gezeigten Videosequenzen stimmen mit den zeitgleich wiedergegebenen Fahrgeräuschen überein. Nachfolgende Studien zur lautsprachlichen Kommunikation können somit diese Umgebungssimulation unter verbesserten Bedingungen nutzen.

7. Literatur

- [1] Boersma, P.: Praat, a system for doing phonetics by computer. *Glott International* 5 (2001), 341-345 <http://www.fon.hum.uva.nl/praat/>
- [2] Davis, C., J. Kim, K. Grauwinkel, H. Mixdorff: Lombard speech: auditory (A), visual (V) and AV effects. *Proc. of the 3rd International Conference of Speech Prosody* (2006), 361-365. Dresden, Deutschland
- [3] Dreher, J. J., J. O'Neill: Effects of ambient noise on speaker intelligibility for words and phrases. *Journal of the Acoustical Society of America* 29 (1957), 1320-1323
- [4] Garnier, M., M. Dohen, H. Lœvenbruck, P. Welby, L. Bailly: The Lombard Effect: a physiological reflex or a controlled intelligibility enhancement? *Proc. of 7th International Seminar on Speech Production* (2006), 255-262. Ubatuba, Brasilien
- [5] John, T., O. Niebuhr, G. Schmidt, A. Theiß: Phonetic analysis vs. dirty signals: Fixing the paradox. *Proc. der 24. Konferenz zur Elektronischen Sprachsignalverarbeitung (ESSV)* (2013), 1-8. Bielefeld, Deutschland
- [6] John, T., R. Landgraf, C. Lüke, S. Rohde, G. Schmidt, A. Theiß, J. Withopf: Über die Verbesserung der Sprachkommunikation in geräuschbehafteten Umgebungen. In: Oliver Niebuhr (Hrsg.): *Formen des Nicht-Verstehens*, 185-203. Peter Lang, Frankfurt, 2014
- [7] Lane, H., B. Tranel: The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research* 14 (1971), 677-709
- [8] Lau, P.: The Lombard Effect as a communicative phenomenon. *UC Berkeley Phonology Lab Annual Report* 39 (2008), 1-9
- [9] Lüke, C., A. Theiß, G. Schmidt, O. Niebuhr, T. John: Creation of a Lombard Speech Database using an Acoustic Ambiance Simulation with Loudspeakers. *6th Biennial DSP Workshop for In-Vehicle Systems* (2013). Seoul, Korea
- [10] Mixdorff, H., U. Pech, C. Davis, J. Kim: Map Task dialogs in noise – a paradigm for examining Lombard speech. *Proc. 16th International Congress of Phonetic Science* (2007), 1329-1332. Saarbrücken, Deutschland
- [11] Niebuhr, O., B. Peters, R. Landgraf, G. Schmidt: The Kiel Corpora of "Speech & Emotion" - A Summary. 41. Jahrestagung für Akustik, DAGA (2015)
- [12] Peters, B.: 'Video Task' oder 'Daily Soap Szenario' - Ein neues Verfahren zur kontrollierten Elizitation von Spontansprache (2001). http://www.ipds.uni-kiel.de/pub_exx/bp2001_1/Linda21.html
- [13] Summers, W V., D. B. Pisoni, R. H. Bernacki, R. I. Pedlow, M. A. Stokes: Effects of noise on speech production: Acoustic and perceptual analyses. *Journal of the Acoustical Society of America* 84 (1988), 917-928