More Noise, Less Talk – The Impact of Driving Noise and In-Car Communication Systems on Acoustic-Prosodic Parameters in Dialogue

Rabea Landgraf[1], Gerhard Schmidt[2], Johannes Köhler-Kaeß, Oliver Niebuhr[3], Tina John[1]

[1] *General Linguistics, ISFAS, Kiel University, Germany, Email: landgraf@isfas.uni-kiel.de; tjohn@isfas.uni-kiel.de*

[2] *Digital Signal Processing and System Theory (DSS), Kiel University, Germany, Email: gus@tf.uni-kiel.de*

[3] *SDU Technology Entrepreneurship and Innovation (TEI), IRCA, University of Southern Denmark, Email: olni@sdu.dk*

Introduction

First described in 1911, the Lombard effect became an umbrella term for the changes in speech production and speech behaviour that take place in loud environments. Over the last hundred years, research revealed that a large number of communicative parameters are modified in Lombard speech. Examples are the increase of intensity, fundamental frequency, and word duration, a modification of the formant frequencies, and a reduction of the voice onset time [6] [11]. The duration of not or less sonorous consonants like plosives is reduced in noise, while vowels and more sonorous consonants like nasals are lengthened [1]. In addition, the amount of speech produced in a dialogue increases [4] and Mixdorff (2007) found that subjects doing map tasks are solving it faster when talking at loud background noise [8].

The noisy interior of a driving car is an example of a Lombard-inducing environment. That is, loud noises cause stress [3] and make it difficult to maintain an effective and intelligible communication. Furthermore, the inside of a car is characterised by an unfavourable acoustic energy radiation, because subjects in the back seat talk towards their interlocutors in the front seat, who, in turn, talk towards the windshield. In-Car Communication systems (ICC) are intended to improve this situation by transmitting the speech signals with microphones and loudspeakers from passenger to passenger. The present study aims at providing a linguistic evaluation of such an ICC system. For this purpose and because the Lombard effect that emerges in a driving car has hardly been studied to date, the following questions have to be answered: [5]

- How does speech communication in a car change under Lombard conditions (without ICC)?
- How does the Lombard effect in a car change by using an ICC system?

The SPID speech corpus

To answer the above research questions, a speech corpus SPID (SPontaneous In-car Dialogues) has been developed [13] and analysed along the acoustic-phonetic and communicative parameters that are known to be modified by the Lombard effect. The corpus and its analyses will be described in the next sections.

ICC at Kiel University

Engineers at Kiel University developed an ICC system with microphones placed inside the seat belts and loudspeakers in the ceiling of the car. Because of the unfavourable acoustic energy radiation in the vehicle mentioned before, it transmits signals from the front louder to the back than in the opposite direction. This system was used to conduct the present investigation. [9]

Map Tasks

For the elicitation of spontaneous speech, the Map Task paradigm was used [10]. That is, subjects of a dialogue pair took the roles of instruction giver and instruction follower. The maps they received were almost identical, but partially differed in the names of persons, streets, places, house numbers and in their compositions. During the recordings, the giver had to describe a route drawn on his/her map to the follower, who had to retrace it on his/her own map. The map design was chosen to support the elicitation of spontaneous speech by controversial discussions between the dialogue partners.

Ambience simulation

Recordings were made in an acoustic and visual ambiance simulation developed at Kiel University. In this way, speech signals can be obtained and analysed under controlled and reproducable conditions. Inside a stationary vehicle, loudspeakers are placed on 2 seats and in 6 windows. Driving noises that are played during speech recordings can be removed from the speech signals afterwards to ensure that these noises do not interfere with acoustic measurements. This technique makes it unnecessary to elicit Lombard speech with the use of headphones. The visual ambiance simulation is achieved by a screen that shows different driving situations, which are conform with the acoustic signals. [2] [7] [12]

Figure 1: The ambiance simulation at Kiel University. Left: The acoustic simulation. Right: The visual simulation.

Corpus Conditions

For each recording condition of SPID, another map task was constructed and solved by the dialogue partners so that no task had to be used twice. 4 male and 4 female dialogue pairs participated in the recordings, all of them native standard German speakers. Participants were prevented from making eye contact, because it often cannot be maintained in a car. That is, one subject was placed in the front passenger seat, and the other one behind the dialogue partner. During the whole experiment, seating positions and speech roles were maintained by the participants. In a first experiment session, speech recordings took place without ICC in the above described ambiance simulation at driving speeds and noises of 50 km/h, 130 km/h, as well as in a silent reference condition. The driving noises were randomized for each dialogue pair to avoid learning effects in the analyses. Afterwards, recordings were made in the driving car at 50 km/h and 130 km/h, where the driving conditions were also randomized. About 3 to 4 months later, a second recording condition was conducted, which was almost identical to the first session, except for the use of an ICC system. This was to prevent a learning effect concerning speech task and experimental ambiance. For the present investigation, only the recordings made inside the ambiance simulation were considered. [5]

The whole corpus of SPID contains about 15 hours 47 minutes of speech recordings in the laboratory and 9 hours 55 minutes of uncutted material in the real driving conditions. The uncutted recordings also contain breaks in the speech tasks during stillstand at traffic lights or induced by speed limitations.

Analyses

A linear least squares regression model (LRM) was used for statistically analysing the speech signals. Using the software environment R [15], the influence of the factors *ICC* (without, with), *noise* (silence, 50 km/h, 130 km/h), *seat* (front, rear), and *role* (giver, follower) was examined on the independent variables mean pitch, mean intensity, signal duration, and amount of speech in a dialogue.

After collecting the data from praat scripts [14], a within-subject normalisation was conducted for the measurements of the independent variables following equation 1, where s denote each individual subject and i labels every single recording condition. Thereby, the holistic behaviour of all subjects can be compared. Subsequently, the LRM was conducted for every independent variable and a linear transformation was applied to the estimated values of the LRM, by which the value of the condition *without ICC, silence (rear) (giver)*, respectively, was subtracted from the values of all other experimental conditions. In this way, all combinations of factor levels can be easily compared to each other. [5]

$$x'_{s,i} = x_{s,i} - \bar{x}_s. \quad (1)$$

Results

Intensity and fundamental frequency

Figure 2 shows the estimated intensity values for every experimental condition calculated by the LRM. A similar picture emerges for both the mean intensity and the mean pitch of the speech signals. Therefore, only the mean intensity is depicted in figure 2. There is a Lombard effect without and with the ICC system, in accord with previous findings, because measurements are increasing with higher noise levels ($p=0.0055$ for intensity; $p<0.0001$ for pitch). The ICC system also has a highly significant influence ($p<0.0001$ for both parameters). With ICC, the measurements are reduced in all factor combinations excepting *silence, front*. There is also an interaction between the two factors *ICC* and *noise*, which is only marginally significant for mean intensity ($p=0.0513$) and significant for mean pitch ($p=0.0009$). That is, the Lombard effect still occurs with ICC, but its magnitude is reduced. [5]

Furthermore, the factor *seat* strongly influences the loudness ($p=0.0002$) and marginally influences the pitch level ($p=0.0836$) of the speech signals. That is, without ICC subjects in the back seat speak louder than those in the front seat. Additionally, there is a significant interaction with *ICC*, again more pronounced for mean intensity ($p<0.0001$) than for mean pitch ($p=0.0152$). With ICC switched on, subjects in the front seat speak louder than their dialogue partners in the back seat. *Role* shows no significant influence on the measurements. As can be seen in figure 2, in the silence condition with ICC the mean intensity and mean pitch strongly decrease in the back seat, whereas they slightly increase for front passengers. [5]

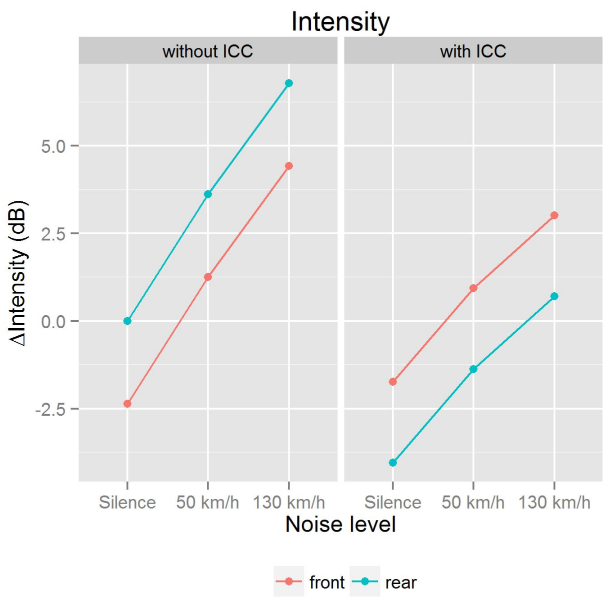

Figure 2: Estimated and transformed values of the LRM for Δ mean intensity.

Signal duration

Regarding the total duration of the dialogues in the laboratory conditions of SPID, *ICC* has only little influence on this parameter (p=0.0388). Thus, subjects are doing their speech task much faster with activated ICC system, as can be seen in figure 3. *Noise* has no impact on the measurements. However, the combination of both factors shows a significant influence on dialogue duration (p=0.0198). Without ICC, the dialogue partners become faster in solving their map tasks at higher noise levels. This Lombard effect is consistent again with findings of former studies. When using ICC, no systematic modification of this parameter shows up for different driving noises. In addition, the factors *seat* and *role* have no significant influence on the measurements.

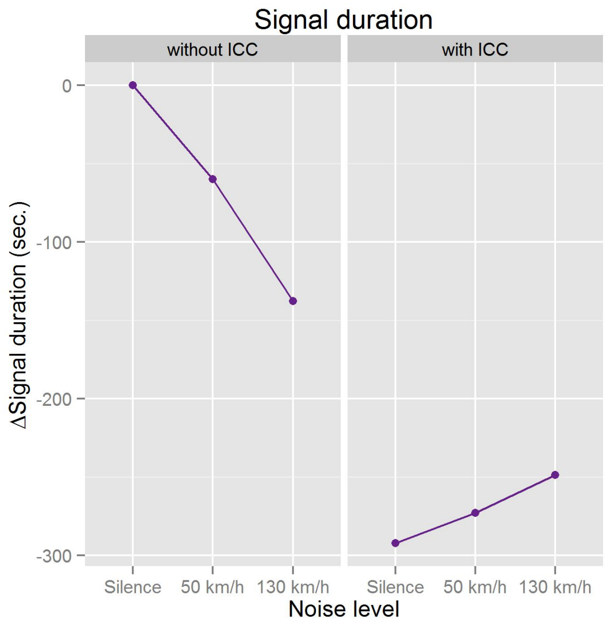

Figure 3: Estimated and transformed values of the LRM for Δ signal duration.

Amount of speech

This parameter denotes the percentage of spoken language within the recorded speech signal of a single subject. For this purpose, only those parts of the speech signals were analysed, which have been orthographically annotated. As can be seen from figures 4 and 5, *noise* has a significant influence on this parameter (p=0.0010). The Lombard effect emerges in the form of a change in the amount of speech as driving noises increase. The factor *ICC* also has a significant effect (p=0.0022). But with activated ICC system, only a shift of the level of the Lombard effect can be noticed. There are no systematic effects of the ICC system on the noise induced changes.

The seating position has a significant effect (p=0.0002). And, regarding the above mentioned shift of measurements, there is a highly significant interaction between the factors *ICC* and *seat* with p<0.0001. That is, with ICC, the amount of speech in the back seat decreases,

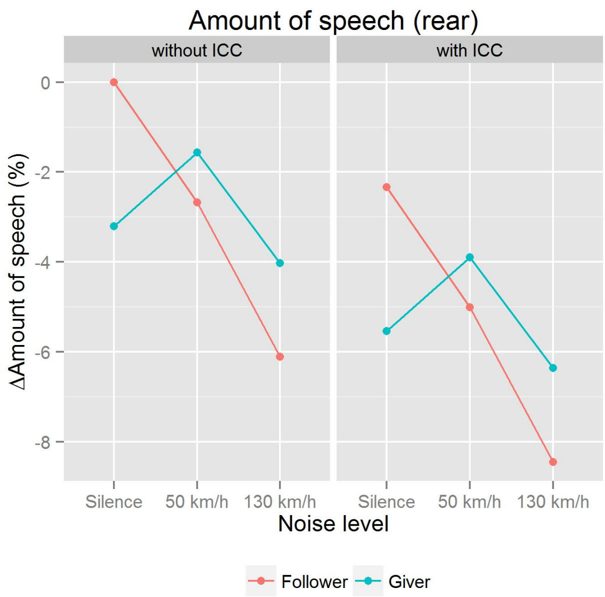

Figure 4: Estimated and transformed values of the LRM for Δ amount of speech in the back seat.

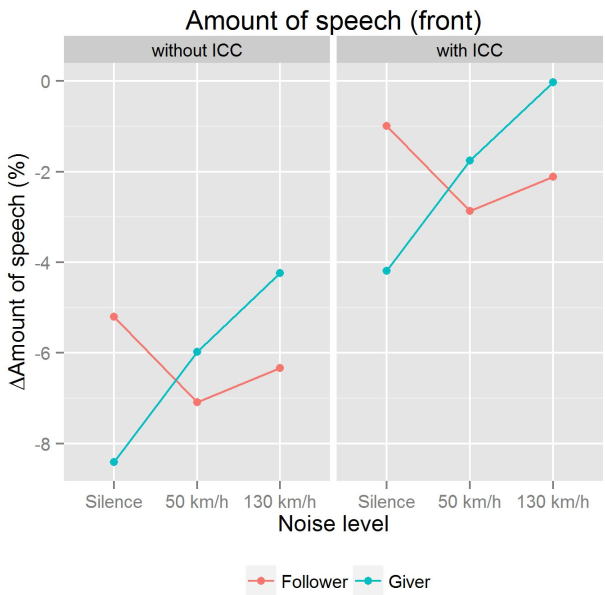

Figure 5: Estimated and transformed values of the LRM for Δ amount of speech in the front seat.

while it increases in the front seat, both without any considerable changes of the Lombard effect described earlier. Furthermore, there is a significant interaction between *noise* and *seat* (p=0.0119). *Noise* has an increasing effect on the amount of speech in the front seat, while the opposite applied to speakers in the back seat.

In addition, the factor *role* has no significant influence on the measurements, but the interaction between *noise* and *role* does with p=0.0078. Giver and follower adapt differently to the increasing noise level. For instance, without

ICC, followers in the back seat (figure 4, red line) are decreasing their amount of speech at higher noise levels, whereas their dialogue partners in the front seat, who had the role of the giver (figure 5, blue line), are increasing their amount of speech at higher driving noises.

Discussion

It has been shown that with the increasing driving noise levels, the measurements of speech production and behaviour are modified. Some of these modifications are already described for Lombard-inducing environments in the literature. With the use of ICC, the Lombard effect is reduced for mean intensity and mean pitch. No effect was found for dialogue duration. Regarding the amount of speech in a dialogue, speech production was shifted to the front seat position.

Regarding loudness and pitch, the measurements in the ICC condition decrease more strongly in the back seat. This could be explained by the unequal supporting effect of the ICC system in the two seating positions, as mentioned above. In addition, the decrease of both parameters in the back seat in silence with ICC could be caused by the unfavourable acoustic energy radiation that already exists in a car without any noise. It can also be compensated by ICC. [5]

Furthermore, the strong decrease of dialogue duration with ICC is a striking new finding. Reasons could be the use of ICC, a learning effect that occurs despite the break between two recording sessions, or a placebo effect, because subjects knew when the ICC system was used. The amount of speech in a speech signal shifts from the back seat without ICC to the front seat with ICC. An explanation could be again the adjustment of ICC in the experiment, which means that with an ICC system, subjects in the back seat need less speaking time, because the communication situation has been improved by ICC.

Conclusion

The results show that ICC can be regarded as a new strategy for compensating the Lombard effect. So, if noise causes stress, Lombard speech can be seen as an expression of this stress. Its reduction by ICC then means a reduction of stress and an improvement of the speech communication in a car. This is undoubtly an advantage of the situation in any vehicle, because it not only enhances speech communication between the passengers, but also improves driving safety when the driver participtes in a dialogue.

References

[1] Fónagy, I., Fónagy, J.: Sound pressure level and duration. Phonetica: international journal of phonetic science 15 (1966), 14-21

[2] John, T., Niebuhr, O., Schmidt, G., and Theiß, A.: Phonetic analysis vs. dirty signals: Fixing the paradox. Proc. der 24. Konferenz zur Elektronischen Sprachsignalverarbeitung (ESSV) (2013), 1–8

[3] Junqua, J.-C.: The influence of acoustics on speech production: A noise-induced stress phenomenon known as the Lombard reflex. Speech Communication 20 (1996), 13-22

[4] Landgraf, R.: Simulating complex speech-production environments. Niebuhr, O., and Skarnitzl, R. (eds): Tackling the Complexity of Speech. Epocha, Prague (2015), 97–110

[5] Landgraf, R., Köhler-Kaeß J., Lüke, C., Niebuhr, O., Schmidt, G.: Can you hear me now? Reducing the Lombard effect in a driving car using an In-Car Communication system. Proceedings of the Speech Prosody (2016), 479-483

[6] Lau, P.: The Lombard Effect as a communicative phenomenon. UC Berkeley Phonology Lab Annual Report 39 (2008), 1–9

[7] Lüke, C., A. Theiß, G. Schmidt, O. Niebuhr, and T. John: Creation of a Lombard speech database using an acoustic ambiance simulation with loudspeakers. Proc. of the 6th Biennial Workshop on Digital Signal Processing for In-Vehicle Systems (2013), 1–8

[8] Mixdorff, H., Pech, U., Davis, C., Kim, J.: Map Task dialogs in noise – a paradigm for examining Lombard speech. Proc. 16th International Congress of Phonetic Science (2007), 1329-1332

[9] Theiß, A., Schmidt, G., Withopf, J., Lüke, C.: Instrumental Evaluation of In-Car Communication Systems. Proc. ITG Fachtagung Sprachkommunikation (2014), 1–4

[10] Thompson, H. S., Anderson, A., Bard, E. G., Doherty-Sneddon, G., Newlands, A., and Sotillo, C. F.: The HCRC Map Task Corpus: Natural dialogue for speech recognition. Proc. Human Language Technology Workshop (1993), 25–30

[11] Van Summers, W., Pisoni, D. B., Bernacki, R. H., Pedlow, R. I., Stokes, M. A.: Effects of noise on speech production: Acoustic and perceptual analyses. Journal of the Acoustical Society of America 84 (1988), 917–928

[12] Warhadpande, A., C. Lüke, A. Theiß, G. Schmidt: Improvement by Adding Video Feature in an Acoustic Ambiance Simulation for Automobiles. 7th Biennial Workshop on DSP for In-Vehicle Systems and Safety (2015), 1-6

[13] Speech Corpus SPID, URL: `http://dss.kirat-online.de/index.php/media-center/data-bases/spid-corpus/`

[14] Praat, URL: `http://www.fon.hum.uva.nl/praat/`

[15] Free software R, URL: `https://www.r-project.org/`