German Alcohol Language Corpus - the Question of Dialect

Florian Schiel, Thomas Kisler

Bavarian Archive for Speech Signals (BAS), Ludwig-Maximilians-Universität München,
Schellingstr. 3, 80799 München, Germany
schiel|kisler@phonetik.uni-muenchen.de

Abstract

Speech uttered under the influence of alcohol is known to deviate from the speech of the same person when sober. This is an important feature in forensic investigations and could also be used to detect intoxication in the automotive environment. Aside from acoustic-phonetic features and speech content which have already been studied by others in this contribution we address the question whether speakers use dialectal variation or dialect words more frequently when intoxicated than when sober. We analyzed 300,000 recorded word tokens in read and spontaneous speech uttered by 162 female and male speakers within the German Alcohol Language Corpus. We found that contrary to our expectations the frequency of dialectal forms decreases significantly when speakers are under the influence. We explain this effect with a compensatory over-shoot mechanism: speakers are aware of their intoxication and that they are being monitored. In forensic analysis of speech this "awareness factor" must be taken into account.

Keywords: alcohol, dialect, forensic

1. Introduction

Alcohol consumption has various effects on the drinker, for example impaired balance, coordination problems, and slow reaction time (e.g. (Chin and Pisoni, 1997)). Another well-known effect is a change in speech production, hence speech might act as a modality through which intoxication could be detected. In contrast to other test methods for intoxication, speech has the advantage that it can be observed without any obtrusive interaction with the potentially intoxicated person. Since voice controlled applications already exist in the automotive environment, the possibility of automatic detection of alcoholic intoxication based on speech is of interest. If the linguistic parameters of spoken language that change under the influence of alcohol were known, it might be possible to automatically detect intoxication in a built-in vehicle computer to prevent driving under the influence.

A number of previous studies have examined the effect of alcohol on the acoustic properties of the speech signal, predominantly fundamental frequency (f0). But findings for f0 are inconsistent: they vary from a significant increase (e.g. (Klingholz et al, 1988), (Hollien et al, 2001)) to a decrease (e.g. (Watanabe et al, 1994), (Aldermann, 1995)) to a change dependent on the breath alcohol concentration (BRAC) (Künzel et al, 2003) or even to no change at all (e.g. (Sobell and Coleman, 1982), (Pisoni et al, 1985), (Cooney et al, 1998)). Similar results can be observed for other acoustic features or feature sets; for a good overview see (Chin and Pisoni, 1997). Some of these contradictory findings might be due to varying experimental setups and low participant numbers: the number of speakers ranges from 4 to 35 in earlier studies, and most studies were conducted with male participants only. In another study (Baumeister et al, 2012) we conducted an f0 analysis of intoxicated speech based on the German Alcohol Language Corpus (ALC) (Schiel et al, 2012) which was designed to provide a publicly available, large and statistically sound corpus for speech recorded in an automotive environment. The majority (79.1%) of speakers within the ALC corpus increase their median f0 when intoxicated.[1]

Studies regarding non-acoustical effects caused by alcoholic intoxication are few: (Smith et al, 1975) investigated number of words, dialog acts, initiations, speech overlaps and interruptions in dyadic communication (female-male). In summary the authors state that "alcohol appeared to make social communication more disorganized, and intoxicated subjects seemed less likely to follow conventional rules of etiquette in their speech" ((Smith et al, 1975), p. 1397). (Künzel et al, 1992) analyzed linguistic errors in a standard reading task (male subjects only, but with varying alcohol levels) among other features. They found a significant increase in word pronunciation errors for higher BAC levels ($> 0.16\%$, measured in BRAC), and subjects were less inclined to correct their own errors. (Barfüsser, 2010) analyzed a subset of the ALC (the same data set as in the present study) for hesitations, pauses, pronunciation errors, corrections and repetitions. She reported a significant increase in the number and duration of hesitations, number of pronunciation errors and repetitions.

The aim of the present study is to test whether the usage of dialectal variants or dialect words is affected by intoxication. More specifically, we want to test the hypothesis that speakers under the influence of alcohol use significantly more dialectal terms than when sober. The rationale behind this hypothesis is the fact that intoxicated persons in general show more unrestrained behavior than when sober. Most speakers in Germany use a dialect: we expect speakers to use their dialect more freely when under the influence. If this hypothesis can be confirmed with data of the ALC, it would be a useful tool for forensic investigations and/or methods to automatically detect intoxication from the speech signal.

[1] In this study only speakers with a blood alcohol concentration (BAC) higher than 0.05% were analyzed, resulting in a total of 148 speakers.

2. Speech data

The data used in this study are taken from the German Alcohol Language Corpus (ALC) which comprises recordings of intoxicated and sober speech of 162 German speakers (77 female, 85 male) in three different speech styles: the “read speech” part comprises telephone numbers, spelling of city names, addresses and German tongue twisters; the “spontaneous speech” parts consists of image descriptions (monologues) and question answering, e.g. “Which was the best gift you’ve ever received?”; the “command and control” (C&C) speech part contains verbal commands typically used for vehicle navigation and edutainment systems. C&C recordings were either read from the screen or prompted by a technique called “situational prompting” (Mögele et al, 2006) where the speaker is given a description of a real-world context, and is asked to think of and utter a command within that context. This results in “semi-spontaneous” speech in the sense that the wording of the targeted command is not prompted.

Each speaker was recorded sober and with one self-selected BAC level varying from 0.023% to 0.175% (median 0.089%). An important feature of the ALC recordings is that the speakers were aware that they were being recorded in a forensic project. All recordings were transcribed orthographically and tagged for dialectal word usage (among other tags) by a group of three annotators (inter-labeller agreement was not checked). Since the definition of a dialectal word (in contrast to a “normal” word) is far from clear, it was decided to tag all word forms as “dialectal” that are not listed in the German Duden of 2011 (standard of German orthography) and are not proper names. So, for instance colloquial words like “ausm” (“aus dem”) or “grade” (“gerade”) are not tagged as “dialectal” (because they appear in the Duden) while words like “bissel” (“bisschen”) or “vui” (“viel”) are tagged. Note that in German the context often determines whether a word form is in fact dialectal or not. For instance the regular article “des” (singular, male/neutral, genitive) is used instead of “das” (singular, neutral, nominative/accusative) in some southern German dialects. It follows that a dialectal tagging cannot be achieved by a simple lookup in a list of regular word forms without regarding the syntactic context. For further information about the recordings, the annotation process and the ALC in general see (Schiel et al, 2012); for free academic access to the corpus please refer to (CLARIN-ALC, 2013).

For this study recordings of read name spellings were filtered from the data set because here it is unlikely that dialectal variants occur. The remaining transcripts were word-tokenized and the number of dialectal tags was determined within each recording; to calculate the dialectal density these counts were normalized by the word count (repetitions and filled pauses such as “uhm” were counted as words). Analyzed recordings were assigned to two speaking styles: prompted addresses, tongue twister sentences and read C&C as “read”; monologues, dialogs and spontaneously uttered C&C as “spontaneous”. This procedure resulted in a total number of 12634 recordings, 8586 stemming from sober speakers, and 4049 stemming from intoxicated speakers. Table 1 shows the distribution of recorded word tokens across the factors intoxication, speaker gender and speaking style. The word distribution is far from uniform across factors. However, the amount of data is large enough to run logistic tests, which do not require uniform distributions within the data partitions. Figure 1 shows the histogram of the blood alcohol concentration (BAC) across speakers; note that each speaker was recorded with only one BAC level and when being sober.

Table 1: Word counts of analyzed spoken data across intoxication, speaker gender and speaking style (“spont” = “spontaneous”)

recorded word tokens			
302527			
intoxicated		sober	
92317		210210	
female	male	female	male
45098	47219	99445	110765
read	spont	read	spont
25079	67238	53212	156998

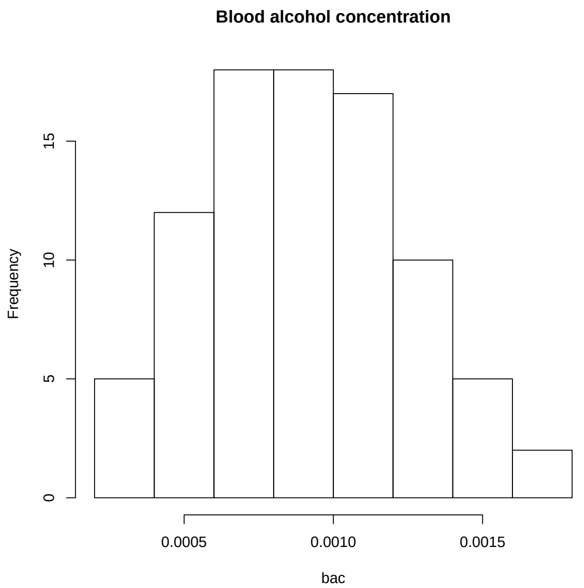

Figure 1: Histogram of blood alcohol concentration across speakers

3. Results

In this sections the dialectal density (in dialectal words per word) is tested against various factors. The main hypothesis here is that density significantly increases with speaker intoxication.

Since the data set contains multiple counts of dialectal words uttered by the same speaker under different conditions (intoxication, speaking style), a simple logistic test like Chi-Square would give misleading results caused by statistical dependencies within the individual speaker’s recording sets (see for instance (Baayen, 2008) for a discussion of within-speaker dependencies). Therefore a logistic Mixed Effect Model (MEM, (Baayen, 2008)) was applied that allows us to model the speaker as a random factor, and also allows testing of more than one independent factor and their possible interactions. To test for statistical dependencies on speaker characteristics we also included speaker gender (female/male) and the speakers' dialect region (defined as the German state where the speaker spent the majority of her/his childhood; 13 classes) as independent factors in our analysis. To verify that the obtained results are not caused by annotator idiosyncrasies we additionally tested for interactions with the factor annotator (3 classes).

3.1. Dialectal Density and Intoxication

The logistic MEM test for dialectal density yielded a highly significant effect ($z = -13.15, p(> |z|) < 0.0001$) on the (binary) intoxication state (BAC level zero or above zero). But the effect is in the opposite direction to our hypothesis: speakers utter significantly *less* dialectal words when intoxicated. Looking at the results for individual speakers, we found that 10 of the 162 speakers follow a reverse trend by uttering significantly *more* dialectal words in the intoxicated state. We carefully studied the meta-data (including BAC) of these 10 speakers but no common feature was found to explain their opposite behavior.
The (binary) state "intoxicated" involves a different BAC level for each speaker (from 0.023% to 0.175%), while "sober" means a BAC level of zero. The question is, do the dialectal density and the BAC level of a speaker correlate? To test for this we correlated the average dialectal density per speaker while intoxicated against their BAC level. A simple linear model yielded a weak, but significant negative correlation ($r^2 = 0.054, F[1, 160] = 9.14, p = 0.003$). Figure 2 shows the scatter plot and the linear regression between BAC and dialectal density.

3.2. Influence of Speaking Style

Spontaneous speech is much more affected by dialectal density than read speech ($z = 41.11, p(> |z|) < 0.0001$), which is not surprising: approx. 0.7% of read speech words are dialectal, while 4.9% are dialectal in spontaneous speech (measured across sober and intoxicated speech). But there is also a highly significant interaction between intoxication and speaking style ($z = -7.21, p(> |z|) < 0.0001$): while in read speech the dialectal density increases only slightly with intoxication (from 0.4% to 0.9%), the density drops significantly from 5.6% to 4.3% in spontaneous speech. The latter has a higher impact on the total data set, since the number of spontaneously uttered words (224236) was higher than read words (78291). The observed main effect, that dialectal density is decreased with intoxication, is therefore mainly caused by the changes in intoxicated spontaneous speech.

3.3. Influence of Speaker Gender and Dialect

No significant differences in speakers' gender or dialect regarding dialectal density were found in our data, nor any interactions with the intoxication state.

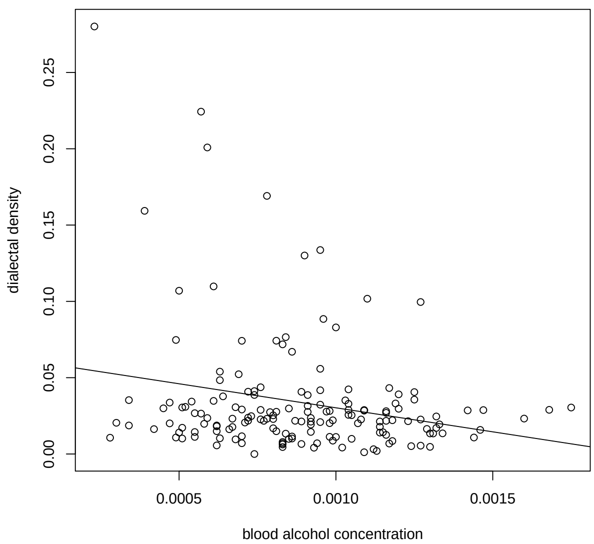

Figure 2: Correlation between average dialectal density and BAC across 162 speakers

3.4. Influence of Annotator

The transcripts of the three annotators do not differ significantly with regard to dialectal density, so there is no evidence for annotators' idiosyncrasies in the data. Also, we found no significant interaction with the speakers' intoxication state. It follows that no annotator showed a skewed tagging behavior when transcribing speech of an intoxicated speaker.

4. Discussion and Conclusion

In contrast to other linguistic effects caused by intoxication (see Introduction) the density of dialectal terms seems to *decrease* significantly with intoxication. For spontaneous speech we found a significant *negative* correlation between dialectal density and blood alcohol concentration. On the other hand we observe non-significant *increase* of dialectal density in intoxicated read speech.

One possible explanation of this observation may be social awareness of the speakers in combination with a deficient ability to control their speech production: intoxicated speakers recorded in ALC were aware of the fact that they were being recorded in a forensic experiment, and they were of course also aware of their intoxication. The speakers reduced their usual dialectal density when recorded in order to appear sober. But reduced ability to control speech production (because of intoxication) causes them to overcompensate, i.e. to produce a lower dialectal density than in sober speech.

There are two interesting lessons to be learned from this study: firstly, dialectal density seems to be a linguistic feature that speakers are aware of and one that can still be controlled by intoxicated speakers, in contrast to other linguistic factors such as word/hesitation/pronunciation error density, hesitation duration, and repetitions. Secondly, speakers may over-compensate for dialectal density when they are aware they are being observed in the intoxicated state. This might be of particular interest to forensic evaluations. Our findings suggest that this effect only happens in spontaneous speech.

5. Acknowledgments

We like to thank Matthias Lemser (BA) for providing the data extraction routines and performing the statistical tests. We also thank the ALC project group at Ludwig-Maximilians-Universität München for providing the Alcohol Language Corpus, and the CLARIN project for making this valuable resource available to the scientific community (CLARIN-BAS, 2014).

6. References

Alcohol Language Corpus - CLARIN Repository. Munich, Germany:Ludwig-Maximilians-Universität, accessed March 3, 2014, http://hdl.handle.net/11858/00-1779-0000-001A-1C5D-1.

Aldermann, G. A., Hollien, H., Martin, C., DeJong, G. (1995), "Shifts in fundamental frequency and articulation resulting from intoxication". In: Journal of the Acoustical Society of America 97, pp. 3363-3364.

Barfüsser, S. (2010), "Unstetigkeiten bei alkoholisierter Sprache". Master thesis, Ludwig-Maximilians-universität München, Germany.

Baayen, R. H. (2008), "Analysing Linguistic Data: A Practical Introduction to Statistics Using R". Cambridge University Press:Cambridge, pp. 263-328.

Baumeister, B., Heinrich, C., Schiel, F. (2012), "The influence of alcoholic intoxication on the fundamental frequency of female and male speakers". In: Journal of the Acoustical Society of America 132, pp. 442-451.

Chin S.B., Pisoni D.B. (1997), "Alcohol and Speech". San Diego, CA:Academic Press.

CLARIN Repository Bavarian Archive of Speech Signals. Munich, Germany:Ludwig-Maximilians-Universität, accessed March 3, 2014, http://hdl.handle.net/11858/00-1779-0000-0006-BF00-E.

Cooney, O. M., McGuigan, K. G., Murphy, P. J. P. (1998), "Acoustic analysis of the effects of alcohol on the human voice". In: Journal of the Acoustical Society of America 103, pp. 2895-2895.

Hollien, H., DeJong, G., Martin, C. A., Schwartz, R., Liljegren, K. (2001), "Effects of ethanol intoxication on speech suprasegmentals". In: Journal of the Acoustical Society of America 110, pp. 3198-3206.

Klingholz, F., Penning, R., Liebhardt, E. (1988), "Recognition of low-level alcohol intoxication from speech signal". In: Journal of the Acoustical Society of America 84, pp. 929-935.

Künzel, H.J., Braun, A., Eysholdt, U. (1992), "Einfluss von Alkohol auf Sprache und Stimme". Heidelberg: Kriminalistik Verlag.

Künzel, H. J., Braun, A. (2003), "The effect of alcohol on speech prosody". In: Proceedings of the ICPhS2003, Barcelona, Spain, pp. 2645-2648.

Martin, C. S., Yuchtman, M. (1986), "Using speech as an Index of Alcohol-Intoxication". Research on Speech Production No. 12, pp. 413-426.

Mögele, H., Kaiser, M., Schiel, F. (2006), "SmartWeb UMTS Speech Data Collection: The SmartWeb Handheld Corpus". Proc. of the LREC 2006, Genova, Italy.

Pisoni, D. B., Hathaway, S. N., Yuchtman, M. (1985), "Effects of alcohol on the acoustic-phonetic properties of speech: Final report to GM research laboratories". In: Research on Speech Perception Progress Report No. 11 (Indiana University, Bloomington, IN), pp. 109-171.

Schiel, F. (2011), "Perception of Alcoholic Intoxication in Speech". In: Proceedings of the Interspeech 2011, Florence, Italy, pp. 3281-3284

Schiel, F., Heinrich, C., Barfüsser, S. (2012), "Alcohol Language Corpus: The first public corpus of alcoholized German speech". In: Language Resources and Evaluation 46(3), pp. 503-521.

Smith, R.C., Parker, E.S., Noble, E.P. (1975), "Alcohol's Effect on Some Formal Aspects of Verbal Social Communication". In: Archives of General Psychiatry, 32, pp. 1394-1398.

Sobell, L. C., Sobell, M. B., Coleman, R. F. (1982), "Alcohol-induced dysfluency in nonalcoholics". In: Folia Phoniatrica 34, pp. 316-323.

Watanabe, H., Shin, T., Matsuo, H., Okuno, F., Tsuji, T., Matsuoka, M., Fakaura, J., Matsunaga, H. (1994), "Studies on vocal fold injection and changes in pitch associated with alcohol intake". In: Journal of Voice 8, pp. 340-346.