

Disfluencies in the speech of intoxicated speakers

Florian Schiel and Christian Heinrich

Abstract

Our hypothesis is that speakers under the influence of alcohol produce more linguistic/phonetic errors because of the negative effect of ethanol on cognitive processes and speech motor control. We examined the speech of 150 German speakers of both genders with regard to rates of 6 types of disfluencies and 2 durational measures. The intoxication of speakers ranged from 0.050% to 0.175% blood alcohol concentration; other factors investigated are speaker gender and speaking style (read, spontaneous, command&control). We found that most rates of disfluencies as well as durations increase with intoxication – but not for command&control speech; gender has no influence; individual speakers deviate from the general trend frequently. We conclude that for forensic investigations disfluency rates should be applied with greatest care (i.e. to individual speaker only), and command&control speech as typically used in automotive systems is not suitable for the automatic detection of intoxication on the basis of disfluency rates.

KEYWORDS ALCOHOL SPEECH, SPEECH DISFLUENCIES

Affiliations

Institut für Phonetik und Sprachverarbeitung, LMU München
email: schiel@phonetik.uni-muenchen.de heinrich@phonetik.uni-muenchen.de

1. Introduction

Alcohol intoxication, that is high levels of ethanol in the blood, reduces the blood flow to the cerebellum (Volkow, Mullani, Gould, Adler, Guynn, Overall and Dewey 1988), acts as a general depressant of the central nervous system and may cause euphoria, reduced inhibition, erratic behaviour, impaired balance, delayed reaction times and gradual loss of muscle coordination (ataxia) (Naranjo and Bremner 1993; Dawson and Reid 1997). The operational deterioration of mental processing and ataxia leads to the perceived phenomenon commonly known as 'alcohol speech', which not only refers to a less precise articulatory control, but also to deficient speech-planning processes. The latter might manifest themselves in the form of disfluencies such as repetitions, false starts, word and sentence breaks, changes in speaking rate, increased rate of unfilled and filled pauses, unusual prolongations of phones, and filler sounds. In the following we will refer to such observable speech events as disfluencies of speech, in the sense that each disfluency is a deviation from an 'ideal' fluent speech production, of the type a professional speaker would produce when reading from a script. This study explores the hypothesis that such observable disfluency events occur in the speech of sober and intoxicated speakers, but that the frequency correlates with the amount of impairment in the speech production process, and hence should also correlate with the amount of alcohol intoxication. If this were the case, it would imply that alcohol intoxication can be predicted more or less reliably from observed disfluency events, for instance in forensic investigations of speech recordings.

To verify the hypothesis, we defined a set of observable disfluency categories and labelled their occurrence in a large speech corpus of sober and intoxicated speakers. We then estimated labelled disfluency rates and durations of events that can also be regarded as typical for disfluent speech for the same individual speakers when sober and when intoxicated, and analysed the results with regard to additional factors of gender and speaking style in a generalised linear model (rates, e.g. Dobson 1990) and ANOVA (durations) respectively.

The next section discusses some earlier studies regarding speech disfluencies in spontaneous, non-intoxicated speech relevant for our present study. Section 3 discusses earlier studies regarding disfluencies in intoxicated speech. Sections 4 and 5 describe the speech data, the tagging methods and the statistical analysis. Sections 6 and 7 present and discuss the results on our data set and possible implications for forensic investigations.

2. Disfluencies in speech

Although there exists a number of recent studies about observable speech events that can be summarised as 'speech disfluencies', the term 'disfluency' (sometimes

also ‘dysfluency’) itself is not very well defined. For this study we follow a rather narrow definition, where ‘disfluency’ in the context of spoken language is ‘any of various breaks or irregularities that occurs within the flow of otherwise fluent speech’ (modified from Wikipedia 2014). That is, we do not consider pronunciation errors or slips-of-the-tongue as disfluencies, if they are not corrected by the speaker and do not cause any sort of interruption in the rhythmic structure. It follows that, for the purpose of this article, only disfluencies caused by impaired mental and motor control performance are relevant. Therefore we restrict this introduction about disfluencies to the following three studies.

Shriberg (2001) analysed the rates of filled pauses, repetitions, omissions, replacements and insertions of words, and pronunciation errors in spontaneous speech exchanged between human interlocutors as well as between a human and a computerised speech interface. Shriberg found significant differences in rates between individual speakers, as well as according to the position within the phrase/sentence (initial more frequent than medial or final), and higher rates for men than for women. Shriberg attributed the higher rate of disfluencies in initial position to the higher mental load associated with speech planning processes at the beginning of a phrase/sentence (Shriberg 2001). If this is true, we would also expect a positive correlation between intoxication and rate of disfluencies, since alcohol intoxication is known to inhibit mental processing. All of Shriberg’s analysed events could be possible candidates for our investigation. However, the insertion/deletion/repetition of words is only observable in read speech (where there is a reference). Therefore, we omitted these events in our analysis, since we investigate several speech styles. In addition, we do not classify (uncorrected) pronunciation errors as disfluencies. Shriberg also reported a significant influence of the conversational partner on the rate of disfluencies.

Bortfeld, Leon, Bloom, Schober and Brennan (2001) counted disfluencies in conversations in American English. Telephone conversations yielded higher rates (8.83 disfluencies per 100 words) than face-to-face (5.5). Older speakers produced slightly more disfluencies than younger speakers, and the conversational partner had no influence on the counts. The authors conclude that disfluencies may be used by speakers to signal help requests, or as a prohibitive turn taking signal (‘do not interrupt me’), which would be more likely to occur in telephone conversations where visual cues are not present. The same effect was reported by Clark and Fox Tree (2002) for filled pauses ‘uh’ and ‘um’ (2.9 per 100 words in telephone vs 2.3 in face-to-face conversations). Clark and Fox Tree (2002) also found that these hesitations improved the speed at which following words were recognised, possibly by signalling to the listener to focus on the following word. The role of filled pauses is undoubtedly more complex than that of other disfluencies. For now we assume that there are two (maybe more) possible causes for

a speaker to use a filled pause: either as an active communication signal or as a default production to mask performance/planning delays in his/her speech. It follows that the prediction of filled pause rates under the influence of alcohol is ambivalent: if hesitations are mainly used as active communicative markers, the rate might decrease in intoxicated speakers because of impaired speech planning including (meaningful) hesitations; if filled pauses rather indicate performance problems (for instance because the speaker needs more time for planning processes), the rate might increase under the influence of alcohol; it is also possible that we might encounter a mixed effect whereby both cancel each other out in our data.

To our knowledge no studies have yet addressed durations of filled and unfilled pauses concerning speech under the influence of alcohol.

3. Earlier studies of disfluencies in speech of intoxicated speakers

It is a common hypothesis that the intake of ethyl alcohol as well as other factors such as emotional state, fatigue, stress and illness all influence the way a person speaks. A number of studies during the last decades have investigated this hypothesis from different points of view: looking for reliable acoustic (Künzel and Braun 2003; Cooney, McGuigan, Murphy and Conroy 1998) or behavioural (Hollien, DeJong, Martin, Schwartz and Liljegren 2001; Behne, Rivera and Pisoni 1991; Trojan and Kryspin-Exner 1968) features that may indicate intoxication, studying the physiological effects of alcohol on the articulators (Watanabe, Shin, Matsuo, Okuno, Tsuji, Matsuo, Fukaura and Matsunaga 1994) or pursuing forensic questions (Künzel and Braun 2003; Braun 1991; Klingholz, Penning and Liebhardt 1988; Martin and Yuchtman 1986), such as in the infamous case of the captain of the oil tanker Exxon Valdez that caused an oil spill in Prince William Sound, Alaska, on 24 March 1989 (Johnson, Pisoni and Bernacki 1990).

With respect to intoxication, to our knowledge only four studies exist that report empirically based rates/measures of disfluencies, as this term was defined in the previous section (Künzel, Braun and Eysholdt 1992; Hollien, Liljegren, Martin and DeJong 1999; Hollien et al. 2001; Christenfeld and Creager 1996; Tisljár-Szabó, Rossu, Varga and Pléh 2013).

Künzel et al. (1992) studied read and semi-spontaneous (picture description) speech of 33 male speakers with varying intoxication levels verified by breath alcohol concentration (BrAC). Along with many other features, the authors counted insertions, omissions and substitutions (as in Shriberg 2001) but on several linguistic levels: phones, syllables, words and phrases. Additionally, repetitions and pause rates were counted. Aside from insertions (in read speech only), no feature showed a clear tendency to increase under the influence of alcohol

(Künzel et al. 1992: 19). There was a weak tendency for omission and substitution rates to increase for BrAC > 0.08%. Repetition and pause rates did not change significantly with intoxication. However, when the detected unfilled pauses were divided into two semantically different groups ‘content pauses’ and ‘hesitations’, the authors found a significant increase in the latter for speakers with more than 0.08% BrAC.

Hollien et al. (1999 and 2001) is the only previous study regarding disfluencies to deal with speakers of both genders: read speech from 16 female and 23 male speakers was analysed with regard to insertion/substitution/repetition of phones/words (including filled pauses), voicing, de-voicing, lengthening of phones, and pronunciation errors. Speakers were recorded when sober and at three different BrAC levels ranging from 0.04% to 0.13%. The different types of disfluencies were pooled into one single disfluency index. Only rates for this index were reported: pooled rates increase significantly with BrAC (the effect was up to 150% from sober to the highest BrAC). Hollien et al. (1999 and 2001) conclude that their pooled disfluency counts are a reliable indicator of alcohol intoxication.

Christenfeld and Creager (1996) interviewed 108 male speakers in pubs and counted the number of tokens of the filled pause ‘um’. Speakers’ blood alcohol concentration (BAC) was estimated from the number of drinks they had before the interview (11 sober, the majority with three drinks). The authors report a weak but significant negative correlation of hesitations with intoxication (number of drinks), i.e. the number of ‘ums’ decreases under the influence of alcohol. Clark and Fox Tree (2002) offer two possible explanations for this surprising finding: if hesitations are communication signals that need to be planned, intoxicated speakers may be increasingly less able to plan for these signals, or alternatively, they may be less concerned with making their speech intelligible for the listener and may therefore drop these helpful communication signals.

Tisljár-Szabó et al. (2013) analysed the number of unfilled pauses in speech from 15 Hungarian under-graduate speakers of both genders. They found an average increase in the number of unfilled pauses in intoxicated speech. Since the applied tests (paired t-test) were not appropriate for logistic values and the methodology is unclear, it remains uncertain whether these results were significant.

4. Speech Data

All analysed data were derived from the German Alcohol Language Corpus (ALC) (Schiel, Heinrich and Barfußler 2012), which is a collection of speech audio samples of 77 female and 85 male speakers in both intoxicated and sober condition (about 5 minutes speech material in intoxicated and 10 minutes in sober condition per speaker). In contrast to most other studies investigating speech under the influence of alcohol, blood samples (BAC) were taken, and

ALC comprises three different speaking styles: read speech (numbers, tongue twisters, addresses), spontaneous speech (picture description, interview) and command&control speech (typically used in an automotive communication system to operate devices like navigation system, radio or climate control unit, for example: 'Temperature 23 degrees Celsius').

Participants were asked to drink a certain number of alcoholic beverages within a maximum time period of about two hours. Ethical restrictions required that participants could choose their desired BAC level for the experiment individually. Therefore the number of beverages – depending on the body height, weight, age and gender of the participant – was calculated, targeting a BAC which was chosen beforehand by every participant in the range of 0.05%–0.15%. The lower bound of 0.05% was motivated by German law which does not allow driving in public traffic with a BAC higher than 0.05%; the upper bound was set to protect the health of the participants. The calculation is based on a combination of formulas established by Erik M. P. Widmark (Widmark 1932) and P. E. Watson (Watson, Watson and Batt 1980).

After the consumption of the alcoholic beverages and a wait of a further 20 minutes, which is necessary for the residual alcohol to evaporate in the mouth cavity, both breath and blood alcohol concentrations were determined. Immediately after these measurements, speech samples were recorded and speaker specific metadata were collected, such as gender, age, dialect background, smoking habits and drinking habits.

The measured BAC levels of the 162 participants ranged from 0.023% to 0.175% in an approximate Gaussian distribution with a mean at 0.085%; all recordings under the influence of alcohol are labelled as 'intoxicated state' in the following, regardless of their individual BAC level. After a minimum of two weeks a second group of speech samples from the same speaker were recorded in sober condition. To avoid the influence of the interview partner as reported by Shriberg 2001, all 162 speakers were interviewed by the same partner in both recordings.

A subset of 20 speakers (all showing a BAC above 0.05% in the first experiment) were invited a third time to repeat the first recording under exactly the same conditions but without intoxication. This so-called control set is intended to allow unknown influential factors in the experimental setup to be identified. All recorded audio data were transcribed orthographically by trained phoneticians using the online annotation tool Webtranscribe (Draxler 2005). Filled pauses, repetitions, false starts, word interruptions, unusual phone lengthenings and other linguistic events according to the Verbmobil transliteration standard (Burger, Weilhammer, Schiel and Tillmann 2000) were tagged in the orthographic transcription. Based on the orthographic transcripts, signals were automatically segmented and labelled into phonetic units including unfilled pauses using

the MAUS technique (Schiel 1999). The MAUS tool was configured to detect potential unfilled pauses between word units but not within word units (which is treated as the disfluency class ‘word interruption’); the minimum length of a detected unfilled pause was set to 50 msec, since smaller silence intervals are not perceived as pauses; detected unfilled pauses of less than 50 msec length were equally spread across the adjacent phone segments.

For more detailed information about the corpus and the recording/annotation procedures, see Schiel et al. 2012.

5. Method

The present study deals with observable deviations from the normal rhythmic speech production of sober speakers. Based on previous studies and the given annotation and automatic segmentation of the ALC, the investigated disfluency categories include, therefore, the rate of filled pauses, the rate of inter-word unfilled pauses of greater length than 50 msec, the rate of repetitions, the rate of false starts (often including word breaks), the rate of word interruptions (mostly by inserted filled or unfilled pauses) and the rate of unusual phone lengthenings (judged by the labeller). Table 1 describes the manually tagged disfluency classes and corresponding markers and gives examples for their transcription. Pronunciation errors/slips-of-the-tongue in otherwise fluent speech which are often subsumed under ‘disfluencies’ are not investigated here. Due to the sparsity of observed disfluency events, metadata (except speaker gender) are not tested as independent factors in this study, nor are disfluency counts correlated against the BAC level.

Table 1: Manually tagged disfluency classes in the ALC transcription used in this study

Disfluency	Marker	Example
Filled pause	<"ah>, <hm>, <"ahm>, <hes>	in <"ah> Rom
Repetition	+/.../+	+/ein/+ +/ein/+ Haus ...
False start	-/.../-	-/der Mann/- die Frau ...
Word interruption	_..._	Urlaubs_<hm>_budget Haupt_<P>_bahnhof
Phone lengthening	<Z>	und dann<Z> sind wir ... gehen wir na<Z>ch ...

Speech data of 150 selected speakers of ALC were analysed. These 150 speakers all exhibited a blood alcohol concentration minimum of 0.05% while intoxicated (0.05% is the legal limit for driving in Germany). Speech data were divided into 900 bins where each bin is defined by speaker (150), intoxication state (2) and speaking style read, spontaneous and command&control (3). Disfluency markers based on the manual tagging (see Table 1) were counted automatically for

every bin. Filled pauses and phone lengthening counts were then normalised to the total number of syllables in the bin, while false starts, repetitions and word interruptions counts were normalised to the total number of words in the bin to compensate for the differing amounts of speech per bin. Rates of unfilled pauses (also normalised to total number of words), durations of unfilled pauses and durations of filled pauses were derived from the automatic phonetic segmentation and averaged within each bin; durations were not normalised to speech rate. This procedure results in a total of 7200 rates/measurements (6 rates + 2 duration measures in 900 bins).

Statistical analysis on rates was carried out using generalised linear models in R (function `glm()` in R Version 3.1.1) with the respective disfluency rate as the dependent variable, intoxication and gender as independent variables (both binary factors); measured durations were tested using a linear model and ANOVA (functions `lm()`, `anova()` in R version 3.1.1) with intoxication and gender as independent variables. Since models including speaking style as a fixed factor in all cases indicated strong interactions with speaking style (which is to be expected), we fitted models for all three speaking styles separately; this results in only one rate/measure per speaker and intoxication state, and therefore does not require additional modelling of speakers as a random variable to prevent statistical correlations within the data of a single speaker.

All rates/measures were cross-verified to a control group recording without intoxication to rule out hidden factors.

6. Results

Table 2 lists the change in rate of the measured disfluencies (in percent) and durations (in milliseconds) from sober to intoxicated speech averaged across all 150 speakers and for the three speaking styles read, spontaneous and command&control speech. Significance levels from the corresponding generalised linear model are given in brackets. For example, for read speech the average rate of filled pauses changes from 0.28% in sober speech to 0.42% in intoxicated speech, while the average duration of filled pauses does not change substantially (279 msec to 280 msec). The gender factor was never found to interact with intoxication; the results are therefore assumed to be gender-independent. Figure 1 illustrates the averaged rates/durations for intoxicated (i) and sober (s) speech and the three speaking styles read (read), spontaneous (spont) and command&control (comm) as bar plots.

There is a general problem when testing very low rates: standard tests like chi-squared or the generalised linear model are not reliable with rates that are very close to 0% or 100%. In our case this involves mainly repetitions, word interruptions and phone lengthenings. For example, the rate of phone lengthenings in

Table 2: Average change of absolute disfluency rates in percent and durations in milliseconds from sober to intoxicated state for 150 speakers of ALC (significance level of generalised linear model)

Disfluency	Read speech	Spontaneous speech	Command&control speech
Rate of filled pauses	0.28% → 0.42% (p < 0.001)	2.23% → 2.5% (p < 0.05)	0.5% → 0.6% (n.s.)
Rate of unfilled pauses >50msec	21.44% → 27.46% (p < 0.001)	22.81% → 24% (p < 0.001)	18.12% → 18.24% (n.s.)
Rate of false starts	0.31% → 0.38% (p < 0.01)	1.19% → 1.15% (n.s.)	0.59% → 0.82% (p < 0.05)
Rate of repetitions	0.31% → 0.33% (n.s.)	0.72% → 0.46% (p < 0.001)	<u>0.09% → 0.15%</u> (n.s.)
Rate of word interruptions	0.33% → 0.55% (p < 0.001)	<u>0.028% → 0.029%</u> (n.s.)	<u>0.065% → 0.064%</u> (n.s.)
Rate of phone lengthening	0.13% → 0.25% (p < 0.001)	0.31% → 0.47% (p < 0.001)	<u>0.06% → 0.1%</u> (n.s.)
Duration filled pauses	279 ms → 280 ms (n.s.)	424 ms → 452 ms (p < 0.05)	353 ms → 344 ms (n.s.)
Duration unfilled pauses >50msec	172 ms → 205 ms (p < 0.001)	367 ms → 402 ms (p < 0.05)	256 ms → 265 ms (n.s.)

command&control speech changes on average from 0.06% to 0.10%, that is, the rate is increased by 40%. However, the corresponding model reports a p-level above 0.05, which is not significant by common standards. The reason for these unexpected high p-levels is the very low number of observed events; in very rare events one single random fluctuation may have a very large effect. Based on our data it is therefore undecidable whether these disfluencies are influenced by intoxication regardless of their high absolute rate changes. We underline these unreliable test results to alert the reader, and in the following we consider all non-significant test results, including these unreliable p-levels, as for the time being 'unchanged'.

Only two average disfluency rates decrease when the speaker is intoxicated: false starts and repetitions, both only for spontaneous speech. All remaining average disfluency rates increase, when there is a significant change from sober to intoxicated.

Average durations of filled and unfilled pauses generally increase, when there is a significant change.

To ensure that the displayed effects are in fact caused by the intoxication and not another, yet unknown, factor, we also examined disfluency rates/durations for the control subset. (Recall from Section 4 that speakers were sober in both, the second and third recording sessions.) All rates/measurements showed non-significant changes between the second and third recording sessions. As an example Figure 2 shows the change in phone lengthening counts sorted across

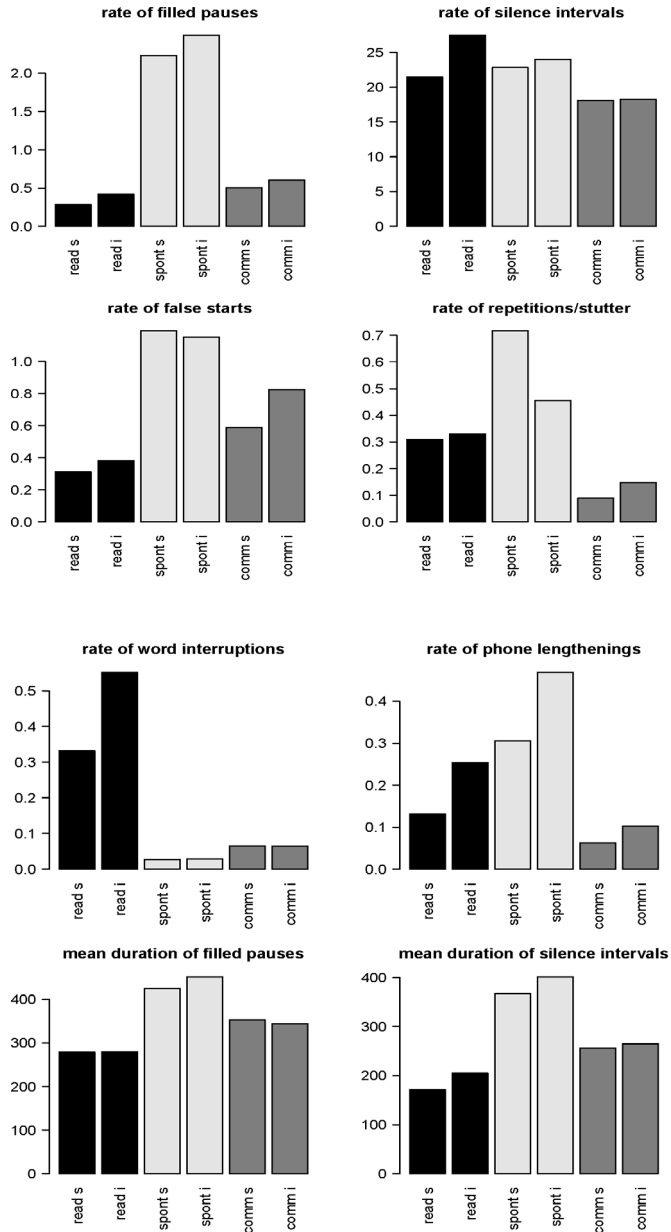


Figure 1: Barplots displaying the average change of disfluency rates in percent and durations in milliseconds from sober (s) to intoxicated state (i) for the three speaking styles

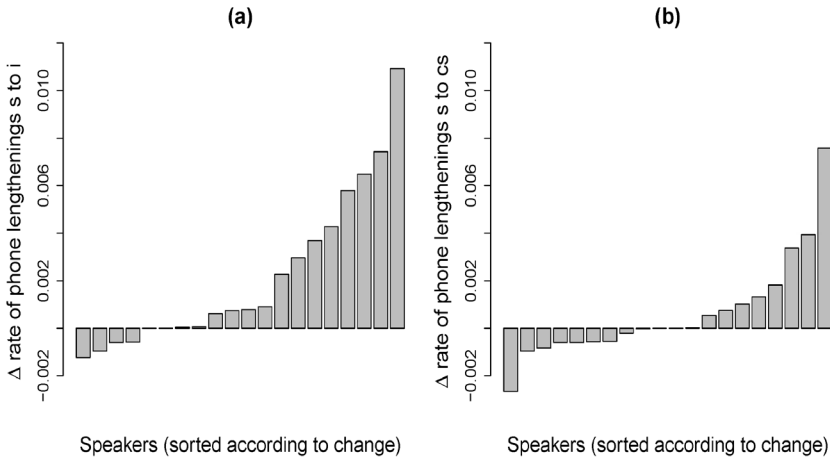


Figure 2: Sorted barplots across the 20 speakers of the control set displaying the average change of the rate of phone lengthenings in percent from sober (s) to intoxicated (i) state(a) and from sober (s) to control condition sober (cs) state(b)

the 20 speakers of the control subset between sober and intoxicated state (a) and for the control condition (b). While the majority of speakers increase the number of phone lengthenings when intoxicated (a), in the control condition the proportion of speakers who increase is about equal to the proportion of speakers who decrease their count, which indicates random behaviour in the control experiment. We therefore conclude that the overall increase in rate of phone lengthenings in the main experiment is caused by intoxication.

7. Discussion

In general, with a few exceptions, the rates of disfluencies observed in this study rise with intoxication in read and spontaneous speech, which is in line with our main hypothesis and with most earlier studies, as discussed in Section 3.

There is a tendency for a higher percentage of filled pauses in speech under the influence of alcohol compared to speech in sober condition, especially for read speech. This contradicts Clark and Fox Tree (2002), who reported a decrease in the number of filled pauses. We also found that the number and length of unfilled pauses increase in speech under the influence of alcohol (except for command&control speech). These three effects could indicate planning difficulties when intoxicated, for which the speaker compensates by inserting silent and filled pauses.

Repetitions seem to occur less often in spontaneous speech under the influence of alcohol, which contradicts our main hypothesis that disfluency rates would

increase with rising intoxication level. At the moment we do not have a conclusive explanation for this observation.

The average duration of filled and unfilled pauses rises with intoxication mainly for spontaneous speech. However, these effects can partly be explained by the reduced speaking rate under the influence of alcohol. Heinrich and Schiel (2011) found that speaking rate decreases on average by 4–5% in intoxicated speech, which would be sufficient to explain most of the observed changes, with the possible exception of unfilled pause length in read speech, which shows a 19% increase on average.

As is often the case with correlations of measurable observations against speaker states, there is a heterogeneous picture across speakers: while the majority of speakers may increase a certain linguistic/phonetic rate/measure with intoxication, other speakers decrease or do not change the same rate/measure at all. This idiosyncratic behaviour can be observed in the results of this study as well as in many other linguistic/phonetic features in combination with alcohol (e.g. Johnson et al. 1990; Hollien et al. 2001; Künzel and Braun 2003; Heinrich and Schiel 2011). For example in Figure 2, twelve speakers increase the rate of phone lengthenings, four decrease, and four do not change at all with intoxication; patterns for other rates/measures analysed in this study are very similar. Speaker-individual disfluency rates (as shown for example in Figure 6.2) seem not to be correlated to the measured BAC level (although it must be kept in mind that such correlations cannot be calculated reliably due to the very low occurrence rate of disfluencies). It follows that there can be no generally valid expectation for intoxicated speakers' behaviour regarding disfluencies; forensic investigations as well as automatic detection systems for intoxication should take this into consideration.

The command&control speaking style shows no significant changes with intoxication for disfluency rates or durations (except a weak increase of the rate of false starts). The reason for this result is probably the very simple linguistic nature of these commands, which do not require major planning and production effort on the part of the speaker. This result seems to indicate that (simple) speech directed to a speech-driven interface is not suitable for the automatic detection of speaker intoxication by means of disfluencies alone.

Acknowledgments

Sabine Barfüsser (MA) assisted the authors in the first experiments regarding disfluencies in the ALC corpus (unpublished). The authors would like to thank the association 'Bund gegen Alkohol und Drogen im Strassenverkehr e.V.' and the 'Bavarian Archive for Speech Signals' for providing the speech corpus 'Alcohol Language Corpus'. We also thank the European CLARIN initiative for making the corpus available for academic research.

Declaration of interest

There exists no conflict of interests relevant to the contents of this article regarding both authors. There is no involvement of any commercial company in the funding of this study.

About the authors

Florian Schiel received his Dipl.-Ing. and Dr.-Ing. degrees from the Technical University in Munich in 1990 and 1993 respectively, both in electrical engineering. Since 1993 he has been affiliated to the Institute of Phonetics, Ludwig-Maximilians-Universität Munich (LMU), leading the German VERBMOBIL, SmartKom, BITS and SmartWeb project groups. In 1994 and 1997 he spent 6 months of each year as a research fellow at the International Computer Science Institut (ICSI), Berkeley, California. Since 2001 he has held the chair of Phonetic Speech Processing at LMU. From 1996 to 2015 he acted as founding director of the Bavarian Archive for Speech Signals (BAS) at the LMU.

Christian Heinrich received his MA in 2007 and his Dr.phil. degree in 2014 from Ludwig-Maximilians University Munich, both in Phonetics and Speech Processing. His doctoral thesis is concerned with the rhythmical structure of speech under the influence of alcohol. Since 2015 he has been product manager and part of the integrated speech solutions team at DFC-Systems in Munich Germany.

References

- Behne, D. M., Rivera, S. M. and Pisoni, D.B. (1991) Effects of alcohol on speech: durations of isolated words, sentences and passages. *Research on Speech Perception* 17: 285–301.
- Braun, A. (1991) Speaking while intoxicated: phonetic and forensic aspects. *Proceedings of the XIIth International Congress of Phonetic Sciences*, Aix-en-Provence: 146–149.
- Bortfeld, H., Leon, S., Bloom, J., Schober, M. and Brennan, S. (2001) Disfluency rates in conversation: effects of age, relationship, topic, role, and gender. *Language and Speech* 44: 123–147. <http://dx.doi.org/10.1177/00238309010440020101>
- Burger, S., Weilhammer, K., Schiel, F. and Tillmann, H. G. (2000) Verbmobil data collection and annotation. In W. Wahlster (ed.) *Verbmobil: Foundations of Speech-to-Speech Translation* 539–551. Berlin: Springer. http://dx.doi.org/10.1007/978-3-662-04230-4_39
- Christenfeld, N. and Creager, B. (1996) Anxiety, alcohol, aphasia and ums. *Journal of Personality and Social Psychology* 70: 451–460. <http://dx.doi.org/10.1037/0022-3514.70.3.451>
- Clark, H. H. and Fox Tree, J. E. (2002) Using *uh* and *um* in spontaneous speaking. *Cognition* 84: 73–111. [http://dx.doi.org/10.1016/S0010-0277\(02\)00017-3](http://dx.doi.org/10.1016/S0010-0277(02)00017-3)

- Cooney, O. M., McGuigan, K., Murphy, P. and Conroy, R. (1998) Acoustic analysis of the effects of alcohol on the human voice. *Journal of the Acoustical Society of America* 103: 2895. <http://dx.doi.org/10.1121/1.421829>
- Dawson, D. and Reid, K. (1997) Fatigue, alcohol and performance impairment. *Nature* 388: 235. <http://dx.doi.org/10.1038/40775>
- Dobson, A. J. (1990) *An Introduction to Generalized Linear Models*. London: Chapman and Hall.
- Draxler, C. (2005) WebTranscribe – an extensible web-based speech annotation framework. *Proceedings of TSD 2005*: 61–68. http://dx.doi.org/10.1007/11551874_8
- Heinrich, C. and Schiel, F. (2011) Estimating speaking rate by means of rhythmicity parameters. *Proceedings of the Interspeech 2011*: 1873–1876.
- Hollien, H., DeJong, G., Martin, C. A., Schwartz, R. and Liljegen, K. (2001) Effects of ethanol intoxication on speech suprasegmentals. *Journal of the Acoustical Society of America* 110: 3198–3206. <http://dx.doi.org/10.1121/1.1413751>
- Hollien, H., Liljegen, K., Martin, C. A. and DeJong, G. (1999) Prediction of intoxication levels by speech analysis. In A. Braun (ed.) *Advances in Phonetics*, 40–50. Stuttgart: Steiner Verlag.
- Johnson, K., Pisoni, D. B. and Bernacki, R. H. (1990) Do voice recordings reveal whether a person is intoxicated? A case study. *Phonetica* 41: 215–237. <http://dx.doi.org/10.1159/000261863>
- Klingholz, F., Penning, R. and Liebhardt, E. (1988) Recognition of low-level alcohol intoxication from speech signal. *Journal of the Acoustical Society of America* 84: 929–935. <http://dx.doi.org/10.1121/1.396661>
- Künzel, H. J. and Braun, A. (2003) The effect of alcohol on speech prosody. *Proceedings of the ICPhS*, Barcelona: 2645–2648.
- Künzel, H. J., Braun, A. and Eysholdt, U. (1992) *Einfluß von Alkohol auf Sprache und Stimme*. Heidelberg: Kriminalistik Verlag.
- Martin, C. S. and Yuchtman, M. (1986) Using speech as an index of alcohol-intoxication. *Journal of the Acoustical Society of America* 79: 413–426. <http://dx.doi.org/10.1121/1.2023207>
- Naranjo, C. A. and Bremner, K. E. (1993) Behavioural correlates of alcohol intoxication. *Addiction* 88(1): 31–41. <http://dx.doi.org/10.1111/j.1360-0443.1993.tb02761.x>
- Schiel, F. (1999) Automatic phonetic transcription of non-prompted speech. *Proceedings of the ICPhS*: 607–610.
- Schiel, F., Heinrich, C. and Barfüßer, S. (2012) Alcohol language corpus. *Language Resources and Evaluation* 46 (3): 503–521. <http://dx.doi.org/10.1007/s10579-011-9139-y>
- Shriberg, E. E. (2001) To 'errrr' is human: ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association* 31(1): 153–169. <http://dx.doi.org/10.1017/s0025100301001128>
- Tisljár-Szabó, E., Rossu, R., Varga, V. and Pléh, C. (2013) The effect of alcohol on speech production. *Journal of Psycholinguistic Research* 43(6): 737–748. <http://dx.doi.org/10.1007/s10936-013-9278-y>

- Trojan, F. and Kryspin-Exner, K. (1968) The decay of articulation under the influence of alcohol and paraldehyde. *Folia Phoniatica* 20: 217–238. <http://dx.doi.org/10.1159/000263201>
- Volkow, N. D., Mullani, N., Gould, L., Adler, S. S., Guynn, R. W., Overall, J. E. and Dewey, S. (1988) Effects of acute alcohol intoxication on cerebral blood flow measured with PET. *Psychiatry Research* 24: 201–209. [http://dx.doi.org/10.1016/0165-1781\(88\)90063-7](http://dx.doi.org/10.1016/0165-1781(88)90063-7)
- Watanabe, H., Shin, T., Matsuo, H., Okuno, F., Tsuji, T., Matsuoka, M., Fukaura, J. and Matsunaga, H. (1994) Studies on vocal fold injection and changes in pitch associated with alcohol intake. *Journal of Voice* 8: 340–346. [http://dx.doi.org/10.1016/S0892-1997\(05\)80282-6](http://dx.doi.org/10.1016/S0892-1997(05)80282-6)
- Watson, P. E., Watson, I. D. and Batt, R. D. (1980) Total body water volumes for adult males and females estimated from simple anthropometric measurements. *The American Journal of Clinical Nutrition* 33(1): 27–39.
- Widmark, E. M. P. (1932) *Die theoretischen Grundlagen und die praktische Verwendbarkeit der gerichtlich-medizinischen Alkoholbestimmung*. Berlin: Urban und Schwarzenberg.