NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Emotional communication in long-term abstained alcoholics

Chelsea Harmsworth (MSc) and Silke Paulmann (PhD)

University of Essex

Department of Psychology and Centre for Brain Science

Corresponding author:

Silke Paulmann

Department of Psychology and Centre for Brain Science

Wivenhoe Park, Colchester, CO4 3SQ

Phone: +44-1206-873422

Email: paulmann@essex.ac.uk

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Abstract

Background: Alcoholism is associated with difficulties in perceiving emotions through non-verbal channels including prosody. The question whether these difficulties persist to long-term abstinence has, however, received little attention. **Methods:** In a two-part investigation, emotional prosody production was investigated in long-term abstained alcoholics and age- and education matched healthy controls. First, participants were asked to produce semantically neutral sentences in different emotional tones of voice. Samples were then acoustically analyzed. Next, naïve listeners were asked to recognize the emotional intention of speakers from a randomly collected subset. Voice quality indicators were also assessed by the listeners. **Results:** Findings revealed emotional prosody production differences between the two groups. Differences were particularly apparent when looking at pitch use. Alcoholics' mean and variability of pitch differed significantly from controls' use. The use of loudness was affected to a lesser extent. Crucially, naïve raters confirmed that the intended emotion was more difficult to recognize from exemplars produced by alcoholics. Differences between the two groups were also found with regard to voice quality. **Conclusions:** These results suggest that emotional communication difficulties can persist long after alcoholics have quit drinking.

Keywords: Alcoholism; Social Cognition; Emotional Prosody; Vocal Emotion

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

INTRODUCTION

Non-verbal emotion signals form a crucial part of social interactions: we can encode a range of emotional states based on others' use of facial expressions, body postures, or prosody (sometimes referred to as "tone of voice"). Alcoholism is often associated with deficits in processing these kinds of emotional signals. Specifically, recently detoxified alcoholics demonstrate difficulties in perceiving emotions through a range of non-verbal channels including facial expressions (Frigerio et al., 2002; Philippot et al., 1999), body postures (Maurage et al., 2009), and prosody (Monnot et al., 2001; Uekermann et al., 2005). Some research suggests that these perception difficulties are long lasting as they have been found to persist through to mid- and long-term abstinence (Foisey et al., 2007; Kornreich et al., 2001; Valmas et al., 2014).

Accurate recognition of emotional signals is, however, only one part of successful social interactions. Properly and authentically *expressing* emotional states is just as important. This is particularly true for *vocal* emotional communication as listeners rely heavily on prosody to make inferences about the speaker's intentions and feelings in cases where verbal messages are ambiguous or lack emotional content (e.g., "I'll see you next week" can be said in a happy, cheerful tone of voice suggesting that the speaker is looking forward to this event, or it can be said in an annoyed, angry tone of voice suggesting quite the opposite). Clearly, both failure to detect and failure to express vocal emotional intentions effectively can lead to interpersonal communication breakdown. However, while an increasing number of studies have tried to describe the role of alcoholism in emotional prosody *perception* (Oscar-Berman et al., 1990; Monnot et al., 2001; Uekermann et al., 2005), research on emotional prosody *production* in alcoholics has been largely neglected. The present investigation aims to start fill this gap in the literature by exploring how long-term abstainers[1] express vocal emotions and, crucially, how these emotional intentions are perceived by naïve listeners. When expressing how we feel, we modulate various acoustic cues, such as fundamental frequency (perceived as pitch), loudness, or tempo. For instance, it has been shown that we increase our mean and range of pitch and

---

[1] Here, we follow conventions in the literature (e.g., Kornreich et al., 2001; Fein et al., 2010) who use the term "long-term" abstainers for individuals who have abstained from alcohol for more than six months.

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

loudness when expressing anger (as opposed to, for instance, neutral) and we also speak considerably faster when angry. Further acoustic cue profiles are associated with other emotions (for example, when expressing sadness, speakers use a smaller range of pitch and loudness and decrease their speech rate; see Banse & Scherer, 1996). Inadequate acoustic cue use is likely to lead to difficulties in listeners' abilities to recognise how the speaker feels. To the best of our knowledge, there is only one previous study that has focused on the production of vocal emotions in alcoholics. Monnot and colleagues (2003) asked 24 detoxified alcoholics and 15 healthy controls to intone sentences in one of five emotions (happiness, sadness, anger, boredom, surprise) and in a neutral tone. Four researchers were then asked to identify the expressed emotions. Detailed acoustical analyses of produced speech were not provided in this study, limiting our ability to specify how alcoholics might differ in their emotional expressions from healthy controls. Also, judges' exact accuracy rates were not reported, leaving it unclear as to how difficult listeners might find it to recognize emotions expressed through speech from detoxified alcoholics. However, the authors report that pitch was positively linked to how accurately the four judges rated the intended emotion, suggesting that pitch is particularly important when encoding emotional speech in alcoholics. Moreover, this research highlights that adequate pitch variations are key to expressing vocal emotions. Given the lack of information about other acoustic cues used in this sample, it remains unclear which additional parameters listeners relied on when judging emotions expressed by detoxified alcoholics and it is also not possible to comment on potential cue use differences between detoxified alcoholics and controls. Finally, the question of whether a history of alcohol abuse can have long-term effects on emotional prosody production cannot be answered with data from recently detoxified alcoholics. This is, however, an important question to address given evidence that emotional *perception* deficits can still be observed in mid-term to long-term abstainers (e.g., Fein et al., 2010; Foisey et al., 2007; Kornreich et al., 2001; Valmas et al., 2014). Thus, to address these questions, two studies were conducted. Study 1 explored acoustic cue use in emotional prosody production in a sample of long-term abstainers and healthy controls. In particular, we investigated how speakers use pitch, tempo (duration), and loudness to express six basic emotions and neutral to infer whether long-term abstainers use acoustic cues similarly to controls and

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

speakers described in the wider emotional prosody production literature (e.g., Banse and Scherer, 1996; Paulmann and Uskul, 2014). If emotional prosody cue use is not affected in long-term abstainers, we expect them to show similar acoustic cue use profiles to healthy controls and reports of speakers in the literature; however, if a history of alcohol abuse can impact on emotional prosody production abilities, altered profiles should be expected. Based on evidence reported by Monnot et al. (2003) we specifically expect to find differences between groups with regard to pitch production.

Although descriptions of acoustic parameter use are vital for exploring emotional prosody production in abstained alcoholics, they do not provide a holistic picture. In particular, we need to also assess how speech samples are perceived by naïve listeners. Can they detect which emotion abstainers are trying to express? And, do listeners judge emotional speech samples from abstainers differently to samples spoken by healthy controls? In other words, can we estimate the potential social ramifications for abstained alcoholics? As mentioned before, this part of emotional social interactions has been overlooked in the research community so far. There is, however, limited evidence that couples with one alcoholic member report more difficulties expressing emotions as well as feeling as if their emotions are not understood in contrast to non-alcoholic couples (Philippot et al., 2003). Whether this perceived difficulty can be confirmed experimentally will be tested here. Thus, in Study 2 we explore whether emotional speech produced by abstained alcoholics is recognized with a similar success rate as emotional speech produced by controls when judged by naïve listeners. Crucially, listeners are also asked how much they thought speakers actually *felt* the emotion they tried to express. Moreover, to get a more informed picture about the emotional speech produced, we also explored the role of perceived voice quality in emotional prosody production. Voice quality refers to the characteristics of produced speech and can include features such as how rough, melodic, or nasal a voice sounds. Here, we focused on two qualitatively different voice qualities and asked raters to indicate how “husky” (linked to a rough or strained sounding voice) or “flat” a voice sounds. Latter quality has been linked to abulia, or to being perceived as sounding indifferent. In short, Study 2 reports empirical data which allows exploring how emotional speech samples produced by abstainers and controls are perceived by naïve listeners. If true that abstainers have difficulties expressing emotions in speech, listeners should find it more difficult to accurately judge emotional utterances from them than those of controls. Also, if true that abstainers' speech is less emotionally expressive and of a different voice quality, we expect to find rating differences between groups. Combined, Studies 1 and 2 will thus allow describing, for the very first time, how a history of alcohol abuse can impact on emotional speech production abilities and how these effects can impact on listeners' judgements about the speakers.

STUDY 1

MATERIALS AND METHODS

Participants

Fifteen long-term abstained alcoholics and the same number of age and education matched healthy controls were recruited. Independent samples t-tests showed that abstained alcoholics and controls did not differ in age ($t(14)=.12$, p=.903) and years of education ($t(14)=1.50$, $p=.154$). Participants in the alcoholic group had a past medical diagnosis and met the DSM-IV criteria for alcohol dependence. Each abstainer had abstained from alcohol for at least one year (range 1-18.1 years). None of them reported having any other addiction in the past (full participant information can be found in Table 1). All participants were right-handed native English speakers. They were recruited via newspaper, radio adverts and leafleting in Alcohol Anonymous and other self-help groups (alcoholics only). Participants gave full informed consent before the start of the experimental session and were financially compensated for their participation. The study was approved by the Ethical Committee of the Science and Health Faculty of the University of Essex.

Assessments

We pre-screened participants for depression (Patient Health Questionnaire; PHQ-9, Kroenke et al., 2002) and anxiety (Generalized Anxiety Disorder 7-item (GAD-7), Spitzer et al., 2006). While the two groups did not differ on scores for depression ($t(14)=1.59$, $p=.134$), the scores for general anxiety disorder differed between groups ($t(14)=-3.65$, $p=.003$). Abstainers displayed higher general

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

anxiety levels than healthy controls. We did not recruit participants who self-reported use of psychotropic medication or those who reported a history of diagnosed neurological problems. We also asked participants to fill out the Revised Life Orientation (LOT-R, Herzberg et al., 2006) monitoring individuals' differences in generalized optimism versus pessimism.

- place Table 1 about here -

Procedure

All participants were tested individually. Before the start of the emotional speech recording session, all participants completed the questionnaires listed above. In the main emotional speech production task, participants were asked to intone 20 semantically neutral sentences (e.g., "*The book was green*") in one of six emotional (angry, disgust, fear, happy, sad and surprised) and a neutral tone of voice. For baseline recordings, all participants started with the neutral category. After this, participants were allowed to choose which category to express next. For each emotional category, participants were presented with written scenarios that represented a situation in which this emotion would commonly be elicited. In addition, we also asked participants to describe a time when they had felt that particular emotion in the past. It has been shown that reliving and reacting emotional situations in this kind of task lead to changes in voice patterns in speakers (e.g., Velten-Technique, 1968). No exemplars of how a specific emotion should sound were given to participants. After the emotion induction procedure, participants were presented with the list of 20 semantically neutral sentences. Each participant was asked to repeat each sentence three times in a specific emotion to ensure clear, artefact- and error-free recordings (only error- and artefact free recordings entered our statistical analysis). Therefore, each participant produced 420 utterances (6 emotions plus neutral x 20 sentences x 3 repetitions of each sentence). Sentences were recorded with Audacity, using a high-quality clip-on microphone. The recordings were digitized at a mono, 16 bit, 44,100 Hz sampling rate. Each testing session lasted approximately 40 minutes.

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

RESULTS

Acoustic data was analysed using Praat software (Boersma and Weenink, 2013). Parameters of interest were pitch (measured in semitones and calculating the interval between F0 mean and 16.35 Hz), amplitude (measured in dB) perceived as loudness, and duration (seconds) perceived as speech rate. We measured pitch on the logarithmic semitone scale as opposed to Hertz to account for potential differences between groups as they slightly differed in their male/female ratio. It has been suggested that there are no measurable differences between genders in pitch variability when expressed in semitones (Traunmüller & Eriksson, 1995; Bird, 2013). Previous findings suggest that differences between neutral and emotional prosody should be between one and five semitones (Lolli, Lewenstein, Basurto, Winnik, Loui, 2015).

Table 2 shows means and standard deviations for each extracted parameter for all emotional categories and both groups separately. To investigate whether the two groups used acoustical cues differently, we conducted several Analyses of Variance (ANOVAs) in which *speaker group* (abstainers/controls) was treated as between-subjects variable, *emotion* (anger, disgust, fear, happiness, sadness, surprise and neutral) as within-subjects variable, and each acoustic variable (pitch, duration, amplitude) served as dependent variable.

- place Table 2 about here -

Pitch

Result revealed a significant main effect of *Emotion*, ($F(6,168)$= 38.885, $p<.001$, η^2*.581*, suggesting that different emotions were expressed using different pitch as expressed in semitones. For instance, surprised was expressed using the highest mean pitch, followed by anger, happiness, fear, disgust and sadness. Neutral utterances were intoned with a lower mean pitch than all emotions (see Table 2). This main effect was qualified by a significant *Speaker Group* x *Emotion* interaction, $F(6,168)$=4.

896, *p*<.001, η^2.149, confirming that the two groups differed in how they used pitch to express specific emotions. Post-hoc pairwise comparisons revealed that healthy controls used a higher pitch when expressing fear (*p*=.024) and surprise (*p*=.030) when compared to abstainers. We also looked at the effects for each group separately and compared emotional sentence production to neutral sentence production. This analysis indicated that abstainers spoke with an increased mean pitch when expressing anger (*p*=.001), disgust (*p*=.023), happy (*p*=.001) and surprise (*p*=.001), but not when expressing fear (*p*=.129) or sadness (*p*=.627). In contrast, healthy controls expressed all emotions with higher pitch when compared to neutral sentences (all *p*s<.001) except from sadness (*p*=.597).

To confirm that pitch use differences were not due to the groups having slightly different male/female ratios, we ran the same analysis for male and female participants separately. Contrasts again confirmed that male abstainers modulated pitch differently when comparing neutral and angry sounding sentences (*p*=.001) as well as neutral and happy sounding expressions (*p*=.002). In contrast, male control participants modulated pitch differently for neutral vs anger (*p*=.015), neutral vs fear (*p*=.001), neutral vs happiness (*p*=.007) and neutral vs. surprise (p=.001). Similarly, for female abstainers, only the contrasts between neutral and happiness (*p*=.007) and neutral and surprise (*p*=.029) reached significance, while a range of emotions were uttered with a different pitch than neutral for female controls (anger (*p*=.003), disgust (*p*=.001), fear (*p*=.001), happy (*p*=.001), surprise (*p*=001)). These patterns thus confirm pitch usage differences when expressing emotions by abstained alcoholics compared to healthy controls. [2]

Pitch Variability

There was a significant main effect of *Emotion* for pitch variability (standard deviation of pitch as expressed in semitones re: 16.35Hz), *F*(6,168)=19.755, *p*<.001, η^2.414, showing a wider use of pitch when expressing surprise followed by anger and followed by disgust, happiness, fearful and neutral. Utterances intoned in a sad tone of voice showed the smallest pitch variability. There was also a significant main effect for *Speaker Group*, *F*(1,28)=5. 595, *p*=.032, η^2.153, showing that healthy controls showed more varied use of pitch than abstainers. The two main effects did not interact.[3]

132
153

Mean Amplitude

Result for mean amplitude only revealed a significant main effect of *Emotion*, $F(6,168)=50.631$, $p<.001$, $\Omega^2= .64$, showing that angry sentences were spoken in the loudest voice followed by surprise, happy, fear, disgust and neutral. Sadness was spoken more quietly than all other emotions. No main effect of *Speaker Group* ($p=.621$) or interaction between *Emotion* x *Speaker Group* (p=.084) was found.

Amplitude Range

Results revealed a different amplitude range use for different emotions, $F(6,168)=50.631$, $p<.001$, $\Omega^2=.69$. As can be seen from Table 2, angry sentences were intoned using a wider amplitude range than sad sentences. The main effect of *Speaker Group* did not reach significance $p=.093$, $\Omega^2=.10$, but looking at the amplitude range means revealed that healthy controls tended to use a slightly wider amplitude range than abstainers (34.18 dB vs 32.35 dB).

Utterance Duration

For utterance length, only a main effect of *Emotion* was found, $F(6,168)=5.583$, p<.001, $\Omega^2=.75$. Means showed that fear was spoken with a faster speech rate than disgust (1.35 seconds vs 1.51 seconds).

Leave-one-out Analysis

Following conventions from other fields that report results from relatively small sample sizes, we ran so-called jackknifing analyses to confirm that the differences in pitch use between groups were not largely driven by one individual (c.f. Paulmann et al., 2010). We thus re-ran analyses for mean pitch as well as for pitch variability 14 times, always leaving out one abstainer at the time. F- and p-values were monitored. Results for the mean semitones analyses showed that statistical findings were stable for the interaction between speaker group and emotion (all F's > 4.31) and the main effect of group (all *F*s> 1.5, all *p*s>.084). Similarly, results for the analyses looking at the variability of semitones revealed stable effects confirming that results were unlikely due to be connected to only one individual in the data set.

Influence of Anxiety on acoustic variable modulation

As shown in Table 1, a group comparison revealed that abstainers and controls differed with regard to their baseline anxiety levels. Thus, to investigate the potential influence of anxiety scores on acoustic measures Pearson's correlations were calculated for the abstainers. No significant correlations were found (all p's>.05), suggesting that anxiety levels did not impact on production of emotions.

Overall, results revealed that participants used different acoustic patterns for the different categories expressed mirroring previous results from untrained speakers (e.g., Paulmann et al., 2016). Group differences between healthy controls and abstainers were particularly apparent for pitch use. In particular, healthy controls used a higher pitch when expressing emotional as opposed to neutral prosody while the same pattern was not observed in abstainers. They failed to show a pitch increase when expressing fear and sadness. Moreover, healthy controls used a more varied pitch approach than abstainers. Finally, healthy controls also appeared to use a wider range of loudness though this effect failed to reach significance. Taken together, results showed differences between healthy controls and abstainers in modulating pitch parameters when intoning emotional sentences.

STUDY 2

Study 2 set out to explore whether sentences intoned by abstainers and healthy controls in Study 1 are perceived differently by naïve listeners. In particular, our goal was to investigate whether the emotional intention of speakers could be reliably determined. We also investigated if speakers differed with regard to voice quality attributes. In particular, we asked listeners to judge how much they felt the expressed emotion, how much they felt the speech sounded husky to them and how much it sounded inexpressive, or flat.

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

MATERIALS AND METHODS

Participants

A group of 24 (11 male & 13 female) native English speakers were recruited through campus and online advertisement. The listener group had a mean age of 28 (range 19-62) and mean number of years in education was 17 (range 13 - 27). Exclusion criteria included a history of mental health (e.g. depression), neurological problems (e.g. stroke), or a history of substance abuse all of which were measured by self-reporting. None of the participant's self-reported any biological family members who had a known history of substance abuse. The listener group self- reported normal or corrected-to-normal vision, and no hearing impairments.

Materials

To avoid bias judgements in the selection of stimuli for the recognition study, a discriminant analysis was first performed to predict emotional category membership of all stimuli collected in Study 1 (c.f. Paulmann et al., 2016 for similar approach). In this analysis, acoustical parameters (pitch, intensity, and duration) were entered as independent variables while the intended emotional category (anger, disgust, fear, happiness, pleasant surprise, sadness, and neutral) served as dependent variables. Results revealed that based on these three acoustic parameters, 29.5% of abstained alcoholics' speech samples and 36.5% of healthy controls' utterances could be classified accurately. From these correctly classified utterances we decided to present 15 sentences for each of the seven categories meaning that 210 sentences were randomly selected for Study 2. 105 sentences came from the correctly identified the healthy control group samples and 105 from the abstained alcoholics.

Procedure

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Participants were tested individually in booths at the University of Essex. Listeners were first asked to read and sign a consent form and then fill out a background questionnaire. Before the start of the study, listeners were informed of the procedure. They were told that they would be presented with spoken materials on a computer running Superlab software. Participants were instructed that they would hear utterances spoken by different speakers. Their first task was to identify the emotional category they believed the speaker was trying to convey. They were advised to answer as quickly and accurately as possible. On-screen categories were labelled as “angry”, “disgust”, “fear”, “happy”, “sad”, “surprise”, and “neutral”. Their second task was to make three assessments about the utterance: First, they were asked to indicate on a scale from 1 (not at all) to 7 (very much) how much they thought the speaker sounded as if he/she really felt the emotion, how much they felt the speaker sounded flat (explained as inexpressive), and how rough/husky the speaker sounded. A trial sequence was thus as follows: a fixation cross was presented for 200ms followed by the presentation of the utterance, followed by a seven box response screen. After participants provided their emotional assessment, they were presented with the three rating scale screens, which also contained the question at hand. A blank screen was presented for 500 ms as an inter-stimulus interval. After five practice trials, participants had the chance to ask the experimenter for help. The main experiment contained a total of 210 utterances which was divided into seven blocks that consisted of 30 trials each. Each block was followed by a short break. Testing time lasted around one hour and listeners were compensated £6 for their time.

RESULTS

Statistical analysis

The statistical package SPSS (version 21) was used to analyze the data. To investigate whether utterances from controls were better recognized than those from abstainers, we conducted a 2 (*speaker group*) x 7 (*emotion*) within-subjects ANOVA for which listeners' emotion recognition scores served as dependent variable. Rating of voice quality indicators were analyzed with separate within-subjects ANOVAs. All responses were averaged for each participant and emotion before carrying out the analyses. Effect size was measured using omega-square (Ω^2). According to Olejnik and Algina (2003) and treated effect size values between 0.0009 – 0.048 as small, values between 0.048 and 0.138 as medium, and values above 0.138 as large.

Emotion recognition accuracy

Figure 1 shows mean (and standard deviations, SD) recognition accuracy rates of utterances intoned by abstainers and healthy control speakers for each emotional category separately. Utterances expressed by healthy controls resulted in higher recognition rates for all categories. This was confirmed by the statistical analysis which revealed a main effect of *speaker group*, $F(1,23)=63.838$, $p<.001$, $\Omega^{2=}$.74., showing that listeners were more accurate at identifying emotions spoken by healthy controls as opposed to abstainers (42% v 31%). There was also a significant main effect of *emotion,* F(6,138)=31.242, P<.001, $\Omega^{2=}$.58. Neutral prosody was best recognised (53%), followed by utterances intended to express pleasant surprise (52%), sadness (49%), angry (44%), disgust (22%), fear (20%) and happiness (15%). Post-hoc comparisons revealed a significant difference between recognition rates for neutral utterances and utterances spoken in a disgusted, fearful and happy tone of voice (all ps<.001). A significant two-way interaction between *speaker* and *emotion* was also found $F(6,138)=13.323$, P<.001, $\Omega^2=.37$. Looking at each emotion separately, results revealed that listeners were significantly better at identifying utterances expressed in an angry (p<.001), fearful (p<.001) and surprised (p<.001) tone of voice when spoken by healthy controls compared to abstained alcoholics.

\- Place Figure 1 about here –
\-

Voice quality: Emotional Expressiveness

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Results showed a significant main effect of *speaker group*, *F*(1,23)=71.143, *P*<.001, Ω^2=.77. Listener's perceived healthy controls' utterances as more emotionally expressive than abstained alcoholics (4.22 v 3.84). A significant main effect of *emotion*, *F*(6,138)=23.877, *P*<.001, Ω^2=.51, showed that listeners perceived utterances spoken in a surprised tone of voice (4.69) as most expressive and neutral (3.51) utterances were rated as least expressive. Post-hoc comparisons revealed a significant difference between neutral utterances and all other emotional utterances in terms of how much the listeners thought the speaker felt the emotion (all *p*s<.01). Results also revealed a significant *emotion* x *speaker* interaction, *F(6,138)=6.975, p<.001,* $\Omega^{2=}$.03, showing that utterances expressing anger, disgust, fear, happy or surprised prosody by controls were perceived as sounding more "felt" than the same emotions expressed by abstainers (*p*<.001).

Voice quality: Huskiness

The ANOVA revealed a significant main effect of *speaker group, F*(1,23)=8.095, *p*=.009, $\Omega^{2=}$.26. Listeners rated utterances spoken by abstainers as sounding rougher than utterances spoken by healthy controls (3.00 vs. 2.80). There was also a significant main effect of *emotion*, *F*(6,138)=9.673, *p*<.00, Ω^2=.30. Listeners rated sad utterances as sounding most rough or husky (3.24) and surprise utterances as sounding the least rough (2.47). Post-hoc comparisons revealed that rating scores for fear (p=.016), happy (p=.001), sad (p=.033) and surprise (p=.001) sentences differed significantly from rating scores for neutral utterances. There was also a significant two-way interaction between *speaker group* x *emotion, F(6,138)=* 2.231, *p*=.044, $\Omega^{2=}$.09. Post-hoc comparisons by emotion revealed that sentences intoned in angry and neutral tone of voice by abstainers were rated as sounding significantly huskier than those uttered by healthy controls (p<.05).

Voice quality: Flatness

The analysis revealed a significant main effect of *speaker group*, *F(1,23)=75.362* , *p*=.001, $\Omega^{2=}$.77. Abstainers' utterances were rated as sounding more flat than those spoken by controls (4.00

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

vs. 3.52). A significant main effect of *emotion* also emerged, $F(6,138)=32.956$, $p<.001$, $\Omega^{2}=.59$. Sad utterances were rated as sounding most flat (4.79), while surprised sounding sentences were rates as sounding least flat (2.80). Planned pairwise comparisons between neutral and emotional utterances showed that all emotions were rated as sounding less flat in comparison to neutrally intoned utterances (all *p*s<.01). The *speaker* x *emotion* interaction was also significant, $F(6,138)=7.771$, $p<.001$, $\Omega^{2}=.25$. Post-hoc comparisons revealed that sentences intoned in an angry, disgust, fearful, neutral or surprised tone of voice by abstainers were rated as significantly more flat than utterances intoned by healthy controls ($p<.05$).

Overall, results of Study 2 showed that listeners blind to the group manipulation assessed randomly selected emotional speech exemplars as sounding significantly different. In particular, we found that naïve listeners found it harder to accurately recognize the intended emotions when uttered by abstainers in comparison to those intoned by healthy controls. Listeners also perceived exemplars spoken by abstainers to sound less emotionally expressive, more flat and rougher sounding than speech produced by healthy controls.

GENERAL DISCUSSION

The present investigation explored emotional vocal expressions in long-term abstinent alcoholics. In Study 1, it was shown that abstinent alcoholics control mean and variability of pitch differently than healthy controls when communicating emotions through tone of voice. In Study 2, it was shown that naïve listeners judged randomly selected samples spoken by abstainers as sounding less emotionally expressive than samples produced by controls. Crucially, the emotional intentions of abstainers were also more difficult to recognize. Taken together, these results suggest that emotional prosody production problems associated with alcoholism can persist even after individuals have (long) stopped drinking.

Emotional Prosody Production Differences

The data reported here uniquely lend empirical support to the notion that a history of alcohol abuse can have long term effects on emotional tone of voice production. The most prominent difference between long-term abstinent alcoholics and the control group was the way that mean and variability of pitch was modulated when trying to express an emotion. Abstainers did not increase pitch when expressing fear or sadness; moreover,, the results also confirmed that controls generally used a more varied pitch than abstinent alcoholics. The adequate modulation of pitch has repeatedly been shown to play a vital role in communicating emotions through speech (Frick, 1985; Monnot et al., 2003; Scherer, 2003; Scherer et al., 1972). In fact, low or monotonic pitch has been linked to depressive speech, suggesting lacking affect (e.g., Moore et al., 2004). The results here suggest that although abstinent alcoholics alter their pitch when expressing emotions, they do so less effectively than controls. Thus, our data provide evidence that dry alcoholics' pitch production differs from "normal" usage, suggesting a limited ability to express emotional prosody in these individuals. This is in line with results reported for recently detoxified alcoholics (Monnot et al., 2003). Several accounts may explain this production difference: First, it has been shown that alcoholism can lead to severe right hemisphere brain changes (see Oscar-Berman & Marinkovic, 2003, for review). Interestingly, pitch-related processes have repeatedly been linked to right hemisphere brain structures (e.g., Sidtis and Van Lancker Sidtis, 2003) and lack of pitch control has been reported for patients with right hemisphere brain lesions (Ross & Monnot, 2008; Shapiro & Danly, 1985). Similarly, alcohol-related brain changes have also been linked to the frontal lobes, limbic system, and the cerebellum (Oscar-Berman & Marinkovic, 2003), often seen as key players in an emotional prosody network (c.f. Kotz & Paulmann, 2011). Thus, it can be speculated that alcohol-related brain changes contribute to the effects observed here. Moreover, the role of the cerebellum has been tied to motor co-ordination and control over vocal tract muscles involved in pitch production in particular (Ackermann, Mathiak, Riecker, 2007). Interestingly, cerebellar dysfunctions have additionally been shown to lead to harsh sounding voice quality (Darley, Aronson, Brown, 1975), a phenomenon observed here, too. Finally, problems in expressing emotional prosody might also be linked to physical alterations of the vocal apparatus caused by heavy drinking. For instance, alcohol consumption can lead to inflammation of laryngeal mucosa which can affect vocal fold vibration patterns. This alteration may influence both pitch production as well as voice quality (e.g., making the voice sound harsh; c.f. Kreiman & Sidtis, 2013). Similarly, some research suggests a strong link between smoking and alcoholism (e.g., Difranza and Gurrera, 1990) and voice production mechanisms are altered by smoking (e.g., Aronson and Bless, 2009). Future studies should thus aim to control for smoking history of participants. It is beyond the scope of the present investigation to pinpoint the underlying mechanisms of the pitch production differences between alcoholics and controls but the accounts summarized here merit testing in future studies.

Perception of Emotional Prosody

The first part of this investigation suggested that abstinent alcoholics can fail to properly control and execute their vocal apparatus leading to fluctuations in pitch use. Timing and loudness control was not affected as prominently. While differences in production are meaningful to explore in their own right, the more pressing question is whether the inability to use pitch adequately could actually lead to difficulties in listeners recognising the intended emotion. Arguably, not controlling and modulating pitch cues appropriately could lead to production of less “stereo-typical” emotion exemplars; in other words, making it more difficult for listeners to gauge the emotional intention. This was directly tested in Study 2.

Study 2 used the materials produced in Study 1. Acoustic analyses of these materials confirmed that different emotional expressions were characterized by varying acoustic profiles (c.f. Table 1) which for the most part mirrored those observed in previous studies using acted speech (e.g., Banse & Scherer, 1996; Paulmann & Uskul, 2014). Not surprisingly, recognition rates for emotional exemplars obtained here were largely lower than recognition rates obtained for materials intoned by actors (e.g., Banse and Scherer, 1996), but they were still above chance level (14%) and resembled recognition rates reported for materials spoken by untrained speakers (e.g., Paulmann et al., 2016).

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Exemplars were initially selected based on a discriminant analysis and only materials that were correctly identified by this analysis were used in Study 2. Still, results suggest that naïve listeners found it generally more difficult to decode emotions from abstainers' speech compared to utterances produced by controls. In particular, results suggest that emotional utterances expressing anger, fear, or surprise were most difficult to recognize when intoned by abstainers. Generally speaking, these emotions are also those expressed with *higher* pitch than neutral expressions. Thus, combined results suggest that inadequate use of pitch when expressing emotions in speech may lead to a failure in the listener to detect the intended emotion. Clearly, a difficulty in deciphering what a speaker is trying to express can potentially lead to social misunderstandings or possibly interaction breakdowns.

Next to finding it more difficult to judge the emotionality of speech produced by abstainers when compared to controls, listeners also judged speech samples differently on a variety of dimensions linked to the perception of voice quality. In particular, abstainers' utterances were rated as sounding huskier, more flat and, crucially, less emotionally expressive. Latter finding, that is the fact that abstainers speech was perceived as less emotionally expressive might again be linked to the differences in pitch (and possibly intensity) variability modulations observed in Study 1. It also directly links with the result that abstainers' emotional speech is more difficult to recognize. As discussed above, several explanations to account for voice quality differences seem plausible; however, cerebellar dysfunctions as well as changes of the mucosa lining the larynx seem to be among the most likely candidates at this point. Taken together, the present findings, for the first time, highlight how a history of alcohol abuse can affect emotional tone of voice production in the long-term. We also showed that the expressive differences between abstainers and controls has effects on naive listeners, leading to lower recognition rates, lower emotional expressiveness scores and higher ratings of harshness and flatness of the voice.

Future Directions

To the best of our knowledge, this is the first investigation exploring the *long-term* effects of alcohol abuse on communicating emotions through the tone of voice. An inability to express emotions vocally

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

can have severe impacts on social interactions. Knowing more about which factors contribute to abstainers' problems in conveying emotions non-verbally can potentially help to develop strategies that target how emotional tone of voice use can be improved in affected individuals. Here, we explored acoustic parameters which have long been known to play a prominent role in successful emotional prosody production. Analyses revealed that abstainers and controls differed with regard to their pitch use, while durational parameters (speech rate) seemed to be unaffected. A more detailed picture of which other parameters (e.g., frequency bands) are used differently will lead to a broader understanding of why emotional speech of abstainers lacks emotional expressiveness and is considered to be more difficult to recognize than speech by controls.

For therapeutic purposes, it will be important to explore whether observed pitch use differences stem from an inability to fully control the vocal apparatus (e.g., caused by brain damage to areas linked to motor control and/or emotional prosody processes), or through damage to the vocal folds or muscles surrounding them (Aronson and Bless, 2009). Ideally this will include a combination of neuroimaging and vocal production techniques that allow studying the mechanisms underlying emotional prosody production difficulties in alcoholics more systematically.

Finally, the current study tested eight female and seven male speakers who had abstained from drinking alcohol for at least one year. Future studies should try to determine in how far gender and length of abstaining can play a moderating role in emotional speech communication by testing larger sample sizes and including abstaining length as a co-variate in the analysis.

Conclusion

The ability to communicate emotions through voice is an important and necessary aspect of social relationships. In fact, prosody has been self-reported as the most common method of distinguishing emotions in real-life situations (Planalp, 1998). Knowing more about the long-term effects of alcohol abuse in emotional prosody production is thus crucial for abstainers to help with their interpersonal communication. If abstinent alcoholics and those with no alcohol abuse history differ in the way they express emotions in speech, it may be necessary to create social skills training programs that help mitigate conflicts between different parties before they blow out of proportion. The current investigation provides a first step in trying to understand how abstainers' differ in emotional tone of voice production and the effect that this has on listeners. Clearly, future work is needed to fully unravel the underlying mechanisms of this usage difference.

Reference List

Ackermann, H., Mathiak, K., & Riecker, A. (2007). The contribution of the cerebellum to speech production and speech perception: clinical and functional imaging data. *The Cerebellum*, *6*(3), 202-213.

Aronson, A. E., Bless, D.M. (2009). "Voice Disorders of Structural Origin." *Clinical Voice Disorders. 4th ed. New York: Thieme Med Pub,* 24-38.

Banse, R. and Scherer, K.R. 1996. Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology, 70*, 614-636.

Boersma P, Weenink D. (2012). Praat: Doing Phonetics by Computer (computer program), version 5.5.

Bird, C. (2013). The prosody of riddle openings. In: Attardo, S., Wagner, M. M., & Urios-Aparisi, E. (Eds.). (2013). *Prosody and humor* (Vol. 55). John Benjamins Publishing.

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Audio seminars in speech pathology: Motor speech disorders*. Philadelphia: WB Saunders.

DiFranza, J. R., & Guerrera, M. P. (1990). Alcoholism and smoking. *Journal of Studies on Alcohol*, *51*(2), 130-135.

Fein, G., Key, K. and Szymanski, M.D. (2010). ERP and RT Delays in Long-Term Abstinent Alcoholics in Processing of Emotional Facial Expressions During Gender and Emotion Categorization Tasks. *Alcoholism: Clinical and Experimental Research, 34*, 1127-1139.

Foisy, M.L., Kornreich, C., Fobe, A., D'Hondt, L., Pelc, I., Hanak, C., Verbanck, P. and Philippot, P. (2007). Impaired emotional facial expression recognition in alcohol dependence: do these deficits persist with midterm abstinence? *Alcoholism: Clinical and Experimental Research, 31*, 404-410.

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Frick, R.W., (1985). Communicating emotion: The role of prosodic features. *Psychological Bulletin, 97,* 412-429.

Frigerio, E., Burt, D.M., Montagne, B., Murray, L.K. and Perrett, D.I. (2002). Facial affect perception in alcoholics. *Psychiatry Research, 113,*161-171.

Herzberg, P.Y., Glaesmer, H. and Hoyer, J., 2006. Separating optimism and pessimism: a robust psychometric analysis of the revised Life Orientation Test (LOT-R). *Psychological assessment*, *18*, 433-438.

Kornreich, C., Blairy, S., Philippot, P., Hess, U., Noël, X., Streel, E., Le Bon, O., Dan, B., Pelc, I., Verbanck, P. (2001). Deficits in recognition of emotional facial expression are still present in alcoholics after mid-to long-term abstinence. *Journal of studies on alcohol, 62,* 533-542.

Kotz, S. A., & Paulmann, S. (2011). Emotion, language, and the brain. *Language and Linguistics Compass*, *5*(3), 108-125.

Kreiman, J., & Sidtis, D. (2011). *Foundations of voice studies: An interdisciplinary approach to voice production and perception*. John Wiley & Sons.

Kroenke, K. and Spitzer, R.L. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric annals, 32*, 509-515.

Lolli, S., Lewenstein, A. D., Basurto, J., Winnik, S., & Loui, P. (2015). Sound frequency affects speech emotion perception: Results from congenital amusia. *Frontiers in Psychology*, *6*, 1340.

Maurage, P., Campanella, S., Philippot, P., Charest, I., Martin, S. and de Timary, P. (2009). Impaired emotional facial expression decoding in alcoholism is also present for emotional prosody and body postures. *Alcohol and alcoholism, 44*, 476-485.

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Monnot, M., Nixon, S., Lovallo, W. and Ross, E. (2001). Altered emotional perception in alcoholics: deficits in affective prosody comprehension. *Alcoholism: Clinical and Experimental Research*, *25*, 362-369.

Monnot, M., Orbelo, D., Riccardo, L., Sikka, S., & Ross, E. (2003). Acoustic analyses support subjective judgments of vocal emotion. *Annals of the New York Academy of Sciences*, *1000*(1), 288-292.

Moore, E. I. I., Clements, M., Peifer, J., & Weisser, L. (2004, September). Comparing objective feature statistics of speech for classifying clinical depression. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE* (Vol. 1, pp. 17-20). IEEE.

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychological methods*, *8*(4), 434-447.

Oscar-Berman, M., Hancock, M., Mildworf, B., Hutner, N. and Weber, D.A. (1990). Emotional perception and memory in alcoholism and aging. Alcoholism: Clinical and Experimental Research, 14(3), pp.383-393.

Oscar-Berman, M. and Marinkovic, K., (2003). Alcoholism and the brain: an overview. *Alcohol Research and Health*, *27*, 125-133.

Paulmann, S., Seifert, S., & Kotz, S. A. (2010). Orbito-frontal lesions cause impairment during late but not early emotional prosodic processing. *Social Neuroscience*, *5*(1), 59-75.

Paulmann, S. and Uskul, A.K. (2014). Cross-cultural emotional prosody recognition: Evidence from Chinese and British listeners. *Cognition & Emotion*, *28*, 230-244.

Paulmann, S., Furnes, D., Bøkenes, A. M., & Cozzolino, P. J. (2016). How Psychological Stress Affects Emotional Prosody. *PloS one*, *11*(11), e0165022.

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Philippot, P., Kornreich, C., Blairy, S., Baert, I., Dulk, A.D., Bon, O.L., Streel, E., Hess, U., Pelc, I. and Verbanck, P. (1999). Alcoholics' deficits in the decoding of emotional facial expression. *Alcoholism: Clinical and Experimental Research*, *23*, 1031-1038.

Planalp, S. (1998). Communicating emotion in everyday life: Cues, channels, and processes. *Handbook of communication and emotion: Research, theory, applications, and contexts*, pp. 29-48.

Ross, E. D., & Monnot, M. (2008). Neurology of affective prosody and its functional–anatomic organization in right hemisphere. *Brain and language*, *104*(1), 51-74.

Scherer, K.R., Koivumaki, J. and Rosenthal, R. (1972). Minimal cues in the vocal communication of affect: Judging emotions from content-masked speech. *Journal of Psycholinguistic Research*, 1, 269-285.

Scherer, K.R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech communication*, *40*, 227-256.

Spitzer, R.L., Kroenke, K., Williams, J.B. and Löwe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*, *166*,1092-1097.

Shapiro, B. E., & Danly, M. (1985). The role of the right hemisphere in the control of speech prosody in propositional and affective contexts. *Brain and language*, *25*(1), 19-36.

Sidtis, J. J., & Sidtis, D. V. L. (2003). A neurobehavioral approach to dysprosody. In *Seminars in speech and language* (Vol. 24, No. 02, pp. 093-106). Copyright© 2003 by Thieme Medical Publishers, Inc., New York, NY, USA.

Traunmüller, H., & Eriksson, A. (1995). The frequency range of the voice fundamental in the speech of male and female adults. *Consulté Le*, 12(2).

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Uekermann, J., Daum, I., Schlebusch, P. and Trenckmann, U.(2005). Processing of affective stimuli in alcoholism. *Cortex, 41*, 189-194.

Valmas, M.M., Mosher Ruiz, S., Gansler, D.A., Sawyer, K.S. and Oscar-Berman, M., 2014. Social cognition deficits and associations with drinking history in alcoholic men and women. *Alcoholism: Clinical and Experimental Research, 38*, 2998-3007.

Velten, E. (1968). A laboratory task for induction of mood states. *Behaviour research and therapy*, *6*(4), 473-482.

NOTE: THIS VERSION MAY NOT BE IDENTICAL TO THE PUBLISHED VERSION

Figure Legend

Figure 1: Accuracy (%) of mean emotional recognition responses for each speaker group. Bars show correct responses for each emotional category (error bars represent standard deviations).

Table 1: Demographic and patient information for participants (mean, SD)

Variable	Abstained Alcoholics	Healthy Controls
Sex (F/M)	5/10	8/7
Age NS	51.87 (12.98)	51.27 (13.32)
Age Range	33 to 76	35 to 76
Education (in years)	13.91 (3.42)	15.8 (3.56)
Disease duration (in years)	13.7 (7.55)	N/A
Disease duration range (in years)	5 to 27	N/A
Abstinence duration (in years)	9 (9.10)	N/A
Abstinence range (in years)	1 to 18.1	N/A
Number of alcoholic drinks per week	N/A	2.33 (3.2)
GAD-7 **	6.73 (4.53)**	2.6 (3.6)**
PHQ-9 NS	4.93 (3.61)	3.07 (2.66)
LOT-R NS	13.33 (5.01)	15.07 (4.25)

*N/A = not applicable; **= difference between was significant at p<.05; Scores 0-5 for the GAD-7 represent mild anxiety, 6-10 moderate, 11-15 moderately severe anxiety, 16-21 severe anxiety. PHQ-9 scores from 0-5 represents mild depression, 6-10 moderate depression, 11-15 moderately severe depression, 16-21 severe depression. A score of over 7 on the GAD-7 represents clinical anxiety and over 9 on the PHQ-9 clinical depression. For the LOT-R higher scores represent higher optimism. The number of years of education for each group was worked out from the number of completed years in education from primary school.*

Table 2. *Means (SD) for each acoustic variable displayed per group. Originally, pitch was measured in Hertz and then converted using praat's function "convert Hz to semitones" using the formula (12*log2(F0mean/16.35). Duration was measured in seconds and amplitude in decibel.*

Group	**Emotion**	**Log F0 (SD)**	**Pitch range variability (SD)**	**Mean amplitude (SD)**	**Amplitude range (SD)**	**Utterance duration**
AA	**Anger**	839.33 68 (3.83)	311.97 01 (34.382)	68.55 (1.38)	34.27 (1.05)	1.46 (.05)
	Disgust	6.6538.00 (4.19)	2.469.73 (5.275)	61.68 (1.12)	33.10 (.85)	1.46 (.07)
	Fear	6.4537.80 (4.49)	-3.633.43 (7.667.62)	62.28 (1.34)	31.18 (.94)	1.35 (.04)
	Happiness	8.0339.38 (5.10)	2.7910.65 (6.166.56)	64.40 (1.09)	33.25 (.89)	1.46 (.05)
	Neutral	5.4036.75 (4.03)	-2.094.85 (8.629.11)	59.40 (.88)	31.04 (.82)	1.37 (.05)
	Sadness	5.2236.58 (4.03)	-2.304.67 (8.5517)	57.98 (.96)	30.66 (.95)	1.46 (.04)
	Surprise	9.4340.78 (5.27)	1312.6947 (3.804.98)	65.71 (1.27)	32.91 (1.05)	1.38 (.06)
HC	**Anger**	9.89 41.36 (5.52)	714.20 67 (5.66.84)	66.49 (1.38)	36.67 (1.05)	1.47 (.05)
	Disgust	9.0040.36 (4.95)	7.8115.24 (8.2342)	60.83 (1.12)	36.35 (.85)	1.55 (.07)
	Fear	10.7642.21 (5.8060)	1.379.44 (7.067.33)	64.26 (1.33)	32.58 (.94)	1.34 (.04)
	Happiness	39.5638 (5.3214)	4.4212.25 (6.7063)	62.66 (1.09)	34.37 (.89)	1.47 (.05)
	Neutral	6.2637.97 (3.99)	0.457.80 (6.8665)	58.09 (.88)	33.47 (.82)	1.47 (.05)
	Sadness	36.528.16 (4.47)	1.227.76 (6.907.14)	56.83 (.96)	31.94 (.95)	1.45 (.04)
	Surprise	1438.16.17 (4.47)	16.1220.11 (26.6463)	66.15 (1.27)	33.90 (1.05)	1.43 (.06)