How context affects perception: judging distress & linguistic content in forensic audio recordings

Lisa Roberts, Peter French
Department of Language & Linguistic Science, University of York, UK.
J. P. French Associates, York, UK.

Abstract
In order to assess the role of context on the perception of speech sounds, extracts of speech material from authentic forensic cases were played to two groups of listeners (experienced vs inexperienced forensic phoneticians) in two conditions (with and without sequential context). Listeners were asked to categorise each extract according to a four-way scale of perceptible distress and rate each extract according to a five-point scale of perceptible linguistic content. Findings show that extracts played without context are perceived as reflecting higher degrees of distress and lower levels of linguistic content. Experienced phoneticians performed more consistently as a group and were less prone to changing their responses across the two conditions.

Key words: perception, distress, forensic, context, linguistic content.

Introduction
Forensic phoneticians are often asked to transcribe and interpret recordings that are brief and of poor quality. To avoid preconceptions entering into the transcription, a ‘bottom up’ approach is encouraged (Fraser, 2003). This entails the phonetician undertaking the task ‘blind’ in the first instance, i.e., without context information and working only from the recorded sound. Some phoneticians working in this way have considered their task to have been impeded by the lack of context information (Hirson & Howard, 1994). Others, however, have written reports that may give the impression that that the significance or linguistic content of an unclear utterance was resolved mainly or wholly from examining its internal phonetic and acoustic properties (Rose 2009). Furthermore, forensic phoneticians may be asked to make assessments of the psychological states of victims in recordings. Such evaluations are currently not advocated by the International Association of Forensic Phonetics and Acoustics (IAFPA), most likely owing to a lack of sufficient research evidence to inform such an opinion (IAFPA Code of Practice, clause 9). This study examines the influence of context (sequential context together with background information) on forensic phoneticians’ perceptions of the sounds uttered by victims undergoing violent attacks.
Experiment

Stimuli
Thirty-two stimulus utterances containing productions from victims experiencing violent attack were extracted using Sony Sound Forge from eight authentic forensic recordings. All victims originated from England/Wales and spoke English as their native language.

Subjects
Sixteen members of the Forensics Research Group in the Department of Language and Linguistic Science, University of York, participated in the experiment. Six participants (four male, two female) had extensive forensic casework experience and had all acted as expert witnesses (the ‘experienced group’); ten participants (three male, seven female) were postgraduate student forensic phoneticians with no casework experience (the ‘inexperienced group’). The inexperienced group was aged between 22 years and 25 years, whereas the ages of the experienced group ranged from 27 years to 57 years.

Method
Subjects were presented with brief stimulus sounds of which twenty four (the experimental trials) were played under two conditions: (i) without context - the sound was heard in isolation; (ii) with context – the sound was heard with preceding and following phonetic context, and background information concerning the circumstances of the recorded attack was provided. A further eight extracts acted as controls and were played devoid of context. They were repeated towards the end of the experiment to check subjects’ behaviour did not change with continued exposure to the experiment. The experiment was delivered via PowerPoint Presentation 2007 and audio files were played through closed cup Sanako Tandberg Educational headphones (headset model SLH-07). Each listener was asked to: (a) categorise each excerpt in terms of a 4-way classification corresponding to the degree of distress they perceived the stimulus as representing; (b) rate the excerpt on a 5-point scale specifying the degree to which they perceived the stimulus material as having linguistic content. Two experiments were run, varying the order of presentation, in control for a possible order effect.

Findings
This section reports only on the experimental trials.\(^1\) ANOVA showed a significant effect with respect to both the distress categorisations and the linguistic content ratings for context and experience, with context exercising the stronger effect. The distress categorisations showed: for context:
F(1,707) = 6.29, p < .012; and experience: F(1,707) = 4.58, p < .033). The linguistic content ratings showed: for context: F(1,714) = 11.94, p < .001; for experience: F(1,714) = 5.24, p < .022). The interaction between experience and context was not significant for either of the two scales. Owing to the small and unbalanced number of participants in each group; age, sex and native language were not included in the statistical analysis. The effect of order, i.e. whether the extract was first heard with or without context, was not significant.

As illustrated in the pie charts below, where participants altered their responses across the two conditions, extracts were categorised as reflecting lower degrees of distress and rated as having higher degrees of linguistic content when heard with context. This pattern is evident for both groups of forensic phoneticians, although the experienced group was less prone to changing responses across the different conditions.

Table 1: Pie charts showing the change in direction of listener responses from ‘without context’ trials to ‘with context’ trials.

<table>
<thead>
<tr>
<th>Distress Category</th>
<th>Experienced Group</th>
<th>Inexperienced Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>No change</td>
<td>5%</td>
<td>10%</td>
</tr>
<tr>
<td>Less distressed</td>
<td>22%</td>
<td>24%</td>
</tr>
<tr>
<td>More distressed</td>
<td>73%</td>
<td>66%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Linguistic Content</th>
<th>Experienced Group</th>
<th>Inexperienced Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>No change</td>
<td>5%</td>
<td>11%</td>
</tr>
<tr>
<td>More content</td>
<td>28%</td>
<td>34%</td>
</tr>
<tr>
<td>Less content</td>
<td>64%</td>
<td>55%</td>
</tr>
</tbody>
</table>

**Conclusion**

These findings highlight the fact that the attribution of linguistic content to a victim’s brief utterance, scream or vocalisation is unlikely to be the product of a consideration of the internal properties of the sound itself. Higher order information – including sequential context and background story – may play a pivotal part in listeners’ decisions that a person is attempting to
communicate linguistic information. Similarly, the impression of distress is only partly conveyed by the sound itself; contextual information is again an important factor. The resources used by forensic phoneticians in interpreting difficult recordings may share more with those of lay-listeners (c.f., Jefferson 1978) than they might be aware.

Notes
1. ANOVA did not show a significant effect for changes in categorisation and rating of extracts from participants’ control trials, thus the possibility that changes in participants’ responses were due to increased exposure of the experiment is discounted.

Acknowledgements
We are very grateful to Paul Foulkes, Tamar Keren-Portnoy, and Eytan Zweig for their assistance with this research, and especially the Forensics Research Group for their participation in the experiment.

References