Profiling the anonymous authors of malicious forensic texts

Dr Andrea Nini
andrea.nini@manchester.ac.uk
www.andreanini.com

13th Biennial Conference of the International Association of Forensic Linguists
University of Porto
Any text can be a forensic text

Legal genres:
courtroom interaction
witness examination
interrogation
will

Evidence in a case:
text messages
books
email
suicide notes

_Malicious forensic text_ (MFT): a text that is a piece of evidence in a forensic case and that also commits a crime.

threatening letter
ransom demand
extortion note
hate/abusive letter
stalking message
etc

Corpus of fabricated texts

95 subjects

Three 300 word samples:
- Letter of complaint to travel agency
- Letter to Prime Minister about crisis
- Threatening letter to boss

Addressor-addressee personal knowledge

Corpus of authentic texts

Gender (biological)

Age

Social class

\[
\text{Index} = \frac{(\text{Occupation}_{\text{father}} + \text{Occupation}_{\text{mother}})}{2} + \text{Occupation}_{\text{subject}}
\]

About 130 features from several disciplines grouped in general linguistic patterns

- **Gender**
  - (Deep) Formality

- **Age**
  - (Deep) Formality

- **Social class**
  - (Deep) Formality
About 130 features from several disciplines grouped in general linguistic patterns

Gender

(Deep) Formality

Negative/ positive emotions

Age

Social class
### Logistic regression model

Gender = M or F?

<table>
<thead>
<tr>
<th>Task</th>
<th>Letter of complaint</th>
<th>Letter to Prime Minister</th>
<th>Threat to employer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Task 1</td>
<td>✔️ Deep Formality</td>
<td>✗ Deep Formality</td>
<td>✗ Deep Formality</td>
</tr>
<tr>
<td>Task 2</td>
<td>✔️ Deep Formality</td>
<td>✗ Deep Formality</td>
<td>✗ Deep Formality</td>
</tr>
<tr>
<td>Task 3</td>
<td>✔️ Deep Formality</td>
<td>✗ Deep Formality</td>
<td>✗ Deep Formality</td>
</tr>
<tr>
<td></td>
<td>✔️ Swear words</td>
<td>✔️ Swear words</td>
<td>✔️ Swear words</td>
</tr>
<tr>
<td></td>
<td>✔️ Positive emotion words</td>
<td>✗ Positive emotion words</td>
<td>✗ Positive emotion words</td>
</tr>
</tbody>
</table>

57%  

71%

Your behaviour has **pissed** many of us off and we're going to **fuck** your car up if you don't stop. The way you make me feel is like **shit** and I will not take your bullshit anymore.

This used to be a place we were **happy** to come to every day, now it is somewhere we dread. **Treat** us with the **respect** and **care** we deserve, or the whole town will see what insults we put up with displayed all over your **pride** and **joy**.
About 130 features from several disciplines grouped in general linguistic patterns

- **Gender**
  - (Deep) Formality
  - Expletives and powerless register

- **Age**
  - (Deep) Formality
  - Emotionality
  - Syntactic complexity

- **Social class**
Better writers or less memory?

On our arrival to the airport, which was delayed by an hour <because the coach to the airport <you'd provided> was unapologetically late>, I was appalled to discover that, <having overbooked the business economy seats on the plane>, my family and I had the choice of <waiting 22 hours for the next flight> or <to take seats in economy class>.

I require full reimbursement, within seven days, of the £1500 paid for the holiday, plus €1000 for the expenses detailed above, plus £1000 compensation for inconvenience and distress.
About 130 features from several disciplines grouped in general linguistic patterns

Gender
- (Deep) Formality
- Expletives and powerless register

Age
- (Deep) Formality
- Syntactic complexity
- Change in world view

Social class
- (Deep) Formality
- Vocabulary size
- Syntactic complexity
Figure 6-4. Distribution of subordination and left-branching indices by social class for 10 Montreal subjects.

Logistic regression model
Class = Working or Middle?

Addressor-addressee personal knowledge

Task 1
Letter of complaint

- ✔ Deep Formality
- ✗ Advanced Guiraud 1000
- ✔ Relative clauses

78%

Task 2
Letter to Prime Minister

- ✗ Deep Formality
- ✔ Advanced Guiraud 1000
- ✗ Relative clauses

74%

Task 3
Threat to employer

- ✗ Deep Formality
- ✗ Advanced Guiraud 1000
- ✗ Relative clauses

NS

and I really think you could do a lot more to help. I really do think that it is in your power to represent us and get this country (England) back on its feet as the job industry is going and stop all the work going to other Countries this country used to be a nation of working sorry hard working people.

I can assume you receive many letters about how the recent decline has affected the lives of many fellow voters. But as a lifelong loyal Conservative I have always held the thought that my interests would be at the heart of the party's fundamental belief that sectors such as local policing should be held in high regard.
Register defines the space of variation

All the effects are in the predicted directions

<table>
<thead>
<tr>
<th>Personal Knowledge</th>
<th>Class (%)</th>
<th>Age (%)</th>
<th>Gender (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High personal knowledge</td>
<td>69%</td>
<td>71%</td>
<td></td>
</tr>
<tr>
<td>Some personal knowledge</td>
<td>69%</td>
<td></td>
<td>77%</td>
</tr>
<tr>
<td>No personal knowledge</td>
<td></td>
<td>68%</td>
<td>57%</td>
</tr>
</tbody>
</table>

The Register Axiom

If a linguistic feature is distributed across social groups and communicative situations or registers, then the social groups with greater access to the situations and registers in which the feature occurs more frequently will exhibit more frequent use of this feature in their social dialects.

Profiling the anonymous authors of malicious forensic texts

Dr Andrea Nini
andrea.nini@manchester.ac.uk
www.andreanini.com

13th Biennial Conference of the International Association of Forensic Linguists
University of Porto