Gender Differences in Emotion Language
A Sociolinguistic Analysis of Emotives in Irish English

Martin Schweinberger
Universität Kassel, Institut für Anglistik und Amerikanistik
martin.schweinberger@uni-kassel.de

Introduction
This study investigates systematic differences in emotional language use, i.e. frequency differences of words (emotives) that are associated with one of eight emotional states (ANGER, ANTICIPATION, DISGUST, FEAR, JOY, SADNESS, SURPRISE, TRUST).

Lexical words are correlated with the age and gender of speakers in the Irish component of the International Corpus of English (ICE) in order to determine whether emotives are socially stratified.

Given social stereotyping (Coates 2015; Holmes 1997; Lakoff 1973), according to which younger speakers, and young female speakers in particular, over-proportionately use emotional language, the present study aims to determine the accuracy of such claims with respect to the use of emotives.

Methodology
The analysis makes use of a Word-Emotion Association Lexicon (Mohammad and Turney 2013), comprising 10,170 terms, in which lexical elements are assigned scores based on ratings gathered through the crowd-sourced Amazon Mechanical Turk service. The associations between terms and emotions are based on 38,726 ratings from 2,216 raters who answered a sequence of questions for each word which were then fed into the emotion association rating. Each term was rated 5 times. For 85% of words, at least 4 raters provided identical ratings.

The data to which the emotion-ratings are applied is the spoken section of the Irish component represented private and public dialogue encompassing speech of 552 speakers uttering 33,200 lexical words.

Each utterance is split up into individual words and aligned with the speaker biographic data (age, gender, etc.). Stop words, i.e. function words which lacked semantic content, were removed.

Next, Sentiment Analysis was performed during which each word was given a score on the eight basic emotions. The resulting table contained one word per row, the age and gender of speakers as well as the emotion category ratings. The final data set represented private and public dialogue encompassing speech of 552 speakers uttering 33,200 lexical words.

Finally, mixed-effects binomial logistic regressions were fit to the data applying a step-wise step-up procedure. Fixed effects included the age and sex of speaker as well as their interaction while speakers were included as a random effect.

Data Summary and Visualization

<table>
<thead>
<tr>
<th>Gender</th>
<th>Age</th>
<th>Speakers</th>
<th>ANGER</th>
<th>ANTICIPATION</th>
<th>FEAR</th>
<th>DISGUST</th>
<th>JOY</th>
<th>SADNESS</th>
<th>SURPRISE</th>
<th>TRUST</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>female</td>
<td>26-33</td>
<td>69</td>
<td>3,952</td>
<td>373</td>
<td>37</td>
<td>59</td>
<td>73</td>
<td>144</td>
<td>1,451</td>
<td>4,735</td>
<td>1,068</td>
</tr>
<tr>
<td>male</td>
<td>26-33</td>
<td>68</td>
<td>3,952</td>
<td>373</td>
<td>37</td>
<td>59</td>
<td>73</td>
<td>144</td>
<td>1,451</td>
<td>4,735</td>
<td>1,068</td>
</tr>
<tr>
<td>female</td>
<td>34-49</td>
<td>48</td>
<td>2,238</td>
<td>55</td>
<td>21</td>
<td>94</td>
<td>39</td>
<td>50</td>
<td>117</td>
<td>2,742</td>
<td>422</td>
</tr>
<tr>
<td>male</td>
<td>34-49</td>
<td>69</td>
<td>2,990</td>
<td>63</td>
<td>128</td>
<td>94</td>
<td>37</td>
<td>59</td>
<td>64</td>
<td>360</td>
<td>1,384</td>
</tr>
<tr>
<td>female</td>
<td>50+</td>
<td>54</td>
<td>3,745</td>
<td>36</td>
<td>160</td>
<td>61</td>
<td>31</td>
<td>140</td>
<td>80</td>
<td>73</td>
<td>599</td>
</tr>
<tr>
<td>male</td>
<td>50+</td>
<td>86</td>
<td>3,219</td>
<td>69</td>
<td>129</td>
<td>90</td>
<td>34</td>
<td>72</td>
<td>81</td>
<td>56</td>
<td>1,451</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>552</td>
<td>33,200</td>
<td>1,384</td>
<td>748</td>
<td>385</td>
<td>1,068</td>
<td>697</td>
<td>1,451</td>
<td>4,054</td>
<td>1,068</td>
</tr>
</tbody>
</table>

The analysis could not detect significant gender or age stratification of emotives associated with ANTICIPATION, DISGUST, SADNESS, or SURPRISE.

The study focuses exclusively on lexical items (words) and neglects contextual factors such as fixed expressions, negation or coercion. The study has not controlled for word types or frequency effects, i.e. investigating whether the effects were caused by an over-use of individual lexical items.

Conclusions
The analysis contradicts social stereotypes according to which young female speakers are more likely to use emotional language compared to other societal cohorts while showing that men – regardless of age – are more likely to use emotives associated with ANGER and FEAR compared to women who are significantly more likely to use emotives associated with JOY.

References