

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 of the *International Corpus of English (ICE)* (Kirk and Kallen 2008). Each utterance is split up into individual words and aligned with the speaker bio-data (age, gender, etc.). Stop words, i.e. function words which lacked semantic content, were removed.

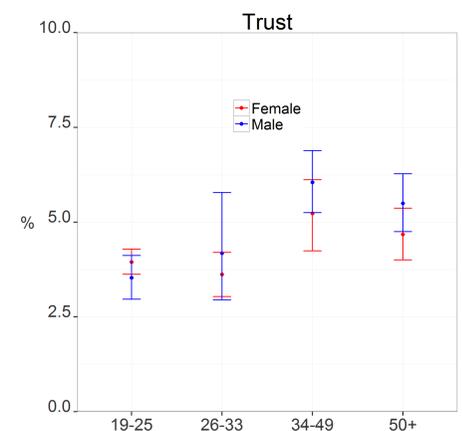
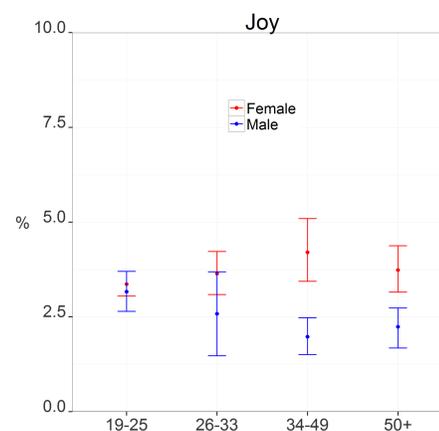
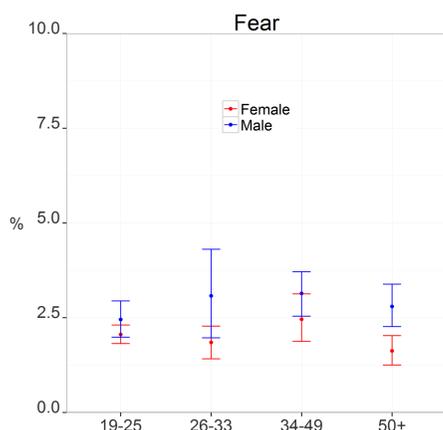
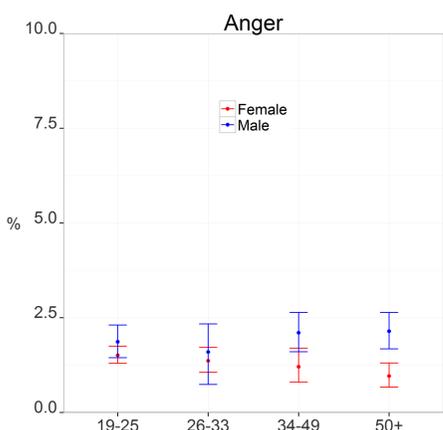
Next, Sentiment Analysis was performed during which each word was given a score on the eight basic emotions. The resulting table contained one word per row, the age and gender of speakers as well as the emotion category ratings. The final data set represented private and public dialogue encompassing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

Age	Gender	Speakers	Words	ANGER	ANTICIPATION	FEAR	DISGUST	JOY	SADNESS	SURPRISE	TRUST	Total
19-25	female	151	12,166	184	517	250	143	409	249	212	480	14,610
19-25	male	58	4,077	76	151	100	61	129	94	82	144	4,914
26-33	female	69	3,952	54	173	73	52	144	68	76	143	4,735
26-33	male	18	813	13	25	25	6	21	22	14	34	973
34-49	female	48	2,238	27	101	55	21	94	39	50	117	2,742
34-49	male	68	2,990	63	128	94	37	59	64	42	181	3,658
50+	female	54	3,745	36	160	61	31	140	80	73	175	4,501
50+	male	86	3,219	69	129	90	34	72	81	50	177	3,921
Total		552	33,200	522	1,384	748	385	1,068	697	599	1,451	40,054

Table 1: Overview of the final data set



Results

The analysis shows that...

- men are significantly more likely to use words associated with ANGER (Odds Ratio 1.5) and FEAR (OR 1.38) compared to women while being less likely to use words associated with JOY (OR 0.93).
- speakers between 34 and 49 are more likely to use words associated with FEAR compared to speakers between 19 and 25 (OR 1.35).
- men between 34 and 49 and above 50 are less likely to use words associated with JOY compared to women between 19 and 25 (OR 0.51, 0.63).
- speakers between 34 and 49 and above 50 are less likely to use words associated with TRUST compared to speakers aged between 19 and 25 (OR 1.57, 1.32).

Emotion	Predictor	χ^2 -value	DF	Significance
ANGER	Sex	11.9	1	<.001***
	Age	11.3	3	<.05*
FEAR	Sex	9.5	1	<.01**
	Age	9.5	3	<.01**
JOY	Sex	20.3	1	<.001***
	Age:Sex	28.1	4	<.001***
TRUST	Age	34.4	3	<.001***

Table 2: Overview of significant predictors for emotive use in Irish English

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

Coates, J. (2015). *Women, men and language: A sociolinguistic account of gender differences in language*. London and New York: Routledge.
 Field, A., J. Miles, and Z. Field (2012). *Discovering statistics using R*. Sage.
 Holmes, J. (1997). Women, language and identity. *Journal of Sociolinguistics* 1(2), 195-223.
 Kirk, J. M. and J. L. Kallen (2008). Ice Ireland 1.2.2. The Ireland component of the International Corpus of English Version 1.2.2. url: <http://www.qub.ac.uk/sites/ICE-Ireland/>
 Lakoff, R. (1973). Language and womans place. *Language in Society* 2(1), 45-79.
 Mohammad, S. M. and P. D. Turney (2013). Crowd sourcing a word-emotion association lexicon. *Computational Intelligence* 29(3), 436-465.