

A SOCIO-LINGUISTIC ANALYSIS OF EMOTIVES IN IRISH ENGLISH

Martin Schweinberger

Home: www.martinschweinberger.de

Mail: martin.schweinberger.hh@gmail.com

University of Hamburg | Universität Kassel

Introduction

This study makes use of the *Word-Emotion Association Lexicon* (Mohammad and Turney 2013) in which lexical elements are aligned with scores based on emotion ratings gathered through the crowd-sourced *Amazon Mechanical Turk service* (Mohammad and Turney 2013) to investigate whether there are systematic differences in emotional language use, i.e. the use 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 extra-linguistic factors in the Irish component of the *International Corpus of English* (ICE) in order to determine whether *emotives* are socially stratified.

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

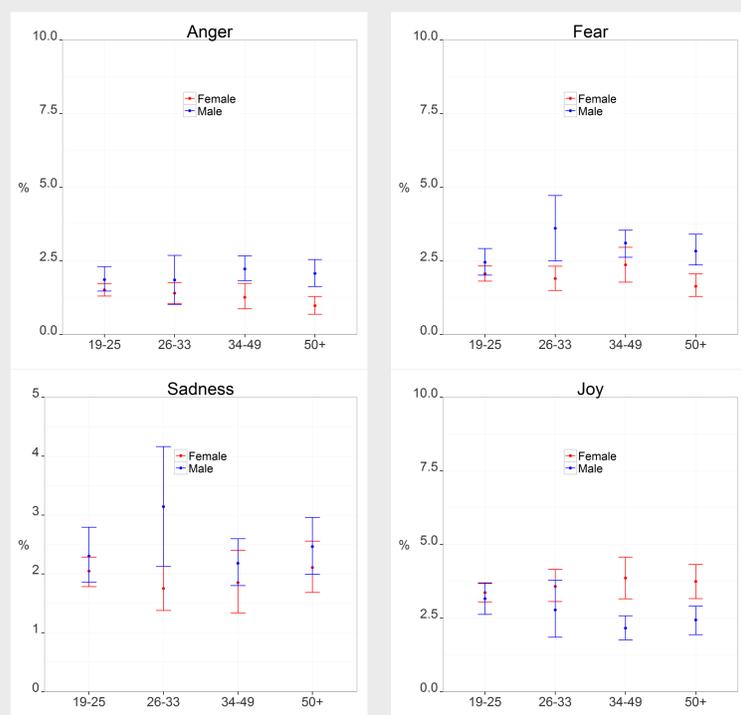
Emotion and language

The concept of EMOTION in the present study is based on Plutchik (1980, 1994) who argues for the existence of eight basic emotions – JOY, SADNESS, ANGER, FEAR, DISGUST, SURPRISE, TRUST, and ANTICIPATION.

The underlying idea is that certain words are associated with one (or more) of these basic emotions while other words lack such associations.

For instance, the words *cry* and *tragedy* are more readily associated with SADNESS while words like *happy* or *beautiful* are indicative of JOY. Higher use of *emotives* is regarded an indicator of emotional language.

Visualization



Methodology

The analysis makes use of an emotion-association lexicon comprising 10,170 terms. 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 (Mohammad and Turney 2013). 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 biodata (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 4 distinct text types (private and public dialogue, unscripted and scripted monologues) encompassing speech of 725 speakers uttering 32,843 lexical words.

Finally, mixed-effects binomial logistic regressions were fit to the data applying an AIC-based step-wise step-up procedure. Fixed effects included the age and sex of speakers, the text type, the number and sex (same vs mixed) of interlocutors, and all second-level interactions while speakers were included as a random effect.

Results

The analysis shows that...

- males are significantly more likely to use words associated with ANGER and SADNESS compared to female speakers while being less likely to use words associated with JOY.
- speakers above 50 are less likely to use words associated with DISGUST and JOY compared to speakers between the ages of 19 and 25.
- words associated with FEAR and TRUST are more likely to occur in public dialogue and in scripted monologues compared to private dialogues.

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.

Statistical Analysis

Predictor	Emotion					
	ANGER	DISGUST	FEAR	JOY	SADNESS	TRUST
sex:male	1.59***	1.29+	1.2+	0.69+	1.22*	
age:26-33		1.01				0.92
age:34-49		0.78				1.15
age:50+		0.65*				1.15
txtyp:S1B			1.75***	1.46		1.39***
txtyp:S2A			0.72	1.14		1.24
txtyp:S2B			1.61**	1.35		1.5***
ints				0.99		
smsx:samesex						1.18
ints:txtypS1B				0.85**		
ints:txtypS2A				0.77		
ints:txtypS2B				0.9		
sex:male:ints				1.03		

Table: Odds Ratio plus significance level; red = sig., pink = marginally sig, gray = not. sig. (insignificant predictors were included, if their inclusion sig. reduced AIC.)

References

References

- Coates, J. (2015). *Women, men and language: A sociolinguistic account of gender differences in language*. London & New York: Routledge.
- 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*.
- Lakoff, R. (1973). Language and woman's place. *Language in society* 2(01), 45–79.
- Mohammad, S. M. and P. D. Turney (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence* 29(3), 436–465.
- Plutchik, R. (1980). *A general psychoevolutionary theory of emotion*. Volume 1 of *Theories of emotion*. New York, NY: Academic Press.
- Plutchik, R. (1994). *The psychology and biology of emotion*. HarperCollins College Publishers.