

Source:

Jacobs, A.M., Schuster, S., Xue, s., & Lüdtke, J. (2017). What's in the brain that ink may character: A Quantitative Narrative Analysis of Shakespeare's 154 Sonnets for Use in (Neuro-)Cognitive Poetics. *Scientific Study of Literature*, 6(2).

Table A1 *Overview of all variables of the quantitative narrative analyses reported in this paper. The raw data table can be obtained from the 1st author on demand: ajacobs@zedat.fu-berlin.de.*

Variable	Explanation
S_ID	sonnet number/name
S_type	sonnet category (<i>young Man</i> vs. <i>dark lady</i>)
Word based Variables	
word_frq_mean	mean value of word frequency according to the Subtlex-database (Brysbaert & New, 2009)
word_frq_sd	standard deviation of word frequency according to the Subtlex-database (Brysbaert & New, 2009)
word_valence_mean	mean value of all single word valence ratings according to the database of Warriner et al. (2013)
word_valence_sd	standard deviation of all single word valence ratings according to the database of Warriner et al. (2013)
word_valence_sum	sum of all single word valence ratings according to the database of Warriner et al. (2013)
word_valence_span	difference between the lowest and highest single word valence rating according to the database of Warriner et al. (2013)
word_arousal_mean	mean value of all single word arousal ratings according to the database of Warriner et al. (2013)
word_arousal_sd	standard deviation of all single word arousal ratings according to the database of Warriner et al. (2013)
word_arousal_sum	sum of all single word arousal ratings according to the database of Warriner et al. (2013)
word_arousal_span	difference between the lowest and highest single word arousal rating according to the database of Warriner et al. (2013)
word_emo_potential_mean	mean value of the product of the valence and arousal rating for all words according to the database of Warriner et al. (2013)
word_emo_potential_sd	standard deviation of the product of the valence and arousal rating for all words according to the database of Warriner et al. (2013)
word_concreteness_mean	mean value of all single word concreteness ratings according to the Subtlex-database (Brysbaert & New, 2009)

word_concreteness_sd	standard deviation of all single word concreteness ratings according to the Subtlex-database (Brysbaert & New, 2009)
word_processing_fluency_mean	mean value of the product of the concreteness rating and word frequency for all words according to the Subtlex-database (Brysbaert & New, 2009)
word_processing_fluency_sd	standard deviation of the product of the concreteness rating and word frequency for all words according to the Subtlex-database (Brysbaert & New, 2009)
word_surprisal_subtlex_mean	mean value of all single word surprisal values computed with the algorithm of Willems et al. (2015) on the Subtlex-Corpus
word_surprisal_subtlex_sd	standard deviation of all single word surprisal values computed with the algorithm of Willems et al. (2015) on the Subtlex-Corpus
word_surprisal_shakespeare_mean	mean value of all single word surprisal values computed with the algorithm of Willems et al. (2015) on the Shakespeare-Corpus
word_surprisal_shakespeare_sd	standard deviation of all single word surprisal values computed with the algorithm of Willems et al. (2015) on the Shakespeare-Corpus
RID_primary_p_cognition	number of words related to primary process cognition according to the Regressive Imagery Dictionary / RID (Martindale, 1975)
RID_secondary_p_cognition	number of words related to secondary process cognition according to the Regressive Imagery Dictionary / RID (Martindale, 1975)
RID_emotions	number of words related to emotions according to the Regressive Imagery Dictionary / RID (Martindale, 1975)

Coh-metrix

Seven Descriptive Indices	e.g. <i>number of words</i> or <i>word length</i> (see http://cohmetrix.com/ for more details)
76 Coh-metrix features	e.g. Noun overlap or results from Latent Semantic Analysis (see http://cohmetrix.com/ for more details)
Flesch Reading Ease	according to Coh-metrix
Flesch Kincaid Grade Level	according to Coh-metrix
L2 Readability	according to Coh-metrix
Eight Text Easability Principal Component Scores	e.g. Narrativity, Syntactic complexity (see http://cohmetrix.com/ for more details)

Symbolic Imagery/ Thematic Richnes/Semantic Association

Symbolic_Imagery_Index	count of symbolic imagery categories according to Meireles (2005)
Simonton_Thematic Richness_Index	count of 24 specific topics identified by Simonton (1989)
EAT_number_unique_associated_WRDs_sum	sum of the number of unique associations for each word according to the Edinburgh Associative Thesaurus (EAT; Kiss et al., 1973)

SEANCE

20 SEANCE components	e.g. component 1 (Negative adjectives) or component 6 (Affect friends and family component, Crossley et al., 2016)
SEANCE_Composite_Score	sum of the values of the 20 SEANCE components (cf. Crossley et al., 2016)
SEANCE_Negative_Mood_Score	sum of the values of the SEANCE components 1 and 7 (cf. Crossley et al., 2016)
SEANCE_Positive_Mood_Score	sum of the values of the SEANCE components 4, 5, 12 and 19 (cf. Crossley et al., 2016)
