

Linguistic features in depression: a meta-analysis

Katrine Bønneland Tølbøll
Aarhus University
katrine.tolboll@gmail.com

Abstract

Recent research on depression suggests that speech can reveal underlying processes in the mind of the depressed. This paper systematically reviews the literature on linguistic features in depression. A corpus of 26 papers investigating the relation between depression and one of the three linguistic features, first-person singular pronouns, positive emotion words, or negative emotion words, were analysed. Three meta-analyses were performed on the three linguistic features. The meta-analyses identify differences in first-person singular pronoun use, negative emotion word use, and positive emotion word use between depressed individuals and healthy controls (Cohen's d of 0.44, 0.72 and -0.38). Furthermore, the meta-analyses identify correlations for severity of depression and first-person singular pronoun use, negative emotion word use, and positive emotion word use (Pearson's r of 0.19, 0.12 and -0.21). All three linguistic features produced small to medium effect sizes thus suggesting a relation between the use of the linguistic features and depression. The effect was not moderated by age or type of task the respondents completed.

Keywords: Depression, Meta-analysis, First-person singular pronouns, Negative emotion words, Positive emotion words.

1. Introduction

Depression is a mental health disorder that affects numerous people worldwide. It is currently estimated that 300 million people globally are suffering from depression (World Health Organization, 2018) and unfortunately, this number is only increasing. The World Health Organization has estimated that by 2030 depression will be the most disabling disease worldwide (Mathers, Fat, Boerma, & World Health Organization, 2008). Thus, it is important to put effort into understanding depression and how it effectively can be diagnosed (Mundt, Snyder, Cannizzaro, Chappie, & Geralts, 2007). Currently, most clinical practices rely on a clinician to evaluate a patient's symptoms, however, researchers are currently directing their attention to finding an effective way of measuring depression which does not rely on clinicians (Cummins et al., 2015; Mundt et al., 2007). Research has been conducted using several methods e.g. genetic biomarkers (Strawbridge, Young, & Cleare, 2017), mobile phone usage (Alhassan et al., 2018) and functional and structural brain imaging (Busatto, 2013; Drevets, 2000).

The present meta-analysis seeks to investigate how depression affects one's use of language. The assumption is that patients suffering from depression will exhibit linguistic behaviour which differs from that of healthy individuals. This is based on the principle of cognitive commitment, which suggests that knowledge about the linguistic structures of human language should be consistent with knowledge about human cognition (Evans, 2012). Hence, natural language acts as a linguistic marker of cognitive processes (Rodriguez, Holleran, & Mehl, 2010). Under this assumption, language acts as a lens of the mind (Evans, 2012). When looking into how depression affects language use, one can focus either on *how* things are being said or on *what* is being said. When looking at how things are being said, researchers look at paralinguistic features like pitch, intensity, speech rate, etc. For example, speakers suffering from depression tend to speak at a slower rate (Cummins et al., 2015). This meta-analysis will only focus on what is being said, more specifically on linguistic features in semantics and morphology.

There has until now – to the knowledge of the author – not been a comparison of multiple linguistic features in depression. A meta-analysis is needed to evaluate whether there is a relation between linguistic features and depression. And a comparison is needed to evaluate which of the linguistic features has the strongest relation to depression. This knowledge can be used as a guideline for future research in this field. After a thorough informal screening of the literature, three linguistic features seemed to be the ones most commonly investigated, thus yielding enough materials for a meta-analysis. These were: first-person singular pronouns, negative emotion words, and positive emotion words. These three linguistic features and their relation to depression were investigated in oral, written and social media domains. Some studies have focused on a clear distinction between language use in depressed individuals and controls (Bernard, Baddeley, Rodriguez, & Burke, 2016; Jarrold et al., 2011; Rude, Gortner, & Pennebaker, 2004). Others have investigated a continuous spectrum, focusing on language use and severity of depression (Jaeger, Lindblom, Parker-Guilbert, & Zoellner, 2014; Pulverman, Lorenz, & Meston, 2015; Rodriguez et al., 2010).

1.1 What is depression and how is it measured?

Depression affects how you feel, think, and act in a negative and destructive way. Furthermore, it is known to be affecting the depressed individual's cognitive processes, such as memory and concentration (Kircanski, Joormann, & Gotlib, 2012). Consequently, depression affects everyday life tremendously.

As earlier mentioned, the most common method of diagnosing depression is for the patient to contact their general practitioner. If the general practitioner suspects that the patient is depressive, the general practitioner will match the patient's pattern of thinking with a series of official agreed-upon criteria in order to diagnose the severity of the depression. This matching is performed via interviews and self-report questionnaires (Cummins et al., 2015).

The most used self-report scale for evaluating the severity of depression is the Beck Depression Inventory (BDI-II) (Maust et al., 2012). BDI was created by Aaron Beck and colleagues in 1961 (A. T. Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). The BDI-II contains 21 items. The items cover cognitive, affective, and somatic symptoms of depression, such as disturbed sleep, tearfulness,

hopelessness, etc. Each item allows the patients four choices ranging from no symptoms to severe symptoms. The patients are asked to report how they felt during the past week. A total score of 0-13 is considered minimal range, 14-19 is mild, 20-28 is moderate, and 29-63 is severe depression. BDI-II's construct validity has been established, and the scale can be used to differentiate depressed from non-depressed patients (Maust et al., 2012).

Another technique of assessing the severity of depression is by using a semi-structured interview with a trained clinician. A popular method is The Hamilton Scale for Depression (Ham-D). Ham-D was created by Max Hamilton in the 1950s. Originally the scale was created to assess whether anti-depressant medicine worked (Hamilton, 1960). Thus, the questionnaire was designed to be used by the clinician during an interview with an already diagnosed depressed patient. Here the clinician rates the severity of depressive symptoms on a scale from 0 to 4 on 17 items. Scores of 0-7 are considered normal, 8-16 mild depression, 17-23 moderate depression, and scores above 24 indicate severe depression.

Several other questionnaires and tests exist, however, studies were only included in this meta-analysis if the participants had completed a commonly accepted tool for assessment of depression such as BDI-II and Ham-D, following guidelines from Pignone et al. (2002).

Self-reports and semi-structured interviews are often criticised for not being objective. Research shows that general practitioners can only correctly recognise depression in 50% of the cases with an already diagnosed depressed person. Further, general practitioners would falsely diagnose non-depressed as depressed in 19% of the cases (Mitchell, Vaze, & Rao, 2009). These findings are a clear indicator of the need for a new way of handling diagnosis of depression.

1.2 Linguistic features in depression

One way to investigate how depression affects language use is to look at linguistic features. To do so, several studies have used the LIWC tool (Linguistic Inquiry and Word Count) (Pennebaker, Chung, Irel, Gonzales, & Booth, 2007). LIWC is a linguistic analysis system, which counts how frequently a person uses words in numerous categories in a text. An example of a category is first-person singular pronouns. Here, the LIWC would categorise *I*, *I'd*, *I'll*, *I'm*, *I've*, *me*, *mine*, *my*, and *myself* as a first-person singular pronoun. The creation of the program sparked a renewed interest in the linguistic style of people with depression (Rude et al., 2004). In 2004, Rude et al. found that depressed individuals use more first-person singular pronouns, more negative emotion words, and fewer positive emotion words than healthy individuals when completing the same task. This has inspired several researchers to conduct further analyses of the three linguistic features and depression. Thus, the three linguistic features are the most commonly researched (Bernard et al., 2016; Edwards & Holtzman, 2017; Fast & Funder, 2010; Rodriguez et al., 2010; Tackman et al., 2019). In the following, the three linguistic features will be elaborated in turn.

1.3 Depression and first-person singular pronouns

Psychologists Jeff Greenberg and Tom Pyszczynski (1987) consider self-awareness of the depressed individual to be important when dealing with depression. Their theory of self-awareness states that a depressed individual becomes trapped in a bad self-regulatory cycle, which they cannot escape from. This self-regulatory cycle can happen when they lose a significant part of their self-esteem e.g. when losing a parent or life-partner (Pyszczynski & Greenberg, 1987). When experiencing such a loss, one enters the self-regulatory cycle to process the loss. But, if one cannot escape the self-regulatory cycle, one may not be able to recover from the loss, consequently leaving one in a high self-focused state (Pyszczynski & Greenberg, 1987). This self-focus causes one to generate unreachable high standards for oneself, which will enhance one's negative feelings and thoughts about oneself.

This theory creates the foundation for the assumption, that a depressed individual will use more first-person singular pronouns since the depressed individual is in a high self-focused state. Several studies have examined the linguistic style of depressed individuals hypothesising an increased use of first-person singular pronouns compared to healthy individuals, finding significant results (Bernard et al., 2016; Dunnack & Park, 2009; Fast & Funder, 2010; M. L. Molendijk et al., 2012; Pulverman et al., 2015; Rude et al., 2004; Sanders, 2013; Sloan, 2005; Zimmermann, Wolf, Bock, Peham, & Benecke, 2013). Some studies, however, have not been able to find evidence for this hypothesis (Rodriguez, Holleran, & Mehl, 2010; Sonnenschein, Hofmann, Ziegelmayer, & Lutz, 2018). Conclusively, the overall picture is unclear, and a more systematic assessment of the evidence is needed.

1.4 Depression and emotion words

Aaron Beck's cognitive theory of depression (1967) states that depression-prone people appraise events in a negative way. Depressive-prone people have latent depressive schemas, which make them predisposed to become depressed (Beck, 1967). The depression-prone person's schemas are mostly inactive until the onset of a depression. The schemas can be activated by certain stressors like being rejected. Beck argues, that the depressive schemas affect the way the depression-prone person process information. Stimuli not congruent to the depressive schema will simply be filtered out, making the attention shift towards stimuli that are schema-congruent (Rude et al., 2004). Beck (1967) described how a dysfunctional thinking pattern generates three forms of automatic negative thinking: negative thoughts of the self, the world, and the future. For example, a depressed person would view themselves as inadequate, thinks that others view them as inadequate as well, thus making the future seem hopeless since the inadequacy will prevent the situation of improving.

Following Beck's cognitive theory of depression (1967), a depressed person experiences themselves, others, and the future in a negative way, thus the depressed person will express themselves using more negative emotion words. Similar, it is also theorised that a depressed individual would use less positive emotion words (Rude et al., 2004).

Much research within the field of depression has taken a starting point in Beck's cognitive theory of depression. Rude et al. (2004) found evidence for the hypothesis that depressed people use significantly more negative emotion words than non-depressed people, this has later largely been

supported by other studies (M. L. Molendijk et al., 2012; Settanni & Marengo, 2015; Sohn, 2017). However, multiple studies have not been able to replicate this relation (Capecehatro, Sacchet, Hitchcock, Miller, & Britton, 2013; Rodriguez, Holleran, & Mehl, 2010; Sloan, 2005).

Positive emotion words have also been examined thoroughly. Several studies report significant results that depressed individuals use fewer positive emotion words than healthy individuals (Baikie et al., 2006; Bernard et al., 2016; Jaeger et al., 2014; M. L. Molendijk et al., 2012; Rodriguez et al., 2010; Sohn, 2017; Sørensen, 2015). In summary, the overall picture is unclear. A more systematic evaluation of the evidence is therefore needed.

1.5 Purpose and aim of meta-analysis

The goal for this meta-analysis is to test three questions: (1) whether there is an overall positive relation between depression and first-person singular pronoun use, (2) whether there is an overall positive relation between depression and negative emotion word use, (3) whether there is an overall negative relation between depression and positive emotion word use, and further to test whether the relations between depression and linguistic features are moderated by the text type produced by the participant, e.g. if the participants participated in a written or oral task.

2. Methods

2.1 Inclusion and exclusion criteria

Papers were only included in the meta-analysis if (1) they investigated the relation between depression and the use of first-person singular pronouns or positive emotion words or negative emotion words, (2) the study used an accepted method of assessment of depression (Pignone et al., 2002) e.g. the BDI and the Ham-D.

Studies investigating this relation in non-clinical and clinical studies were included. Studies reporting relations between depression and the use of first-person singular pronouns or positive and negative emotion words, were excluded if the participants had not completed an assessment of depression. These studies included analyses of e.g. tweets from Social Media posts (Cheng, Li, Kwok, Zhu, & Yip, 2017; Seabrook, Kern, Fulcher, & Rickard, 2018) or suicide notes from depressed individuals (Baddeley, Daniel, & Pennebaker, 2011; Fernández-Cabana, García-Caballero, Alves-Pérez, García-García, & Mateos, 2013).

2.2 Study selection

A formal literature search was conducted on Google Scholar and PubMed on September 25, 2018, on October 29, 2018 and updated on July 23, 2019. The used search terms were: (depress*) AND (pronoun OR negative emotion OR positive emotion), (depress*) AND (linguistic style)¹. These keywords produced 240 potentially relevant articles of which 104 were duplicates. 136 papers were

¹ The use of search-term "linguistic style" was chosen in the aim of including papers reporting studies which report the use of the LIWC, which has a high prevalence in studies within this field.

then screened and 57 were assessed for eligibility based on the abstracts of the papers. Ultimately 26 of these papers met the inclusion criteria and were included in the meta-analysis. Further, an additional search for unpublished studies was performed by asking for references from colleagues who already had researched within this field. This yielded two additional relevant papers.

Conclusively, the literature-search resulted in 28 papers. In these, 22 studies reported statistics for first-person singular pronouns (total of 1842 participants), 17 studies for positive emotion words (total of 1387 participants), and 13 studies for negative emotion words (total of 1149 participants). The remaining 31 papers were used as background literature and are cited when relevant.

2.3 Data extraction

For this meta-analysis, both differences in linguistic style between depressed and controls, and correlations between linguistic style and severity of depression are investigated. For the literature examined, the overall research question was: “Is there a relation between depression and language use?”. To answer this question, three methods are commonly used in the literature.

1. The experimenter collects linguistic data from two groups: an already diagnosed depressed group, and a healthy control group. All participants are then asked to complete an assessment of depression. The data from this assessment is then used to answer the question: Is there a difference in the use of linguistic features between the two groups (Baddeley, Pennebaker, & Beevers, 2013; Brockmeyer et al., 2015; Marc L. Molendijk et al., 2010; Sohn, 2017; Sørensen, 2015).
2. The experimenter asks undiagnosed participants (e.g. college students) to complete an assessment of depression. Based on their scores, the participants are then categorised into two groups, healthy and depressed (following the diagnostic guidelines given within the assessment). The two groups are then tested for differences in the use of linguistic features (Capececiatro, Sacchet, Hitchcock, Miller, & Britton, 2013; De Choudhury, Counts, Horvitz, & Hoff, 2014; Jarrold et al., 2011; Rude et al., 2004; Sloan, 2005).
3. The experimenter asks undiagnosed participants (e.g. college students) to complete an assessment of depression. The students’ score on the assessment and the use of linguistic features are then investigated, to see whether there is a correlation. These studies investigate the correlation between the use of linguistic features and depression severity (Bernard et al., 2016; Dirkse, Hadjistavropoulos, Hesser, & Barak, 2015; Dunnack & Park, 2009; Fast & Funder, 2010; Holmes et al., 2007; Humphreys, King, Choi, & Gotlib, 2018; Jaeger et al., 2014; Pulverman et al., 2015; Rodriguez et al., 2010; Sanders, 2013; Schoch-Ruppen, Ehlert, Uggowitzer, Weymerskirch, & La Marca-Ghaemmaghami, 2018; Settanni & Marengo, 2015).

Studies reporting differences between the use of linguistic features for depressed and healthy individuals were scanned for sample sizes for depressed and healthy groups, characteristics of the task, and age. Further, the statistical estimates (mean, standard deviation and test statistics) were extracted. For papers reporting correlations between the use of linguistic features and severity of depression, estimates of Pearson’s r , characteristics of the task, age and sample sizes were extracted.

To compare, the studies were arranged into two groups based on the task participants were performing. The first category, *Written*, includes studies where the product of the tasks was written material such as essays (Bernard et al., 2016; Holmes et al., 2007; Marc L. Molendijk et al., 2010; Pulverman et al., 2015; Rude et al., 2004; Schoch-Ruppen et al., 2018; Sloan, 2005), diaries (Baikie et al., 2006; Dunnack & Park, 2009; Rodriguez et al., 2010) and Social Media posts (De Choudhury et al., 2014; Settanni & Marengo, 2015). In this category, the participants had a chance to be more deliberate in their word usage, e.g. by censoring themselves and putting much thought into their answers. The second category, *Oral*, includes studies in which the linguistic material was collected from “tasks” such as therapy sessions (Sohn, 2017; Sonnenschein, Hofmann, Ziegelmayer, & Lutz, 2018; Sørensen, 2015; Zimmermann, Brockmeyer, Hunn, Schauenburg, & Wolf, 2017), audio recordings from everyday life (Baddeley et al., 2013) or answering questions orally (Brockmeyer et al., 2015; Capececiatro et al., 2013; Fast & Funder, 2010; Jarrold et al., 2011; Zimmermann et al., 2013). Participants had less opportunity to censor themselves making the answers less deliberate. In this category, the oral answers were transcribed by the experimenter or an associated research assistant.

2.4 Data analysis

To investigate the magnitude of linguistic differences in depression, the effect size for each study is calculated. This meta-analysis focuses on the effect sizes i.e. the size of the relationship between measures of depression and the linguistic features; Cohen’s d and Pearson’s r . Cohen’s d can be classified into small effect, medium effect, and large effect. A small effect is defined as $d = .2$, a medium effect is defined as $d = .5$, and a large effect is defined as $d = .8$ (Cohen, 1988). Furthermore, Cohen (1988) classifies Pearson’s r into small effect, medium effect, and large effect. A small effect is defined as $r = .1$, a medium effect is defined as $r = .3$, and a large effect is defined as $r = .5$. Following Cohen (1988), estimates for Pearson’s r and Cohen’s d are not directly comparable, however, the classifications are comparable. Consequently, papers reporting linguistic differences between depressed and healthy individuals, and papers reporting correlations between the frequency of linguistic features and severity of depression, are examined separately.

For studies reporting differences between the use of linguistic features for depressed and non-depressed individuals, Cohen’s d , statistical significance, and the overall variance of the observed effects across the studies were calculated. To calculate the overall effect size for all the studies, a random-effects regression model was made.

A random effects model allows the size of the relation between depression and linguistic features to differ due to random factors. The chosen random factor for this model was Article, this allows the effect size to differ from study to study, due to the participants who participated in the different studies, had individual differences, which affected the results in different articles.

Thus, the random effects models were conducted using Cohen’s d as the outcome variable predicted only by the intercept with Article as random intercept and weighted by the variance of the effect size.

Using the metafor package (Viechtbauer, 2010) in the statistical program R (R Development Core Team, 2010), Cohen's d and variance of Cohen's d were calculated for all the studies reporting means and standard deviations (Jarrold et al., 2011; M. L. Molendijk et al., 2012; Rude et al., 2004; Sloan, 2005). Not all studies reported their results using the same statistical terms, thus these results have been converted into Cohen's d and variance of Cohen's d . Three studies (Baddeley et al., 2013; Brockmeyer et al., 2015; Sonnenschein et al., 2018) reported F -values and were converted following procedures from Thalheimer & Cook (2002). Three studies reported t -statistics (Capecelatro et al., 2013; De Choudhury et al., 2014; Zimmermann et al., 2017) and were converted following procedures from Thalheimer & Cook (2002). Two studies (Sohn, 2017; Sørensen, 2015) reported estimates for Poisson mixed-effects models. Beta-estimates and standard errors were converted into t -statistics and finally converted to Cohen's d (Field, Miles, & Field, 2012).

Following the same procedure as for calculating Cohen's d , the metafor package (Viechtbauer, 2010) was used to calculate the variance of Pearson's r , for studies reporting correlations between the use of linguistic features and severity of depression (Baikie et al., 2006; Bernard et al., 2016; Brockmeyer et al., 2015; Dirkse et al., 2015; Dunnack & Park, 2009; Fast & Funder, 2010; Holmes et al., 2007; Humphreys et al., 2018; Jaeger et al., 2014; Pulverman et al., 2015; Rodriguez et al., 2010; Sanders, 2013; Settanni & Marengo, 2015). A random-effects model was constructed using Pearson's r as the outcome variable predicted only by the intercept with Article as random intercept and weighted by the variance of the effect size.

Since all studies have not been conducted under the exact same conditions, Cochran's Q (Cochran, 1954) and I^2 (Higgins, Thompson, Deeks, & Altman, 2003) were calculated to see if this affected the overall results. I^2 indicates the percentage of variance within the meta-analysis that can be attributed to study heterogeneity, thus explaining how much of the overall variance can be explained by the within-study variance (Higgins et al., 2003).

It was assessed whether the task-categories, written or oral, could contribute to explain the overall variance. To assess this, a mixed-effects regression model was created with effect size as the outcome, task as fixed effect and Article as a random effect, thus assuming different random intercepts for each article. Moreover, it was assessed whether the respondents' age could explain the overall variance.

Lastly, the effect of influential studies and publication bias was investigated. An influential study strongly drives the overall results thus influencing the results. Publication bias is the tendency to only publish significant results. Possible influential studies and publication bias was estimated using rank correlation test, here assessing whether lower sample sizes were related to bigger effect sizes. A significant rank correlation indicates a possible publication bias.

3. Results

3.1 First person singular pronouns

The overall literature search yielded 28 papers consisting of 57 studies. Of these 57 studies, 22 reported sufficient statistical estimates for the category first-person singular pronouns. Of the 22 studies, 12 studies used a written task. (See table 1.1 and 1.2).

Table 1.1 Studies comparing use of first-person singular pronouns in depressed and healthy individuals

Article	Year	Sample size Depressed	Sample size Control	Mean age	Depression measure	Task
Rude, Gortner and Pennebaker	2004	31	67	18.34 years	BDI-II	Written – Essay about starting university
Jarrold et al.	2011	14	12	69.75 years	CES-D	Oral – Structured interview
Brockmeyer et al.	2015	59	29	38.75 years	BDI-II	Oral – Question about saddest and happiest days
Molendijk et al.	2010	110	108	37.90 years	SCL-90	Written – Essay about life
Baddeley et al.	2013	29	28	34.22 years	BDI-II	Oral – Audio-recordings from everyday lives of depressed people
De Choudhury et al.	2014	28	137	30.37 years	PHQ-9	Written – Facebook posts
Sloan	2005	18	17	19.20 years	BDI-II	Written – Essay about the previous day
Sonnenschein et al.	2018	27	58	37.40 years	SCID-I	Oral – Interview with a therapist
Sohn	Unpublished	42	43	32.54 years	Ham-D	Oral – Interview with a therapist
Sørensen	Unpublished	16	16	32.05 years	Ham-D	Oral – Interview with a therapist

Table 1.2 Studies comparing use of first-person singular pronouns and severity of depression

Article	Year	Sample size	Mean age	Depression measure	Task
Bernard et al.	2016	136	18.80 years	CES-D	Written – Essay about starting university
Fast & Funder a	2010	90	NA	BDI-II	Oral – Semi-structured interview
Fast & Funder b	2010	90	NA	BDI-II	Oral – Semi-structured interview
Sanders	2013	34	22.84 years	BDI-II	Written – Essay about the past
Rodriguez et al. a	2010	57	18.70 years	BDI-II	Written – Diary about themselves
Rodriguez et al. b	2010	57	18.70 years	BDI-II	Written – Blog posts about themselves
Dunnack & Park	2009	120	NA	CES-D	Written – Journal about a serious loss
Pulverman et al.	2015	133	34.11 years	BDI-II	Written – Essay about deepest thoughts about sex
Humphreys et al.	2018	54	34.86 years	CES-D	Oral – Speech sample on participants' child
Ziemer & Korkmaz	2017	93	49.60 years	CES-D	Written – Expressive writing session on thoughts and feelings
Dirkse et al.	2015	59	NA	PHQ-9	Written – Expressive writing exercise by email

3.1.1 META-ANALYSIS

Cohen's d and variance of Cohen's d were calculated for all the studies reporting either means and standard deviations, F -values, or t -statistics. Pearson's r and variance of Pearson's r were calculated for all the studies reporting correlations between first-person singular pronouns and severity of depression.

Two random effects models were constructed to estimate the overall effect sizes and their variance. To visualize the effect sizes across papers, forest plots were produced (figure 1).

As shown on the forest plot (figure 1, right-hand side), the estimated effect size (Cohen's d) for the 10 included studies which compared use of first-person singular pronouns between depressed and healthy individuals, does not overlap zero (Cohen's $d = 0.44$, $CI = [0.30, 0.59]$)², hence, indicating a distinctive pattern of increased first-person singular pronoun use in depression. Following Cohen (1988) this effect is classified as a small to medium effect, thus the effect is non-trivial. From the data, an overall variance (τ^2) of 0.0013 ($SE = 0.0234$) was found. Most of the variance ($I^2 = 2.36\%$) could be reduced to random sample variability between studies (Q -stats = 10.21, $p = 0.33$). There were no significant outliers or any obvious publication bias (Kendall's tau = 0.29, $P = 0.29$).

² If studies do not find a positive relation between linguistic features and depression, the effect size (Cohen's d and Pearson's r) will be below 0 (above 0 in the case of positive emotion words since we expect a negative relation). The model labeled "RE model" is the combined estimate of the all effect sizes from all the studies.

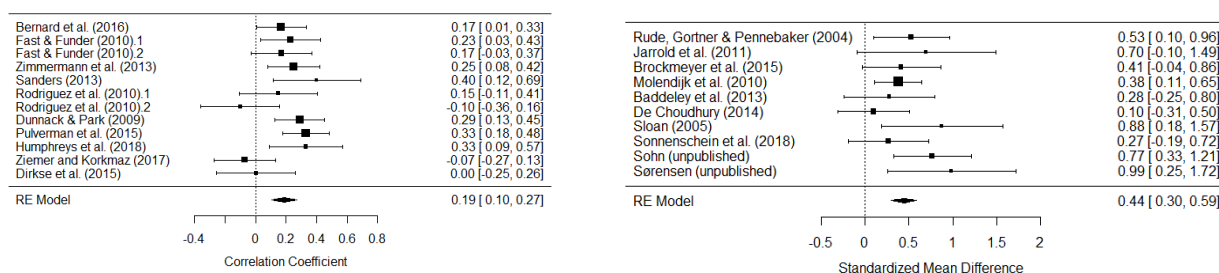


Figure 1. Forrest plots of effect sizes (Cohen's d and Pearson's r) in use of first-person singular pronouns.

The x-axis reports the effect size (positive values indicate larger use of first-person singular pronouns in depression, while negative indicate lower use). The y-axis reports the studies for which statistical estimates of use of first-person singular pronouns were provided.

The estimated effect size (Pearson's r) for 12 included studies which compare use of first-person singular pronouns and severity of depression, does not overlap zero (Pearson's $r = 0.19$, CI = [0.10, 0.27]) (See figure 1, left-hand side). The estimated effect size is considered small and thus being non-trivial (Cohen, 1988). This indicates a positive correlation between severity of depression and first-person singular pronoun use. The overall variance (τ^2) is 0.0104 (SE = 0.0091). Much of the variance ($I^2 = 49.66\%$) could not be reduced to random sample variability between studies (Q-stats = 21.95, $p = 0.025$). There were no significant outliers or any obvious publication bias (Kendall's tau = -0.33, $P = 0.15$).

3.2 Negative emotion words

13 studies reported statistical estimates for negative emotion words. Of the 13 studies, 9 studies used a written task. (See table 2.1 and table 2.2)

Table 2.1 Studies comparing use of negative emotion words in depressed and healthy individuals

Article	Year	Sample size Depressed	Sample size Control	Mean age	Depression measure	Task
Rude, Gortner and Pennebaker	2004	31	67	18.34 years	BDI-II	Written – Essay about starting university
Capecelatro et al.	2013	27	25	47.40	BDI-II	Oral – Recall Photographs
Molendijk et al.	2010	110	108	37.90 years	SCL-90	Written – Essay about life
Baddeley et al.	2013	29	28	34.22 years	BDI-II	Oral – Audio-recordings from everyday lives of depressed people
Sloan	2005	18	17	19.20 years	BDI-II	Written – Essay about the previous day
Sohn	Unpublished	42	43	32.54 years	Ham-D	Oral – Interview with a therapist

Table 2.2 Studies comparing use of negative emotion words and severity of depression

Article	Year	Sample size	Mean age	Depression measure	Task
Bernard et al.	2016	136	18.80 years	CES-D	Written – Essay about starting university
Settanni & Marengo	2015	201	28.40 years	DASS-21	Written – Facebook posts
Rodriguez et al. a	2010	57	18.70 years	BDI-II	Written – Diary about themselves
Rodriguez et al. b	2010	57	18.70 years	BDI-II	Written – Blog posts about themselves
Holmes et al.	2007	25	36.30 years	BDI-II	Written – Expressive writing on the most traumatic life event
Jaeger et al.	2014	35	NA	BDI-II	Oral – Interview on one neutral and one traumatic life event
Ziemer & Korkmaz	2017	93	49.60 years	CES-D	Written – Expressive writing on thoughts and feelings

3.2.1 META-ANALYSIS

Following the same procedure, Cohen's d , the variance of Cohen's d , Pearson's r and variance of Pearson's r was calculated. To visualise the effect sizes across papers, forest plots were produced (figure 2).

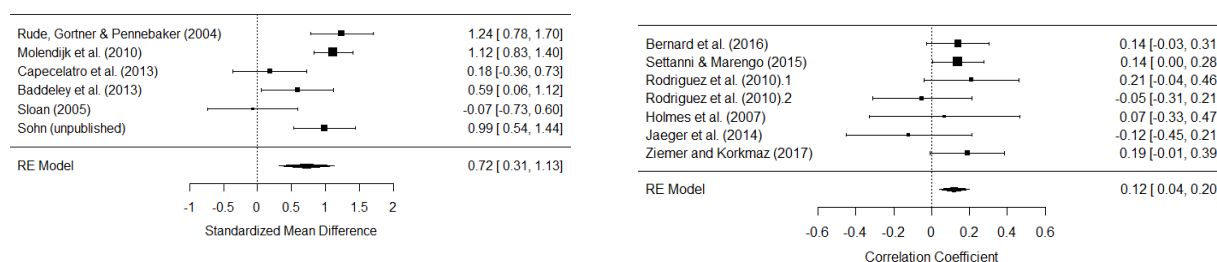


Figure 2. Forrest plots of effect sizes (Cohen's d and Pearson's r) in use of negative emotion words

As shown on the forest plot (figure 2, left-hand side), the estimated effect size (Cohen's d) for the 6 included studies which compare differences in use of negative emotion words in depressed and healthy individuals, does not overlap zero (Cohen's $d = 0.72$, $CI = [0.32, 1.13]$), indicating a pattern of increased negative emotion words used in depression. Following Cohen (1988) this is classified as a medium to large effect. From the data, an overall variance (τ^2) of 0.1956 ($SE = 0.1629$) was found. Much of the variance ($I^2 = 78.19\%$) could not be reduced to random sample variability between studies (Q -stats = 20.41, $p = 0.001$). There were some significant outliers thus a possible publication bias (Kendall's tau = -0.87, $P = 0.017$), meaning mostly studies that found an effect have been published.

The estimated effect size (Pearson's r) for 7 included studies which compare differences in negative emotion words and severity of depression, does not overlap zero (Pearson's $r = 0.12$, $CI = [0.04, 0.20]$) (see figure 2, right-hand side). The estimated effect size is considered small and thus being non-trivial (Cohen, 1988), thus, indicating a positive correlation between use of negative emotion

words and severity of depression. The overall variance (τ^2) is 0.0000 (SE = 0.0071). All the variance ($I^2 = 0.00\%$) could be reduced to random sample variability between studies (Q-stats = 4.82, $p = 0.57$). There were no significant outliers or any obvious publication bias (Kendall's tau = -0.43, $P = 0.24$).

3.3 Positive emotion words

17 studies reported statistical estimates for positive emotion words. Of the 17 studies, 12 studies used a written task (see table 3.1 and table 3.2).

Table 3.1 Studies comparing use of positive emotion words in depressed and healthy individuals

Article	Year	Sample size Depressed	Sample size Control	Mean age	Depression measure	Task
Rude, Gortner and Pennebaker	2004	31	67	18.34 years	BDI-II	Written – Essay about starting university
Capecelatro et al.	2013	27	25	47.40 years	BDI-II	Oral – Recall Photographs
Molendijk et al.	2010	110	108	37.90 years	SCL-90	Written – Essay about life
Baddeley et al.	2013	29	28	34.22 years	BDI-II	Oral – Audio-recordings from everyday lives of depressed people
Sloan	2005	18	17	19.20 years	BDI-II	Written – Essay about the previous day
Sohn	Unpublished	42	43	32.54 years	Ham-D	Oral – Interview with a therapist
Sørensen	Unpublished	16	16	32.05 years	Ham-D	Oral – Interview with a therapist

Table 3.2 Studies comparing use of positive emotion words and severity of depression

Article	Year	Sample size	Mean age	Depression measure	Task
Bernard et al.	2016	136	18.80 years	CES-D	Written – Essay about starting university
Settanni & Marengo	2015	201	28.40 years	DASS-21	Written – Facebook posts
Rodriguez et al. a	2010	57	18.70 years	BDI-II	Written – Diary about themselves
Rodriguez et al. b	2010	57	18.70 years	BDI-II	Written – Blog posts about themselves
Pulverman et al.	2015	133	34.11 years	BDI-II	Written – Essay about deepest thoughts about sex
Holmes et al.	2007	25	36.30 years	BDI-II	Written – Expressive writing on the most traumatic life event
Jaeger et al.	2014	35	NA	BDI-II	Oral – Interview on one neutral and one traumatic life event
Ziemer & Korkmaz	2017	93	49.60 years	CES-D	Written – Expressive writing on thoughts and feelings
Baikie et al.	2006	14	NA	DASS-21	Written – Diary entries
Dirkse et al.	2015	59	NA	PHQ-9	Written – Expressive writing exercise by email

3.3.1 META-ANALYSIS

Random effects models were conducted to estimate the overall effect sizes and their variance. Following forest plots (figure 3) were produced to visualise the effect size across papers.

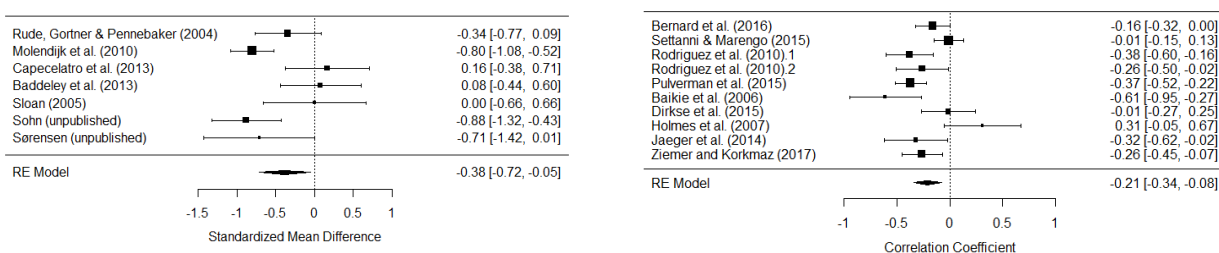


Figure 3. Forrest plots of effect sizes (Cohen's d and Pearson's r) in use of positive emotion words.

As shown on the forest plot (figure 3, left-hand side), the estimated effect size (Cohen's d) for the 7 included studies which compare differences in use of positive emotion words in depressed and healthy individuals, does not overlap zero (Cohen's $d = -0.38$, CI = [-0.72, -0.05]), indicating a pattern of decreased use of positive emotion words in depression. According to Cohen (1988), this effect size is classified as a small to medium effect. However, it is worth noting the large confidence intervals and thus the relatively uncertain estimates. From the data, an overall variance (τ^2) of 0.12 (SE = 0.12) was found. Much of the variance ($I^2 = 68.89\%$) could not be reduced to random sample variability between studies (Q-stats = 15.58, $p = 0.008$). There were no significant outliers or any obvious publication bias (Kendall's tau = 0.33, $P = 0.47$).

The estimated effect size (Pearson's r) for 10 included studies which compare use of negative emotion words and severity of depression, does not overlap zero (Pearson's $r = -0.21$, CI = [-0.34, -0.08]) (see figure 3, right-hand side). The estimated effect size is considered small to medium and thus being non-trivial (Cohen, 1988). The overall variance (τ^2) is 0.0304 (SE = 0.0208). The variance ($I^2 = 72.95\%$) could not be reduced to random sample variability between studies (Q-stats = 31.40, $p = 0.0003$). There were no significant outliers or any obvious publication bias (Kendall's tau = -0.02, $P = 1.00$).

4. Discussion

4.1 Overview and interpretation of the results

Research has suggested a distinctive pattern of increased first-person singular pronoun use, increased use of negative emotion words, and decreased use of positive emotion words in depressed individuals. This paper has systematically reviewed the evidence for this suggestion. Here, 57 studies were identified, 22 reporting first-person singular pronouns, 13 negative emotion words, 17 positive emotion words.

By the usual standards as suggested by Cohen (1988), a small to medium positive relation between the use of first-person singular pronouns and depression was found (Cohen's $d = 0.44$ and Pearson's $r = 0.23$). This suggests that the link between first-person singular pronouns use and depression is present even though it is considered a small effect. The lower effect size for correlations (Pearson's r) between first-person singular pronouns and severity of depression might indicate that while a strong contrast (depressed vs. control) shows a stronger difference, more nuanced contrasts, e.g. different levels of depressive traits in controls, do not show the same differences. In other words, the differences between the frequency of first-person singular pronouns, when having a depression severity score of for example 12 and 13 on the BDI scale, is small.

Conclusively, depressed individuals use more frequently first-person singular pronouns than healthy individuals. Furthermore, the more depressed an individual is, the more first-person singular pronouns they use. The current finding is consistent with Pyszczynski and Greenberg's theory (1987), that depressed individuals tend to perseverate about themselves because they are stuck in the self-regulatory cycle. This also corresponds with earlier done meta-analyses investigating the relation (Edwards & Holtzman, 2017; Tackman et al., 2019).

A medium positive relation between the use of negative emotion words and depression was found (Cohen's $d = 0.72$ and Pearson's $r = 0.12$). The effect is concluded to be medium due to the contradictory classification of the effect sizes (i.e. relatively high Cohen's d , but relatively low Pearson's r). As with first-person singular pronouns, the lower effect size for correlations between negative emotion words and severity of depression might indicate that while a strong contrast (depressed vs. control) shows a stronger difference, more nuanced contrasts, e.g. different levels of depressive traits in controls, do not show the same differences. Furthermore, it was tested whether the interpretation of the results could have been affected by a publication bias or an influential paper, which could possibly have skewed the results. A possible publication bias was found in papers reporting differences between the use of negative emotion words in depressed and healthy individuals. This publication bias could possibly explain the large effect size which was found. In short, depressed individuals use more negative emotion words compared to healthy individuals. Further, the more depressed an individual is, the more negative emotion words the individual would use. However, the effect found here might have been affected by a possible publication bias. The finding supports the theory of Beck's cognitive theory of depression, which states that depressed individuals express themselves negatively (1967).

A small to medium negative relation between the use of positive emotion words and depression was found (Cohen's $d = -0.38$ and Pearson's $r = -0.21$). In line with first-person singular pronouns and negative emotion words, the small Pearson's r might indicate that a smaller effect might be due to the small contrasts. As a result, depressed individuals use fewer positive emotion words compared to healthy individuals. Further, the more depressed an individual is, the fewer positive emotion words the individual would use. The finding corresponds with Beck's cognitive theory of depression; depressed individuals express themselves in a less positive way (1967).

To the knowledge of the author, no meta-analyses have carried out an investigating of the relation between depression and use of negative- and positive emotion words.

Lastly, the variance of results presented here for all three linguistic features could not be explained by either age or task category. The results were affected neither by the age of the participants or whether they participated in an oral or written task.

4.2 Conceptual and methodological limitations in the definition of depression and severity

Scales and questionnaires are the most used tools by practical clinicians since they investigate multiple known symptoms of depression such as hopelessness, guilt feelings, weight loss (Aaron T. Beck, 1967; Edwards & Holtzman, 2017). Yet, general practitioners have only a 50% chance of correctly diagnosing a patient already suffering from depression (Mitchell et al., 2009). Diagnosing a healthy individual as depressed might lead the person to take unneeded anti-depressives, which can lead to unwanted side effects (Santarsieri & Schwartz, 2015). Not diagnosing a depressed individual as depressed, leaves the individual suffering unnecessarily. Research shows that different depression rating scales do not overlap for symptoms, meaning some rating scales only investigate certain symptoms (Fried, 2017). Patients suffering from depression might not experience all typical symptoms, so some scales might not investigate the symptoms the patient experiences (Fried, 2017). This could be a possible explanation for the low success rates for general practitioners. Fried (2015) has suggested that future research on depression should not rely on diagnosis but on individual symptoms. An example of why research should look at individual symptoms are given with the linguistic feature; positive emotion words. When investigating the use of positive emotion words and depression, an important but rarely often considered effect, is the fact that when a very depressed individual has decided to commit suicide, they often experience a sense of relief (Rudd, 2008). Depressed people who are suicidal often see their own death as a solution to end their suffering and can therefore experience an increase in hopefulness (Rudd, 2008). This increase in hopefulness is seen in an increase in positive emotion word use (Baddeley et al., 2011). Studies show how depressed individuals just prior to their suicide use more positive emotion words, and fewer negative emotion words (Lester, 2010).

Currently, research only focuses on diagnosis (depressed vs not depressed). The diagnosis is reached by looking at multiple symptoms of depression and sums symptoms-scores (Pignone et al., 2002). When analysing linguistic features in depression, studies only look at the diagnosis and not the underlying symptoms, thus generalising that all depressed have the same symptoms. The positive correlation between depressed suicidal individuals and use of positive emotion words is a good argument for why future research on depression should rely on individual symptoms. When looking at people diagnosed with depression, we see an overall negative relation between positive emotion words and depression. However, suicide has the opposite relation to positive emotion words. Consequently, research might gain a more nuanced perspective on depression by investigating the different underlying symptoms and their possible relations with different linguistic features. To the knowledge of the author, no research on linguistic markers of depression has adopted this approach yet.

4.3 Future directions

In line with positive emotions words and suicide, future research could benefit from investigating individual symptoms in depression, which potentially could give a more refined perspective (Fried, 2015). Further, machine learning systems could be used to analyse and predict symptoms of depression. Several studies have already used machine learning trying to predict depressed individuals based on linguistic cues. In 2013, De Choudhury et al. were able to predict depression with 70% classification accuracy using social media. As impressively, Moore et al. (2008) were able to predict the depression using paralinguistic features with a classification accuracy of 91% for males and 96% for females. Sohn (2017) and Sørensen (2015) were able to predict depression using linguistic features with a classification of 72% and 87% respectively. These unarguably high success rates yield for further research within the field of language use in depression

5. Conclusion

After systematically reviewing the literature on language use in depression, it is concluded that there is an overall small but positive relation between depression and first-person singular pronouns, and depression and negative emotion word use. Further, a small overall negative relation between depression and positive emotion words was found. Studies implementing findings trying to predict depression from linguistic features yield promising results, therefore more research should be conducted trying to predict depression from language use.

References

- Alhassan, A. A., Alqadhib, E. M., Taha, N. W., Alahmari, R. A., Salam, M., & Almutairi, A. F. (2018). The relationship between addiction to smartphone usage and depression among adults: A cross sectional study. *BMC Psychiatry*, *18*. <https://doi.org/10.1186/s12888-018-1745-4>
- Baddeley, J. L., Daniel, G. R., & Pennebaker, J. W. (2011). How Henry Hellyer's use of language foretold his suicide. *Crisis*, *32*(5), 288–292. <https://doi.org/10.1027/0227-5910/a000092>
- Baddeley, J. L., Pennebaker, J. W., & Beevers, C. G. (2013). Everyday social behavior during a major depressive episode. *Social Psychological and Personality Science*, *4*(4), 445–452. <https://doi.org/10.1177/1948550612461654>
- Baikie, K. A., Wilhelm, K., Johnson, B., Boskovic, M., Wedgwood, L., Finch, A., & Huon, G. (2006). Expressive writing for high-risk drug dependent patients in a primary care clinic: A pilot study. *Harm Reduction Journal*, *3*, 34. <https://doi.org/10.1186/1477-7517-3-34>
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*, 561–571.
- Beck, Aaron T. (1967). *Depression: Clinical, Experimental, and Theoretical Aspects*. University of Pennsylvania Press.
- Bernard, J. D., Baddeley, J. L., Rodriguez, B. F., & Burke, P. A. (2016). Depression, Language, and Affect: An Examination of the Influence of Baseline Depression and Affect Induction on Language. *Journal of Language and Social Psychology*, *35*(3), 317–326. <https://doi.org/10.1177/0261927X15589186>
- Brockmeyer, T., Zimmermann, J., Kulesa, D., Hautzinger, M., Bents, H., Friederich, H.-C., ... Backenstrass, M. (2015). Me, myself, and I: Self-referent word use as an indicator of self-

- focused attention in relation to depression and anxiety. *Frontiers in Psychology*, 6, 1564. <https://doi.org/10.3389/fpsyg.2015.01564>
- Busatto, G. F. (2013). Structural and Functional Neuroimaging Studies in Major Depressive Disorder With Psychotic Features: A Critical Review. *Schizophrenia Bulletin*, 39(4), 776–786. <https://doi.org/10.1093/schbul/sbt054>
- Capecehatro, M. R., Sacchet, M. D., Hitchcock, P. F., Miller, S. M., & Britton, W. B. (2013). Major depression duration reduces appetitive word use: An elaborated verbal recall of emotional photographs. *Journal of Psychiatric Research*, 47(6), 809–815. <https://doi.org/10.1016/j.jpsychires.2013.01.022>
- Cheng, Q., Li, T. M., Kwok, C.-L., Zhu, T., & Yip, P. S. (2017). Assessing Suicide Risk and Emotional Distress in Chinese Social Media: A Text Mining and Machine Learning Study. *Journal of Medical Internet Research*, 19(7), e243. <https://doi.org/10.2196/jmir.7276>
- Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, 10, 101–129. <https://doi.org/10.2307/3001666>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J: L. Erlbaum Associates.
- Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication*, 71, 10–49. <https://doi.org/10.1016/j.specom.2015.03.004>
- De Choudhury, M., Counts, S., Horvitz, E. J., & Hoff, A. (2014). Characterizing and predicting postpartum depression from shared facebook data. *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 626–638. ACM.
- De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *ICWSM*, 13, 1–10.
- Dirkse, D., Hadjistavropoulos, H. D., Hesser, H., & Barak, A. (2015). Linguistic analysis of communication in therapist-assisted internet-delivered cognitive behavior therapy for generalized anxiety disorder. *Cognitive Behaviour Therapy*, 44(1), 21–32. <https://doi.org/10.1080/16506073.2014.952773>
- Drevets, W. C. (2000). Functional anatomical abnormalities in limbic and prefrontal cortical structures in major depression. In *Cognition, Emotion and Autonomic Responses: The Integrative Role of the Prefrontal Cortex and Limbic Structures: Vol. 126. Progress in Brain Research* (pp. 413–431). [https://doi.org/10.1016/S0079-6123\(00\)26027-5](https://doi.org/10.1016/S0079-6123(00)26027-5)
- Dunnack, E. S., & Park, C. L. (2009). The Effect of an Expressive Writing Intervention on Pronouns: The Surprising Case of I. *Journal of Loss and Trauma*, 14(6), 436–446. <https://doi.org/10.1080/15325020902925084>
- Edwards, T., & Holtzman, N. S. (2017). A meta-analysis of correlations between depression and first person singular pronoun use. *Journal of Research in Personality*, 68, 63–68. <https://doi.org/10.1016/j.jrp.2017.02.005>
- Evans, V. (2012). Cognitive linguistics. *Wiley Interdisciplinary Reviews: Cognitive Science*, 3(2), 129–141. <https://doi.org/10.1002/wcs.1163>
- Fast, L. A., & Funder, D. C. (2010). Gender Differences in the Correlates of Self-Referent Word Use: Authority, Entitlement, and Depressive Symptoms. *Journal of Personality*, 78(1), 313–338. <https://doi.org/10.1111/j.1467-6494.2009.00617.x>
- Fernández-Cabana, M., García-Caballero, A., Alves-Pérez, M. T., García-García, M. J., & Mateos, R. (2013). Suicidal Traits in Marilyn Monroe's Fragments. *Crisis*, 34(2), 124–130. <https://doi.org/10.1027/0227-5910/a000183>

- Field, A. P., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. London ; Thousand Oaks, Calif: Sage.
- Fried, E. I. (2015). Problematic assumptions have slowed down depression research: Why symptoms, not syndromes are the way forward. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00309>
- Fried, E. I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *Journal of Affective Disorders*, 208, 191–197. <https://doi.org/10.1016/j.jad.2016.10.019>
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23, 56–62.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ (Clinical Research Ed.)*, 327(7414), 557–560. <https://doi.org/10.1136/bmj.327.7414.557>
- Holmes, D., Alpers, G. W., Ismailji, T., Classen, C., Wales, T., Cheasty, V., ... Koopman, C. (2007). Cognitive and emotional processing in narratives of women abused by intimate partners. *Violence Against Women*, 13(11), 1192–1205. <https://doi.org/10.1177/1077801207307801>
- Humphreys, K. L., King, L. S., Choi, P., & Gotlib, I. H. (2018). Maternal depressive symptoms, self-focus, and caregiving behavior. *Journal of Affective Disorders*, 238, 465–471. <https://doi.org/10.1016/j.jad.2018.05.072>
- Jaeger, J., Lindblom, K. M., Parker-Guilbert, K., & Zoellner, L. A. (2014). Trauma Narratives: It's What You Say, Not How You Say It. *Psychological Trauma: Theory, Research, Practice and Policy*, 6(5), 473–481. <https://doi.org/10.1037/a0035239>
- Jarrold, W., Javitz, H. S., Krasnow, R., Peintner, B., Yeh, E., Swan, G. E., & Mehl, M. (2011). Depression and Self-Focused Language in Structured Interviews with Older Men. *Psychological Reports*, 109(2), 686–700. <https://doi.org/10.2466/02.09.21.28.PR0.109.5.686-700>
- Kircanski, K., Joormann, J., & Gotlib, I. H. (2012). Cognitive Aspects of Depression. *Wiley Interdisciplinary Reviews. Cognitive Science*, 3(3), 301–313. <https://doi.org/10.1002/wcs.1177>
- Lester, D. (2010). The Final Hours: A Linguistic Analysis of the Final Words of a Suicide. *Psychological Reports*, 106(3), 791–797. <https://doi.org/10.2466/pr0.106.3.791-797>
- Mathers, C., Fat, D. M., Boerma, J. T., & World Health Organization (Eds.). (2008). *The global burden of disease: 2004 update*. Geneva, Switzerland: World Health Organization.
- Maust, D., Cristancho, M., Gray, L., Rushing, S., Tjoa, C., & Thase, M. E. (2012). Psychiatric rating scales. In *Handbook of Clinical Neurology* (Vol. 106, pp. 227–237). <https://doi.org/10.1016/B978-0-444-52002-9.00013-9>
- Mitchell, A. J., Vaze, A., & Rao, S. (2009). Clinical diagnosis of depression in primary care: A meta-analysis. *Lancet (London, England)*, 374(9690), 609–619. [https://doi.org/10.1016/S0140-6736\(09\)60879-5](https://doi.org/10.1016/S0140-6736(09)60879-5)
- Molendijk, M. L., van Tol, M.-J., Penninx, B. W. J. H., van der Wee, N. J. A., Aleman, A., Veltman, D. J., ... Elzinga, B. M. (2012). BDNF val66met affects hippocampal volume and emotion-related hippocampal memory activity. *Translational Psychiatry*, 2, e74. <https://doi.org/10.1038/tp.2011.72>
- Molendijk, Marc L., Bamelis, L., van Emmerik, A. A. P., Arntz, A., Haringsma, R., & Spinhoven, P. (2010). Word use of outpatients with a personality disorder and concurrent or previous major depressive disorder. *Behaviour Research and Therapy*, 48(1), 44–51. <https://doi.org/10.1016/j.brat.2009.09.007>

- Moore, E., Clements, M. A., Peifer, J. W., & Weisser, L. (2008). Critical analysis of the impact of glottal features in the classification of clinical depression in speech. *IEEE Transactions on Bio-Medical Engineering*, 55(1), 96–107. <https://doi.org/10.1109/TBME.2007.900562>
- Mundt, J. C., Snyder, P. J., Cannizzaro, M. S., Chappie, K., & Geralt, D. S. (2007). Voice acoustic measures of depression severity and treatment response collected via interactive voice response (IVR) technology. *Journal of Neurolinguistics*, 20(1), 50–64. <https://doi.org/10.1016/j.jneuroling.2006.04.001>
- Pennebaker, J. W., Chung, C. K., Irel, M., Gonzales, A., & Booth, R. J. (2007). *LIWC2007 Manual The Development and Psychometric Properties of LIWC2007*.
- Pignone, M. P., Gaynes, B. N., Rushton, J. L., Burchell, C. M., Orleans, C. T., Mulrow, C. D., & Lohr, K. N. (2002). Screening for Depression in Adults: A Summary of the Evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine*, 136(10), 765. <https://doi.org/10.7326/0003-4819-136-10-200205210-00013>
- Pulverman, C. S., Lorenz, T. A., & Meston, C. M. (2015). Linguistic changes in expressive writing predict psychological outcomes in women with history of childhood sexual abuse and adult sexual dysfunction. *Psychological Trauma: Theory, Research, Practice and Policy*, 7(1), 50–57. <https://doi.org/10.1037/a0036462>
- Pyszczynski, T., & Greenberg, J. (1987). Self-regulatory perseveration and the depressive self-focusing style: A self-awareness theory of reactive depression. *Psychological Bulletin*, 102(1), 122–138. <http://dx.doi.org.ez.statsbiblioteket.dk:2048/10.1037/0033-2909.102.1.122>
- R Development Core Team. (2010). *a language and environment for statistical computing: Reference index*. Retrieved from <http://www.polsci.wvu.edu/duval/PS603/Notes/R/fullrefman.pdf>
- Rodriguez, A. J., Holleran, S. E., & Mehl, M. R. (2010). Reading Between the Lines: The Lay Assessment of Subclinical Depression From Written Self-Descriptions. *Journal of Personality*, 78(2), 575–598. <https://doi.org/10.1111/j.1467-6494.2010.00627.x>
- Rudd, M. D. (2008). Suicide warning signs in clinical practice. *Current Psychiatry Reports*, 10(1), 87–90. <https://doi.org/10.1007/s11920-008-0015-4>
- Rude, S., Gortner, E.-M., & Pennebaker, J. (2004). Language use of depressed and depression-vulnerable college students. *Cognition and Emotion*, 18(8), 1121–1133. <https://doi.org/10.1080/02699930441000030>
- Sanders, S. (2013). Exploring the linguistic styles of students with a propensity for alcoholism and students with symptoms of depression. *HIM 1990-2015*. Retrieved from <https://stars.library.ucf.edu/honorstheses1990-2015/1460>
- Santarsieri, D., & Schwartz, T. L. (2015). Antidepressant efficacy and side-effect burden: A quick guide for clinicians. *Drugs in Context*, 4. <https://doi.org/10.7573/dic.212290>
- Schoch-Ruppen, J., Ehlert, U., Uggowitz, F., Weyerskirch, N., & La Marca-Ghaemmaghami, P. (2018). Women's Word Use in Pregnancy: Associations With Maternal Characteristics, Prenatal Stress, and Neonatal Birth Outcome. *Frontiers in Psychology*, 9, 1234. <https://doi.org/10.3389/fpsyg.2018.01234>
- Seabrook, E. M., Kern, M. L., Fulcher, B. D., & Rickard, N. S. (2018). Predicting Depression From Language-Based Emotion Dynamics: Longitudinal Analysis of Facebook and Twitter Status Updates. *Journal of Medical Internet Research*, 20(5), e168. <https://doi.org/10.2196/jmir.9267>
- Settanni, M., & Marengo, D. (2015). Sharing feelings online: Studying emotional well-being via automated text analysis of Facebook posts. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01045>

- Sloan, D. M. (2005). It's All About Me: Self-Focused Attention and Depressed Mood. *Cognitive Therapy and Research*, 29(3), 279–288. <https://doi.org/10.1007/s10608-005-0511-1>
- Sohn, J. (2017). *A Quantitative Investigation of Depressed Language Use*. Aarhus University.
- Sonnenschein, A. R., Hofmann, S. G., Ziegelmayr, T., & Lutz, W. (2018). Linguistic analysis of patients with mood and anxiety disorders during cognitive behavioral therapy. *Cognitive Behaviour Therapy*, 47(4), 315–327. <https://doi.org/10.1080/16506073.2017.1419505>
- Sørensen, K. (2015). *When Depression Speaks: A computational analysis of word use in Danish individuals with first-episode major depressive disorder*. Aarhus University.
- Strawbridge, R., Young, A. H., & Cleare, A. J. (2017). Biomarkers for depression: Recent insights, current challenges and future prospects. *Neuropsychiatric Disease and Treatment*, 13, 1245–1262. <https://doi.org/10.2147/NDT.S114542>
- Tackman, A. M., Sbarra, D. A., Carey, A. L., Donnellan, M. B., Horn, A. B., Holtzman, N. S., ... Mehl, M. R. (2019). Depression, negative emotionality, and self-referential language: A multi-lab, multi-measure, and multi-language-task research synthesis. *Journal of Personality and Social Psychology*, 116(5), 817–834. <https://doi.org/10.1037/pspp0000187>
- Thalheimer, W., & Cook, S. (2002). How to calculate effect sizes from published research articles: A simplified methodology. Retrieved 16 November 2018, from http://www.bwgriffin.com/gsu/courses/edur9131/content/Effect_Sizes_pdf5.pdf
- Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the **metafor** Package. *Journal of Statistical Software*, 36(3). <https://doi.org/10.18637/jss.v036.i03>
- World Health Organization. (2018, March 22). Depression. Retrieved 25 November 2018, from World Health Organization website: <http://www.who.int/news-room/fact-sheets/detail/depression>
- Ziemer, K. S., & Korkmaz, G. (2017). Using text to predict psychological and physical health: A comparison of human raters and computerized text analysis. *Computers in Human Behavior*, 76, 122–127. <https://doi.org/10.1016/j.chb.2017.06.038>
- Zimmermann, J., Brockmeyer, T., Hunn, M., Schauenburg, H., & Wolf, M. (2017). First-person Pronoun Use in Spoken Language as a Predictor of Future Depressive Symptoms: Preliminary Evidence from a Clinical Sample of Depressed Patients. *Clinical Psychology & Psychotherapy*, 24(2), 384–391. <https://doi.org/10.1002/cpp.2006>
- Zimmermann, J., Wolf, M., Bock, A., Peham, D., & Benecke, C. (2013). The way we refer to ourselves reflects how we relate to others: Associations between first-person pronoun use and interpersonal problems. *Journal of Research in Personality*, 47(3), 218–225. <https://doi.org/10.1016/j.jrp.2013.01.008>