Credit and Capital Markets, 56 (2023) 3 / 4: 353 – 388
https://doi.org/10.3790/ccm.2023.1436301
Scientific Papers

Using Negations in Analyzing German Texts in Finance

Matthias Pöferlein*

Abstract

Domain-specific dictionaries have prevailed, when conducting the dictionary-based approach to measure the sentiment of textual data in finance. Through the contributions of *Bannier* et al. (2019a) and *Pöferlein* (2021), two versions of a dictionary suitable for analyzing German finance-related texts are available (BPW dictionary). This paper conducts and tests further improvements of the given word lists by calculating the sentiment of German-speaking annual reports to forecast future return on assets and future return on equity. This corrected and expanded version provides more significant results. Despite the broad usage of negations, this type of improvement in combination with the BPW dictionary has not yet been tested when conducting the dictionary-based approach. Therefore, this paper additionally tests different negation lists to show that implementing negations can improve results.

Keywords: Textual Analysis, Textual Sentiment, Sentiment Analysis, Content Analysis, Negations, Annual Reports

JEL Classification: G14, G17

I. Introduction

Public companies use annual reports as a tool of external communication with investors. Investors use these reports as a basis for their investment decisions. In addition to business figures, these reports contain a large amount of text, which is purely qualitative information. By using methods of textual analysis, the quantitative information encoded in these texts can be obtained and further processed. Therefore, obtaining annual reports' textual sentiment to prove correlations with financial ratios or share prices, represents an established field in accounting and finance research (*Chakraborty/Bhattacharjee* 2020; *Kang* et al. 2018; *Kearney/Liu* 2014; *Loughran/McDonald* 2011). We focus our paper on the two variables future return on assets (*FROA*) and future return on equity (*FROE*)

* Matthias Pöferlein (M. A.), University of Bayreuth, Chair of Finance and Banking, Universitätsstr. 30, 95447 Bayreuth, E-Mail: poeferlein@web.de.

We would like to thank an anonymous referee for the valuable comments and suggestions, which substantially improved the quality of this paper.

Duncker & Humblot · Berlin

one year ahead, which are frequently used as an independent performance measure in relevant studies (*Daniel* et al. 2004; *King* et al. 2004; *Koelbl* 2020; *Mysková/Hájek* 2020; *Vojinović* et al. 2020).

Algaba et al. (2020) define that "sentiment is the disposition of an entity toward an entity, expressed via a certain medium." The specified disposition can be conveyed quantitatively through numbers although it is primarily expressed qualitatively, using text, audio, or visual media (*Algaba* et al. 2020). This sentiment provides a measure of the degree of positivity or negativity and can potentially offer an additional perspective in the process of stock price formation. As a result, it can help address key questions in the field of behavioral finance (*Kearney/Liu* 2014).

The two most common textual analysis methods for obtaining sentiment from qualitative data are the dictionary-based approach (or bag-of-words) and machine learning (*Chakraborty/Bhattacharjee* 2020; *Kearney/Liu* 2014). Using a mapping algorithm, the dictionary-based approach utilizes predefined word lists to assign words into positive, negative, or other sentiment categories like uncertainty. By counting these classified words, several measurements of sentiment can be calculated (*Li* 2010; *Loughran/McDonald* 2015; *Rice/Zorn* 2019). The machine learning approach uses a subset of linguistic labeled texts to train complex models. These models are then used to predict the sentiment of a given set of texts (*Rice/Zorn* 2019; *Shapiro* et al. 2022). Contributions like *Frankel* et al. (2022) and *Mishev* et al. (2020) show that to measure the sentiment of financial text, machine learning approaches can be superior. However, this advantage has the additional disadvantage that machine learning approaches are often a blackbox and are therefore almost unreplicable and difficult to explain (*Algaba* et al. 2020; *Krause* et al. 2016). To prevent these challenges and provide a replicable approach for future research, this paper focuses on the dictionary-based approach.

When using the dictionary-based approach, domain-specific dictionaries have proven to be superior and prevailed in analyzing financial texts (*Kang* et al. 2020; *Kearney/Liu* 2014; *Loughran/McDonald* 2015; *Luo/Zhou* 2020; *Shapiro* et al. 2022). The newly developed finance word lists by *Bannier* et al. (2019a) (BPW_O) have been improved by *Pöferlein* (2021) (BPW_N). Due to the novelty of those dictionaries, the first hypothesis of this paper is that further correcting and expanding the BPW_N dictionary, to get an expanded BPW_E dictionary, improves the results of forecasting future ROAs and ROEs from the sentiment of annual reports. One possible improvement that has not yet been tested in the context of the BPW word lists is the use of negations. Due to their potential high impact and widespread usage (*Bochkay* et al. 2020; *Borochin* et al. 2018; *Loughran/McDonald* 2011; *Shapiro* et al. 2022), the additional hypothesis of this paper is that accounting for negations additionally improves results.

The contribution of this paper to the literature on analyzing German-speaking financial texts is the further extension and optimization of the edited version of the BPW dictionary. Additionally, this paper is the first contribution using different negations combined with the two versions of the BPW dictionary. Therefore, future research in analyzing the sentiment of German-speaking texts in finance can be conducted more precisely.

This paper proceeds as follows. In the second part, we provide a short review of the relevant literature on textual analysis, focusing on analyzing financial texts with and without using negations. The third section presents the data and the applied parsing procedure, in addition to the usage and creation of the dictionaries. The fourth section highlights the empirical approach used to obtain the results presented in section five. Lasty, the sixth section concludes.

II. Literature Review

Several contributions like *Chakraborty/Bhattacharjee* (2020), *Kearney/Liu* (2014), and *Luo/Zhou* (2020) provide an excellent overview of the extensive field of textual analysis in finance. Moreover, certain overview papers provide additional information about specific areas of caution (*Algaba* et al. 2020; *Loughran/McDonald* 2016) and ideas for future research (*Kaya* et al. 2020). Due to the above-mentioned reasons this paper and therefore the following literature review focuses on the dictionary-based approach.

One of the first steps in measuring the sentiment of a text is selecting a dictionary or word list (*Loughran/McDonald* 2015). According to *Loughran/McDonald* (2016), four different word lists have been primarily used by researchers in classifying English finance-related texts. These can be divided into two general dictionaries, namely “General Inquirer” (*Stone* et al. 1966) and “DICTION” (*Hart* 2000), and two word lists generated for finance-related texts by *Henry* (2006, 2008) and *Loughran/McDonald* (2011).

Through the contributions of *Henry* (2006, 2008) and *Loughran/McDonald* (2011), the usage of general word lists for different forms of finance-related textual content like news (*Tetlock* 2007; *Tetlock* et al. 2008), earnings press releases (*Davis* et al. 2012; *Davis/Tama-Sweet* 2012) or annual reports (*Feldman* et al. 2008; *Yuthas* et al. 2002) was widely criticized in favor of domain-specific word lists (*Algaba* et al. 2020; *Chakraborty/Bhattacharjee* 2020; *Lewis/Young* 2019; *Loughran/McDonald* 2015; *Mishev* et al. 2020; *Price* et al. 2012).

In the field of finance, the word lists provided by *Loughran/McDonald* are primarily used (*Kearney/Liu* 2014; *Loughran/McDonald* 2016), for different kinds of finance-related textual data. These lists were used to analyze news (*Ferguson* et al. 2015; *Hillert* et al. 2018), conference calls (*Da Tonin/Scherer* 2022; *Druz* et al. 2020), and annual reports (*Berns* et al. 2022; *Kang* et al. 2018).

The above-mentioned domain-specific problems regarding the German language were also present. Research was primarily limited to general dictionaries like SentiWS (*Remus* et al. 2010) and LIWC (*Meier* et al. 2018; *Wolf* et al. 2008). In order to rectify this problem, *Bannier* et al. (2019a) introduced a German domain-specific dictionary in the field of finance. After the usage of the original word lists in different contributions (*Bannier* et al. 2017, 2019b; *Röder/Walter* 2019; *Tillmann/Walter* 2018, 2019), a reformed and extended version was introduced by *Pöferlein* (2021).

An essential element in the approach introduced by *Loughran/McDonald* (2011) is the use of negations. They account for simple negations for their list of positive words using the six negations "no, not, none, neither, never, nobody" occurring within three words preceding a positive word (*Loughran/McDonald* 2011). In accordance with the work of Loughran and McDonald, negations are widely used in the textual analysis of business texts. These are either used in the form proposed by Loughran and McDonald (*Huang* et al. 2014; *Renault* 2017), as an extended version of the six negations (*Borochin* et al. 2018; *Brau* et al. 2016; *Correa* et al. 2021) or in other forms (*Jandl* et al. 2014; *Jegadeesh/Wu* 2013).

Despite having the contribution by *Loughran/McDonald* (2011) as a theoretical foundation (*Bannier* et al. 2019a), *Bannier* et al. (2017, 2019a, 2019b) and other authors using the BPW have not yet accounted for negations in their papers (*Pöferlein* 2021; *Röder/Walter* 2019; *Tillmann/Walter* 2018, 2019).

III. Data

1. Data Source

We get the initial sample of relevant companies and all the financial variables from the Amadeus database provided by Bureau van Dijk. Hereby we focus on stock-listed companies from three German-speaking countries, Austria, Germany, and Switzerland. Additionally, we only select companies with available reports for at least one year between 2010 and 2020. From the initial sample of 893 companies, 740 companies published at least one annual report on their web page. We were able to find and manually download 6,275 annual reports[1]. Table 1 provides an overview of the Amadeus search strategy, and the following sample creation. We obtained all other variables from Amadeus.

[1] 620 annual reports have a different fiscal year. Due to available data in Amadeus those reports weren't removed from the sample.

Table 1

Sample Creation

Source/Filter	Sample Size
Active companies in Amadeus	3,105,008
Country: Austria, Germany, Switzerland	480,282
Stock listed companies	10,738
At least one available annual report in the years 2010 to 2020	893
Company with annual report available on Homepage	740
Final sample of annual reports	6,275

2. Used Dictionaries

We use the BPW_N dictionary proposed by *Pöferlein* (2021) to analyze the annual reports. These word lists also build the foundation for constructing the BPW_E word lists. Additionally, we use the original word lists by *Bannier* et al. (2019a) (BPW_O) to compare results.

To get the extended version of the BPW_N (BPW_E) we manually check all word lists and delete words with a different or ambiguous meaning (e.g. "prolongiert" (English: prolonged) on the negative word list). During the review of all three relevant lists, we delete 22 words on the positive, 141 words on the negative and 259 words on the stop words list.

In order to find missing words in all three word lists, we use the German news corpus 2020 from Universität Leipzig (2022) to check every word for missing basic forms and variations. Additionally, we account for synonyms, their basic forms, and variations. We manually check all words found for their plausibility regarding the different word lists. Out of the 35,254 basic forms found, we add 1,911 positive, 3,157 negative, and 779 stop words. Through the 17,630 synonyms found, we are able to add another 746 positive, 2,389 negative, and 85 stop words. Finally, we add an alternative spelling of mutated vowels according to *Pöferlein* (2021). A summary of the conducted steps and the resulting alteration of the three word lists is presented in Table 2.

Table 2

Updating the BPW_N Dictionaries

	Positive	Negative	Stop words
BPW_O total words	2,223	10,147	3,682
BPW_N total words	2,849	12,661	4,132
Delete words with a different meaning	–22	–141	–259
Adding basic forms	+1,911	+3,157	+779
Adding synonyms	+746	+2.389	+85
Adding mutated vowels	+692	+1,336	+84
BPW_E total words	6,176	19,402	4,821

We use four different lists of negations. Firstly, we obtain the two German lists of the Linguistic Inquiry and Word Count LIWC2001 and LIWC2015 in their original form (*Meier* et al. 2018; *Wolf* et al. 2008), containing 13 and 39 negations. Additionally, we generate two own lists based on the six negations given by *Loughran/McDonald* (2011)[2]. Furthermore, we account for the criticism of *Picault/Renault* (2017) by adding the word "lower", resulting in seven negations[3]. To obtain the German version of these two lists, we screen 30 corresponding annual statements of the DAX companies in 2017 for the negations given by *Loughran/McDonald* (2011) and *Picault/Renault* (2017) and their matching German translations. This approach is based on *Bannier* et al. (2019a) where they evaluated their dictionary by using corresponding German and English quarterly and annual reports from DAX and MDAX companies. Overall, we find 8,063 translations of the Loughran and McDonald negations, resulting in 25 individual negations. Due to the additional word "lower", 9,201 translations can be found for the Picault and Renault negations, resulting in 316 individual negations (including mutated vowels).

We apply the above-described approach of obtaining the extended version of the three word lists to the four negation lists resulting in 26 LIWC2001, 49 LIWC2015, 28 LMD, and 916 PR negations. Altogether we manually check 2,525 basic forms and 1,151 synonyms for their plausibility. Finally, we add the above used alternative spelling of mutated vowels. Table 3 summarizes all steps and the resulting alterations.

[2] Negation list LMD.

[3] Negation list PR.

Table 3

Creating and Updating Negations

	LIWC 2001	LIWC 2015	LMD	PR
Basic form / Translation (BPW_N)	13	39	25	316
Delete words with a different meaning				–6
Adding basic forms	+12	+8	+1	+397
Adding synonyms	+1	+2	+2	+84
Adding mutated vowels				+125
BPW_E total words	26	49	28	916

3. *Parsing*

Based on the criticism of *Loughran/McDonald* (2015), we follow *Pöferlein* (2021) in giving a detailed overview of performed text manipulation. Owing to this approach, difficulties in replicating this study due to unspecified parsing rules are avoided.

First and foremost, we convert the manually collected PDFs to UTF-8 encoded TXT files (*Bannier* et al. 2017, 2019b; *Kang* et al. 2020; *Meier* et al. 2018). We conduct the following parsing procedure in accordance with *Pöferlein* (2021) using an automated parser programmed in Python. We replace typographic ligatures (*Bannier* et al. 2017, 2019b), hyphens (*Loughran/McDonald* 2011), and convert all words to lowercase (*Pengnate* et al. 2020; *Picault/Renault* 2017; *Tillmann/Walter* 2018). Furthermore, we remove irrelevant content in the form of special characters (*Allee/Deangelis* 2015; *Fritz/Tows* 2018), numbers (*Ferris* et al. 2013; *Gentzkow* et al. 2019), punctuation (*Iqbal/Riaz* 2022; *Picault/Renault* 2017), and multiple whitespaces (*González* et al. 2019; *Schmeling/Wagner* 2016). Eventually, we follow *Bannier* et al. (2017, 2019b) and delete all words with less than three characters. Depending on the dictionary, we use the associated stop word list (BPW_O, BWP_N or BPW_E).

Following *Pöferlein* (2021), we include an automated alteration of the words "betrug" and "sorgen" prior to the parsing procedure when using the BPW_N word lists. Additionally, when using the BPW_E, we add the word "bremse" from the BPW_N word list and the two words "stahl" and "sucht" from the BPW_E dictionary to the automated alteration. When written in lowercase the words "betrug", "sorgen" and "sucht" are changed to "betrugnoneg", "sorgennoneg" and "suchtnoneg". Additionally, the words "bremse" and "stahl" are changed to "bremsenoneg" and "stahlnoneg" when written with a first capital letter. These alterations are due to the change in meaning of certain words when written with a first capital or lowercase letter. Due to peculiarities of the German language, in addition to the approach of *Pöferlein* (2021), occurrences of the word "betrug" at the beginning of a sentence are changed to "betrugnoneg". Table 4 displays an overview of these different meanings. Due to this pre parsing procedure, we are able to additionally reduce the stated exaggeration of negative words in *Pöferlein* (2021).

Table 4

Differences Between Capital and Lowercase Letters

Words with a first capital letter	Translation	Words with a first lowercase letter	Translation
Betrug	fraud	**betrug**	amounted
Bremse	brake	bremse	slow down
Sorgen	sorrow	**sorgen**	care
Stahl	steel	stahl	steal
Sucht	addiction	**sucht**	search

Note: German words altered using the suffix "noneg" are bold.

IV. Methodology

1. *Measurement of Sentiment and Implementation of Negations*

We use Python to count the occurrence of positive (p) and negative (n) words from each of the three dictionaries. We use the relative measurement of Net-Tone (*NTone*), which is the most common measurement regarding the BPW-Dictionary (*Bannier* et al. 2017, 2019b; *Tillmann/Walter* 2018) and has proven to be superior to other measurements (*Pöferlein* 2021). This measurement solely focuses on the number of positive and negative words and is not altered by the length of analyzed documents:

$$NTone = \frac{p - n}{p + n} \tag{1}$$

In the existing literature, negations are considered in two different ways. In order to provide a fully comprehensive analysis of the influence of negations, this paper uses both approaches. We follow *Druz* et al. (2020), *Loughran/McDonald* (2011), and *Shapiro* et al. (2022) in counting words as negated if there is a negation among the three preceding words. In handling negated words, we use two different approaches. In accordance with *Bushman* et al. (2016) and *Druz* et al. (2020), negated words are not counted. Measurements using this approach are marked with the suffix "_ig" (for ignore). Additionally, the more common approach of handling negations is term shifting (*Algaba* et al. 2020; *Bochkay* et al. 2020; *Jandl* et al. 2014; *Taboada* et al. 2011). Here the negated word is counted as a word from the opposite dictionary. Measurements using this approach are marked with the suffix "_ts" (for term shifting). Depending on the respective dictionaries, the corresponding negation lists are used.

Following *Bannier* et al. (2017), *Davis* et al. (2015), and *Pöferlein* (2021), all words found are weighted equally. Due to this, other researchers can replicate and further develop the results of this paper. *Henry/Leone* (2016) also support this approach and the superiority of equal weighting.

2. Empirical Approach

The most common approach for measuring the impact of sentiment on future profitability using a bag-of-words model is linear regression (*Bannier* et al. 2019b; *Boudt/Thewissen* 2019; *Henry* et al. 2021; *Patelli/Pedrini* 2014). Therefore, we apply the following linear regression model using two different dependent variables:

$$Dep_j = \alpha_0 + \alpha_1 NTone_j + \sum_{k=1}^{K} \alpha_k Control_{kj} + \varepsilon_j \tag{2}$$

Dep represents the two different variables, future return on assets (*FROA*) and future return on equity (*FROE*) one year ahead. Both variables are used frequently as an independent performance measure (*Daniel* et al. 2004; *King* et al. 2004; *Koelbl* 2020), even though ROA is considered to be more accurate and less influenced by accounting (*Myšková/Hájek* 2020; *Vojinović* et al. 2020).

We use five different control variables (*Control*) as well as year and industry fixed effects based on relevant research findings (*Alshorman/Shanahan* 2022; *Aly* et al. 2018; *Boudt/Thewissen* 2019; *Davis/Tama-Sweet* 2012; *González* et al. 2019; *Kang* et al. 2018). These include the age of the company (*AGE*), a dummy variable to identify loss firms (*LOSS*), the leverage (*LEV*), the current return on assets (*ROA*) and the current return on equity (*ROE*). When using *FROA* as a dependent variable *ROE* is excluded from the regression. The same applies for using *FROE* and *ROA*. The calculation of all variables can be found in the appendix (Table 16).

V. Results

According to *Loughran/McDonald* (2011), we exclude annual reports with less than 2,000 words from the sample. Additionally, we eliminate reports with less than 200 individual words to remove corrupted data. Due to different stop word lists connected with the particular dictionaries, the numbers of excluded reports and, therefore, the numbers of analyzed annual reports vary. A possible alternative of considering the following analyses on a uniform data sample is not carried out, as this contradicts the general basic logic of using different dictionaries.

1. Summary Statistics

The following three tables report the summary statistics for all three dictionaries used. Table 5 provides descriptive statistics for all variables used to analyze the original dictionary by *Bannier* et al. (2019a) (BPW_O). It can be observed that the future and present return variables have a high standard deviation, with values ranging from highly negative to highly positive.

Table 5

Descriptive Statistics for BPW_O Variables (N = 4,168)

Statistic	Mean	St. Dev.	Min	Max	Pctl. (25)	Pctl. (75)
FROA	4.303	11.967	–93.678	100.000	1.367	8.676
FROE	6.592	43.348	–783.269	372.161	3.029	18.954
NTone	–0.075	0.183	–0.750	0.703	–0.195	0.035
AGE	49.641	45.224	0.000	555.000	16.000	92.000
LOSS	0.178	0.383	0.000	1.000	0.000	0.000
LEV	1.648	3.015	0.000	111.411	0.581	1.881
ROA	4.534	11.764	–91.969	90.525	1.596	8.856
ROE	9.089	40.780	–783.269	924.023	3.639	19.430

As shown in Table 6, the mean *NTone* using BPW_N slightly increases, while the standard deviation and minimum values remain the same. Additionally, the maximum value slightly decreases. The additional usage of negations leads to higher values of *NTone*, where using a combination of PR negations and term shifting creates a positive mean.

Table 6

Descriptive Statistics for BPW_N Variables (N = 4,112)

Statistic	Mean	St. Dev.	Min	Max	Pctl. (25)	Pctl. (75)
FROA	4.309	11.928	−93.678	100.000	1.383	8.684
FROE	6.659	43.503	−783.269	372.161	3.064	19.035
NTone	−0.051	0.183	−0.750	0.696	−0.172	0.062
NTone_LIWC01_ig	−0.043	0.185	−0.745	0.711	−0.165	0.070
NTone_LIWC15_ig	−0.032	0.186	−0.733	0.730	−0.156	0.081
NTone_LMD_ig	−0.039	0.184	−0.739	0.702	−0.161	0.074
NTone_PR_ig	−0.024	0.188	−0.733	0.723	−0.146	0.092
NTone_LIWC01_ts	−0.031	0.177	−0.708	0.708	−0.147	0.075
NTone_LIWC15_ts	−0.008	0.172	−0.630	0.719	−0.122	0.096
NTone_LMD_ts	−0.024	0.178	−0.679	0.698	−0.143	0.085
NTone_PR_ts	0.010	0.165	−0.630	0.673	−0.093	0.112
AGE	49.803	45.309	0.000	555.000	16.000	92.750
LOSS	0.177	0.382	0.000	1.000	0.000	0.000
LEV	1.663	3.032	0.000	111.411	0.593	1.898
ROA	4.545	11.737	−91.969	90.525	1.606	8.874
ROE	9.163	40.985	−783.269	924.023	3.738	19.498

Table 7 shows that further extending the three word lists leads to an increase in *NTone*, resulting in a positive mean. In contrast to using the BPW_N dictionary, the usage of negations also leads to positive means.

Table 7

Descriptive Statistics for BPW_E Variables (N = 4,116)

Statistic	Mean	St. Dev.	Min	Max	Pctl(25)	Pctl(75)
FROA	4.310	11.922	–93.678	100.000	1.384	8.680
FROE	6.660	43.482	–783.269	372.161	3.068	19.015
NTone	0.139	0.141	–0.558	0.740	0.046	0.227
NTone_LIWC01_ig	0.148	0.143	–0.553	0.764	0.053	0.239
NTone_LIWC15_ig	0.149	0.144	–0.553	0.764	0.054	0.240
NTone_LMD_ig	0.146	0.142	–0.553	0.756	0.053	0.234
NTone_PR_ig	0.154	0.146	–0.554	0.781	0.058	0.247
NTone_LIWC01_ts	0.147	0.137	–0.484	0.767	0.055	0.231
NTone_LIWC15_ts	0.148	0.138	–0.484	0.767	0.056	0.232
NTone_LMD_ts	0.148	0.140	–0.537	0.758	0.056	0.235
NTone_PR_ts	0.147	0.133	–0.463	0.731	0.060	0.227
AGE	49.826	45.322	0.000	555.000	16.000	93.000
LOSS	0.177	0.382	0.000	1.000	0.000	0.000
LEV	1.662	3.031	0.000	111.411	0.592	1.898
ROA	4.546	11.732	–91.969	90.525	1.606	8.869
ROE	9.162	40.965	–783.269	924.023	3.738	19.491

To compare the alteration of *NTone* when using BPW_N and BPW_E, we conduct a dependent-samples t-test. There is a significant difference between *NTone*, when using BPW_N (*Mean* = –0.078, *St. Dev.* = 0.210) or BPW_E (*Mean* = 0.123, *St. Dev.* = 0.156), $t(6247) = -161.77$, $p < .001$.[4]

As highlighted in Table 8, regarding all 6,275 analyzed reports, the editing of stop words leads to an alteration of total and individual words found. Interestingly in contrast to the BPW_N individual words using BPW_E decrease, while the total number of words increase. Expanding the positive and negative word lists of the BPW_N lead to an immense increase in total and individual words.

[4] For conducting the t-test, all 6,275 data points are used.

Table 8

Total Number of words

	BPW_O	BPW_N	BPW_E
	All words		
Number of words	156,966,254	127,408,125	129,692,675
Individual words	1,143,403	1,143,083	1,142,806
	Positive words		
Number of words	2,169,243	2,219,778	5,709,076
Individual words	1,702	1,718	4,075
	Negative words		
Number of words	2,488,910	2,436,004	4,323,617
Individual words	5,013	5,028	8,341

After correcting for dictionary-specific stop word lists, Table 9 displays the cumulative fraction of the ten most frequent positive words used. Despite having minor differences in fractions, the positive words used in BPW_O and BPW_N are identical. In contrast, the ten most frequently used words of the BPW_E are entirely different. This shows the high impact the above-described extension has.

Table 9

Ten most Frequent Positive words

BPW_O		BPW_N		BPW_E	
word	cum %	word	cum %	word	cum %
ertrag	2.06 %	ertrag	2.01 %	erträge	2.31 %
erreicht	3.79 %	erreicht	3.70 %	chancen	4.08 %
erfolg	5.50 %	erfolg	5.38 %	zusammen	5.63 %
zusammenarbeit	7.04 %	zusammenarbeit	6.88 %	wachstum	7.12 %
erfolgreich	8.56 %	erfolgreich	8.37 %	wert	8.59 %
erreichen	10.05 %	erreichen	9.82 %	führen	9.77 %
positiven	11.49 %	positiven	11.23 %	vermögens	10.85 %
positiv	12.92 %	positiv	12.62 %	bedeutung	11.74 %
positive	14.34 %	positive	14.01 %	sicherheit	12.62 %
möglichkeit	15.76 %	möglichkeit	15.40 %	aktiven	13.46 %

Note: We obtained frequencies from the complete sample of 6,275 annual reports.

Considering the most frequent negative words in Table 10, the main difference between BPW_O and BPW_N is the above-described correction of the word “betrug”, accounting for 2.36 % of all negative words. Due to the extension of the word list, the results for BPW_E show three new words accounting for 25.43 % of all negative words and therefore have a higher fraction than the ten most frequent words on the other lists. Due to their meaning, some words appear both on the lists of negative words and on the corresponding lists of negations. This is particularly clear in the case of the word “nicht”, which is the most frequently used negative word in the BPW_E dictionary. All duplications were checked and, in our view, represent both negations and words to be counted as negative.

Table 10

Ten most Frequent Negative words

BPW_O		BPW_N		BPW_E	
word	cum %	word	cum %	word	cum %
gegen	3.72 %	gegen	3.80 %	nicht	17.10 %
verpflichtungen	7.26 %	verpflichtungen	7.42 %	risiken	23.11 %
verluste	10.20 %	verluste	10.42 %	risiko	25.43 %
betrug	12.56 %	wertminderungen	12.51 %	gegen	27.57 %
wertminderungen	14.61 %	verfügung	14.47 %	verpflichtungen	29.61 %
verfügung	16.53 %	wertminderung	16.31 %	verluste	31.30 %
wertminderung	18.33 %	wertberichtigungen	17.99 %	wertminderungen	32.48 %
wertberichtigungen	19.97 %	ermittlung	19.65 %	verfügung	33.58 %
ermittlung	21.60 %	rückgang	21.29 %	wertminderung	34.62 %
rückgang	23.20 %	verpflichtung	22.90 %	wertberichtigungen	35.56 %

Note: We obtained the frequencies from the complete sample of 6,275 annual reports.

These findings are consistent with *Shapiro* et al. (2022), stating that apart from domain specificity, the size of the word list is important. A translation of the words used in Table 9 and 10 can be found in the appendix (Table 25).

To test the suitability of the three word lists, we apply the assumption of *Loughran/McDonald* (2011) that the value of sentiment has a direct impact on the particular dependent variable in Figure 1. Moreover, higher values in sentiment should lead to higher values in the dependent variables. All three word lists show different and ascending values for *FROA* and *FROE* in all quintiles. Therefore, the necessary assumptions can be considered as given for all three dictionaries.

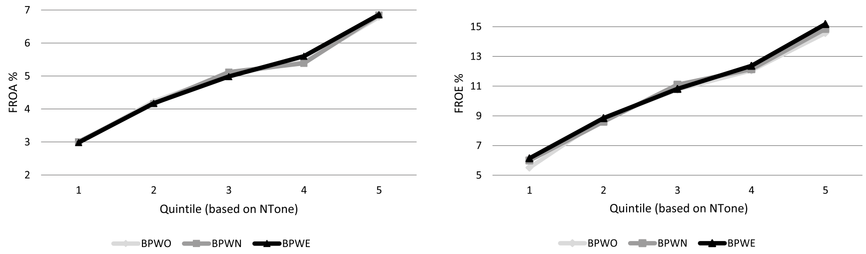

Figure 1: Dependent Variables by Quintile

Additionally, we conduct Kruskal-Wallis Tests for all six measurements shown in Figure 1. The tests show a statistically significant difference between the quintiles of each measurement. Detailed test statistics can be found in the appendix (Table 17).

In addition, we create two groups with above and below median *NTone*, to compare the average *FROA* and *FROE*. For every pair given in Figure 2, we perform an independent-samples t-test. All pairs are significantly different from one another. In addition to the given results for the BPW_E in Figure 2, we conduct the same tests for below and above measurement for BPW_O and BPW_N. These additional tests show that all pairs for all three word lists are significantly different. The results for all t-tests can be found in the appendix (Tables 18 to 20).

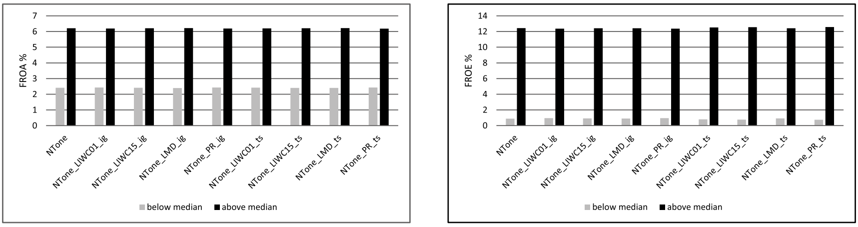

Figure 2: FROA and FROE Grouped by above and below Median Sentiment (BPW_E)

2. *Significance of Results*

Table 11 presents the results for the relation between the two dependent variables (future ROA and future ROE) and *NTone* for all three used dictionaries in a multivariate context, as described in section IV.2.

Table 11

Regression of NTone and the three Dictionaries (BPW_O, BPW_N, BPW_E)

	Dependent variable:					
	FROA (BPW_O) (1)	FROA (BPW_N) (2)	FROA (BPW_E) (3)	FROE (BPW_O) (4)	FROE (BPW_N) (5)	FROE (BPW_E) (6)
NTone	1.346	1.517	2.231*	10.397**	10.474**	15.987***
	(0.947)	(0.937)	(1.215)	(4.054)	(4.092)	(5.266)
AGE	−0.0002	−0.001	−0.001	−0.012	−0.013	−0.013
	(0.003)	(0.003)	(0.003)	(0.015)	(0.016)	(0.016)
LOSS	−1.427*	−1.377*	−1.356*	−20.234***	−20.266***	−20.041***
	(0.754)	(0.757)	(0.754)	(3.849)	(3.888)	(3.868)
LEV	0.076	0.076	0.079	−0.005	−0.003	0.018
	(0.057)	(0.057)	(0.058)	(0.947)	(0.946)	(0.949)
ROA	0.586***	0.590***	0.590***			
	(0.043)	(0.043)	(0.043)			
ROE				0.273***	0.272***	0.270***
				(0.071)	(0.071)	(0.072)
Constant	1.547	1.380	0.945	12.613	12.276	9.146
	(3.680)	(3.682)	(3.686)	(11.940)	(11.988)	(11.879)
Observations	4,168	4,112	4,116	4,168	4,112	4,116
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES	YES
R2	0.405	0.411	0.411	0.174	0.174	0.174
Adjusted R2	0.396	0.402	0.402	0.162	0.161	0.162
Residual Std. Error	9.299 (df = 4106)	9.222 (df = 4050)	9.216 (df = 4054)	39.692 (df = 4106)	39.846 (df = 4050)	39.811 (df = 4054)
F Statistic	45.819*** (df = 61; 4106)	46.360*** (df = 61; 4050)	46.425*** (df = 61; 4054)	14.164*** (df = 61; 4106)	13.939*** (df = 61; 4050)	14.015*** (df = 61; 4054)

Significance levels are based on robust standard errors (given in parantheses) and are indicated by
* p < 0.1; ** p < 0.05; *** p < 0.01.

The displayed results show a significant relationship between the dependent variables and *NTone* using the extended BPW dictionary (estimation (3) and (6)). Based on those findings, we can confirm the first hypothesis that further correcting and expanding the BPW dictionary improves its ability to forecast future ROAs and ROEs. This shows that the *NTone* of annual reports seems to contain relevant information for future ROAs and ROEs. An increase in *NTone* by the interquartile change of 0.181 for the BPW_E word lists leads to an increase of 40.38 % in *FROA* and 289.36 % in *FROE*. Similar relationships were also found while using the dictionaries *Henry* (2006, 2008) and *Ruscheinsky* et al. (2018) on English-speaking annual reports (*Henry* et al. 2021; *Koelbl* 2020). When analyzing conference calls *Druz* et al. (2020) stated that managers could possibly reveal information about future earnings through their usage of sentiment. Although this is a possible reason, we are unable to confirm such a relationship based on the given data.

Additionally, there is a highly significant relationship between the two dependent variables and the current parameters of those variables (*ROA* and *ROE*). The binary variable *LOSS* also shows a significant impact on *FROA* and *FROE*. These results are consistent with *Davis/Tama-Sweet* (2012), *Davis* et al. (2012), and *Henry* et al. (2021).

Table 12 and Table 13 display the results for using the four different negation lists separated for *FROA* and *FROE*, when using the BPW_E dictionary. The usage of the two LIWC negation lists and the PR negation list improves the significance of results for *FROA* when using the approach of term shifting negated words. The already highly significant results for *FROE* kept their level of significance when using negations. Therefore, we can confirm the second hypothesis that using negations further improves results. The other significant relationships regarding *ROA*, *ROE* and *LOSS* remain unchanged.

Table 12

Regression of NTone and FROA for BPW_E (term shift Negated words)

	Dependent variable:				
	FROA (7)	FROA (8)	FROA (9)	FROA (10)	FROA (11)
NTone	2.231*				
	(1.215)				
NTone_LIWC01_ts		2.739**			
		(1.255)			
NTone_LIWC15_ts			2.636**		
			(1.254)		
NTone_LMD_ts				2.294*	
				(1.235)	
NTone_PR_ts					2.619**
					(1.318)
AGE	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LOSS	−1.356*	−1.324*	−1.330*	−1.350*	−1.343*
	(0.754)	(0.752)	(0.752)	(0.753)	(0.754)
LEV	0.079	0.080	0.080	0.079	0.080
	(0.058)	(0.059)	(0.059)	(0.058)	(0.059)
ROA	0.590***	0.589***	0.589***	0.590***	0.589***
	(0.043)	(0.044)	(0.043)	(0.043)	(0.044)
Constant	0.945	0.891	0.905	0.933	0.919
	(3.686)	(3.688)	(3.690)	(3.687)	(3.669)
Observations	4,116	4,116	4,116	4,116	4,116
Year Fixed Effects	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES
R2	0.411	0.412	0.411	0.411	0.411
Adjusted R2	0.402	0.403	0.403	0.402	0.403
Residual Std. Error (df = 4054)	9.216	9.214	9.215	9.216	9.215
F Statistic (df = 61; 4054)	46.425***	46.476***	46.463***	46.430***	46.451***

Significance levels are based on robust standard errors (given in parantheses) and are indicated by
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 13

Regression of NTone and FROE for BPW_E (term shift Negated words)

	Dependent variable:				
	FROE (12)	FROE (13)	FROE (14)	FROE (15)	FROE (16)
NTone	15.987***				
	(5.266)				
NTone_LIWC01_ts		17.400***			
		(5.302)			
NTone_LIWC15_ts			17.053***		
			(5.297)		
NTone_LMD_ts				15.719***	
				(5.212)	
NTone_PR_ts					17.286***
					(5.375)
AGE	–0.013	–0.013	–0.013	–0.013	–0.013
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
LOSS	–20.041***	–19.915***	–19.941***	–20.045***	–20.011***
	(3.868)	(3.864)	(3.868)	(3.872)	(3.860)
LEV	0.018	0.025	0.025	0.020	0.028
	(0.949)	(0.949)	(0.949)	(0.949)	(0.950)
ROE	0.270***	0.270***	0.270***	0.270***	0.270***
	(0.072)	(0.071)	(0.071)	(0.072)	(0.071)
Constant	9.146	8.999	9.053	9.146	9.109
	(11.879)	(11.883)	(11.904)	(11.894)	(11.773)
Observations	4,116	4,116	4,116	4,116	4,116
Year Fixed Effects	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES
R2	0.174	0.174	0.174	0.174	0.174
Adjusted R2	0.162	0.162	0.162	0.162	0.162
Residual Std. Error (df = 4054)	39.811	39.803	39.806	39.813	39.808
F Statistic (df = 61; 4054)	14.015***	14.045***	14.035***	14.005***	14.025***

Significance levels are based on robust standard errors (given in parantheses) and are indicated by
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

We also performed all regression using BPW_E and negations with the approach of ignoring the negated words. Only the usage of the LIWC 2001 negations was able to improve results. Due to the minor level of improvement compared to term shifting, the results are given in the appendix (Table 21 and 22) and are not discussed further.

Additional proof of the importance of implementing negations is given in in Table 14 and Table 15. When using BPW_N, negations and the approach of term shifting, the levels of significance are equal to the usage of the superior BPW_E word lists. These results underline the importance of implementing negations. Based on the results visible, the word list PR, developed specifically for the financial context, should be used. These results are consistent with *Shapiro* et al. (2022) and their findings, which claim that using negations improves the prediction of human sentiment ratings.

Table 14

Regression of NTone and FROA for BPW_N (term shift Negated words)

	Dependent variable:				
	FROA (17)	FROA (18)	FROA (19)	FROA (20)	FROA (21)
NTone	1.517				
	(0.937)				
NTone_LIWC01_ts		1.977**			
		(0.965)			
NTone_LIWC15_ts			2.028**		
			(0.996)		
NTone_LMD_ts				1.635*	
				(0.970)	
NTone_PR_ts					2.168**
					(1.051)
AGE	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LOSS	−1.377*	−1.345*	−1.340*	−1.369*	−1.347*
	(0.757)	(0.756)	(0.755)	(0.757)	(0.756)
LEV	0.076	0.076	0.077	0.077	0.077
	(0.057)	(0.057)	(0.058)	(0.058)	(0.058)
ROA	0.590***	0.590***	0.590***	0.590***	0.590***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Constant	1.380	1.410	1.381	1.367	1.321
	(3.682)	(3.673)	(3.677)	(3.684)	(3.674)
Observations	4,112	4,112	4,112	4,112	4,112
Year Fixed Effects	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES
R2	0.411	0.411	0.411	0.411	0.411
Adjusted R2	0.402	0.403	0.403	0.402	0.403
Residual Std. Error (df = 4050)	9.222	9.219	9.219	9.221	9.219
F Statistic (df = 61; 4050)	46.360***	46.413***	46.413***	46.369***	46.421***

Significance levels are based on robust standard errors (given in parantheses) and are indicated by
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 15

Regression of NTone and FROE for BPW_N (term shift Negated words)

	Dependent variable:				
	FROE (22)	FROE (23)	FROE (24)	FROE (25)	FROE (26)
NTone	10.474**				
	(4.092)				
NTone_LIWC01_ts		11.706***			
		(4.169)			
NTone_LIWC15_ts			11.697***		
			(4.241)		
NTone_LMD_ts				10.715***	
				(4.156)	
NTone_PR_ts					11.961***
					(4.255)
AGE	–0.013	–0.013	–0.013	–0.013	–0.013
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
LOSS	–20.266***	–20.149***	–20.154***	–20.247***	–20.218***
	(3.888)	(3.883)	(3.889)	(3.891)	(3.876)
LEV	–0.003	–0.0001	0.005	0.002	0.005
	(0.946)	(0.945)	(0.946)	(0.947)	(0.946)
ROE	0.272***	0.271***	0.271***	0.272***	0.271***
	(0.071)	(0.071)	(0.071)	(0.071)	(0.071)
Constant	12.276	12.258	12.057	12.124	11.681
	(11.988)	(11.914)	(11.932)	(11.999)	(11.905)
Observations	4,112	4,112	4,112	4,112	4,112
Year Fixed Effects	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES
R2	0.174	0.174	0.174	0.173	0.174
Adjusted R2	0.161	0.161	0.161	0.161	0.161
Residual Std. Error (df = 4050)	39.846	39.839	39.841	39.846	39.843
F Statistic (df = 61; 4050)	13.939***	13.966***	13.957***	13.937***	13.950***

Significance levels are based on robust standard errors (given in parantheses) and are indicated by
* p < 0.1; ** p < 0.05; *** p < 0.01.

The improvement shown above also partially applies when using the approach of ignoring negated words. The relevant tables are given in the appendix (Tables 23 and 24).

VI. Conclusion

This paper uses the dictionary-based approach to compute the sentiment of German-speaking annual reports. Due to the novelty of the used dictionary, the aim of this paper is to improve the given BPW_O and BPW_N word lists by further correction and expansion. Additionally, we test the use of different negations to further improve the results.

The expansion of the BPW_N word lists leads to an immense increase in total words found (positive: 157 %, negative: 77 %). Additionally, the ten most frequent positive and negative words found underwent an enormous change. This leads to a significant change in NTone calculated by using BPW_E. Despite the fundamental alteration, we successfully test basic assumptions visually and statistically. By using the new and extended BPW_E, we are able to improve regression results compared to the two previous versions and therefore confirm the first hypothesis. Additionally, we can show that negations should be implemented because they are able to improve results. A deterioration of results caused by the usage of negations could not be observed and should therefore be implemented in the form of term shifted PR negations.

Furthermore, by successfully improving the second version of the BPW dictionary and testing the implementation of negations, this paper contributes immensely to the existing literature on analyzing German corporate disclosures.

Due to this successful improvement of the BPW dictionary, further research on finance related texts should be conducted by using the BPW_E. Based on the novelty of this dictionary, other types of corporate disclosure should be analyzed, and a comparison to general German dictionaries should be conducted.

Appendix

Table 16

Description of Variables

Variable	Description
AGE	Age of the Company: Difference between the year of observation and the date of incorporation
FROA	Future Return on Assets: Return on Assets (ROA) one year ahead
FROE	Future Return on Equity: Return on Equity (ROE) one year ahead
LEV	Leverage: Sum of non-current liabilities and current liabilities, divided by shareholders funds
LOSS	LOSS equals one if the Profit and Loss before tax is negative, zero otherwise
NTone	Net Tone: Difference between the number of positive and negative words, divided by the sum of positive and negative words
ROA	Current Return on Assets: Profit and Loss before tax divided by total assets times 100
ROE	Current Return on Equity: Profit and Loss before tax divided by shareholders funds times 100

Table 17

Kruskal-Wallis test Statistics

	FROA			FROE		
	BPW_O	BPW_N	BPW_E	BPW_O	BPW_N	BPW_E
Kruskal-Wallis-H	207.201	210.450	249.461	242.842	240.057	256.486
df	4	4	4	4	4	4
Asymp. Sig.	< .001	< .001	< .001	< .001	< .001	< .001

Table 18

Independent Samples t-test for below and above Median Sentiment (BPW_O)

Dependent variable	Sentiment measure	Statistics	df	p	Mean below	Mean above
FROA	NTone	–8.377	4072	<.001	2.763	5.844
FROE	NTone	–7.975	3187	<.001	1.277	11.910

Table 19

Independent Samples t-test for below and above Median Sentiment (BPW_N)

Dependent variable	Sentiment measure	Statistics	df	p	Mean below	Mean above
FROA	NTone	–9.516	3992	<.001	2.558	6.060
	NTone_LIWC01_ig	–10.890	3879	<.001	2.312	6.307
	NTone_LIWC15_ig	–10.300	3939	<.001	2.417	6.201
	NTone_LMD_ig	–9.506	4009	<.001	2.560	6.059
	NTone_PR_ig	–10.080	3929	<.001	2.456	6.162
	NTone_LIWC01_ts	–11.120	3878	<.001	2.271	6.348
	NTone_LIWC15_ts	–10.800	3914	<.001	2.328	6.291
	NTone_LMD_ts	–9.363	3996	<.001	2.586	6.033
	NTone_PR_ts	–9.846	3931	<.001	2.499	6.120
FROE	NTone	–8.312	3140	<.001	1.066	12.250
	NTone_LIWC01_ig	–9.025	3024	<.001	0.595	12.720
	NTone_LIWC15_ig	–8.760	3029	<.001	0.770	12.550
	NTone_LMD_ig	–8.230	3153	<.001	1.120	12.200
	NTone_PR_ig	–8.669	3026	<.001	0.830	12.490
	NTone_LIWC01_ts	–9.123	3020	<.001	0.530	12.790
	NTone_LIWC15_ts	–8.996	3019	<.001	0.614	12.700
	NTone_LMD_ts	–8.341	3080	<.001	1.047	12.270
	NTone_PR_ts	–8.400	3015	<.001	1.008	12.310

Table 20

Independent Samples t-test for below and above Median Sentiment (BPW_E)

Dependent variable	Sentiment measure	Statistics	df	p	Mean below	Mean above
FROA	NTone	−10.350	3875	<.001	2.411	6.209
	NTone_LIWC01_ig	−10.240	3877	<.001	2.430	6.190
	NTone_LIWC15_ig	−10.350	3874	<.001	2.411	6.208
	NTone_LMD_ig	−10.410	3874	<.001	2.400	6.220
	NTone_PR_ig	−10.240	3873	<.001	2.430	6.190
	NTone_LIWC01_ts	−10.300	3852	<.001	2.419	6.201
	NTone_LIWC15_ts	−10.380	3853	<.001	2.406	6.213
	NTone_LMD_ts	−10.390	3884	<.001	2.403	6.217
	NTone_PR_ts	−10.230	3872	<.001	2.433	6.187
FROE	NTone	−8.607	3105	<.001	0.878	12.440
	NTone_LIWC01_ig	−8.489	3102	<.001	0.956	12.360
	NTone_LIWC15_ig	−8.554	3100	<.001	0.912	12.410
	NTone_LMD_ig	−8.581	3104	<.001	0.895	12.420
	NTone_PR_ig	−8.501	3099	<.001	0.947	12.370
	NTone_LIWC01_ts	−8.728	3016	<.001	0.798	12.520
	NTone_LIWC15_ts	−8.787	3016	<.001	0.759	12.560
	NTone_LMD_ts	−8.553	3103	<.001	0.913	12.410
	NTone_PR_ts	−8.798	3008	<.001	0.752	12.570

Table 21

Regression of NTone and FROA for BPW_E (Ignore Negated words)

	Dependent variable:				
	FROA (27)	FROA (28)	FROA (29)	FROA (30)	FROA (31)
NTone	2.231*				
	(1.215)				
NTone_LIWC01_ig		2.392**			
		(1.212)			
NTone_LIWC15_ig			2.322*		
			(1.208)		
NTone_LMD_ig				2.228*	
				(1.214)	
NTone_PR_ig					2.216*
					(1.202)
AGE	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LOSS	−1.356*	−1.340*	−1.344*	−1.353*	−1.351*
	(0.754)	(0.753)	(0.753)	(0.754)	(0.754)
LEV	0.079	0.079	0.079	0.079	0.079
	(0.058)	(0.058)	(0.058)	(0.058)	(0.058)
ROA	0.590***	0.589***	0.589***	0.590***	0.589***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Constant	0.945	0.917	0.925	0.939	0.928
	(3.686)	(3.687)	(3.688)	(3.686)	(3.679)
Observations	4,116	4,116	4,116	4,116	4,116
Year Fixed Effects	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES
R2	0.411	0.411	0.411	0.411	0.411
Adjusted R2	0.402	0.403	0.402	0.402	0.402
Residual Std. Error (df = 4054)	9.216	9.215	9.216	9.216	9.216
F Statistic (df = 61; 4054)	46.425***	46.447***	46.439***	46.427***	46.431***

Significance levels are based on robust standard errors (given in parantheses) and are indicated by * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 22

Regression of NTone and FROE for BPW_E (Ignore Negated words)

	Dependent variable:				
	FROE (32)	FROE (33)	FROE (34)	FROE (35)	FROE (36)
NTone	15.987***				
	(5.266)				
NTone_LIWC01_ig		16.281***			
		(5.184)			
NTone_LIWC15_ig			16.014***		
			(5.173)		
NTone_LMD_ig				15.707***	
				(5.205)	
NTone_PR_ig					15.547***
					(5.048)
AGE	–0.013	–0.013	–0.013	–0.013	–0.013
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
LOSS	–20.041***	–19.967***	–19.984***	–20.037***	–20.020***
	(3.868)	(3.866)	(3.868)	(3.869)	(3.864)
LEV	0.018	0.022	0.022	0.019	0.023
	(0.949)	(0.949)	(0.949)	(0.949)	(0.949)
ROE	0.270***	0.270***	0.270***	0.270***	0.270***
	(0.072)	(0.072)	(0.072)	(0.072)	(0.072)
Constant	9.146	9.038	9.074	9.133	9.065
	(11.879)	(11.879)	(11.889)	(11.885)	(11.827)
Observations	4,116	4,116	4,116	4,116	4,116
Year Fixed Effects	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES
R2	0.174	0.174	0.174	0.174	0.174
Adjusted R2	0.162	0.162	0.162	0.162	0.162
Residual Std. Error (df = 4054)	39.811	39.807	39.808	39.812	39.810
F Statistic (df = 61; 4054)	14.015***	14.032***	14.025***	14.012***	14.019***

Significance levels are based on robust standard errors (given in parantheses) and are indicated by
* p < 0.1; ** p < 0.05; *** p < 0.01.

Table 23

Regression of NTone and FROA for BPW_N (Ignore Negated words)

	Dependent variable:				
	FROA (37)	FROA (38)	FROA (39)	FROA (40)	FROA (41)
NTone	1.517				
	(0.937)				
NTone_LIWC01_ig		1.648*			
		(0.929)			
NTone_LIWC15_ig			1.624*		
			(0.934)		
NTone_LMD_ig				1.527	
				(0.942)	
NTone_PR_ig					1.638*
					(0.931)
AGE	−0.001	−0.001	−0.001	−0.001	−0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
LOSS	−1.377*	−1.363*	−1.363*	−1.375*	−1.365*
	(0.757)	(0.757)	(0.757)	(0.757)	(0.757)
LEV	0.076	0.076	0.077	0.076	0.077
	(0.057)	(0.057)	(0.058)	(0.058)	(0.058)
ROA	0.590***	0.590***	0.590***	0.590***	0.590***
	(0.043)	(0.043)	(0.043)	(0.043)	(0.043)
Constant	1.380	1.394	1.380	1.373	1.368
	(3.682)	(3.677)	(3.680)	(3.683)	(3.677)
Observations	4,112	4,112	4,112	4,112	4,112
Year Fixed Effects	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES
R2	0.411	0.411	0.411	0.411	0.411
Adjusted R2	0.402	0.402	0.402	0.402	0.402
Residual Std. Error (df = 4050)	9.222	9.221	9.221	9.222	9.221
F Statistic (df = 61; 4050)	46.360***	46.379***	46.377***	46.362***	46.381***

Significance levels are based on robust standard errors (given in parantheses) and are indicated by
* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 24

Regression of NTone and FROE for BPW_N (Ignore Negated words)

	Dependent variable:				
	FROE (42)	FROE (43)	FROE (44)	FROE (45)	FROE (46)
NTone	10.474** (4.092)				
NTone_LIWC01_ig		10.706*** (4.032)			
NTone_LIWC15_ig			10.595*** (4.066)		
NTone_LMD_ig				10.446** (4.115)	
NTone_PR_ig					10.383*** (3.956)
AGE	–0.013 (0.016)	–0.013 (0.016)	–0.013 (0.016)	–0.013 (0.016)	–0.013 (0.016)
LOSS	–20.266*** (3.888)	–20.207*** (3.886)	–20.205*** (3.888)	–20.253*** (3.888)	–20.231*** (3.883)
LEV	–0.003 (0.946)	–0.001 (0.946)	0.001 (0.946)	–0.001 (0.947)	0.001 (0.946)
ROE	0.272*** (0.071)	0.271*** (0.071)	0.271*** (0.071)	0.272*** (0.071)	0.271*** (0.071)
Constant	12.276 (11.988)	12.280 (11.951)	12.202 (11.961)	12.210 (11.995)	12.087 (11.995)
Observations	4,112	4,112	4,112	4,112	4,112
Year Fixed Effects	YES	YES	YES	YES	YES
Industry Fixed Effects	YES	YES	YES	YES	YES
R2	0.174	0.174	0.174	0.174	0.174
Adjusted R2	0.161	0.161	0.161	0.161	0.161
Residual Std. Error (df = 4050)	39.846	39.843	39.843	39.846	39.844
F Statistic (df = 61; 4050)	13.939***	13.951***	13.949***	13.940***	13.945***

Significance levels are based on robust standard errors (given in parantheses) and are indicated by
* p < 0.1; ** p < 0.05; *** p < 0.01

Table 25

Translation of ten most Frequent words (all three Dictionaries)

Positive words		Negative words	
German	English	German	English
chancen	chances	betrug	fraud, amounted
erfolg	success	ermittlung	investigation
erfolgreich	successful	gegen	against
erreichen	achieve	nicht	not
erreicht	achieved	risiken	risks
ertrag	return, revenue	risiko	risk
erträge	returns, revenues	rückgang	decline
führen	lead	verfügung	decree
positiven	positive	verluste	losses
vermögens	assets	verpflichtung	obligation
wachstum	growth	verpflichtungen	obligations
wert	value	wertberichtigungen	value adjustments
zusammen	together	wertminderung	impairment
zusammenarbeit	cooperation	wertminderungen	impairments

References

Algaba, A./*Ardia*, D./*Bluteau*, K./*Borms*, S./*Boudt*, K. (2020): Econometrics meets sentiment: An overview of methodology and applications. Journal of Economic Surveys, Vol. 34(3), 512 – 547.

Allee, K. D./*Deangelis*, M. D. (2015): The Structure of Voluntary Disclosure Narratives: Evidence from Tone Dispersion. Journal of Accounting Research, Vol. 53(2), 241 – 274.

Alshorman, S. A. A./*Shanahan*, M. (2022): The voice of profit: exploring the tone of Australian CEO‘s letters to shareholders after the global financial crisis. Corporate Communications: An International Journal, Vol. 27(1), 127 – 147.

Aly, D./*El-Halaby*, S./*Hussainey*, K. (2018): Tone disclosure and financial performance: evidence from Egypt. Accounting Research Journal, Vol. 31(1), 63 – 74.

Bannier, C. E./*Pauls*, T./*Walter*, A. 2017): CEO-Speeches and Stock Returns.

Bannier, C. E./*Pauls*, T./*Walter*, A. (2019a): Content analysis of business specific text documents: Introducing a German dictionary. Journal of Business Economics, Vol. 89(1), 79 – 123.

Bannier, C. E./*Pauls*, T./*Walter*, A. (2019b): The Annual General Meeting revisited: The role of the CEO speech.

Berns, J./*Bick*, P./*Flugum*, R./*Houston*, R. (2022): Do changes in MD&A section tone predict investment behavior? Financial Review, Vol. 57(1), 129 – 153.

Bochkay, K./*Hales*, J./*Chava*, S. (2020): Hyperbole or Reality? Investor Response to Extreme Language in Earnings Conference Calls. The Accounting Review, Vol. 95(2), 31 – 60.

Borochin, P. A./*Cicon*, J. E./*DeLisle*, R. J./*McKay Price*, S. (2018): The effects of conference call tones on market perceptions of value uncertainty. Journal of Financial Markets, Vol. 40, 75 – 91.

Boudt, K./*Thewissen*, J. (2019): Jockeying for Position in CEO Letters: Impression Management and Sentiment Analytics. Financial Management, Vol. 48(1).

Brau, J. C./*Cicon*, J./*McQueen*, G. (2016): Soft Strategic Information and IPO Underpricing. Journal of Behavioral Finance, Vol. 17(1), 1 – 17.

Bushman, R. M./*Hendricks*, B. E./*Williams*, C. D. (2016): Bank Competition: Measurement, Decision-Making, and Risk-Taking. Journal of Accounting Research, Vol. 54(3), 777 – 826.

Chakraborty, B./*Bhattacharjee*, T. (2020): A review on textual analysis of corporate disclosure according to the evolution of different automated methods. Journal of Financial Reporting and Accounting, Vol. 18(4), 757 – 777.

Correa, R./*Garud*, K./*Londono*, J. M./*Mislang*, N. (2021): Sentiment in Central Banks' Financial Stability Reports. Review of Finance, Vol. 25(1), 85 – 120.

Da Tonin, J. M. F./*Scherer*, L. M. (2022): Market reaction to the tones of earnings conference calls. Journal of Business Management, Vol. 62(1), 1 – 18.

Daniel, F./*Lohrke*, F. T./*Fornaciari*, C. J./*Turner*, R. (2004): Slack resources and firm performance: a meta-analysis. Journal of Business Research, Vol. 57(6), 565 – 574.

Davis, A. K./*Piger*, J. M./*Sedor*, L. M. (2012): Beyond the Numbers: Measuring the Information Content of Earnings Press Release Language. Contemporary Accounting Research, Vol. 29(3), 845 – 868.

Davis, A. K./*Tama-Sweet*, I. (2012): Managers' Use of Language Across Alternative Disclosure Outlets: Earnings Press Releases versus MD&A. Contemporary Accounting Research, Vol. 29(3), 804 – 837.

Druz, M./*Petzev*, I./*Wagner*, A. F./*Zeckhauser*, R. J. (2020): When Managers Change Their Tone, Analysts and Investors Change Their Tune. Financial Analysts Journal, Vol. 76(2), 47 – 69.

Feldman, R./*Govindaraj*, S./*Livnat*, J./*Segal*, B. (2008): The Incremental Information Content of Tone Change in Management Discussion and Analysis.

Ferguson, N. J./*Philip*, D./*Lam*, H. Y. T./*Guo*, J. M. (2015): Media Content and Stock Returns: The Predictive Power of Press. Multinational Finance Journal, Vol. 19(1), 1 – 31.

Ferris, S. P./*Hao*, Q./*Liao*, M.-Y. (2013): The Effect of Issuer Conservatism on IPO Pricing and Performance. Review of Finance, Vol. 17(3), 993 – 1027.

Frankel, R./*Jennings*, J./*Lee*, J. (2022): Disclosure Sentiment: Machine Learning vs. Dictionary Methods. Management Science, Vol. 68(7), 5514 – 55.

Fritz, D./*Tows*, E. (2018): Text Mining and Reporting Quality in German Banks: A Cooccurrence and Sentiment Analysis. Universal Journal of Accounting and Finance, Vol. 6(2), 54 – 81.

Gentzkow, M./*Kelly*, B./*Taddy*, M. (2019): Text as Data. Journal of Economic Literature, Vol. 57(3), 535 – 574.

González, M./*Guzmán*, A./*Téllez*, D. Fernando/*Trujillo*, M. Andrea (2019): What you say and how you say it: Information disclosure in Latin American firms. Journal of Business Research, Vol. 127(3), 427 – 443.

Hart, R. P. (2000): DICTION 5.0.

Henry, E. (2006): Market Reaction to Verbal Components of Earnings Press Releases: Event Study Using a Predictive Algorithm. Journal of Emerging Technologies in Accounting, Vol. 3, 1 – 19.

Henry, E. (2008): Are Investors Influenced By How Earnings Press Releases Are Written? Journal of Business Communication, Vol. 45(4), 363 – 407.

Henry, E./*Leone*, A. J. (2016): Measuring Qualitative Information in Capital Markets Research: Comparison of Alternative Methodologies to Measure Disclosure Tone. The Accounting Review, Vol. 91(1), 153 – 178.

Henry, E./*Thewissen*, J./*Torsin*, W. (2021): International Earnings Announcements: Tone, Forward-looking Statements, and Informativeness. European Accounting Review, 1 – 35.

Hillert, A./*Jacobs*, H./*Müller*, S. (2018): Journalist disagreement. Journal of Financial Markets, Vol. 41, 57 – 76.

Huang, X./*Teoh*, S. H./*Zhang*, Y. (2014): Tone Management. The Accounting Review, Vol. 89(3), 1083 – 1113.

Iqbal, J./*Riaz*, K. (2022): Predicting future financial performance of banks from management's tone in the textual disclosures. Quality & Quantity, (56), 2691 – 2721.

Jandl, J.-O./*Feuerriegel*, S./*Neumann*, D. (2014): Long- and Short-Term Impact of News Messages on House Prices: A Comparative Study of Spain and the United States, Auckland.

Jegadeesh, N./*Wu*, D. (2013): Word power: A new approach for content analysis. Journal of Financial Economics, Vol. 110(3), 712 – 729.

Kang, Y./*Cai*, Z./*Tan*, C.-W./*Huang*, Q./*Liu*, H. (2020): Natural language processing (NLP) in management research: A literature review. Journal of Management Analytics, Vol. 7(2), 139 – 172.

Kang, T./*Park*, D.-H./*Han*, I. (2018): Beyond the numbers: The effect of 10-K tone on firms' performance predictions using text analytics. Telematics and Informatics, Vol. 35(2), 370 – 381.

Kaya, D./*Maier*, C./*Böhmer*, T. (2020): Empirische Kapitalmarktforschung zu Conference Calls: Eine Literaturanalyse. Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, Vol. 72, 183 – 212.

Kaye, J./*Druin*, A./*Lampe*, C./*Morris*, D./*Hourcade*, J. P. (Hg.)): Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, New York, NY, USA 2016.

Kearney, C./*Liu*, S. (2014): Textual Sentiment in Finance: A Survey of Methods and Models. International Review of Financial Analysis, (33), 171 – 185.

King, D. R./*Dalton*, D. R./*Daily*, C. M./*Covin*, J. G. (2004): Meta-analyses of post-acquisition performance: indications of unidentified moderators. Strategic Management Journal, Vol. 25(2), 187 – 200.

Koelbl, M. (2020): Is the MD&A of US REITs informative? A textual sentiment study. Journal of Property Investment & Finance, Vol. 38(3), 181 – 201.

Krause, J./*Perer*, A./*Ng*, K. (2016): Interacting with Predictions, in: Kaye, Jofish/Druin, Allison/Lampe, Cliff/Morris, Dan/Hourcade, Juan Pablo (Hrsg.): Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, New York, NY, USA 2016, 5686 – 5697.

Lewis, C./*Young*, S. (2019): Fad or future? Automated analysis of financial text and its implications for corporate reporting. Accounting and Business Research, Vol. 49(5), 587 – 615.

Li, F. (2010): Textual Analysis of Corporate Disclosures: A Survey of the Literature. Journal of Accounting Literature, (29), 143 – 165.

Loughran, T./*McDonald*, B. (2011): When Is a Liability Not a Liability?: Textual Analysis, Dictionaries, and 10-Ks. The Journal of Finance, Vol. 66(1), 35 – 65.

Loughran, T./*McDonald*, B. (2015): The Use of Word Lists in Textual Analysis. Journal of Behavioral Finance, Vol. 16(1), 1 – 11.

Loughran, T./*McDonald*, B. (2016): Textual Analysis in Accounting and Finance: A Survey. Journal of Accounting Research, (54), 1187 – 1230.

Luo, Y./*Zhou*, L. (2020): Textual tone in corporate financial disclosures: a survey of the literature. International Journal of Disclosure and Governance, Vol. 17, 101 – 110.

Meier, T./*Boyd*, R. L./*Pennebaker*, J. W./*Mehl*, M. R./*Martin*, M./*Wolf*, M./*Horn*, A. B. (2018): "LIWC auf Deutsch": The Development, Psychometrics, and Introduction of DE-LIWC2015.

Mishev, K./*Gjorgjevikj*, A./*Vodenska*, I./*Chitkushev*, L. T./*Trajanov*, D. (2020): Evaluation of Sentiment Analysis in Finance: From Lexicons to Transformers. IEEE Access, Vol. 8, 131662 – 131682.

Myšková, R./*Hájek*, P. (2020): Mining risk-related sentiment in Corporate annual reports and its effect on financial performance. Technological and Economic Development of Economy, Vol. 26(6), 1422 – 1443.

Patelli, L./*Pedrini*, M. (2014): Is the Optimism in CEO's Letters to Shareholders Sincere? Impression Management Versus Communicative Action During the Economic Crisis. Journal of Business Ethics, Vol. 124, 19 – 34.

Pengnate, S./*Lehmberg*, D. G./*Tangpong*, C. (2020): Top management‘s communication in economic crisis and the firm‘s subsequent performance: sentiment analysis approach. Corporate Communications: An International Journal, Vol. 25(2), 187 – 205.

Picault, M./*Renault*, T. (2017): Words are not all created equal: A new measure of ECB communication. Journal of International Money and Finance, (79), 136 – 156.

Pöferlein, M. (2021): Sentiment Analysis of German Texts in Finance: Improving and Testing the BPW Dictionary. Journal of Banking and Financial Economics, Vol. 16(2), 5 – 24.

Price, M. S./*Doran*, J. S./*Peterson*, D. R./*Bliss*, B. A. (2012): Earnings conference calls and stock returns: The incremental informativeness of textual tone. Journal of Banking & Finance, Vol. 36(4), 992 – 1011.

Remus, R./*Quasthoff*, U./*Heyer*, G. (2010): SentiWS – a Publicly Available German-language Resource for Sentiment Analysis. Proceedings of the 7th International Language Ressources and Evaluation (LREC'10), 1168 – 1171.

Renault, T. (2017): Intraday online investor sentiment and return patterns in the U.S. stock market. Journal of Banking & Finance, Vol. 84, 25 – 40.

Rice, D. R./*Zorn*, C. (2019): Corpus-based dictionaries for sentiment analysis of specialized vocabularies. Political Science Research and Methods, Vol. 67, 1 – 16.

Röder, F./*Walter*, A. (2019): What drives Investment Flows into Social Trading Portfolios? The Journal of Financial Research, Vol. 42(2), 383 – 411.

Ruscheinsky, J. R./*Lang*, M./*Schäfers*, W. (2018): Real estate media sentiment through textual analysis. Journal of Property Investment & Finance, Vol. 36(5), 410 – 428.

Schmeling, M./*Wagner*, C. (2016): Does Central Bank Tone Move Asset Prices?

Shapiro, A. H./*Sudhof*, M./*Wilson*, D. J. (2022): Measuring news sentiment. Journal of Econometrics, Vol. 228(2), 221 – 243.

Stone, P. J./*Dunphy*, D. C./*Smith*, M. S./*Ogilvie*, D. M. (1966): The General Inquirer: A Computer Approach to Content Analysis. The M.I.T. Press, Cambridge, Massachusetts, London.

Taboada, M./*Brooke*, J./*Tofiloski*, M./*Voll*, K./*Stede*, M. (2011): Lexicon-Based Methods for Sentiment Analysis. Computational Linguistics, Vol. 37(2), 267 – 307.

Tetlock, P. C. (2007): Giving Content to Investor Sentiment: The Role of Media in the Stock Market. The Journal of Finance, Vol. 62(3), 1139 – 1168.

Tetlock, P. C./*Saar-Tsechansky*, M./*Macskassy*, S. (2008): More Than Words: Quantifying Language to Measure Firms' Fundamentals. The Journal of Finance, Vol. 63(3), 1437 – 1467.

Tillmann, P./*Walter*, A. (2018): ECB vs Bundesbank: Diverging Tones and policy Effectiveness.

Tillmann, P./*Walter*, A. (2019): The effect of diverging communication: The case of the ECB and the Bundesbank. Economics Letters, Vol. 176©, 68 – 74.

Universität Leipzig (2022): Deutscher Wortschatz: Korpus News 2020. Accessed: 15.04.2022, https://corpora.uni-leipzig.de/de?corpusId=deu_news_2020.

Vojinović, Ž./*Milutinović*, S./*Leković*, B. (2020): Micro-specific Profitability Factors of the Serbian Insurance Industry: A Panel Data Estimation. Ekonomie a Management, Vol. 23(1), 135 – 155.

Wolf, M./*Horn*, A. B./*Mehl*, M. R./*Haug*, S./*Pennebaker*, J. W./*Kordy*, H. (2008): Computergestützte quantitative Textanalyse: Äquivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count. Diagnostica, Vol. 54(2), 85 – 98.

Yuthas, K./*Rogers*, R./*Dillard*, J. F. (2002): Communicative Action and Corporate Annual Reports. Journal of Business Ethics, Vol. 41(1-2), 141 – 157.