Journal of Language and Social Psychology
2019, Vol. 38(2) 170–193

DOI: 10.1177/0261927X18808562
journals.sagepub.com/home/jls


The Relative Power of Negativity: The Influence of Language Intensity on Perceived Strength

Christine Liebrecht[1], Lettica Hustinx[2], and Margot van Mulken[2]

Abstract

Negative utterances and words have been found to be stronger than positive utterances and words, but what happens if positive and negative utterances are intensified? Two online experiments were carried out in which participants judged the strength of (un)intensified positive and negative evaluations in written dialogues. Both studies showed intensified language was perceived as stronger than unmarked language (i.e., language that was not intensified), and negative evaluations were stronger than positive evaluations. What is more, intensification and polarity interact; the increment of perceived strength for intensified positive adjectives (Study 1) and purely intensified adverbs (*really, very*; Study 2) was bigger than the increment in perceived strength of intensified negative adjective and adverbs. When a meaningful intensifier (*deliciously, disgustingly*) was used, the negativity effect remained. The findings were discussed within cognitive frameworks such as relevance theory, theory of mind, and theory on verbal aggression.

Keywords

polarity, language intensity, stylistics, Pollyanna principle, negativity bias, word-of-mouth

Many consumers who seek information about a hotel's or a restaurant's services and products read online reviews or other evaluations via social media (electronic

[1]Tilburg University, Tilburg, Netherlands
[2]Radboud University Nijmegen, Nijmegen, Netherlands

Corresponding Author:
Christine Liebrecht, Department of Communication and Cognition, Tilburg University, Warandelaan 2, 5037 AB Tilburg, Netherlands.
Email: c.c.liebrecht@tilburguniversity.edu

word-of-mouth, eWOM; Hennig-Thurau, Gwinner, Walsch, & Gremler, 2004). Depending on their positive or negative expressions, these messages will affect consumers' attitudes, behavioral intentions, and (buying) behaviors (e.g., Ha, 2004; Reichheld & Schefter, 2000; Ward & Lee, 2000).

Negative and positive messages seem to have different effects. Previous studies showed negative evaluations had a greater impact on people than positive evaluations did. For example, if consumers read a set of online reviews containing equal amounts of positive and negative reviews, their attitude toward the reviewed topic was more negative (e.g., Purnawirawan, De Pelsmacker, & Dens, 2015). The imbalance in relative strength of negativity and positivity was observed not only for eWOM but also for other fields such as journalism (e.g., Soroka, 2006), organizational communication (e.g., Kahneman & Tversky, 1984), and (political) advertising (e.g., Newhagen & Reeves, 1991). Although the majority of the findings indicates negativity outweighs positivity, some experimental studies and meta-analyses nuanced the robustness of this *negativity effect*, showing that inter alia setting (e.g., consumer environments: Ahluwalia, 2002; political advertising: Allen & Burrell, 2002) and textual factors (e.g., type of information: Skowronski & Carlston, 1989; message framing: Block & Keller, 1995; O'Keefe & Jensen, 2006) could reduce the impact of negativity.

In their meta-analysis concerning gain and loss frames, O'Keefe and Jensen (2006) discussed linguistic examples of prior framing studies and observed a mixture of multiple concepts in the experimental materials that might have affected experimental outcomes. We observed another influencing factor as well: Operationalization of the experimental materials was not equally balanced with respect to positivity and negativity. Items such as *a healthy heart* or *healthy skin* are not true antonyms of *a heart disease* or *skin cancer* (O'Keefe & Jensen, 2006).

Based on O'Keefe and Jensen (2006), we suggest the large variety of language use in prior experimental materials is an alternative explanatory factor of the equivocal findings of the negativity effect. In order to examine this assumption, we focused on the role of language use in negative and positive messages by using true antonyms in the experimental materials. Moreover, we also investigated whether an intensification of language affected the perception of negative and positive messages. These so-called language intensifiers are known as linguistic elements that strengthen evaluative utterances (e.g., Liebrecht, 2015), and can enhance attitude change (e.g., M. A. Hamilton & Hunter, 1998; M. A. Hamilton, Hunter, & Burgoon, 1990).

Evaluations containing language intensifiers are generally perceived as stronger than evaluations without these elements (i.e., unmarked evaluations; e.g., J. W. Bowers, 1963; Burgers & De Graaf, 2013, Study 2; Liebrecht, 2015). These elements might affect the perception of the strength of negative or positive evaluations. Since we know messages containing either language intensity or negativity are perceived as stronger than unmarked evaluations or positive evaluations, we investigate whether the combination of negativity and intensification has a cumulative effect, or whether the intensification affects polarity differently, in the sense that intensification moderates the effect of polarity.

To our knowledge, the combination of the two phenomena has not yet been systematically investigated. Because research material is generally not fully included in the method sections or appendices of the reported experiments, it is difficult to discern a clear pattern of the effects of polarity and intensity in previous work (Liebrecht, 2015). Moreover, the ambiguous findings of both factors in prior research require more attention. We anticipate that language intensity will shed a new light on the negativity effect. *Bad* is stronger than *good*, but is *worse* equally stronger than *better*? Is there still an imbalance in the perception of the strength of negative and positive evaluations if intensifiers are used? This article focuses on the role of polarity and intensification and, more specifically, on evaluative words and their true antonyms and the intensified equivalents of these words.

Understanding the Negativity Effect

The negativity effect can be explained with the help of cognitive psychological research that showed negative utterances are perceived as stronger than positive ones (e.g., Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; D. L. Hamilton & Huffman, 1971; Wyer, 1974). Negative utterances get more attention, arouse more emotions, have more influence on recipient behavior, and are stored better and longer in memory (Baumeister et al., 2001; Dijksterhuis & Aarts, 2003; Lagerwerf, Boeynaems, van Egmond-Brussee, & Burgers, 2015; Peeters & Czapinski, 1990; Pratto & John, 1991).

Two theories explain the strong effects of negativity. First, negative messages are powerful for evolutionary reasons. Because negativity is often associated with fear or danger, and positivity with security and safety, at a cognitive level, a person automatically pays more attention to unpleasant (negative) than to pleasant (positive) information. This psychological phenomenon is called the Negativity bias (Jing-Schmidt, 2007; Parkinson, 1995; Pratto & John, 1991).

Second, negative messages are perceived as stronger at a social level. The Pollyanna principle states there is a universal human tendency to use and expect positive evaluative words (Boucher & Osgood, 1969; Jing-Schmidt, 2007). Language users expect a positive message because of their learned behavior based on conventions such as politeness and *face* protection (Brown & Levinson, 1987). According to Taylor (1991), people have a universal tendency to see and tell things from the bright side of life. Even when people receive a negative message, they will try to weaken it or to make something positive out of it (Taylor, 1991). However, when the message appears to be irrefutably negative, then the violation of the (positive) expectation makes the (negative) message more intense. Since a positive message is default, the negative message is marked and, therefore, perceived as stronger. This so-called contrast effect was already described by Skowronski and Carlston (1989). Because the stronger effects of negativity can be hypothesized with both the Negativity bias and the Pollyanna principle, we consider these two theories to be two sides of the same coin.

The Strength of Language Intensity

An evaluation can range from (moderately) positive (*a good hotel*) to strongly positive (*a great hotel*) or from (moderately) negative (*a bad hotel*) to strongly negative (*a terrible hotel*). An evaluation is the linguistic expression of personal feelings, attitudes, or value judgments toward an object, situation, or process (Thompson & Hunston, 2000). The stylistic devices that strengthen evaluations are language intensifiers.

The definition of language intensity is somewhat problematic. J. W. Bowers (1963) is often cited as the founder of language intensity research. Language intensity is defined by him as "the quality of language which indicates the degree to which the speaker's attitude toward a concept deviates from neutrality" (p. 345). This definition would require us to define neutrality. Typical of intensified language, however, is that it occurs in evaluative speech. Evaluative language cannot be neutral, nor is every evaluative utterance intense (for an extensive discussion, see Van Mulken & Schellens, 2012). Furthermore, an evaluation is intensified if it contains an element that adds strength to it. The strength of an utterance is the force of the impact an utterance has on the perceiver, either in a positive direction or in a negative direction.

A better definition is suggested by Liebrecht (2015), who claims a linguistic element in an evaluation is an intensifier when the omission or replacement of the same element results in a less strong evaluation. Intensification can be achieved in many ways: Words can be replaced (*great* instead of *good*), words can be added (*extraordinarily good* instead of *good*), or stylistic figures (e.g., metaphors, *He drowned in a sea of grief* instead of *He was very sad*) and typographical elements (*nice!!!* instead of *nice*) can be used (Renkema, 1997). Language intensifiers change the stylistic strength (cf. M. A. Hamilton & Hunter, 1998) of an evaluation.

The effect of language intensity on recipients has often been studied by means of experiments (for an overview, see Liebrecht, 2015; Van Mulken & Schellens, 2012). The results were not unequivocal. In some studies, language intensity seemed to increase the clarity and credibility of the message (M. A. Hamilton et al., 1990) with a positive effect on the attitude of the receiver (e.g., Buller et al., 2000; M. A. Hamilton & Stewart, 1993; Rogan & Hammer, 1998). In other experiments, no—or opposite—differences were found between intensified and unintensified evaluations (e.g., J. W. Bowers, 1963; Burgers & De Graaf, 2013, Study 1; Burgoon & Chase, 1973; Hornikx, Pieper, & Schellens, 2008).

There are three explanations for these mixed findings. The first main explanation relates to the definition and operationalization of language intensity (for an extensive discussion, see Liebrecht, Hustinx, Van Mulken, & Schellens, 2016; Van Mulken & Schellens, 2012). Based on J. W. Bowers's (1963) definition, scholars operationalized the concept in their experimental materials differently. The experimental texts differed in evaluative nature and genre (such as news articles; e.g., Burgers & De Graaf, 2013, or advertisements; e.g., Hornikx et al., 2008). Moreover, the intensified texts were compared with texts containing less strong evaluations (e.g., *ill-considered* instead of *stupid*; cf. M. A. Hamilton & Hunter, 1998), texts with extensifiers (e.g., *absolutely* instead of *probably*; cf. Burgers & De Graaf, 2013, Study 2), or texts without

evaluations (e.g., *these essential questions* instead of *these questions*; cf. Anderson & Blackburn, 2004). The variety of evaluation levels in the experimental texts may account for the ambiguous findings.

Regarding the operationalization of language intensity, it appeared the manipulation of intensity was not always concerned with only the *stylistic strength* of the utterance but sometimes affected the *content* of the utterances as well. We acknowledge, for example, the *death* of a person is a more forceful and impactful event than the *injury* of a person (cf. M. A. Hamilton et al., 1990). The modification of *injury* into *death* is what we call a content-related modification. The difference between *good* and *fantastic* is what we would call a stylistic intensification. In order to investigate the effects of language intensity more systematically, we will only focus on stylistic intensity instead of content-related strength.

The second explanation concerns the influence of other variables scholars included in their studies. In particular, Hamilton and colleagues addressed the effects of language intensity in combination with other variables such as source credibility (e.g., M. A. Hamilton & Hunter, 1998; M. A. Hamilton et al., 1990), attitude discrepancy between source and language user (e.g., Aune & Kikuchi, 1993; M. A. Hamilton et al., 1990; McEwen & Greenberg, 1970), and gender (Burgoon, Jones, & Stewart, 1975; M.A. Hamilton & Stewart, 1993). All these variations in language intensity research might be accountable for the mixed effects; therefore, we propose to investigate the effects of language intensity more systematically by excluding confounding factors.

The third explanation for the mixed findings in language intensity research could be related to the polarity of the texts. Although the valence of the texts has never been taken into account as a factor in experimental studies, and although the amount of studies that reported the valence of the experimental materials is low, in studies where positive messages were used (e.g., arguing for a statement), persuasive effects of language intensity were more often found than in studies where negative messages were used (e.g., arguing against a statement; Liebrecht, 2015). In addition, the meta-analysis of M. A. Hamilton and Hunter (1998) indicated that findings of language intensity across studies were homogeneous in both high-discrepancy messages (i.e., more negative from the receiver's point of view) and in low-discrepancy messages (i.e., more positive from the receiver's point of view). These findings might indicate polarity plays a significant role in the effects of language intensity since recipients react differently to positive messages than they do to negative messages. To put it differently, it is worthwhile to investigate whether message polarity interacts with the *intensity* of the evaluation.

Polarity and Stylistic Strength

Literature in which the relationship between polarity and stylistic strength is addressed is scarce and varies in methodological quality. In linguistics, scholars have mainly investigated the perceived strength of (English) single words with perception studies in which participants assessed positive and negative words on Likert-type scales. These studies showed rankings from the most positive to the most negative words and

indicate that the most negative words were perceived more negatively than the positive words were perceived as positive (e.g., Feldman, 1966; Jones & Thurstone, 1955). Furthermore, various lists have been constructed to show the effects of words or pictures on factors such as valence (i.e., polarity) and arousal (see the *Affective Norms for English Words* by Bradley & Lang, 1999 and the *International Affective Picture System* by Lang, Bradley, & Cuthbert, 1997). Such scores can be used to compose standardized wordlists with strength values (Warriner, Kuperman, & Brysbaert, 2013) that have already been applied in the field of (automatic) sentiment analysis. In this field, the sentiment of a message is most often based on the calculation of the sentiment scores of individual (English) words in wordlists (e.g., De Rijke et al., 2013; Na, Lee, Nam, & Lee, 2009). However, the reliability of crowd-sourced approaches is highly questionable, and these measures did not take into account the contexts in which these words appeared.

Evidence for the effect of language use on polarity was found in research involving questionnaires. Kamoen, Holleman, and Van den Bergh (2007) studied wording effects in surveys. They found that people were more hesitant to endorse a negatively formulated question (such as *How boring was this text?*) than a positively formulated question (*How interesting was this text?*). Subsequent studies also showed that the participants were more likely to disagree with negatively formulated questions and statements than to agree with the positive counterparts (Kamoen, Holleman, Mak, Sanders, & Van den Bergh, 2011). These results indicated the participants perceived the negatively formulated questions as more negative than the positively formulated questions as positive. In these studies, the unmarked negative and positive evaluative words were taken into account (*boring/interesting*); the effect of language intensity (*mind-numbing/fascinating*) was left out of consideration.

In this article, we aim to explore the effects of polarity and language intensity. Will the positive–negative asymmetry remain, or will language intensity nullify the negativity effect? It is possible that adding language intensity to negative and positive evaluations will not affect the negativity effect at all, which would mean both phenomena simply have a cumulative effect. Put statistically, we would then expect to find only two main effects. On the other hand, polarity and intensity may also interact. This can be explained by cognitive frameworks such as relevance theory (Sperber & Wilson, 1994) and theory of mind (Lieberman, 2013); both describe the ability to attribute mental states to others.

Relevance theory (Sperber & Wilson, 1994) argues that utterances convey a number of conscious and unconscious implicatures that receivers should decode to understand their communicative intentions correctly. Messages are, therefore, only perceived to be relevant when the sender manifestly marks them as having a communicative intention, thus being *ostensive* (Carston, 2008; Tendahl, 2009). Following this theory, language intensifiers could be seen as ostensiveness markers. The addition of intensifiers could acquire more attention in positive evaluations than in negative evaluations, because positivity is the default. Deviation from the default means either negativity or intensified positivity. Therefore, when using an intensified positive message, the sender manifestly signals the positivity of the message, whereas

the use of an intensified negative message calls attention to the intensity of the message. Since unmarked negative messages are perceived to be stronger than unmarked positive messages, the impact of positive intensity will be relatively greater than the impact of negative intensity.

A similar expectation can be formulated with the help of theory of mind, a cognitive theory that states people are able to attribute mental states such as beliefs, desires, and intentions to themselves and others, and understand them (e.g., Lieberman, 2013). For language use, it can be reasoned that language users choose their words consciously in order to ensure the recipient will interpret the message correctly (in other words, they *mentalize* how recipients might interpret the message; Lieberman, 2013). This is especially so if the sender wants to share a sincere positive evaluation in a social context where positivity is the default; here, the message must be strengthened in order to mark the deviation from the default. Furthermore, following the principles of theory of mind and relevance theory, recipients also expect language users will try to communicate their messages as clearly as possible. A recipient will perceive the usage of intensifiers as intentional language use—especially in positive evaluations—as a marker that the positive evaluation deviates from the default. This might enhance the perceived strength of intensified positivity as more than that of intensified negativity. Our line of reasoning is also in accordance with language expectancy theory (Burgoon, Denning, & Roberts, 2002) and defaultness theory (Giora, Givoni, & Fein, 2015).

Hypotheses

In order to investigate the perceived strength of intensified and unmarked negative and positive evaluations, we conducted two online experiments in which we specifically focused on the perception of strength of evaluations. In order to enhance the generalizability of our findings, we operationalized language intensity in two different ways.

First, we verified whether negative evaluations were indeed perceived to be stronger than positive evaluations. Since the majority of prior studies confirm this (e.g., Baumeister et al., 2001; Feldman, 1966; D. L. Hamilton & Huffman, 1971; Jones & Thurstone, 1955; Wyer, 1974), we expect that negative polarity will be perceived as stronger than positive polarity (Hypothesis 1 [H1]).

Second, we checked whether intensified language was indeed perceived as stronger than unmarked evaluative language. Scholars assume the underlying goal of intensification is to enhance the strength of a standpoint (e.g., Aune & Kikuchi, 1993; Renkema, 1997; Van Mulken & Schellens, 2012). In our study, it is verified this is indeed the case. The hypothesis, therefore, is intensified evaluations, on both the positive and negative sides, will be perceived as stronger than unmarked positive and negative evaluations (H2).

Third, based on relevance theory and theory of mind, we expect an interaction effect of polarity and language intensity. The interaction will show intensification enhances the perceived strength of positive evaluations more than the perceived strength of negative evaluations (H3).

Study 1

Method

The first study focused on the intensification of adjectives in evaluations in conversational discourse. Participants judged the perceived strength of intensified and unmarked positive and negative evaluative utterances in Dutch.

Design. The experiment had a 2 × 2 within-participant design; intensification (unmarked/intensifier) was crossed with polarity (positive/negative). Each participant was exposed to all four conditions. For each condition, five different evaluative domains were used: Taste, Beauty, Quality, Interest, and Intelligence. The 20 adjectives were presented twice to the participants resulting in 40 different dialogues to assess per participant. Four versions of the questionnaire were constructed based on a balanced Latin square design. The order of the items differed per questionnaire.

Materials and Pretest. Since we wanted to examine the difference in strength of positive and negative intensifiers systematically, we used antonyms in order to optimize the different scalar positions (intensified negative—unmarked negative—unmarked positive —intensified positive). For example: if a person evaluates the beauty of a painting, he or she can be either positive or negative. The painting can be called *pretty* (positive) or *ugly* (negative antonym). These positive or negative adjectives can be intensified by other adjectives, such as *wonderful* (intensified positive) or *horrible* (intensified negative).

In order to select the adjectives for Study 1, two pretests were carried out; the first was designed for the selection of unmarked positive adjectives to set the baseline. It was intended that the baseline adjectives be of comparable strength. For instance, the strength of *clever* (domain Intelligence) should be relatively comparable to *pretty* (domain Beauty). For each domain, we compiled a list of approximately 20 different adjectives based on dictionaries and Internet listings. Participants ($N = 21$) rated the perceived strength of various adjectives on an 11-point Likert-type scale (comparable to the study of Jones & Thurstone, 1955). The pretest resulted in five moderately positive adjectives whose perceived strength was comparable (mean strength about 2.30, $M_{min} = 1.90$ and $M_{max} = 2.52$ on a 5-point scale) to the evaluation domains: *mooi* [*pretty*] (domain Beauty), *slim* [*clever*] (Intelligence), *lekker* [*nice*] (Taste), *interessant* [*interesting*] (Interest), and *goed* [*good*] (Quality).

A second pretest was carried out in order to select antonyms and intensifiers of the unmarked positive adjectives. Students in Dutch linguistics ($N = 22$) were asked to write down the most suitable antonym of the unmarked positive adjectives. In addition, they had to produce two intensifiers—the strongest possible single-word intensification—for the positive adjectives. In order to ensure the produced antonyms were true antonyms, we carried out the same test with the most common unmarked negative adjectives. In this test, other students in linguistics ($N = 16$) were asked to produce the most suitable antonym and two intensifiers of the unmarked negative adjectives *lelijk*

[*ugly*], *dom* [*stupid*], *vies* [*nasty*], *saai* [*boring*], and *slecht* [*bad*]. Finally, the most frequently mentioned unmarked and intensifying positive and negative adjectives were selected for the experiment. The intensified positive items were (in order of domains): *prachtig* [*beautiful*], *geniaal* [*brilliant*], *heerlijk* [*delicious*], *boeiend* [*fascinating*], and *geweldig* [*great*]; the intensified negative adjectives were: *afschuwelijk* [*horrible*], *achterlijk* [*retarded*], *smerig* [*revolting*], *slaapverwekkend* [*mind-numbing*], and *vreselijk* [*terrible*].

In short, a 2 × 2 factor (Intensity: unmarked/intensifier; Polarity: positive/negative) repeated measures mixed design was used. Each participant judged 40 dialogues for five evaluation domains: Beauty, Intelligence, Taste, Interest, and Quality. They thus evaluated two instances of each combination of factors. An additional practice conversation preceded the experimental dialogues.

Following Rokeach (1968), we used dialogues to set a context for the adjectives. In a dialogue, person A asked person B's opinion on various topics (referring to the five domains). The answer of person B was operationalized with one of the four conditions.

In order to avoid gender-effects (cf. Burgoon et al., 1975), we did not use first names but referred to the dialogue partners with the expression *person A* and *person B*. By choosing this way of referring to dialogue partners, we were able to avoid differences in source credibility (cf. M. A. Hamilton et al., 1990; M. A. Hamilton & Hunter, 1998). In the dialogue, we made sure person A did not advocate his or her own opinion or attitude, that he or she did not try to alter the opinion of person B, but simply asked person B to give an opinion about an event, object, or situation that B had recently experienced. The participant had to consider this dialogue and assess only the evaluative response of person B. Consequently, no possible discrepancy between existing attitudes of persons A and B could interfere with the participants' judgments (cf. M. A. Hamilton & Hunter, 1998). An example of the experimental materials is shown below.

A: Heb je al eens bij Da Vinci gegeten?

[Have you ever eaten in the Da Vinci restaurant?]

B: Toevallig ben ik er vorig jaar geweest. Het eten was *smerig/vies/lekker/heerlijk.*

[I happened to go there last year. The food was *revolting/nasty/nice/delicious.*]

Participants. A total of 93 Dutch participants took part in the online experiment: 13 of them were left out; 6 participants appeared to be nonnative Dutch speakers; and 7 participants answered *yes* when asked whether they were interrupted during the questionnaire. The remaining 80 participants had an average age of 19.4 years ($SD = 2.19$). More women (79%) than men (21%) participated. The participants were students in the first year of their education at the faculty of Arts at a Dutch university.

Instrumentation. Each of the 40 items was followed by a 21-point Likert-type scale, ranging from very negative (−10) via neutral (0) to very positive (+10). To investigate the perceived strength of each item, the participants were asked to rate the strength of the expressed opinion of person B on the scale ("According to you, how strongly worded is the opinion of B?" *hoe krachtig vind je de mening van persoon B verwoord?*). The scale itself was unnumbered and quite long in order to encourage the participants to score the items based on their first impression, and not to compare the scores of prior answers. At the end of the questionnaire, personal data were asked: gender, age, education level, and mother tongue.

Procedure. The participants were invited by a research assistant during lectures at the faculty of Arts at a Dutch university and via e-mail. The participants took part voluntarily by clicking on the hyperlink, which randomly assigned each participant to one of the four versions of the questionnaire in the software program *NetQuestionnaire*. About half of the respondents received no reward, and half of the group (38 persons) received a scratch card (for the value of 1 euro) as a stimulus to participate in the study.[1] The whole procedure took 12 to 15 minutes.

Data Analysis. We used the R statistical package (R Core Team, 2014). Mixed-model analyses were used with participants and items as random variables. A mixed-model analysis is a statistical technique that offers the advantage of directly assessing the significance of predictors as well as the interaction of other factors with these predictors (Baayen, Davidson, & Bates, 2008; Quené & Van den Bergh, 2008). Moreover, it offers the researcher the possibility to include covariates and random slopes. Visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality. We also report R^2 scores, which are an approximation of explained variance for mixed models and can be directly compared to η^2 measures for analyses of variance. The use of this measure is common in regression analyses: It helps interpret the importance of effect sizes (Nakagawa & Schielzeth, 2013).

All our models were checked to see if the addition of Gender or Age improved the quality of the model. This was not the case. Polarity and intensity were chosen as within-participants fixed factors. In the statistical models reported below, we, therefore, included only fixed and random predictors that were significant. Both models were controlled for the within-subjects nature of the task by including random effects for items with a variance components covariance structure and restricted maximum likelihood estimation.

Results

First, we verified whether participants used the strength scale correctly. Since the scale is unconventional, and since it is not inconceivable that a person would interpret the strength of a negative word in positive terms, we anticipated mistakes in the polarity of the strength. However, in only a few cases, participants rated a positive item with a negative score, or vice versa. We removed 188 of the total 3,019 data points (6.23%).

Table 1. Deviance Values (−2 Log likelihood) and Degrees of Freedom (*df*) for Mixed-Model Analyses in Study 1.

Deviance	Null model (*df*)	Complete model (*df*)	χ^2 test
Perceived strength	13180 (4)	12620 (7)	$\chi^2(3) = 560$*

*$p < .001$.

Table 2. Mean Scores on Perceived Strength Broken Down for Polarity and Intensity.

	Positive		Negative	
Polarity/intensity	*M*	*SD*	*M*	*SD*
Unmarked	4.31	2.14	5.15	2.32
Intensifier	6.41	2.31	6.84	2.34

Note. We used the absolute values of the reported scores on the 21-point Likert-type scale from −10 to +10 to answer the research question.

Table 3. Statistical Model for Perceived Strength in Study 1.

Fixed effects	*b*	*t* (*df*)	*p*
Intercept	5.13	29.68 (142.8)	<.001
Polarity	−.81	−7.83 (2887.2)	<.001
Intensity	1.67	15.86 (2885.4)	<.001
Intensity * Polarity	0.42	3.03 (2892.2)	<.001

Note. df = degrees of freedom. The Intercept represents a negative unmarked adjective. Effect size R^2 for fixed effects with standardized generalized variance = 0.29 (based on Nakagawa & Schielzeth, 2013).

Analysis of the removed values revealed some participants had more missing values than others, probably indicating they had more difficulty with the scale. For items, the missing values were more evenly distributed.

Information criteria were calculated between an empty model with only the random factors and the complete model including all factors. Table 1 presents the information criteria for this model with a χ^2 test to test whether decreases are significant. Table 2 shows the mean scores on Perceived Strength broken down for Polarity and Intensity.

We included Polarity, Intensity, and the interaction of both as predictors in the model, with participants and items as random factor. R^2 for this model was .29.

Table 3 shows the final model for Perceived Strength in adjectives. Polarity significantly predicted Perceived Strength, $b = -.81$; $t(2887.2) = -7.83$, $p < .001$. Negative polarity was perceived as more forceful than positive polarity. This is what was expected. Intensity also significantly predicted Perceived Strength, $b = 1.67$, $t(2885.4) = 15.86$, $p < .001$. As expected, an increase in intensity was also perceived as

stronger. It shows the effect for intensity was twice as strong as the effect of polarity. There was also a small ordinal interaction effect of Polarity and Intensity, $b = .42$, $t(2892.2) = 3.03$, $p < .001$, supporting our hypothesis: The increment in strength for intensified positive adjectives was more substantial than for intensified negative adjectives.

Discussion

Negative evaluations were perceived to be stronger than positive evaluations; intensified evaluations were perceived to be stronger than unmarked evaluations; and this effect was perceived to be twice as strong. So far so good. However, we also found some evidence that intensifiers add more strength to evaluations with positive adjectives than to evaluations with negative adjectives. Thus, the effect of language intensity seems to be somewhat greater on the positive side than on the negative side.

However, although our materials were thoroughly selected with the help of pretests, a caveat should be taken into account: Intensifiers vary in semantic meaning (Liebrecht, Hustinx, Van Mulken, & Schellens, 2018). We cannot close our eyes to the fact that an adjective such as *brilliant* also adds luminosity to the evaluation or an intensifier such as *excellent* adds outstandingness. These semantic modifications, although small, may have influenced the outcome of our study. Therefore, a second study was carried out in which the meaningfulness of intensifiers was taken into account. For this, we selected adverbs instead of adjectives. In this word class, a clear distinction can be made between pure adverbial intensifiers (*very*) that do not contain any semantic meaning, and meaningful adverbial intensifiers (*wonderfully*), that do contain semantic meaning (e.g., "stuff to wonder about"). Furthermore, by examining adverbial intensifiers instead of adjectives, our findings in Study 1 can be replicated for other modes of intensification.

More important, by making a distinction in the semantic modifications of intensifiers, we can investigate more deeply how deviation from the default works. Adverbs allow us to keep the grammatical category constant and at the same time permit us to stipulate whether it is indeed intensity that moderates negativity. If the third hypothesis also holds true for adverbs, then it certainly must hold true for pure intensifiers, where only intensification is added to the equation.

Study 2

The first hypotheses in Study 2 remained the same as in Study 1: Negative evaluations are perceived to be stronger than positive ones (H1), and intensification leads to increased perceived strength (H2). We also expected intensifiers that contain semantic meaning to have a stronger effect than intensifiers that only intensify (pure intensifiers). To put it differently: an adverb such as *very* or *really* will be perceived as less powerful than adverbs such as *disgustingly*. Therefore, we predict meaningful intensifiers are perceived as stronger than pure intensifiers (H2a).

We expected to again find the incremental impact of positive intensification is larger than the effect of negative intensification. We, therefore, hypothesized that negative evaluations with pure intensifiers are perceived to be lower in strength than positive evaluations with pure intensifiers (H3a) and that meaningful intensifiers in negative evaluations are perceived as less strong than meaningful intensifiers in positive evaluations (H3b).

Method

Design. Study 2 had a 3 × 2 within-participant design, with intensification (unmarked/pure intensifier/meaningful intensifier) and polarity (positive/negative) as factors. Each participant saw all six conditions in all five evaluative domains. So, in total, each participant rated 30 items. Six versions of the questionnaire were constructed based on a balanced Latin square. The questionnaires differed in order of the items.

Materials and Pretest. The materials and written dialogues of Study 1 served as the basis for the materials of Study 2. However, the factor intensification was operationalized differently. Instead of adjectives, adverbs were used to intensify the unmarked evaluations of Study 1. For example, an unmarked positive adjective was strengthened with the addition of the pure intensifier *zeer* [*very*] or the meaningful intensifier *verrukkelijk* [*deliciously*] (producing "*verrukkelijk lekker [deliciously nice*]").

The meaningfulness of the intensifying adverbs was pretested in a manner comparable to the first pretest in Study 1. Participants ($N = 27$) rated the meaningfulness of positive and negative intensified evaluations that differed in semantic modification (e.g., *het eten was lekker* [*the meal was nice*] and was intensified with e.g., *zeer/enorm/verrukkelijk* [*very/tremendously/deliciously*]) on 7-point Likert-type scales (3 items, $\alpha = .86$). Based on the results, the meaningful and pure intensifiers were selected for the experiment.

Participants. Participants who took less than 5 minutes to complete the questionnaire were removed from the sample; 91 participants remained. However, one participant was a nonnative speaker of Dutch, and was, therefore, removed from the sample as well. The remaining 90 Dutch participants had an average age of 29.8 years ($SD = 13.34$) and 78% of them were female. Almost all participants were students or had already obtained a bachelor's or master's degree (98.9%).

Instrumentation. Each item was followed by a 21-point scale, ranging from very negative (−10) via neutral (0) to very positive (+10). For this study, a slider was used (unavailable for Study 1). The same operationalization of perceived strength was used. At the end of the questionnaire, personal data were requested (gender, age, level of education, mother tongue).

Procedure. Participants voluntarily enrolled in the experiment after their attention was drawn during lectures at the faculty of Humanities at a Dutch university (a different

Table 4. Information Criteria (2 Restricted Log Likelihood) for Mixed-Model Analyses in Study 2.

Deviance	Null model (*df*)	Complete model (*df*)	χ^2 test
Perceived strength	11728 (4)	10548 (9)	$\chi^2(5) = 1193$*

Note. *df* = degrees of freedom.
*$p < .001$.

Table 5. Mean Scores on Perceived Strength Broken Down for Polarity and Type of Intensifier in Study 2.

	Positive		Negative	
Polarity/type of intensifier	*M*	*SD*	*M*	*SD*
Unmarked	3.59	2.01	4.64	2.49
Pure intensifier	5.73	2.63	6.56	2.65
Meaningful intensifier	7.14	2.88	8.59	2.58

Note. *df* = degrees of freedom. We used the absolute values of the reported scores on the 21-point Likert-type scale from −10 to +10 to answer the research question.

university than Study 1) and via social media. By clicking on a hyperlink, the participant was randomly assigned to one of the six versions of the questionnaire. Answering all questions took from 12 to 15 minutes.

Results

Before testing the hypotheses, we verified whether participants used the strength scale correctly. Unfamiliar with the rating scale, some participants rated a positive item with a negative score, or vice versa. We removed 271 of the total 2,700 data points (10.04%). Similar to Study 1, analysis of the removed values revealed some participants had more missing values then others, but for items the missing values were more evenly distributed.

In order to test our hypotheses for Study 2, information criteria were calculated between an empty model with only participant and items as random factors, and the complete model including all factors (see Table 4). Table 5 shows the mean scores on Perceived Strength broken down for Polarity and Type of Intensifier (levels: Unmarked, Pure Intensifier, Meaningful Intensifier).

The full model (Table 6) included Polarity, Type of Intensifier, and the interaction of Type of Intensifier and Polarity as predictors, items, and participants as random factor. R^2 for this model was .76.

Polarity significantly predicted perceived strength, $b = -1.33$, $t(540.2) = 8.07$, $p < .001$. Again, as expected, negative polarity was perceived as stronger than positive polarity. Type of Intensifier also significantly predicted Perceived Strength. Pairwise

Table 6. Statistical Model for Perceived Strength in Study 2.

Fixed effects	*b*	*t* (*df*)	*p*
Intercept	8.38	38.06 (159.3)	<.001
Type of intensifier (pure)	−2.08	−14.48 (2333)	<.001
Type of intensifier (unmarked)	−3.9	−27.68 (2330.4)	<.001
Polarity	−1.33	−8.07 (540.2)	<.001
Pure Intensifier * Polarity	0.63	3.22 (2332.3)	.001
Meaningful Intensifier * Polarity	0.3	1.53 (2327.9)	n.s.

Note. *df* = degrees of freedom. The intercept represents a negative meaningful evaluation. Effect size R^2 for fixed effects with standardized generalized variance = .76 (based on Nakagawa & Schielzeth, 2013).

comparisons (Bonferroni) revealed all three types of intensifiers differed significantly from each other, and evaluations without intensifiers were perceived as less strong than evaluations with meaningful and pure intensifiers. In turn, pure intensifiers were found to be less strong than meaningful intensifiers: unmarked intensifiers, $b = -3.9$, $t(2330.4) = 27.68$, $p < .001$; pure intensifiers, $b = -2.08$, $t(2333) = -14.48$, $p < .001$. Again, there was also a small interaction effect of polarity and type of intensifier. Inspection (Bonferroni) revealed only positive pure intensifiers differed significantly from negative pure intensifiers, $b = 0.63$, $t(2332.3) = 3.22$, $p = .001$). The increase in experienced strength for positive pure intensifiers compared to the increase in strength of the negative pure intensifiers was slightly more substantial.

Discussion

The findings of Study 2 supported and clarified the findings of Study 1. In all cases, negative evaluations were perceived to be stronger than positive evaluations, and unmarked evaluations were less strong than intensified evaluations. We again found a small interaction effect. Intensification affected the effect of polarity but only for pure intensifiers. The discrepancy between the strength of negative and positive utterances was slightly moderated by intensification: The increment of perceived strength for purely intensifying adverbs was bigger than the increment in perceived strength of intensifying negative adverbs. This is in accordance with H3a. Meaningful intensification, on the other hand, did not show this expected pattern; therefore, we found no support for H3b in the current study. In the General Discussion section, we will elaborate on these findings.

General Discussion

In this article, we reported two experimental studies in which we measured the perceived strength of positive and negative evaluations in written dialogues. Our studies confirmed the negativity effect, as predicted by the negativity bias and the Pollyanna principle (e.g., Jing-Schmidt, 2007; Parkinson, 1995; Pratto & John,

1991), and the intensification effect (J. W. Bowers, 1963; Burgers & De Graaf, Study 2; Liebrecht, 2015).

We also found the negativity effect can be modified by intensification. Although the effect was small, on two occasions we found intensification affects the effect of polarity. We found this in Study 1, with intensified adjectives; we found it again in Study 2, with intensified adverbs, although only for pure intensifiers (such as *very* or *really*). The increment of perceived strength for intensified positive adjectives and purely intensified adverbs was bigger than the increment in perceived strength of intensified negative adjectives and adverbs. Cognitive theories such as relevance theory (Sperber & Wilson, 1994) or the theory of mind (e.g., Lieberman, 2013) may explain these findings. Since an unmarked positive message is the default, intensification of such an unmarked message comes across as stronger than an intensified negative message when compared to an unmarked negative message. Positive intensification stands out.

However, pure intensification is rare. In most cases, the words that are used to intensify contain more meaning than only intensification (Liebrecht et al., 2018). When words like *really, very, extremely*, and *quite* are placed in front of an adjective or adverb, it makes the meaning of that adjective or adverb more intense or more powerful. The same is true for some types of typography: Adding exclamation marks or using capital letters only has the function of intensifying the message. Most words do more than that: They not only intensify the message but also contain extra meaning, and this entails a content-related modification. Words like *brilliant* or *dazzling* add a shining quality to what is good, and *glorious* and *magnificent* add grandeur to what is good. Every meaningful intensifier contains extra meaning. Since the extra meaning differs per word, this may explain why meaningful positive intensification does not outperform meaningful negative intensification. Thus, the increase in perceived strength of negative evaluations with a meaningful intensifier (*disastrously bad*) is as large as the increase in strength of positive evaluations with a meaningful intensifier (*extraordinarily good*).

In short, meaningful intensifiers have an additional function. It can be assumed that meaningful intensifiers reattract the recipient's attention to the polarity that overrules the ostensiveness of intensifiers. As a result, the language user signals primarily the polarity of the evaluation, which activates the principles of the Negativity bias again. We can conclude intensifiers are able to mark a sender's deviation from the default, although the strengthening effects seem to depend on the semantic modification.

Our findings bridge two existing views on language intensity. On the one hand, language intensity is shaped by stylistic features employed by the speaker. Words like *very* or *delicious* mark the unmarked evaluation *nice* (e.g., Bolinger, 1972; Fletcher, 1980). Aune and Kikuchi (1993) call this *actual language intensity*. On the other hand, language intensity is a matter of perception, which was the dependent variable of the current study. Aune and Kikuchi (1993) call this *perceived language intensity*. In accordance with our theoretically based operationalization (actual intensity), the intensified evaluations were perceived as stronger than the unintensified evaluations (perceived intensity). The question is whether these two perspectives can be viewed

separately. Based on theory of mind, we suggest that senders might strategically use intensifiers to ensure recipients will interpret the message correctly, and the recipients would expect language users to do so as well.

The findings are based on evaluations containing true antonyms. We intentionally excluded other variables that were investigated extensively in prior language intensity research, as we first wanted to focus on and determinate the language factor. To our mind, the effects of source credibility (e.g., M. A. Hamilton et al., 1990; M. A. Hamilton & Hunter, 1998), attitude discrepancy between source and language user (e.g., Aune & Kikuchi, 1993; M. A. Hamilton et al., 1990; McEwen & Greenberg, 1970), and gender (Burgoon et al., 1975; M. A. Hamilton & Stewart, 1993) can now be investigated in combination with language intensity and polarity.

We also found the strength of intensifiers cannot be based on standardized word lists with strength values such as the ones that are used with sentiment analysis (e.g., Warriner et al., 2013). The perspective on both polarity and language intensity is often lacking in previous research on language intensity (e.g., Aune & Kikuchi, 1993; Renkema, 1997; Van Mulken & Schellens, 2012), or on the strength of positive and negative words (e.g., Feldman, 1966; Jones & Thurstone, 1955). The same is true for disciplines such as eWOM (e.g., Ha, 2004; Purnawirawan et al., 2015; Ward & Lee, 2000), journalism (e.g., Soroka, 2006), organizational communication (e.g., Kahneman & Tversky, 1984), and (political) advertising (e.g., Allen & Burrell, 2002). We, therefore, believe future research should address both factors; polarity as well as intensification need to be included in the assessment of evaluative language.

There are some considerations and suggestions for improvement that should be taken into account in future research. In Study 2, the pure intensifiers were operationalized with three Dutch adverbial intensifiers. These words indeed contain hardly any semantic content, but they are frequently used in Dutch (which is also the case for the English alternatives *very* and *really*). High word frequency may have contributed to the relative impact of the pure intensifiers. High-frequency words are known to bleach in meaning, and they are usually perceived as less strong than low-frequency words (Liebrecht et al., 2018). To investigate the robustness of our results, it is relevant to take word frequency into account.

Our measure of perceived strength is another caveat that should be taken into account. One of the disadvantages of using online questionnaires is the unnatural manner of assessing items on a conscious level (e.g., J. S. Bowers & Schacter, 1990; Newell & Shanks, 2014). In future research, we will use an alternative approach by means of an oral production task. If our interpretation of the results of both experiments is correct, then the same pattern should be found in spoken language (following the findings of Wilson & Wharton, 2006). Deviation from the default will be marked by intentional stress on pure intensifiers in positive evaluations in comparison to pure intensification in negative evaluations by which the speaker will mark his intention to highlight the ostensive character of pure intensification. We, therefore, expect speakers to stress pure positive intensification more often than pure negative intensification. For meaningful intensifiers, however, this pattern should not be observed, and speakers will stress meaningful positive intensification as often as

meaningful negative intensification. Otherwise, a less controlled but more ecological valid method would be a corpus study of intensified positive and negative evaluations in natural spoken language, such as the Spoken Dutch Corpus project that comprises about 1,000 hours of speech by Dutch adults or the British National Corpus with over 100 million samples of written and spoken language. By means of these studies, not only the third-party observations of written conversations are examined but also the perceptions of actual communicators in spoken conversations. Spoken conversations are known to differ in many communicative factors such as visibility, audibility, simultaneity (Clark & Brennan, 1991), and paralinguistic cues, such as intonation, prosody, and nonverbal communication (e.g., Poyatos, 1993).

Another suggestion for future research is to include other types of intensification in positive and negative evaluations. For example, what is the effect of stacking intensifiers (*a super rich tycoon*)? The speed of bleaching of some intensifiers is worth investigating (cf. Tagliamonte & Roberts, 2005). Compare the contrasting combinations of originally negative and positive words (*terribly good; abominably fun*), and novel intensifiers in youth language, like in Dutch *ziek mooi* (lit. *sick pretty*) and *kapot leuk* (lit. *damaging amusing*), they deserve to be included in future effect research or corpus investigation. Culture might also be included in future corpus research. Dutch is fairly similar to English regarding the use of pure intensifiers such as *very* or *really* (although Dutch has more choice), and Dutch is similar to English regarding the cultural preference for hedges and nuanced modifications of opinions. Cultures that are known for their more expressive communication styles may have a different perception of the intensifiers used in our studies and may perceive the meaningful intensifiers as less forceful than our Dutch participants. They may also prefer a broader range of extreme intensifiers. Most Western languages do make the difference between pure intensifiers and meaningful intensifiers, which gives rise to the expectation we will find similar results for other languages and cultures. This, however, remains to be investigated.

What is more, the effects of intensification and polarity should be seen in a wider and more social perspective. In some circumstances, language intensity in a negative message might be qualified as verbal aggression, and could, therefore, be included in verbal aggression frameworks. These frameworks are built on the basis of the actual usage of aggressive language and one's attitude to aggressive language, which can be assessed on the Verbal Aggressiveness Scale (VAS; Infante & Wigley, 1986). This scale ranges from the most negatively worded items to the most benevolently worded items (Levine et al., 2004; Tafoya & Hamilton, 2012). However, Levine et al. (2004) noticed the 20 VAS items may be problematic: The most negative and positive end of the VAS scale appear to be loaded on not one but two factors. Such an asymmetry in assessing negative and positive antonyms ties in with our findings that also suggest an imbalance between positive and negative intensified language. Thus, positive polarity could also play a role in verbal aggressiveness frameworks, especially when it invites collaborative behavior. As "verbal aggression represents an imminent danger to civilized society" (M. A. Hamilton, 2012, p. 6), it is

worthwhile to investigate if and how the usage of intensifiers can contribute to a better understanding of verbal aggressiveness.

Conclusions

The relevance of our studies is threefold. First, the consistent effect we have found suggests the polarity of the texts, indeed, can be an explanation for the mixed results in previous (experimental) studies regarding the effects of language intensity. It turns out intensifiers in positive and negative contexts function in different ways, dependent on the type and meaningfulness of the intensifier.

Second, our study has shown it is necessary to investigate the effects of positive and negative words and expressions in context. In previous research, the strength of evaluation in a text was mostly based on isolated expressions or single words. In automatic sentiment analysis, for example, the sentiment of a message is often based on the calculation of the sentiment scores of individual words in wordlists (e.g., De Rijke et al., 2013; Na et al., 2009). The perceived strength by real language users, however, is based on context and on the combination of specific words. Therefore, it is best not to attribute a fixed score to a text that contains an intensifier; the polarity of the text influences the strength of the evaluation.

Last, the negativity effect was found to be strong, but it can be modified. The Negativity bias and the Pollyanna principle predict that negative expressions are perceived as stronger than positive expressions. The results of our studies have confirmed this prediction, but the type and meaningfulness of the intensifiers also play a role. The negativity effect can be moderated by language intensity. Apparently, cognitive theories such as relevance theory or theory of mind are indispensable when one explains how language functions, and why it functions as it does. Based on our findings, we conclude intensity is a matter of the recipient's perception, which is shaped by stylistic features employed by the speaker. Language users vary their utterances according to the context and to the receiver of their messages. Expectancy, ostensiveness, and knowing what the other knows are key attributes in analyzing language, and in predicting what the effect of language choice will be. Whenever language is evaluative, that is, whenever language is not descriptive or objective, but subjective and personal, it is of the utmost importance to include context and perspective, and to account for the large and sometimes unique effects of stylistic and meaningful connotations.

In future sentiment analyses, opinion mining will be combined with machine learning; then, consumer reviews can be treated as singular and context-bound opinions, hence doing justice to the fact that what a person means with a certain word may vary from individual to individual, and from context to context.

Acknowledgments

The authors would like to thank Peter Jan Schellens, Roeland van Hout, and Henk Kempff as well as the two anonymous reviewers for their valuable feedback on earlier drafts of this article.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Note

1. No differences were found between the paid and unpaid participants on their scores on the dependent variable, $t(78) = 0.543$, $p = .59$.

References

Ahluwalia, R. (2002). How prevalent is the negativity effect in consumer environments? *Journal of Consumer Research*, *29*, 270-279.

Allen, M., & Burrell, N. (2002). The negativity effect in political advertising: A meta-analysis. In J. P. Dillard & M. Pfau (Eds.), *The persuasion handbook: Development in theory and practice* (pp. 83-96). Thousand Oaks, CA: Sage.

Anderson, P. A., & Blackburn, T. R. (2004). An experimental study of language intensity and response rate in email surveys. *Communication Reports*, *17*, 73-84.

Aune, R. K., & Kikuchi, T. (1993). Effects of language intensity similarity on perceptions of credibility, relational attributions, and persuasion. *Journal of Language and Social Psychology*, *12*, 224-238.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*, 323-370.

Block, L. G., & Keller, P. A. (1995). When to accentuate the negative: The effects of perceived efficacy and message framing on intentions to perform a health-related behavior. *Journal of Marketing Research*, *32*, 192-203.

Bolinger, D. (1972). *Degree words*. The Hague, Netherlands: Mouton.

Boucher, J., & Osgood, C. E. (1969). The Pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *8*, 1-8.

Bowers, J. S., & Schacter, D. L. (1990). Implicit memory and test awareness. *Journal of Experimental Psychology*, *16*, 404-416.

Bowers, J. W. (1963). Language intensity, social introversion, and attitude change. *Speech Monographs*, *30*, 345-352.

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Retrieved from http://www.uvm.edu/pdodds/teaching/courses/2009-08UVM-300/docs/others/everything/bradley1999a.pdf

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge, England: Cambridge University Press.

Buller, D. B., Burgoon, M., Hall, J. R., Levine, N., Taylor, A. M., Beach, B. H., . . . Bergen, A. (2000). Using language intensity to increase the success of a family intervention to protect children from ultraviolet radiation: Predictions from language expectancy theory. *Preventive Medicine*, *30*, 103-114.

Burgers, C., & De Graaf, A. (2013). Language intensity as a sensationalistic new feature: The influence of style on sensationalism perceptions and effects. *Communications*, *38*, 167-188.

Burgoon, M., & Chase, L. J. (1973). The effects of differential linguistic patterns in messages attempting to induce resistance to persuasion. *Speech Monographs*, *40*, 1-7.

Burgoon, M., Denning, V. P., & Roberts, L. (2002). Language expectancy theory. In J. P. Dillard & M. Pfau (Eds.), *The persuasion handbook: Developments in theory and practice* (pp. 117-136). Thousand Oaks, CA: Sage.

Burgoon, M., Jones, S. B., & Stewart, D. (1975). Toward a message-centered theory of persuasion: Three empirical investigations of language intensity. *Human Communication Research*, *1*, 240-256.

Carston, R. (2008). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford, England: Blackwell.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127-149). Washington, DC: American Psychological Association.

De Rijke, M., Jijkoun, V., Laan, F., Weerkamp, W., Ackermans, P., & Geleijnse, G. (2013). Generating, refining and using sentiment lexicons. In P. Spyns & J. Odijk (Eds.), *Essential speech and language technology for Dutch, theory and applications of natural language processing* (pp. 359-377). Berlin, Germany: Springer.

Dijksterhuis, A., & Aarts, H. (2003). On wildebeests and humans: The preferential detection of negative stimuli. *Psychological Science*, *14*, 14-18.

Feldman, S. (1966). *Cognitive consistency: Motivational antecedents and behavioral consequents*. New York, NY: Academic Press.

Fletcher, W. H. (1980). "BLOOD-HOT," "STONE-GOOD": A preliminary report on adjective-specific intensifiers in Dutch. *Leuvense Bijdragen*, *69*, 445-472.

Giora, R., Givoni, S., & Fein, O. (2015). Defaultness reigns: The case of sarcasm. *Metaphor and Symbol*, *30*, 290-313.

Ha, H. Y. (2004). Factors influencing consumer perceptions of brand trust online. *Journal of Product & Brand Management*, *13*, 329-342.

Hamilton, D. L., & Huffman, L. J. (1971). Generality of impression-formation processes for evaluative and nonevaluative judgments. *Journal of Personality and Social Psychology*, *20*, 200-207.

Hamilton, M. A. (2012). Verbal aggression: Understanding the psychological antecedents and social consequences. *Journal of Language and Social Psychology*, *31*, 5-12.

Hamilton, M. A., & Hunter, J. E. (1998). The effect of language intensity of receiver evaluations of message, source and topic. In M. Allen & R. W. Preiss (Eds.), *Persuasion: Advances through meta-analysis* (pp. 99-138). Cresskill, NJ: Hampton Press.

Hamilton, M. A., Hunter, J. E., & Burgoon, M. (1990). An empirical test of an axiomatic model of the relationship between language intensity and persuasion. *Journal of Language and Social Psychology*, *9*, 235-255.

Hamilton, M. A., & Stewart, B. L. (1993). Extending an information processing model of language intensity effects. *Communication Quarterly*, *41*, 231-246.

Hennig-Thurau, T., Gwinner, K. P., Walsch, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing*, *18*, 38-52.

Hornikx, J., Pieper, M., & Schellens, P. J. (2008). Versterkende, afzwakkende en numerieke markeringen in claims over cosmeticaproducten: Maken ze claims overtuigender?

[Strengthening, weakening and numerical markers in claims about cosmetics: Do they enhance the claims' persuasiveness?]. *Tijdschrift voor Communicatiewetenschap*, *36*, 3-13.

Infante, D. A., & Wigley, C. J. (1986). Verbal aggressiveness: An interpersonal model and measure. *Communication Monographs*, *53*, 61-69.

Jing-Schmidt, Z. (2007). Negativity bias in language: A cognitive-affective model of emotive intensifiers. *Cognitive Linguistics*, *18*, 417-443.

Jones, L. V., & Thurstone, L. L. (1955). The psychophysics of semantics: An experimental investigation. *Journal of Applied Psychology*, *39*, 31-36.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*, 341-350.

Kamoen, N., Holleman, B., Mak, P., Sanders, T., & Van den Bergh, H. (2011). Agree or disagree? Cognitive processes in answering contrastive survey questions. *Discourse Processes*, *48*, 355-385.

Kamoen, N., Holleman, B., & Van den Bergh, H. (2007). Hoe makkelijk is een niet moeilijke tekst? Een meta-analyse naar het effect van vraagformulering in tekstevaluatieonderzoek [How easy is a non-difficult text? A meta-analysis of the wording effect in questionnaires]. *Tijdschrift voor Taalbeheersing*, *29*, 314-332.

Lagerwerf, L., Boeynaems, A., van Egmond-Brussee, C., & Burgers, C. (2015). Immediate attention for public speech: Differential effects of rhetorical schemes and valence framing in political radio speeches. *Journal of Language and Social Psychology*, *34*, 273-299.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). *International affective picture system (IAPS): Technical manual and affective ratings*. Retrieved from https://www2.unifesp.br/dpsicobio/adap/instructions.pdf

Levine, T. R., Beatty, M. J., Limon, S., Hamilton, M. A., Buck, R., & Chory-Assad, R. M. (2004). The dimensionality of the verbal aggressiveness scale. *Communication Monographs*, *71*, 245-268.

Lieberman, M. (2013). *Social: Why our brains are wired to connect*. New York, NY: Broadway Books.

Liebrecht, C. (2015). *Intens krachtig. Stilistische intensiveerders in evaluatieve teksten* [*Intensely powerful. Language intensity in evaluative texts*] (Unpublished doctoral dissertation). Radboud University Nijmegen, Netherlands.

Liebrecht, C., Hustinx, L., Van Mulken, M., & Schellens, P. J. (2016). Krachtige taal. Een literatuurstudie naar taalintensivering in vier onderzoeksvelden [Powerful language. A literature review of language intensity in four disciplines]. *Tijdschrift voor Taalbeheersing*, *38*, 57-79.

Liebrecht, C., Hustinx, L., Van Mulken, M., & Schellens, P. J. (2018). Two dimensions of language intensity in evaluative discourse: Contextuality and semantic richness. In M. de los Ángeles Gómez González & L. Mackenzie (Eds.), *The construction of discourse as verbal interaction* (pp. 273-296). Amsterdam, Netherlands: John Benjamins.

McEwen, W. J., & Greenberg, B. S. (1970). The effect of message intensity on receiver evaluations of source, message and topic. *Journal of Communication*, *20*, 340-350.

Na, S.-H., Lee, Y., Nam, S.-H., & Lee, J.-H. (2009). Improving opinion retrieval based on query-specific sentiment lexicon. In M. Boughanem, C. Berrut, J. Mothe & C. Soule-Dupuy (Eds.), *Proceedings of the 31th European conference on IR research advances in information retrieval* (pp. 734-738). Berlin, Germany: Springer.

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133-142.

Newell, B. R., & Shanks, D. R. (2014). Unconscious influences on decision making: A critical review. *Behavioral and Brain Sciences*, *37*, 1-19.

Newhagen, J. E., & Reeves, B. (1991). Emotion and memory responses for negative political advertising: A study of television commercials used in the 1988 presidential election. In F. Biocca (Ed.), *Television and political advertising: Vol. 1. Psychological processes* (pp. 197-220). Hillsdale, NJ: Lawrence Erlbaum.

O'Keefe, D. J., & Jensen, J. D. (2006). The advantages of compliance or the disadvantages of noncompliance? A meta-analytic review of the relative persuasive effects of gain-framed and loss-framed messages. *Annals of the International Communication Association*, *30*, 1-43.

Parkinson, B. (1995). *Ideas and realities of emotion*. London, England: Routledge.

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, *1*, 33-60.

Poyatos, F. (1993). *Paralanguage: A linguistic and interdisciplinary approach to interactive speech and sounds*. Amsterdam, Netherlands: John Benjamins.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*, 380-391.

Purnawirawan, N., De Pelsmacker, P., & Dens, N. (2015). The impact of managerial responses to online reviews on consumers' perceived trust and attitude. In I. B. Banks, P. De Pelsmacker & S. Okazaki (Eds), *Advances in advertising research: Extending the boundaries of advertising* (Vol. 5, pp. 63-74). Wiesbaden, Germany: Springer Gabler.

Quené, H., & Van den Bergh, H. (2008). Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language*, *59*, 413-425.

R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/

Reichheld, F. F., & Schefter, P. (2000). E-loyalty: Your secret weapon on the web. *Harvard Business Review*, *78*, 105-113.

Renkema, J. (1997). Geïntensiveerd taalgebruik: Een analyseschema [Intensified language: An analysis scheme]. In H. Van den Bergh, D. Janssen, N. Bertens & M. Damen (Eds.), *Taalgebruik ontrafeld* [Language use explained] (pp. 495-504). Dordrecht, Netherlands: Foris.

Rogan, R. G., & Hammer, M. R. (1998). An exploratory study of message affect behavior: A comparison between African Americans and Euro-Americans. *Journal of Language and Social Psychology*, *17*, 449-465.

Rokeach, M. (1968). *Beliefs, attitudes, and values: A theory of organization and change*. San Francisco, CA: Jossey-Bass.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*, 131-142.

Soroka, S. N. (2006). Good news and bad news: Asymmetric responses to economic information. *Journal of Politics*, *68*, 372-385.

Sperber, D., & Wilson, D. (1994). *Relevance. Communication and cognition*. Oxford, England: Blackwell.

Tafoya, M. A., & Hamilton, M. A. (2012). Relational dynamics and the expression of aggression and comforting between siblings. *Journal of Language and Social Psychology*, *31*, 49-74.

Tagliamonte, S., & Roberts, C. (2005). So weird; so cool; so innovative: The use of intensifiers in the television series Friends. *American Speech*, *80*, 280-300.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, *110*, 67-85.

Tendahl, M. (2009). *A hybrid theory of metaphor: Relevance theory and cognitive linguistics*. Berlin, Germany: Springer.

Thompson, G., & Hunston, S. (2000). Evaluation: An introduction. In S. Hunston & G. Thompson (Eds.), *Evaluation in text. Authorial stance and the construction of discourse* (pp. 1-27). Oxford, England: Oxford University Press.

Van Mulken, M., & Schellens, P. J. (2012). Over loodzware bassen en wapperende broekspijpen. Gebruik en perceptie van taalintensiverende stijlmiddelen [About lead-heavy basses and flapping trouser legs. Use and perception of intensifying stylistic devices]. *Tijdschrift voor Taalbeheersing*, *34*, 28-55.

Ward, M. R., & Lee, M. J. (2000). Internet shopping, consumer search and product branding. *Journal of Product & Brand Management*, *9*, 6-20.

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*, 1191-1207.

Wilson, D., & Wharton, T. (2006). Relevance and prosody. *Journal of Pragmatics*, *38*, 1559-1579.

Wyer, R. S. (1974). *Cognitive organization and change: An information processing approach*. New York, NY: Wiley.

Author Biographies

Christine Liebrecht (PhD, Radboud University Nijmegen, the Netherlands) is an assistant professor at Tilburg Center for Cognition and Communication at Tilburg University (The Netherlands). Her research centers on the usage and effects of linguistic devices in written (online) business-to-consumer and consumer-to-consumer communication.

Lettica Hustinx (PhD, Radboud University Nijmegen, the Netherlands) is an assistant professor at the Centre for Language Studies at Radboud University Nijmegen (the Netherlands). Her research focusses on the linguistic realization as well as the persuasive effects of stylistic variations such as concreteness, vividness, and language intensity.

Margot van Mulken (PhD, VU University Amsterdam, the Netherlands) is a full professor of international business communication at the Centre for Language Studies at Radboud University Nijmegen (the Netherlands). Her research focusses on the use and effects of various stylistic devices such as irony, puns, language intensity, and visual metaphors and on the role of language choice in international encounters.