Negativity induced differences in the semantic-pragmatic behavior of expressive adjectives

Kalle GLAUCH — *Ruhr-Universität Bochum*

Abstract. Expressive Adjectives (EAs) (e.g., *damn*, *blasted*) express a negative speaker attitude toward a discourse referent. EAs are typically treated as a homogeneous class in semantic and pragmatic literature, mainly focusing on *damn* as the prototypical EA. The presumed behavior of *damn* is then generalized to all other EAs, neglecting the potential impact of meaning variations among EAs on their interpretation. This article addresses this shortcoming by investigating how different degrees of negativity in EAs affect their interpretation and acceptability. We report three studies: a valence rating study that establishes a negativity hierarchy of common EAs; a forced-choice study showing that EA negativity interacts with utterance-internal emotional cues in EA-interpretation; and an acceptability rating study showing that differently negative EAs vary in their compatibility with positive referents. While some results remain inconclusive, the findings challenge earlier generalizations on EA behavior and highlight the necessity for future research on EA processing to take these distinctions into account.

Keywords: expressive adjectives, emotional valence, negativity, experimental pragmatics.

1. Introduction

Expressive Adjectives (EAs) (e.g., *damn, frigging*) express a negative speaker attitude. Unlike descriptive attributive adjectives (e.g., *blue*), the interpretation of EAs is flexible and not strictly tied to their syntactic position, allowing them to target (be interpreted relative to) any discourse referent within an utterance. Drawing on (Potts, 2004, 2007), (Frazier et al., 2015) propose the 'speech-act hypothesis' to account for the interpretational flexibility of EAs. The speech-act hypothesis is a pragmatic account of EA-interpretation that treats EAs as quasi-independent non-at-issue speech-acts. As independent speech-acts, EAs can target referents denoted by constituents other than their syntactic sister. For instance, in sentence (1), the EA *damn* can target the subject-referent, the object-referent (local reading), or the event-referent, despite being realized object-internally.[1]

(1) The carpenter broke the damn glass.

☹ : (The waiter)	subject-reading
☹ : (The glass)	object-reading
☹ : (The waiter shattering the glass)	event-reading

EAs are typically treated as a homogeneous class in semantic and pragmatic literature. This

[1]The sad smiley notation to indicate a negative speaker attitude is taken from Gutzmann (2019b). The current paper differentiates between different degrees of negative speaker attitudes expressed by EAs as follows: the smiley ☹ represents an intermediate negativity associated with the EA *damn*, the smiley 😕 represents low negativity associated with the EA *darn*, and the smiley 😠 represents high negativity associated with the EA *fucking*.

perspective is prevalent in both theoretical (e.g., McCready, 2012; Padilla Cruz, 2022; Guercio and Orlando, 2022) and experimental works (e.g., Frazier et al., 2015; Gutzmann, 2019b, a; Bross, 2021; Ronderos and Domaneschi, 2023), which predominantly focus on *damn* as the prototypical EA. The presumed behavior of *damn* is then extended to all other EAs, neglecting how meaning variations among EAs could affect their processing and interpretation. Specifically, differences in the degree of negativity EAs express, and how this may interact with other emotional or attitudinal cues within an utterance to derive the EA target, remain unexplored.

This article addresses this research gap by empirically investigating how differences in the negativity of EAs influence their semantic-pragmatic behavior. The article will adopt a highly exploratory approach, owing to the absence of previous studies in this area, and will therefore abstain from formulating specific predictions beforehand.

The rest of the article is structured as follows: Section 2 reports a pre-study that empirically examines how the EAs *fucking*, *frigging*, *damn*, *darn*, *goddamn* and *blasted* differ from each other in the degree of negativity they express. Building on this, Section 3 reports a forced-choice study designed to explore how these differences in negativity affect the preferred interpretation of EAs in sentences manipulated by the valence of discourse referents. Section 4 presents an acceptability rating study testing whether EAs differ in acceptability depending on the valence of potential EA targets. Finally, Section 5 summarizes the key findings and concludes the paper.

2. Study 1: Differences in EA negativity

2.1. Research question

Study 1 investigates how EAs differ in terms of the degree of negativity they express as a prerequisite for examining their influence on EA interpretation in Studies 2 and 3. For some EAs, intuitive judgments appear sufficient to rank them hierarchically based on their expressed degree of negativity. For example, the EA *fucking* is intuitively perceived as considerably more negative than *blasted*. However, extensive empirical studies comparing the degrees of negativity conveyed by different EAs to validate these intuitions are still lacking. While sociolinguistic research on the use of EAs across demographic variables like age and gender (e.g., McCloskey and Coleman, 1992; Kaye and Sapolsky, 2004; Murphy, 2009) provides initial insights into the nuanced speaker attitudes conveyed by various EAs, these studies fall short of offering sufficient evidence to formulate definitive hypotheses regarding their negativity.

2.2. Method

Study 1 investigates the influence of EAs on the perceived valence of referents they modify to infer the expressed degree of negativity using a valence rating method. The underlying assumption is that a higher degree of EA negativity should lead to a more negative valence for modified nouns.

In a 1×6 within-subject repeated measures design, participants were presented with EA-Noun combinations manipulated by the EAs *damn*, *darn*, *goddamn*, *fucking*, *frigging*, *blasted* (e.g.

The EA view) and prompted to rate the perceived valence of the discourse referent. This specific subset of EAs was chosen for three reasons. First, they are pure expressive adjectives that convey only expressive meaning, differing from hybrid EAs (e.g., *shitty*, *bloody*) that may convey both expressive and descriptive meaning or either of the two depending on the context (e.g., Potts, 2004; Gutzmann, 2019b). Secondly, among the purely expressive EAs, they are among the most frequently used. Thirdly, they intuitively differ in the expressed degree of negativity, which is essential for investigating the interaction between EA negativity and referent valence in Studies 2 and 3.

Participants provided valence ratings on a 1-to-9 scale, with higher values indicating a higher degree of positivity. A 1-to-9 scale is the typical measuring scale for the emotional dimension of valence used in linguistic research, especially in sentiment dictionaries (e.g., Kanske and Kotz, 2010; Warriner et al., 2013; Stadthagen-Gonzalez et al., 2017). Using the same scale in the current study allows for the comparison of the modified valence both across different levels of EA and against the baseline inherent valence ratings for words in isolation taken from the sentiment dictionary of Warriner et al. (2013). This approach enables the measurement of the precise valence change induced by applying an EA to discourse referents with differing inherent valence ratings. Figure 1 shows an experimental item from the participants' perspective.

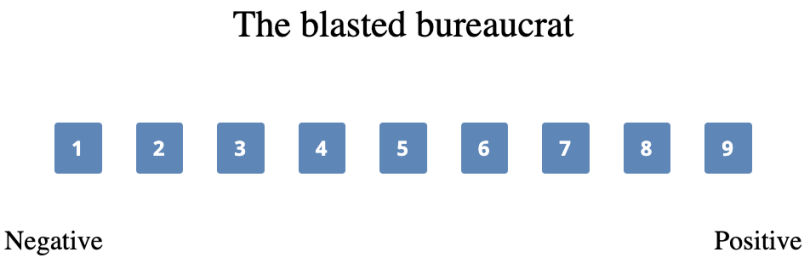

Figure 1: Example experimental item from participants' perspectice

2.3. Items

In total, 96 experimental items were constructed in all 6 conditions. The items were evenly distributed across six lists using a Latin-Square Design, ensuring that each participant encountered every item only once and that each condition was processed with equal frequency. The selection of nouns used as experimental items followed this procedure: First, 2775 tokens in the valence range between 3 and 7 were manually selected from the sentiment dictionary of Warriner et al. (2013). Next the list was segmented into distinct bins based on the valence values of the nouns in increments of 0.5. Lastly, a random selection of 12 words to be used as experimental items was performed within each bin. This method ensured a balanced and randomized selection of words across the whole range of inherent valence ratings.

Besides the experimental items, each list included 100 filler items composed of combinations of descriptive adjectives and nouns (e.g., *The heavy rain*, *The blue car*). The filler items aimed at concealing the study's objective and encouraging participants to utilize the full range of the rating scale throughout the experiment. At the beginning of each list, 3 items serving as the reference points for highly positive (*The wonderful Christmas*), highly negative (*The vicious*

murderer) and neutrally valenced discourse referents (*The white machine*) were presented. Afterwards, the experimental and filler items were successively displayed in an order uniquely randomized for each participant.

2.4. Participants

Participants (n = 78, Age: M = 44.8, SD = 15.2; Gender: male = 47, female = 31) were recruited through Prolific. Participants were prescreened for: English as a native language; currently residing in the US; a submission rejection rate of $\leq$ 1% ; and prior participation in at least 100 experiments on Prolific. Participants were randomly distributed to one of the 6 experimental lists. Participants received remuneration averaging £10 per hour[2].

2.5. Data exclusion

Before conducting the data analysis, participants were excluded if $\leq$ 25% of their response times per item were 650ms or shorter. A preliminary trial involving three reliable participants indicated such brief reaction times to be unlikely. As a result of this criterion, 5 participants were excluded from the study. After this participant level filtering, 41 individual observations with a response time of 650ms or shorter were also removed. 47 observations that deviated by ±2.5 standard deviations from the mean rating for the specific combination of EA and noun were identified as outliers and removed. In total, 6,920 observations were included in the data analysis.

2.6. Results

A generalized additive mixed effects model (GAMM) was fitted to the data using the R package `mgcv` (Wood and Wood, 2015). The model predicts the modified valence of the referents as a function of the fixed effects *Expressive Adjective* and *Inherent Valence*, allowing for non-linear effects of EA on modified valence across different ranges of inherent valence values. The model's random effect structure includes random intercepts for participants and experimental items, accounting for idiosyncratic differences in response patterns that are unrelated to the independent variables. The model translates to

```
gam(response ~ s(valence_war) + ea + s(itemname, bs = "re")
+ s(participant_id, bs = "re"), data = data)
```

Figure 2 visualizes the predicted distributions of the modified valence as a function of EA and inherent valence. The 95% confidence intervals are presented as shaded confidence bands around the estimate.

[2]The same for the other two studies reported in this article.

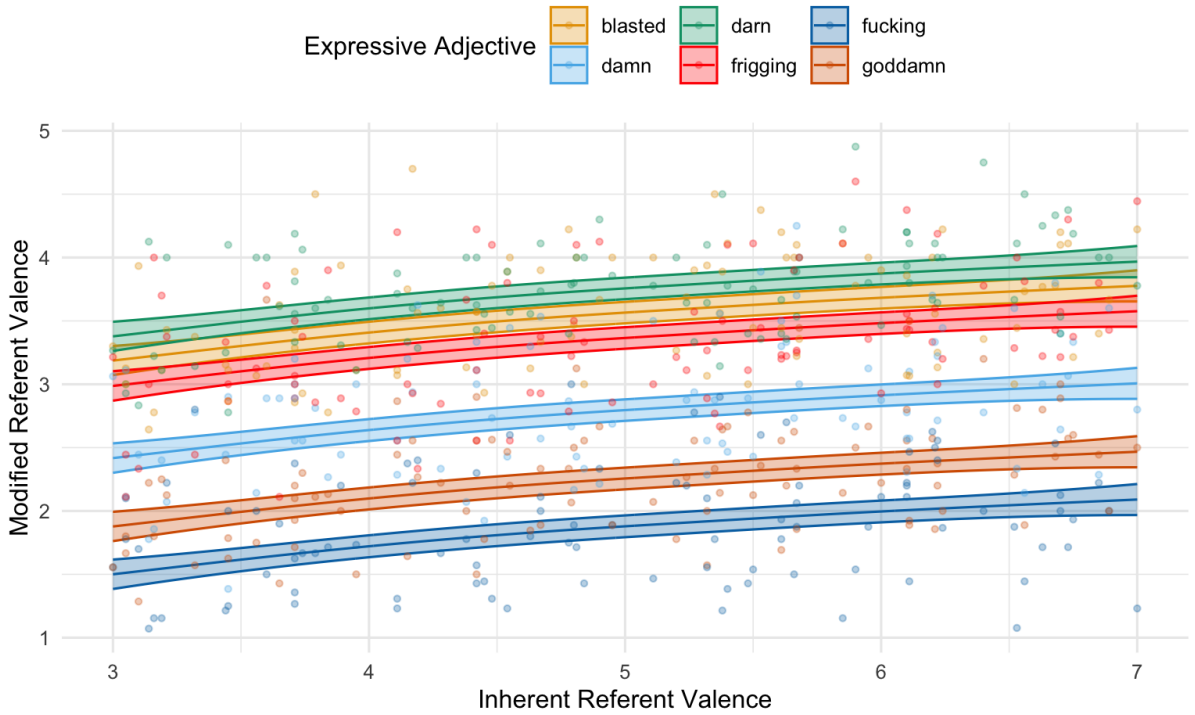

Figure 2: GAMM predicted distributions of modified valence by EA and inherent valence

2.7. Discussion

The findings establish a clear hierarchy for the degree of negativity expressed by the EAs investigated. The EA *darn* conveys the lowest degree of negativity, followed by *blasted*, *frigging*, *damn*, *goddamn* and *fucking*, which expresses the most negative speaker attitude. The GAMM-predicted distributions not being completely horizontal indicates that the inherent valence of discourse referents significantly influences their modified valence. A different outcome would occur only if the negativity of the EA completely overrode the referent's inherent valence, making EA the sole determinant in valence computation for EA-Noun combinations. However, the influence of inherent valence is weak, leading to greater EA-induced valence negativization for inherently positive referents than for inherently negative ones. The comparably weak effect of inherent-valence on the modified valence indicates that EA modification might involve two different processes depending on the valence of the inherent valence.

When a sentence contains a negative referent functioning as an appropriate EA target, EA modification primarily involves a matching operation. The negative EA targets the compatibly negative referent, potentially increasing its negativity slightly, depending on the intensity of the EA and the valence of the referent. This kind of operation is observed in (Ronderos and Domaneschi, 2023), who demonstrate that in a visual-world paradigm, listeners automatically focus on a contextually negative discourse referent as soon as they process an EA. Such a process is similar to the cue-based retrieval mechanism proposed for subject-verb agreement (e.g., Wagers et al., 2009; Dillon et al., 2013; Jäger et al., 2017), except that, instead of syntacto-semantic factors like case and animacy, negative valence serves as the primary cue making a discourse referent a target for EAs.

However, when the context contains no negative discourse referent, this kind of direct matching does not work, as valence cannot function as a cue for EA interpretation. Nonetheless, the negative attitude expressed by the EA still needs to be applied to a target. In this case, addressees have to actively reason about which referent the speaker could potentially express a negative attitude toward, based on other pragmatic cues such as EA-placement. This reasoning process might be considerably more demanding, especially when it involves a valence change induced by a highly negative EA that not only alters the polarity of a discourse referent (whether the valence is positive or negative) but also sets an extreme value. As seen in Figure 2 the valence changing operation for very negative EAs like *fucking* or *goddamn* is considerable. Even very positive referents are perceived as highly negative after the EA *fucking* is applied to them. For weakly negative EAs like *darn*, the valence change is much less pronounced.

The method employed in the current study does not provide direct evidence for an online processing mechanism of EAs beyond the one observed for sentences containing clearly negative referents in the study by (Ronderos and Domaneschi, 2023). However, a different kind of processing for sentences with no saliently negative referents seems compatible with the results of the current study and is theoretically plausible. Additional online studies on EA interpretation in sentences with differently valenced referents are needed to test this hypothesis.

3. Study 2: EA negativity and interpretation

3.1. Research question

Study 2 investigates whether the observed differences in negativity levels of the EAs in Study 1 influence their preferred interpretations. Given the absence of existing research on this topic, we will once again avoid making definitive predictions. To construct an experiment potentially capable of revealing insights regarding the Research question, we posit that EAs, as expressions of negative speaker attitudes, are susceptible to emotional cues. Consequently, they may interact variably with differently valenced referents within the same utterance.

3.2. Method

Study 2 employed a three-option forced-choice task within a 2×3 within-subjects repeated measures design. The six conditions result from fully crossing the two factors *EA* (*fucking* vs. *darn*) and *Subject-Referent Valence* (negative vs. neutral vs. positive). Participants were presented with short sentences following the structure 'The [negative | neutral | positive]-noun verbed the [fucking | darn] noun." and prompted to select the most likely target of the speaker's negative attitude. Inferring the target of a negative speaker attitude within an utterance is the typical operationalization of EA interpretation in experimental research (e.g., Bross, 2021; Gutzmann, 2019a, b; Frazier et al., 2015; Ronderos and Domaneschi, 2023). Participants were instructed to respond quickly based on their initial intuitions. The response options in the forced-choice task consisted of the subject referent, the object referent, and the event referent. Figure 3 illustrates a sample experimental item from the participants' perspective.

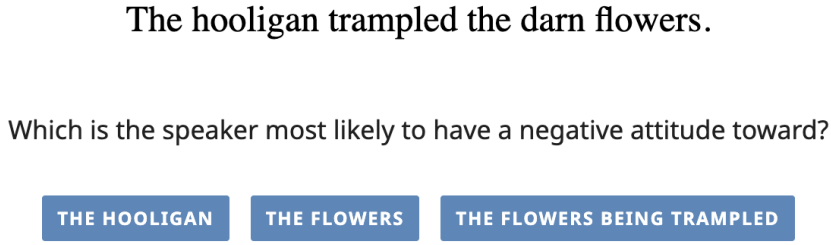

Figure 3: Example experimental item from participants' perspective

3.3. Items

In total, 24 experimental items were constructed in all six conditions. The items were evenly distributed across six lists using a Latin-Square Design.

The selection of nouns used as subject referents followed this procedure: First, 746 tokens denoting animate referents were manually taken from Warriner et al. (2013). These nouns were then categorized into three bins based on their valence values: negative (2 to 3.5), neutral (4.25 to 5.75), and positive (6.5 to 8). To ensure the reliability of the valence ratings, the values were compared against those of the sentiment dictionary of Mohammad (2018). From each bin, the 24 nouns showing the highest agreement in valence ratings across both sources (r = 0.995) were selected for the subject-referent manipulation. The 72 nouns were randomly assigned into groups of three and allocated to a specific item.

For the manipulation of EA negativity, the two EAs *fucking* and *darn* were selected. As illustrated in Figure 2, they express the most and least extreme negative speaker attitude respectively, making them the most likely pair to reveal potential influences of EA negativity on interpretation. The combinations of verb and object-noun were manually constructed to denote a slightly negative event (e.g., *knock over the milk*, *trample the flowers*). The object-nouns have a neutral to slightly positive valence (e.g., *window*, *concert*). Table 1 shows a sample experimental item in each condition.

Table 1: Example experimental item in all conditions

	Fucking	**Darn**
Negative	The jerk knocked over the fucking milk.	The jerk knocked over the darn milk.
Neutral	The houseguest knocked over the fucking milk.	The houseguest knocked over the darn milk.
Positive	The helper knocked over the fucking milk.	The helper knocked over the darn milk.

In addition to the experimental items, each list contained 51 filler items and 3 trial items. The filler items followed the same sentence structure as the experimental items, with half featuring a subject-internal adjective (e.g., *The rude guest complained about the dinner menu*) and the other half an object-internal adjective (e.g., *The janitor cleaned the filthy bathroom*). The trial items were presented at the start of each list to familiarize the participants with the task. Afterwards, the experimental and filler items were sequentially displayed in an order uniquely randomized for each participant.

3.4. Participants

Participants (n = 80, Age: M = 41.5, SD = 14.4; Gender: female = 49, male = 31) were recruited through Prolific. Prescreening adhered to the same criteria used in the pre-study. Participants were randomly distributed to one of the 6 experimental lists.

3.5. Data exclusion

Observations were excluded if the response time was ≤1500ms. This lower limit was set based on a trial run with three reliable participants, indicating reaction times below this threshold to be unlikely. Based on this criterion, 147 observations were removed. In total, 1773 observations were included in the data analysis.

3.6. Results

Table 2 displays the absolute and relative frequencies of subject interpretations, along with their 95% confidence intervals for each condition.

Table 2: Descriptive statistics for subject readings of EAs by subject valence

EA	Valence	n	Abs. Freq	Rel. Freq.	CI Lower	CI Upper
darn	negative	289	173	0.599	0.548	0.646
darn	neutral	299	164	0.548	0.499	0.596
darn	positive	294	155	0.527	0.477	0.576
fucking	negative	293	206	0.703	0.656	0.747
fucking	neutral	298	159	0.534	0.484	0.582
fucking	positive	300	139	0.463	0.414	0.512

Figure 4 presents an interaction plot visualizing the data. The plot depicts the proportion of subject interpretations on the y-axis against the EA-placement on the x-axis. Different levels of subject-referent valence are distinguished by varying line types. The red points represent the proportion of subject interpretations, and the error bars indicate the 95%-CI for each distinct factor combination.

A generalized linear mixed-effects model (GLMM) using a binomial distribution with a logit link function was fitted to the data with the R package `lme4` Bates et al. (2015). Dummy coding was used to obtain binary response values, by treating the object-reading and the event-reading uniformly as the non-subject response category. The model predicts EA interpretation as a function of the fixed effects *EA*, *Subject-Referent Valence* and their interaction. The model's random effect structure consists of random intercepts for participants and experimental items. The whole model translates to:

```
glmer(interpretation ~ EA * Valence + (1 | participant_id) + (1 | item),
data = data, family = binomial(link = "logit"))
```

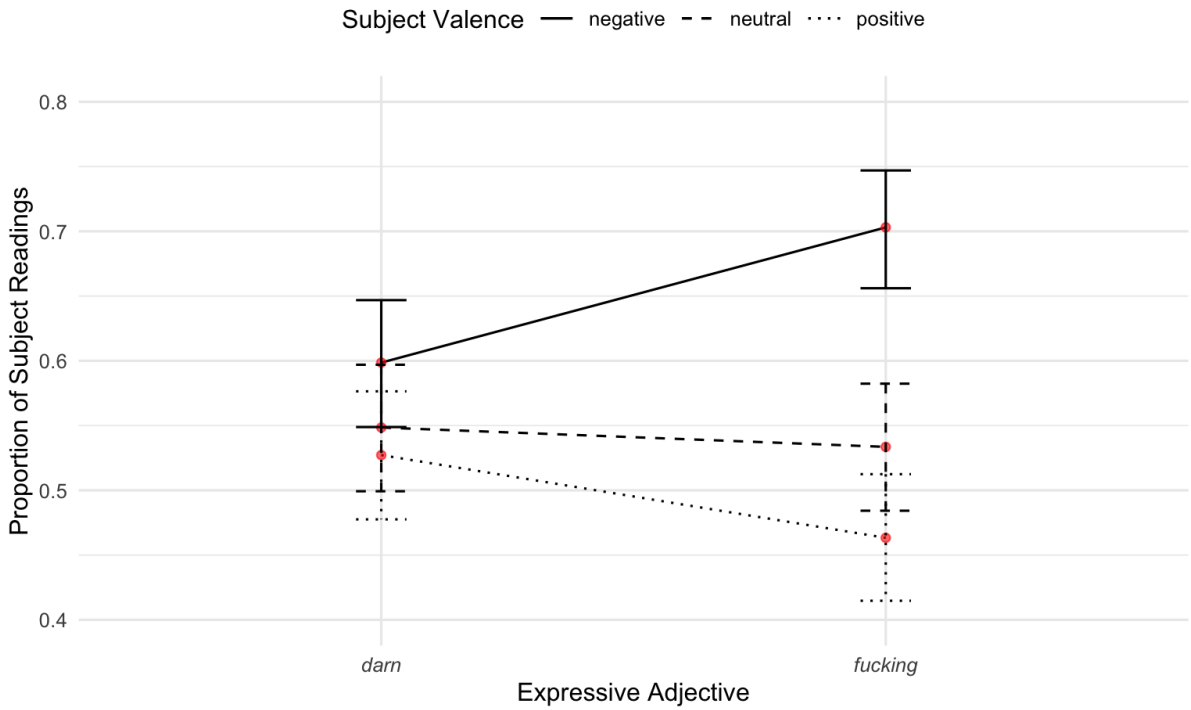

Figure 4: Observed proportions of subject readings by EA and subject valence with 95%CI

Table 3 presents the model output with *positive darn* sentences as the reference category (Intercept).

Table 3: Model summary for fixed effects

Term	Estimate	Std. Error	z Value	p Value
(Intercept)	0.195	0.189	1.032	0.30204
EAfucking	−0.330	0.183	−1.803	0.07139
Valencenegative	0.382	0.186	2.053	0.04003*
Valenceneutral	0.077	0.183	0.418	0.67605
EAfucking:Valencenegative	0.849	0.267	3.182	0.00146**
EAfucking:Valenceneutral	0.255	0.258	0.989	0.32266

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The model reveals a significant main effect of Subject-Referent Valence with negatively valenced subject referents featuring a significantly higher proportion of subject-readings compared to positive subject referents (Estimate = 0.382, Std. Error = 0.186, z = 2.053, p = 0.04*). No significant difference in EA interpretation is found between positive and neutral subject referents (Estimate = 0.077, Std. Error = 0.183, z = 0.418, p = 0.67). No significant main of EA is found (Estimate = -0.330, Std. Error = 0.183, z = -1.803, p = 0.071). However, the model indicates an interaction effect between EA and Subject-Referent Valence, with a significantly higher likelihood of subject-readings for negatively valenced subject referents compared to positive subject referents (Estimate = 0.869, Std. Error = 0.267, z = 3.182, p = 0.00146**) when the EA *fucking* is used.

3.7. Discussion

The presented study investigated whether EAs varying in the expressed degree of the negative speaker attitudes show differences in interpretational preferences, and whether these differences are modulated by the valence of the discourse referents realized in an utterance. Specifically, the study tested if the EAs *darn* and *fucking* behave differently when the valence of the subject-referent within in same clause is manipulated as negative, neutral or positive. The EAs *darn* and *fucking* were chosen because they show the largest disparity regarding the negativity they express, as shown in Study 1. While *darn* conveys a slightly negative speaker attitude, *fucking* conveys an extremely negative attitude.

The study reveals a significant interaction between the factors EA and subject-valence, with a higher probability of subject-readings for negatively valenced referents, compared to neutral and positive ones, when *fucking* is used. In contrast, for *darn*, the proportions of subject-readings across different valence levels do not significantly differ. To explain the observed difference in the behavior of the EAs, we propose that EAs respond to emotional cues within the discourse, such as the inherent valence of discourse referents. *Fucking*, expressing an extremely negative speaker attitude, is more aligned with negative referents (e.g., *hooligan*) that agree with its negativity. In contrast, with neutral or positive referents (e.g., *helper*), *fucking*'s negativity appears excessive, reducing compatibility. Conversely, *darn*, expressing more moderate negativity, agrees with referents in a broader valence range, leading to a consistent interpretation across different levels of subject-valence.

Such an account predicts that strongly negative EAs should be perceived as increasingly infelicitous in utterances lacking a similarly negative referent, with the acceptability of the sentence decreasing as the EA's negativity increases. This may be due to the addressee struggling to identify a suitable target for the extremely negative speaker attitude conveyed by the EA and the accompanying substantial valence change of the referent. For instance, sentence (2) containing the extremely negative EA *fucking* should be less felicitous than (3) containing the slightly less negative EA *damn* which in turn should be less felicitous than (4), containing the least negative EA *darn*.

(2) The student ate the fucking cookie.

😠 : (#The student)	subject-reading
😠 : (#The cookie)	object-reading
😠 : (#The student eating the cookie)	event-reading

(3) The student ate the damn cookie.

☹️ : (#?The student)	subject-reading
☹️ : (#?The cookie)	object-reading
☹️ : (#?The student eating the cookie)	event-reading

(4) The student ate the darn cookie.

🙁 : (The student)	subject-reading
🙁 : (The cookie)	object-reading
🙁 : (The student eating the cookie)	event-reading

While the proposal is speculative and will be investigated further in Study 3, a more concrete insight from Study 2 is that EAs are less homogeneous than typically presumed in terms of their interpretational behavior due to differences in expressive content. This suggests that the uniform treatment of EAs in prior research is insufficient, casting doubt on previously established generalizations. Although the current study exclusively analyzed the extreme examples of *fucking* and *darn*, more nuanced interpretational differences between EAs may exist.

4. Study 3: EA negativity and acceptability

4.1. Research question

Study 2 further investigates if EAs differentially interact utterance-internal emotional cues. Specifically, it examines whether EAs differ in perceived acceptability when embedded in contexts that either contain a saliently negative referent or not. As shown in (2) to (4), we predict that the acceptability of EAs in purely positive sentences (without an obviously negative referent) negatively correlates with the negativity of EAs due to the incompatibility between extremely negative EAs and positive referents. The application of negative EAs to positive referents involves not only a valence switch from positive to negative but also results in extremely negative valence values. This might be perceived as infelicitous given the initial positivity of the referent prior to EA modification.

4.2. Method

Study 2 employed an acceptability rating method in a 3×2 within-subject repeated measures design, resulting in 6 conditions derived from fully crossing the two factors EA (*darn* vs. *damn* vs. *fucking*) and Event Valence (positive vs. negative). Participants were presented with short sentences following the structure "The positive-noun [negative | positive]-verb the positive-noun." and selected the acceptability of the sentence on a 9-point Likert-Scale. Figure 5 illustrates a sample experimental item from the participants' perspective.

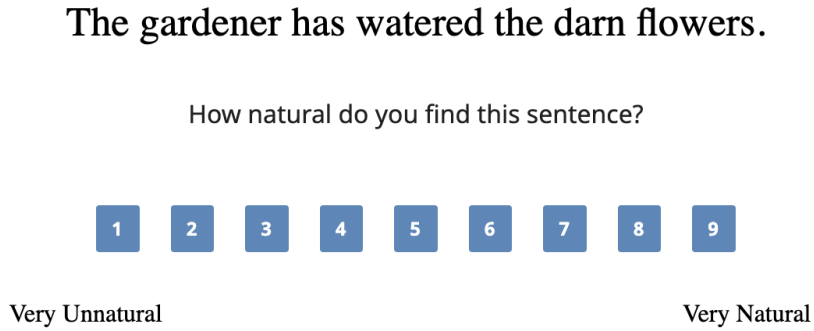

Figure 5: Example experimental item from participants' perspective

4.3. Items

In total, 24 experimental items were constructed in all six conditions. The items were evenly distributed across six lists using a Latin-Square Design. The items were manually formulated to contain a positive subject referent, a positive object referent and either a positive or negative event referent. For the manipulation of EA negativity, the EAs *darn*, *damn*, and *fucking* were used. *Damn* conveys an intermediate degree of negativity, allowing for the testing of gradient effects of EA negativity on its interaction with referent valence. Table 4 shows a sample experimental item in each condition.

Table 4: Example Experimental Item in all Conditions

	Positive Event	**Negative Event**
Darn	The teacher has praised the darn student.	The teacher has humiliated the darn student.
Damn	The teacher has praised the damn student.	The teacher has humiliated the damn student.
Fucking	The teacher has praised the fucking student.	The teacher has humiliated the fucking student.

In addition to the experimental items, each list contained 48 filler items and 2 trial items. The filler items followed the same sentence structure as the experimental items, with the descriptive adjective being realized either subject-internally or object-internally. The fillers differed regarding the semantic fit between the adjective and the modified noun with either low fit (e.g., *The aquatic accountant has crunched the numbers.*), intermediate fit (e.g., *The ancient skateboarder has performed a trick.*) or high fit (e.g., *The jeweler has crafted the intricate necklace.*). This ensured that participants made use of the full range of the rating scale throughout the experiment.

4.4. Participants

Participants (n = 80, Age: M = 41.5, SD = 14.2; Gender: female = 51, male = 27, non-binary = 2)[3] were recruited through Prolific. Prescreening adhered to the same criteria used in the pre-study. Participants were randomly distributed to one of the 6 experimental lists.

4.5. Data exclusion

Observations were excluded if the response time was ≤2000ms. This lower limit was set based on a trial run with three reliable participants, indicating reaction times below this threshold to be unlikely. Based on this criterion, 220 observations were removed. Additionally, 2 observations that deviated by ±2.5 standard deviations from the mean rating for a specific combination of EA and verb were identified as outliers and removed. In total, 1698 observations were included in the data analysis.

[3]In the other experiments, a non-binary response category was not offered. However, in response to a comment from one participant, it was included in the final study.

4.6. Results

Table 5 displays the mean, standard deviation, standard error, and the 95% confidence interval of the mean for each condition.

Table 5: Descriptive Statistics for Acceptability by EA and Event-Valence

EA	Valence	Mean	SD	n	SE	CI Lower	CI Upper
damn	negative	6.136	2.312	273	0.140	5.905	6.366
damn	positive	4.972	2.389	288	0.141	4.740	5.205
darn	negative	5.484	2.378	283	0.141	5.251	5.717
darn	positive	4.996	2.228	282	0.133	4.777	5.215
fucking	negative	5.683	2.551	284	0.151	5.433	5.933
fucking	positive	4.472	2.479	288	0.146	4.231	4.713

Figure 6 presents an interaction plot visualizing the data. The plot depicts the acceptability value on the y-axis against the EA on the x-axis. Different levels of event-referent valence are distinguished by varying line types. The red points represent the mean value, and the error bars indicate the 95%-CI for each distinct factor combination.

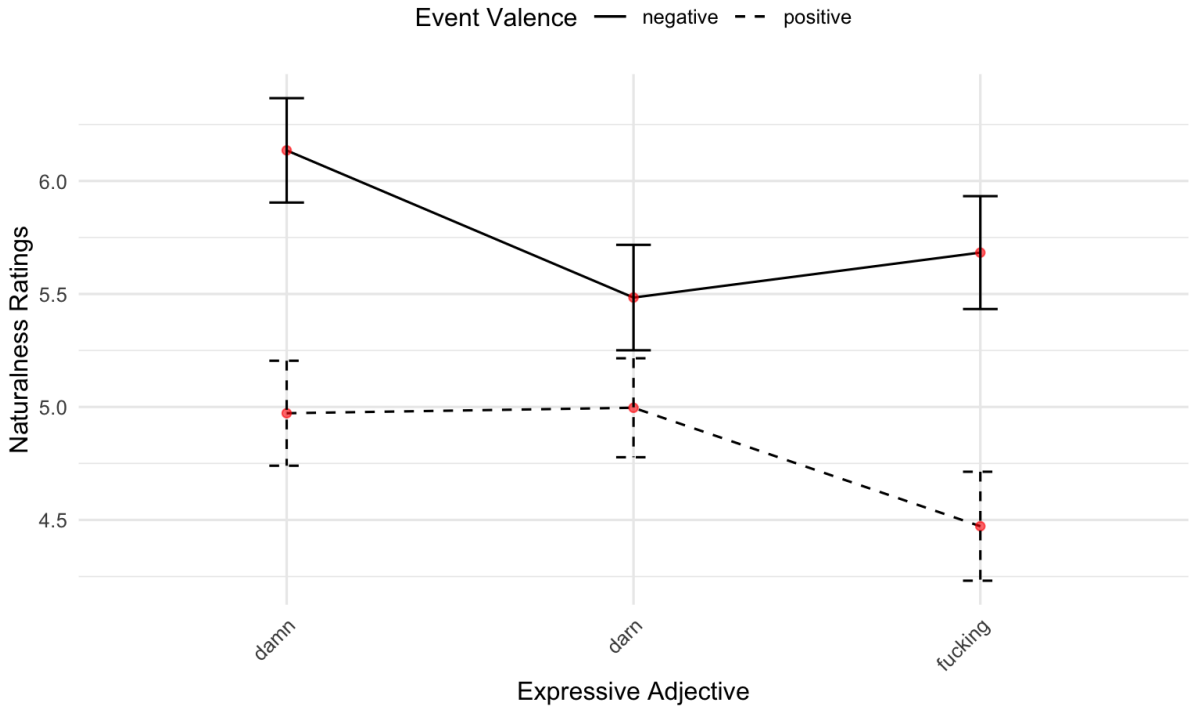

Figure 6: Observed Mean Acceptability by EA and Event Valence with 95%CI

A linear mixed-effects model (LMM) was fitted to the data with the R package `lme4` (Bates et al., 2015). The model predicts the acceptability as a function of the fixed effects *Expressive Adjective* and *Event Valence* and their interaction. The model's random effect structure consists of random intercepts for participants and items. The whole model translates to:

```
lmer(rating ~ EA * event-valence + (1 | participant_id) + (1 | itemname),
```

```
data = data)
```

Table 6 presents the model output with *negative darn* sentences as the reference category (Intercept).

Table 6: Model Summary for Fixed Effects

Term	Estimate	Std. Error	df	t Value	p Value
(Intercept)	5.6033	0.2400	121.6554	23.345	< 2e-16***
EAdamn	0.5345	0.1531	1600.2161	3.490	0.000495***
EAfucking	0.1410	0.1505	1595.5816	0.937	0.348881
valencepos	−0.6122	0.1526	1602.7451	−4.012	6.3e-05***
EAdamn:valencepos	−0.5227	0.2183	1607.2844	−2.395	0.016752*
EAfucking:valencepos	−0.6847	0.2132	1597.1925	−3.211	0.001348**

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The model reveals a significant main effect for EA, with significantly higher overall ratings for *damn* compared to *darn* (Estimate = 0.534, Std. Error = 0.153, t = 3.49, p < 0.001***). The overall rating difference between *fucking* and *darn* does not reach statistical significance (Estimate = 0.141, Std. Error = 0.150, t = 0.937, p = 0.34). The main effect for event valence reveals that overall acceptability is significantly lower in sentences containing only positive referents compared to sentences containing a negative event referent (Estimate = -0.612, Std. Error = 0.152, t = -4.01, p < 0.001***). Furthermore, the model shows an interaction effect between EA and Event Valence, with a significantly larger drop in acceptability for *damn* (Estimate = -0.522, Std. Error = 0.218, t = -2.395, p = 0.016*) and *fucking* (Estimate = -0.684, Std. Error = 0.213, t = -3.21, p = 0.001**) compared to *darn* when the level of event valence changes from negative to positive.

4.7. Discussion

The presented study investigated whether EAs varying in negativity show differences in acceptability when used in contexts that either contain a negative discourse referent they can readily target or not. It was predicted that acceptability negatively correlates with EA negativity in purely positive sentences.

The results cannot be interpreted straightforwardly due to the differing baseline acceptability for the EAs in sentences containing a negative referent, where all EAs should be similarly appropriate. *Damn*, with a rating of 6.2, is rated as significantly more acceptable than *darn* and *fucking*, which have ratings of 5.5 and 5.7, respectively. The baseline acceptability differences might be due to frequency differences between the EAs, although this is difficult to ascertain for *fucking*, which is rarely used in written computer-mediated communication (CMC). In 50,000,000 randomly selected review headlines from the Amazon Review Corpus[4], the EA *fucking* is used only 70 times, while the EA *damn* is used 8514 times and *darn* 7001 times.

[4]The Amazon Reviews Corpus is a collection of multiple hundred million reviews from Amazon.com across various product categories. This dataset includes reviews written by customers, review-headings, ratings, and other metadata about the products.

These frequency differences clearly do not reflect the actual use of the respective words across all spoken and written communication channels. Interestingly, tabooness does not seem to influence the baseline acceptability. Otherwise, *fucking*, which belongs to the taboo category of *sexual obscenities* (e.g., Jay, 2009; Allan, 2020), should be more unacceptable than *darn* and *damn*, irrespective of the valence of the discourse referents in the given context.

Two potential measures could effectively demonstrate the interaction between EA negativity and referent valence on acceptability. The first measure is the decrease in rating within each EA when the event valence shifts from negative to positive. A greater compatibility of the EA with a negative referent compared to a positive referent should result in a higher drop in rating. The second measure, primarily considered during hypothesis formulation, is the acceptability difference for purely positive sentences across different levels of EA negativity. The lower the compatibility of the EA with positive referents, the lower the acceptability should be.

Regarding the decrease in acceptability across the levels of event valence for the EAs, the results partly confirm the predictions. The drop in acceptability is significantly smaller for *darn* compared to the more negative EAs *damn* and *fucking*. However, contrary to predictions, the negativity differences between *damn* and *fucking* do not result in a higher drop in acceptability for *fucking*. Based on this measure, the results suggest a binary distinction between EA negativity and compatibility with positively valenced referents, with *damn* and *fucking* on one side, and *darn* on the other. One important conclusion from this interpretation would be that the acceptability of EAs with positive referents is not determined by cues like tabooness which are strongly correlated but not identical to valence, which applies to *fucking* but not to *damn*. Given the similar degree of negativity observed for *frigging* and *blasted* to *darn* in the pre-study, it might be speculated that they pattern similarly to *darn*, although the current study does not provide direct evidence for this.

Also regarding the acceptability of the EAs with purely positive sentences as the second potential measure, the hypotheses are only partly confirmed; however, in a different manner than suggested by the valence decrease for different levels of event valence discussed before. While the EAs *darn* and *damn* pattern the same with an acceptability around 5, *fucking* is significantly more negative with an acceptability of 4.5. According to this measure, it could be concluded that only *fucking* patterns differently from other EAs regarding compatibility with positive referents. This difference could be due to the significantly higher degree of negativity or its tabooness.

Similarly to Study 2, Study 3 finds a significant main effect of valence, with EAs being preferred in contexts that contain a negative referent. However, this time the main effect is not reducible to interaction effect between EA and Valence, as *darn* itself also shows a significant acceptability difference between the two levels of event valence. Given that Study 1 at least descriptively shows that *darn* preferentially targets negative discourse referents rather than neutral or positive ones, we prefer the tentative conclusion that all EAs preferentially target negative referents, although this effect is considerably modulated by the negativity of the EA. Further research is required to confirm these predictions.

Overall, Study 3 supports the conclusion of Study 2 that EAs differentially interact with the valence of discourse referents within the context. *Darn* differs from *fucking* in both acceptability differences across levels of event valence and acceptability in purely positive sentences, while it contrasts with *damn* only in terms of the first measure. Conversely, *damn* differs from *fucking*

only in acceptability in purely positive sentences. As mentioned earlier, it is not entirely clear how these results should be interpreted regarding the gradedness of compatibility between differently negative EAs and positive referents, as well as the potential role of tabooness. However, given that *darn* differs from the other EAs in at least one of the two measures of interest, we postulate that the compatibility difference is not binary and cannot be solely attributed to tabooness. This conclusion is somewhat speculative and requires further research, particularly by investigating how other EAs, such as *goddamn* or *frigging*, behave.

5. Conclusion

The studies presented in this paper demonstrate that EAs are less uniform than previously suggested, challenging prior claims that generalize the behavior of *damn* to all EAs. The findings reveal marked variation in the expressed degree of negativity among EAs, which significantly influences their semantic-pragmatic behavior, including interpretation and acceptability. These variations emerge from interactions with utterance-internal emotional cues, such as the valence of potential discourse referents that the EA can target. While the results clearly demonstrate that the least negative EA, *darn*, differs significantly from the most negative EA, *fucking*, further research is needed to investigate whether finer-grained distinctions in semanto-pragmatic behavior exist among EAs. Such research should aim to more rigorously control for related factors, such as tabooness, to better disentangle their effects from purely emotional cues, such as valence.

References

Allan, K. (2020). Taboo words and language: An overview. In K. Allan (Ed.), *The Oxford Handbook of Taboo Words and Language*, pp. 1–27. Oxford University Press.

Bates, D., M. Maechler, B. Bolker, S. Walker, R. H. B. Christensen, H. Singmann, B. Dai, G. Grothendieck, P. Green, and M. B. Bolker (2015). Package 'lme4'. *convergence 12*(1), 2.

Bross, F. (2021). On the interpretation of expressive adjectives: pragmatics or syntax? *Glossa: a journal of general linguistics 6*(1).

Dillon, B., A. Mishler, S. Sloggett, and C. Phillips (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language 69*(2), 85–103.

Frazier, L., B. Dillon, and C. Clifton (2015). A note on interpreting damn expressives: Transferring the blame. *Language and Cognition 7*(2), 291–304.

Guercio, N. L. and E. Orlando (2022). Expressives and argument extension. In *Semantics and Linguistic Theory*, pp. 44–62.

Gutzmann, D. (2019a). The dog ate the damn cake! the syntax of expressive adjectives. In *New Frontiers in Artificial Intelligence: JSAI-isAI 2018 Workshops, JURISIN, AI-Biz, SKL, LENLS, IDAA, Yokohama, Japan, November 12–14, 2018, Revised Selected Papers*, pp. 240–255. Springer.

Gutzmann, D. (2019b). *The Grammar of expressivity*, Volume 72. Oxford University Press.

Jäger, L. A., F. Engelmann, and S. Vasishth (2017). Similarity-based interference in sentence

comprehension: Literature review and bayesian meta-analysis. *Journal of Memory and Language 94*, 316–339.

Jay, T. (2009). The utility and ubiquity of taboo words. *Perspectives on psychological science 4*(2), 153–161.

Kanske, P. and S. A. Kotz (2010). Leipzig affective norms for german: A reliability study. *Behavior research methods 42*, 987–991.

Kaye, B. K. and B. S. Sapolsky (2004). Offensive language in prime-time television: Four years after television age and content ratings. *Journal of Broadcasting & Electronic Media 48*(4), 554–569.

McCloskey, L. A. and L. M. Coleman (1992). Difference without dominance: Children's talk in mixed-and same-sex dyads. *Sex Roles 27*, 241–257.

McCready, E. (2012). Emotive equilibria. *Linguistics and Philosophy 35*, 243–283.

Mohammad, S. M. (2018). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 english words. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, Melbourne, Australia.

Murphy, B. (2009). 'she's a fucking ticket': the pragmatics of fuck in irish english–an age and gender perspective. *Corpora 4*(1), 85–106.

Padilla Cruz, M. (2022). Is free enrichment always free? revisiting ad hoc-concept construction. *Journal of Pragmatics 187*, 130–143.

Potts, C. (2004). *The logic of conventional implicatures*, Volume 7. OUP Oxford.

Potts, C. (2007). The expressive dimension.

Ronderos, C. R. and F. Domaneschi (2023). Anticipating the damn referent: How comprehenders rapidly retrieve the speaker's attitude when processing negative expressive adjectives. *Cognitive Science 47*(5), e13295.

Stadthagen-Gonzalez, H., C. Imbault, M. A. Pérez Sánchez, and M. Brysbaert (2017). Norms of valence and arousal for 14,031 spanish words. *Behavior research methods 49*, 111–123.

Wagers, M. W., E. F. Lau, and C. Phillips (2009). Agreement attraction in comprehension: Representations and processes. *Journal of memory and language 61*(2), 206–237.

Warriner, A. B., V. Kuperman, and M. Brysbaert (2013). Norms of valence, arousal, and dominance for 13,915 english lemmas. *Behavior Research Methods 45*(4), 1191–1207.

Wood, S. and M. S. Wood (2015). Package 'mgcv'. *R package version 1*(29), 729.