Semantic prosody of extended lexical units: A case study

Chariton Charitonidis
ORCID iD: 0000-0003-0298-2629
dr.chariton.charitonidis@gmail.com

> There are three girls and a boy, all with a fondness for bright red jackets – and when I don't see them, when I oversleep, I actually feel blue. Bluer. That'd be the word my mom would use, not something as dramatic as depressed. I've had the blues for twenty-four years.
>
> (Gillian Flynn: "Dark places")

Abstract

Semantic prosody is typically referred to as an evaluative function of certain words or multiword items appearing within collocates of positive or negative meaning. The present study deals with the semantic prosody (context properties) of extended lexical units (ELUs) according to the psycholinguistic variables 'valence' (emotional positivity), 'arousal' (excitement, mood-enhancement), and 'concreteness'. The object of investigation are the verbal phrases *feel blue* (unambiguous idiomatic ELU, without a literal counterpart) and *see red* (ambiguous ELU, idiomatic or literal). The study builds on Snefjella & Kuperman (2016) who propose context norms for English words on the basis of a USENET mega-corpus. For the detection of ELU representations, a questionnaire-based survey was conducted with speakers of American English. For the detection of the context values of ELUs, a corpus research was carried on by using the Corpus of Contemporary American English (COCA) and the News on the Web corpus (NOW). The results suggest that ELU contexts largely conform to the averaged context norms of ELU constituents. ELU representations are strongly dissociated from contexts.

Keywords: context norms; crowdsourcing; emotion; multiword expressions; evaluative prosody.

1. Introduction

According to Louw (1993: 157), *semantic prosody* is "a consistent aura of meaning with which a form is imbued by its collocates". The key variable re-

ferred to in the literature in connection with semantic prosody is 'positivity' (or 'valence' in psycholinguistic terms, see later in this section).[1] For instance, the adverb *utterly* has an overwhelmingly 'bad' prosody. It shows up with a great number of 'bad' right-collocates, cf. *confused, demolishing, destroying, ridiculous,* etc., but only with a small number of 'good' right-collocates, cf. *secure, dedicated,* etc. (Louw 1993: 160). The vast majority of words with semantic prosody examined in the literature are not inherently positive or negative, cf. *utterly* mentioned above*,* or the verbs *set in, produce, cause,* etc. (Hauser & Schwarz 2016: 883).

Hunston (2007: 250) and Morley & Partington (2009: 144–148) point out two main research threads in the study of semantic prosody.[2] According to the first research thread ("the discourse perspective", Partington 2009) and its most well-known representative John McHardy Sinclair, semantic prosody is regarded as a property of a longer sequence of cooccurring items comprising a 'unit of meaning' (Sinclair 2004). *Units of meaning* are standardly assigned to templates and refer to broad attitudinal discourse functions, such as 'difficulty', 'reluctance', 'inability', etc. In (1) three contexts assigned to the template ' {visible} + [negative] + *naked eye* (core)' are given, taken from Sinclair (2004: 103).[3]

(1)   even though nothing is visible to the **naked eye.** We should trust…
        human ovum is barely visible to the **naked eye.** The corpus…
        plants that you can see with the   **naked eye** just as much as…

The semantic prosody assigned to all units of meaning in (1) is 'difficulty'. This prosody is evident in 85% of the full set of contexts including 'naked eye' (Sinclair 2004: 87–88). According to the second research thread ("the lexical-priming perspective", Partington 2009), semantic prosody is thought of as a property of a word or item, evident in collocates having a *positive* or *negative* attitudinal meaning (Louw 1993; Partington 2004; Hoey 2005; etc.). Semantic prosody distinguishes near-synonyms as a feature, see the near synonyms *hap-*

---

[1] Alternative terms for 'semantic prosody' are 'discourse prosody' or 'evaluative prosody'. It should be noted that all three terms are dissociated from phonology or phonetics.

[2] For a critical evaluation of the research on semantic prosody, see Stewart (2010).

[3] In this template, {visible} is an index of 'semantic preference' and [negative] a 'colligation' index. *Colligation* is the relation of co-occurrence between the core and abstract grammatical categories, e.g. past participles, quantifiers, negatives, etc. *Semantic preference* controls the collocational and colligational patterns (Stubbs 2009: 124).

pen, set in, occur, come about, and *take place* in (2) – bold letters indicate higher frequency of semantic-prosody tokens.

(2)   *happen*   unfavourable/**neutral** prosody
         set in   unfavourable prosody
         occur   unfavourable/**neutral** prosody
         come about   (emphasis on process)
         take place   unfavourable prosody

In this second research thread, semantic prosody is associated to expectations and reapplies in different contexts. For instance, if the context of a core item is typically positive, the appearance of this item in a context other than positive will call for „an additional attitudinal meaning, derived intertextually" (Hunston 2007: 250). Accordingly, in "prosodic clashes", irony is most commonly produced, see (3) and (4), taken from Morley & Partington (2009: 146).

(3)   *an outbreak of* (the expectation is for something bad)
         – *sanity* (at the EU)
         – *of honesty* (among Italian journalists)
         – *of good taste*

(4)   *there's much to be said for* (the expectation is of something good, or at least neutral)
         – *failure*
         – *acrimony*
         – *envy*
         – *death*

Whereas latest advances in computational linguistics have facilitated the detection of recurrent multi-word units and their collocates (Rundell 2018; Tang & Liu 2018), it still remains unclear (a) what is the full range of semantic prosody, i.e. which set of semantic variables are referred to by this term, and (b) what are the workings of semantic prosody in complex utterances, e.g. compounds, verbal phrases, etc. In particular, does semantic prosody of complex utterances consider the semantic prosody of *constituents* as stand-alone units? These issues become much more difficult to address when regarding *idioms* as targets of semantic prosody. Semantic dimensions vary within different classes of idioms. For instance, Citron et al. (2016) report that German

unambiguous idioms (i.e. idioms without a literal counterpart) are rated as more positive and arousing than ambiguous ones (i.e. idioms with a literal counterpart); or that German unambiguous idioms are rated as less concrete than ambiguous ones, etc.

The present study is the first attempt in the literature to define the semantic prosody of extended lexical units (ELUs)[4] with joint reference to the psycho-linguistic variables 'valence' (emotional positivity), 'arousal' (excitement, mood-enhancement), and 'concreteness'.[5] In the following, I give Kuperman's (2013) description of these variables.

> Valence, or emotional positivity, gages the amount of pleasantness or discomfort that a person feels when reading the word, and is measured on a scale from 1 (sad, unhappy) to 9 (happy). Words with extreme average valence ratings are *pedophile* (1.26) and *vacation* (8.53). Arousal assesses the level of excitement that raters associate with the read word, and is measured on a scale from 1 (calm) to 9 (excited). Words with extreme average arousal ratings are *grain* (1.6) and *insanity* (7.79)… Concreteness assesses, on a scale from 1 to 5, how easily the referent of the word can be seen, heard, felt, smelled, or tasted... Words with extreme average concreteness ratings are: *essentialness* (1.04) and *flashlight* (5.00).
>
> (Kuperman 2013: 3)

Before I proceed to the object of investigation, I would like to present the research on which this study builds.

2. Context norms for English words (Snefjella & Kuperman 2016)

By conducting extensive questionnaire-based surveys with speakers of American English on the emotional content of English words, Warriner et al. (2013) compiled, among others, large datasets of valence and arousal ratings. Similarly, Brysbaert et al. (2014) compiled a large dataset of concreteness ratings. In Snefjella & Kuperman (2016), the application of ratings to the 7 billion

---

[4] In this paper, the term 'lexical unit' is dissociated from Sinclair's (2004) notion of 'lexical item', i.e. a longer n-gram corresponding to a 'unit of meaning' (see the discussion earlier in this section).

[5] In the literature, the valence and arousal variables are referred to as 'affective', and the concreteness variable as 'sensorimotor' (Warriner et al. 2013; Brysbaert et al. 2014).

token USENET corpus (Shaoul & Westbury 2013) resulted in mean valence, arousal, and concreteness values for word contexts. Each context was confined from five content words before to five content words after a target word.[6] Contexts in which fewer than three words matched with ratings were excluded.[7] Accordingly, 14,853 words entered the analysis for which Snefjella & Kuperman (ibid) had semantic estimates for both individual words and their contexts. In Table 1, a sample context for the word *evidence* is given. Blanks indicate the absence of ratings for specific words.

Table 1. A sample context for the word *evidence* (Snefjella & Kuperman 2016: 137).

Word	Valence	Arousal	Concreteness
always			1.71
offer	5.94	3.42	2.23
zero			2.86
factual	5.89	3.05	2.41
logical	6.60	4.11	2.11
evidence	–	-	–
false			2.36
claims	5.15	3.90	
unless			1.54
stupid	2.65	4.68	1.75
unable	2.96	3.76	1.77
Mean	4.87	3.82	2.82

At the next stage, Snefjella & Kuperman (ibid.) averaged all context means across all occurrences of each word in the corpus. The resulting *norms* refer to three meta-variables, i.e. 'context valence', 'context arousal', and 'context

---

[6] In Snefjella & Kuperman's (2016) study, the term 'content words' is equivalent to the term 'non-stopwords'. Stopwords correspond to the default English stopword list of the R tm-package (personal communication).

[7] Excluded were also 493 words whose overall context values were more than three standard deviations above or below the mean of the respective variable (Snefjella & Kuperman 2016: 136).

concreteness'. The full list of norms can be found in the supplementary dataset of Snefjella & Kuperman (2016). These norms serve as "indices of the overall tendency of a word to occur in positive, exciting, or concrete contexts" (Snefjella & Kuperman 2016: 137).

Snefjella & Kuperman (2016: 139) show that "words tend to favour the company of words with similar affective and sensorimotor connotations." For instance, the verb *produce* has a valence of 7 and a context valence of 5.63, i.e. positive values in both cases, etc. In Table 2, the moderate to strong positive correlations of context and word ratings refer to this tendency.

Table 2. Correlations of context and word ratings (Snefjella & Kuperman 2016: 139).

Context valence vs. word valence	.58***
Context arousal vs. word arousal	.48***
Context concreteness vs. word concreteness	.72***

*** $p < .001$

Snefjella & Kuperman (2016: 136) report that "a word considered to have some quality (high concreteness, low valence etc.) does not necessarily occur in contexts that share that quality". There are also positive words that occur in negative contexts, e.g. *freeing, innocence, patriotic,* etc., and negative words that occur in positive contexts, e.g. *blinded, delinquent, motherless,* etc. (Snefjella & Kuperman 2016: 140).

The same authors conclude:

> It appears that "semantic prosody" encompasses dimensions other than valence; arousal and concreteness participate as well. Furthermore, the observed correlations indicate that the words that corpus linguists have identified as semantically prosodic – most often neutral words found in negative contexts – are the notable *exceptions* to the general pattern of words keeping company with words similar to themselves… In response to the questions of Whitsitt (2005), regarding whether the effects of semantic prosody carry over from one context (i.e. 'experimental task', CC) to the next, we answer with an emphatic yes, as evidenced by the correlations of contextual valence, arousal, and concreteness with lexical decision RTs, measures of serial recall ability, and age of acquisition ratings.
>
>   (Snefjella & Kuperman 2016: 144)

3. The present study

The present study extends the analysis in Snefjella & Kuperman (2016) to include multi-word expressions. In particular, this study seeks to determine (a) the context values of the ELUs *feel blue* and *see red,* and (b) the relations of these context values to (i) the averaged context norms of ELU constituents (Snefjella & Kuperman 2016), and (ii) the ELU representations, i.e. the valence, arousal, and concreteness values of the ELUs as unique lexical units. *Feel blue* and *see red* intersect semantically as idioms, cf. the meanings 'be sad or depressed' and 'become very angry', respectively, whereas *see red* has also a literal sense, i.e. 'see the colour red'. Both ELUs are verbal phrases that contain a Basic Colour Term (Berlin & Kay 1969) in argument position, and express negative emotions. Most notably, these ELUs refer to a contradictory pattern. While they express *negative* emotions, they contain verbs and Basic Colour Terms with *positive* context valence (*feel:* 5.74, *blue:* 5.86, *see:* 5.64, *red:* 5.65; Snefjella & Kuperman 2016).[8] It is thus worth seeing whether contexts target the positive context norms of constituents or the ELUs as negative representations.

This study will test the following working hypotheses:

(1) The averaged context norms of ELU constituents are ignored in ELU contexts, i.e. ELU representations suppress the context norms of constituents entirely.

(2) *Feel blue* and the idiomatic sense of *see red* have as negative idiomatic ELUs the same or a very similar relation to contexts. Literal *see red* is essentially different from idiomatic *see red*.

(3) Given that contexts can be strong cues as to the affective and sensorimotor qualities of a word (Snefjella & Kuperman 2016: 140), the relationships of affective and sensorimotor variables in ELU representations should be mirrored, for the most part, in ELU contexts.[9]

---

[8] The valence of these terms (word valence) is also positive, i.e. *feel:* 6.27, *blue:* 6.53, *see:* 6.27, *red:* 5.67 (Warriner et al. 2013).

[9] Admittedly, this is the most daring and path-searching hypothesis, however of considerable theoretical interest.

I will test these predictions by means of (a) a questionnaire-based survey, and (b) corpus analysis. In particular, the ELU representations will be detected by means of online interviews with speakers of American English. The context values of ELUs will be obtained by using the Corpus of Contemporary American English (COCA) and the News on the Web corpus (NOW).

The analysis will proceed as follows. Section 4 detects the valence, arousal, and concreteness representations of the ELUs *feel blue* and *see red*. Section 5 detects the valence, arousal, and concreteness values of ELU contexts. Section 6 tests how these contexts relate to (a) the averaged context norms of ELU constituents (Snefjella & Kuperman 2016), and (b) the ELU representations. Section 7 tests in both ELUs (a) the correlations of valence, arousal, and concreteness representations, and (b) the correlations of context valence, context arousal, and context concreteness. Section 8 summarizes the research findings with reference to the working hypotheses. Section 9 points out an issue for future research.[10]

4. ELU representations (online survey)

4.1. Materials and design

Table 3 below displays the valence, arousal, and concreteness assignments used in the online survey. Each assignment included the two critical items, together with three control words (fillers) of very low (filler-L), intermediate (filler-M), and very high (filler-H) values.[11] The control words for valence and arousal were taken from Warriner et al. (2013). The control words for concreteness were taken from Brysbaert et al. (2014).[12] Each participant was asked to evaluate all five items within her/his assignment. Accordingly, a statistical repeatedmeasures design was followed.

---

[10] For reporting statistical significance the following letters and symbols will be used: ns: $p > 0.05$, *: $p \leq 0.05$, **: $p \leq 0.01$, ***: $p \leq 0.001$.

[11] In the valence and arousal assignments, filler-M refers to a zero (neutral) value.

[12] According to these databases, the mean and standard deviation values of these words are [valence:] *abuse* 1.53 (1.07), *reservation* 5 (0.82), *delight* 8.21 (0.92), [arousal:] *level* 2.15 (1.63), *reputation* 5 (1.95), *insanity* 7.79 (1.44), [concreteness:] *belief* 1.19 (0.68), *introduction* 3 (1.31), *horse* 5 (0).

Table 3. Valence, arousal, and concreteness assignments.

Assignment	Critical item 1	Critical item 2	filler-L	filler-M	filler-H
Valence	feel blue	see red	abuse	reservation	delight
Arousal	feel blue	see red	level	reputation	insanity
Concreteness	feel blue	see red	belief	introduction	horse

4.2. Survey participants

Each assignment in Table 3 was evaluated 20 times by different Amazon Mechanical Turk "Workers". Accordingly, 60 different Workers participated in the study (36 male, 24 female; age range: 24–70). The participation requirements were: location in US, HIT Approval Rate greater than 95%, number of HITs approved greater than or equal to 1000. Each Worker participated in exchange for $0.75.

4.3. Procedure

The survey was announced by the author as "Requester" on the crowdsourcing platform Amazon Mechanical Turk in the first quarter of 2019. Workers were invited to rate five words or phrases according to their emotional content. The time allotted per Worker was 15 minutes. Each request page contained a unique survey link that redirected participants to the webform platform Google Forms. After reading and accepting the disclaimer's conditions, participants were asked to fill out a brief personal information form about age, gender, first language, country/state resided in most between birth and age 7, and educational level.[13]

The instructions in the valence and arousal assignments were adopted from Warriner et al. (2013), and the instructions in the concreteness assignments were adopted from Brysbaert et al. (2014), with minor modifications. In the instructions for valence and arousal, participants were asked to evaluate five words or phrases on a nine-point scale. Participants were acquainted with the respective variables by means of example words of high and low values. Participants were asked not to spend too much time on each word or phrase,

---

[13] These entries are identical to the personal information entries in Warriner et al. (2013) and Brysbaert et al. (2014).

but to respond spontaneously. In the instructions for concreteness, participants were asked to evaluate five words or phrases on a five-point scale. Participants were acquainted with the respective variable by means of example words and situations. The full instructions can be found in Appendix 1.

In the main evaluation form, critical items (=ELUs) and fillers were presented on a single page in random order. Right below each critical item or filler, ten calibrator words were displayed to help participants reconstruct the full evaluation space. The calibrator words for valence and arousal were taken from Kuperman et al. (2013) and the calibrator words for concreteness from Brysbaert et al. (2014). They were presented in the same order within an evaluation form and in random order between evaluation forms. The scale for valence depicted the following values and glosses: 1 'completely unhappy', 2, 3, 4, 5 'neutral', 6, 7, 8, 9 'completely happy'. The scale for arousal depicted the following values and glosses: 1 'completely calm', 2, 3, 4, 5 'neutral', 6, 7, 8, 9 'completely aroused/excited'. The valence and arousal values were arranged vertically from the most negative or non-arousing value (lower end) to the most positive or arousing value (higher end). This arrangement was suggested by the common association of low positions with negative or non-arousing emotions, and elevated positions with positive or arousing emotions, expressed in language in many different ways (Lakoff & Johnson 1980: 14–21). The scale for concreteness depicted the following values and glosses: 1 'abstract (language based)', 2, 3, 4, 5 'concrete (experience based)'. The concreteness values were also arranged vertically from the most abstract (lower end) to the most concrete (higher end).

Next to each answer option there was a checkbox. For an unknown word or phrase an additional checkbox was placed at the bottom. Participants were asked to select one of the checkboxes. In Appendix 2 the reader can find excerpts from the valence and concreteness forms.

Each interview lasted approximately 5–10 minutes. All interviews were anonymous.

4.4. Results

4.4.1. Valence

Figure 1 shows the overall results for valence in a scale from 1 ('completely unhappy') to 9 ('completely happy').
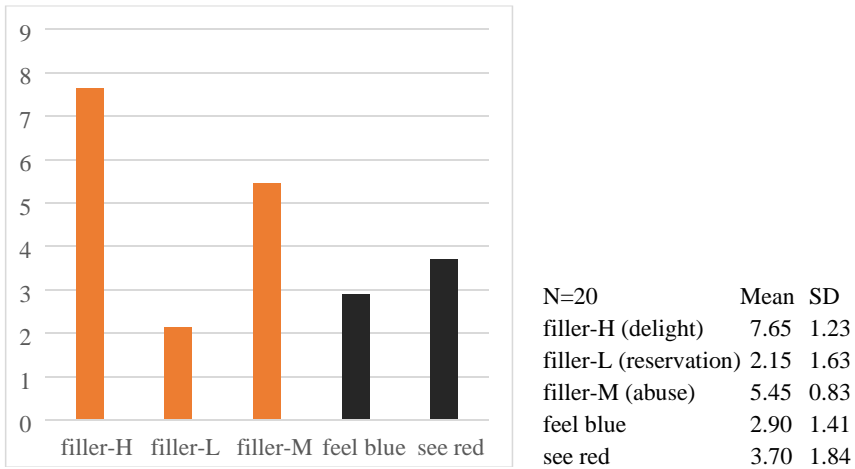
Figure 1. Representations of ELUs and fillers (valence).

The samples of all three fillers violated the assumption of normality (Shapiro-Wilk test, $p < .05$). For all five lexical units (=conditions), Mauchly's test showed that the condition of sphericity was also violated because the variances of differences between all combinations of related groups were not equal ($\chi^2(9) = 29.39$, $p = .001$). For these reasons, instead of one-way related ANOVA, Friedman's ANOVA (Friedman 1937) was conducted. This test revealed significant effects of condition ($\chi^2(4) = 56.435$, $p < .001$). The subsequent Wilcoxon signed rank tests (post hoc tests) referred to 10 pairwise comparisons at a .01 level of significance for each comparison ($.05/10 = .01$, Bonferroni correction). These comparisons revealed significant differences throughout, i.e. between all fillers and ELUs and between the ELUs themselves ($p < .01$ in all comparisons, exact one-tailed significance). Overall, a highly distinctive pattern emerged in which the negatively-valenced *feel blue* and *see red* showed up between filler-L and filler-M, whereas *feel blue* was significantly more negative than *see red*.

4.4.2. Arousal

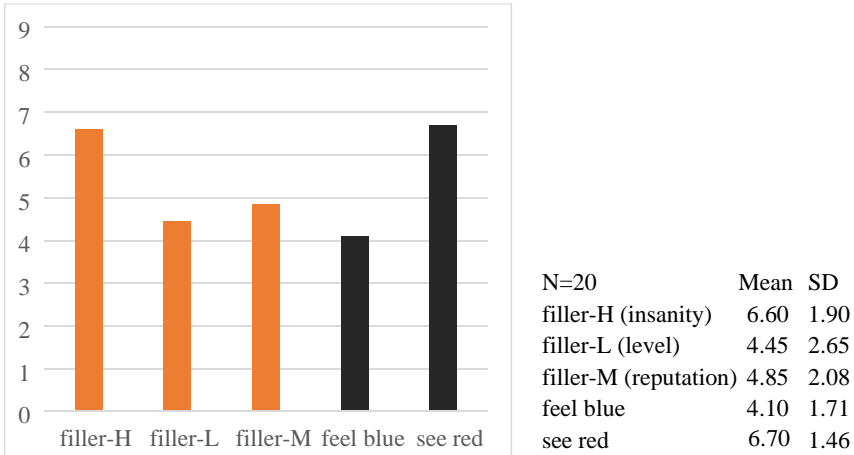Fig. 2 shows the overall results for arousal in a scale from 1 ('completely calm') to 9 ('completely aroused/excited').

Figure 2. Representations of ELUs and fillers (arousal).

The sample of filler-L violated the assumption of normality (Shapiro-Wilk test, p < .05). As in the valence patterns, Mauchly's test showed that the condition of sphericity was also violated because the variances of differences between all combinations of related groups were not equal ($\chi^2(9) = 21.758$, p = .010). Friedman's ANOVA revealed significant effects of condition ($\chi^2(4) = 25.425$, p < .001). The subsequent Wilcoxon signed-rank tests (post hoc tests) referred to 10 pairwise comparisons at a .01 level of significance for each comparison (.05/10 = .01, Bonferroni correction). In the following, I report the results with reference to the exact one-tailed significance: filler-L was not significantly different from filler-M, p = .260.

However, filler-H was significantly more arousing than both filler-L (p = .004) and filler-M (p = .000). *feel blue* was not significantly different from filler-L (p = .374) and filler-M (p = .130), while being significantly less arousing than filler H (p = .001). *see red* was significantly more arousing than filler-L (p = .003) and filler-M (p = .001), while not being significantly different from filler-H (p = .418). *feel blue* and *see red* were significantly different from one another (p = .000). Overall, *feel blue* and *see red* referred to a moderately distinct baseline of low and high arousal, respectively.[14]

---

[14] Filler-L, i.e. *level,* is one of the key factors of this moderately distinct baseline. As opposed to the results of the present survey, in Warriner et al. (2013) *level* had an explicit low-arousal status

4.4.3. Concreteness

Fig. 3 shows the overall results for concreteness in a scale from 1 ('abstract') to 5 ('concrete').
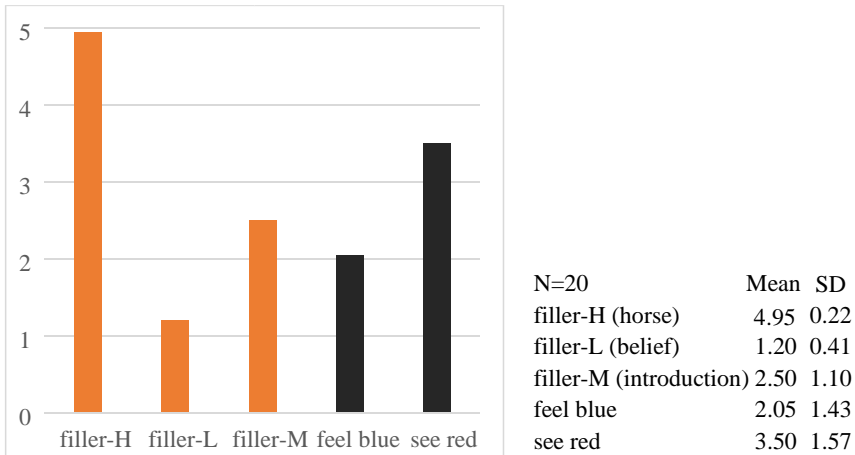

N=20	Mean	SD
filler-H (horse)	4.95	0.22
filler-L (belief)	1.20	0.41
filler-M (introduction)	2.50	1.10
feel blue	2.05	1.43
see red	3.50	1.57

Fig. 3: Representations of ELUs and fillers (concreteness).

All samples violated the assumption of normality (Shapiro-Wilk test, p < .05). Mauchly's test showed that the condition of sphericity was also violated because the variances of differences between all combinations of related groups were not equal ($\chi^2(9) = 36.796$, p = .000). Friedman's ANOVA revealed significant effects of condition ($\chi^2(4) = 52.228$, p < .001). The subsequent Wilcoxon signed-rank tests (post hoc tests) referred to 10 pairwise comparisons at a .01 level of significance for each comparison (.05/10 = .01, Bonferroni correction). In the following, I report the results with reference to the exact one-tailed significance: All fillers were significantly different from one another, p = .000. *feel blue* was not significantly different from both filler-L (p = .017) and filler-M (p = .108). However, *feel blue* was significantly less concrete than filler-H (p = .000). *see red* was significantly different from filler-L (p = .000)

(M: 2.15, SD: 1.63). It should be noted, however, that the arousal ratings in Warriner et al. (2013) show split-half reliability.

and filler-H (p = .000). However, *see red* was not significantly different from filler-M (p = .025). *feel blue* and *see red* were significantly different from one another (p = .001). Overall, a distinct baseline emerged, against which *feel blue* appeared as an ELU of low to moderate concreteness, and *see red* as an ELU of moderate concreteness.

Summarizing, the representations of *feel blue* and *see red* are given in Table 4. SD values are given in parentheses. As can be seen, *feel blue* appears with lower mean values throughout. It should be noted that the valence and arousal patterns are opposed to the results in Citron et al. (2016) in which unambiguous idioms in German were rated as more emotionally valenced and arousing than ambiguous ones. In contrast, the concreteness patterns conform to the results in Citron et al. (ibid) according to which unambiguous idioms in German were rated as less concrete than ambiguous ones.

Table 4. Representations of *feel blue* and *see red*.

	Valence	Arousal	Concreteness
feel blue (N = 20)	2.90 (1.41)	4.10 (1.71)	2.05 (1.43)
see red (N = 20)	3.70 (1.84)	6.70 (1.46)	3.50 (1.57)

5. ELU contexts (corpus analysis)

We can now proceed to the corpus-analysis part of this study querying the affective and sensorimotor contexts of *feel blue* and *see red*. The Corpus of Contemporary American English (COCA) and the News on the Web corpus (NOW) were accessed through the website <https://corpus.byu.edu>. According to the website's information (accessed 29 May 2019), COCA contains more than 560 million words of text, divided among spoken, fiction, popular magazines, newspapers, and academic texts. NOW contains 7.8 billion words of data from web-based newspapers and magazines from 2010 to the present time.

For the strings 'feel blue' (COCA) and 'see red' (COCA/NOW) 17 concordance lines for *feel blue,* 43 concordance lines for idiomatic *see red,* and

35 concordance lines for literal *see red* were obtained (=three samples).[15] Following the analysis in Snefjella & Kuperman (2016), the arrays of ten content words (i.e. non-stopwords) surrounding the core ELUs, were matched with ratings taken from Warriner et al.'s (2013) and Brysbaert et al.'s (2014) norming studies. For each of these context arrays (excluding the core ELUs), mean values of context valence, context arousal, and context concreteness were calculated.[16] Finally, the mean values were averaged across samples. All concordance lines, context arrays, and their means can be found in the supplementary-data file on the website Researchgate.net.[17] Henceforth, context arrays are simply referred to as 'contexts'.

The corpus analysis resulted in specific values for ELU contexts, see Table 5 below. Notwithstanding that the sample of *feel blue* was small (N=17), the assumption of normality was met for all three context variables, p > .05 (Shapiro-Wilk). Regarding the samples of *see red,* hyperlexeme is the union of idiomatic and literal senses.

Table 5. ELU contexts (Descriptives).

Test variables	feel blue (N=17)	see red: hyperlexeme (N=78)	see red: idiomatic (N=43)	see red: literal (N=35)
Context valence	5.88 (0.41)	5.68 (0.50)	5.55 (0.46)	5.83 (0.52)
Context arousal	4 (0.23)	4.05 (0.30)	4.03 (0.30)	4.08 (0.30)
Context concreteness	3.08 (0.33)	3.34 (0.44)	3.19 (0.43)	3.53 (0.37)

As Table 5 shows, in all four samples context valence was slightly positive, context arousal was slightly negative/low, whereas context concreteness referred approximately to a middle point between abstract and concrete.

---

[15] Concordance lines in which ambiguous uses of *see red* showed up were not included. The 17 concordance lines for idiom *feel blue* corresponded to the full set of COCA results. It should be noted that, originally, the analysis of *feel blue* was thought of as a pilot survey for validating the context norms in Snefjella & Kuperman (2016).

[16] None of the outliers were removed because they did not exceed the three standard-deviations mark (see Section 2).

[17] This file also contains the respective extended contexts. <https://www.researchgate.net/publication/353237301> (accessed 15 July 2021).

6.   Results

6.1.   Comparison of context values to test values

For detecting the role of ELU contexts in the demarcation of ELU meanings, a set of t-tests were conducted. The context values in Table 5 were compared to two hypothetical population values (test values), i.e. μ^1: the averaged context norms of ELU constituents (Snefjella & Kuperman's 2016 norms), and μ^2: the mean values of ELU representations obtained in Section 4, see Table 6 below.

Table 6. Hypothetical population values for *feel blue* and *see red*.

H_0	feel blue			see red		
μ^1	V: 5.81	A: 4.03	C: 3.11	V: 5.64	A: 4.06	C: 3.21
μ^2	V: 2.9	A: 4.10	C: 2.05	V: 3.7	A: 6.7	C: 3.5

μ^1: Averaged context norms of constituents (Snefjella & Kuperman 2016)
μ^2: Representations (Online survey).

Tables 7 and 8 below display the results of these comparisons, respectively (t-values).

Table 7. ELU contexts vs. averaged context norms of ELU constituents (t-values).

Test variable	feel blue	see red: hyperlexeme	see red: idiomatic	see red: literal
	t(16)	t(77)	t(42)	t(34)
Context valence	.736 ns	.653 ns	−1.228 ns	2.161*
Context arousal	−.570 ns	−.281 ns	−.712 ns	.376 ns
Context concreteness	−.378 ns	2.654**	−.336 ns	5.058***

(Two-tailed significance.)

Table 8. ELU contexts vs. ELU representations (t-values).

Test variable	feel blue	see red: hyperlexeme	see red: idiomatic	see red: literal
	t(16)	t(77)	t(42)	t(34)
Context valence	29.878***	34.725***	26.380***	24.355***
Context arousal	−1.569 ns	−77.527***	−57.630***	−51.537***
Context concreteness	12.725***	−3.225**	−4.789***	.454 ns

(Two-tailed significance.)

As can be seen in Table 7, ELU contexts conformed to or were strongly associated to the averaged context norms of ELU constituents. Most notably, idiomatic ELUs showed *absolute conformity* to the averaged context norms of ELU constituents, see the non-significant differences in *feel blue* and idiomatic *see red* across all three test variables. In contrast, as can be seen in Table 8, ELU contexts were strongly dissociated from representations.

To discuss more details, I will spell out the patterns in Tables 7 and 8 with reference to the standardized measure *Cohen's d.* Cohen's d standardizes the difference between a sample mean and a hypothetical population value with reference to the standard deviation of the sample mean (Formula for d: $H_1(M) - H_0(\mu)$ / SD. Reference values for d: 0.2/small-sized effect, 0.5/medium-sized effect, 0.8/large-sized effect. A d-value greater than 1 suggests a dissociation of dimensions).

6.1.1. ELU contexts vs. averaged context norms of ELU constituents

As already mentioned and as can be seen in Table 9 below, idiom *feel blue* shows up in contexts that are not significantly different from the averaged context norms of constituents. The same pattern holds for idiomatic *see red*. The contexts of literal *see red* have a significantly higher valence and concreteness than the averaged context norms of constituents, respectively, whereby only the difference in concreteness represents a large-sized effect. The contexts of

hyperlexeme *see red* also have a significantly higher concreteness than the av-
eraged context norms of constituents. However, this difference represents ap-
proximately a small-sized effect.

Table 9. ELU contexts vs. averaged context norms of ELU constituents (Cohen's d).

Test variable	feel blue	see red: hyperlexeme	see red: idiomatic	see red: literal
Context valence	0.17 ns	0.08 ns	−0.2 ns	0.37*
Context arousal	−0.13 ns	−0.03 ns	−0.1 ns	0.07 ns
Context concreteness	−0.09 ns	0.3**	−0.05 ns	0.86***

(Two-tailed significance.)

6.1.2. ELU contexts vs. ELU representations

As can be seen in Table 10 below, with the exception of the arousal patterns of
feel blue and the concreteness patterns of literal *see red,* ELU contexts are
significantly different from ELU representations. In addition, with the excep-
tion of the arousal patterns of *feel blue* and the concreteness patterns in all
three samples of *see red,* the d-values are disproportionally high and do not
qualify for effect-size considerations. As an exception, the contexts of literal
see red refer to a distinct pattern by both showing a non-significant difference
from representations and referring to an almost null size effect.

Table 10. ELU contexts vs. ELU representations (Cohen's d).

Test variable	feel blue	see red: hyperlexeme	see red: idiomatic	see red: literal
Context valence	7.27***	4***	4.02***	4.1***
Context arousal	−0.43 ns	−8.83***	−8.9***	−8.73***
Context concreteness	3.12***	−0.36**	−0.72***	0.08 ns

(Two-tailed significance.)

6.1.3. Discussion

The magnitude of standardized differences in Tables 9 and 10 suggests that the averaged context norms of ELU constituents and the ELU representations refer to totally different populations (dimensions) and are linked to ELU contexts in a fundamentally different way.

As regards the relation of ELU contexts to the averaged context norms of ELU constituents in Table 9, Cohen's d does not exceed '1' while pointing to the affinity of test variables to test values. On the other hand, the disproportionally high values of Cohen's d in Table 10 (d > 1) suggest that ELU contexts and ELU representations are strongly dissociated. Accordingly, the first working hypothesis set out in section 3 must be rejected. ELU representations do not suppress the context norms of ELU constituents.

The negatively-valenced representations of idiom *feel blue* and idiomatic *see red* call for the respective *negative* senses, as opposed to the averaged context norms of constituents that are *positive* showing absolute conformity to *positive* contexts.[18] These patterns verify the first part of the second working hypothesis in section 3. *feel blue* and the idiomatic sense of *see red* have as negatively-valenced ELUs a very similar relation to contexts.

In both hyperlexeme and idiomatic *see red,* the relations of context valence and arousal to the averaged context norms of constituents and to representations are blatantly similar. On top of this, in both hyperlexeme and idiomatic *see red,* the relations of contexts to representations are more consistent, showing similar significant differences in all dimensions. These patterns suggest that idiomatic senses of ambiguous ELUs have cognitive priority over literal ones. As an exception, the non-significant relation between context concreteness and representation concreteness in literal *see red* sets apart this sense from hyperlexeme and idiomatic *see red,* as one should expect.

6.2. Contexts of idiomatic and literal *see red:* Direct comparison of samples

As could be seen in Section 6.1.1, the contexts of literal *see red* were significantly more valenced and concrete than the averaged context norms of con-

---

[18] At the same time, the dissociation of context valence and representation valence is much stronger in *feel blue* (*feel blue:* 7.27 vs. *see red:* 4.02, Cohen's d), suggesting more robust effects of use in context (see Section 9).

stituents. However, in idiomatic *see red* and for the same relations, no significant differences were detected. It is now worth seeing whether the attested differences are mirrored within *see red*'s contexts without reference to the averaged context norms of constituents. In this way, the exact role of context in the demarcation of idiomatic and literal senses will become apparent. In the following, the results of independent samples t-tests are reported.

Context valence. On average, contexts surrounding literal *see red* (M = 5.83, SD = .52, N = 35) had a higher valence than contexts surrounding idiomatic *see red* (M = 5.55, SD = .46, N = 43). This difference was significant, t(76) = -2.483, p (two-tailed) = .015, and represented approximately a medium-sized effect, r = .27.

Context arousal. Contexts surrounding literal *see red* (M = 4.08, SD = .30, N = 35) were, on average, more arousing than contexts surrounding idiomatic *see red* (M = 4.03, SD = .30, N = 43). However, this difference was not significant, t(76) = -.757, p (two-tailed) = .451, and represented a rather small-sized effect, r = .09.

Context concreteness. On average, contexts surrounding literal *see red* (M = 3.53, SD = .37, N = 35) were more concrete than contexts surrounding idiomatic *see red* (M = 3.19, SD = .43, N = 43). This difference was highly significant, t(76) = -3.705, p (two-tailed) = .000, and represented approximately a large-sized effect, r = .39.

Concluding, both the results from the one sample t-tests in section 6.1 and the results from the independent samples t-tests in this section suggest the crucial role of context in the demarcation of idiomatic and literal *see red*. In particular, contexts surrounding literal *see red* were more valenced and more concrete than contexts surrounding idiomatic *see red*. At the same time, the difference in context concreteness was more relevant, referring approximately to a large-sized effect. In a nutshell, the second part of the second working hypothesis set out in section 3 is cross-validated.

7. Correlations

Let us now examine the third working hypothesis set out in section 3. This hypothesis predicted that the relationships of affective and sensorimotor variables in ELU representations should be mirrored, for the most part, in ELU contexts.

7.1. Correlations in ELU representations

In order to detect genuine effects between ELU representations, Pearson partial correlations were calculated. The respective tests aimed at assessing the relation between two variables by controlling for the effects of the remaining third variable. Given the lack of normality in the concreteness data, a bootstrapping procedure was followed to check the results (1,000 samples, BCa 95% confidence interval). This procedure determines that a significant correlation coefficient should strictly refer to either a positive or negative distribution of the resampled data – at the same time, it is not allowed that this distribution contains 0. In the following analysis, significant partial correlations up to ± .1 are referred to as "small correlations", between ± .1 and ± .3 as "moderate correlations", and between ± .3 and ± .5 as "large correlations" (Citron et al. 2015: 99).

Regarding *feel blue,* no significant relationship was found between representation valence and representation arousal, r = −.066, p (one-tailed) = .395. Similarly, representation valence did not correlate with representation concreteness, r = .021, p (one-tailed) = .466. On the other hand, representation arousal correlated negatively with representation concreteness, r = −.429, p (one-tailed) = .034. For this large-sized correlation, the Bootstrap confidence interval [−.679, −.081] was negative and did not contain 0. Table 11 below summarizes the results. It should be noted that in Citron et al.'s (2016) study of German idioms, arousal correlated *positively* with concreteness, as opposed to these results.

Table 11. Partial correlations between ELU representations (feel blue).

	Representation valence	Representation arousal	Representation concreteness
Representation valence	1	−.066 ns	.021 ns
Representation arousal	20	1	−.429*
Representation concreteness	20	20	1

(One-tailed significance.)

As regards *see red,* no significant correlations were found, see Table 12.

Table 12: Partial correlations between ELU representations (see red).

	Representation valence	Representation arousal	Representation concreteness
Representation valence	1	−.214 ns	.059 ns
Representation arousal	20	1	−.212 ns
Representation concreteness	20	20	1

(One-tailed significance.)

7.2. Correlations in ELU contexts

For *word contexts,* Snefjella & Kuperman (2016: 139) report three small-to-moderate and highly significant correlations between context variables, see Table 13.

Table 13: Correlations of context variables (Snefjella & Kuperman 2016: 139).

Context valence vs. context arousal	−.33***
Context valence vs. context concreteness	.21***
Context arousal vs. context concreteness	−.19***

*** p < .001.

This section produces the correlation patterns of context valence, context arousal, and context concreteness, in the idiom *feel blue,* and in hyperlexeme, idiomatic, and literal *see red*. As in section 7.1, Pearson partial correlations were calculated.

A. feel blue

There were no significant correlations between the three context variables, see Table 14 below. It should be noted, however, that the negative correlation 'context valence vs. context arousal' is a large one, its significance value being very close to the alpha level, r = −.421, p = .052.

Table 14: Partial correlations between context variables (feel blue).

	Context valence	Context arousal	Context concreteness
Context valence	1	−.421 ns	.174 ns
Context arousal	17	1	−.167 ns
Context concreteness	17	17	1

(One-tailed significance.)

B. see red: hyperlexeme

Context valence correlated negatively with context arousal, r = −.279, p = .007 (a moderate correlation). Table 15 contains the full set of relationships.

Table 15: Partial correlations between context variables (see red: hyperlexeme).

	Context valence	Context arousal	Context concreteness
Context valence	1	−.279**	.081 ns
Context arousal	78	1	−.166 ns
Context concreteness	78	78	1

(One-tailed significance.)

C. see red: idiomatic

Context valence correlated negatively with context arousal, r = −.264, p = .046 (a moderate correlation). Context concreteness correlated negatively with context arousal, r = −.265, p = .045 (a moderate correlation). Table 16 contains the full set of relationships.

D. see red: literal

Context valence correlated negatively with context arousal, r = −.423, p = .006 (a large-sized correlation). Table 17 contains the full set of relationships.

Table 16: Partial correlations between context variables (see red: idiomatic).

	Context valence	Context arousal	Context concreteness
Context valence	1	−.264*	−.120 ns
Context arousal	43	1	−.265*
Context concreteness	43	43	1

(One-tailed significance.)

Table 17: Partial correlations between context variables (see red: literal).

	Context valence	Context arousal	Context concreteness
Context valence	1	−.423**	.024 ns
Context arousal	35	1	−.228 ns
Context concreteness	35	35	1

(One-tailed significance.)

Let us now discuss the results. The significant correlations detected in the samples of *see red,* i.e. 'context valence vs. context arousal' and 'context arousal vs. context concreteness' are *negative*. Both correlations are reported by Snefjella & Kuperman (2016) for *word* contexts (Table 13). The negative correlation 'context valence vs. context arousal' is evident in all three samples of *see red*. However, the large-sized correlation in literal *see red* puts apart this sense from hyperlexeme and idiomatic *see red* that both show moderate correlations.

7.3. Comparison of correlations in ELU representations and ELU contexts

Let us now compare the partial correlations in ELU representations tested in section 7.1 with the partial correlations in ELU contexts tested in Section 7.2.

It should be noted that in both ELU representations and ELU contexts, all partial correlations (whether significant or not) were *negative*. In the following, I will refer to *significant* partial correlations alone.

As regards *feel blue,* representations referred to a large-sized correlation between arousal and concreteness, r = −.429, p (one-tailed) < .05. This significant correlation was not mirrored in ELU contexts, r = −.167, p (one-tailed) > .05.

As regards *see red,* no significant correlations were found in representations, as opposed to the contexts of hyperlexeme that referred to a medium-sized correlation between valence and arousal, r = −.279, p (one-tailed) < .01.

Accordingly, the third working hypothesis set out in Section 3 must be rejected. The relationships of affective and sensorimotor variables in ELU representations are *not* mirrored in ELU contexts. ELU contexts rather show general correlation patterns that also *word* contexts show (Table 13).

8. Conclusion

In line with Snefjella & Kuperman's (2016) assessments set out in Section 2, this study has shown that semantic prosody involves not only valence (positivity) but also arousal and concreteness. For the most part, ELU *constituents* keep company with words similar to themselves. However, this study has also shown that ELUs escape this pattern as unique mental representations. In the following, I would like to give the main results of this paper with reference to the working hypotheses set out in Section 3.

(1) ELU representations do not suppress the context norms of constituents. In particular, idioms (or idiomatic senses of ELUs) are strictly grounded in the context norms of their constituents. ELU representations are strongly dissociated from contexts.

(2) *feel blue* and the idiomatic sense of *see red* have as negatively-valenced ELUs the same or a very similar relation to contexts, for that matter. Literal *see red* is essentially different from idiomatic *see red*.

(3) The relationships of affective and sensorimotor variables in ELU representations are not mirrored in ELU contexts. This pattern may be regarded as evidence confirming the dissociation of representations and contexts.

The patterns detected in this paper should be verified taking into account many more ELUs. Further affective variables should be considered, e.g. 'dominance' (Warriner et al. 2013), or various sensorimotor variables (Lynott et al. 2019), to obtain context values and check their impact on ELUs. On top of this, a method should be developed for accessing ELU representations automatically, i.e. without reference to ratings of native speakers. Supervised learning algorithms could be used for this task, cf. the automatically generated norms for German lemmas in Köper & Schulte im Walde (2016), etc.

9.   Context valence and attitudinal functions[19]

In this last section, I would like to point out a salient pattern in the contexts of *feel blue* and *see red* that refers jointly to the discourse and lexical-priming perspective of semantic prosody mentioned in the Introduction. I will thereby refer exclusively to valence (positivity), by following the respective focus in the literature. In particular, I will argue that ELUs are associated to "typical scenarios in the everyday world" (Stubbs 2009: 133) at the interface of their lexical representation and context.

The analysis detected two thresholds of context valence that change or switch attitudinal functions, i.e. 5.65 for *feel blue* and 5.62 for idiomatic *see red*. The negatively-valenced *feel blue* refers to the general attitudinal function 'prevention'. In particular, *feel blue* typically shows up in positively-valenced contexts evoking the pragmatic scenario 'comfort', 'provision', etc., see (5).[20]

(5)   [feel blue: context no. 17, context valence: 6.23]
         Often small gestures open <u>large doors of feeling. Sometimes in the morning we</u> **feel blue**, <u>but not for long</u> <u>if there's a flower on the night table or next</u> to the bathtub to look at you when you wake up. (COCA/NEWS: Christian Science Monitor. "Petals and Stems Are His Art". 1992)

---

[19] The examples (5)–(8) in this section contain context numbers referring to contexts in the supplementary-data file (Researchgate.net). Context arrays are underlined. It should be noted that for detecting the attitudinal functions to which *feel blue* is associated, three additional contexts from the NOW corpus were considered (context nos. 18–20, supplementary-data file).

[20] As an exception, the positively-valenced contexts no. 19 and no. 20 are associated to the second scenario, i.e. 'emotional rejection', etc.

Contexts with a valence mean below 5.65 refer to a different scenario, i.e. 'emotional rejection', 'discomfort', etc., see (6).

(6)   [feel blue: context no. 13, context valence: 5.40]
        ... and those times Lymon would <u>slip and make some remark about him would make Lily</u> **feel blue**. <u>Lymon Jr. was drafted into the Army in' 66. He went</u> to Viet Nam and died fighting a war he neither believed in nor understood. (COCA/FIC: Ebony: Christian Science Monitor. Carter, Juanita "Lymon And Lily". 1997)

The explicit bifurcation of scenarios around a threshold may be due to the high number of exposures to positive uses of *feel blue* in early ages ("time weighting" of encounters, see Morley and Partington 2009: 148).

    Similarly, when the negatively-valenced *see red* shows up in contexts with a valence mean below 5.62, it refers to the default attitudinal function (strong) reluctance, aversion, or intolerance for things or situations regarded as "bad", see (7).

(7)   [see red: context no. 25, context valence: 4.47]
        Finally, her father took custody and moved her to New York. Diandre's and <u>Anthony's fathers are both in</u> prison. # „ For awhile, SharLinda would **see red** <u>and be so upset, „ Melinda recalls. „ I even kept</u> her red Crayola out of the way. (COCA/NEWS: Denver Post. Kevin Simpson: "Grandma copes with cross to bear 'Forever-baby' hers after daughter's hard life". 1997)

In contexts with a valence mean of/above 5.62, a different scenario shows up, *though not always,* in which aversion, perhaps as "envy", is directed towards others' benefits, privileges, success, etc., see (8).

(8)   [see red: context no. 31, context valence: 5.62]
        John Stossel;s report will make you **see red**. Are we going too far to protect inmates' rights? (COCA/SPOK: ABC_2020: "The Great Prison Pastime; Beyond Belief; Clinton Health Care Plan". 1993)

It should be noted that, for both *feel blue* and idiomatic *see red,* the suggested thresholds are almost identical, i.e. 5.65 and 5.62, respectively. The unstable

attitudinal switch in idiomatic *see red* may be due to the ambiguity of hyper-lexeme that refers to heterogeneous contexts (idiomatic and literal).[21]

As opposed to idiomatic *see red*, literal *see red* refers to totally different notions, e.g. 'vision', 'physical environment', 'capacity', etc., whereby there is no threshold of contextual valence that demarcates specific sub-scenarios.

Appendix 1: Online-survey: Instructions

The valence and arousal instructions were adopted from Warriner et al. (2013). The concreteness instructions were adopted from Brysbaert et al. (2014). Minor modifications were made.

A. Valence instructions

You are invited to take part in the study that is investigating emotion, and concerns how people respond to different types of words or phrases. You will use a scale to rate how you felt while reading each word or phrase. There will be five words or phrases. The scale ranges from 1 (unhappy) to 9 (happy). At one extreme of this scale, you feel completely unhappy, annoyed, unsatisfied, melancholic, despaired, or bored. You can indicate feeling completely unhappy by selecting 1. The other end of the scale is when you are happy, pleased, satisfied, contented, hopeful. When you feel completely happy you should indicate this by choosing rating 9. The numbers also allow you to describe intermediate feelings of pleasure, by selecting any of the other feelings. If you feel completely neutral, neither happy nor sad, select the middle of the scale (rating 5). Please don't spend too much time thinking about each word or phrase. Rather, make your ratings based on your first and immediate reaction as you read each word or phrase.

B. Arousal instructions

You are invited to take part in the study that is investigating emotion, and concerns how people respond to different types of words or phrases. You will use a scale to rate how you felt while reading each word or phrase. There will be five words or phrases. The scale ranges from 1 (calm) to 9 (excited). At one extreme of this scale, you are relaxed, calm, sluggish, dull, sleepy, or unaroused. You can indicate feeling

---

[21] Another factor may be the magnitude of dissociation between representation and context. This dissociation is considerably weaker in *see red* as compared to *feel blue,* i.e. 4 vs. 7.27, respectively (Cohen's d, see section 6.1.2).

completely calm by selecting 1. The other end of the scale is when you feel completely stimulated, excited, frenzied, jittery, wideawake, or aroused. When you feel completely aroused you should indicate this by choosing rating 9. The numbers also allow you to describe intermediate feelings of calmness/arousal, by selecting any of the other feelings. If you feel completely neutral, neither calm nor at all excited, select the middle of the scale (rating 5). Please don't spend too much time thinking about each word or phrase. Rather, make your ratings based on your first and immediate reaction as you read each word or phrase.

C. Concreteness instructions

Some words or phrases refer to things or actions in reality, which you can experience directly through one of the five senses. We call these words or phrases concrete words or phrases. Other words or phrases refer to meanings that cannot be experienced directly but which we know because the meanings can be defined by other words. These are abstract words or phrases. Still other words or phrases fall in-between the two extremes, because we can experience them to some extent and in addition we rely on language to understand them. We want you to indicate how concrete the meaning of each word or phrase is for you by using a 5-point rating scale going from abstract to concrete. A concrete word or phrase comes with a higher rating and refers to something that exists in reality; you can have immediate experience of it through your senses (smelling, tasting, touching, hearing, seeing) and the actions you do. The easiest way to explain a word or phrase is by pointing to it or by demonstrating it (e.g. To explain 'sweet' you could have someone eat sugar; To explain 'jump' you could simply jump up and down or show people a movie clip about someone jumping up and down; To explain 'couch', you could point to a couch or show a picture of a couch). An abstract word or phrase comes with a lower rating and refers to something you cannot experience directly through your senses or actions. Its meaning depends on language. The easiest way to explain it is by using other words (e.g. There is no simple way to demonstrate 'justice'; but we can explain the meaning of the word by using other words that capture parts of its meaning). Always think of how concrete (experience based) the meaning of the word or phrase is to you. In all likelihood, you will encounter several words or phrases you do not know well enough to give a useful rating. This is informative to us too, as in our research we only want to use words or phrases known to people. Please indicate when you don't know a word or phrase by using the letter N.

Appendix 2: Online-survey: Evaluation forms (excerpts)

(Valence)   (Concreteness)

How do you feel when reading the following word or phrase?

Please read the example words under each word or phrase for getting the full emotional range.

abuse *

Emotional range: free, icebox, jail, grin, joke, kitten, industry, insecure, invader, hat

○ 9 Completely happy

○ 8

○ 7

○ 6

○ 5 Neutral

○ 4

○ 3

○ 2

○ 1 Completely unhappy

○ Unknown word or phrase

How do you feel when reading the following word or phrase?

Please read the example words under each word or phrase for getting the full concreteness range.

introduction *

Concreteness range: infinity, grasshopper, marriage, polite, whistle, theory, gas, shirt, kick, sugar

☐ 5 Concrete (experience based)

☐ 4

☐ 3

☐ 2

☐ 1 Abstract (language based)

☐ N I do not know this word well enough to give a rating

References

Berlin, Brent & Paul Kay. 1969. *Basic color terms. Their universality and evolution.* Berkeley, CA: University of Los Angeles Press.

Brysbaert, Marc, Amy Beth Warriner & Victor Kuperman. 2014. Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods* 46. 904–911.

Citron, Francesca M.M., Cristina Cacciari, Michael Kucharski, Luna Beck, Markus Conrad & Arthur M. Jacobs. 2016. When emotions are expressed figuratively: Psycholinguistic and Affective Norms of 619 Idioms for German (PANIG). *Behavior Research Methods* 48. 91–111.

Friedman, Milton. 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association* 32. 675–701.

Hauser, David J. & Norbert Schwarz. 2016. Semantic prosody and judgment. *Journal of Experimental Psychology: General* 145(7). 882–896.

Hoey, Michael. 2005. *Lexical priming: A new theory of words and language*. London/New York: Routledge.

Hunston, Susan. 2007. Semantic prosody revisited. *International Journal of Corpus Linguistics* 12(2). 249–268.

Kay, Paul & Chad K. McDaniel. 1978. The linguistic significance of the meanings of basic color terms. *Language* 54(3). 610–646.

Köper, Maximilian & Sabine Schulte im Walde. 2016. Automatically Generated Norms of Abstractness, Arousal, Imageability and Valence for 350,000 German Lemmas. In: *Proceedings of the 10th Conference on Language Resources and Evaluation* (LREC). Portorož, Slovenia, May 2016.

Kuperman, Victor. 2013. Accentuate the positive: Diagnostics of semantic access in English compounds. *Frontiers in Language Sciences* 4(203). 1–10.

Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Louw, Bill. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In Mona Baker, Gill Francis & Elena Tognini-Bonelli (eds.), *Text and Technology: In Honour of John Sinclair*, 157–176. Amsterdam: Benjamins.

Lynott, Dermot, Louise Connell, Marc Brysbaert, James Brand & James Carney. 2020. The Lancaster Sensorimotor Norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods* 52. 1271–1291.

Morley, John & Alan Partington. 2009. A few *Frequently Asked Questions* about semantic – or *evaluative* – prosody. *International Journal of Corpus Linguistics* 14(2). 139–158.

Partington, Alan. 2004. "Utterly content in each other's company": semantic prosody and semantic preference. *International Journal of Corpus Linguistics* 9(1). 131–156.

Regier, Terry, Paul Kay & Richard S. Cook. 2005. Focal colors are universal after all. In *Proceedings of the National Academy of Sciences* 102(23). 8386–8391.

Rundell, Michael. 2018. Searching for extended units of meaning – and what to do when you find them. *Lexicography* 5(1). 5–21.

Shaoul, Cyrus & Chris Westbury. 2013. A reduced redundancy usenet corpus (2005–2011). Edmonton, AB: University of Alberta. www.psych.ualberta.ca

Sinclair, John. 2004. *Trust the text: Language, Corpus, and discourse*. London/New York: Routledge.

Snefjella, Bryor & Victor Kuperman. 2016. It's all in the delivery: Effects of context valence, arousal, and concreteness on visual word processing. *Cognition* 156. 135–146.

Stewart, Dominic. 2010. *Semantic Prosody: A Critical Evaluation*. New York: Routledge.

Stubbs, Michael. 2009. The search for units of meaning: Sinclair on empirical semantics (memorial article). *Applied Linguistics* 30(1). 115–137.

Tang, Xuri & Gaixiang Liu. 2018. Solving contradictions in semantic prosody analysis with prosody concord. *International Journal of Corpus Linguistics* 23(4). 437–466.

Warriner, Amy Beth, Victor Kuperman & Marc Brysbaert. 2013. Norms of valence, arousal, and dominance for 13,915 English lemmas. Behavior Research Methods 45(4). 1191–1207.

Whitsitt, Sam. 2005. A critique of the concept of semantic prosody. *International Journal of Corpus Linguistics* 10(3). 283–305.

Corresponding author:
Chariton Charitonidis
ORCID iD: 0000-0003-0298-2629
dr.chariton.charitonidis@gmail.com