CONTEXT NORMS AND MULTIWORD EXPRESSIONS: REVISITING JOHN SINCLAIR'S ATTITUDINAL FUNCTIONS

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

> "Most divorces and break ups are over money. Few people have sat down with pen and paper and written out their true feelings about money and handed it to their spouse to read."
>
> ("Sani Peyarchi palangal for Rasi Meenam", vasiyam-mandregam.blogspot.com)

Abstract
This study deals with the attitudinal functions 'difficulty' and 'reluctance/inability' that, according to Sinclair (2004), are mapped onto word sequences containing the phrases *naked eye* and *true feelings*, respectively. The methods build on Snefjella & Kuperman (2016) who proposed context norms for English words calculated on ratings of valence, arousal, and concreteness. 149 concordances for *naked eye* from the British National Corpus (BNC) and 2607 concordances for *true feelings* from the English Web Corpus (enTenTen15) were considered. Using the software R as a text-mining tool, values of context valence, arousal, and concreteness were computed. The binary logistic method was applied on deviant subclasses (stages). The results show that a categorical switch of attitudinal function occurs in *true feelings* alone, and only regarding context valence and concreteness.

Keywords: multiword expressions, semantic/evaluative prosody, context norms, emotion

1. Introduction

1.1 Attitudinal functions

According to Sinclair (2004), core words or phrases often appear within longer word sequences to comprise lexical 'units of meaning'. *Units of meaning* refer to attitudinal discourse functions that are "on the pragmatic side of the semantics/pragmatics continuum" (ibid.: 87).

In particular, units of meaning are assigned to templates. In (1) three concordances containing the template '{visible} + [negative] + *naked eye* (core)' are given, taken from Sinclair (2004: 103).[1]

(1) even though nothing is visible to the naked eye. We should trust our patients
human ovum is barely visible to the naked eye. The corpus luteum forms in the
plants that you can see with the naked eye just as much as those for which

The attitudinal function, assigned to all units of meaning in (1) is 'difficulty'. This attitudinal function is evident in 85% of the full set of *naked eye*'s contexts (ibid.: 87-88).

Similarly, the template '{expression} + [possessive adjective] + *true feelings* (core)' bears the attitudinal function 'reluctance-inability' (ibid.: 90), see (2).

(2) we try to communicate our true feelings to those around as
Mary confesses her true feelings to John in the school library
I would never tell anyone my true feelings in fear of being seen as weak

According to Hoey (2005), Louw (1993), Partington (2004), etc., core words or phrases may bear an attitudinal function that is evident in collocates with a *positive* or *negative* attitudinal meaning. If the context of a lexical 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). The same applies if lexical items typically appearing within negative contexts show up in contexts other than negative.

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)

[1] 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).

– *honesty* (among Italian journalists)
– *good taste*

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

1.2 Context norms

In recent years, there has been a considerable focus regarding the interface of lexical meaning and emotion (for a review of relevant studies, see Citron et al. 2016 and Yao et al. 2017). In papers on multiword expressions, special attention was paid to the psycholinguistic variables *valence, arousal,* and *concreteness,* and their interrelations, see Kuperman (2013) on English compounds, Citron et al. (2016) on German idioms, Lindstromberg (2019) on collocations, etc.

In the following, I give Kuperman's (2013: 3) 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)."

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. By mining the 7 billion token USENET corpus compiled by Shaoul & Westbury (2013), Snefjella & Kuperman (2016) obtained valence, arousal, and concreteness values of word

contexts. Each context was confined from five "content words" before to five "content words" after a target word.[2] Contexts in which fewer than three words matched with ratings were excluded.[3] 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 words for which no ratings were available.

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, all context means were averaged across all occurrences of each word in the corpus. The resulting *norms* referred to three meta-variables, i.e. 'context valence', 'context arousal', and 'context concreteness', and can be found in the dataset accompanying Snefjella & Kuperman (2016).

[2] In Snefjella and Kuperman (2016) 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).

[3] 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).

2. Attitudinal functions and context valence

As already implied in section 1.1, attitudinal functions, such as 'difficulty', 'reluctance-inability', 'irony', etc. are not always present. Most notably, however, there are no *empirical* investigations on cancelling or switching attitudinal functions.[4]

In Charitonidis (2021) it was pointed out that the metaphorical (senses of) the phrases *feel blue* 'feel sad or depressed' and *see red* 'get very angry' related to default attitudinal functions ("scenarios"). These attitudinal functions were modulated or switched according to a certain threshold of *context valence*. In literal *see red* 'see the colour red' no attitudinal switch was evident. To be more specific, I would like to provide a few examples from Charitonidis (2021).[5]

Feel blue typically appears in positively-valenced contexts mapped onto the attitudinal function 'comfort', 'provision', etc., see (5).[6]

(5) [feel blue, context valence: 6.23]
Often small gestures open [[large doors of feeling. Sometimes in the morning we **feel blue**, but not for long if there's a flower on the night table or next]] to the bathtub to look at you when you wake up. (COCA\NEWS: Christian Science Monitor. "Petals and Stems Are His Art". 1992)

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 valence: 5.40]
... and those times Lymon would [[slip and make some remark about him would make Lily **feel blue**. Lymon Jr. was drafted into the Army in '66. He went]] 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)

[4] For a theoretical treatise on cancelling or "smoothing" attitudinal functions, see Louw & Chateau (2010).
[5] In Charitonidis (2021), the Corpus of Contemporary American English (COCA) and the News on the Web corpus (NOW) were used (https://www.english-corpora.org).
[6] The arrays of ten "content words" (i.e. non-stopwords) around the core expressions are included in double square brackets. Henceforth, context arrays are also referred to simply as 'contexts'.

Similarly, when metaphorical *see red* appears 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 valence: 4.47]
Finally, her father took custody and moved her to New York. Diandre's and [[Anthony's fathers are both in prison. "For awhile, SharLinda would **see red** and be so upset," Melinda recalls. "I even kept]] 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 attitudinal function 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 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)

In literal *see red,* as already mentioned, no attitudinal switch is evident.

In a nutshell, expressions referring to the physical world seem to call for attitudinal functions in a fundamentally different way than phrases referring to emotions. The latter are prone to attitudinal switches according to thresholds of context valence as opposed to the former.

3. The present study

The present study tests the findings in Charitonidis (2021) empirically, while extending the scope of the analysis to two additional variables. In particular, this study seeks to determine whether or how extreme or moderate values of context valence, arousal, and concreteness correlate to (different) attitudinal functions.

The following working hypotheses will be tested:

1. Phrases referring to the physical world relate to attitudinal functions that are *not switched* by context.
2. Phrases referring to emotion relate to attitudinal functions that *are switched* by context.

I will test these predictions by means of a text-mining and statistical procedure. The objects of investigation will be the phrases *naked eye* (physical world) and *true feelings* (emotion). I have chosen these particular phrases because (a) they are repeatedly referred to in the literature as standard examples of attitudinal function or "semantic prosody" (Louw 1993),[7] and (b) they immediately correspond to the literal and emotion-laden (senses of) phrases already examined in Charitonidis (2021). My methods will largely follow Snefjella & Kuperman (2016) who proposed norms of context valence, arousal and concreteness for English words. The analysis will proceed as follows. Section 4 presents the text-mining and statistical methods used in the present study. Section 5 tests the contexts and attitudinal functions of the phrases *naked eye* and *true feelings*. Section 6 discusses the results. Section 7 explores the patterns of most frequent terms within *naked eye*'s and *true feelings*' contexts. Section 8 gives the conclusions.

4. Methods

The raw concordances for *naked eye* and *true feelings* were obtained using the British National Corpus (BNC) and the English Web corpus 2015 (enTenTen15), respectively, as crawled and processed by Sketch Engine (www.sketchengine.eu).[8] For the string 'true feelings', BNC provided only a statistically insufficient sample of,

[7] In Louw (1993) and in the vast bulk of literature, the term "semantic prosody" is typically referred to as an evaluative function of certain words or multiword expressions appearing within collocates of *positive* or *negative* meaning. In the following analysis, I will avoid this term because it focuses on 'positivity' (or 'valence' in psycholinguistic terms) alone.

[8] According to Sketch Engine's website information [accessed 15 April 2020], "BNC is a 100-million-word collection of samples of a written and spoken language of British English from the later part of the 20th century". The English Web Corpus 2015 (enTenTen15) is an English corpus made up of texts collected from the Internet during the last 10 years. It contains 15 billion words.

approximately, 53 concordances and was not used (www.sketchengine.eu, accessed July 12, 2021).

After removing duplicates, misaligned, or non-identifiable text, 149 concordances for 'naked eye' and 2607 concordances for 'true feelings' were considered.

4.1 Text mining

The text-mining task was accomplished by using the open-source software R (cran.r-project.org). First, all characters were converted to lowercase. The following elements were removed in successive order: (a) punctuation except for apostrophes and intra-word dashes, (b) stopwords, see the default English stopword list of the R *tm*-package, (c) sequences of one or two characters, (d) all apostrophes/quotes, and (e) hyphens/dashes at the beginning and end of words. The output strings were finally tokenized by using the word tokenizer of the R *tokenizers*-package. Following the analysis in Snefjella & Kuperman (2016), the arrays of ten "content words" (i.e. five non-stopwords before and five non-stopwords after the core phrases) were matched with ratings taken from Warriner et al.'s (2013) and Brysbaert et al.'s (2014) norming studies. In *naked eye* the lookup of values (looping) was made with reference to word lemmas. The word lemmas were obtained with reference to the file "lemmatization-en.txt" (www.lexiconista.com). In *true feelings* the lookup of values was firstly made with reference to the given word forms and subsequently to the word lemmas by replacing the NAs obtained from the first loop.

4.2 Independent variables

For each of the context arrays (excluding the core phrases), mean values of context valence, arousal, and concreteness were calculated. Outliers exceeding the three standard-deviations mark were removed. For each context variable, three subsamples of extremely low, moderate, and extremely high values were selected (deviant sampling). The context means in each of the three subsamples were binned into three levels or *stages* of a single ordinal variable (=the predictor). Tables 2 and 3 below show the lower and upper bounds of stages, for *naked eye* and *true feelings*, respectively. The lower bounds of the middle stages of context valence, i.e. 5.69 for *naked eye* and 5.70 for *true feelings*, begin after the value 5.65, suggested by

Charitonidis (2021) as a threshold for an attitudinal switch (see section 2). The lower bounds of the middle stages of context arousal and concreteness begin after the mean/median of the source samples (see Tables 4 and 5 in section 5).

Table 2. Stages of context variables (naked eye)

		Context valence	Context arousal	Context concreteness
Low stage (N=20)	Minimum	4.62	3.33	2.14
	Maximum	5.28	3.60	2.69
Middle stage (N=15)	Minimum	5.69	3.93	3.22
	Maximum	5.78	3.98	3.29
High stage (N=20)	Minimum	6.15	4.24	3.63
	Maximum	6.63	4.74	4.13

Table 3. Stages of context variables (true feelings)

		Context valence	Context arousal	Context concreteness
Low stage (N=30)	Minimum	4.19	3.16	1.90
	Maximum	4.53	3.39	2.10
Middle stage (N=30)	Minimum	5.70	4.11	2.87
	Maximum	5.72	4.12	2.88
High stage (N=30)	Minimum	6.82	4.95	3.69
	Maximum	7.22	5.17	3.98

4.3 Dependent variables

The attitudinal functions 'difficulty' for *naked eye* and 'reluctance/inability' for *true feelings*, were assessed as dependent *categorical* variables (=the outcome). The positivity/negativity of attitudinal functions was determined by manually inspecting the context arrays at all stages, including stopwords and core phrases. In total, 165 context arrays for *naked eye* and 270 context arrays for *true feelings* were examined.[9] During this procedure, extended contexts were ignored.[10]

[9] These totals redundantly include a very small number of context arrays that were referred to by more than one context variable.

[10] All context arrays, together with their means and their indications regarding positive or negative attitude, can be found in the supplementary-data files on the website Researchgate.net.

In (9) and (10) below, examples of positive and negative outcomes are given, for *naked eye* and *true feelings*, respectively.

(9) *naked eye*

a. +difficulty

[[bigger than his little fingernail, were totally meaningless to the **naked eye.** Altogether there were thirty-eight certificates, each for somebody who had died]] (Concordance 18)

(Gavin Lyall: "The Conduct of Major Maxim")

b. -difficulty

[[finest clusters of its type; it is easily visible with the **naked eye** and is well seen with × 7 binoculars, while with × 12 and × 20 it is truly glorious.]] (Concordance 81)

(Patrick Moore: "Exploring the Night Sky with Binoculars")

(10) *true feelings*

a. +reluctance/inability

[[Cooper is the one most smitten, though he hides his **true feelings** until the inevitable clinch. When gangster Andrews and his torpedo]] (Concordance 1331)

(Review of Howard Hawks' film "Ball Of Fire (Bola de fuego)", sasquatchvideo.net)

b. -reluctance/inability

[[talks about being so in love with someone that only God knows the **true feelings** of how that person feels. I love singing along]] (Concordance 197)

(Review of The Beach Boys' song "God only knows", www.retro-daze.org)

In the statistical analysis proper, the binary logistic regression method was used. This method was suggested by the dichotomous categorical output and the three-stage predictor.

5. Results

Tables 4 and 5 below show the descriptives of *naked eye*'s and *true feelings*' contexts, respectively. In *naked eye* (Table 4), context valence was slightly positive (5.73), context arousal was slightly negative/low (3.92), whereas context concreteness referred, approximately, to a middle point between abstract and concrete (3.18). Similarly, in *true feelings* (Table 5), context valence was slightly positive (5.76), context arousal was slightly negative/low (4.11), whereas context concreteness referred, approximately, to a middle point between abstract and concrete (2.86).

In a nutshell, the mean (and median) values for both phrases were very similar. It should be noted, however, that in *naked eye* the range of context valence, i.e. 2.00, was considerably smaller than the range of context valence in *true feelings*, i.e. 3.04.

Table 4. Descriptive statistics of context variables (naked eye)

		Context valence	Context arousal	Context concreteness
N	Valid	148	149	149
	Outliers	1	0	0
Mean		5.73	3.92	3.18
Median		5.77	3.90	3.19
Std. Deviation		0.37	0.28	0.39
Range		2.00	1.41	1.98
Minimum		4.62	3.33	2.14
Maximum		6.63	4.74	4.13

(Pairwise exclusion)

Table 5. Descriptive statistics of context variables (true feelings)

		Context valence	Context arousal	Context concreteness
N	Valid	2596	2595	2599
	Outliers	11	12	8
Mean		5.76	4.11	2.86
Median		5.78	4.10	2.85
Std. Deviation		0.51	0.34	0.37
Range		3.04	2.01	2.08
Minimum		4.19	3.16	1.90
Maximum		7.22	5.17	3.98

(Pairwise exclusion)

Table 6 below shows the proportions of ±difficulty in *naked eye*'s contexts and Table 7 shows the proportions of ±reluctance/inability in *true feelings*' contexts, with reference to the particular stages. As can be seen, in *true feelings* there was an explicit preponderance of negative outcomes at the high stages of both context valence and concreteness, i.e. 21 negative vs. 9 positive outcomes for both variables.

Table 6. +difficulty vs. -difficulty (naked eye)

	Low stage (N=20)		Middle stage (N=15)		High stage (N=20)		All stages (N=55)	
Difficulty	+	–	+	–	+	–	+	–
Context valence	14	6	9	6	9	11	32	23
Context arousal	7	13	10	5	10	10	27	28
Context concreteness	11	9	10	5	15	5	36	19

Table 7. +reluctance/inability vs. -reluctance/inability (true feelings)

	Low stage (N=30)		Middle stage (N=30)		High stage (N=30)		All stages (N=90)	
Reluctance/Inability	+	–	+	–	+	–	+	–
Context valence	20	10	17	13	9	21	46	44
Context arousal	10	20	13	17	13	17	36	54
Context Concreteness	17	13	17	13	9	21	43	47

Let us now proceed to the main results of this study.

5.1 Naked eye

Table 8 below displays the results of the binary logistic regression analysis for *naked eye*. *B* (the regression coefficient) indicates change in log odds and is not directly interpretable. It should be considered as a directional index of greater or lower likelihood (positive b vs. negative b, respectively). In simple terms, a positive b value indicates that the likelihood of difficulty increases between stages, and a negative b value indicates that the likelihood of difficulty decreases between stages.

Table 8. Context valence, arousal, and concreteness vs. difficulty (naked eye)

Stage comparisons	Context valence	**B**	Context arousal	**B**	Context concreteness	**B**
All		-0.53 ns		0.30 ns		0.45 ns
Low-High		-1.05 ns		0.62 ns		0.90 ns
Middle-High		-0.61 ns		-0.69 ns		0.41 ns
Low-Middle		-0.44 ns		1.31 ns		0.49 ns

ns = p > .05

As can be seen, all stage comparisons were not significant, p > .05. As to context valence, - difficulty was more likely between the low and high stage. This difference, however, did not reach statistical significance, b = -1.05 ns. Regarding context arousal, +difficulty was more likely between the low and middle stage. Again, this difference did not reach statistical significance, b = 1.31 ns. Finally, regarding context concreteness, +difficulty was more likely between the low and high stage. Again, this difference did not reach statistical significance, b = 0.90 ns.

5.2 True feelings

Table 9 below displays the results of the binary logistic regression analysis for *true feelings*. A zero b value indicates that the predictor and the outcome were independent, i.e. the proportions of positive and negative outcomes were equal, see Table 7 earlier in this section.

Table 9. Context valence, arousal, and concreteness vs. reluctance/inability (true feelings)

Stage comparisons	Context valence	**B**	Context arousal	**B**	Context concreteness	**B**
All		-0.77**		0.21 ns		-0.55*
Low-High		-1.54**		0.43 ns		-1.12*
Middle-High		-1.17*		0.00 ns		-1.12*
Low-Middle		-0.43 ns		0.43 ns		0.00 ns

* p < .05, ** p < .01, ns = p > .05

Regarding context valence, all differences between stages were significant except for the difference between the low and middle stage, b = -0.43, ns. -reluctance/ inability was more likely at the high stage as compared to both low stage, b = -1.54, p

= .006, and middle stage, b = -1.17, p = .040. Regarding context arousal, all differences between stages were not significant, p > .05. As to context concreteness, all differences between stages were significant except for the difference between the low and middle stage, 0.00 ns. -reluctance/inability was more likely at the high stage as compared to both low stage, b = -1.12, p = .040, and middle stage, b = -1.12, p = .040. Most notably, the significant effects of both context valence and concreteness emerged from the same stage comparisons. However, the effect of context valence was considerably stronger at the extreme stages.

6. Discussion

The working hypotheses set out in section 3 were, for the most part, confirmed. The results of this study are: (a) *naked eye*, as a phrase referring to the physical world, relates to an attitudinal function that is not switched categorically by context valence, arousal, or concreteness (section 5.1), and (b) *true feelings*, as a phrase referring to emotion, relates to an attitudinal function that is switched categorically by context valence and concreteness (section 5.2). Let us now discuss the particulars of these results.

A. *Context valence.* Regarding *true feelings*, the difference between the low and high stage referred to a higher significance level, i.e. -1.54, p = .006, than the difference between the middle and high stage, i.e. -1.17, p = .040. Accordingly, an incremental pattern of context valence toward -reluctance/inability is highly probable. It is also likely that a value around 5.65 qualifies as a threshold for an attitudinal switch (see section 2). In simple terms, and perhaps overgeneralizing, positive situations seem to call for "genuine emotions" (-reluctance/inability), whereas negative situations seem to suppress "genuine emotions" (+reluctance/inability).

In *naked eye*, as already mentioned, no interaction of context valence and attitudinal function was detected.

B. *Context arousal.* In both *naked eye* and *true feelings*, all differences between stages were not significant. What can be deduced from this is that (a), the difficulty seeing an object is dissociated from an arousing or non-arousing situation (*naked eye*), and (b) the likelihood of freely expressing "genuine emotions" is equal in both

situations that enhance one's mood and situations that weaken one's mood (*true feelings*).

C. *Context concreteness.* In *true feelings,* the comparisons of both low stage with high stage and middle stage with high stage showed the same significant difference, i.e. b = -1.12, p = .040. This flat pattern is opposed to the incremental pattern of *context valence* in which the difference between the low and high stage referred to a higher significance level than the difference between the middle and high stage (see A. above). Accordingly, again with reference to context concreteness, a shift toward -reluctance/inability seems to be both less consistent and most likely to occur at a later point of the three-stage paradigm, perhaps close before the high stage. Further research is needed to clarify this issue.

In *naked eye,* as already pointed out, no interaction of context concreteness and attitudinal function was detected.

7. Exploring stages: Term frequencies

Sections 5 and 6 have shown that the attitudinal function expressed by *true feelings* switches categorically according to stages of context valence and concreteness. In particular, context valence referred to an incremental pattern of significant stage differences, whereas context concreteness was a significant but weaker predictor. In *naked eye,* no interactions were found. The aim of this section is to justify these patterns independently. Accordingly, the substratum of context variables will be examined by means of a usage-based task.

Before I proceed to the main task, I would like to produce a relevant set of frequent content words (=non-stopwords, see the default English stopword list of the R *tm*-package) at the extreme stages of context valence and concreteness (FREQ≥3, where 'FREQ' indicates token frequency. Henceforth, word-level valence, arousal, and concreteness are abbreviated as V, A, and C, respectively.)[11]

(a) *naked eye.* At the high stage of context valence, 19 out of 33 tokens were explicitly positive, cf. the terms *see* (V=6.27, FREQ=9), *good* (V=7.89, FREQ=5), and *star* (V=7.47, FREQ=5). Similarly, at the high stage of context concreteness, 19

[11] It is reminded that the valence and arousal ratings come from Warriner et al. (2013), and the concreteness ratings come from Brysbaert et al. (2014).

out of 35 tokens were explicitly positive, cf. the terms *binoculars* (C=5, FREQ=4), *can* (C=4.55, FREQ=4), *glass* (C=4.82, FREQ=4), *star* (C=4.69, FREQ=4), and *egg* (C=4.97, FREQ=3).

(b) *true feelings.* At the low stage of context valence, 15 out of 33 tokens were explicitly negative, cf. the terms *fear* (V=2.93, FREQ=5), *hate* (V=1.96, FREQ=4), *anger* (V=2.5, FREQ=3) and *deny* (V=3.81, FREQ=3). At the high stage of the same variable, 40 out of 61 tokens were explicitly positive, cf. the terms *love* (V=8, FREQ=11), *good* (V=7.89, FREQ=9), *give* (V=7.73, FREQ=8), *courage* (V=7.76, FREQ=3), *family* (V=7.25, FREQ=3), *live* (V=7.95, FREQ=3), and *relationship* (V=7.83, FREQ=3). Regarding context concreteness, 44 out of 47 tokens at the low stage were abstract having a concreteness value below '3' (exception: *base,* C=3.86, FREQ=3). In contrast, *all* most frequent terms at the high stage, i.e. 42 tokens, had a concreteness value above '3'.

The patterns in (a) and (b) show that in *true feelings* opposed terms of very low and very high values of valence and concreteness tend to cluster under corresponding stages of very low and very high values of context valence and context concreteness. In *naked eye,* these aligned patterns did not occur. For a fully-fledged picture of the most frequent terms, the reader is referred to the Appendix.

Tables 10 and 11 below display the descriptive statistics for content words at the three stages of each context variable, for *naked eye* and *true feelings*, respectively. 'N' refers to the original number of content words in the context arrays of every stage, excluding the core phrases. 'Σ' indicates grand sums of valence, arousal, and concreteness for all content words at each stage. 'FREQ' refers to the totals of tokens for content words with FREQ≥3. 'SUM' refers to the totals of valence, arousal, and concreteness for content words with FREQ≥3. The proportions of valence, arousal, and concreteness at each stage are indicated by 'SUM/Σ' and spelled out in percentages.

Table 10. Descriptive statistics for content words in the contexts of *naked eye*

	Low Stage (N=200)				Middle Stage (N=150)				High Stage (N=200)			
	Σ	FREQ	SUM	**SUM/Σ (%)**	Σ	FREQ	SUM	**SUM/Σ (%)**	Σ	FREQ	SUM	**SUM/Σ (%)**
V	724.65	23	90.09	**12.43**	591.26	13	41.85	**7.08**	853.31	33	158.10	**18.53**
A	504.62	39	106.24	**21.05**	435.45	18	66.86	**15.35**	618.5	32	116.17	**18.78**
C	458.46	35	84.29	**18.39**	453.26	23	88.21	**19.46**	702.1	35	142.93	**20.36**

Table 11: Descriptive statistics for content words in the contexts of *true feelings*

	Low Stage (N=300)				Middle Stage (N=300)				High Stage (N=300)			
	Σ	FREQ	SUM	**SUM/Σ (%)**	Σ	FREQ	SUM	**SUM/Σ (%)**	Σ	FREQ	SUM	**SUM/Σ (%)**
V	906.93	33	103.59	**11.42**	1289.96	53	251.62	**19.51**	1560.8	61	449.00	**28.77**
A	644.13	57	156.42	**24.28**	967.34	26	115.02	**11.89**	1062.85	18	92.90	**8.74**
C	556.21	47	97.51	**17.53**	808.13	26	62.35	**7.72**	947.83	42	159.72	**16.85**

The proportions of valence, arousal, and concreteness of most frequent terms (see 'SUM/Σ' in Tables 10 and 11) were compared by means of N-1 chi-square tests. Tables 12 and 13 below display the results, respectively.

Table 12. Proportions of most frequent terms: Stage comparisons (naked eye)

	Low-Middle		Middle-High		Low-High	
	Difference (%)	$\chi^2(1)$	Difference (%)	$\chi^2(1)$	Difference (%)	$\chi^2(1)$
V	5.35	0.25 ns	11.45	0.92 ns	6.10	0.37 ns
A	5.70	0.25 ns	3.43	0.09 ns	2.27	0.06 ns
C	1.07	0.01 ns	0.90	0.01 ns	1.97	0.04 ns

FREQ≥3. ns = p > .05

Table 13. Proportions of most frequent terms: Stage comparisons (true feelings)

	Low-Middle		Middle-High		Low-High	
	Difference (%)	$\chi^2(1)$	Difference (%)	$\chi^2(1)$	Difference (%)	$\chi^2(1)$
V	8.09	0.96 ns	9.26	1.30 ns	17.35	3.64*
A	12.39	1.67 ns	3.15	0.11 ns	15.54	2.00 ns
C	9.81	1.32 ns	9.13	1.14 ns	0.68	0.01 ns

FREQ≥3. * p = .0565, ns = p > .05

The tests for *naked eye* in Table 12 showed that all differences in proportions between stages and across all three variables were non-significant, in accord with the effects of context variables (see section 5.1).

The results for *true feelings* in Table 13 referred to a single significant difference in valence between the low and high stage, i.e. $\chi^2(1) = 3.64$, p = .057. The significant difference between the middle and high stage of context valence was not addressed by these results (see section 5.2). Nonetheless, this single difference suggests that valence of most frequent terms is a relevant indication of (a) extreme

stages of the respective context variable, and (b) opposite values of attitudinal function.

The absence of concreteness effects in Table 13 may be associated with the flat significant differences between the stages of context concreteness pointed out in section 5.2. It should be noted, however, that the opposed terms regarding word concreteness (see (b) above) could perhaps serve as an indication of extreme stages of context concreteness and, concomitantly, opposite values of attitudinal function.

8. Conclusion

This case study was the first attempt in the literature to investigate empirically how attitudinal functions of multiword expressions are cancelled or switched categorically according to affective (valence, arousal) and sensorimotor (concreteness) context norms. The objects of investigation were the phrases *naked eye* and *true feelings*, originally dealt with by John McHardy Sinclair.

By following the methods in Snefjella & Kuperman (2016), this paper extended (a) the scope of attitudinal function from Sinclair's (2004) 'units of meaning' to context arrays of five content words before to five words after a core expression, and (b) the semantic range of attitudinal function from valence (positivity) to two additional variables, i.e. arousal and concreteness.

In the multiword expression *true feelings* expressing emotion, the significant effects of both context valence and concreteness emerged at the same stages, while switching the attitudinal function categorically. At the extreme stages, however, the effects of context valence were stronger. In the multiword expression *naked eye* referring to the physical world, a categorical attitudinal switch was not detected.

An additional usage-based task with most frequent terms showed that in *true feelings* the significant effects of word valence corresponded to the significant effects of context valence at the extreme stages. The absence of word concreteness effects is likely to relate to the flat significant differences between the stages of context concreteness.

In both *naked eye* and *true feelings*, no interaction between context arousal and attitudinal function was detected. There were also no term-frequency effects.

Concluding, future research should extend the present analysis to a large set of multiword expressions associated with attitudinal functions. The findings of such a research could enhance our understanding of the attitudes, emotions, and intentions of discourse agents.

References

Brysbaert, M., A.B. Warriner & V. Kuperman (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods* 46: 904-911. https://doi.org/10.3758/s13428-013-0403-5

Charitonidis, C. (2021). Semantic prosody of extended lexical units: A case study. *Yearbook of the Poznań Linguistic Meeting* 7(1): 43-74. https://doi.org/10.14746/yplm.2021.7.2

Citron, F.M.M., C. Cacciari, M. Kucharski, L. Beck, M. Conrad & A.M. Jacobs (2016). When emotions are expressed figuratively: Psycholinguistic and affective norms of 619 idioms for German (PANIG). *Behavior Research Methods* 48: 91-111. https://doi.org/10.3758/s13428-015-0581-4

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

Hunston, S. (2007). Semantic prosody revisited. *International Journal of Corpus Linguistics* 12(2): 249-268. https://doi.org/10.1075/ijcl.12.2.09hun

Kuperman, V. (2013). Accentuate the positive: Diagnostics of semantic access in English compounds. *Frontiers in Psychology* 4: 203. https://doi.org/10.3389/fpsyg.2013.00203

Lindstromberg, S. (2019). Ratings of the emotional valence and arousal of collocations and their constituent words: How can they be useful in L2 vocabulary research? *System* 87. https://doi.org/10.1016/j.system.2019.102144

Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis & E. Tognini-Bonelli (eds.), *Text and technology: In honour of John Sinclair.* Amsterdam: Benjamins, 157-176.

Louw, B. & C. Chateau (2010). Semantic prosody for the 21st century: Are prosodies smoothed in academic contexts? A contextual prosodic theoretical perspective. In S. Bolasco, I. Chiari & I. Luca (eds.), *JADT 2010: Statistical Analysis of Textual Data – Proceedings of the 10th International Conference.* Rome: LED Edizioni Universitarie.

Morley, J. & A. Partington (2009). A few Frequently Asked Questions about semantic – or evaluative – prosody. *International Journal of Corpus Linguistics* 14(2): 139-158. https://doi.org/10.1075/ijcl.14.2.01mor

Partington, A. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference. *International Journal of Corpus Linguistics* 9(1): 131-156. https://doi.org/10.1075/ijcl.9.1.07par

Shaoul, C. & C. Westbury (2013). *A reduced redundancy usenet corpus (2005-2011).* Edmonton, AB: University of Alberta. http://www.psych.ualberta.ca/~westburylab/downloads/usenetcorpus.download.html [accessed 23 January, 2021].

Sinclair, J.M. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge.

Snefjella, B. & V. Kuperman (2016). It's all in the delivery: Effects of context valence, arousal, and concreteness on visual word processing. *Cognition* 156: 135-146. https://doi.org/10.1016/j.cognition.2016.07.010

Stubbs, M. (2009). The search for units of meaning: Sinclair on empirical semantics (memorial article). *Applied Linguistics* 30(1): 115-137. https://doi.org/10.1093/applin/amn052

Warriner, A.B., V. Kuperman & M. Brysbaert (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods* 45(4): 1191-1207. https://doi.org/10.3758/s13428-012-0314-x

Yao, Z., J. Wu, Y. Zhang & Z. Wang (2017). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behavior Research Methods* 49(4): 1374-1385. https://doi.org/10.3758/s13428-016-0793-2

Corpora

British National Corpus (BNC) (www.sketchengine.eu)

English Web corpus 2015 (enTenTen15) (www.sketchengine.eu)

Appendix: Most frequent terms

V1, V2, V3: Patterns of word valence at the stages "Low", "Middle", and "High" of context valence.
A1, A2, A3: Patterns of word arousal at the stages "Low", "Middle", and "High" of context arousal.
C1, C2, C3: Patterns of word concreteness at the stages "Low", "Middle", and "High" of context concreteness.
N: Original sample sizes of all content words in the context arrays of each stage.
Σ: Grand sums of valence, arousal, or concreteness for all content words at each stage.
FREQ: Token frequencies of content words (FREQ≥3).
SUM: Sums of valence, arousal, or concreteness for context words with FREQ≥3.
SUM/Σ: Proportions of valence, arousal, and concreteness for context words with FREQ≥3.

A. naked eye

V1 (N=200, Σ=724.65)	FREQ	SUM	V2 (N=150, Σ=591.26)	FREQ	SUM	V3 (N=200, Σ=853.31)	FREQ	SUM
visible (5.76)	6	34.56	visible (5.76)	4	23.04	see (6.27)	9	56.43
just (NA)	4	NA	even (NA)	3	NA	binoculars (5.57)	6	33.42
little (5.91)	4	23.64	less (NA)	3	NA	good (7.89)	5	39.45
large (5.77)	3	17.31	see (6.27)	3	18.81	star (7.47)	5	37.35
long (NA)	3	NA				visible (5.76)	5	28.80
worm (4.86)	3	14.58				easily (NA)	3	NA
Total:	**23**	**90.09**		**13**	**41.85**		**33**	**158.10**
SUM/Σ:		**12.43%**			**7.08%**			**18.53%**

A1 (N=200, Σ=504.62)	FREQ	SUM	A2 (N=150, Σ=435.45)	FREQ	SUM	A3 (N=200, Σ=618.5)	FREQ	SUM
look (3.76)	6	22.56	visible (4.15)	5	20.75	visible (4.15)	8	33.2
binoculars (2.58)	5	12.9	binoculars (2.58)	4	10.32	star (5.5)	5	27.5
can (3.14)	5	15.7	cluster (4.21)	3	12.63	easily (NA)	4	NA
pair (3.35)	5	16.75	easy (3.82)	3	11.46	bright (5)	3	15
visible (4.15)	5	20.75	see (3.9)	3	11.7	cluster (4.21)	3	12.63
eta (NA)	4	NA				even (NA)	3	NA
area (2.19)	3	6.57				million (5.38)	3	16.14
even (NA)	3	NA				see (3.9)	3	11.7
make (3.67)	3	11.01						
Total:	**39**	**106.24**		**18**	**66.86**		**32**	**116.17**
SUM/Σ:		**21.05%**			**15.35%**			**18.78%**

C1 (N=200, Σ=458.46)	FREQ	SUM	C2 (N=150, Σ=453.26)	FREQ	SUM	C3 (N=200, Σ=702.1)	FREQ	SUM
see (3.21)	8	25.68	star (4.69)	6	28.14	visible (3.08)	9	27.72
good (1.64)	6	9.84	visible (3.08)	5	15.4	binoculars (5)	4	20
even (2.79)	5	13.95	binoculars (5)	3	15	can (4.55)	4	18.2
visible (3.08)	4	12.32	invisible (2.83)	3	8.49	glass (4.82)	4	19.28
almost (1.66)	3	4.98	nova (3.85)	3	11.55	see (3.21)	4	12.84
know (1.68)	3	5.04	see (3.21)	3	9.63	star (4.69)	4	18.76
look (2.96)	3	8.88				egg (4.97)	3	14.91
though (1.2)	3	3.6				faint (3.74)	3	11.22
Total:	**35**	**84.29**		**23**	**88.21**		**35**	**142.93**
SUM/Σ:		**18.39%**			**19.46%**			**20.36%**

B. true feelings

V1 (N=300, Σ=906.93)	FREQ	SUM	V2 (N=300, Σ=1289.96)	FREQ	SUM	V3 (N=300, Σ=1560.8)	FREQ	SUM
fear (2.93)	5	14.65	express (5.53)	5	27.65	love (8)	11	88
hide (4.9)	5	24.5	find (6.45)	4	25.8	good (7.89)	9	71.01
hate (1.96)	4	7.84	hide (4.9)	4	19.6	give (7.73)	8	61.84
tell (5.27)	4	21.08	may (5.55)	4	22.2	able (6.64)	3	19.92
anger (2.5)	3	7.5	really (NA)	4	NA	art (6.85)	3	20.55
conjurer (NA)	3	NA	show (5.91)	4	23.64	courage (7.67)	3	23.01
deny (3.81)	3	11.43	word (5.77)	4	23.08	face (6.36)	3	19.08
express (5.53)	3	16.59	can (6.41)	3	19.23	family (7.25)	3	21.75
israel (NA)	3	NA	even (NA)	3	NA	live (7.95)	3	23.85
			feel (6.27)	3	18.81	relationship (7.83)	3	23.49
			keep (6.32)	3	18.96	see (6.27)	3	18.81
			know (6.82)	3	20.46	show (5.91)	3	17.73
			put (5.09)	3	15.27	talk (6.64)	3	19.92
			sakura (NA)	3	NA	understand (6.68)	3	20.04
			try (5.64)	3	16.92			
Total:	**33**	**103.59**		**53**	**251.62**		**61**	**449.00**
SUM/Σ:		**11.42%**			**19.51%**			**28.77%**

A1 (N=300, Σ=644.13)	FREQ	SUM	A2 (N=300, Σ=967.34)	FREQ	SUM	A3 (N=300, Σ=1062.85)	FREQ	SUM
will (2.9)	10	29	come (3.57)	4	14.28	sex (7.6)	5	38
let (2.71)	7	18.97	reveal (4.14)	4	16.56	discover (5.7)	4	22.8
know (3.24)	6	19.44	emotion (4.75)	3	14.25	confront (6.11)	3	18.33
one (2.67)	6	16.02	express (4.79)	3	14.37	face (4.59)	3	13.77
make (3.67)	5	18.35	give (4.57)	3	13.71	just (NA)	3	NA
way (2.9)	5	14.5	heart (5.07)	3	15.21			
hide (3.52)	3	10.56	hide (3.52)	3	10.56			
keep (3.43)	3	10.29	love (5.36)	3	16.08			
may (3.19)	3	9.57						
realize (3.24)	3	9.72						
sakura (NA)	3	NA						
though (NA)	3	NA						
Total:	**57**	**156.42**		**26**	**115.02**		**18**	**92.90**
SUM/Σ:		**24.28%**			**11.89%**			**8.74%**

C1 (N=300, Σ=556.21)	FREQ	SUM	C2 (N=300, Σ=808.13)	FREQ	SUM	C3 (N=300, Σ=947.83)	FREQ	SUM
good (1.64)	5	8.2	express (2.21)	6	13.26	can (4.55)	6	27.3
just (1.52)	5	7.6	good (1.64)	5	8.2	people (4.82)	5	24.1
get (2.38)	4	9.52	hide (3.21)	3	9.63	write (4.22)	4	16.88
love (2.07)	4	8.28	reveal (2.79)	3	8.37	heart (4.52)	3	13.56
need (1.69)	4	6.76	say (2.58)	3	7.74	hide (3.21)	3	9.63
really (1.44)	4	5.76	tell (2.9)	3	8.7	read (3.56)	3	10.68
base (3.86)	3	11.58	without (2.15)	3	6.45	receive (2.69)	3	8.07
begin (2.56)	3	7.68				sakura (NA)	3	NA
beyond (1.72)	3	5.16				short (3.61)	3	10.83
courage (1.52)	3	4.56				talk (4.07)	3	12.21
other (2.04)	3	6.12				woman (4.46)	3	13.38
reveal (2.79)	3	8.37				world (4.36)	3	13.08
will (2.64)	3	7.92						
Total:	**47**	**97.51**		**26**	**62.35**		**42**	**159.72**
SUM/Σ:		**17.53%**			**7.72%**			**16.85%**