Triangulating the GRID:

A corpus-based cognitive linguistic analysis of five Greek emotion terms

Marina Terkourafi

University of Illinois at Urbana-Champaign

Efthymia C. Kapnoula

University of Athens, Greece

Penny Panagiotopoulou

University of Patras, Greece

Athanassios Protopapas

Institute for Language & Speech Processing, Greece

Terkourafi, M., Kapnoula, E.C., Panagiotopoulou, P., & Protopapas, A. (in press). Triangulating the GRID: A corpus-based cognitive linguistic analysis of five Greek emotion terms. In J. Fontaine, K. R. Scherer, & C. Soriano (Eds.), *Components of emotional meaning: A sourcebook*. Oxford University Press.

Abstract

This chapter discusses two different approaches to the meaning of emotion terms, corpus-based linguistic analysis and feature profiling as carried out within the GRID, in order to compare their results and explore the potential for cross-fertilisation between them. To this end, we present an analysis of five Greek emotion terms (*aghonia*, 'anguish', *erotas*, 'romantic love', *siginisi*, 'being touched or moved by sth', *lipi*, 'sorrow', and *stenokhorja*, 'chagrin') investigated using both the GRID and a corpus-based methodology. For the needs of the latter, occurrences of each term were located in the Hellenic National Corpus and analysed within Halliday's framework of systemic functional grammar, paying particular attention to his notion of grammatical metaphor (Halliday 1998), in virtue of which human experience is re-packaged into grammatical functions within the clause (subject, complement, etc.). We found that the corpus analyses offer independent support for the GRID results. In addition, they afford us with original insights into the meanings of emotion terms, and in particular into differences between terms that remained indistinguishable on the basis of the GRID (near-synonyms). In conclusion, we argue for the usefulness of combining the two approaches as a way of arriving at a more comprehensive picture of the meanings of emotion terms.

1. Introduction

"Although language is abstracted from human experience, it must correspond to human experience and represent important human concerns." (Fontaine *et al.* 2007: 1056)

The aim of correlating subjective experience with its lexicalisation by means of particular emotion terms in a language is at the heart of the GRID (Fontaine et al. 2007). To achieve this, the GRID builds semantic profiles of prototypical emotion terms by asking questionnaire respondents to rate each term on over a hundred features capturing several dimensions of variation in emotional experience across languages/cultures (2007:1050).

Problems can, nevertheless, arise if the initial process of translation/back-translation used to identify the translation equivalents of the 24 prototypical emotion terms delivers more than one translation equivalent for the same term in different languages. In our data, this was the case with the Greek terms *lipi* ('sorrow') and *stenokhorja* ('chagrin'), both suggested as translation equivalents of the English term 'sadness' by our bilingual translators.

Our search for a solution to this practical problem prompted us to explore an alternative approach to the meaning of emotion terms, namely, a corpus-based cognitive linguistic approach that probes the meaning of emotion terms in context. In this chapter, we present the results of that exploration and compare them with those obtained by the GRID. Our evidence illustrates the different strengths and weaknesses of the two approaches and suggests that they should be seen as complementary. In other words, a more comprehensive picture of the meaning of emotion terms emerges if the two methodologies are used *in tandem*.

2. Two approaches to the meaning of emotion terms

A major theme of emotion research to date has been establishing a small set of basic dimensions that underlies the conceptualisation of emotions in different languages/cultures. This has led to

an important motivation for proposing the GRID methodology, namely "[t]o obtain definitive evidence concerning the optimal low-dimensional space" against which emotions are conceptualised cross-linguistically (Fontaine et al. 2007: 1050). By enabling large-scale comparison of the semantic profiles of prototypical emotion terms based on empirical data, the GRID methodology represents an important step toward reaching this goal and has already made significant contributions to this debate, most notably by revealing that four, rather than two, dimensions are necessary in order to account for variation in emotional experience cross-linguistically (Fontaine et al. 2007, this volume?).

To analyse the meaning of emotion terms, the GRID builds on the linguistic semantic tradition of componential analysis (Goodenough 1956, Lounsbury 1956) or decompositional semantics (e.g., Katz and Fodor 1963, Jackendoff 1972), in which word meanings are analytically broken down into smaller components (or features) intended to capture the atomic elements of their meaning.

Recently, an alternative view of word meaning has emerged within cognitive linguistics, on which there is no easy distinction between lexical meaning and encyclopaedic knowledge (Peeters 2000, Taylor 2003). On the cognitive linguistic view, linguistic semantics (encoded meaning) emerges out of pragmatics (use), and word meanings are "a network of shared, conventionalised, and to some extent perhaps idealised knowledge, embedded in a pattern of cultural beliefs and practices" (Taylor 2003: 86). Moreover, word meanings are dynamic cognitive structures whose meaning can only be understood in the context of other cognitive structures (such as domains, frames and schemata) that also extent beyond the language system itself (2003: 87). In other words, linguistic semantic information seriously underdetermines word meaning, leaving significant scope for contextual inference to fill in the gap (Recanati 2003).

This view has been gaining ground in linguistic circles and is currently being explored in the rapidly growing field of lexical pragmatics (Wilson 2003, Wilson and Carston 2007).

The difference between contextualised and decontextualised approaches to word meaning is also highlighted by the proponents of the GRID. Discussing a distinction (proposed by Robinson and Clore, 2002) "between current emotion, which is episodic, experiential, and contextual, and beliefs about emotions which are semantic, conceptual, and decontextualised," they state that "clearly, by design, our data on semantic profiles belong to the latter category" (Fontaine et al. 2007: 1056). This remark leaves open the possibility that an approach which analyses the meaning of emotion terms in context may be a welcome addition on the side of the GRID that could reveal interesting new generalisations about the meaning of emotion terms.

In setting out to discover what a contextualised approach may have to add to the analysis of emotion terms undertaken by the GRID, we adopt a corpus-based methodology that exploits the increasing availability of natural language corpora as extensive repositories of longer stretches of discourse produced spontaneously by a variety of speakers/authors. In this way, two desiderata of the cognitive linguistic enterprise can be met at once: (a) the study of the meaning of emotion terms in context, and (b) the investigation of a sufficiently diverse pool of informants, since the only prerequisite to considering a piece of (encyclopaedic) knowledge as part of the meaning of a word is that it be "shared by a sufficient number of people" (Taylor 2003: 93).

Corpus-based methodologies present us with the opportunity of analysing the meaning of emotion terms by studying the grammatical relationships that they contract with other constituents within the clause. To find out what these relationships can reveal about their semantic structure, we adopt Halliday's framework of systemic functional grammar as applied to emotion research in English (Halliday 1998) and Greek (Lascaratou 2007; Terkourafi and Bali

2007). The central insight of this approach is that grammatical structure re-packages experience into relations within the clause. This is what Halliday calls the 'experiential' function of the clause, that is, "its guise as a way of representing patterns of experience" (1985/1994: 106). This is accomplished through grammatical metaphor (ibid.: 340-367), a term coined by Halliday to refer to variation in how meanings are *expressed* rather than in the meanings themselves (ibid: 341).

To understand how grammatical metaphor works, an example by Halliday himself may be useful. In analysing the meaning of *I have a headache*, he explains:

> Here the grammar constructs an entity, a kind of thing, called an *ache*; it then uses a part of the body to assign this ache to a class, *head+ache*, which it constructs into a composite thing called a *headache*. ... The grammar then sets up a structural configuration of possession [...]. Some person (usually the speaker) becomes the owner of this thing [...] and someone else can ask them *how's your headache?*, with *you* as possessive Deictic. (Halliday 1998: 3-4)

But why should the grammar favour wordings like *I have a headache* over *My head aches*? The answer to this second question, according to Halliday, lies in information structure preferences, specifically the default preference in English to present in initial clause position the 'theme' of the message (that which is being talked about) and use the rest of the message to say something about it. In *I have a headache,*

> …the setting of this unpleasant experience, is not my head, it is me — my self as a whole. So the grammatical Theme of the clause ought to be 'me'. Therefore, since it is the first element of clause structure [...] that is thematic, this 'me' has to figure by itself as a nominal group; and the unmarked way of getting a nominal group into thematic position in English, given that the clause is declarative, is to map it onto the Subject. Hence the preferred form of expression will be that with Subject *I.* (Halliday 1998: 4-5)

Similar considerations lead to equivalent structures in several other languages (e.g., French, Russian, and Chinese) which share the information structure preferences of English: "In all these

languages it is the person rather than the body part which is typically selected as Theme in expressions of pain" (1998: 5).

Grammatical metaphor is thus responsible for meaning attributed to a word in virtue of the grammatical category to which it belongs (Verb, Noun, etc.) and the role this plays in the clause (Subject, Complement, etc.) rather than because of its lexical content. In the above example, 'pain' is conceptualised as a possession in virtue of its lexicalisation as a noun that fulfils the grammatical role of complement of the verb 'to have'. On this view, processes are typically realised by the verbal group of the clause, participants by nominal groups, and circumstances by adverbial or prepositional phrases (for a concise summary of Halliday's process types and their linguistic reflexes, see Lascaratou 2007: 37-44).

3. Design and methodology of the corpus study

To explore the potential of a corpus-based methodology to contribute to the analysis of the meaning of emotion terms in context, five Greek emotion terms were selected for analysis. Three of these terms (*aghonia*, 'anticipation, anguish', *erotas*, 'passion, romantic love', and *siginisi,* 'yearning, being touched or moved')[1] were not among the basic set of 24 prototypical emotion terms of the GRID. They were selected because they appear to be specific to Greek, having no precise translational equivalent in English, and were expected to be important and meaningful to

[1] The original Greek spellings of these words are *αγωνία, έρωτας, συγκίνηση, λύπη* and *στενοχώρια*. Their Roman transliterations are used throughout this text for the convenience of the readers and do not constitute accurate representations of pronunciation. The canonical pronunciation of these 5 terms in IPA notation is /aɣoˈnia/, /ˈɛrotas/, /siˈɟinisi/, /ˈlipi/, and /stɛnoˈxorʝa/, respectively.

Greek respondents based on preliminary frequency counts.[2] The final pair (*lipi,* 'sorrow', and *stenokhorja*, 'chagrin') were proposed as translation equivalents of the same English term, 'sadness' by our bilingual translators. We thus hoped that a corpus-based methodology might help us to shed some more light onto the meaning of these emotion terms for Greek speakers.

These five emotion terms were investigated using both the GRID questionnaire and a corpus-based methodology. For the needs of the GRID, data collection was carried out on a sample of 40 Greek university students (see results in chapter X). In what follows, we draw on the GRID results for information about the component scores of the five terms on the four dimensions, about term scores on selected individual items from the questionnaire, and about correlations among feature ratings between terms, in order to identify maximally similar and dissimilar terms from the basic set.

For the purposes of the corpus study, occurrences of the five terms were located in the Hellenic National Corpus, a 47 million word written corpus drawn mainly from journalistic sources (HNC, available from hnc.ilsp.gr; Hatzigeorgiu et al. 2000). To ensure comparability with the results of the GRID, which used the citation form of words (i.e., nominative singular in Greek), only nominative singular and accusative singular occurrences of these terms were considered, nominative and accusative being morphologically indistinguishable for feminine nouns in Greek (i.e. four of our terms: *aghonia, siginisi, lipi,* and *stenokhorja*). For the fifth term, *erotas*, which is masculine and so has distinct nominative and accusative forms, we ran separate searches for nominative (*erotas*) and accusative (*erota*) occurrences and included all of these in the analysis.

---

[2] Notice that in the corpus all three are more frequent (N=1,986, 1,593, and 707 respectively) than either of the pair of terms (N=585 and 156) that correspond to English 'sadness,' a term that *was* among the 24 terms of the original GRID set.

A number of occurrences thus obtained were unsuitable for our purposes and excluded from further analysis. Specifically, out of the total number of occurrences of the five terms in the corpus we excluded appearances in book or movie titles, or when the terms were the object of lexicographical definition, as these did not represent spontaneous usage by the author. In addition, cases in which the exact same clause was repeated in the corpus, and cases in which *erotas*, "romantic love," referred to the act of lovemaking (as in *kano erota*, lit. 'make love') rather than the emotion were also excluded. Our final sample thus consisted of 1,986 tokens of *aghonia*, 1,593 of *erotas* (including accusative *erota*), 707 of *siginisi*, 585 of *lipi*, and 156 of *stenokhorja* (including the alternative spelling *stenakhorja*). Manual annotation and statistical analysis of this sample using SPSS and R was carried out by one annotator in consultation with the other authors, with whom problematic cases were also discussed.

As was mentioned earlier, in Halliday's framework of systemic functional grammar, lexicalisation by verb is generally thought to indicate conceptualisation as a process, while Noun Phrases (NP) are thought to lexicalise the roles of participants in a process, and Adverbial Phrases (AP) or Prepositional Phrases (PP) optional circumstances surrounding the process. Since the current analysis — in order to yield results comparable with those of the GRID — focused only on lexicalisation of emotions by means of nouns,[3] the relevant grammatical functions out of this list are those that can be fulfilled by nouns, i.e. verbal subject, verbal complement (traditional 'direct object'), and complement of a preposition (as part of a PP). Moreover, it is well known from the cognitive linguistic literature on transitive clauses as a type of construction (Taylor 2003) that the prototypical transitive clause involves a human agent

[3] Lexicalisation by noun is of course not the only possible means by which emotional experience may be lexicalised (for some other possibilities, see, for instance, Lascaratou's 2007 and Terkourafi and Bali's 2007 analyses of various expressions lexicalising 'pain' in Greek).

volitionally acting on a patient whose free will and possibility for autonomous action is correspondingly curtailed, making it appear more object-like as a result. Building on this literature as well as in line with previous work on emotion terms in Greek (Lascaratou 2007, Terkourafi and Bali 2007), we took nouns functioning as verbal subjects to indicate the agent of a process (ex. 1), while we considered nouns functioning as verbal complements to be potential objects of possession (ex. 2), and nouns functioning as part of a PP to indicate optional circumstances associated with the process (quality, location, etc.; ex. 3).

(1) *Ki oso pernun i meres o **erotas** mu olo fundoni* ("And as the days pass, my **passion** grows stronger.")

(2) *Dhen ekho **aghonia** ghia to melon* ("I have no **anxiety** about the future.")

(3) *Tileorasis ki efimeridhes estiazun se thimata apo ti Sumatra os ti Stokholmi ke olos o kozmos skivi **me siginisi** pano apo tin kini tu simfora* ("Television and newspaper reports focus on victims from Sumatra to Stockholm and the entire world bows **with sympathy** over its shared tragic fate.")

In addition, potentially significant dimensions of semantic variation were noted. The relevant indications were provided by the semantics of the verb or by means of (adjectival or adverbial) modification (the relevant clues are indicated in bold in the examples below). Specifically, we coded for intensity (ex. 4), duration (ex. 5), repetition (ex. 6), and potential metaphorical construal of the emotion in physical terms (e.g. as a liquid; ex. 7),[4] as we expected

[4] In coding for metaphorical construal of the emotions, we follow Conceptual Metaphor theory (e.g., Lakoff and Johnson 1980, Lakoff and Kövecses 1987), according to which metaphor is a necessary mode of thought that

all of these to provide additional insights into the type of eventuality or entity the emotion was conceptualised as.

(4) (Intensity): *I **aghonia** ton anthropon ikhe **apokorifothi*** ("**Anticipation** was at its **highest point.**")

(5) (Duration): *Omos i **stenokhorja** tu **kratise** ligho* ("But his **chagrin** didn't **last** long.")

(6) (Repetition): *I **lipi** pu **enalasete** me ti khara* ("**Sorrow** that keeps **alternating** with joy.")

(7) (Physical construal/liquid): *I **siginisi** ekhi **plimirisi** tis kardhies olon ton theaton* ("**Yearning** has **flooded** the hearts of all the spectators.")

Finally, we assessed each clause for the kind of affect expressed, distinguishing between 'very positive' 'positive', 'negative', 'very negative', or 'undecided' (ex. 8-13). By 'affect' we mean the attitude or emotional stance that the speaker adopts or conveys through his/her utterance. At first sight, affect is the only of our corpus annotation categories that maps relatively straightforwardly onto one of the GRID dimensions, namely 'valence' — even here, however, the mapping should not be assumed *a priori* but rather subject to confirmation by the empirical data, for it is possible that whether a term is assessed in context or in isolation could drastically alter the affect felt to be expressed by it. To assess affect in a replicable manner, we referred to semantic information expressed by the surrounding linguistic context *excluding* the emotion term itself. Potential sources included other nouns conjoined with or dependent on the emotion term (ex. 8, 9, 10, 11), semantic contrast (ex. 12), as well as prepositional phrases, adjectives, adverbs,

---

involves the mapping of abstract concepts on to concrete ones and enables us to reason about the latter in virtue of this mapping.

and verbs present in the clause. When the surrounding linguistic context provided conflicting indications (ex. 12) or did not provide any indications as to affect (ex. 13), this was classified as 'undecided'.

(8) (Very positive): *Etsi kharaktirizi o idhios tin tenia tu epimenondas omos ke* ***stin aghapi, ton erota kai ti sindrofikotita*** *pu ta enoni.* ("This is the way that he himself characterises his movie, insisting also on **the love, the passion and the companionship** that binds it all together.")

(9) (Positive): *Thimame* ***ti siginisi ke ti khara*** *pu eniosa otan kitaksa ti fotografia aftu tu pedhiu* ("I remember **the yearning and the joy** that I felt when I looked at the photograph of this child.")

(10) (Negative): *Dhen prepi na dhiksi* ***fovo i stenokhorja*** ("He must not show **fear or chagrin**.")

(11) (Very negative): *Pote dhen idha toso polus skinothetes ke tosi* ***aghonia. Apoghnosi*** *skhedhon.* ("I never saw so many (film) directors and so much **anticipation.** Almost to the point of **despair**.")

(12) (Undecided/conflicting cues): ***Aghapes, entiposis, apoghoitefsis kai nostalghies*** *[…] ksanazondanevun […] me ena lirizmo pu bori na feri tin piitiki* ***siginisi*** *os ta dhakria* ("**Passions, impressions, disappointments and yearnings** come to life with a lyricism that can bring poetic **yearning** to the point of tears.")

(13) (Undecided/lack of clues): *Simera mia olokliri khora perimeni me* ***aghonia*** *tin etimighoria ghia enan mono anthropo* ("Today an entire country is waiting with **anticipation** for the verdict for a single man.")

In sum, the following categories were used to annotate occurrences of the five terms in the sample of sentences selected from the corpus: (a) grammatical function (=subject/ object/ prepositional phrase), (b) significant dimensions of semantic variation (=intensity/ duration/ repetition/ physical construal), and (c) affect (=very positive/ positive/ negative/ very negative/ undecided). A record of all the relevant clues used during the annotation process was kept for the purposes of verification and future analysis. In the next section, we present the results of the corpus analysis and compare them with those of the GRID.

4. Results

4.1 Aghonia (=anguish, anxiety, anticipation, suspense)

We begin by discussing the results for the most frequent of the five terms, *aghonia*. In English, *aghonia* can correspond to 'anguish, anxiety' as well as more positively-loaded 'anticipation, suspense.' In total, 1,986 occurrences of *aghonia* in the corpus were included in the final analysis (for a summary of the results, see Table A.1 in the Appendix). Some examples are provided below.

(14) *I **aghonia** ke o fovos mu meghalonan* ("My **anguish** and my fear grew bigger.")

(15) *S'ola ta pedhia aresi i dhrasi ke i **aghonia** ke to kalo istoriko mithistorima ekhi bolika ke apo ta dhio*. ("All children love action and **suspense**, and a good historical novel has lots of both.")

(16) *Metakinume dhistaktika, ja na min po me **aghonia*** ("I move with hesitation, not to say with **apprehension.**")

In terms of grammatical function within the clause, occurrences of *aghonia* were evenly split between subject (32.2%, ex. 4, 14) and complement (32.4%, ex. 2) positions ($\chi^2(1) = 0.01$, $p = 0.911$), and although there was a slight trend for it to occur as part of a PP (35.4%, ex. 16), this trend was not statistically significant ($\chi^2(1) = 2.67$, $p = 0.102$). This means that if we want to locate the core meaning of *aghonia*, we must look elsewhere.

Some useful indications in this regard are provided by its co-occurrence with lexical items indicating long duration (e.g., verbs *perimeno/anameno*, 'to wait'; 28.2%) and high intensity (e.g., *meghali*, 'great'; 25.3%) which were relatively more frequent than lexical items indicating physical construal of the emotion (e.g., physical appearance: *emfanis/faneri*, 'visible/obvious'; 7.6%) and repetition (e.g., *pali* 'again'; 4.1%). These differences were statistically significant (e.g., proportion of intensity vs. proportion of physical construal, $\chi^2(1) = 1841.91$, $p < 0.0001$), suggesting that duration and intensity are significant dimensions of semantic variation for the conceptualisation of *aghonia*. The importance of intensity to the conceptualisation of *aghonia* is also supported by its co-occurrence with lexical items indicating verticality (*katakorifo,* 'vertical, sheer', *korifonete*, 'soars'; 2.6%, ex. 4) suggesting a metaphorical mapping akin to that of heat, which exploits a combination of two metaphors, INTENSITY IS QUANTITY and MORE IS UP (Lakoff 1987). Finally, in terms of affect, occurrences of

aghonia tend to be negatively (52.2%) rather than positively coloured (13.4%) or undecided (34.4%): these proportions are significantly unevenly distributed ($\chi^2(2) = 448.91$, $p < 0.0001$) and the proportion of negative occurrences is marginally significantly greater than the positive and undecided proportions combined ($\chi^2(1) = 3.72$, $p = 0.054$), whereas the proportion of positive occurrences is significantly less than the negative and undecided proportions combined ($\chi^2(1) = 1064.51$, $p < 0.0001$).

Taken jointly, these results suggest that the core meaning of *aghonia* is to indicate a state of excited expectation, with the possibility of negative affect foregrounded in the context of other terms occurring in the same clause (e.g., *apognhosi,* 'despair', ex. 11, *fovos*, 'fear', ex. 14). These findings are consistent with the Greek GRID results. Specifically, in the GRID ratings for the 144 emotion features, *aghonia* correlated highly positively with anxiety (*agkhos*, 0.91), stress (*stres*, 0.89), and fear (*fovos*, 0.86). Moreover, *aghonia* was second only to *stres* on the high end of the Arousal dimension (3rd factor), with a factor score of 1.59. The heightened degree of arousal conveyed by *aghonia* is also consistent with its low absolute scores on the other two dimensions, 0.01 on Power and 0.28 on Unpredictability, while the potential for negative affect is consistent with its somewhat negative score (−0.38) on the Evaluation dimension.

4.2 Erotas (=passion, romantic love)

Next we discuss results for *erotas*, another Greek-specific term that may be rendered in English as 'romantic love or passion.' In total, 1,593 occurrences of *erotas* in the corpus were analysed (for a summary of the results, see Table A.2 in the Appendix). In terms of affect, *erotas* occurred predominantly in positively-loaded contexts (54.7%) as opposed to negatively-loaded ones

(10.3%), while undecided occurrences of *erotas* were also non negligible (35%). The difference between positive and non-positive contexts (negative and undecided combined) was statistically significant, $\chi^2(1) = 13.94$, $p = 0.0002$, allowing us to infer that *erotas* is a predominantly positive emotion. This is in accordance with the Greek GRID results, where *erotas* correlated positively with love (*aghapi*, 0.88), pleasure (*apolafsi*, 0.77), and joy (*khara* 0.76), and scored positively very high on Evaluation (1.33).

However, a closer look at the corpus data suggests that this is not the end of the story for *erotas*. The most frequent grammatical function of *erotas* is as a verbal complement (40.8%), a proportion that is statistically significantly greater than the proportion of its occurrences as a verbal subject (28.5%), $\chi^2(1) = 34.80$, $p < 0.0001$, or as part of a PP (27.4%).[5] However, occurrences of *erotas* in these three syntactic positions are not equally distributed with respect to affect (see Table 1).

Insert Table 1 around here

As table 1 shows, the proportion of negative affect is significantly higher in subjects than in complements, $\chi^2(1) = 4.90$, $p = 0.027$, while the proportion of positive affect is significantly higher in complements than in subjects, $\chi^2(1) = 5.06$, $p = 0.025$. Examples 17 and 18 illustrate these two possibilities.

[5] In the remaining 3.3% of occurrences, *erotas* was not integrated into the syntactic frame of the sentence (e.g., enumerations).

(17) *Dhioti o erotas tifloni ke ine kakos odhighos*. ("Because **romantic love** blinds and is a treacherous guide.")

(18) *Latrepses ti zoi, tin aghapi, ton erota*. ("You worshipped life, love, **romantic love.**")

The difference in the affect expressed by *erotas* in the two positions is statistically significant and may be interpreted as indicative of two opposing conceptualisations of romantic love in Greek: when conceptualised as an object (ex. 18), *erotas* is positively loaded, leading to a view of romantic love as a prized possession, one that can enrich one's life; conversely, when conceptualised as an agent, *erotas* is negatively loaded, with examples such as 17 leading to an image of romantic love as a perpetrator that torments those afflicted by it.

This second facet of *erotas* is one that cannot be detected from the GRID results in which *erotas* scored positively very high on Evaluation (1.33). How can we reconcile these two findings? One possible suggestion is that the predominance of occurrences of *erotas* in positively-loaded contexts, as seen in the corpus data, may make this the default context for the term, biasing interpretation toward the first of the two conceptualisations presented above. In other words, it is possible that, when presented with the term in isolation, questionnaire respondents implicitly call up a positively-loaded context in which they assess the term. Nevertheless, this contextual association does not capture the totality of the semantic potential of *erotas*—for this, a contextualised approach is necessary. Indeed, focusing on the positive conceptualisation alone could lead to the simplifying assumption that romantic love is always positively conceptualised in Greek, contrary to what we see in the corpus and leaving us with little to say about its negatively loaded occurrences.

An analysis of significant dimensions of semantic variation completes this picture. Here, *erotas* is most frequently characterised by high intensity (17.6%), long duration (14.5%), and physical realisation of the emotion (11.5%) but only rarely by repetition (2.8%) — a finding which is statistically significant, with the proportion of repetition being significantly less than the next semantic dimension, the proportion of physical construal, $\chi^2(1) = 90.33$, $p < 0.0001$. Moreover, frequent references to the lifecycle of romantic love with verbs such as *ksekino*, 'start', *arkhizo*, 'begin', *ghenieme,* 'be born', and *anthizo*, 'blossom', suggest that romantic love in Greek is characterised by an upbeat feeling and a sense of vitality.

These findings provide further support for those of the GRID. Specifically, the intensity and vitality seen in the corpus analysis are consistent with the very high score of *erotas* on Arousal (1.44), while the gradually evolving durative aspect of the term is consistent with its very low score on Unpredictability (−1.07). Its intermediate position on the negative side of the Power dimension (−0.35), on the other hand, suggests that power is not a defining feature of this term. Finally, it is worth mentioning that *erotas* is immediately adjacent or next-to-adjacent in scores to *aghapi* ('love') on three of the GRID dimensions (*aghapi* was located at 1.39 on Evaluation, −0.27 on Power, and −1.63 on Unpredictability), but further towards the high end on Arousal (cf. *aghapi* at 0.32), indicating both the similarity between the two terms as well as the crucial distinction between them — a distinction whose significance becomes fully apparent only once the corpus data are taken into account. According to these, the higher degree of Arousal that characterises *erotas* may be attributed to the agentivity associated with negatively-loaded occurrences of *erotas* in subject position.

4.3 Siginisi (=sympathy, yearning, being touched or moved by sth)

The third term we analyzed was *siginisi*, a Greek-specific term with no obvious English translation; depending on context, *siginisi* may be rendered in English as 'sympathy', 'yearning,' 'being touched,' or 'deeply moved' (ex. 3, 7, 9, 12). The lack of an obvious translation for *siginisi* is paralleled by an absence of any immediately recognizable distinct emotional content conveyed by it: in the GRID ratings, *siginisi* did not correlate significantly with any other emotion term; its highest correlation, with compassion (*simbonja*), reached barely 0.55.

In total, 707 occurrences of *siginisi* in the corpus were analysed (for a summary of the results, see Table A.3 in the Appendix). In these, *siginisi* tended to occur in positively affective clauses (57.9%) as opposed to either negatively affective (23.3%) or undecided (18.8%) ones, a difference that was statistically significant, $\chi^2(1) = 17.43$, $p < 0.0001$. This indicates that *siginisi* is generally a positively loaded emotion and agrees with its moderately high score (0.75) on the positive side of the Evaluation dimension in the GRID results.

In terms of grammatical function, in the corpus *siginisi* tended to occur primarily as part of a PP (37.6%, ex. 3) or as a verbal complement (36.4%, ex. 9), and less frequently in subject position (25.6%, ex. 7). The proportion of occurrences of *siginisi* in subject position is significantly less than the proportion of its occurrences as a complement, $\chi^2(1) = 13.19$, $p = 0.0003$, suggesting that *siginisi* tends not to be conceptualised as an agent that acts wilfully on those that experience it.

What *siginisi* does tend to be conceptualised as becomes clearer if we look at the semantic dimensions along which it tends to vary. Here, *siginisi* is mostly characterised by high intensity (35.4%) and long duration (16%), and only rarely by repetition (2.8%; significantly less than the duration proportion, $\chi^2(1) = 70.25$, $p < 0.0001$). Moreover, when *siginisi* is construed in

physical terms (7.8%, ex. 7), this is almost always as a liquid (51 occurrences, or 7.3%), as can be inferred from its co-occurrence with verbs such as *plimirizo*, 'overflow', *ghemizo*, 'fill', and *pnigho*, 'drown,' nouns such as *khimari*, 'torrents,' and adjectives such as *vathia*, 'deep' (which can indicate a liquid by metonymy, as in 'deep sea') and *dhiakhiti,* 'pervasive, diffuse' (from Greek *dhia+kheō* = 'pour through'). In this, *siginisi* parallels several other emotions, which are similarly metaphorically conceptualised as contained within the human body.[6]

The metaphorical construal of *siginisi* as a liquid that can overflow its container (i.e. the human body and, by metonymy, its human bearer) is consistent with its very low score on the Power dimension (−1.03) in the GRID results, suggesting that *siginisi* is an emotion over which Greek speakers feel they have little or no control. However, this result would seem to be at odds with the finding cited just above, based on its grammatical function within the clause, that *siginisi* does not tend to be conceptualised as an agent that acts wilfully on those that experience it. Nevertheless, these two results are not necessarily irreconcilable. The key lies in the conceptualisation of *siginisi* as a liquid: once we have established that *siginisi* is metaphorically conceptualised in this way, we are in a position to see that *siginisi* is not so much conceptualised as an agent that acts out of free will but rather as a force of nature that simply cannot be resisted. This interpretation gains support from the moderately positive score of *siginisi* on the Unpredictability dimension (0.49) in the GRID results, which underscores the relative unexpectedness of this emotion, as well as by its frequent modification on the intensity dimension in the corpus data: as with all natural phenomena, man has no control over their timing or intensity. In summary, *siginisi* offers a prime example of an emotion term where the

---

[6] On the metaphor THE BODY IS A CONTAINER FOR THE EMOTIONS, see Kövecses 1990.

corpus and the GRID results combine to yield a fuller picture than what either of them can capture in isolation.

4.4 Lipi and stenokhorja (=sorrow and chagrin)

The final pair of terms analysed are *lipi*, 'sorrow,' and *stenokhorja*, 'chagrin'. These were selected not because they are in any way specific to Greek but because of their semantic proximity to English 'sadness', one of the GRID's 24 prototypical emotion terms, of which both are translation equivalents in Greek. The results of the Greek GRID confirmed this intuition, with *stenokhorja* being highly correlated with *lipi* (0.94) and both being highly correlated with 'sadness' from the English sample (0.92 for *stenokhorja* and 0.91 for *lipi*). The relative positions of the two terms along the four basic dimensions also suggest a near identity, as they are adjacent or next-to-adjacent in Power (both very low: −1.23 for *stenokhorja* vs. −1.37 for *lipi*), Arousal (both somewhat low: −0.37 vs. −0.57, respectively), and Unpredictability (both in the middle region: 0.02 vs. −0.22, respectively). There is a minor difference in Evaluation only, where they both occupy the middle region but are separated by five intervening terms along this dimension, with *stenokhorja* (−0.47) being slightly more negative than *lipi* (−0.18)

A comparison of the feature profiles of the two terms using MANOVA over the 144 features by the distinct groups of raters showed them to be statistically indistinguishable: Wilks' lambda = 0.019, $F(1, 80) = 0.641$, $p = 0.785$. In simple between-subjects contrasts, only two of the individual features were significantly different for the two terms after Bonferroni correction ($\alpha = 0.000035$), while for 116 features there was no trace of a difference ($p > 0.1$). The two features distinguishing *lipi* from *stenokhorja* were q026 "incongruent with own standards and

ideals," with *stenokhorja* scoring higher than *lipi*, and q028 "centre of attention," with *lipi* scoring higher than *stenokhorja*. Both of these significant differences were in Appraisal items. The two terms were also identical in the other terms with which they correlated: high positive correlations for *lipi* included being hurt (*na plighonese* 0.90), disappointment (*apoghoitefsi* 0.88), and despair (*apoghnosi* 0.86), with the corresponding coefficients for *stenokhorja* being 0.94, 0.93, and 0.93, respectively.

In total, 585 occurrences of *lipi*, and 156 of *stenokhorja* (or *stenakhorja*) were included in the analysis (for a summary of the results, see Table A.4 in the Appendix). An analysis of the affect of the clauses in which the two terms appeared in the corpus data showed them to be nearly synonymous and overwhelmingly negatively loaded (86.8% for *lipi* vs. 85.9% for *stenokhorja*, significantly greater than the corresponding alternatives combined, $p < 0.0001$), with undecided occurrences ranking a distant second (11.5% vs. 13.5%), and positively loaded ones being extremely rare (1.7% vs. 0.6%).

Where differences between the two terms began to emerge, however, was in their grammatical function within the clause. Here, the preference for *lipi* in complement position was clear (65.6%); only secondarily (and statistically significantly less frequently, $\chi^2(1) = 114.77$, $p < 0.0001$) did *lipi* function as part of a PP (23.8%), and even less frequently as subject (10.4%). Occurrences of *stenokhorja,* on the other hand, were much more evenly distributed across the three grammatical functions (31.4% in subject position vs. 30.8% in complement position, and 38.5% as part of a PP, a distribution not statistically different from uniform, $\chi^2(2) = 1.69$, $p = 0.429$).

Zooming in on occurrences of *lipi* as a complement, we find that in a large majority of cases (77.6%) *lipi* collocates with the verb *ekfrazo,* 'express', a finding whose significance

becomes obvious when we consider that the next two verbs *esthanome* and *niotho* (both 'feel') taken jointly account for only 6.51% of occurrences of *lipi* in complement position. "Express sorrow" (*ekfrazo lipi*), in other words, constitutes a collocation or formula (Wray 2002) in Greek, frequently found in formal contexts as in (19):

(19) *I elvetiki kivernisi eksefrase ti vathia tis lipi ke ta silipitiria tis stus sigenis ton thimaton*
("The Swiss government **expressed** its deep **sorrow** and its condolences to the relatives of the victims.")

Closer examination of the corpus data reveals that the full form of the collocation is as in (20):

(20) *$(PRO)_i$ ekfrazi (ti) (vathia) $(POSS)_i$ lipi $(POSS)_i$ (ghia)*
$(PRO)_i$ expresses (DET) (deep) $(POSS)_i$ sorrow $(POSS)_i$ (for)

i.e., "X expresses X's deep sorrow for", where PRO stands for pronoun, DET for determiner (here, the definite article), POSS for possessive, the subscript index $_i$ shows referent identity (the person that expresses sorrow is the same one that possesses the sorrow) and parentheses indicate optional items.

The formula in (20) accounts for over three quarters of occurrences of *lipi,* 'sorrow', in the corpus data. In over three quarters of instances, that is, occurrences of *lipi* in the corpus are not creative but part of a set phrase. As part of a set phrase, *lipi* in these instances is largely bleached of its full-fledged semantic content (Lehman 1985), its use driven mainly by situational considerations and the exigencies of particular settings for the expression of certain socially

prescribed feelings. This hypothesis is supported by the patterns of semantic variation found with the two terms. Of the two, *stenokhorja* is consistently creatively modified more than *lipi*, sometimes two to three times as much, scoring 34.6% vs. 17.6% for intensity ($\chi^2(1) = 20.33$, $p < 0.0001$), 15.4% vs. 5.6% for duration ($\chi^2(1) = 15.12$, $p = 0.0001$), 10.3% vs. 3.6% for physical construal ($\chi^2(1) = 10.18$, $p = 0.0014$), and a non-significant 1.9% vs. 1.5% for repetition ($\chi^2(1) < 0.001$, $p = 0.985$). In all these ways, *stenokhorja*, 'chagrin', offers opportunities for foregrounding the subjectivity of the speaker, something which *lipi* seems ill-suited to achieve by virtue of its situational exigency-driven formulaic outlook.

To the apparent identity of the two terms in the GRID results, then, the corpus data add an interesting twist. On the one hand, consistent with a slightly more negative position on the Evaluation dimension of the GRID, the corpus analysis indicated that *stenokhorja* is the more creatively used, and so more subjectively loaded, term of the two. On the other hand, and beyond formulaic usage, corpus analysis revealed that the two words, albeit both being potential translation equivalents of English 'sadness', are not absolutely synonymous. They are separated by differences in register (formal vs. informal) and in semantic strength (degree of foregrounding of the speaker's subjectivity). In fact, these two types of differences are not independent but mutually reinforce each other: decreased subjectivity and emotional detachment are exactly what one might expect to find in formal settings, while the opposite might be expected to be true in informal, more personable circumstances. Since translation equivalence is not only a matter of lexical semantics but also of stylistic and situational appropriateness, a corpus-based methodology seems particularly well-suited in this case to help select the most appropriate translation equivalent each time.

5. Conclusions

We outlined a corpus-based methodology that builds on Halliday's framework of systemic functional grammar, in particular his notion of grammatical metaphor, to analyse the semantics of emotion terms, and compared our findings to those of the GRID. We investigated two different sets of terms. First, terms deemed to be specific to Greek socio-cultural experiences were investigated. Analysis of three such terms (*aghonia*, 'anguish', *erotas*, 'romantic love', and *siginisi*, 'being touched') confirmed the results of the GRID and added new dimensions to their meaning (e.g., the findings regarding two types of romantic love, and metaphorical construal of *siginisi* as a liquid). Second, terms constituting translation equivalents of a single English term were investigated. Here, two terms, *lipi* ('sorrow') and *stenokhorja* ('chagrin'), were compared. The results of the corpus analysis once again confirmed those of the GRID but this time they additionally revealed stylistic and semantic differences between the two terms that remained undetected by the GRID and could be crucial in determining the best translation equivalent of a term in context. In conclusion, we propose that the two methodologies are complementary and that using them *in tandem* can significantly enhance the analytical breadth and depth of future analyses of emotion terms.

References

Goodenough, W. (1956). Componential analysis and the study of meaning. *Language*, 32, 195–216.

Halliday, M.A.K. (1985/1994). *An Introduction to Functional Grammar*. London: Arnold.

Halliday, M.A.K. (1998). On the grammar of pain. *Functions of Language,* 5:1, 1–32.

Hatzigeorgiu, N., et al. (2000). Design and implementation of the online ILSP corpus. In: *Proceedings of the second international conference of language resources and evaluation (LREC), Athens, 31 May-2 June 2000*, 3, 1737–1740.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Katz, J. and Fodor, J. A. (1963). The structure of a semantic theory. *Language,* 39, 170–210.

Kövecses, Z. (1990). *Emotion Concepts*. Berlin/New York: Springer Verlag

Lakoff, G. (1987). *Women, Fire and Dangerous Things*. Chicago: Chicago University Press.

Lakoff, G. and Kövecses, Z. (1987). The cognitive model of anger inherent in American English. In D. Holland and N. Quinn (eds.) *Cultural Models in Language and Thought,* pp. 195–221. Chicago: University of Chicago Press.

Lascaratou, C. (2007). *The Language of Pain: Expression or Description?* Amsterdam: John Benjamins.

Lehman, C. (1985). Grammaticalisation: synchronic variation and diachronic change. *Lingua e Stile,* 20, 303–18.

Lounsbury, F. (1956). A semantic analysis of Pawnee kinship usage. *Language,* 32, 158–94.

Peeters, B. (ed.) (2000). *The Lexicon-Encyclopedia Interface*. Oxford: Elsevier Science.

Recanati, F. (2003). *Literal Meaning*. Cambridge: Cambridge University Press.

Taylor, J.R. (2003). *Linguistic Categorisation: Prototypes in Linguistic Theory.* Oxford University Press. Third edition.

Terkourafi, M. and Bali, P. (2007). Speaking of pain in Greek: Implications for the cognitive permeation of emotions. *Cognition and Emotion,* 21, 1745–79.

Wilson, D. and Carston, R. (2007). A unitary approach to lexical pragmatics: Relevance, inference and ad hoc concepts. In N. Burton-Roberts (ed.) *Advances in Pragmatics*, pp. 230–59. London: Palgrave.

Wilson, D. (2003). Relevance and lexical pragmatics. Italian *Journal of Linguistics/Rivista di Linguistica,* 15.2, 273–91.

Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Appendix: Summary of results of the corpus analysis*

Insert Table A.1 here

Insert Table A.2 here

Insert Table A.3 here

Insert Table A.4 here

* The sum of all percentages is not equal to 100% due to absence of adequate syntactic information in some of the clauses. This applies to all similar cases.

Table 1

Grammatical functions of "erotas" in the corpus by type of affect

	Affect		
	Negative	Undecided	Positive
Verbal subject	**12.00%**	36.00%	52.00%
Prepositional phrase	10.09%	34.63%	55.28%
Complement	7.71%	33.26%	**59.03%**

Table A.1

Classification percentages for "aghonia" in the corpus (N=1,986)

Grammatical function			
Subject	Complement	Prepositional phrase	
32.18%	32.38%	35.4%	
Semantic variation			
Intensity	Duration	Repetition	Metaphorical construal
25.33%	28.20%	4.13%	7.60%
Affect			
Negative	Positive	Undecided	
52.17%	13.39%	34.44%	

Table A.2

Classification percentages for "erotas/erota" in the corpus (N=1,593)

Grammatical function			
Subject	Complement	Prepositional phrase	
40.80%	28.50%	27.37%	
Semantic variation			
Intensity	Duration	Repetition	Metaphorical construal
17.64%	14.50%	2.76%	11.49%
Affect			
Negative	Positive	Undecided	
10.30%	54.68%	35.03%	

Table A.3

Classification percentages for "siginisi" in the corpus (N=707)

Grammatical function			
Verbal subject	Complement	Prepositional phrase	
25.60%	36.35%	37.62%	
Semantic variation			
Intensity	Duration	Repetition	Metaphorical construal
35.36%	15.98%	2.83%	7.78%
Affect			
Negative	Positive	Undecided	
23.34%	57.85%	18.81%	

Table A.4

Classification percentages for "lipi" (N=585) and "stenokhorja" (N=156) in the corpus

	Grammatical function		
	Verbal subject	Complement	Prepositional phrase
lipi	10.43%	65.64%	23.76%
stenokhorja	31.41%	30.77%	38.46%
	Semantic variation		
	Intensity	Duration	Repetition	Metaphorical construal
lipi	17.61%	5.64%	1.54%	3.59%
stenokhorja	34.62%	15.38%	1.92%	10.26%
	Affect		
	Negative	Positive	Undecided
lipi	86.84%	1.71%	11.45%
stenokhorja	85.90%	0.64%	13.46%