ExELL Explorations in English Language and Linguistics

UDC 81'373.612.2=111
Original scientific article

ANXIETY between mind and society: a corpus-driven cross-cultural study of conceptual metaphors

Henrik Nordmark[1] & Dylan Glynn[2]
[1]*Lund University*
[2]*Université Paris 8*

Abstract

This study examines the possibility of using corpus-driven quantitative techniques to describe emotion concepts. It examines the concept of ANXIETY in American English, British English, Japanese and Swedish. In Cognitive Linguistics, the description of emotion concepts, based on lexical semantics, is done with the analytical framework of the Idealised Cognitive Model and the Theory of Conceptual Metaphors. Despite the descriptive power of this approach, it does not produce falsifiable results and does not account for social variation. Multifactorial Usage-Feature Analysis takes the theory and analytical assumptions of this tradition and provides a means for empirically testing proposed conceptual structures as well as interpreting them relative to social-cultural variation. The case study focuses on four conceptual metaphors associated with the concept of anxiety and a range of causes of the emotion state. It examines the relationship between the different causes and the metaphors relative to the four cultures. Although the metaphors and the causes exist in all four cultures, the use of multivariate statistics in the form of correspondence analysis, factor analysis and multinomial logistic regression, produce distinctive profiles for the cultures in question. The use of the conceptual metaphors in the three languages shows that British and American are essentially identical. Although distinct, relative to Japanese, Swedish is similar to English. Japanese's profile is the most distinct of the three in its metaphoric structuring of the emotion concept.

Key words: conceptual metaphors; emotion concepts; ANXIETY; corpus linguistics; multivariate statistics; Multifactorial Usage-Feature Analysis (Behavioural Profile Approach).

1. Introduction

Cognitive Linguistics has long sought to identify, describe, and explain abstract culturally determined, yet perceptually based, concepts. Emotions, as embodied and experienced states that result from our socially determined world represent some of the most complex examples of such conceptual structures (Kövecses, 1986, 1991). Within the field, the methodology has thus far been largely restricted to introspective research. Although introspection can access subtle conceptual dis-

tinctions, it does not produce results that are readily falsifiable. Moreover, the nature of the description is 'idealised' in that it does not account for social variation, which is an inherent part of the usage-based model of language. The Idealised Cognitive Model (Lakoff 1987) represents a powerful analytical heuristic in cross-cultural and cross-linguistics analysis, but until it can be used to produce falsifiable results and to account for social variation, its role will remain purely theoretical.

This study draws on the corpus-driven Cognitive Linguistics tradition and adopts the multifactorial usage-feature methodology (also called behavioural profile approach) developed by Dirven *et al.* (1982), Rudzka-Ostyn (1989), Geeraerts *et al.* (1994), Gries (2003), Divjak (2010) and Glynn (forthc.). This methodology has been successfully applied to a range of morpho-syntactic, semantic-pragmatic and sociolinguistic questions (cf. Stefanowitsch & Gries, 2006; Glynn & Fischer, 2010, Glynn & Robinson, in press) but has not yet been applied to the metaphoric structuring of emotion concepts cross-culturally.

The concept of ANXIETY is a particularly poignant emotion in the contemporary world. It is a difficult emotion to pin down, a category which designates a response to socio-emotional stress. The stereotypes of the three cultures, English, Swedish and Japanese, could not be more different with regards to the concept: the loud competitive English and Americans, embracing the stress of industrial urban life, not so different from each other save for the sometimes understated demeanour of the English. This is contrasted by the calm, careful, yet confident Swede with a respectful diffidence to capitalist work practices and then again by the Japanese stereotype of a highly traditional and regulated society where norms, hierarchy and work performance dominate. The validity of such stereotypes does not interest us, but they do make for a good starting point for the cross-cultural analysis of the emotion concept ANXIETY. What is shared, and therefore, basic to this emotion concept and what is different between the cultures?

Moreover, the lexemes examined in this paper are the technical terms used to refer to the emotional state within the psychological literature. A purview of the literature reveals that there is little or no concern for how this concept might vary culturally or how this could influence clinical practice. In psychology, the emotional state of ANXIETY is understood as a physiological response to fear. Assuming such emotional states are universal does not entail that they are conceptualised in the same manner. The implications for assuming that the emotional response associated with a given stimulus is parallel to how it is conceptualised are serious. Although it is likely that a large degree of correlation can be observed between cultures and languages, obviously, divergence is also to be expected. Such divergences may be important both in clinical and social psychology.

2. Data and analysis

2.1. Lexemes

The study is based on the use of three comparable lexemes designating the concept ANXIETY in American and British English, Japanese and Swedish. These lexemes are *anxiety*, *fuan*, and *ångest* respectively.

Anxiety

The first lexeme, *anxiety*, can be glossed as "the state of feeling nervous or worried that something bad is going to happen" (Oxford Advanced Learners Dictionary 2005). More precise definitions in British and American English suggest there is little semantic variation between the two dialects.

Merriam Webster's Dictionary of American English

1. a. painful or apprehensive uneasiness, usually over an impending or anticipated ill
1. b. fearful concern or interest
2. an abnormal and overwhelming sense of apprehension and fear often marked by physiological signs (as sweating, tension, and increased pulse), by doubt concerning the reality and nature of the threat, and by self-doubt about one's capacity to cope with it)

Oxford English Dictionary

1. feeling of worry, nervousness, or unease about something with an uncertain outcome (count and mass noun)
2. Strong desire or concern to do something or for something to happen (restricted to the NP + Infinitive construction - *anxiety to please. These examples were omitted from the study.*

Despite the obvious similarity to *anger* and *angst*, its origins are the Latin *anxietas* via French *anxiété. The word has medical connotations and, as stressed in the preceding section, is used as a technical term in psychology. The word's use in English dates back to the early 16th century and it is not clear why the lexeme did not generalise in its usage in a way similar to stress has in the 20th century. Its use remains relatively specific in contemporary English.*

Fuan

The dictionary entries for *fuan* are brief but offer similar descriptions to those of *anxiety*. The dictionary *Meikyo Kokugo Jiten* gives the following definition: "A worry that something bad will happen, as well as that feeling." The entry is not expanded further than this though it should be noted that it is also found in compounds,

such as *fuanshougai (anxiety disorder)*, *fuanhossa (anxiety attack)* and *fuanshinkeishou (anxiety neurosis)*. The encyclopaedia *Sekai Dai-hyakka Jiten* restricts its description to the technical use in psychology, confirming that this is indeed the equivalent of the English lexeme.

The most important difference lies in the word class. *Fuan* belongs to a word class commonly referred to as adjectival nouns or nominal adjectives (Shibatani, 2005). As the name implies, adjectival nouns possess behavioural aspects of both adjectives and nouns. They can stand as pre-modifiers in noun phrases, playing a similar role to adjectives in European languages. Moreover, they are also felicitous in copula constructions, in a way that is similar to predicative adjectives. In Japanese, however, adjectives do not normally function as predicatives (Shibatani, 2005). Finally, some adjectival nouns can also function as a subject or an object in a sentence, acting as a head in a noun phrase. *Fuan* is an example of such an 'adjective' that can also function as subject or object in a clause.

For the study conducted in this paper, the instances where *fuan* is functioning as a pre-modifier in a noun phrase have been omitted since in those cases the word is clearly an adjective. The instances where *fuan* functions as predicative have, however, been retained. Physical as well as mental diseases in Japanese are typically profiled by predicatives, whether they are nouns or adjectival nouns, in contrast to English and Swedish where they typically function as objects.

(1) a. *kare* *wa* *fuanshougai* *da*.
he TOP mark anxiety disorder COP
'He has an anxiety disorder.'

b. *kare* *wa* *gan* *da*.
he TOP mark cancer COP
'He has cancer.'

Another point that warrants noting is that the lexeme is compositionally complex. The first component, *fu*, is comparable to the English *non-* or *un-*, and the second, *an*, signifies 'relax'. So, the word can also be used to express unease, or uneasiness, which both are common ways of translating *fuan* to English.

Ångest

In Swedish, two dictionaries give almost identical definitions. The *Norstedts svenska ordbok* defines the lexeme as a "strong negative feeling of being exposed to pressure or (unspecified) threat. Often lasting a long time and affects the whole life". The *Nationalencyklopedins Ordbok* offers the same definition but adds "often connected to palpitations, difficulties in breathing, feeling of dizziness, etc." *Ångest* is also similar to the Swedish word for regret, *ånger*, and etymologically they have been used in a similar way. Currently, however, the word is restricted to the description of the emotional state comparable to that described by *anxiety* in English. It can be found in compounds such as *ångeststörning (anxiety disorder)*,

ångestattack (anxiety attack) and *ångestneuros (anxiety neurosis)*.

2.2. Data. Online personal diaries

All of the data used in this study are extracted from online personal diaries or 'blogs'. The language and the subject matter of blogs are especially appropriate for this study. Not only is the language spontaneous and informal, the subject matter is that of personal daily concerns. Quasi-narrational in nature, since, unlike a traditional diary, readers can respond, the diaries chart the daily tribulations of everyday lives. The occurrences of the ANXIETY lexemes, therefore, constitute natural contextualised uses focused on the personal expression of emotional states. This is exactly the kind of description the use of the lexemes in psychology is designed to categorise. Data taken from a broader source would run the risk of containing a large number of technical uses, which, almost by definition, would be the same across the languages. Moreover, a broader data source would have introduced more sociolinguistic complexity. One of the most important issues when comparing cultures and languages is assuring the data used are comparable. By restricting the genre in this specific way, we can be fairly certain that issues such as age, register and so forth are not affecting the results.

In order to obtain a sample with the maximal number of occurrences per language/culture, a more or less equal number of examples were extracted for each. The English data were taken from the LiveJournal Corpus (Speelman & Glynn, 2005), 47 British and American 99 occurrences making 146 occurrences in total. The Swedish data were obtained by searching for the lexeme *ångest* using Google's blog search service. A total of 166 Swedish samples were extracted, each with a considerable amount of context. The Japanese data were gathered by running a search for *fuan* on the blogging and social network site Ameba. Altogether, 165 examples of *fuan* were extracted with context.

2.3 Method. Multifactorial usage-feature analysis

Usage feature analysis was developed by Dirven *et al.* (1982), Rudzka-Ostyn (1989) and Geeraerts (1990). However, the application of multivariate statistics to the results of the analysis is the step that gives the method its descriptive power. The resulting multifactorial usage-feature analysis (also called behavioural profile analysis) draws on established quantitative methods in sociolinguistics. Geeraerts *et al.* (1994, 1999) and Gries (1999, 2003) developed such categorical multivariate techniques and applied them to the results of feature-analysis. In recent years, the method has gained popularity. Heylen (2005), Gries (2006), Divjak (2006, 2010), Gries and Divjak (2006, 2009), Grondelaers *et al.* (2008, 2009) Glynn (2009, 2010a,

2010b, 2013, in press a), Janda & Solovyev (2009), Glynn & Fischer (2010), Krawczak & Glynn (2012, in press), Krawczak & Kokorniak (2012), and Glynn & Robinson (in press) amongst others are representative of this growing field of research. The application of this method to specifically emotion concepts is a relatively new endeavour and is represented by Glynn (2013, in press b), Krawczak (in press, submitted) and Glynn & Nordmark (submitted).

In simple terms, the multifactorial usage-feature method consists of the repeated analysis of a range of semantic, pragmatic, and social characteristics of speech events. A large sample of a given phenomenon, here emotion concept key words, are extracted from a corpus with their context. These occurrences are manually annotated for whatever usage features are hypothesised to be indicative of conceptual structure. The results of this analysis provide a behavioural profile of the linguistic form. Due to its complexity, this profile needs to be interpreted with the aid of multivariate statistics, which permits the identification usage-patterns across the data. If sufficient data are available, the statistics can also be used to determine the descriptive accuracy of the analysis by testing its predictive power.

2.4. Analysis. Metaphor and Cause of ANXIETY

The usage-feature analysis seeks to capture the conceptual structuration of ANXIETY through an analysis of source concepts in metaphoric profiling, but also through the analysis of a range of characteristics of the event that triggers the emotional state. This section lists and explains each of these analytical dimensions.

Source Concept

The metaphoric structuring of emotion concepts is well-documented, just as the lexical methods for identifying those metaphors (Kövecses 1986, 2000; Lakoff 1987). Since it is not possible to search a concept in observational data (searches being restricted to strings of letters or sounds), usage-based methods face inherent limitations in the description of metaphoric structures. Each instance was carefully checked for indications metaphoric structuring. The source of this evidence was the content verb of the utterance. For each utterance, this verb was identified and listed. In total, some 265 (out of 477) metaphoric uses were identified. Of these, 233 appeared to be indicative of a conceptual metaphor. The verbs that profiled the ANXIETY event metaphorically were then grouped into sets determined by semantic similarity. This step, of course, is highly subjective and suffers from the same vagaries as 'traditional' conceptual metaphor analysis. Determining what constitutes semantic similarity and therefore demarcating the 'source concept' can be a difficult task. After thorough consideration, most of the metaphoric instances

could be categorised into four relative distinct conceptual relations. The source concepts identified include: MANIPULABLE OBJECT; OBSTACLE; OPPONENT; and PAIN. Examples of the predicates that were used to identify conceptual metaphors include *get over*, *pass*, *go through* for AXIETY is an OBSTACLE, *bring*, *give*, *take for* ANXIETY is a MANIPULATABLE OBJECT, *attack*, *control*, *suppress*, *overcome* ANXIETY is an OPPONENT and *ease*, *dissipate*, *feel* for ANXIETY is PHYSICAL PAIN. The metaphors based on these source concepts are exemplified below:

MANIPULABLE OBJECT

(2) a. *Hetsäter och har ångest så snart skär jag ihjäl mig*.
'Binge eating and have anxiety so soon I'll cut myself to death.'

b. *大丈夫かいな??と一抹の不安を抱え、優雅にタクシーで病院に向かう私。*
'[Is he] alright?? [I thought and] held some anxiety while I gracefully headed towards the hospital in a taxi.'

OPPONENT

(3) *Känner hur ångest attackerar mig*.
'Feel how anxiety attacks me.'

PHYSICAL PAIN

(4) *... my anxiety won't dissipate at all and continues to haunt my dreams with cold sweats and worrisome undertones.*

OBSTACLE

(5) […]
今日中に帰れるん…??よぎる不安…まさか。兵庫県ごときの距離で日帰りが難しく.
'[…] Can I return today…?? Anxiety passing by…Surely it's hard to go on a day trip for such a long distance as Hyogo prefecture.'

Example (6) represents a typical literal example, grammatically profiled with a copula construction:

(6) *不安は沢山ありましたが*、[…].
'There was a lot of anxiety, but [...].'

Cause of ANXIETY

The Event Cause can be argued to be essential to the understanding of the ANXIETY experience. Different Causes are not only indicative of distinct cultural patterns but can be interpreted as an operationalisation of the emotion experience itself. We can assume that the emotion concept, designated by the various terms for ANXIETY,

is neither discrete nor uniform in nature. However, using observational data, it is difficult, if not impossible, to categorise the actual emotional experience of the patient of the event. The Cause can, therefore, be seen as a proxy or perhaps approximate index of kind of experience in question. Such an interpretation entails that different causes lead to different experiences. This would mean that, for example, a life threatening illness would lead to a different emotional experience of ANXIETY than some untoward comments by an employer. This assumption is not to be taken lightly. It is just as possible to argue that the fact that the lexeme *anxiety* has been used to profile an experience, regardless of the Cause, means that it is this particular profiling that is pertinent to the experiencer or perceiver. While overtly acknowledging that it is precisely the kind of experience profiled by the choice of the lexeme and not necessarily the Cause that may be indicative of the conceptual structure, this dimension was carefully annotated in the sample. It should be remembered that even if this analysis and annotation does not inform the study of conceptual structure, it does inform the socio-cultural dimension of the description.

Six broad Event Causes were identified in the data. These causes were 'body image'/'health', 'state of affairs'/'circumstances', 'activities', 'events', 'occupation-education', and 'emotion'/'relationships'. The first feature, 'body image and health' includes Causes such as eating, drinking, weight and general health. In example (7), below, the cause of the ANXIETY is the consumption of crisps and the weight gain and health problems relating to it.

Cause-Event: 'body-health'/'food-drink'

(7) *Hade det väldigt trevligt men åt lite chips och fick ångest och ville spy upp allt.*
'Had a very nice time but ate some crisps and got anxiety and wanted to throw everything up.'

The Cause 'state of affairs'/'circumstance' is a category of causes where no specific event or thing is identifiable as responsible for the experience. In (8a), the cause of the ANXIETY is the fact that the speaker is alone on the weekend. Examples (8b) and (8c) are also typical of this category. These Causes are in line with the kind of Cause Event with which the technical use of the word is associated.

Cause-Event: 'state of affairs'/'circumstance'

(8) a. *Att sitta hemma SJÄLV en lördag- eller fredagkväll ger mig sån jävla ångest* [...].
'To sit home ALONE on a Friday or Saturday night gives me such fucking anxiety […].'

b. *my anxiety gets borderline crippling on the one-block walk back to BART because I'm alone in Oakland in the dark.*

c. *また、週末で病院の救急外来の対応に若干不安がありますが* [...].

'And there is also some anxiety over how they treat emergency outpatients at the hospital during the weekend […].'

The Event Cause 'activity' is more specific than 'state of affairs'. As can be seen in (9a) and (9b), the ANXIETY is caused by a specific activity performed by the speaker, namely driving.

Cause-Event: 'activity'

(9) a. *Har lite ångest över att köra in i garaget så jag* [...].
'Have a bit of anxiety over driving into the garage so I […].'

b. *i just realized that more often than not, driving gives me anxiety.*

Closely related to 'activity' is 'event', the difference here being that instead of the speaker performing the activities, the anxiety was caused by what someone else did or will do.

(10) a. *Låg vaken och hade ångest över att vår dotter kommer trilla i vattnet i sommar.*
'Lay awake and had anxiety over the fact that our daughter is going to fall in the water this summer.'

'Occupation'/'education' were annotated in instances where ANXIETY resulted from school, work and related activities. Examples (11a) and (11b), below, refer to student marks whereas (11c) is ANXIETY over a new workplace.

Cause-Event: 'occupation-education'

(11) a. *... and so started my anxiety over grades* […].

b. *Ny arbetsplats, ny ångest.*
'New workplace, new anxiety.'

c. *すでに２年のクラス替えの不安でいっぱいなのは私だけでしょうか.*
'Am I the only one already filled with anxiety over the class change in the second year?'

Causes of ANXIETY related to feelings and personal relationships were annotated as 'emotion - relationships'. In (12a), the ANXIETY is caused by fear of being rejected by another person and in (12b), the marriage is causing the ANXIETY. Example (12c) sees the patient missing loved ones resulting in an experience of ANXIETY.

Cause-Event: 'emotion - relationships'

(12) a. *The anxiety I get before I arrive at his gets less and less each time I go. It's mainly because I get scared of being rejected by his Mum.*

b. *この結婚正しかったのか？と不安が*[…]
'Was this marriage really right [for me]? The anxiety is […]

c. *... ångest. Imorgon är det fem månader sedan jag lämnade Loviseholm. Tror inte någon kan förstå hur mycket jag saknar livet och människorna där*, [...].

'... anxiety. Tomorrow five months will have passed since I left Loviseholm. Don't think anyone can understand how much I miss life and the people there, […].'

Lastly, a category 'unspecified' was needed for instances where the Cause was not discernable, even with considerable context. It is important to note that this does not mean that it has no Cause. The category is used to identify those examples where unspecified fears or apprehensions are the Cause of the emotional state. Consider examples (13a) – (13c).

(13) a. *My anxiety level is up again. I feel like something is going to happen*.

b. *Fan ångesten är hög, magen gör ont, hjärtat slår dubbelt, andningen ökar, kroppen skakar*.
'Fuck[,] the anxiety is high, my stomach hurts, my heart beats twice as fast, my breathing intensifies, my body is shaking.'

c. *この抑圧と理不尽と苦悩と不安の数年間をそれぞれに経験し*,[...].
'All these years of suppression and outrage and pain and anxiety [we have] experienced respectively […].'

3. Results and interpretation. Cross-cultural metpahoric structuring

Firstly, let us consider the distribution of the metaphors across the four cultures. Table 1 presents the frequencies of the conceptual metaphors observed relative to each of the languages. In terms of simple frequency, the most noticeable result is how relatively infrequent metaphoric examples are in the Japanese data. Since the Japanese part of speech is not directly comparable to the Swedish and English, one should be careful not to interpret this as indicative of anything other than an epiphenomenal effect of the grammatical profiling. However, the very fact that the figurative uses of *fuan* are so infrequent deserves further research. Turning to British English, American English and Swedish, there are no noteworthy differences between the British and American data and the general trends across English and Swedish seem comparable. In both cases, MANIPULABLE OBJECT is the most common source domain followed by PAIN and OPPONENT. In both cases, OBSTACLE is extremely rare. However, MANIPULABLE OBJECT accounts for the overwhelming majority of examples in Swedish, where the English examples are a little more evenly spread across the three metaphors. The difference between the languages is statistically significant (df = 2, *p*-value = 7.078e-08).

Table 1. Frequency of conceptual metaphors relative to language.

	Language					
Metaphor	English	(UK)	(US)	Swedish	Japanese	Total
Literal	82	29	53	76	150	308
OBJECT	29	7	22	79	2	110
OBSTACLE	1	-	1	1	2	4
OPPONENT	14	3	11	5	7	26
PAIN	20	8	12	5	4	29
Total Metaphors	64	18	46	90	15	233

On its own, this result is not particularly informative. Although the difference is significant, all three metaphors occur in both languages and in all three cultures. That the Swedish conceptualise ANXIETY as a MANIPULABLE OBJECT more systematically than English speakers or that the source concepts of PAIN and OPPONENT are slightly more common in English could easily be due to a myriad of factors external to the actual conceptual structures typical of each culture/language. The advantage of multifactorial feature analysis is that we can consider these metaphoric occurrences relative to other dimensions of usage. To these ends, let us now consider the interaction of the metaphors and the Cause of the ANXIETY event.

3.1 Identifying cross-cultural conceptual structure of ANXIETY

In order to capture the interaction of these two different usage dimensions, the Metaphor and the Cause, both relative to the culture in question, we employ multiple correspondence analysis. Using a Chi-squared algorithm, the analysis calculates the relative associations of every possible combination across the Languages, Metaphors and Causes. It represents the relative degree of association or disassociation between features through their relative proximity on a two-dimensional plot. The features that appear close together on the plot are highly associated in use where those far apart are distinctly not associated in use. The size of the data points indicates the relative importance of that feature in explaining the overall behaviour of the data. For a more complete description of how to interpret the results of this technique, see Glynn (in press b). The method used is the 'adjusted' method, proposed by Greenacre (2007). Before we consider the results of the analysis, we need to determine how reliable those results are.

Table 2. Numerical results of the joint multiple correspondence analysis of Metaphor for and Cause of ANXIETY in American English, British English, Japanese and Swedish.

```
Principal inertias (variance):
 dim    value        %   cum%  screen plot
 1      0.124247  49.9   49.9  *************************
 2      0.078397  31.5   81.4  ****************
 3      0.011704   4.7   86.2  **
 4      0.003197   1.3   87.4  *
 ...
 Features:          name  quality  contrib. dim. 1  contrib. dim. 2
 1 |          ENGLISH UK |    610 |             1 |              87 |
 2 |          ENGLISH US |    794 |             0 |             215 |
 3 |            JAPANESE |    821 |           198 |              90 |
 4 |             SWEDISH |    856 |           226 |              46 |
 5 |             LITERAL |    893 |            75 |              16 |
 6 | Met. MANIP. OBJECT |    903 |           241 |               3 |
 7 |  Metaphor OBSTACLE |    217 |             1 |               4 |
 8 |  Metaphor OPPONENT |    813 |             0 |              75 |
 9 |       Metaphor PAIN |    708 |             1 |              88 |
10|     Cause Activity |    651 |            12 |              19 |
11| Cause Circumstance |    685 |            22 |              28 |
12|         Cause Event |    593 |            17 |               0 |
13|   Cause Food/Drink |    932 |           157 |              39 |
14|   Cause Occupation |    915 |            36 |              68 |
15| Cause Relationship |    468 |            10 |              26 |
16|       Cause Unspec. |    829 |             4 |             196 |
```

In Table 2, the principal inertias show that it is possible to accurately explain the structure of the data in two dimensions. Indeed, 81.4% of the dispersion of the data can be represented along two the axes visualised below in the biplot (Figure 2). Indeed, adding a third dimension would only improve the explanation by 4.7%. An explained inertia score of 81.4% represents an extremely stable representation of the structure of the data. The second half of Table 2 presents the quality of representation for each of features, or data points, as well as their contribution to structuring along the two dimensions visualised. Any quality score of 500 or more can be interpreted as being accurately represented. Features that score lower than this are too infrequent to make for interpretation or their behaviour is spread out across otherwise distinct associations making their representation in two-dimensions less reliable. Despite the stable results, it must be remembered that correspondence analysis is an exploratory technique and its generalisations are only valid for the sample examined. Below, we consider confirmatory statistics,

which give us probabilities about how representative these findings are more generally.

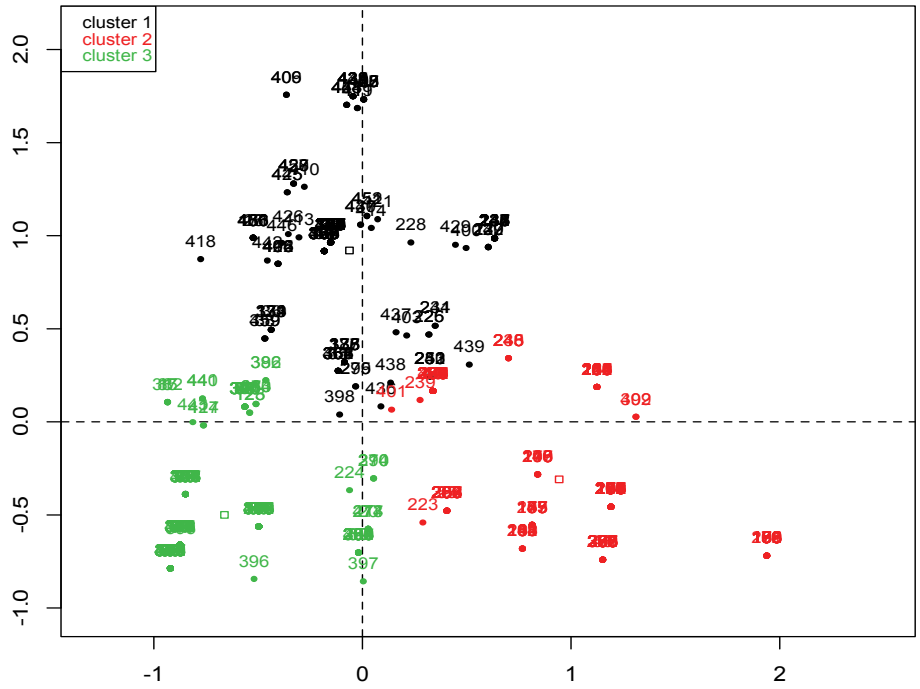

Figure 1. Factor map for the results of the multiple correspondence analysis.

Before we interpret the correspondence analysis, we submit the results of that analysis to what is termed a factor analysis. The factor analysis determines the number of underlying factors, or structures, in the data. Although the behaviour of the data can be visually represented in two-dimensions, it appears to be actually structured along three dimensions. This is not unusual, since the representation is determined by the ability to depict the associations in the data in a two-dimensional plot. This tells us nothing of the underlying structure of the data, which appears to cluster neatly into three sub-structures. On a two-dimensional plane, Figure 1 depicts the location of each of the individual examples and colours them to show which of the three underlying factors the example belongs to. If the clustering is clear, then there will be little overlap between the clusters. In Figure 1, the clustering is so clear it is almost discrete. These results suggest that, indeed, three quite distinct structures are present.

Returning to the examples, it is evident that this structuring correlates with the three languages. In figure 2, below, we see the distribution of the associations in the data. The language data points and the features with which they are associated map systematically onto the three structures identified in the factor analysis.

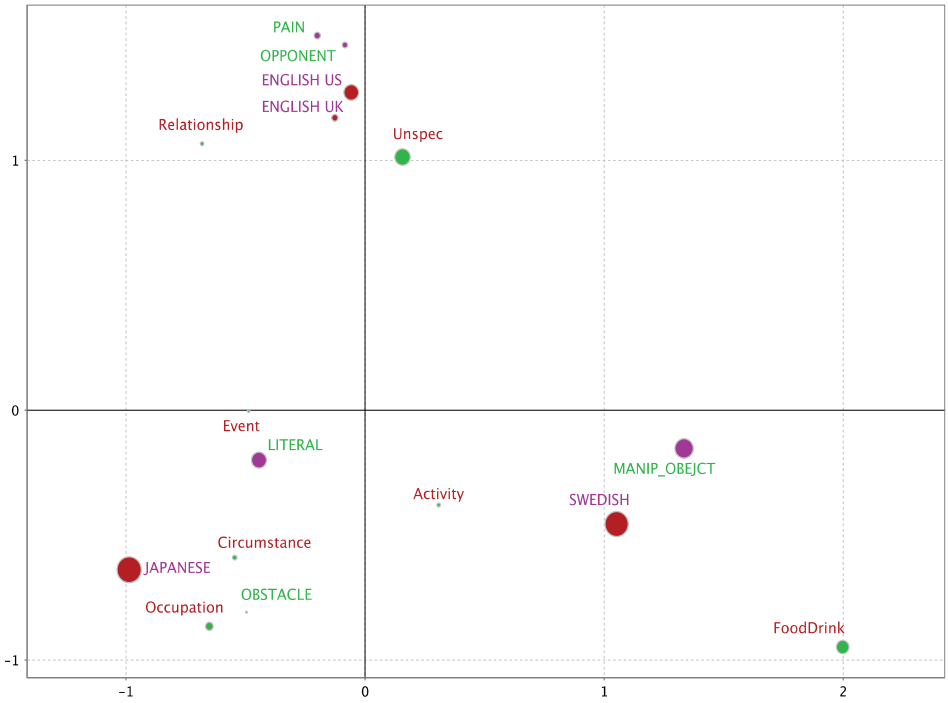

Figure 2. Metaphor for and Cause of ANXIETY in American English, British English, Japanese and Swedish (multiple correspondence analysis).

The cluster of features at the top encircles both the British and American data points, confirming their similarity relative to the usage-features in question. Two metaphors are found to be associated with the British and American English usage. Although we know that these metaphors do also occur with Swedish, in combination with the Causes that also cluster around the British and American data points, these metaphors are distinctly associated with the English language. The Causes in question are equally distinctly associated with the English usage patterns. From this, we have a clear profile of English – ANXIETY as a result of 'undetermined' abstract malaise or 'emotion' stress resulting from personal 'relationships'. This type of stress is conceptualised as an OPPONENT or as PAIN. Although there is always the risk of *post hoc* interpretation, this profile of ANXIETY does match intuitive expectations of the concept in these cultures. It also partially fits the technical use of the term, which underlines the idea of uneasiness resulting from non-specific causes or fears of future uncertainties.

Turning to the bottom right hand quadrant of the plot, we find the Swedish data point. We know from table 1 that MANIPULABLE OBJECT is the metaphor most highly associated with Swedish usage and we see this remains true relative to the

most typical Cause for ANXIETY in Swedish. This Cause is 'body-health'/'food-drink'. Indeed, returning to the data, we find that this cause is exclusively associated with the Swedish data, explaining its representation as highly and distinctly associated in the plot. Another Cause, 'activity', lies somewhat between the Japanese and the Swedish clusters, suggesting it is a point of overlap between the two cultures/languages. However, it appears slightly more typical of Swedish. Again, we have a clear and distinctive profile for Swedish ANXIETY, concern over 'self-image' and 'health', conceptualised as A MANIPULABLE OBJECT.

Lastly, we turn to the Japanese cluster in the bottom left quadrant. Although the metaphor of ANXIETY is an OBSTACLE was uncommon, when combined with the Causes 'occupation' and 'circumstance', it appears highly associated with Japanese usage. However, it does lie between the Japanese and Swedish data points and so its association is not distinct, associated to some extent also with Swedish usage and the Cause of 'activity' and/or perhaps 'food-drink'. It must be remembered that the metaphoric associations represented are not only the association with the language in question but the language in combination with specific Causes. It is the Causes of 'occupation' and 'circumstance' that are associated with Japanese and the metaphor of OBSTACLE offering an even clearer and distinct profile. 'Work' and 'study' as well as social 'states of affairs' are what typically cause ANXIETY, which is conceptualised as an OBSTACLE in the Japanese data set. Just as for the English data, this profile matches a received stereotype of the culture in question, where work is extremely important and strict social rules apply. Hypothetically, these results would corroborate this stereotype.

3.2 Confirming cross-cultural conceptual patterns

Despite the informative results provided by the correspondence analysis, the technique is designed for identifying patterns in complex data and it does not permit us to make generalisations beyond the sample. Multinomial logistic regression is not yet widely used in the field but it suits our needs here. It is an extension of binary logistic regression analysis, popular in Cognitive Linguistics and described in Speelman (in press). Logistic regression models the data in order to predict its behaviour. The rationale is simple: if one can accurately predict the behaviour of a given linguistic, social or conceptual phenomenon, then one knows that whatever was used to predict its behaviour is accurate and descriptively adequate. More specifically, it offers three pieces of information. First, it tells us which features are significantly associated with whatever phenomena one wishes to model. Second, it ranks these features in terms of relative importance in predicting the behaviour of the phenomenon. Third, it calculates how accurately a combination of these features predicts the use or occurrence of the phenomenon in question. Since we are interested in the conceptual profiles of ANXIETY in three languages/cultures, we

will attempt to predict which language a given example belongs to, based on a combination of the conceptual metaphor and the Cause of the event.

One of the clearest results is the Cause of 'health-body' – 'food-drink' which is uniquely associated with the Swedish data. These examples will be omitted from the regression analysis. Including them would greatly increase the accuracy of the models in predicting the language of the examples but would defeat confirmatory purpose of the analysis. The analysis is intended to determine to what extent the features, which are less clearly associated with a given language/culture, come together to predict and, therefore, determine the usage profile of that language/culture.

Multinomial logistic regression uses one of the outcomes, which for us is English, Japanese or Swedish, as a reference level against which to compute the prediction of the others. This means we will have three models, each with one of the languages as the base line for the others. Table 3 presents the results of the analyses.

Before we interpret the results, we need to report the diagnostics of the models. For an explanation of these diagnostics, see Agresti (2013) and Hosmer & Lemeshow (2013). The models were checked for multicollinearity, which was a concern, especially seeing associations observed between Cause and Metaphor in the correspondence analysis. However, none of the variance inflation factors were above 4.0 and the assumption of orthogonality is met. The models are parsimonious, the analysis of deviance, shown at the bottom of the table, indicates that both factors significantly increase the explanatory power of the models. However, one concern was data sparseness. There were 15 cells (20%) with zero frequencies. This is borderline acceptable. Despite this, the fit of the model is reasonable, indeed excellent given the size of the sample, and overdispersion does not appear to be more problematic than one would expect. Some observations were identified as being highly influential, but their removal did not have a substantial impact on the results and so they are retained for the models presented above. The examples of the Metaphor OBSTACLE and the Cause 'activity' were omitted from the analysis because they were too infrequent to submit to the model. Combined with the removed instances of the 'food-drink' Cause, mentioned above, this reduced the dataset from 477 observations to 441. Nevertheless, the chi-squared statistics for goodness-of-fit, which are both sensitive to our small sample size, suggest reasonable fit (neither the deviance nor Person's statistics should show significance). Although neither are extremely high, both are substantially higher than alpha-level of .05. Pseudo R^2 scores are less sensitive to sample size and are, therefore, better indices for the model. The R^2 scores are listed at the bottom of the table and all are much higher than the accepted rules of thumb for goodness-of-fit. The models and their statistics were produced in R, using nnet (Venables & Ripley 2002) and `mlogit` (Croissant 2012). The diagnostics were performed with a range of packages including `pscl` (Jackman 2012), `epicalc` (Chongsuvivatwong 2012), `perturb` (Hendrickx 2012) and `polytomous` (Arppe 2012).

Table 3. Coefficients for Multinomial Logistic Regression, predicting Language with Metaphor and Cause.

English Reference Level		Japanese Reference Level		Swedish Reference Level	
Coefficients:	Estimate	Coefficients:	Estimate	Coefficients:	
JAPANESE : OBJECT	-3.575573 ***	ENGLISH : OBJECT	3.575573 ***	ENGLISH : OBJECT	-1.018584 ***
SWEDISH : OBJECT	1.018584 ***	SWEDISH : OBJECT	4.594156 ***	JAPANESE : OBJECT	-4.594156 ***
JAPANESE : OPPONENT	-0.784565	ENGLISH : OPPONENT	0.784565	ENGLISH : OPPONENT	0.982367
SWEDISH : OPPONENT	-0.982367	SWEDISH : OPPONENT	-0.197802	JAPANESE : OPPONENT	0.197802
JAPANESE : PAIN	-2.613683 ***	ENGLISH : PAIN	2.613683 ***	ENGLISH : PAIN	1.159592 *
SWEDISH : PAIN	-1.159592 *	SWEDISH : PAIN	1.454092 *	JAPANESE : PAIN	-1.454092 *
JAPANESE : Cause-Circumstance	1.081796 *	ENGLISH : Cause-Circumstance	-1.081796 *	ENGLISH : Cause-Circumstance	0.082415
SWEDISH : Cause-Circumstance	-0.082415	SWEDISH : Cause-Circumstance	-1.164211 *	JAPANESE : Cause-Circumstance	1.164211 *
JAPANESE : Cause-Event	0.049982	ENGLISH : Cause-Event	-0.049982	ENGLISH : Cause-Event	0.893630 .
SWEDISH : Cause-Event	-0.893630 .	SWEDISH : Cause-Event	-0.943611 *	JAPANESE : Cause-Event	0.943611 .
JAPANESE : Cause-Occupation	1.136672 *	ENGLISH : Cause-Occupation	-1.136672 *	ENGLISH : Cause-Occupation	-0.966186 .
SWEDISH : Cause-Occupation	0.966186 *	SWEDISH : Cause-Occupation	-0.170486	JAPANESE : Cause-Occupation	0.170486
JAPANESE : Cause-Relationship	-1.578717 *	ENGLISH : Cause-Relationship	1.578717 *	ENGLISH : Cause-Relationship	1.828029 *
SWEDISH : Cause-Relationship	-1.828029 *	SWEDISH : Cause-Relationship	-0.249312	JAPANESE : Cause-Relationship	0.249312
JAPANESE : Cause-Unspec.	-2.066845 ***	ENGLISH : Cause-Unspec.	2.066845 ***	ENGLISH : Cause-Unspec.	1.015264 **
SWEDISH : Cause-Unspec	-1.015264 **	SWEDISH : Cause-Unspec.	1.051581 **	JAPANESE : Cause-Unspec.	-1.051581 **
Concordance statistic:	0.770760	Concordance statistic:	0.844717	Concordance statistic:	0.761168

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Log-Likelihood: -363.88	Goodness-of-Fit		Accuracy	Overall: 64.6	Analysis of Deviance Table
McFadden R^2: 0.247		Chi-Square df Sig		English: 59.3	LR Chisq df Sig
Cox and Snell: 0.417	Pearson	34.025 28 .200		Japanese: 81.6	Metaphor 136.10 6 < 2.2e-16 ***
Nagelkerke R^2: 0.470	Deviance	34.179 28 .195		Swedish: 49.6	Cause 163.96 12 < 2.2e-16 ***

Likelihood ratio test : chisq = 238.14 (p.value = < 2.22e-16)

To interpret the models, we need to consider both the statistical significance of the predictors and their estimate. Only predictors that are significant can be interpreted. They are indicated by asterisks; the relative alpha-levels explained beneath the table. The estimates are the numbers listed before the asterisks. These figures are expressed as log-odds and should only be interpreted relatively. They rank the importance in predicting the outcome of one language over the others. Negative numbers predict against and positive for.

The first thing to note in the data is that the metaphor OPPONENT is not significantly predictive of any of the languages. Although its association with English is distinctive in the correspondence analysis, the lack of significance is to be expected since there are relatively few occurrences and even if they are most frequent in the English data, they are relatively evenly dispersed across the three languages. Nevertheless, although not significant, in the first model, where English is the reference, both estimates are negative and in the other models, the non-significant estimates point towards English. Although we cannot interpret these results *per se*, they do corroborate what we observed above.

The other two metaphors are significant predictors for language across all three models. Relative to English, we see that PAIN is significantly and strongly disassociated with both Japanese and Swedish, confirming what we saw in the correspondence analysis. However, relative to Japanese, we see there is a significant and relatively strong degree of association between the use of the PAIN metaphor and Swedish. Of all the predictors, the OBJECT metaphor is the most important. We see an extremely strong association with Swedish in all models. Moreover, relative to Japanese, it is also strongly associated with English. In all cases, it is disassociated with Japanese. These results confirm what we see in the correspondence analysis for the metaphors.

Turning to the Cause features, it should firstly be mentioned that interactions between Cause and Metaphor were not found to be significant. It is possible this is merely a result of the relatively small sample. Overall, the findings in the exploratory analysis are confirmed. Every feature is significant in predicting the outcome of at least one language. In all three models, we see circumstance predicting Japanese with a reasonably strong effect size (estimate). Where Japanese is the reference value, we see it significantly and strongly predicting against English and Swedish. The Cause 'event' is only clearly significant predicting against Swedish relative to Japanese. It is borderline significant in predicting for Japanese, relative to Swedish and a borderline significant prediction for English. In any case, the effect size in both instances is relatively small. This clarifies what we saw in the correspondence analysis: the Cause 'event' is not distinctive and is equally shared by the two languages. The role of both the Causes 'relationship – emotion' and 'unspecified' is systematically significant and relatively important in predicting English. Only relative to Japanese, do we see that is also a significant predictor of Swedish, although its predictive effect size is much smaller than that for English.

The profiles that have been confirmed statistically are exemplified in (14) – (16). These profiles are tendencies observed in the data, which can be claimed to represent tendencies in the broader language context, at least in the language of online personal diaries. To the extent that one believes that such diaries are representative of culture, one can make extrapolations about cultural differences.

ANXIETY is an OPPONENT, caused by unspecified and or abstract concerns in English

(14) […] *I also want to handle and control my depression, anxiety and mood swings on my own...*

Anxiety caused by 'circumstance'- 'state of affairs' and 'occupation'-'education'

(15) a. *また、週末で病院の救急外来の対応に若干不安がありますが*[...].
'And there is also some anxiety over how they treat emergency outpatients at the hospital during the weekend […].'

b. *すでに２年のクラス替えの不安でいっぱいなのは私だけでしょうか.*
'Am I the only one already filled with anxiety over the class change in the second year?'

ANXIETY is a MANIPULABLE OBJECT, caused by 'health'/'body' concerns in Swedish

(16) *Det här ger mig ångest [...] så måste jag se över vad jag äter och jag måste komma ut på promenader,…*
'This gives me anxiety [...] I have to look over what I eat and I have to get out for walks.'

Having interpreted the predictors, their significance and relative impotence, we can now consider the overall predictive strength of the model. For each occurrence, the model makes a prediction of the outcome (the language of the example). However, this prediction is not discrete. Its prediction takes the form of the odds that the example will be language A, B, or C. This can be converted to proportions of likelihood, or percentages, but the model will never get a given example wholly right or wrong. Therefore, expressing the accuracy of the model in terms of 'how many right and wrong' can be extremely misleading. If we simplify the prediction in this way, we need to take into account an accurate prediction of one outcome occurring (true positive) as well as an accurate prediction of it not occurring (true negative). We have also, of course, the inverse of this. The accuracy of a model is calculated in the following way:

$$accuracy = \frac{true\ positives + true\ neagtives}{false\ positive + false\ negatives + true\ positives + true\ negatives}$$

These scores are listed for each language and an overall score calculated for all three is listed in table 3. Remembering that right or wrong is calculated with the arbitrary cut-off of 50% and that we are predicating between three outcomes, we see an accuracy for Swedish of 49%, which is not that high. However, English is

predicted somewhat better at 59.3 and at over 80%, the prediction rate for Japanese is exceptionally high. This would suggest that the profile for Japanese is the most distinct. It appears the models struggle to distinguish English and Swedish, which may be a sign of the cultural similarity between these two languages.

Finally, a simpler, and perhaps more widely used way of determining the predictive strength of a logistic model is to plot the percentage of true positives against the false positives and calculate the area under the resulting curve (Hanley & McNeil 1982). The resulting score, the concordance statistic, can be understood as the percentage of cases where the model is assigned the highest probability to the correct outcome. The score ranges between 0.5, where model offers nothing more than pure chance in predicting the outcome, and 1.0, where one the model always assigns the highest odds to the correct outcome. The rules of thumb are: 0.5 > 0.6 little or no predictive strength; 0.6 > 0.7 weak prediction; 0.7 > 0.8 reasonable prediction; 0.8 > 0.9 strong prediction; over 0.9 extremely strong predictive accuracy to the point where the model assumptions should be re-checked as well as the possibility that the data have been over-fit. In table 3, we see that using Japanese as the reference level, we obtain a strongly predictive result and using English and Swedish; we also obtain reasonably predictive results. A combination of the concordance statistics and accuracy scores leave little doubt that Metaphor and Cause together are sufficient to explain the difference between Japanese and Swedish/ English and largely sufficient to distinguish the English and Swedish use of the lexemes that designate ANXIETY.

4. Summary

Let us summarise the confirmed results. It must be remembered that logistic regression modelling allows one to make claims beyond the sample about the language more broadly. At least for the language of on-line personal diaries, the following generalisations are statistically extremely likely to be valid. It should also be noted that with the exception of the Cause of body-health' / 'food-drink', the profiles identified are tendencies, not discrete distinctions.

English:	Metaphor – PAIN; Cause – 'relationship'/'emotion' and 'unspecified'/'abstract uncertainty'
Japanese:	Metaphor – no significant associations found; Cause – 'occupation' and 'social circumstance'/'state of affairs'
Swedish:	Metaphor – OBJECT; Cause – 'body-health'/'food-drink'

Although based upon only a small sample, multifactorial usage-feature analysis has been shown to be capable of offering insights into how conceptual metaphors

are structured and a means for analysing their variation across languages and cultures. Importantly, it is not just which metaphors occur, but how they are used – in this instance, what causes they are associated with. Indeed, the pairing of the Causes 'emotion' and 'unspecified fear' with the metaphors of OPPONENT and PAIN is far from counterintuitive. Just as in Swedish, the rather 'down to earth' Cause of body image is conceptualised as a MANIPULABLE OBJECT. Lastly, although the metaphor OBSTACLE was too infrequent to make any strong claims, it also intuitively matches the causes of 'work' and 'circumstance' Of course, further research would be needed to ascertain if this interpretation of an inherent association between these causes and metaphors is valid. The identification of a correlation does not, obviously, indicate a causal relation. A larger sample and a richer, more detailed, feature analysis are needed to answer such questions, perhaps combined with an elicitation-based methodology.

References

Agresti, Alan (2013). *Categorical Data* (3rd edn.). Hoboken: John Wiley.

Arppe, Antti (2012). polytomous: Polytomous logistic regression for fixed and mixed effects. R package version 0.1.4. http://CRAN.R-project.org/package=polytomous.

Chongsuvivatwong, Virasakdi (2012). epicalc: Epidemiological calculator. R package version 2.15.1.0. http://CRAN.R-project.org/package=epicalc.

Croissant, Yves (2012). mlogit: multinomial logit model. R package version 0.2-3. http://CRAN.R-project.org/package=mlogit.

Dirven, René, Louis Goossens, Yvan Putseys, Emma Vorlat (1982). *The Scene of Linguistic Action and its Perspectivization by SPEAK, TALK, SAY and TELL*. Amsterdam: John Benjamins.

Divjak, Dagmar (2006). Ways of intending: A corpus-based cognitive linguistic approach to near-synonyms in Russian. Gries, Stefan Th., & Anatol Stefanowitsch, eds. *Corpora in Cognitive Linguistics. Corpus-Based Approaches to Syntax and Lexis*. Berlin: Mouton de Gruyter, 19–56.

Divjak, Dagmar (2010). *Structuring the Lexicon: A Clustered Model for Near-Synonymy*. Berlin: Mouton de Gruyter.

Geeraerts, Dirk (1990). The lexicographical treatment of prototypical polysemy. Tsohatzidis, Savas, ed. *Meanings and Prototypes. Studies in Linguistic Categorization*. London: Routledge, 195-210.

Geeraerts Dirk, Stefan Grondelaers, & Peter Bakema (1994). *The Structure of Lexical Variation. Meaning, naming, and context*. Berlin: Mouton de Gruyter.

Geeraerts, Dirk, Stefan Grondelaers, & Dirk Speelman (1999). *Convergentie en Divergentie in de Nederlandse Woordenschat*. Amsterdam: Meertens Instituut.

Glynn, Dylan (2009). Polysemy, syntax, and variation. A usage-based method for Cognitive Semantics. Vyvyan Evans, & Stephanie Pourcel, eds. *New Directions in Cognitive Linguistics*. Amsterdam: John Benjamins, 77–104.

Glynn, Dylan (2010a). Lexical fields, grammatical constructions and synonymy. A study in usage-based Cognitive Semantics. Hans-Jörg Schmid, & Suzanne Handl, eds. *Cogni-*

tive Foundations of Linguistic Usage-Patterns. *Empirical Studies*. Mouton: Berlin: Mouton de Gruyter, 89-118.

Glynn, Dylan (2010b). Testing the Hypothesis. Objectivity and verification in usage-based Cognitive Semantics. Dylan Glynn, & Kerstin Fischer, eds. *Quantitative Cognitive Semantics. Corpus-driven approaches*. Berlin: Mouton de Gruyter, 239-270.

Glynn, Dylan (2013). Corpus-driven lexical semantic evidence for conceptual structure of ANGER. Novakova, Iva, Peter Blumenthal, & Dirk Siepmann, eds. *New Directions in Lexical Semantics and Discourse Organization*. Frankfurt: Peter Lang.

Glynn, Dylan (in press a). The many uses of *run*. Corpus methods and cognitive socio-semantics. Dylan Glynn, & Justyna Robinson, eds. *Corpus Methods for Semantics. Quantitative Studies in Polysemy and Synonymy*. Amsterdam: John Benjamins.

Glynn, Dylan (in press b). Correspondence Analysis. An exploratory technique for identifying usage patterns. Dylan Glynn, & Justyna Robinson, eds. *Corpus Methods for Semantics. Quantitative Studies in Polysemy and Synonymy*. Amsterdam: John Benjamins.

Glynn, Dylan (in press). Conceptualisation of HOME in popular Anglo-American texts. A multifactorial diachronic analysis. Díaz, J. Enrique, ed. *Diachronic Studies in Conceptual Metaphor*. Amsterdam: John Benjamins.

Glynn, Dylan (forthc.). *Mapping Meaning. Corpus Methods for Cognitive Semantics*. Cambridge: Cambridge University Press.

Glynn, Dylan, & Henrik Nordmark (submitted). The emotion concept ANXIETY in English, Japanese and Swedish. A corpus-driven lexical analysis. *Linguistica Silesiana.*

Glynn, Dylan, & Kerstin Fischer eds. (2010). *Quantitative Methods in Cognitive Semantics: Corpus-Driven Approaches*. Berlin: Mouton de Gruyter.

Glynn, Dylan, & Justyna Robinson eds. (in press). *Corpus Methods for Semantics. Quantitative studies in polysemy and synonymy*. Amsterdam: John Benjamins.

Greenacre, Michael (2007). *Correspondence Analysis in Practice*. London: Chapman & Hall.

Gries, Stefan Th. (1999). Particle movement: A cognitive and functional approach. *Cognitive Linguistics* 10: 105-145.

Gries, Stefan Th. (2003). *Multifactorial Analysis in Corpus Linguistics: A study of particle placement*. Continuum: London.

Gries, Stefan Th. (2006). Corpus-based methods and cognitive semantics: The many meanings of *to run*. Stefan Th. Gries, Anatol Stefanowitsch eds. *Corpora in cognitive linguistics: Corpus-based approaches to syntax and lexis*. Berlin: Mouton de Gruyter, 57-99.

Gries, Stefan Th. (2010). Behavioral Profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. *Mental Lexicon* 5: 323-346.

Gries, Stefan Th., & Dagmar Divjak (2006). Ways of trying in Russian: Clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2: 23-60.

Gries, Stefan Th., & Dagmar Divjak (2009). Behavioral profiles: A corpus-based approach towards cognitive semantic analysis. Vyvyan Evans, & Stephanie Pourcel, eds.: *New Directions in Cognitive Linguistics*. Amsterdam: John Benjamins, 57-75.

Grondelaers Stefan, Dirk Speelman, & Dirk Geeraerts (2009). National variation in the use of *er* "there". Regional and diachronic constraints on cognitive explanations. Gitte Kristiansen, & René Dirven, eds. *Cognitive Sociolinguistics: Language variation, cultural models, social systems*. Berlin: Mouton de Gruyter, 153-204.

Grondelaers, Stefan, Dirk Geeraerts, Dirk Speelman (2007). A case for a cognitive corpus Linguistics. Gonzalez-Marquez, Monica, Irene Mittleberg, Seanna Coulson, & Mi-

chael Spivey, eds. *Methods in Cognitive Linguistics*. Amsterdam: John Benjamins, 149-169.
Hanley, James, Barbara McNeil (1982). The meaning and use of the Area under a Receiver Operating Characteristic (ROC) Curve. *Radiology* 143: 29–36.
Hendrickx, John (2012). perturb: Tools for evaluating collinearity. R package version 2.05. http://CRAN.R-project.org/package=perturb.
Hosmer, David, & Stanely Lemeshow (2013). *Applied Logistic Regression* (3rd edn.). Hoboken: John Wiley.
Heylen, Kris (2005). A quantitative corpus study of German word order variation. Kepser, Stephan, & Marga Reis, eds. *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*. Berlin: Mouton de Gruyter, 241-264.
Jackman, Simon (2012). pscl: Classes and Methods for R, Developed in the Political Science Computational Laboratory, Stanford University. Department of Political Science, Stanford University. Stanford, California. R package version 1.04.4. URL http://pscl.stanford.edu/.
Janda, Laura, & Valery Solovyev (2009). What constructional profiles reveal about synonymy: A case study of the Russian words for SADNESS and HAPPINESS. *Cognitive Linguistics* 20: 367-393.
Meikyo Kokugo Jiten [Meikyo Japanese dictionary] (2003). Edited by Kitahara Yasuo. Tokyo: Taishukan Shoten.
Kövecses, Zoltán (1986). *Metaphors of Anger, Pride, and Love. A lexical approach to the structure of concepts*. Amsterdam: John Benjamins.
Kövecses, Zoltán (2000). *Metaphor and Emotion. Language, Culture and Body in Human Feeling*. Cambridge: Cambridge University Press.
Krawczak, Karolina, Iwona Kokorniak (2012). A corpus-driven quantitative approach to the construal of Polish *think*. *Poznań Studies in Contemporary Linguistics* 48: 439-472.
Krawczak, Karolina (in press a.) Shame and its near-synonyms in English: A multivariate corpus-driven approach to social emotions. Novakova, Iva, Peter Blumenthal, & Dirk Siepmann, eds. *New Directions in Lexical Semantics and Discourse Organization*. Frankfurt: Peter Lang.
Krawczak, Karolina (submitted). Shame and its near synonyms in English: A multivariate corpus-driven approach to social emotions. *Poznań Studies in Contemporary Linguistics*.
Krawczak, Karolina, & Dylan Glynn (2011). Context and cognition. A corpus-driven approach to parenthetical uses of mental predicates. Kosecki, Krzysztof, & Janusz Badio eds. *Cognitive Processes in Language*. Frankfurt am Main: Peter Lang, 87-99.
Krawczak, Karolina, Dylan Glynn (in press). Operationalising construal. *Of/about* prepositional profiling for cognition and communication predicates. Bretones Callejas, Carmen María, ed. *Construals in Language and Thought: What Shapes What?* Amsterdam: John Benjamins.
Lakoff, George (1987). *Women, Fire, and Dangerous Things. What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
Nationalencyklopedins ordbok [National encyclopaedia] (1996). Höganäs: Bra Böcker.
Norsteds svenska ordbok [Norsteds Swedish dictionary] (2010). Edited by Håkan Nygren. Stockholm: Norstedts.
Greenacre, Michael (2007). *Correspondence Analysis in Practice.* London: Chapman & Hall.

Sekai Dai-hyakka Jiten [Heibonsha's world encyclopaedia] (1988). Tokyo: Heibonsha. 2013-07-05.
Shibatani, Masayoshi (2005). *The Languages of Japan* (8th edn.). Cambridge: Cambridge University Press.
Speelman, Dirk, & Dylan Glynn (2005). LiveJournal Corpus of British and American online personal diaries. University of Leuven.
Speelman, Dirk (in press). Logistic Regression. A confirmatory technique for comparisons in corpus linguistics. Dylan Glynn, Justyna Robinson, eds. *Corpus Methods for Semantics. Quantitative studies in polysemy and synonymy*. Amsterdam: John Benjamins.
Rudzka-Ostyn, Byrgida (1989). Prototypes, schemas, and cross-category correspondences: the case of *ask*. *Linguistics* 27: 613-661.
Venables, William N., & Brian D. Ripley (2002) *Modern Applied Statistics with S* (4th edn.). New York: Springer.

Received July 14, 2013

Accepted for publication August 22, 2013

Authors' address:

Henrik Nordmark
Centre for Languages and Literature
Lund University
Box 201
221 00 Lund
Sweden, E.U.
nordmark.henrik@gmail.com

Prof. Dylan Glynn
Département d'Études des Pays Anglophones
Université Paris 8
2, rue de la Liberté
93526 Saint-Denis
France, E.U.
dglynn@univ-paris8.fr