A corpus-based case study of differences in meaning and usage of two near-synonyms.
American *awesome* vs. British *wicked*

Karin Rasmusson

1. Introduction

"Near-synonymy is the study of semantic relations between lexemes or constructions that possess a similar usage" (Glynn 2010: 2; Lyons 1995:60; Divjak 2006:21). In this study two near-synonyms *awesome* and *wicked* are to be looked more closely at. Near-synonyms have similar meanings but they do not mean exactly the same. The aim of the study is to try to locate whether such small differences in meaning and usage of the two lexemes are present, and if they are what these differences might be.

The first intention of this essay was to compare the different uses of *awesome* and *wicked* in British English and American English using data from personal online diaries. However, there was not enough data of the use of *awesome* in British English. Personal online diaries will still be the source for the data used in this study, but instead of examining each lexeme in both languages, data for *awesome* will be collected from American English online diaries, and data for *wicked* from British English ones.

Language changes over time and cultures, meanings of words change with it and are sometimes weakened or lost. According to the Merriam-Webster online dictionary one of the definitions of *awesome* is:

1. expressive of awe

This definition expressive of awe is in this study assumed to belong to the category of meanings of a word that has been lost due to time and change in the language. This meaning of the word is archaic and would only be found in more ancient texts such as the Bible. Only one of the three definitions listed in previous mentioned dictionary was found in the data used. However, several unlisted meanings could be found for both lexemes.

2. Method

Corpus linguistics implicates that large numbers of examples provided by computerised corpora i.e. large collections of data, are analysed to make

generalisations of frequencies found in authentic language usage (Glynn 2010:2; Gries, Hampe, Schönefeld 2005: 635). The data for this corpus-based study was taken from both British and American personal online diaries from the year 2006, further on referred to as blogs. The first aim to compare the different uses of the lexemes in British and American English was not possible, as mentioned in the introduction. The intention was to analyse 500 examples in total, 125 of each dialect and lexeme, however, there was only about 30 useful examples of the British use of *awesome*. Instead a new aim to compare the use of *awesome* in American English, and *wicked* in British English, was decided on to see if there are any differences between the two near-synonyms and the usage of them. 500 examples, 250 of each lexeme were selected from the data. However, after the data had been copied into Note Pad, and then into an Excel sheet for further cleansing to lose examples not found useful, in total 471 examples remained. 249 belonged to the lexeme *awesome*, and 222 to the lexeme *wicked*. With the examples in order, a coding schema with the intention of finding any differences between the two near-synonyms and the usage of these lexemes was composed. This was the most difficult, troublesome, and time-consuming part of the study. The finished coding schema was then copied into the computer program R Project version 2.12.2, which was used to perform statistic analyses, such as correspondence analyses, and logistic regression analyses. The analyses and the results of these analyses are presented in section 4 below.

3. Analysis

What follows are a closer description of the coding schema.

3.1 Lexeme

The factor *Lexeme* contained the two lexemes; *awesome* and *wicked*. In total 471 examples, 249 of these examples belonged to the lexeme *awesome*, and 222 belonged to the lexeme *wicked*. A reason for there being fewer of the lexeme *wicked* is that not as many applicable examples, as for awesome, could be found within the data.

3.2 Form

Form was a factor not found useful in the analysis due to the fact that there were only two forms found of the two lexemes. Adverbs such as awesomely and wickedly were only found in such small numbers as not to be useful in finding any sufficient statistic results had they been included in the coding schema. The same was true for the nominalised adjective forms, awesomeness and wickedness. One case of *awesome* being used as a noun was also found, but could not be included due to the reason mentioned above. The example with awesome being used as a noun can be seen in section 3.4 below.

3.3 Dialect

Dialect is a factor that would have been useful, had the study dealt with the initial intended thesis. However, since all the data containing *awesome* is taken from American blogs, and the data containing *wicked* are taken from British, this factor has to be considered rather redundant.

3.4 Word class

Both *awesome* and *wicked* are adjectives, and therefore a distinction between predicative adjectives and attributive adjectives was made. 232 attributive adjectives were found and 239 predicative adjective. In the data, one examples of *awesome* being used as a noun was found, but not included in the coding schema since it would not have any significance in an analysis due to the low number of examples.

1. *"Why can't the Japanese save some fucking **awesome** for the rest of us?"* (noun)

2. *summer is gonna be **WICKED**!!* (predicative adjective)

3. *she is the most **awsome** girl i ever could have asked for.* (attributive adjective)

3.5 Emphatic

Under the factor Emphatic two features are to be found: Emphatic, and Non Emphatic. Emphatic includes both expletive words and non-expletive words, but also exclamation marks, smilies and capital letters. Non Empathic is used where the examples lack emphasis. A third feature was thought of initially, but was left out since it did not amount to a sufficient number of examples. This feature was Weak Emphatic, used for a few cases with words such as pretty and quite. It could be debated whether or not they should be looked upon as emphatic at all, or categorised as Non Emphatic. It was decided, on grounds mentioned above, that these examples were to belong to the feature Non Emphatic. In the statistical analyses preformed in the programme R Console, Emphatic did not prove to be significant.

1. *We even recorded a few things,* **such** *a wicked time!* (Emphatic, markers in bold)

2. Its so **fucking** *awesome.* (Emphatic, with the expletive word in bold)

3. *This past week was* **pretty** *awsome.* (An examples of what could have belonged to the feature Weak Emphatic, depending on where the stress is interpreted to be.)

4. *Wicked fun!:)* (Empathic, markers in bold)

3.6 Axiology

Axiology: most of the data was used in a positive connotation (456 examples). A small number of examples containing a negative connotation (15 examples) were found. According to a correspondence analysis consisting only of *Lexeme* and *Axiology* (section 4.1), *wicked* was more closely linked with a negative connotation. This is according to what was discovered during the coding as well, of the 15 examples with a negative connotation 14 belonged to the lexeme *wicked*. It can be assumed that a reason for most of the examples having a positive connotation is due to that the lexemes are most often used with a positive meaning. In an attempt to explain why the examples with a negative connotation belonged to the

British *wicked*, a generalising assumption that the British might fancy having a bit more self-irony than the Americans was made. The positive connotation is close to the core position of the correspondence analysis, and therefore almost equally shared by the two lexemes.

1. *happy new year. another new year spent in the house due to a **wicked** bad dose of the flu*. (Negative connotation)

2. Spain was ***wicked***, it was really hot and sunny. (Positive connotation)

3.7 Intensifier

Features included in the factor Intensifier were; Intensifier, which include intensifiers such as **totally** *awesome/wicked*, **so** *awesome/wicked* etc., Expletive Intensifier, which were mostly represented by different variations of the word fucking, as in; "it was a **fucking** *wicked* night". For those examples lacking any intensifier Non Intensifier was used. This factor, as well as the previously mentioned Emphatic, did not prove to be significant in the analyses.

3.8 Polarity

Polarity proved to be of very little variation, out of 471 examples, only one included a negation. The negation was found among the lexeme *awesome*. Due to this fact it was not used in the analyses.

3.9 Humour

In 419 of the 471 examples no humour was found. If this is due to not mastering the language as a non-native speaker of English, or if there was no humour is difficult to say. However, this makes for a quite subjective analysis. In the British examples(*wicked)* only 37 contained humour, to 185 containing no humour. This outcome was rather unexpected as it can be assumed that the British are quite fond of their self-irony/humour, but as mentioned above, this could have been overlooked as a non-native speaker.

3.10 Theme

Theme is a factor that could have been made a bit finer grained to begin with, but a choice based on time limitation was made, and it is rather coarse grained. Theme includes the features; Holiday, Love, Entertainment, Family Friends, Travel, Leisure, Work/School, Presents, Party, Celebration, Shopping, Appearance, and finally Misc personal and Misc Impersonal. Misc = miscellaneous. Entertainment is a large feature in which many things, such as movies, concerts, music, books, are included. This is one feature that could have started out as much finer grained. Both Misc-features were used when none of the above was a suitable fit. A distinction between whether it could be seen as personal or impersonal was made for the Misc-features. The other features do not need any further presentation, but can be viewed in the coding schema.

3.11 Theme 2

In Theme 2 an even more coarse grained selection was made. The fourteen features were grouped together into six features. Party and Presents were grouped together under Entertainment, Shopping, Holiday, and Travel were grouped together under Leisure, Love under Family Friends, Appearance and Work/School did not fit into any group, and were left as they were before. Misc personal and Misc impersonal created the group Misc. Theme 2 became relevant since there were some features in the factor Theme that contained fewer examples than suitable for statistical analysis.

3.12 Verb

All verbs found in the clause, or phrase that the lexeme was a part of are listed under the factor Verb. This will not be looked into further in this section, but are to be found in the coding schema. NA (non-applicable) was used in those cases where no verb was present.

1. *Wicked fun!:)* (NA)

3.13 Verb Group

A list of all the verbs was made, and all the verbs in the factor Verb were grouped according to meaning similarities, with the exceptions of a few verbs that could not be fitted into a specific group, which were grouped under Other. The copular verb *be* alone created the feature: copular.

3.14 Verb Group 2

N.B. (Grammatically these verbs may not belong to either event, or state verbs. They are just labels used in the coding schema.)

Verb Group	**Verb Group 2**
Sense verb	Event
Utterance verb	Event
Occurrence verb	Event
Exchange verb	Event
Creative verb	Event
Interaction verb	Event
Action verb	Event
Attention verb	Event
Other	Event
Conditional verb	State
Emotion verb	State
Existence verb	State
Desire verb	State
Experience verb	State
Copular	Copular
NA	NA

3.15 Dictionary definitions

A more fine-grained coding of dictionary definitions could have been made. An aspect to take into consideration here is that, previously mentioned, of not being a native speaker, and how easily the sometimes small variations in meaning can be missed due to this fact.

Dictionary definitions of the American senses of the adjective *awesome* according to Merriam Webster Online:

1. expressive of awe
2. a: inspiring awe
   b: terrific; extraordinary

Of the different senses, stated in the Webster Dictionary, only 2b were found to be applicable. Other meanings of *awesome* found in the data were good, great, fun, cool and personal impression of other people.

Dictionary definitions of the British senses of the adjective *wicked* according to The Oxford English Dictionary Online:

1. evil or morally wrong
2. playfully mischievous
3. *informal* excellent; wonderful

All three dictionary definitions of *wicked* were found on more than one occasion in contrast to the dictionary definitions of *awesome*. A few more meanings were also found in the data, such as; fun, cool, crazy, good, and personal impression of others. However, not enough examples of these meanings were found to amount to figures that could be satisfactory for a statistical analysis.

3.16 Animacy

The factor Animacy includes two features; the example is either animate or inanimate.

1. *Shes a great girl, I'm so lucky to have a beauiful girl that's so **awesome***. (Animate)

2. *I'm back from Nottingham. It was **wicked***. (Inanimate)

3.17 Type

Type contains six different features; Human, Concrete Thing, Abstract Thing, Concrete Activity/Event, Abstract State of Affairs, and Period of Time. Concrete and Abstract things are pretty straightforward. Concrete Activity/Event and Abstract State of Affairs, however, caused some confusion at times. Not many examples came to fall under this feature since

it was difficult to apply examples accurately to this concept. Period of time was used for examples including periods of time ranging from a night or day, a weekend, a week to e.g. a summer.

1. *theres also another guy an older one. but he has a g.f . . . but he's **awsome** none the less.* (Human)

2. *…Was so much busier and pumping than last time we went and the atmosphere was **wicked**.* (Abstract Thing)

3. *It was **wicked** to hear different drum fills for all the classic Metallica songs.*(Concrete Activity/Event)

3.18 Ref Type Abstract/Concrete

Under the variable Ref Type Abstract/Concrete, Abstract Thing and Abstract State of Affairs, and Human, Concrete Thing and Concrete Activity/Event were grouped together.

3.19 Ref Type Animacy

This factor is a duplication of the factor Animacy in section 3.16 above. It was not found to be of any greater use since it consists of exactly the same as the factor Animacy just mentioned.

3.20 Ref Type Thing/Process

Concrete Activity/Event and Abstract State of Affairs were grouped together as Process, and Concrete Thing and Abstract Thing were grouped together as Thing, under the factor Ref Thing/Process. 393 examples qualified as thing and 78 as process.

3.21 Subjecthood

Subjecthood includes Subject, Direct Object or SubNA. SubNA was used with attributive adjectives when the subject was missing, and with

predicative adjectives when *awesome* or *wicked* belonged to the phrase or clause in subject predicative position. Stating in the coding schema whether the predicative adjective belonged to a phrase or clause in an adverbial position or in subject or object predicative position, was unfortunately overlooked when coding.

1. *…so I can get all the* **awsome** *layouts.*(Direct Object)

2. *But a* **wicked** *army general type person wouldn't let me go.*(Subject)

4. Results

Two different statistic analyses techniques were applied to investigate whether any differences between the two lexemes do exist, and if they do what these differences might be. The Correspondence Analysis, is a technique which calculates how often different features co-occur and transforms these frequencies into spatial distances. Data in the Correspondence Analysis clustered closely together, are usage-features highly associated with each other and dissociated usage-features are found far from each other. A Correspondence Analysis does not suggest whether the patterns found in the data actually represent the usage of the features or how much is simply chance (Glynn in press: 12). The second technique used is called Logistic Regression. It is a confirmatory technique which models the data and tries to predict a binary outcome based on the analysis of the usage features in the Correspondence Analysis. If the prediction is correct, it can be believed that that the analysis has captured the differences in the data. (Glynn in press: 12)

4.1 Lexeme and Axiology

Figure 1. Lexeme, Axiology
Correspondence Analysis
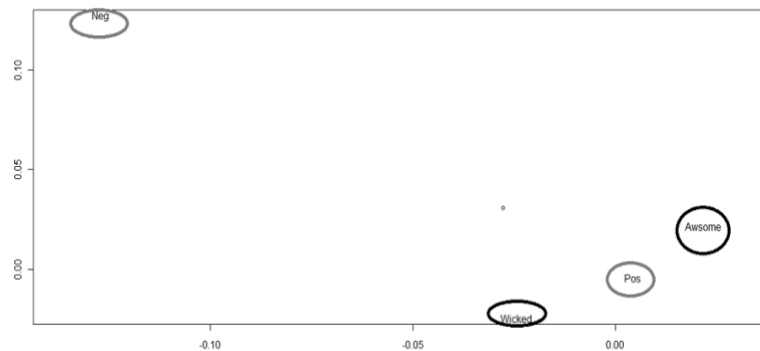
This Correspondence Analysis containing Lexeme and Axiology is mentioned in section 3.6 above. 456 of the examples had a positive connotation (Pos) which is displayed above as a feature shared by both lexemes. That the negative connotation (Neg) proved to be more closely linked to the lexeme *wicked* was not very surprising since 14 of the in total 15 examples with a negative connotation belonged to examples including the lexeme *wicked*. Another fact to take into consideration here is the low amount of examples with a negative connotation. This could be a reason to why Neg is found rather far away from both lexemes.

A Logistic Regression Analysis (enclosed) was performed to see how accurate the results of the Correspondence Analysis were. It showed high values in the S.E. column and significance, but a C-value of 0.53 and an R^2 of 0.043 showed that this outcome cannot be accurately predicted.

4.2 Lexeme Theme 2 and Type

Figure 2. Lexeme, Theme 2, and Type
Correspondence Analysis

The correspondence analysis above shows the results of the analysis of the features of Lexeme, Theme 2, and Type. The lexeme *wicked* is closely associated with the feature Entertainment (Theme), and the two features; Concrete Thing (Type) and Abstract Thing (Type) in the upper left corner of *Figure 2. Wicked* is also closely associated with Period of Time (Type) and Leisure (Theme) in the lower left corner. The three features in the middle occur in a cluster together Work/School (Theme), Appearance (Theme), and Concrete Activity/Event (Type). These features are shared by both lexemes. Misc is in between the cluster belonging to *wicked* and the shared cluster, making it more associated with *wicked* than with *awesome*. In the lower right half of *Figure 2.* Abstract State of Affairs (Type) is closely linked with the lexeme *awesome*. In the upper right corner is a cluster of particular interest. The cluster of Human (Theme) and Family Friends (Theme) are to the right of the lexeme *awesome*; as far as it can come from the lexeme *wicked*. This proves that these features are highly associated with the lexeme *awesome*. It can be concluded that *awesome*, in a much higher degree that *wicked*, is used with human referents, and relations between humans. A hypothesis for this is that the Americans, since the data for *awesome* is taken from American blogs, are more family oriented. It would be interesting, had there been enough data for the British

use of *awesome*, to examine whether *awesome* would be used in a similar way in Britain, or if the hypothesis would be accurate. A hypothesis for *wicked* being used in a higher extent with things, and periods of time is that the British enjoy having a good time, and in many cases do so including alcohol. *Wicked* is not used with human, animate referents, but is rather used in a more impersonal way.

It can also be noticed in *Figure 8.* section 4.6 that Attributive adjectives (adjectives functioning as a premodifier before a noun) are closely associated with *wicked*, which is, as mentioned above, also more associated with things and periods of time than with humans and relations between humans. *Awesome* is in turn, more associated with predicative adjectives (adjectives that occur in the subject predicative position, following a copular verb, often describing a human subject as in:

1. *She is **awesome**.*

where *awesome* is a complement of the subject; she). Due to predicative adjectives often describing a human subject, it was no surprise to find human referents and humans such as family members and friends/lovers in close association with predicative adjectives.

4.3 Logistic Regression Lexeme Theme 2 Type

A list of the usage-features used in the model is found to the left in the Logistic Regression Analysis below. Next in line is the Standard Estimate column, referred to as S.E. A higher figure provides a higher predictive strength of the feature (Glynn in press: 13). If the numbers under the heading Error, are very high there is some kind of problem with the analysis. To the far right of the model the p-value is found. Ultimately the *p*-value should be as close to zero as possible. The closer to zero, the more distinctly possible is it to receive the same or better results at a repeat analysis. The *p*-values followed by dots or stars are more significant in distinguishing differences between the two lexemes. To distinguish the two lexemes a subtraction sign is added in front of the lexeme which comes first in alphabetical order i.e. *awesome*.

Figure 3. Lexeme, Theme 2, and Type
Logistic Regression Analysis (glm)

```
Coefficients:
                         Estimate Std.Error z value Pr(>|z|)
(Intercept)               -0.6283   0.9357   -0.672  0.50188
Theme_2Entertainment      -0.3127   0.6739   -0.464  0.64259
Theme_2Family_Friends     -1.7683   0.7122   -2.483  0.01304 *
Theme_2Leisure            -0.9377   0.6961   -1.347  0.17798
Theme_2Misc               -0.5693   0.6988   -0.815  0.41523
Theme_2Work/School        -0.2991   0.7512   -0.398  0.69050
TypeAbstract_thin          1.8657   0.7105    2.626  0.00865 **
TypeConcrete_Actvty/Event1.0129   0.7109    1.425  0.15422
TypeConcrete_thn           1.1636   0.6870    1.694  0.09032 .
TypeHuman                  0.8235   0.7155    1.151  0.24976
TypePeriod_oftme           1.9474   0.7176    2.714  0.00665 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 651.4  on 470  degrees of freedom
Residual deviance: 585.0  on 460  degrees of freedom
AIC: 607

Number of Fisher Scoring iterations: 4
```

In the model above, Lexeme was used as the Response Variable since the study is trying to find differences in meaning and usage of the two lexemes. Lexeme was analysed here together with Theme 2 and Type. In this model three features received at least one star and one feature received a dot. These stars and dots reveal that these features carry statistical differences in the usage of the two lexemes in the analysis. The figures in the S.E. column for the features which received stars are relatively high, and the figures in the right hand column are relatively close to zero, this can be an indicator of that even if the c-value and the R^2 prove to be quite low, which is to be expected since the two lexemes are near-synonyms, the model can still be looked upon as significant and the features are significantly different. Both Abstract Thing (Type), Period of Time (Type), and Concrete Thing (Type) are significant features for the lexeme *wicked*, as is Family Friends (Theme) for *awesome*. This corresponds to what was seen in the Correspondence Analysis above. The C-value (69.5%) in this model is not as high as desirable (at least 80%), but as mentioned above, the C-value should not be expected to be very high as *awesome* and *wicked* are near-synonyms. In this model the R^2 is 0.175, a desirable R^2 figure would be 0.3 or higher. However, as the figures in the S.E. column are acceptably high along with adequate *p*-values, the differences pass as significant. The

features of Type can be said to carry differences in meaning, and the features of Theme can be said to carry both meaning and cultural differences.

Figure 4. Lexeme, Theme 2, and Type
Logistic Regression (lrm)

```
Frequencies of Responses
Awsome Wicked
  249    222
     Obs  Max Deriv Model L.R.      d.f.   P    C          Dxy
     471     3e-08      66.39              10     0    0.695
0.39
    Gamma      Tau-a       R²      Brier
    0.418      0.195      0.175     0.217

                              Coef    S.E.   Wald Z  P
Intercept                    -0.6283 0.9357 -0.67  0.5019
Theme_2=Entertainment        -0.3127 0.6739 -0.46  0.6426
Theme_2=Family_Friends       -1.7683 0.7122 -2.48  0.0130
Theme_2=Leisure              -0.9377 0.6961 -1.35  0.1780
Theme_2=Misc                 -0.5693 0.6988 -0.81  0.4152
Theme_2=Work/School          -0.2991 0.7512 -0.40  0.6905
Type=Abstract_thing           1.8657 0.7105  2.63  0.0086
Type=Concrete_Activity/Event  1.0129 0.7109  1.42  0.1542
Type=Concrete_thing           1.1636 0.6870  1.69  0.0903
Type=Human                    0.8235 0.7155  1.15  0.2498
Type=Period_of_time           1.9474 0.7176  2.71  0.0067
```

15

4.4 Lexeme Theme 2 Type and Humour

Figure 5. Lexeme, Theme 2, Type, and Humour
Correspondence Analysis
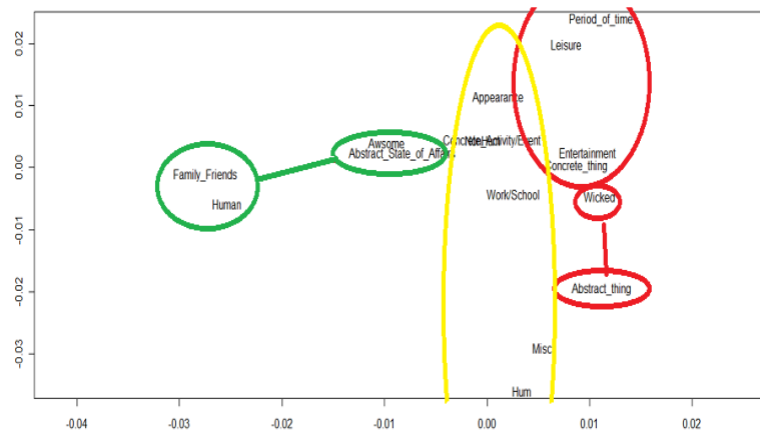
In this Correspondence Analysis, Humour was added to the factors used in the previous analysis (Theme 2, and Type). When Humour was added the patterns changed. Misc (Theme) was included as a member of the cluster in core position containing; Concrete Activity/Event (Type), Appearance (Theme) and Work/School (Theme). Both Hum(our) and Non Hum(our) joined the core cluster. Non Hum is definitely a shared feature since it is very close to core position. Humour is slightly closer to *wicked*, but still a shared feature. As Humour was added to the plot *awesome* and Abstract State of Affairs became much closer associated than in the previous analysis. Another interesting thing is that Period of Time (Type) and Leisure (Theme) joined the same cluster as Concrete Thing (Type) and Entertainment (Theme). The two latter features were closely linked with Abstract Thing in the previous analysis which has now moved to the other side of the lexeme *wicked.*

4.5 Lexeme Theme Type and Humour

Figure 6. Lexeme, Theme, Type and Humour
Logistic Regression (glm)

```
Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)               0.78006   1.02677    0.760 0.447422
TypeAbstract_thing        2.03178   0.73334    2.771 0.005595 **
TypeConcrete_Activ/Event1.07540   0.73214    1.469 0.141874
TypeConcrete_thing        1.40532   0.71413    1.968 0.049082 *
TypeHuman                 0.75951   0.73651    1.031 0.302429
TypePeriod_of_time        2.53170   0.76770    3.298 0.000975 ***
ThemeCelebration         -1.61412   0.96252   -1.677 0.093547 .
ThemeEntertainment       -0.42061   0.69362   -0.606 0.544257
ThemeFamily_Friends      -1.64829   0.73317   -2.248 0.024567 *
ThemeHoliday             -1.67194   0.86546   -1.932 0.053376 .
ThemeLeisure             -2.36434   0.82106   -2.880 0.003981 **
ThemeLove                -2.71418   0.91791   -2.957 0.003107 **
ThemeMisc_impersonal     -1.53919   0.76664   -2.008 0.044674 *
ThemeMisc_personal       -0.50715   0.76639   -0.662 0.508135
ThemeParty               -0.19724   0.75740   -0.260 0.794537
ThemePresents            -1.07754   0.83422   -1.292 0.196473
ThemeShopping            -0.83370   0.86749   -0.961 0.336530
ThemeTravel               0.02273   0.79803    0.028 0.977273
ThemeWork/School         -0.44485   0.76165   -0.584 0.559176
HumourNonHum             -1.51325   0.37265   -4.061 4.89e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 651.40  on 470  degrees of freedom
Residual deviance: 539.43  on 451  degrees of freedom
AIC: 579.43

Number of Fisher Scoring iterations: 4
```

Lexeme was used as the Response Variable for this model as for the previous Logistic Regression. In this model Theme, Type and Humour were also included in the analysis. Non Hum(our) and Period of Time both prove to be significant in distinguishing differences between the two lexemes, which is indicated by the three stars following the *p*-value in the right hand column. Period of Time is an important feature for *wicked* and Non Hum(our) for *awesome*. Even though Non Hum(our) was close to core position in the Correspondence Analysis in *Figure 5.*above, it proves to be more significant for *awesome* in the Logistic Regression above. This result can be linked to the assumption that the British have a bit more self-irony than the Americans, previously discussed. Furthermore, *awesome* can be seen as a word used in a more serious connotation as with family members

and friends. It can be assumed that while talking to, or about, people that one has a close relationship with, things said are regarded as more serious. On the other hand, while talking about things such as concerts and parties one might be seen as subjects more easily joked about. Abstract Thing (Type), Leisure (Theme) and Love (Theme) all received two stars, and all have high figures in the S.E. column indicating high significance for these features as well. Two features; Misc Impersonal (Theme) and Family Friends (Theme) have one star each, but not very high figures in the S.E. column. This model also contains two dots belonging to Celebration (Theme) and Holiday (Theme) both with S.E. figures around 1.6, an acceptable figure. One reason for the two last mentioned features being rewarded a marker of significance could be due to there not being enough examples of each, something that was noticed when the data was inserted into R where the figures for each feature can be displayed. The chance, that the results for Period of Time and Non Hum(our) are accurate and would predict the same or similar results if the model is to be performed repeatedly, are substantially high. The chance for the above mentioned accuracy and probability to predict the same or better results are between 94-97 % for those features with two stars. For the one dot- and one star-feature the chance for an accurate analysis are somewhat lesser than for the former.

In the Logistic Regression (lrm) below a fairly high R² 0.282 and a C-value of 0.768 indicate that this is a strong model in predicting a repeatable result for the features Non Hum(our) and Period of Time.

*Figure 7.*Logistic Regression (lrm)

```
Frequencies of Responses
Awsome Wicked
   249    222

      Obs  Max Deriv Model L.R.       d.f.         P          C
Dxy
      471      1e-09     111.96         19         0      0.768
0.537
     Gamma      Tau-a        R²      Brier
     0.553      0.268     0.282      0.195

                                Coef    S.E.   Wald Z P
Intercept                       0.78006 1.0268  0.76  0.4474
Type=Abstract_thing             2.03178 0.7333  2.77  0.0056
Type=Concrete_Activity/Event    1.07540 0.7321  1.47  0.1419
Type=Concrete_thing             1.40532 0.7141  1.97  0.0491
Type=Human                      0.75951 0.7365  1.03  0.3024
Type=Period_of_time             2.53170 0.7677  3.30  0.0010
Theme=Celebration              -1.61412 0.9625 -1.68  0.0935
Theme=Entertainment            -0.42061 0.6936 -0.61  0.5443
Theme=Family_Friends           -1.64829 0.7332 -2.25  0.0246
Theme=Holiday                  -1.67194 0.8655 -1.93  0.0534
Theme=Leisure                  -2.36434 0.8211 -2.88  0.0040
Theme=Love                     -2.71418 0.9181 -2.96  0.0031
Theme=Misc_impersonal          -1.53919 0.7666 -2.01  0.0447
Theme=Misc_personal            -0.50715 0.7664 -0.66  0.5081
Theme=Party                    -0.19724 0.7574 -0.26  0.7945
Theme=Presents                 -1.07754 0.8342 -1.29  0.1965
Theme=Shopping                 -0.83370 0.8675 -0.96  0.3365
Theme=Travel                    0.02273 0.7980  0.03  0.9773
Theme=Work/School              -0.44485 0.7616 -0.58  0.5592
Humour=NonHum                  -1.51325 0.3727 -4.06  0.0000
```
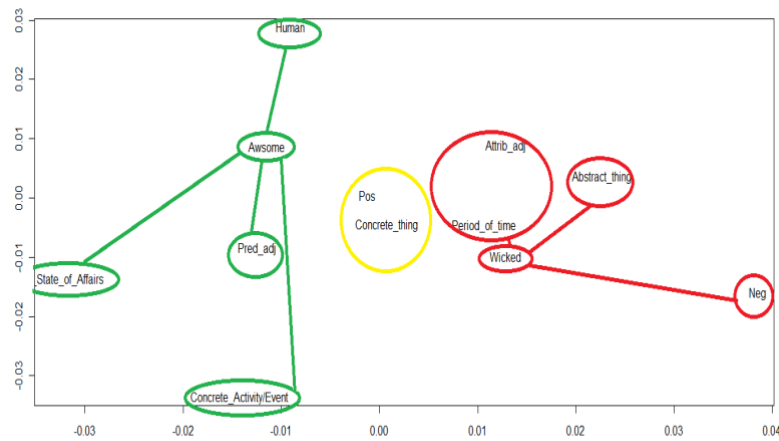
19

4.6 Lexeme Axiology Word Class and Type

Figure 8. Lexeme, Axiology, Word Class, and Type
Correspondence Analysis

The Correspondence Analysis above consists of the features of Lexeme, Axiology, Word Class and Type. The left cluster belongs to the lexeme *awesome*, the right cluster to *wicked,* and the cluster in the middle is the core cluster. The core cluster consists of two features shared by the lexemes, Concrete Thing (Type) and Pos (Axiology). In the left cluster; Human (Type), Pred Adj (Word Class) and Concrete Activity/Event (Type) are all associated with *awesome*. Abstract State of Affairs, seen in the left hand corner of the Correspondence Analysis is placed to the left of the lexeme *awesome*, as far away from *wicked* as it possibly can be, indicating that this feature is highly associated with *awesome*. The right cluster; Period of Time (Type) and Attrib Adj (Word Class) are closely linked with *wicked*. As the left cluster, the cluster to the right also contains features highly associated with the lexeme and found far away from the other lexeme. These features are Abstract Thing (Type) and Neg (Axiology). As mentioned previously, that the negative connotation is to be found highly associated with *wicked* is not surprising since most of the examples with a negative connotation are found with this lexeme. Another reason for Neg to be found so far out to the right can be that there were not many examples in total.

4.7 Lexeme Axiology Word Class and Type

Figure 9. Lexeme, Axiology, Word Class, and Type
Logistic Regression (glm)

```
Coefficients:
                   Estimate Std. Error z value Pr(>|z|)
(Intercept)          1.9533  1.2450   1.569   0.1167
AxiologyPos         -2.8947  1.0575  -2.737   0.0062 **
Word_classPred_adj  -0.4812  0.2086  -2.307   0.0211 *
TypeAbstract_thing   1.5945  0.7087   2.250   0.0245 *
TypeConcrete_Activ/Event0.9075  0.6977  1.301   0.1934
TypeConcrete_thing   1.1928  0.6727   1.773   0.0762 .
TypeHuman            0.0183  0.6929   0.026   0.9789
TypePeriod_of_time   1.7188  0.7032   2.444   0.0145 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 651.4  on 470  degrees of freedom
Residual deviance: 586.0  on 463  degrees of freedom
AIC: 602

Number of Fisher Scoring iterations: 5
```

A Logistic Regression, including the same factors as the Correspondence Analysis in *Figure 8.* was also performed to see whether the patterns in the data were pure chance or if they actually represent the usage of the features. There is approximately 99.3 % chance that the result (**) for Pred Adj (Word Class) are accurate and will be predicted at a repeat analysis of the same factors. Between the other features the percentages vary from approximately 93-98%. The S.E. figures for those features with one or two stars are quite good which proves there to be a predictive strength of these features, even though the R^2 figure is low, and the C-value is 70.5% in *Figure 10.* below. When dealing with near-synonyms the C-value should not be expected to be very high, as mentioned before.

21

Figure 10. Lexeme, Axiology, Word Class, and Type

```
Logistic Regression (lrm)
Frequencies of Responses
Awsome Wicked
  249    222


      Obs  Max Deriv Model L.R.       d.f.       P            C
Dxy
      471       9e-08     65.39          7       0        0.705
0.41
     Gamma       Tau-a         R²        Brier
     0.452       0.205      0.173        0.217


                                Coef    S.E.   Wald Z  P
Intercept                    1.95326  1.2450   1.57   0.1167
Axiology=Pos                -2.89470  1.0575  -2.74   0.0062
Word_class=Pred_adj         -0.48123  0.2086  -2.31   0.0211
Type=Abstract_thing          1.59448  0.7087   2.25   0.0245
Type=Concrete_Activity/Event 0.90753  0.6977   1.30   0.1934
Type=Concrete_thing          1.19282  0.6727   1.77   0.0762
Type=Human                   0.01830  0.6929   0.03   0.9789
Type=Period_of_time          1.71878  0.7031   2.44   0.0145
```

5. Summary

The many problems with the coding schema and confusion about what the aim of this study was going to be made the process much more time-consuming and painful than necessary. Had the intension of the study been more clearly stated at the beginning of the process the coding schema and what it was meant to cover would have been custom made to fit the research question, and hopefully resulted in what could have been thought to be somewhat clearer and satisfactory results. However, interesting results such as; *awesome* being closely associated with human referents and predicative adjectives, and also Non Humour being a highly significant feature for the lexeme *awesome* were found. Non Humour is the feature that received the absolute best results in the Logistic Regression in section 4.5. As for the lexeme *wicked*, it was found that *wicked* is more associated with things, periods of time, and attributive adjectives. *Wicked* was also found to be used in negative connotations. Period of Time proved to be of high significance to the lexeme *wicked*. This all fits very well with assumptions made during the coding process.

Reasons to why the same results might not be obtained if the study were to be remade are as follows: a) that there are not many differences in the usage of meaning between the two near-synonyms, b) there were mistakes in the coding schema causing problems such as with dictionary definitions,

c) the coding schema lacked valuable factors and features, d) the aim of the study was not clearly stated at the beginning of the process, e) the coding of the data is quite subjective, and would probably be carried out differently by another coder.

A few matters would be interesting to look into for further studies; the gender and class of the speaker, and also how the meanings of the lexemes have changed over period of time using other sources along side with blogs for a wider scope. A closer look at the polysemy of both lexemes would also be interesting to take into consideration.

.

Karin Rasmusson

References

Divjak, Dagmar
  2006   Ways of Intending: Delineating and Structuring Near-Synonyms.
           St. Gries & A. Stefanowitsch (edd). Corpora in cognitive
           linguistics. 19-56. Berlin: Mouton
Glynn, Dylan
  2010   Lexical Fields, Grammatical Constructions and Synonymy.
           A study in usage-based Cognitive Semantics
           Berlin:Mouton de Gruyter: 2
Glynn, Dylan
  In press   Corpus-driven Cognitive Linguistics. A case study in
           polysemy
           Lund University: 12, 13
Gries, Stefan Th. Hampe, Beate and Schönefeld, Doris
  2005   Converging evidence: Bringing together experimental and corpus
           data on the association of verbs and constructions.
           Cognitive Linguistics 16-4 (2005), 635-676

Lyons, John
  1995   Linguistic Semantics: An Introduction. Cambridge: Cambridge
           University Press: 60

http://www.oxforddictionaries.com/view/entry/m_en_gb0951....

http://www.merriam-webster.com/dictionary/awesome

Appendix 1.

CoefficientLogistic Regression for 4.1 enclosed

```
            Estimate Std. Error z value Pr(>|z|)
(Intercept)    2.639      1.035   2.550  0.01079 *
AxiologyPos   -2.815      1.039  -2.708  0.00676 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 651.40  on 470  degrees of freedom
Residual deviance: 635.98  on 469  degrees of freedom
AIC: 639.98

Number of Fisher Scoring iterations: 5
```

```
Frequencies of Responses
Awsome Wicked
   249    222

     Obs   Max Deriv Model L.R.        d.f.          P            C
Dxy
     471      3e-08      15.41          1      1e-04         0.53
0.059
     Gamma       Tau-a         R²        Brier
     0.887       0.029       0.043       0.242

           Coef  S.E.  Wald Z P
Intercept    2.639 1.035  2.55  0.0108
Axiology=Pos -2.815 1.039 -2.71  0.0068
```