Cognitive Linguistics:
The Conceptualisation of Happiness

Name: Sara Rundberg
Submitted in partial fulfillment of the C-level English Linguistics option
Supervisor: Dylan Glynn
Centre for Language and Literature
Lund University
Cognitive Linguistics:  
The Conceptualisation of Happiness

“Happiness is a moist potato”

Sara Rundberg

1. Introduction

In today’s society there is an obsession with finding happiness. According to the Coca-Cola Company, all you need to do is open happiness¹, but why stop with a can of coke when you can find your happiness on a dating site. It even seems to come in edible form in your local shop. But what exactly is happiness?

De Saussure spoke of “the arbitrary nature of the sign” as a two part model; “the signified” and “the signifier”. The sign itself is a result of these two parts and can take form as words, sounds, images, smells and more (Saussure 1983:67). “The signified” represents the concept, almost like a mental picture, and “the signifier” represents the form, that is the actual word. Hence, to understand the full concept of happiness not only the words needs to be analyzed but also the mental picture associated with the word.

Through the methods of Cognitive Linguistics this paper will try to examine the concept of happiness by looking at the tree lexemes; happiness, blissful and joyful. I will try to determine if the three lexemes are used differently when speaking of happiness. In order to capture the full concept, the corpus analysis will contain both examples from a blog corpus and images from Google.

2. Method

Cognitive Linguists study how the usage of language can reflect how people perceive the world. This entails that the meaning of a word can be determined by finding patterns in the way it is used (Glynn 2010:2). This should therefore also apply to both written language and visual language such as pictures.

2.1. Collecting data

2.1.1. Blog corpus

To begin with, a vast amount of data was collected from a blog corpus containing the lexemes. All the examples were put into an excel spreadsheet to be cleaned. The initial plan was to look at the lemmas; happy, bliss, joy and content, but while cleaning the data some problems occurred. Many examples had to be deleted due to faulty sentence constructions or duplicated examples. Also most of the examples on joy were women’s names and as for bliss and content, there wasn’t enough usable data. In order to get a reasonable amount of examples for each word the choice was made to use happiness, blissful and joyful. This resulted in that 40 examples for each lexeme and dialect were needed to bring the total to 240 examples.

Table 1. Collected data for each lexeme.

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>UK</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happiness</td>
<td>52</td>
<td>50</td>
</tr>
<tr>
<td>Blissful</td>
<td>29</td>
<td>35</td>
</tr>
<tr>
<td>Joyful</td>
<td>29</td>
<td>45</td>
</tr>
</tbody>
</table>

As seen above there were problems finding usable data for both blissful and joyful, especially from UK blogs. Consecutively, to get the right amount of
examples for each lexeme more examples from blog entries containing *happiness* and *joyful*.

### 2.1.2. Google Images

Collecting data for the image analysis was not as easy as collecting data from the blog entries. Firstly a search was made on Google Images on each lexeme. To specify the type of images a setting on the advanced image search was used to distinguish between photo content and faces. When the images were found the issue arose on how to save each picture individually with its caption. One solution was copying and pasting the image into a word document, but this would be time-consuming and meant that the caption could not be included. Secondly screenshots came to mind but the problem with a PC computer is that it can only take screenshots of the entire screen and not just an individual picture. But still the latter would be a more efficient way. To solve this problem a program was used which enabled a screenshot only of the wanted part of the screen and then saved it on disc.

Furthermore, a simple macro was created to automate the process of inserting the images from disc to an excel spreadsheet.

*Table 2. Collected data for each lexeme.*

<table>
<thead>
<tr>
<th>Lexeme</th>
<th>Photo content</th>
<th>Faces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happiness</td>
<td>57</td>
<td>50</td>
</tr>
<tr>
<td>Blissful</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>Joyful</td>
<td>54</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 2 shows the resulting amount of 313 pictures from Google Images.
2.2. Coding

Once all the data had been cleaned the task of creating a coding scheme began. The coding scheme contains various variables. These variables are determined by looking at the information found in the examples and images.

For every variable there are features that when combined can show patterns in the usage of language (Glynn 2010:7). The actual coding was the most time-consuming task of this research.

2.3 Statistical Program

After all the examples have been coded the statistical program R console was used. More specifically two models were used for this paper; multiple correspondence analysis (MCA) and Logistic Regression (LogReg). MCA is a statistical tool that looks at the features and finds patterns and tendencies. When a tendency is detected it is converted into a distance, where all the features are relative to each other.

The second tool is LogReg which calculates the reliability of the MCA results. In other words it can prove that the results are not chance.

3. Analysis

Due to the fact that this paper examines both written and visual language two different coding schemes were created; one to accommodate the images and one for the words. This would later cause problems in the analysis because the two coding schemes could not be combined.

As mentioned in the previous section a coding scheme is created from looking at the information given in the examples and images.

3.1 Coding words

3.1.1 Lexeme

For this analysis the lexemes examined were happiness, blissful and joyful. There were 102 examples of happiness, 64 of blissful and 74 of joyful represented in this analysis.
3.1.2. Nationality

The first step was to distinguish the dialects American and British, these were coded under the variable Nationality and specified with the features UK and US. In this variable 110 examples were from UK blog entries and 130 from US entries.

3.1.3. Class

Due to the fact that the lexemes are both nouns and adjectives the variable Class was added to distinguish between them. This variable was used to determine the grammatical usage of the lexemes.

3.1.4. Main verb

In this column, the main verb used in association with the lexeme was coded. Here the most represented verb was , be with 95 entries, followed by have (17), make (8) and find (7). This makes sense when according to the OED happiness is “the state of being happy” and “she struggled to find happiness in her life”\(^2\).

3.1.5. Tense

Tense can give information on how the three lexemes are used. It can show if one lexeme is more associated with the past, present or future. For instance it can illustrate that happiness is something not yet achieved or maybe something in the past.

3.1.6. Possessive

This variable shows whether or not the lexeme is used together with a possessive pronoun.

3.1.7 Syntax

Under this category, the syntactic structure was coded. This means that the actual features were the collocations and the lexeme structure.

3.1.8. Subjecthood

Subjecthood indicates the position of the lexeme. In this category the most reoccurring position was Direct Object, with 196 entries.

3.1.9. Polarity

The lexemes polarity is coded with the features positive <pol_pos> or negative <pol_neg>. If the sentence is negated then the polarity would be negative, all the rest would be coded as positive.

3.1.10. Axiology

Axiology was coded to determine whether the lexemes were used in a positive or negative way. If it was neither the variable was coded as neutral <neut>. This variable was somewhat hard to code due to the fact that the lexemes are in themselves positive words. Consequently, the distinction of an intense feeling of happiness was coded as positive (See. 1.a) and a mere statement of the lexeme was then coded as neutral (See.1.b).
(1) a. I could feel the hope and happiness in my chest begin to well up as he just smiled at me and nodded his head.

b. Happiness is someone drawing smiley faces on his toes to make you laugh.

3.1.11. Humour

To indicate whether or not the lexemes were used with humour <hum> or without <nonhum> this variable was added. It was created to try and understand the mood of the way the lexemes were used. If the lexeme was used in a sarcastic or humorous way it was coded as humoristic, all the rest was coded as non humoristic.

3.1.12 Emphatic

This variable indicates whether or not the lexemes are used with emphasis or not.

3.1.13. Cause coarse-grained

This variable was the hardest but most important to code. Cause was divided into two variables in order to get a more coarse-grained (Cc.) and fine-grained (Cf.) analysis. The cause coarse-grained variable was divided into the following features. Furthermore, if the cause could not be determined, it was coded as non applicable (NA).

- Cc.hum_spec
- Cc.event
- Cc.thing_concrete
- Cc.activity
- Cc.thing_abstract
- Cc.state_of_affairs
- Cc.NA
The most frequent cause was a specific human <Cc.hum_spec> with an occurrence number of 89, followed by an event <Cc.event> and a concrete thing <Cc.thing_concrete>. See section 3.1.14 with examples for coding Cause.

3.1.14. Cause fine-grained

It is very important to not have too many features in one variable. If too many features were analyzed in the statistical program the results would be affected in a negative way. This is why the cause was divided into two categories, one that was more detailed than the other. To begin with this variable had too many features; this meant that this category had to be coded again with even fewer features. These are the final features that were used for this variable.

- Cf.another_person
- Cf.everyday_life
- Cf.spatial
- Cf.material_things
- Cf.personal_insight
- Cf.food
- Cf.career
- Cf.celebration
- Cf.NA

The cause variables were one of the most difficult variables to code. Following examples demonstrate the coding process.
1) a. *Guys aren’t suppose to make you feel this way...like you need them to be happy. And that they could take away all of your happiness in an instant. My heart honestly feels like its being torn up right this second.*

b. *First off my title is about New Year’s Day and how unbelievably great that day was just because I was pretty darn happy and blissful and oblivious.*

Example 1a was coded as <Cc.hum_spec> and <Cf.another_person> because the happiness is taken away when “Guys”, or in other words, a person leaves. Therefore the happiness is caused by another person. As for example 1b, <Cc.event> and <Cf.celebration> was coded. The cause of the lexeme blissful in this example, is the event of a New Years Eve celebration.

3.1.15. Theme coarse-grained

Theme was added to capture the bigger picture of the cause. For example, if the cause of the happiness is the activity of praying, then the theme of the cause would be coded as <Tc.spiritual>.

This variable was also divided into two separate variables, one coarse-grained and one fine-grained. Following features were coded.

*Table 3. Features for Theme coarse-grained*

<table>
<thead>
<tr>
<th>Theme coarse-grained</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tc.career</td>
<td>Tc.personal_insight</td>
</tr>
<tr>
<td>Tc.celebration</td>
<td>Tc.possessions</td>
</tr>
<tr>
<td>Tc.entertainment</td>
<td>Tc.relations</td>
</tr>
<tr>
<td>Tc.food</td>
<td>Tc.spiritual</td>
</tr>
</tbody>
</table>

The feature <Tc.personal_insight> stands for a realisation of some sort which brings happiness. <Tc.entertainment> was coded for such activities
as going to the movies, listening to music or playing computer games. \textit{<Tc.relations>} was used when referring to a relation to another person. In the section below a more fine-grained feature specifies what kind of relation it is. The most common feature was \textit{<Tc.relations>} with 93 occurrences, followed by \textit{<Tc.spiritual>} and \textit{<Tc.entertainment>}.

### 3.1.16. Theme fine-grained

The more detailed features of theme were coded under this variable. The table below shows the different features used.

<table>
<thead>
<tr>
<th>Theme fine-grained</th>
<th>T.xactivity</th>
<th>T.xmusic</th>
</tr>
</thead>
<tbody>
<tr>
<td>T.xburden</td>
<td>T.xnonsexual</td>
<td></td>
</tr>
<tr>
<td>T.xcareer</td>
<td>T.xsex</td>
<td></td>
</tr>
<tr>
<td>T.xcelebration</td>
<td>T.xsexual</td>
<td></td>
</tr>
<tr>
<td>T.xdrugs</td>
<td>T.xspiritual</td>
<td></td>
</tr>
<tr>
<td>T.xfood</td>
<td>T.xNA</td>
<td></td>
</tr>
</tbody>
</table>

As mentioned in section 3.1.13, when \textit{<Tc.relations>} is coded the fine-grained feature will clarify what the relation entails. \textit{<Tf.nonsexual>} is used for platonic love such as friendship and relations between family members etc. When coding the love between partners, \textit{<Tf-sexual>} was used.

Finally, for sexual relations where no “feelings” were involved, the features used was \textit{<Tf.sex>}. \textit{<Tf.burden>} was coded for the feeling of having \textit{happiness} but then also fearing to lose it. This meant that actually being e.g. \textit{blissful or joyful} was more a burden because of the fear of losing it. Also unanswered love was coded under this feature, because the
happiness the other person gave was also a burden because the feelings were not mutual.

3.1.17 Duration

This variable was coded to see if any of the lexemes were more or less used for something long-term or short-term.

3.1.18 Dictionary

Under the variable Dictionary, synonyms of the lexemes were used as features.

3.1.19 Referent 1

To begin with this variable was divided into two separate variables, one that coded the noun happiness and one for the adjectives blissful and joyful. This later gave many non applicable (NA) features which affected the statistical results. The choice was then made to try and combine the separate variables and make them into one.

The following features were coded in this variable.

Table 5. Referent 1 features

<table>
<thead>
<tr>
<th>Referent 1</th>
<th>Referent 1 features</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1.activity</td>
<td>R1.state_of_affairs</td>
</tr>
<tr>
<td>R1.event</td>
<td>R1.things_abstract</td>
</tr>
<tr>
<td>R1.hum_gen</td>
<td>R1.things_concrete</td>
</tr>
<tr>
<td>R1.hum_spec</td>
<td>R1.NA</td>
</tr>
</tbody>
</table>
3.1.20 Referent 2

As for Referent 1, Referent 2 also had to be combined. The features were not as many as with Referent 1, namely <R2.hum_gen>, <R2.hum_spec>, <R2.thing_abstract> and <R2.thing_concrete>.

3.1.21. Referent 1 Gender

In this column the gender of the Referent 1 was specified by looking at the information in the examples or by identifying the name of the writer. If the gender could not be determined it was coded as non applicable <RG1.NA>. There were 19 females and 17 males in this variable, the rest were non applicable.

3.1.22. Referent 2 Gender

Same as above, see 3.1.21. There were 72 females and 36 male, the rest were non applicable <RG2.NA>.

3.2 Coding images

The variables were determined by looking at the images and simply coding what was seen.

3.2.1 Type

All images were found on Google Images and with the help of the advanced image search two types of pictures were chosen. These two types are specified under this variable as <photo_content> and <faces>. See section 2.1.2 for the detailed information on the amount of each type.

3.2.2 Lexeme

As with the blog entries, the lexemes happiness, blissful and joyful were analyzed.
3.2.3. Object animacy

The object in focus was coded in this variable, with the distinction between animate and inanimate.

Table 6. features for Object Animacy

<table>
<thead>
<tr>
<th>Object_Animacy</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OA.animate_adult</td>
<td>OA.animate_thing</td>
</tr>
<tr>
<td>OA.animate_adult_child</td>
<td>OA.inanimate_image</td>
</tr>
<tr>
<td>OA.animate_child</td>
<td>OA.inanimate_roman_couple</td>
</tr>
<tr>
<td>OA.animate_roman_couple</td>
<td>OA.scenery</td>
</tr>
<tr>
<td>OA.animate_platonic_couple</td>
<td></td>
</tr>
</tbody>
</table>

See pictures 1 and 2 for examples of coding.

Picture 1. <OA.animate_adult_child>

Picture 2. <OA.scenery>
3.2.4. Object thing type

Here an even more detailed coding of the focused object is added. The most frequent feature was <OT.human> followed by <OT.nature> and <OT.thing>.

Table 7. features for Object thing type

<table>
<thead>
<tr>
<th>Object_Thing_Type</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OT.activity</td>
<td>OT.religious</td>
</tr>
<tr>
<td>OT.animal</td>
<td>OT.thing</td>
</tr>
<tr>
<td>OT.human</td>
<td>OT.NA</td>
</tr>
<tr>
<td>OT.nature</td>
<td></td>
</tr>
</tbody>
</table>

3.2.5 Object scene

This variable was added to capture the background. The features used were; <OS.city>, <OS.indoors>, <OS.nature>, <OS.nobackground>, <OS.patterns>, <OS.people>, <OS.seaside> and finally, <OS.sky>.

3.2.6. Object focus

To determine whether or not the object it looking in the camera, this variable was added. If the focus is straight at the camera it could be interpreted as the viewer being a part of the happiness, opposed to the object facing away from the camera. The feature <OF.camera_regard> was used when the object was looking at the camera, <OF.no_camera_regard> for the opposite. If the faces were not clear or the object was not a human, a flower for example, this was coded as <OF.NA>. 
3.2.7. Object expression

In this variable the aim was to capture the different expressions associated with the lexemes. They were coded accordingly to the following features; <OE.mouth_closed>, <OE.mouth.open>, <OE.smile_teeth>, <OE.smile_no_teeth>, and finally if the face could not be seen, <OE.NA> was used.

3.2.8. Incidental features

After determining the focused object and its background, the other parts of the picture had to be coded. Objects or actions that may not be the main focus but still just as important to understand the whole picture were specified. In this variable there were many features, but the most frequent were <IF.hand_gesture> and <IF.activity>, followed by <IF.material_thing>, <IF.embrace> and <IF.cross_image>. See pictures (3) and (4) for clarity on <IF.hand_gesture> and <IF.cross_image>.

Picture 3. <IF.hand_gesture> Picture 4. <IF.cross_image>

Picture (3) demonstrates one of the types of images that would be coded as <IF.hand_gesture>. Hand gestures such as hand in hair, kissing ones hand to blow a kiss and hand to face would fall under this feature.

Picture (4) demonstrates the reoccurring cross image, this one in a jumping motion. But there are also standing and lying cross images.
3.2.9. Gender

Every focused object was specified by gender. There were 138 females and 46 males coded. Furthermore, there were images of couples and families which were coded as <G.mix.>. The young children were hard to distinguish and were coded with <G.unknown>, the rest were <G.NA>.

3.2.10. Caption

To every image there is an underlining text, here the keyword of that caption was coded.

3.2.11. Dominant colour

The dominant colour of the image was coded in this variable. The most frequent was <blue>, followed by <white> and <green>.

3.2.12. Tone

The tone of the pictures were coded with the features <bright>, <neutral> and <dull>. 
4. Results

The coded words and images were inserted into a statistical programme. More specifically, a Correspondence Analysis designed to find patterns in the usage of language, and Logistic Regression; used to determine the accuracy of the results.

In a Correspondence Analyses the features cluster together if they have similar patterns or meaning. The farther away they are from each other the lesser they have in common.

At this point in the research some problems occurred. As mentioned before the images and words were coded with two separate coding schemes. This blunder meant that they had to be analysed separately in the statistical programmes.

4.1 Words

4.1.1. Cause

*Figure 1. Correspondence analysis of Lexeme and Cause*

Figure (1) shows a clear difference in the usage of the three lexemes. *Joyful*, situated at the far bottom of the plot, is clustered together with
<Cc.thing_abstract>, <Cc.state_of_affairs> and <Cc.event>. This indicates that the closest association of joyful are abstract things. Furthermore, by looking at the blog entries this lexeme tends to be used with reference to spiritual and religious things. Example (3) demonstrates the type of semantic use of joyful.

3) Joyful and cause <Cc.thing_abstract>,
   a. Don’t forget that great blessing that you have. I promise that it will keep you close to the Lord. You are my light and my love. I can not wait until that joyful day when we are again reunited there. May we then never again be separated. In the temple today we did endowments and some initiatory work. How beautiful are the blessing there! How wondrous is God’s love!

The lexeme blissful is clustered together with <Cc.activity>, showing that when speaking of something blissful it is more likely to be an activity.

Happiness, on the other hand, is closely associated with <Cc.hum_spec> and <Cc.thing_concrete>. This indicates that when happiness is used it is used mainly to talk about material things and relations to people, rather than e.g. abstract things. See example (4)

4) Happiness and cause
   a. <Cc.hum_spec>
      I could feel the hope and happiness in my chest begin to well up as he just smiled at me and nodded his head
   b. <Cc.thing_concrete>
      I've lived with my 32” tv and 24” lcd monitor in happiness and isolation too long

4.1.2. Duration and Tense

In Figure (2) the analysis indicates that there are tendencies in how the three lexemes are used in relation to duration and tense. <long_term> and <present> is clustered together with happiness. This could be explained by the fact that the cause of the happiness is either a relation to another person or a material thing. Both these causes are things enjoyed in the present, but at the same time there is no real “end” to the happiness. For example,
buying a material thing will bring happiness, but the reason of buying e.g. a new computer is not to get rid of it and buy a new one the next day, but to keep it for a longer period of time. The same goes for relationships, there is happiness being with the person you love, but the thought that it will end doesn’t really exist. This indicates that the happiness is from now on and forward, i.e. long term.

Figure 2 Correspondence Analysis of Lexeme, Duration and Tense.

<short_term> is clustered in between both blissful and joyful, which means that this feature is equally significant for both lexemes.

As mentioned in the previous section, the cause of something blissful is an activity. This could be an explanation to why blissful is clustered together with <past> and <short_term>. See example (5) below.

5) Blissful and cause  
   a. <Cc.activity>
      Saturday: I was in perfect blissful sleep then all of a sudden I heard girls shuffling their feet talking giggling.

Example (5) demonstrated that there is a clear start and end to the activity. The start of the activity is falling asleep and the end is waking up, and in
between was the “blissful sleep”. Hence, the activity is short term and takes place in the past. Finally, joyful is also clustered with <short_term> and <future>. This can also be explained with the cause; <Cc.thing_abstract>. As mentioned in 4.1.1, the cause for joyful are spiritual or religious things. The future tense can be explained in example (3).

4.2 Images

This analysis shows the images associated with the three lexemes.

4.2.1. Object thing

![Correspondence Analysis of Lexeme and Object thing](image.png)

*Figure 3. Correspondence Analysis of Lexeme and Object thing*

The objects most associated with *happiness* are <OT.human>. This is interesting because the closest associated cause of happiness with the words was also human <Cc.hum_spec>. The lexeme *blissful* is clustered together with <OT.nature> and <OT.thing>. See pictures (5) and (6).
These pictures show the type of objects associated with *blissful*. The pictures tend to show a kind of calmness and serenity.

*Joyful* is associated with `<OT.animal>` and, as with the words, with religious things. But the most interesting thing is that as with the words neither *blissful* nor *joyful* are closely associated with humans.

### 4.2.2. Incidental features

The incidental features were analysed to try and capture the objects not in focus, but equally as important.

*Figure 4. Correspondence Analysis of Lexeme and Incidental features*
In figure (4) <IF.activity> is associated with happiness whereas for the words activity was something closely related to blissful. (See section 4.1.1.). This is interesting because the word blissful is related to <Cc.activity> whereas the word happiness is clustered with <Cc.hum_spec>. 

To try and understand this, pictures (7) and (8) show that it is not the activity itself that is “happy”, but the human. As a result of the happiness they jump or run, it’s not the jumping that causes the happiness. Joyful is again clustered close to something religious, here <IF.cross_image>. See image (9), (10) and (11). It is even clustered with <IF.hand_gesture>, which also has some religious tendencies. See image (12).
As in the section 4.2.1 *blissful* was related to nature, but in this analysis also to *<IF.eye_movement>* and *<IF.embrace>*. The eye movement was usually referred to as closed eyes or looking down, here again giving a picture of calmness and meditation. See image (13)
4.2.3. Object expression

This analysis tries to capture the different expressions associated with the lexemes.

![Figure 5. Correspondence Analysis of Lexeme and Incidental features](image)

Figure (5) shows *joyful* being clustered together with <OE.mouth_open>, indicating that someone is singing or making some kind of a noise. This makes sense when looking at some of the word examples. See example (6)

6)  
   a. *Twelve people should be able to make a joyful noise.*
   b. *Go ahead and make a joyful sound to the Lord in the privacy of your own home behind closed doors and windows.*
   c. *For some reason Christmas doesn’t really bring about as much joyful cheer and anticipation as it used to when I was younger.*

*Blissful* is associated with <IF.mouth_closed> and <IF.smile_no_teeth>. Put these features together with <OE.eye_movement> and the image of calmness comes to mind. See image (13).

The feature <smile_teeth> occurs mostly with *happiness*. See images (14), (15) and (16)
4.2.4 Gender and object animacy

This Correspondence Analysis was done to see what gender was most associated with people.

As mentioned in sections 3.1.21-22 and 3.2.9 the majority of the coded gender has been female. But there are a few males, which, according to figure (6), are closely clustered to children <OA.animate_child>. This means that when thinking of these lexemes, the mental image tends to be mostly of women and children, not adult men.
4.3 Logistic Regression

Logistic Regression is a tool used to calculate the reliability of the results in the Correspondence Analysis (Glynn 2010:11-13). The first LogReg tool is GLM, this determines the statistical significance i.e. proves that the results are not chance. The second LogReg tool is the LRM, which calculates the accuracy of the results.

4.3.1 Logistic Regression Words

Logistic Regression: Lexeme, Duration and Theme fine grained.

Coefficients:

| Term               | Estimate | Std. Error | z value | Pr(>|z|) |
|--------------------|----------|------------|---------|----------|
| (Intercept)        | -2.0144  | 0.6968     | -2.891  | 0.00384 ** |
| Durationshort_term | 1.6576   | 0.4636     | 3.575   | 0.00035 *** |
| Theme_Tf.burden    | -1.0075  | 1.2399     | -0.813  | 0.41646   |
| Theme_Tf.career    | -1.2956  | 1.2296     | -1.054  | 0.29205   |
| Theme_Tf.celebration| 1.9197  | 0.8527     | 2.251   | 0.02436 * |
| Theme_Tf.drugs     | -17.2093 | 1495.2957  | -0.012  | 0.99082   |
| Theme_Tf.food      | -17.2093 | 2284.1018  | -0.008  | 0.99399   |
| Theme_Tf.music     | 1.1677   | 0.8305     | 1.406   | 0.15972   |
| Theme_Tf.nonsexual | 1.6091   | 0.7908     | 2.035   | 0.04186 * |
| Theme_Tf.career    | 0.1213   | 0.7855     | 0.154   | 0.87730   |
| Theme_Tf.sex       | 1.0499   | 1.3523     | 0.776   | 0.43751   |
| Theme_Tf.sexual    | 0.3868   | 0.7923     | 0.488   | 0.62539   |
| Theme_Tf.spiritual | 1.6309   | 0.7264     | 2.245   | 0.02475 * |

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

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 237.32 on 173 degrees of freedom
Residual deviance: 189.03 on 161 degrees of freedom
AIC: 215.03

Number of Fisher Scoring iterations: 16

In this LogReg the lexemes *happiness* and *joyful* were compared. The added features were calculated to be significant to one of the lexemes. If significance is detected a star will appear. The Estimate is the predictor which means that the higher the number the better results. In this column, the negative numbers specify the lexeme first in alphabetic order and positive being the second. This means that in this case the negative stands for *happiness* and the positive represent *joyful*. 
In the column to the left, all the features analysed in the LogReg are listed. On the far right the P-value indicates the significance, the lower the number the stronger significance. In the results above four significances were found, long term and short term duration, as seen in figure (2) and the themes; celebration, nonsexual and spiritual. Duration <long_term> ( Intercept) was significant for the lexeme happiness with a P-value of 0.00384, which means that if this analysis were to be done again, there is a 99.616% chance that the results would be the same. For this variable the Estimate is 2.0144, which is a rather high predictor. Furthermore, <short_term> was highly significant for the lexeme joyful, with a P-value of 0.00035 and a fairly high predictor. The themes; celebration, nonsexual and spiritual, were all significant with joyful.

Frequencies of Responses
happiness   joyful
100     74

<table>
<thead>
<tr>
<th>Obs</th>
<th>Max Deriv Model L.R.</th>
<th>d.f.</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>174</td>
<td>0.003</td>
<td>48.28</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

Dxy Gamma Tau-a R2 Brier
0.551 0.592 0.271 0.326 0.183

<table>
<thead>
<tr>
<th>Coef</th>
<th>S.E.</th>
<th>Wald Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.0144</td>
<td>0.6968</td>
<td>-2.89</td>
</tr>
<tr>
<td>Duration=short_term</td>
<td>1.6576</td>
<td>0.4636</td>
<td>3.58</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.burden</td>
<td>-1.0075</td>
<td>1.2399</td>
<td>-0.81</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.carer</td>
<td>-1.2956</td>
<td>1.2296</td>
<td>-1.05</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.celebration</td>
<td>1.9197</td>
<td>0.8527</td>
<td>2.25</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.drugs</td>
<td>-7.8790</td>
<td>23.2317</td>
<td>-0.34</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.food</td>
<td>-7.8790</td>
<td>35.4808</td>
<td>-0.22</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.music</td>
<td>1.1677</td>
<td>0.8305</td>
<td>1.41</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.nonsexual</td>
<td>1.6091</td>
<td>0.7908</td>
<td>2.03</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.possessions</td>
<td>0.1213</td>
<td>0.7855</td>
<td>0.15</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.sex</td>
<td>1.0499</td>
<td>1.3523</td>
<td>0.78</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.sexual</td>
<td>0.3868</td>
<td>0.7923</td>
<td>0.49</td>
</tr>
<tr>
<td>Theme_fine_grain=Tf.spiritual</td>
<td>1.6309</td>
<td>0.7264</td>
<td>2.25</td>
</tr>
</tbody>
</table>

The accuracy results from the LogReg LRM are listed above. The C-value indicates how many times this analysis is accurate. The wanted value is 0.8 but this result is just under with 0.774. As for the R2-value anything above 0.3 is a good indicator, in this analyses the value is 0.326.
4.3.2 Logistic Regression Images

In this LogReg, the response lexemes were blissful and joyful.

Logistic Regression: Lexeme, Object expression and Incidental features.

```
Deviance Residuals:
     Min       1Q     Median       3Q      Max
-1.80730 -0.67964  0.00021  0.75668  2.39924

Coefficient

                   Estimate Std. Error z value Pr(>|z|)
(Intercept)       -1.9161     1.0423  -1.838  0.06602 .
Object_ExOE.mouth_open  3.3319     1.2038   2.768  0.00564 **
Object_ExOE.NA       1.1929     1.0246   1.164  0.24432
Object_ExOE.smile_no_teeth  1.0760    0.3500   3.005  0.00271 *
Object_ExOE.smile_teeth  2.6156     1.0201   2.564  0.01035 *
Incidental_IF.animal  18.7410    2745.2916   0.007  0.99455
Incidental_IF.cross_image  2.8879     1.2384   2.332  0.01971 *
Incidental_IF.embrace  -2.0971     1.2815  -1.637  0.10173
Incidental_IF.eye_movement  1.0958     0.1753   6.282  3.9e-10 ***
Incidental_IF.hand_gesture  0.7525     0.1486   5.079  3.9e-07 ***
Incidental_IF.material_thin  0.7761     0.5211   1.489  0.13643
Incidental_IF.nature    0.0351     0.8688    0.040  0.96777

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

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 162.12  on 116  degrees of freedom
    Residual deviance: 110.61  on 104  degrees of freedom
    AIC: 136.61

Number of Fisher Scoring iterations: 16
```

As explained in 4.3.1, the stars indicate significance, more specifically <OE.mouth_open>, <smile_teeth> and <IF.cross_image> are significant to the lexeme joyful and <OE.mouth_closed> (Intersect) is significant to blissful.
Frequencies of Responses

<table>
<thead>
<tr>
<th>blissful</th>
<th>joyful</th>
</tr>
</thead>
<tbody>
<tr>
<td>57</td>
<td>60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Obs</th>
<th>Max Deriv Model L.R.</th>
<th>d.f.</th>
<th>P</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>117</td>
<td>0.002</td>
<td>51.5</td>
<td>12</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dxy</th>
<th>Gamma</th>
<th>Tau-a</th>
<th>R2</th>
<th>Brier</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.675</td>
<td>0.697</td>
<td>0.34</td>
<td>0.475</td>
<td>0.155</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coef</th>
<th>S.E.</th>
<th>Wald Z</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-1.916</td>
<td>1.042</td>
<td>-1.84</td>
</tr>
<tr>
<td>Object_Expression=OE.mouth_moving</td>
<td>10.0529</td>
<td>27.0387</td>
<td>0.37</td>
</tr>
<tr>
<td>Object_Expression=OE.mouth_open</td>
<td>3.3319</td>
<td>1.2038</td>
<td>2.77</td>
</tr>
<tr>
<td>Object_Expression=OE.NA</td>
<td>1.1929</td>
<td>1.0246</td>
<td>1.16</td>
</tr>
<tr>
<td>Object_Expression=OE.smile_no_teeth</td>
<td>0.3763</td>
<td>1.0760</td>
<td>0.35</td>
</tr>
<tr>
<td>Object_Expression=OE.smile_teeth</td>
<td>2.6156</td>
<td>1.0201</td>
<td>2.56</td>
</tr>
<tr>
<td>Incidental_features=IF.animal</td>
<td>9.3908</td>
<td>42.2227</td>
<td>0.22</td>
</tr>
<tr>
<td>Incidental_features=IF.cross_image</td>
<td>2.8879</td>
<td>1.2384</td>
<td>2.33</td>
</tr>
<tr>
<td>Incidental_features=IF.embrace</td>
<td>-2.0971</td>
<td>1.2814</td>
<td>-1.64</td>
</tr>
<tr>
<td>Incidental_features=IF.eye_movement</td>
<td>0.1919</td>
<td>1.0957</td>
<td>0.18</td>
</tr>
<tr>
<td>Incidental_features=IF.hand_gesture</td>
<td>0.1115</td>
<td>0.7525</td>
<td>0.15</td>
</tr>
<tr>
<td>Incidental_features=IF.material_thing</td>
<td>0.4047</td>
<td>0.7760</td>
<td>0.52</td>
</tr>
<tr>
<td>Incidental_features=IF.nature</td>
<td>0.0351</td>
<td>0.8687</td>
<td>0.04</td>
</tr>
</tbody>
</table>

The C-value, 0.837, and R2-value, 0.475, are both above the required amount, as mentioned in section 4.3.1.

But because there are some very high numbers on the Error-value on GLM in both the words and images, there might be problems with this result. Because two separate coding schemes were made the analysis had to be done separately. This meant that there was only half the data used for the words and images. Hence, these high Error-values can be a result of too little data.
5. Summary

This analysis has shown that the tree lexemes *happiness*, *blissful* and *joyful* are used for different concepts of happiness. *Happiness* is used mostly in association with other people, more specifically in different types of relationships. There are also indications that when speaking of *happiness* there is an attitude that the feeling is long term, or at least hoping for it to be long term. Usually relationships such as friendship and family do tend to be lifelong. The analysis also shows that the image of someone being happy is a person, often a woman or child, smiling with their teeth and often doing an activity. The activity is most likely jumping or running, implying that when being happy, it is hard to stand still.

Opposed to *happiness*, *blissful* is more a satisfied and calm state. There are also indications that *blissful* is closely associated with nature and activities that took place in the past.

Finally, the images and the words seem to agree that *joyful* is used when referring to something religious or spiritual. It is also often associated with singing and sounds.

As for gender, adult men are not closely associated with any of the lexemes; in fact they are not associated at all. The only males this analysis has shown have been either infants or children.

This paper has tried to examine the concept of happiness but due to some problems and time issues the analysis did not quite go as hoped. There is much more to be investigated with the concept of *happiness*, but to do so more data is needed. In order to get a clearer picture of the concept of both images and words a combined coding scheme would be optimal.
Cognitive Linguistics: The Conceptualisation of Happiness

References

Glynn, Dylan
2010 Corpus-driven Cognitive linguistics. A case study in Polysemy. Lund University: 2,7

Saussure, Ferdinand de
