# Towards an Empirically Grounded Framework for Emotion Analysis

Luna De Bruyne, Orphée De Clercq and Véronique Hoste

LT[3], Language and Translation Technology Team
Ghent University
Groot-Brittanniëlaan 45, 9000 Ghent, Belgium
Email: {luna.debruyne, orphee.declercq, veronique.hoste}@ugent.be

***Abstract*—The first step in training a system for automatic emotion detection consists of manual data annotation. Because there is no consensus on a standard emotion framework, we established a label set which is justified both theoretically and practically. Frequency and cluster analysis of 229 tweet annotations resulted in a label set containing the 5 emotions *Love*, *Joy*, *Anger*, *Nervousness* and *Sadness*. Our label set shows fair resemblance to Ekman's basic emotions, but due to our data-driven approach, our label set is much more grounded in the task (emotion detection) and the domain (Dutch tweets).**

***Keywords*–Emotion Detection; NLP; Emotion Annotation.**

## I. INTRODUCTION

Emotions play a central role in how we perceive the world and how we communicate with it, making them a prominent object of study in many research fields, including psychology [1], linguistics [2], and neurobiology [3]. In Natural Language Processing (NLP), emotion analysis has attracted interest the last decades because of its myriad of applications, including market analysis and customer satisfaction for business intelligence [4], educational and pedagogical applications [5], analysing political tweets and public sentiment [6], crisis communication [7], and mental health applications [8].

Our overall objective is to create an automatic emotion detection system for Dutch. As emotion detection has mainly been studied for English data [9], [10], [11], but only to a limited extent for some other languages, our first step will consist in manually labeling Dutch textual data with emotions following a certain framework. Notwithstanding the long history of theoretical emotion research in psychology and its recent surge in NLP, there is currently no consensus on a standard emotion framework. Categorical models, mostly offering a set of basic emotions, and dimensional representations (see Section II) coexist, and even within those models, different sets of basic emotions and dimensions can be found. This wide spectrum of frameworks impedes the exchange of data and knowledge resources (e.g., annotated datasets and emotion lexicons), which are crucial to train supervised machine learning approaches for emotion detection, and makes it difficult to compare different NLP systems handling emotions. Moreover, the motives on which a particular emotion framework is selected in studies on automatic emotion detection are often unclear, and the frameworks seem to be chosen rather arbitrary.

In this paper, we wish to establish an emotion framework which is justified both theoretically and practically to perform automatic emotion detection on Dutch tweets. To this end, we start from theories about emotion in psychology (more specifically, the work of Shaver et al. [12]), but, contrary to research in the psychological tradition, we work with real-life data instead of words in isolation. This real-life data comes from the same distribution as the data on which we will perform the ultimate task of emotion detection. This ensures that our framework is empirically grounded, which would not be the case if we arbitrary adopted a framework from psychological emotion theory.

To this purpose, we collected a large dataset of Dutch tweets comprising at least one emoji, and we annotated 300 tweets by labeling all possible emotion categories as conveyed by the author of the tweet. First, we performed an Inter-Annotator Agreement (IAA) study by having a small subset of the data (50 tweets) annotated by three different annotators, and we found that for most categories a moderate to substantial agreement could be observed. Subsequently, the remaining 250 tweets were annotated and a cluster analysis was performed. This leads to a reduced, empirically grounded emotion framework consisting of the 5 emotions *Love*, *Joy*, *Anger*, *Nervousness* and *Sadness*.

The remainder of this paper is organized as follows: in Section II, we describe the related work. Section III presents the data collection, IAA study, annotations and explains the clustering technique that was used. In Section IV, we present the results and Section V concludes this paper.

## II. RELATED WORK

In emotion theory, two main approaches for emotional representation coexist, namely, (i) representation based on a categorical model and (ii) based on a dimensional model.

In the dimensional approach, emotions are seen as a vector in a multidimensional space, e.g., with the two dimensions *Valence* (from *Displeasure* to *Pleasure*) and *Arousal* (from *Calmness* to *Excitement*) [13], the three-dimensional *Valence-Arousal-Dominance* model [14] or even a four-dimensional model which also takes *Unpredictability* into account [15].

Categorical representation models, however, involve cognitive labeling of emotions, typically using a set of basic emotions, with the theories of Ekman [16] and Plutchik [17] being the most influential ones. Ekman reports that *Joy*, *Surprise*, *Anger*, *Fear*, *Disgust* and *Sadness* are the six most basic emotions and that these can be linked to universal facial expressions. Plutchik added *Trust* and *Anticipation* to Ekman's set, resulting in a set of eight emotions. Basic emotion frameworks have been provided by many other theorists, ranging from 2 to 14 emotions. Table I gives an overview of different basic emotion frameworks. This list is adapted from [18].

More extensive frameworks exist, but then the categories are not mere basic emotions. Often, they contain secondary emotions, which are more complex categories and can be seen

 

TABLE I. BASIC EMOTION FRAMEWORKS.

| | *Author* | *Basic Emotions* |
|---|---|---|
| [19] | Arnold (1960) | Anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, sadness |
| [16] | Ekman (1992) | Anger, disgust, fear, joy, sadness, surprise |
| [20] | Frijda (1986) | Desire, happiness, interest, surprise, wonder, sorrow |
| [21] | Gray (1982) | Rage, anxiety, joy |
| [22] | Izard (1971) | Anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise |
| [23] | James (1884) | Fear, grief, love, rage |
| [24] | McDougall (1926) | Anger, disgust, elation, fear, subjection, tenderness, wonder |
| [25] | Mowrer (1960) | Pain, pleasure |
| [26] | Oatley & Johnson-Laird (1987) | Anger, disgust, anxiety, happiness, sadness |
| [27] | Panksepp (1982) | Expectancy, fear, rage, panic |
| [17] | Plutchik (1980) | Acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise |
| [28] | Tomkins (1984) | Anger, interest, contempt, disgust, distress, fear, joy, shame, surprise |
| [29] | Watson (1930) | Fear, love, rage |
| [30] | Weiner & Graham (1984) | Happiness, sadness |
| [31] | Epstein (1984) | Fear, anger, sadness, joy, (love, affection) |
| [32] | Roseman (1984) | Surprise, hope, fear, joy, relief, sorrow, discomfort/disgust, frustration, liking, disliking, anger, pride, shame/guilt, regret |

TABLE II. EXTENSIVE EMOTION LISTS.

| | *Author* | *Emotion List* |
|---|---|---|
| [17] | Plutchik (1980) | Aggressiveness, anxiety, awe, contempt, curiosity, cynicism, delight, despair, disapproval, dominance, envy, guilt, hope, love, morbidness, optimism, outrage, pessimism, pride, remorse, sentimentality, shame, submission, unbelief |
| [13] | Russell (1980) | Afraid, alarmed, angry, annoyed, aroused, astonished, at ease, bored, calm, content, delighted, depressed, distressed, droopy, excited, frustrated, glad, gloomy, happy, miserable, pleased, relaxed, sad, satisfied, serene, sleepy, tense, tired |
| [12] | Shaver et al. (1987) | Affection, cheerfulness, contentment, disappointment, disgust, enthrallment, envy, exasperation, horror, irritability, longing, lust, neglect, nervousness, optimism, pride, rage, relief, sadness, shame, suffering, surprise, sympathy, torment, zest |
| [33] | Schröder et al. (2011) | Affection, amusement, anger, annoyance, anxiety, boredom, calmness, contempt, contentment, courage, delight, despair, disappointment, disgust, doubt, elation, embarrassment, empathy, envy, excitement, fear, friendliness, frustration, guilt, happiness, helplessness, hope, humility, hurt, interest, irritation, joy, love, pleasure, politeness, powerlessness, pride, relaxation, relief, sadness, satisfaction, serenity, shame, shock, stress, surprise, tension, trust, worry |

as combinations of basic emotions. The extended version of Plutchik's [17] emotion model, for example, counts thirty-two emotions, of which only eight are basic and the other twenty-four are secondary emotions (e.g., the secondary emotion *Optimism* is defined as the combination of the basic emotions *Optimism* and *Joy*). Also Russel [13] provides a list of some emotion terms, but these are rather stimuli or examples to illustrate the representation model and not basic emotions. Worth mentioning is also the emotion taxonomy proposed by Shaver et al. [12]. This taxonomy was obtained by means of a similarity-sorting task of 135 terms that were experimentally shown to be prototypical emotion words: 213 emotion words were rated by 112 psychology students for prototypicality, resulting in 135 prototypical emotion words. Then, 100 students grouped those emotion words into categories (without a predefined number of categories), which resulted in 100 135x135 co-occurrence matrices that were combined and used as input for the cluster analysis. This resulted in 25 categories, which were subsequently classified under six basic categories: *Love*, *Joy*, *Surprise*, *Anger*, *Sadness* and *Fear*. Table II presents an overview of these extensive emotion lists.

Regarding frameworks used in NLP, categorical frameworks are dominant, and both Ekman's and Plutchik's set of basic emotions are popular (see Table III for an overview of the most used English emotion datasets that use a categorical label set). Sometimes, another framework is chosen specifically based on the task/domain (e.g., [8] employ a set of 15 emotions to investigate signs of suicidal behavior). However, more often the motives on which a particular emotion framework is selected are unclear. [33] expressed the need of a standardized model for emotion detection tasks, resulting in the creation of the Emotion Annotation and Representation Language (EARL). Although this framework originates from the field of Affective Computing, its construction was not data-driven nor experimentally grounded. Moreover, we are not aware of any studies in NLP that make use of this framework.

TABLE III. EMOTION DATASETS.

| | *Dataset* | *Framework* |
|---|---|---|
| [10] | AffectiveText | Ekman |
| [11] | AffectInTweets T1, ST1-4 | Anger, fear, joy, sadness |
| [11] | AffectInTweets T1, ST5 | Plutchik + optimism, pessimism, love |
| [34] | Blogs | Ekman + no emotion + mixed emotion |
| | CrowdFlower | Ekman + enthusiasm, fun, hate, neutral, love, boredom, relief, empty |
| [35] | DailyDialogs | Ekman |
| [6] | Electoral-Tweets | Plutchik |
| [36] | EmoInt | Anger, fear, joy, sadness |
| [37] | Emotion-Stimulus | Ekman + shame |
| [38] | Grounded-Emotions | Happy, sad |
| [39] | ISEAR | E + shame, guilt |
| [40] | SSEC | Plutchik |
| [9] | Tales | Ekman (anger and disgust merged) |
| [41] | TEC | Ekman with posite and negative surprise |

## III. METHOD

In order to establish a framework for emotion detection that is justified both theoretically and practically, we collect a corpus of real-life data which we annotate with an initial extensive emotion label set. Then, we perform a cluster analysis to reveal which categories to merge into one category, resulting in a smaller, empirically grounded label set.

### A. Initial Labels

The initial label set needs to be sufficiently large to capture enough nuances between emotion categories. Most basic emotion sets are rather brief (the most popular ones, [16] and [17], contain only six and eight categories, respectively) and do not capture nuances like *Frustration* and *Envy*, which usually are subsumed under the *Anger* category. However, we think that such differentiations can be useful in certain domains. Moreover, much of the frameworks have a skewed distribution regarding sentiment polarity, with significantly more negative than positive emotions: for example, [22] has only two out of ten emotions that are unambiguously positive. In Ekman's set, only *Joy* is clearly positive, *Anger*, *Disgust* and *Fear* are negative, and *Surprise* can be either negative or positive.

Some frameworks also provide secondary emotions (e.g., [17]), which of course results in a larger emotion set. However,

Figure 1. Emoji's used as queries for collecting tweets.

using such a set of secondary emotions as input for cluster analysis to again obtain a smaller set of categories, seems like a circular approach and is therefore not an option.

Taking this into account, we chose the emotion taxonomy of [12] as a starting point (see Table II). The main advantage of using these 25 emotion words is that the label set is not biased: although one could argue that these are secondary emotion words, they are not deducted from basic emotion categories (on the contrary, it is the other way around). This is a big difference with a framework like Plutchik's [17], where a secondary emotion word is seen as a combination of two or more basic emotion words. Moreover, this label set is independent, and not chosen to fit a certain model (unlike for example [13], where the emotion words are not prototypical, but seem to be chosen to fit the pleasure-arousal circumplex model).

Although the set of 25 emotions was already further clustered into a final set of six basic emotions in the original work of [12], our cluster analysis is no repetition of this approach. While the clustering of [12] was based on the results of a similarity-sorting task of 135 words, our analysis is performed on annotations of real-life data. As this data is similar to the data on which we eventually want to perform emotion detection, our approach ensures that our label set is more grounded in the task of emotion detection and on the domain we are interested in, instead of merely adapted from the psychology field.

### B. Data

We wanted to obtain a collection of Dutch tweets that could be considered high in emotions. To increase the chance of scraping emotional tweets, we used a list of 72 emoji's (see Figure 1) as queries in the database for Dutch tweets Twiqs.nl [42]. We downloaded all tweets that the database returned from the year 2017 and took a random subset of 300 tweets. Before distributing the tweets over the annotators, we removed all duplicates and non-Dutch tweets and replaced them with another random tweet from our overall collection.

The tweets were annotated in a multi-label setup: for each of the 25 emotion words, the annotator needed to indicate whether the emotion is expressed (explicitly or implicitly) or not. Because we are interested in the emotional state of the author while writing the tweet, the annotator was asked to project oneself into the perspective of the tweet's author.

The annotation team consisted of three experienced linguists. In a first round, all three annotators labeled the same 50 tweets according to predefined guidelines. We determined inter-annotator agreement by calculating Cohen's Kappa [43] between each annotator pair and taking the mean of those two scores. IAA varied largely across emotion categories (see Table IV). For most categories, a moderate ($0.4 < \kappa < 0.6$) to substantial ($0.6 < \kappa < 0.8$) agreement can be observed. When leaving the emotions out of consideration for which at least one annotator never indicated it as present (6 categories), the mean Kappa score is 0.498 (moderate agreement). Mean Kappa score between Annotator 1 and 2 was 0.425; 0.465 between Annotator 2 and 3; and 0.533 between 1 and 3.

TABLE IV. IAA SCORES OF FIRST ANNOTATION ROUND

| *Emotion* | $\kappa$ | *Emotion* | $\kappa$ |
|---|---|---|---|
| Anger | 0,619 | Lust | 0,772 |
| Contentment | 0,525 | Nervousness | -0,014 |
| Disappointment | 0,418 | Optimism | 0,473 |
| Disgust | 0,635 | Pity | 0,13 |
| Enthrallment | 0,067 | Pride | 0,772 |
| Enthousiasm | 0,502 | Rejection | nan |
| Envy | nan | Relief | -0,009 |
| Fear | 0,772 | Remorse | -0,007 |
| Frustration | 0,593 | Sadness | 0,427 |
| Irritation | 0,53 | Suffering | 0,219 |
| Joy | 0,492 | Surprise | 0,551 |
| Longing | 0,32 | Torment | nan |
| Love | 0,36 | | |

In the second annotation round, the remaining 250 tweets were distributed among the three annotators. These annotated tweets were merged with Annotator 3's annotations of the tweets of the first round (because she had the highest agreement with both of the other annotators). Tweets for which not a single emotion was indicated as present (and thus were judged as objective), were left aside for further analysis. Our final dataset consists of 229 emotional tweets.

### C. Clustering

We regarded the annotations as vectors per emotion category, resulting in 25 229-dimensional vectors. We construct a 25x25 distance matrix by measuring the Dice dissimilarity [44] between each emotion vector pair. Dice is a common metric for assessing the (dis)similarity between boolean vectors, and contrary to the similar Jaccard metric, it gives a higher weight to double positives:

$$d(A,B) = 1 - \frac{2DP}{2DP+P_A+P_B}$$

with $DP$ the number of double positives (value of 1 in both emotion vector A and B), $P_A$ the number of positives (1-values) in emotion vector A and $P_B$ the number of positives in emotion vector B. Double negatives (value of 0 in both vectors) are not counted. This implies that vector pairs differing only in one value not always get the same distance score: the distance will decrease as more instances have a value of 1 in both vectors. Emotion pairs that were more frequently annotated as present will have a relatively smaller distance.

The resulting distance matrix was then used as input for a hierarchical cluster analysis. We tried seven different linkage methods: single (Nearest Point Algorithm), complete (Farthest Point Algorithm), average (UPGMA Algorithm), weighted (WPGMA), centroid (UPGMC), median (WPGMC) and Ward's linkage (Incremental Algorithm).

## IV. Results

### A. Frequency Analysis

Figure 2 shows that the top 3 most frequent emotions are all positive emotions (*Contentment*, *Joy* and *Enthusiasm*). The negative emotions *Irritation* and *Frustration* complete the top 5. Eight emotions (of which six are negative) appear less than 10 times in the dataset: *Suffering*, *Relief*, *Lust*, *Rejection*, *Envy*, *Fear*, *Remorse* and *Torment*. Interestingly, only three out of six Ekman emotions appear in the top 10 of most frequent emotions, namely *Joy*, *Sadness* and *Anger*. *Fear*, on the other hand, is even the third least frequent emotion in this dataset. This possibly indicates that popular basic emotion frameworks like Ekman's are not always the most suitable to apply in an NLP task like emotion detection, let alone for data coming from a specific domain or genre such as Twitter.

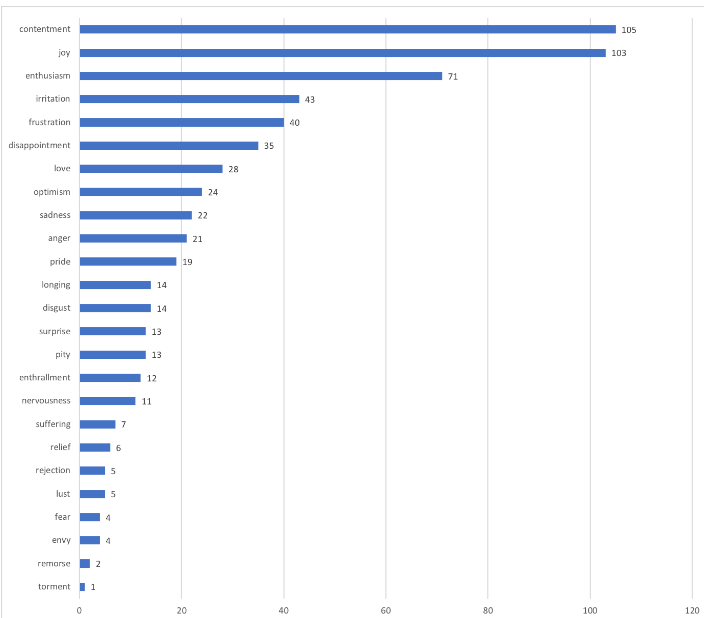


Figure 2. Frequencies of emotion categories.

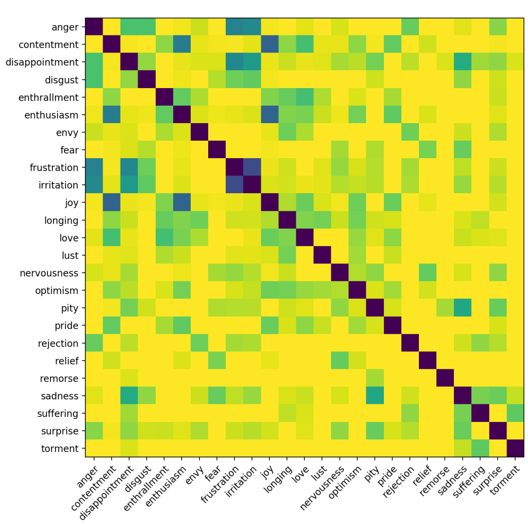


Figure 3. Distance matrix.

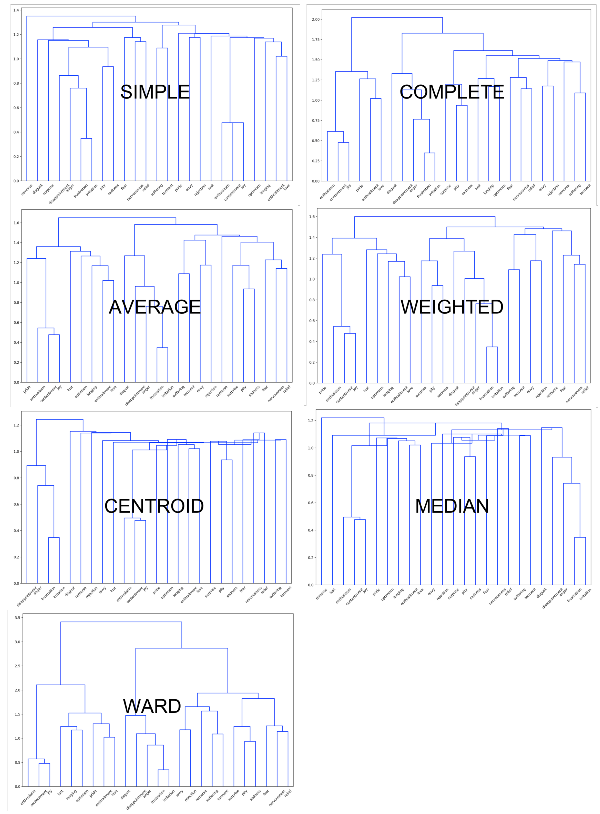


Figure 4. Dendrograms with different linking methods.

### B. Cluster Analysis

Figure 3 shows the distance matrix based on the Dice dissimilarity between each emotion vector pair. This already gives some insight in which emotion categories are more related to each other. *Frustration* and *Irritation*, for example, are the most salient in terms of similarity, but also *Contentment* and *Joy* or *Enthusiasm* and *Joy* show a small distance.

We used this distance matrix as input for a hierarchical clustering algorithm. We tried seven different linkage methods, for which the dendrograms are shown in Figure 4. We asked the annotators to rank the dendrograms based on their own intuition. Their top 3 consisted of the same clusters but the order differed. After discussion, the weighted-linkage clustering was chosen as the most intuitive one. This linkage method is also known as the Weighted Pair Group Method with Arithmetic Mean (WPGMA Algorithm). At each step of the algorithm, the two clusters that have the shortest distance between each other are combined. The distance between clusters is calculated by considering the distance between each pair of elements in the clusters (with one element per cluster) and taking the arithmetic mean of those distances.

Figure 5 plots the WPGMA dendrogram with a distance of 1.3 as cut-off value. This results in eight clusters (of which one only consists of one separate emotion category: *Remorse*). Four of these clusters are related to the basic emotions of Ekman (*Joy*, *Anger*, *Sadness* and *Fear*).

However, as the frequency analysis pointed out, not all emotion categories are equally represented in this dataset. This is why we also performed a second clustering analysis excluding those emotions that were indicated less than ten

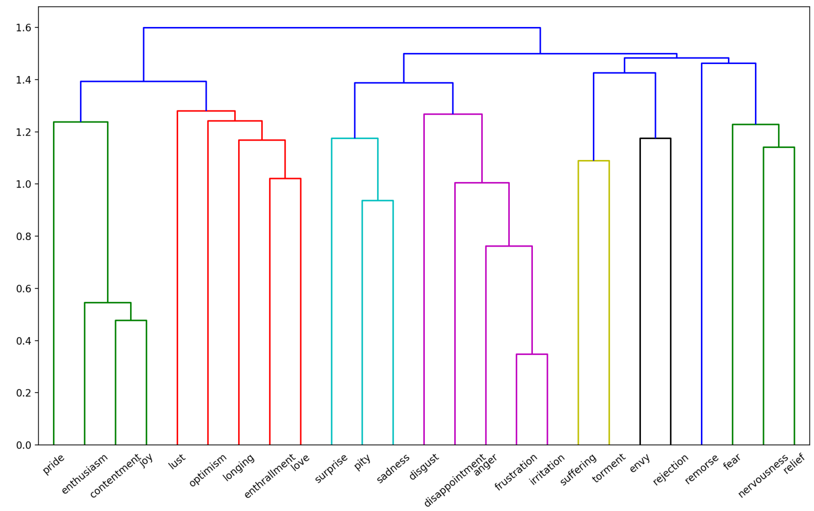


Figure 5. Dendrogram with weighted-linking.

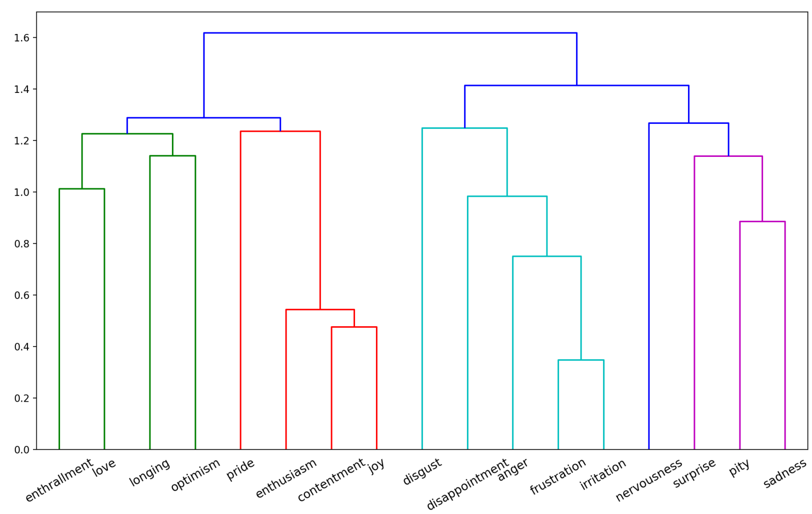


Figure 6. Dendrogram with weighted-linking without infrequent emotions.

times. This dendrogram is depicted in Figure 6. The threshold of ten might seem a bit harsh, especially since this results in the removal of the basic emotion *Fear*. However, based on Figure 5, we hypothesised that *Fear* would still be clustered together with *Nervousness* and the effect of removing it would not be problematic. We tested this and this was indeed the case. As expected, the clusters of Figure 6 are very similar to the ones in Figure 5, though only five clusters remain. Cluster 1 consists of the emotions *Enthrallment*, *Love*, *Longing* and *Optimism*, Cluster 2 comprises *Pride*, *Enthusiasm*, *Contentment* and *Joy*, the emotions *Disgust*, *Disappointment*, *Anger*, *Frustration* and *Irritation* form Cluster 3, *Nervousness* is a category on its own, and *Surprise*, *Pity* and *Sadness* are grouped together under Cluster 5.

The dendrogram nicely shows that the first two and last three categories form two distinct groups (positive versus negative emotions). Although we have a more equal distribution of negative and positive emotion clusters, our final clusters show a fair resemblance to Ekman's basic emotions.

To select an umbrella term per cluster, we take the Ekman emotion if the cluster has a term in common with the Ekman set. Otherwise, we select the emotion word with the highest frequency. This results in a final label set with *Love*, *Joy*, *Anger*, *Nervousness* and *Sadness* as emotion categories.

## V. Conclusion and Future Work

We established an emotion framework which is justified both theoretically and practically to perform automatic emotion detection on Dutch tweets. Frequency and cluster analyses of 229 tweet annotations resulted in a label set containing the 5 emotions *Love*, *Joy*, *Anger*, *Nervousness* and *Sadness*. Unlike many emotion frameworks directly borrowed from psychology, this label set has a more equal distribution over positive and negative emotions and due to our data-driven approach, it is much more grounded in the task (emotion detection) and the domain (Twitter). There is still a fair resemblance between Ekman's basic emotions and our labels, but we are the first that give an empirical motivation for the use of these categories.

For future work, we will use a similar approach to define a label set for two other domains, namely subtitles of reality TV and crisis communication data and verify whether these label sets even deviate more from the popular basic emotion frameworks. In this respect, mainly the crisis communication data will be interesting due to its topic specificity (in contrast to the more general character of tweets). Moreover, we will include dimensional annotations and aggregate these into a varied corpus to be used for Dutch emotion detection.

## ACKNOWLEDGMENT

This research was carried out with the support of the Research Foundation - Flanders under a Strategic Basic Research fellowship.

## REFERENCES

[1] L. F. Barrett, "Are emotions natural kinds?" Perspectives on Psychological Science, vol. 1, no. 1, 2006, pp. 28–58.

[2] J. Harkins and A. E. Wierzbicka, Emotions in Crosslinguistic Perspective. De Gruyter Mouton, 2001.

[3] J. Burgdorf and J. Panksepp, "The neurobiology of positive emotions," Neuroscience & Biobehavioral Reviews, vol. 30, no. 2, 2006, pp. 173–187.

[4] R. Bougie, R. Pieters, and M. Zeelenberg, "Angry customers don't come back, they get back: The experience and behavioral implications of anger and dissatisfaction in services," Journal of the Academy of Marketing Science, vol. 31, no. 4, 2003, pp. 377–393.

[5] S. Mac Kim and R. A. Calvo, "Sentiment analysis in student experiences of learning," in Educational Data Mining, 2010, pp. 111–120.

[6] S. M. Mohammad, X. Zhu, S. Kiritchenko, and J. Martin, "Sentiment, emotion, purpose, and style in electoral tweets," Information Processing & Management, vol. 51, no. 4, 2015, pp. 480 – 499.

[7] Y. Jin, B. F. Liu, D. Anagondahalli, and L. Austin, "Scale development for measuring publics emotions in organizational crises," Public Relations Review, vol. 40, no. 3, 2014, pp. 509–518.

[8] B. Desmet and V. Hoste, "Emotion detection in suicide notes," Expert Systems with Applications, vol. 40, no. 16, 2013, pp. 6351–6358.

[9] C. O. Alm, D. Roth, and R. Sproat, "Emotions from text: Machine learning for text-based emotion prediction," in Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing. Association for Computational Linguistics, 2005, pp. 579–586.

[10] C. Strapparava and R. Mihalcea, "Semeval-2007 task 14: Affective text," in Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007), 2007, pp. 70–74.

[11] S. Mohammad, F. Bravo-Marquez, M. Salameh, and S. Kiritchenko, "Semeval-2018 task 1: Affect in tweets," in Proceedings of The 12th International Workshop on Semantic Evaluation, 2018, pp. 1–17.

[12] P. Shaver, J. Schwartz, D. Kirson, and C. O'Connor, "Emotion knowledge: Further exploration of a prototype approach," Journal of personality and social psychology, vol. 52, no. 6, 1987, pp. 1061–1086.

[13] J. A. Russell, "A circumplex model of affect," Journal of Personality and Social Psychology, vol. 39, no. 6, 1980, pp. 1161–1178.

[14] A. Mehrabian and J. A. Russell, An Approach to Environmental Psychology. MIT Press, 1974.

[15] J. R. Fontaine, K. R. Scherer, E. B. Roesch, and P. C. Ellsworth, "The world of emotions is not two-dimensional," Psychological Science, vol. 18, no. 12, 2007, pp. 1050–1057.

[16] P. Ekman, "An argument for basic emotions," Cognition & Emotion, vol. 6, no. 3-4, 1992, pp. 169–200.

[17] R. Plutchik, "A general psychoevolutionary theory of emotion," in Theories of Emotion, R. Plutchik and H. Kellerman, Eds. Academic Press, 1980, pp. 3–33.

[18] A. Ortony and T. J. Turner, "What's basic about basic emotions?" Psychological Review, vol. 97, no. 3, 1990, pp. 315–331.

[19] M. B. Arnold, Emotion and Personality. Columbia University Press, 1960.

[20] N. H. Frijda, The Emotions. Cambridge University Press, 1986.

[21] J. A. Gray, "Précis of the neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system," Behavioral and Brain Sciences, vol. 5, no. 3, 1982, pp. 469–484.

[22] C. E. Izard, The Face of Emotion. Appleton-Century-Crofts, 1971.

[23] W. James, "What is an emotion?" Mind, vol. 9, no. 34, 1884, pp. 188–205.

[24] W. McDougall, An Introduction to Social Psychology. Methuen & Co, 1926.

[25] O. Mowrer, Learning Theory and Behavior. John Wiley & Sons Inc, 1960.

[26] K. Oatley and P. N. Johnson-Laird, "Towards a cognitive theory of emotions," Cognition and emotion, vol. 1, no. 1, 1987, pp. 29–50.

[27] J. Panksepp, "Toward a general psychobiological theory of emotions," Behavioral and Brain Sciences, vol. 5, no. 3, 1982, pp. 407–422.

[28] S. S. Tomkins, "Affect theory," in Approaches to Emotion, K. R. Scherer and P. Ekman, Eds. Lawrence Erlbaum Associates, 1984, pp. 163–195.

[29] J. B. Watson, Behaviorism. University of Chicago Press, 1930.

[30] B. Weiner and S. Graham, "An attributional approach to emotional development," Emotions, Cognition, and Behavior, 1984, pp. 167–191.

[31] S. Epstein, "Controversial issues in emotion theory," Review of Personality & Social Psychology, vol. 5, 1984, pp. 64–88.

[32] I. J. Roseman, "Cognitive determinants of emotion: A structural theory," Review of Personality & Social Psychology, vol. 5, 1984, pp. 11–36.

[33] M. Schröder, H. Pirker, and M. Lamolle, "First suggestions for an emotion annotation and representation language," in Proceedings of LREC, vol. 6, 2006, pp. 88–92.

[34] S. Aman and S. Szpakowicz, "Identifying expressions of emotion in text," in International Conference on Text, Speech and Dialogue. Springer, 2007, pp. 196–205.

[35] Y. Li, H. Su, X. Shen, W. Li, Z. Cao, and S. Niu, "DailyDialog: A manually labelled multi-turn dialogue dataset," in Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 1: Long Papers), 2017, pp. 986–995.

[36] S. M. Mohammad and F. Bravo-Marquez, "WASSA-2017 shared task on emotion intensity," in Proceedings of the Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA), 2017, pp. 34–49.

[37] D. Ghazi, D. Inkpen, and S. Szpakowicz, "Detecting emotion stimuli in emotion-bearing sentences," in Computational Linguistics and Intelligent Text Processing, A. Gelbukh, Ed. Springer International Publishing, 2015, pp. 152–165.

[38] V. Liu, C. Banea, and R. Mihalcea, "Grounded emotions," in 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE, 2017, pp. 477–483.

[39] K. R. Scherer and H. G. Wallbott, "Evidence for universality and cultural variation of differential emotion response patterning," Journal of Personality and Social Psychology, vol. 66, no. 2, 1994, pp. 310–328.

[40] H. Schuff, J. Barnes, J. Mohme, S. Padó, and R. Klinger, "Annotation, modelling and analysis of fine-grained emotions on a stance and sentiment detection corpus," in Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, 2017, pp. 13–23.

[41] S. M. Mohammad, "#emotional tweets," in Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation. Association for Computational Linguistics, 2012, pp. 246–255.

[42] E. Tjong Kim Sang and A. van den Bosch, "Dealing with big data: The case of twitter," Computational Linguistics in the Netherlands Journal, vol. 3, no. 12, 2013, pp. 121–134.

[43] J. Cohen, "A coefficient of agreement for nominal scale," Educational and Psychological Measurement, vol. 20, 1960, pp. 37–46.

[44] L. R. Dice, "Measures of the amount of ecologic association between species," Ecology, vol. 26, no. 3, 1945, pp. 297–302.