Universität Zürich[UZH]

Bachelorarbeit der Philosophischen Fakultät der Universität Zürich

Klassifikation und Analyse von Emotionswörtern in Tweets für die Sentimentanalyse

Verfasserin: Nadine Stamm

Matrikel-Nr: 10-611-374

Referent: Prof. Dr. Martin Volk

Betreuer: Dr. Manfred Klenner

Institut für Computerlinguistik

29.12.2014

Zusammenfassung

Die Sentimentanalyse, eine Technologie zur Verarbeitung natürlicher Sprache, ist ein aktuelles Thema in der Computerlinguistik. Wie das Wort „Sentiment“ andeutet, beschäftigt sich die Sentimentanalyse mit der Analyse von Gefühlen, Meinungen, und Einstellungen, die in Texten geäussert werden. Während die Sentimentanalyse für Englisch und für Textsorten wie Produktbewertungen schon viel Aufmerksamkeit erhalten hat, ist die Sentimentanalyse für Sprachen wie Deutsch und für andere Textsorten noch etwas unbekannter und auch weniger weit fortgeschritten.
In der vorliegenden Bachelorarbeit befasse ich mich deshalb mit der Sentimentanalyse für das Deutsche und mit einer bisher wenig untersuchten Textsorte: Kurznachrichten von der Mikroblogging-Plattform Twitter. Um herauszufinden, wie Gefühle in Texten ausgedrückt werden, analysiere ich Wörter, die eine Emotion ausdrücken. Das Ziel meiner Arbeit ist es, zu untersuchen, wie Emotionswörter in Tweets verwendet werden und wie sie für die Sentimentanalyse eingesetzt werden können. Dazu erstelle ich in einem ersten Schritt ein Lexikon aus deutschen Emotionswörtern, die ich gemäss der Theorie der Basisemotionen von Robert Plutchik nach acht Basisemotionen annotiere. In einem zweiten Schritt sammle ich Tweets, in denen die zusammengetragenen Emotionswörter in verschiedenen Kontexten vorkommen.
Wie die Analyse dieser Tweets gezeigt hat, gibt es eine Fülle an Schwierigkeiten, die sich für die Sentimentanalyse im Zusammenhang mit Emotionswörtern ergeben. Während Emotionswörter nützliche Hinweise für eine bessere Sentimentanalyse geben können, so müssen sie zuerst einmal korrekt erkannt und annotiert werden. Auch die Struktur der Twitter-Daten stellt die Sentimentanalyse vor Herausforderungen. Die erkannten Auffälligkeiten und Regelmässigkeiten in der Verwendung von Emotionswörtern in Tweets könnten danach in einem nächsten Schritt in Algorithmen verarbeitet werden.

Inhaltsverzeichnis

Zusammenfassung i

Inhaltsverzeichnis ii

Abbildungsverzeichnis iv

Tabellenverzeichnis v

Verzeichnis der zusätzlichen Dateien vi

1 Einleitung 1

2 Theoretische Grundlagen und Fragestellung 3

2.1 Methoden und Ablauf der Sentimentanalyse 3
2.2 Twitter 6
2.3 Emotionstheorien 7
2.4 Sentimentanalyse, Emotion, und Twitter 12
2.5 Verwandte Forschungsansätze 14
2.6 Fragestellung 15

3 Daten und Methoden 16

3.1 Aufbau des Emotionslexikons 16
3.2 Gewinnung und Struktur der Twitter-Daten 21

4 Analyse und Implikationen für die Sentimentanalyse 24

4.1 Probleme mit Emotionswörtern für die Sentimentanalyse 24
4.2 Hinweise auf Emotion in Tweets 31
4.3 Mehrere Emotionswörter in Tweets 33
4.4 Emotionswörter und Emoticons/Emojis 37
4.5 Emotionswörter und Hashtags 39

5 Diskussion 41

6 Schlusswort 44

Literaturverzeichnis **46**

Abbildungsverzeichnis

1 Sentimentanalyse-Ansätze im Überblick 5
2 Beispiele für Emojis 6
3 Plutchiks Modell der Emotionen 9
4 Gesamtansicht von Plutchiks Modell der Emotionen 10
5 Verteilung der Emotionswörter auf Emotionsklassen 19

Tabellenverzeichnis

1 Kappa-Koeffizient bei Zuordnung einer Basisemotion, sechs Annotatoren 20
2 Kappa-Koeffizient bei Zuordnung von ein bis zwei Basisemotionen, sechs Annotatoren 20
3 Probleme mit Emotionswörtern in Tweets 31

Verzeichnis der zusätzlichen Dateien

Die Dateien befinden sich auf der beiliegenden CD.

- Programm für Twitter-Mining (Python-Datei)
- Emotionslexikon (Excel-Datei)
- Korpus mit 800 Tweets (Excel-Datei)
- Korpus mit 650 Tweets (Word-Datei)
- Berechnung des Kappa-Koeffizienten (2 Excel-Dateien)

1 Einleitung

Die Sentimentanalyse, eine Technologie zur Verarbeitung natürlicher Sprache, ist ein aktuelles Thema in der Computerlinguistik. Wie das Wort „Sentiment“ andeutet, beschäftigt sich die Sentimentanalyse mit der Analyse von Gefühlen, Meinungen, und Einstellungen, die in Texten geäussert werden. Nicht überraschend ist die automatische Erkennung solcher Aussagen eine Herausforderung. Während die Sentimentanalyse für Englisch und für Textsorten wie Produktbewertungen schon viel Aufmerksamkeit erhalten hat, ist die Sentimentanalyse für Sprachen wie Deutsch und für andere Textsorten noch etwas unbekannter und auch weniger weit fortgeschritten.

In der vorliegenden Bachelorarbeit befasse ich mich deshalb mit der Sentimentanalyse für das Deutsche und mit einer bisher wenig untersuchten Textsorte, nämlich Kurznachrichten aus der Mikroblogging-Plattform Twitter. Wie bereits erwähnt, beschäftigt sich die Sentimentanalyse mit Gefühlen. Um herauszufinden, wie Gefühle in Texten ausgedrückt werden, analysiere ich deshalb sogenannte Emotionswörter, also Wörter, die eine Emotion ausdrücken. Hinter meiner Themenwahl stehen mehrere Gründe. Für die Untersuchung des Deutschen entschied ich mich zum einen, weil dies meine Muttersprache ist. Das sollte es mir erleichtern, sprachliche Besonderheiten in Tweets zu erkennen. Zum anderen herrscht für die deutsche Sentimentanalyse immer noch ein Mangel an Untersuchungen, Korpora und lexikalischen Ressourcen, was ein guter Grund dafür ist, sich in diese Richtung vorzuwagen. Auch hinter der Entscheidung, Twitter-Daten zu untersuchen, stehen mehrere Gründe. Wie die Entwicklungen der letzten Jahre gezeigt haben, nehmen soziale Medien einen immer höheren Stellenwert für Unternehmen ein. Soziale Medien bieten Firmen die Gelegenheit, mehr über ihre Kunden zu erfahren und direkte Rückmeldungen zu Produkten und Dienstleistungen zu erhalten. Es ist deshalb nicht verwunderlich, dass Unternehmen vermehrt nach Analysetools für Social Media verlangen. Die Sentimentanalyse ist eine Technologie, die für solche Aufgaben gut geeignet ist. Dazu muss sie aber verlässlich funktionieren. Zudem machen die Eigenschaften von Twitter-Daten die Sentimentanalyse spannend. Da die Nachrichten maximal 140 Zeichen lang sein dürfen, fehlt der Kontext. Zudem ist der Sprachgebrauch in Tweets oft umgangssprachlich, oder sehr kreativ. Dies stellt die Sentimentanalyse vor zusätzliche

Schwierigkeiten, bietet aber auch eine spannende Herausforderung. Ein Vorteil von Twitter-Daten ist zudem, dass sie frei verfügbar sind. Somit können sie mit relativ wenig Aufwand gesammelt und zu einem eigenen Korpus zusammengestellt werden. Das Ziel meiner Arbeit ist es, zu untersuchen, wie Emotionswörter in Tweets verwendet werden und wie sie für die Sentimentanalyse eingesetzt werden können. Daraus ergibt sich die folgende Leitfrage: „Wie können Emotionswörter dazu beitragen, die Sentimentanalyse von Tweets zu verbessern?“ Um diese Untersuchung durchzuführen, erstelle ich ein Lexikon aus Emotionswörtern und sammle Tweets, in denen diese Emotionswörter vorkommen. Um die Rolle der Emotionen genauer untersuchen zu können, annotiere ich die gesammelten Emotionswörter nach der Theorie der Basisemotionen von Plutchik [1980]. Meine Analyse der Tweets soll zeigen, welche Probleme sich für die Sentimentanalyse ergeben. Für diese Probleme versuche ich, mögliche Lösungsansätze zu entwickeln. Zudem erläutere ich, ob und wie diese Probleme schon von anderen gelöst wurden.
Der Aufbau meiner Arbeit ist dabei wie folgt: In Kapitel 2 erarbeite ich die theoretischen Grundlagen, welche für meine Untersuchung notwendig sind. In Kapitel 3 beschreibe ich die verwendeten Daten und das methodische Vorgehen. Darauf folgen die Analyse der Tweets und die daraus abgeleiteten Empfehlungen für die Sentimentanalyse in Kapitel 4. Kapitel 5 und 6 enthalten eine Diskussion meiner Ergebnisse, eine Zusammenfassung und einen kurzen Ausblick.

2 Theoretische Grundlagen und Fragestellung

Die Sentimentanalyse ist ein noch jüngeres Forschungsgebiet in der Computerlinguistik, das etwa seit dem Jahr 2000 zunehmend an Aufmerksamkeit gewann [Liu, 2012, 1]. Obwohl diese Forschungsvertiefung noch relativ neu ist – oder vielleicht gerade deswegen? – scheint sie sehr beliebt zu sein, denn es gibt bereits unzählige Studien dazu, wie Feldman [2013, 82] schreibt: „Sentiment analysis is one of the hottest research areas in computer science. Over 7,000 articles have been written on the topic.“ Noch etwas jünger als die Sentimentanalyse ist die Social Media-Technologie Twitter, die seit dem Jahr 2006 existiert. Die ersten Untersuchungen, welche diese beiden Themen verbanden und sich mit Sentimentanalyse für Twitter befassten, wurden um das Jahr 2010 veröffentlicht. Im Folgenden erarbeite ich die für meine Untersuchung notwendigen theoretischen Grundlagen und leite daraus meine Fragestellung ab.

2.1 Methoden und Ablauf der Sentimentanalyse

Die Sentimentanalyse (englisch *Sentiment Analysis*) ist eine Technologie zur Verarbeitung natürlicher Sprache, die analysiert, welche Meinungen oder Gefühle in einem Text geäussert werden. In der Forschungsliteratur werden synonym für *Sentimentanalyse* oft auch die Begriffe *Sentiment Detection, Opinion Mining*, oder *Opinion Extraction* verwendet. Die Bedeutungen dieser Begriffe decken sich allerdings zu einem grossen Teil, weshalb ich im Folgenden nur von der Sentimentanalyse sprechen werde. Liu [2012, 1] formuliert die Aufgabe der Sentimentanalyse als „the field of study that analyzes people's opinions, sentiments, evaluations, appraisals, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes“. Die mittels Sentimentanalyse gewonnenen Meinungen über Produkte, Organisationen und Themen werden vielerorts eingesetzt. Die Hauptanwendungsgebiete liegen dabei im Marketing, der Öffentlichkeitsarbeit, und der Kampagnenführung. Aber auch Einzelpersonen wie

Politiker, Investoren oder Online-Shopper können von der Technologie profitieren [Feldman, 2013, 82]. Weitere Anwendungsmöglichkeiten der Sentimentanalyse liegen in den Bereichen Gesundheit und Sicherheit, z.B. zur Erkennung einer möglichen Gefährdung von Personen [Thelwall et al., 2010, 2544] und allgemein in der Forschung im Bereich (Online-)Linguistik, Kommunikation und Emotion. Ein Vorteil, welcher die Sentimentanalyse gegenüber traditionellen Meinungsbefragungen hat, wird von Shelke et al. [2012] erwähnt: Übliche Erhebungen führen oft dazu, dass aufgrund der vorab definierten Fragen und der wenig motivierten Teilnehmer nur abgefragt wird, was ohnehin schon bekannt ist. Bei der Sentimentanalyse hingegen liegt der Fokus auf dem „Zuhören“ anstatt dem Fragen, was zu realistischeren Ergebnissen führt [Shelke et al., 2012, 30].

Allgemein kann die Sentimentanalyse auf drei Ebenen erfolgen: auf Dokumentebene, auf Satzebene, und auf Phrasenebene [Liu, 2012, 4]. Dabei soll jeweils die in einem Dokument, einem Satz, oder einer Phrase geäusserte Meinung erkannt und als positiv oder negativ eingeordnet werden. Bevor die Polarität eines Satzes erkannt werden kann, muss zudem noch bestimmt werden, ob ein Satz objektiv (Fakten vermittelnd) oder subjektiv (Meinungen vermittelnd) ist [Liu, 2012, 4]. Die Sentimentanalyse auf Dokumentebene ist die gröbste Form der Meinungserkennung, denn dabei wird angenommen, dass ein Dokument genau eine Meinung über ein Thema enthält. Bei der Sentimentanalyse auf Satzebene wird dem Fakt Rechnung getragen, dass zu einem Thema auch mehrere Meinungen bestehen können. Dabei wird angenommen, dass ein Satz jeweils eine Meinung zum besprochenen Thema enthält. Allerdings kann es auch sein, dass in einem Satz mehrere Meinungen zu unterschiedlichen Aspekten vorkommen, z.B. wenn verschiedene Attribute eines Produkts bewertet werden. In diesem Fall ist Sentimentanalyse auf Phrasenebene am besten geeignet [Feldman, 2013, 83-85].

Die in der Literatur präsentierten Sentimentanalyse-Ansätze lassen sich aber noch weiter klassifizieren. Dabei besteht ein grundsätzlicher Unterschied zwischen Machine Learning- und lexikonbasierten Ansätzen.

Beim maschinellen Lernen wiederum unterscheidet man zwischen überwachten (supervised) und unüberwachten (unsupervised) Vorgehensweisen. Überwachtes Lernen bedeutet, dass ein System mittels bereitgestellten Daten und Klassen darauf trainiert wird, neue Instanzen in eine der vorgegebenen Klassen einzuordnen. Für die Sentimentanalyse wären z.B. „positiv“ und „negativ“ naheliegende Klassen [Feldman, 2013, 84]. Im Gegensatz dazu soll ein System beim unüberwachten Lernen durch die Analyse der Lexik und Syntax einer Instanz eine Vorhersage über deren Charakter machen. Dabei bestimmt ein Grenzwert, ab wann eine Instanz als positiv bzw. negativ gilt [Feldman, 2013, 84]. Die Algorithmen für das maschinelle Lernen verwenden sogenannte „Features“, um ihre Entscheidungen zu treffen. Solche Fea-

tures können z.B. Uni- und Bigramme, Part-of-Speech-Tags oder die Position von Wörtern im Satz sein [Shelke et al., 2012, 31].
Bei den lexikonbasierten Ansätzen dagegen werden als Basis ein Wörterbuch oder ein Korpus verwendet, um die Polarität eines Dokuments/Satzes oder einer Phrase zu bestimmen. Dazu wird für jedes Wort im Satz seine „sentiment orientation“ bestimmt und daraus die Polarität des gesamten Dokuments/Satzes abgeleitet [Shelke et al., 2012, 32]. Um Wörterbücher nicht vollständig manuell erstellen zu müssen, verwenden viele Wissenschaftler bestehende lexikalische Ressourcen wie z.B. WordNet. Eine andere Variante ist die Nutzung von bestehenden Korpora, mit denen die emotionale Affinität von Wörtern automatisch bestimmt wird [Shelke et al., 2012, 32].
Die Übersicht in Abbildung 1 illustriert den Zusammenhang zwischen den erwähnten Sentimentanalyse-Ansätzen und nennt jeweils häufig verwendete Algorithmen [Medhat et al., 2014, 1095].

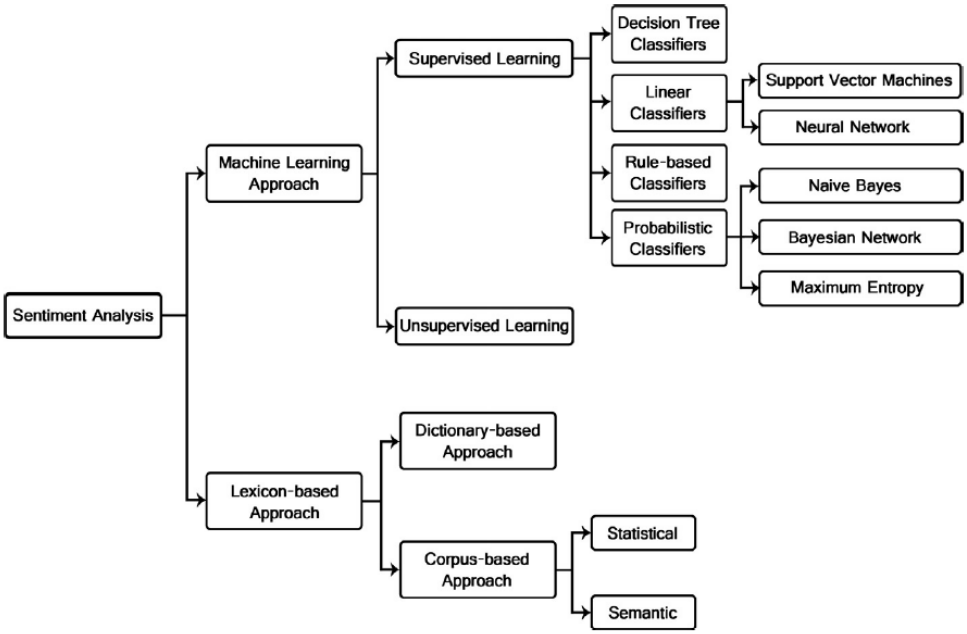

Abbildung 1: Sentimentanalyse-Ansätze im Überblick [Medhat et al., 2014, 1095]

Wenn entschieden wurde, auf welcher Ebene die Sentimentanalyse erfolgt und welche Vorgehensweise dazu verwendet wird, folgt der Ablauf der Sentimentanalyse grundsätzlich immer demselben Schema. Zuerst müssen die zu untersuchenden Daten, welche in den verschiedensten Formaten vorliegen können, in Text umgewandelt werden [Feldman, 2013, 83]. Danach folgt die Vorverarbeitung der Daten. Dazu gehört z.B. das Stemmen, Tokenisieren, PoS-Taggen, und Extrahieren von Entitäten und Relationen. Nun kann die Sentimentanalyse durch das mit den nötigen lexikalischen Ressourcen ausgestattetem System stattfinden [Feldman, 2013, 83].

2.2 Twitter

Twitter [1] ist eine Plattform für Mikroblogging, auf der sich Internetnutzer mittels kurzen Nachrichten, sogenannten *Tweets*, mitteilen und austauschen können. Die Tweets haben eine maximale Länge von 140 Zeichen und können nebst Text auch Bilder, Videos oder Links enthalten. Zudem verwenden viele Nutzer in ihren Nachrichten Emoticons („Smileys“) oder neuerdings auch Emojis (siehe Abbildung 2 für Beispiele), um ihren Emotionen Nachdruck zu verleihen. Die Popularität von Twitter ist enorm: täglich werden 500 Millionen Tweets versandt, pro Monat zählt der Dienst 248 Millionen aktive Nutzer [Twitter, 2014]. Ende 2014 befindet sich Twitter auf Platz 8 der meistbesuchten Webseiten der Welt [Alexa, 2014]. Nebst dem Schreiben von Kurznachrichten können die Benutzer von Twitter anderen Nutzern „folgen“ und damit deren Nachrichten abonnieren, oder fremde Tweets selbst nochmals veröffentlichen (sogenannte *Retweets*). Sollen andere User in Tweets erwähnt werden, so kann dies mithilfe dem @-Zeichen gefolgt von einem Benutzername erfolgen. Ein weiteres bekanntes Merkmal von Twitter sind die *Hashtags* (das #-Zeichen), mit dem das Thema eines Tweets zusätzlich beschrieben werden kann, z.B. #womensfootball. Die Bandbreite der Tweets reicht dabei von persönlichen und unterhaltsamen Nachrichten von Privatpersonen über die tausendfach geteilten Nachrichten prominenter Personen an ihre Fans bis hin zu informativen Tweets wie etwa Schlagzeilen.

Abbildung 2: Beispiele für Emojis

Twitter ist eine öffentliche Plattform. Die Nachteile und Gefahren, die sich dadurch ergeben, scheinen allerdings von vielen Nutzern vergessen oder ignoriert zu werden. Für Forscher wiederum bieten sich durch Twitters öffentliche Natur nahezu unendlich viele Daten, die mit wenig Aufwand gesammelt werden können. Dies verlangt aber auch nach einem verantwortungsvollen Umgang mit den Daten (siehe dazu Abschnitt 3.2).

[1] http://www.twitter.com

2.3 Emotionstheorien

Um die Verwendung von Emotionswörtern für die Sentimentanalyse zu untersuchen, braucht es ein theoretisches Fundament zum Thema Emotion. Dabei liegt der Fokus auf den sogenannten *Basisemotionen*. Robert Plutchiks Theorie zu den Basisemotionen behandle ich zudem etwas detaillierter, da meine Untersuchung auf dieser Theorie beruht.

Grundsätzlich gibt es zwei verschiedene Ansätze, wie Emotionen theoretisch abgebildet werden können: als finite Kategorien, oder als Dimensionen [Devitt and Ahmad, 2013, 477]. Der kategorische Ansatz geht davon aus, dass es eine (finite) Anzahl von Basisemotionen gibt, die in allen Kulturen gleich erlebt werden. Unterstützung erhält dieser Ansatz durch Studien von Gesichtsausdrücken, welche global ähnlich eingeordnet werden [Devitt and Ahmad, 2013, 477]. Im Gegensatz dazu geht der dimensionale Ansatz davon aus, dass Emotionen nicht mittels festgelegten Kategorien, sondern anhand verschiedener, „offener“ Dimensionen gemessen werden. Zu diesen Dimensionen werden z.B. die Achsen „gut-schlecht“ (Valenz), „aktiv-passiv“ (Intensität), und „stark-schwach“ (Dominanz) gezählt. Diese zwei Dimensionen sind wichtig zu erwähnen, weil sie einen Einfluss darauf haben, wie Emotionen in computergestützten Analysen betrachtet werden [478]. Wie Devitt and Ahmad berichten, wird in der Sentimentanalyse im Allgemeinen der dimensionale Ansatz bevorzugt. Dies kann man gut daran erkennen, dass die allermeisten Sentimentanalyse-Technologien und Ressourcen auf die Valenz einer Aussage ausgerichtet sind, also darauf, ob eine Aussage eher positiv oder negativ ist [Devitt and Ahmad, 2013, 478].

Das von mir in dieser Untersuchung befolgte Vorgehen, Emotionswörter nach festen Emotionskategorien zu klassifizieren, folgt dagegen ganz klar einem kategorialen Ansatz, um Emotionen theoretisch abzubilden. Es muss allerdings beachtet werden, dass die von mir angestrebte Klassifizierung nur einen Teilschritt eines gesamten Sentimentanalyse-Systems darstellt. Würde also meine Emotionswörter-Klassifizierung in ein übliches Gesamtsystem eingebettet, so könnte man vielleicht von einem (auf die Darstellung von Emotion bezogenen) hybriden Ansatz sprechen.

Folgend beschreibe ich einige bekannte Emotionstheorien, welche bereits zur Klassifizierung von Emotionswörtern in der Sentimentanalyse-Forschung verwendet wurden.

Die Thematik der primären und sekundären Emotionen wurde bereits vor langer Zeit von Philosophen und Forschern wie Descartes, Spinoza, und Darwin erwähnt [Plutchik, 1980, 131]. Zu den ersten modernen Auffassungen gehört die Theorie von McDougall aus dem Jahr 1921. Der Psychologe beschreibt den Menschen als Besitzer von Instinkten, die, wenn sie aktiviert werden, eine Emotion mit sich ziehen. Der

Instinkt „Flucht“ löst also z.B. das Gefühl „Furcht“ aus. McDougall definiert sieben wichtige Instinkte: Flucht, Abstossung, Wissbegier, Kampflust, Selbsterniedrigung, Selbstbehauptung, und den Elterninstinkt. Aus diesen Instinkten leitet er sieben primäre Emotionen ab: Furcht, Ekel, Verwunderung, Ärger, Unterwerfung, Freude, und Zärtlichkeit [Plutchik, 1980, 132].
Eine weitere Theorie, die sich mit Basis- oder Primäremotionen befasst, stammt von Dahl and Stengel [1978]. Damit verfolgen die Autoren das Ziel, ein Set von Emotionswörtern zu entwickeln, um damit Berichte aus der Psychotherapie inhaltlich zu beschreiben [Dahl and Stengel, 1978, 271]. Ihre Arbeit basiert auf der Annahme, dass Menschen in verschiedenen Situationen unbewusste Entscheidungen treffen, wie sie auf diese Situation emotional reagieren. Diese Entscheidungen können in einem dreistufigen Entscheidungsbaum dargestellt werden, dessen Blätter acht Basisemotionen sind. Dabei ist die erste zu treffende Entscheidung, ob eine Emotion ein Objekt oder das Individuum selbst betrifft [Dahl and Stengel, 1978, 273]. Auf der nächsten Ebene muss entschieden werden, ob ein Objekt Anziehung oder Abstossung auslöst bzw. ob sich eine Person positiv oder negativ fühlt. Auf der dritten Ebene folgt die Auswahl, ob man sich zu einem Objekt hin angezogen/abgestossen fühlt oder ob die Anziehung/Abstossung von dem Objekt selbst ausgeht. Auf der Seite des Individuums muss die Auswahl getroffen werden, ob ein Gefühl aktiver oder passiver Natur ist [273]. Nach Durchlaufen dieses Entscheidungsbaum landet man schliesslich bei einer der acht Grundemotionen Liebe, Überraschung, Ärger, Furcht, Zufriedenheit, Freude, Niedergeschlagenheit, oder Beklemmung [273].
Zwei Jahre später präsentierte Plutchik [1980] eine umfassende Theorie der Emotionen. Plutchik beschreibt seine Ideen zuerst in zehn Postulaten. Über die Basisemotionen schreibt er unter anderem folgendes:

POSTULATE 5 There is a small number of basic, primary, or prototype emotions. POSTULATE 6 All other emotions are mixed or derivative states; that is, they occur as combinations, mixtures, or compounds of the primary emotions [Plutchik, 1980, 129].

Plutchik argumentiert, dass es acht grundsätzliche menschliche (und tierische) Verhaltensweisen gibt, aus denen sich die grundlegenden Dimensionen von jeglichem emotionalen Verhalten ableiten lassen [Plutchik, 1991, 61]. Diese grundlegenden Dimensionen emotionalen Verhaltens sind: Aufnahme/Akzeptanz, Zurückweisung/Ablehnung, Zerstörung, Beschützen, Reproduktion, Mangel, Orientierung, und Entdeckung [63]. Aus diesen eher „funktionalen“ Verhaltenskategorien lassen sich dann die acht „subjektiven“ (Basis-)Emotionen ableiten: Vertrauen, Ekel, Ärger, Furcht, Freude, Traurigkeit, Überraschung, und Erwartung [Plutchik, 1980, 154]. Diese Emotionen können in einem dreidimensionalen Modell dargestellt werden (siehe Abbil-

dung 3), wobei zu jeder Basisemotion zusätzlich eine intensivere und eine schwächere Emotion dazukommen. Die nach unten zulaufende Form des Modells soll zudem andeuten, dass schwächere/"mildere" Emotionen weniger voneinander unterscheidbar sind als intensivere [Plutchik, 1980, 157-158].

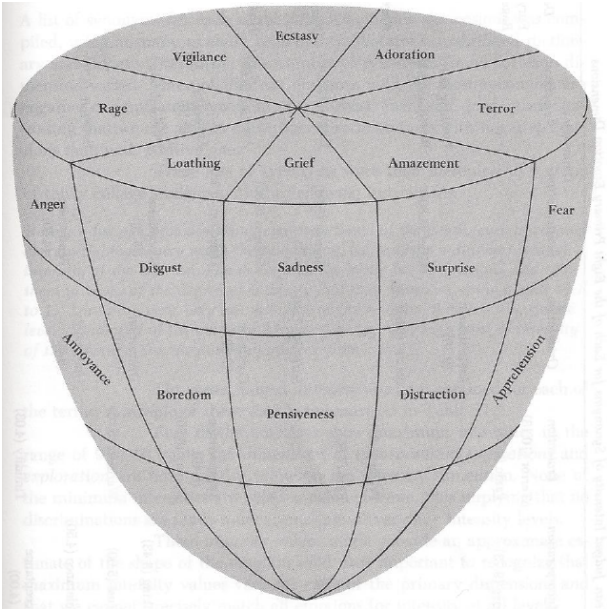

Abbildung 3: Plutchiks Modell der Emotionen [Plutchik, 1980, 157]

In der in Abbildung 3 dargestellten Originalgrafik von Plutchik sieht man allerdings nur eine Seite des Modells. Zum besseren Verständnis folgt deshalb noch eine nicht originale, aber verbreitete Grafik in Abbildung 4, die alle Emotionen enthält[Donaldson, 2011].

In Abbildung 4 kann man zudem auch die in Postulat 6 erwähnten Sekundäremotionen wie z.B. *Liebe* oder *Reue* sehen. Plutchiks Argumentation ist, dass sich durch die Kombination zweier oder dreier primärer Emotionen Dyaden bzw. Triaden bilden lassen, welche eine komplexere Emotion beschreiben. Dies kann wiederum auf den verschiedenen Ebenen der Gefühlsintensität des Modells geschehen [Plutchik, 1980, 161]. Allerdings können einige Emotions-„Mischungen" zu Herausforderungen führen. Entweder fehlt ein passendes Wort, um eine Kombination zweier Emotionen zu beschreiben, oder das gemeinsame Auftreten dieser zwei Emotionen kommt in der menschlichen Erfahrung gar nicht vor.
Auch das Kombinieren zweier Emotionen von unterschiedlichen Intensitätslevels könnte Schwierigkeiten bereiten [161]. Folgend einige Beispiele von Plutchik, wie er aus Primär- zu Sekundäremotionen kommt: „Erwartung + Freude = Optimismus", „Ekel + Erwartung = Zynismus", „Überraschung + Ekel = ?", „Ekel + Freude = Krankhaftigkeit (?)" [Plutchik, 1980, 162]. Diese Dyaden veranschaulichen auch die eben erwähnten möglichen Schwierigkeiten.

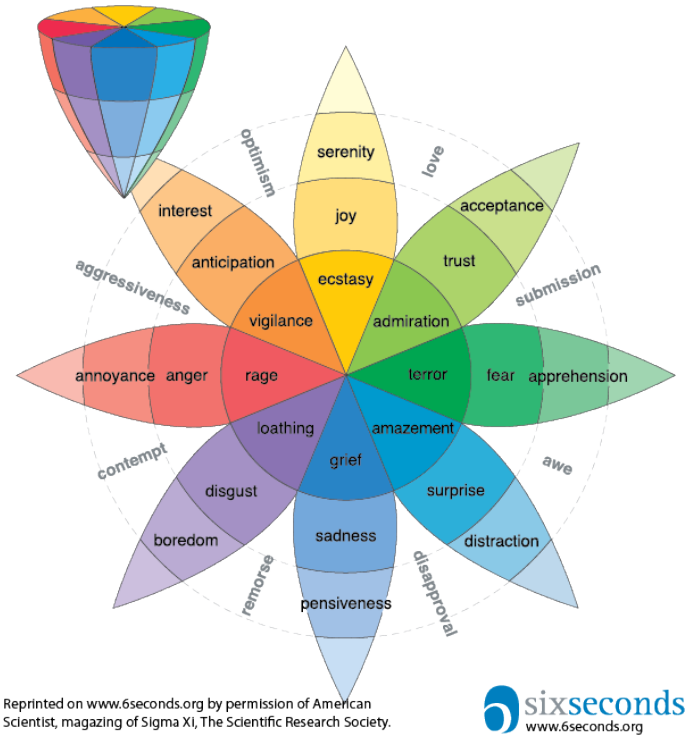

Abbildung 4: Gesamtansicht von Plutchiks Modell der Emotionen [Donaldson, 2011]

Plutchiks Idee, Emotionen in einem dreidimensionalen Modell zueinander in Verbindung zu stellen, scheint im Vergleich zu anderen, bisher genannten Theorien zusätzliche Erkenntnisse zu bringen. Zudem ist seine Praxis, sekundäre Emotionen als Überschneidungen von Basisemotionen zu klassifizieren, ein gutes Vorgehen, um mit komplexeren Emotionen umzugehen. Allerdings erwähnt Plutchik auch die Probleme, die dabei auftauchen. Deshalb scheint mir die Einteilung in sekundäre Emotionen etwas zu unzuverlässig, um für die Sentimentanalyse verwendet zu werden.

Zwei weitere Theorien über Emotionen, auf welche in der Sentimentanalyse-Forschung zurückgegriffen wird, wurden nur wenige Jahre nach Plutchik publiziert.

Ekman [1992] argumentiert wie andere Forscher vor ihm, dass die Unterscheidung zwischen den primären Emotionen aufgrund von menschlichen Gesichtsausdrücken, die für verschiedene Kulturen dieselben sind, gemacht werden kann [Ekman, 1992, 176]. Aus diesen Gesichtsausdrücken ergeben sich folgende sechs Basisemotionen: Furcht, Freude, Ärger, Traurigkeit, Ekel, und Überraschung [175-176].

Die Theorie von Ortony et al. [1988] geht noch einen Schritt weiter, denn die Autoren haben das folgende Ziel: „[W]e would like to lay the foundation for a computationally tractable model of emotion. [...] an account of emotion that could in principle be used in an Artificial Intelligence (AI) system that would, for example, be able to

reason about emotions.“ [Ortony et al., 1988, 2]. Die Autoren sind der Ansicht, dass Menschen auf drei grundlegende Aspekte in der Welt – Ereignisse, Agenten, und Objekte – Reaktionen zeigen, und zwar in Form von Emotionen. Diese Reaktionen werden dann anhand diverser Dimensionen und Bedingungen klassifiziert und in einem Entscheidungsbaum dargestellt, dessen Blätter wiederum einige grundlegende Emotionen sind [19].

Ein weiterer Ansatz zur Klassifizierung von Emotionen, der in der Sentimentanalyse verwendet wird, kommt aus der psychologischen Praxis. Dabei werden allerdings keine Basisemotionen, sondern eher Grundstimmungen unterschieden. Dieses psychometrische Instrument heisst „Profile of Mood States“, kurz POMS. Es wurde 1971 von McNair, Loor und Droppleman entwickelt und besteht aus einem Fragebogen von 65 Adjektiven, für welche die befragte Person angeben muss, wie sehr sie ihrer Stimmung entsprechen [Bollen et al., 2011, 451]. Mit einem Auswertungsschlüssel werden die Antworten dann in einen 6-dimensionalen Stimmungsvektor umgewandelt. Die sechs Grundstimmungen, die mit dem POMS untersucht werden, sind Anspannung, Niedergeschlagenheit, Ärger, Vitalität, Ermüdung, und Verwirrung [Bollen et al., 2011, 451]. Während das Profile of Mood States primär für die Befragung von Menschen gedacht ist, gibt es auch eine erweiterte Version namens POMS-ex, die spezifisch für die Anwendung auf grosse Textmengen gedacht ist. Dabei werden die ursprünglichen 65 Adjektive auf insgesamt 739 Synonyme und verwandte Wörter erweitert [Bollen et al., 2011, 451].

Ein neuerer Versuch, Emotionen für die Verarbeitung natürlicher Sprache zu klassifizieren, stammt von [Hobbs und Gordon 2011]. Dabei ist es ihr Ziel, eine formale Theorie zu entwickeln, welche für die Anwendung in Sprachtechnologien nützlich ist [Hobbs and Gordon, 2011, 27]. Mittels CoreWordNet, einer auf die häufigsten Konzepte beschränkten Version von WordNet, und anderen lexikalischen Ressourcen sammeln Hobbs und Gordon die häufigsten Wörter zum Thema Kognition und Emotion und ordnen diese 140 Wörter in 33 Kategorien ein. Beispiele für solche übergeordneten Emotionskategorien sind „joy emotion“, „sorry-for emotion“, „relief emotion“, und „love emotion“ [29-30]. Dann definieren die Autoren für 26 Basisemotionen eine prädikatenlogische Darstellung. Dazu werden die Emotionen in Beziehung gesetzt zu den Situationen, durch die sie entstehen, und zu den Verhaltensweisen, die durch sie ausgelöst werden können [Hobbs and Gordon, 2011, 31].

Die hier vorgestellten Emotionstheorien, ausgenommen die Ansätze von Ortony et al. [1988] und Hobbs and Gordon [2011], berufen sich alle auf das Konzept der Basisemotionen, nach welchem alle Menschen dieselben grundlegenden Emotionen empfinden. Dass ein solches Konzept besteht, darin sind sich auch alle Autoren einig. Die Frage stellt sich jedoch, weshalb jeder Autor andere Emotionen als Basisemotionen definiert. Zudem gibt es auch keinen Konsens über die Anzahl der Basisemotionen.

Damit fehlt also eine einzige, allgemein anerkannte Theorie der Basisemotionen. Natürlich stösst die Idee der Basisemotionen auch auf Kritik (z.B. durch Ortony et al. [1988, 26]: „talk of ‚basic emotions' is unacceptably vague"). Trotzdem scheint mir eine Einteilung in wenige, primäre Emotionen für meine Untersuchung sinnvoll zu sein. Die Emotionen sind klar voneinander abgegrenzt und die niedrige Anzahl macht es möglich, schnell eine Auswahl zu treffen oder einen Classifier auf die Unterscheidung zu trainieren. Eine Klassifizierung in detailliertere Kategorien wie die von Hobbs and Gordon [2011] ist zwar genauer, könnte aber auch schwieriger umzusetzen sein, und muss deshalb nicht unbedingt zu einer Verbesserung der Sentimentanalyse beitragen.

2.4 Sentimentanalyse, Emotion, und Twitter

Nach der theoretischen Einführung in die Sentimentanalyse, in Twitter und in verschiedene Emotionstheorien möchte ich nun aufzeigen, wie diese Themen in der bestehenden Forschung verknüpft werden und welche Ressourcen dafür bereits existieren. Wie bereits erwähnt, wurden die ersten Forschungsarbeiten, die diese drei Themen kombinieren, um das Jahr 2010 publiziert. Mehrheitlich beschäftigen sich diese mit der englischen Sprache. Die Sentimentanalyse von Twitter-Daten war 2013 ausserdem Thema einer Aufgabe von *SemEval* (Semantic Evaluation Exercises, Task 2 [2]). Dabei sollten die Teilnehmer ein Sentimentanalyse-System entwickeln, das auf Phrasen- und/oder auf Satzebene funktioniert und Tweets als positiv, negativ, oder neutral einordnet [Nakov et al., 2013, 313]. Um die Systeme zu trainieren und zu testen, wurde von SemEval ein Korpus mit ca. 15'000 annotierten Tweets zur Verfügung gestellt [314]. Die besten Systeme erreichten ein F-measure von 88.4% für phrasenbasierte und 69% für satzbasierte Sentimentanalyse [Nakov et al., 2013, 319].
Erfolgreiche Sentimentanalyse-Ansätze für Twitter arbeiten mit verschiedensten Techniken: Ein Beispiel sind Pak and Paroubek [2010], die n-Gramme als Features verwenden und den Naïve Bayes-Algorithmus verwenden, um ihren Classifier zu bauen [Pak and Paroubek, 2010, 1322-1323]. Im Gegensatz dazu verwenden Suttles and Ide [2013] Hashtags und Emoticons als Features und einen Distant Supervision-Algorithmus für ihren Classifier [Suttles and Ide, 2013, 121]. Ein drittes Beispiel ist der Ansatz von [Skinner et al. 2013], in dem die Autoren ein spezifisches Twitter-Sentimentlexikon erstellen, und den Classifier mit dem Support Vector Machines-Algorithmus trainieren [Skinner et al., 2013, 5]. Aus Platzgründen werde ich an späterer Stelle detaillierter auf diese und weitere Ansätze eingehen und mich nach-

[2] http://www.cs.york.ac.uk/semeval-2013/task2/

folgend auf die Beschreibung von Ressourcen für die Sentimentanalyse konzentrieren. Für die allgemeine Sentimentanalyse sind bereits einige (englischsprachige) lexikalische Ressourcen vorhanden. Zu den älteren lexikalischen Ressourcen gehören z.B. das *General Inquirer* Lexikon von Stone et al. (1966) und das *Dictionary of Affect Language* von Whissell (1989) [Devitt and Ahmad, 2013, 481]. Heute werden vermehrt neuere lexikalische Ressourcen wie *WordNetAffect* (Strappavara und Valitutti 2004) und *SentiWordNet* (Esuli und Sebastiani 2006) verwendet [481]. *WordNetAffect* basiert auf der lexikalischen Datenbank *WordNet* und enthält daraus diejenigen Synsets, welche einer affektiven Domäne wie z.B. „attitude“, „emotion“, oder „sensation“ zugeordnet wurden [Devitt and Ahmad, 2013, 485]. *SentiWordNet* dagegen ist eine Art zusätzliche „Schicht“, die über *WordNet* gelegt wurde und die für jedes Synset Polaritäts- und Objektivitätswerte zwischen 0 und 1 angibt [486].
Doch es gibt mittlerweile auch Ressourcen, die spezifisch auf Twitter ausgerichtet sind. Dazu gehört z.B. die Wortliste von Nielsen [2011]. Nielsen argumentiert, dass lexikalische Ressourcen auch typischen Internet-Slang und obszöne Wörter beinhalten sollten, um die Twitter-Sentimentanalyse verbessern zu können [Nielsen, 2011, 93]. Wie der Vergleich mit anderen lexikalischen Ressourcen für die Sentimentanalyse (wie z.B. *General Inquirer* und *SentiStrength*) zeigt, erreicht ein SA-System mit Nielsens Liste den zweiten Platz [5]. Dies ist insbesondere deshalb interessant, weil die Ressource von Nielsen mit weniger als 2500 Wörtern auskommt und damit um einiges kleiner ist als die anderen Ressourcen [Nielsen, 2011, 94].
Nebst lexikalischen Ressourcen gibt es auch wenige frei verwendbare Twitter-Korpora. Dazu gehört etwa das bereits erwähnte 15‘000-Tweet-Korpus von *SemEval 2013*. Ein weiteres ist das *EmpaTweet*-Korpus von Roberts et al. [2012]. Es besteht aus 7000 Tweets, die zu spezifischen Themen wie z.B. Weihnachten oder die Unruhen in Ägypten gesammelt wurden [Roberts et al., 2012, 3807-3808]. Zur Annotation werden Ekmans sechs Basisemotionen und zusätzlich die Emotion „Liebe“ verwendet. Die Auswahl verschiedener Themen soll dafür sorgen, dass im Korpus für alle Emotionen genügend Beispiele vorhanden sind [3807].
Weil für die Sentimentanalyse für Mikroblogging nur wenige Ressourcen vorhanden sind, sind viele Forscher dazu übergegangen, selbst automatisch Korpora zu erstellen. Dies geschieht z.B. mit Hilfe von Hashtags, wie bei Mohammad and Kiritchenko [2014]. In einem ersten Schritt sammelten die Autoren Tweets, die Hashtags mit Ekmans sechs Basisemotionen enthielten (also z.B. „#sadness“) [7]. Danach entfernten sie die Hashtags aus den Tweets und verwendeten sie als Labels für die Emotionsklassen. Ein Classifier wurde nun auf diesen Daten trainiert, sodass neue Tweets ohne Hashtag einer der Basisemotionen zugeordnet werden müssen [Mohammad and Kiritchenko, 2014, 9]. Dies funktioniert erstaunlich gut, obwohl jeder Tweet aus dem Korpus von einer anderen Person mit einem Emotionshashtag „annotiert“ wurde

[10]. Werden die Daten durch die Produzenten selbst (in diesem Fall die Twitterer) mit einer Emotion annotiert, kann man diese also ohne Probleme zusammennehmen und erreicht dabei ähnliche Resultate, wie wenn ein einzelner Annotator alle Tweets selbst annotiert hätte.
Für die deutsche Twitter-Sentimentanalyse gibt es leider kaum Ressourcen. Eine Ressource für die allgemeine Sentimentanalyse ist das *MLSA*-Korpus von Clematide et al. [2012]. Es besteht aus 270 Sätzen, die manuell auf drei Ebenen klassifiziert wurden: Objektivität und Polarität für den gesamten Satz, subjektive/faktuale Sprache auf Phrasen-/Wortebene, und Sprechakte mit Frames [Clematide et al., 2012, 3551]. Zu den *lexikalischen* Ressourcen fürs Deutsche gehören zudem noch *SentiWS* von Remus et al. [2010] und *GermanPolarityCues* von Waltinger [2010]. *SentiWS* enthält 3468 Emotionswörter, die mit einem Wert zwischen -1 (negativ) und 1 (positiv) gewichtet sind. Zudem werden aus den Emotionswörtern alle möglichen Wortformen generiert und zum Lexikon hinzugefügt [Remus et al., 2010, 1168]. Für *GermanPolarityCues* werden bestehende englische Sentimentlexika mit ihren Gewichtungen automatisch ins Deutsche übersetzt, und danach mit deutschen Negationsphrasen ergänzt [Waltinger, 2010, 1639]. Für die deutsche Sentimentanalyse gibt es also nur vereinzelte Korpora und lexikalische Ressourcen. Für die Sentimentanalyse von Mikroblogging-Texten wie z.B. Tweets existieren meines Wissens gar keine deutschen Ressourcen, weder Korpora noch spezifische Wörterbücher. Es besteht hier deshalb noch ein gewisser Aufholbedarf zum Englischen.

2.5 Verwandte Forschungsansätze

In der Literatur finden sich viele Forschungsansätze, die eng mit den Themen meiner Untersuchung verknüpft sind. Verbreitet sind z.B. Anwendungen der Sentimentanalyse, welche sich mit Emotionserkennung im Allgemeinen – nicht auf Twitter reduziert – befassen. Ein beispielhafte Studie dafür stammt von Thelwall et al. [2010], in welcher die Autoren ihren Sentimentanalyse-Algorithmus *SentiStrength* präsentieren. Das Ziel von *SentiStrength* ist es, die Stärke von Emotionen in informellen Texten aus sozialen Netzwerken, Blogs, und Foren zu messen [Thelwall et al., 2010, 2544]. Die Herausforderungen, die sich dabei für Thelwall et al. ergeben, sind sehr ähnlich wie bei der Sentiment- und Emotionsanalyse für Twitter.
Ein weiteres verwandtes Gebiet ist die allgemeine Sentimentanalyse für Twitter-Daten, ohne spezifischen Fokus auf Emotionswörter. Dazu gehört z.B. der Ansatz von Bollen et al. [2011]. Dieser untersucht, inwiefern sich bedeutsame Ereignisse auf die Stimmung der Gesellschaft auswirken, und ob diese Auswirkungen auf Twitter

messbar sind. Die Ergebnisse zeigen, dass die Stimmung der Twitter-Gemeinschaft in deutlichem Zusammenhang mit Grossereignissen wie Präsidentschaftswahlen und Feiertagen steht [Bollen et al., 2011, 452].

2.6 Fragestellung

Aus dem erarbeiteten theoretischen Hintergrund leite ich nun die Fragestellung für meine Untersuchung ab. Mein Hauptziel ist, zu untersuchen, wie Emotionswörter in Tweets verwendet werden und daraus abzuleiten, wie sie für die Sentimentanalyse eingesetzt werden können. Die Leitfrage meiner Untersuchung lautet: „Wie können Emotionswörter dazu beitragen, die Sentimentanalyse von Tweets zu verbessern?“ Um dies herauszufinden, baue ich in einem ersten Schritt ein Lexikon mit deutschen Emotionswörtern auf. Diese klassifiziere ich gemäss der Basisemotionen-Theorie von Plutchik [1980]. Die Erkenntnisse aus der Klassifizierung nutze ich, um Annotationsrichtlinien zu entwerfen. In einem zweiten Schritt sammle ich Twitter-Daten, analysiere das Vorkommen der Emotionswörter in verschiedenen Kontexten und beschreibe die Schwierigkeiten, die sich für die Sentimentanalyse ergeben. Ein weiteres Ziel meiner Arbeit ist es, zu untersuchen, ob und wie für Emotionswörter in Tweets automatisch deren Basisemotion bestimmt werden kann. Ausserdem möchte ich untersuchen, was für ein Zusammenhang zwischen den in Tweets verwendeten Emotionswörtern und der Emotion des Autors besteht. Das Ziel ist dabei, auch die (Basis-)Emotion des Twitterers bestimmen zu können. Danach entwickle ich aus den gewonnenen Erkenntnissen Ideen für Verbesserungsmöglichkeiten für die Sentimentanalyse und stelle Ansätze vor, die schon in der Forschung umgesetzt wurden.

3 Daten und Methoden

Um zu analysieren, wie Emotionswörter in Tweets verwendet werden und wie sie für die Sentimentanalyse eingesetzt werden können, brauche ich zwei Arten von Daten. Erstens benötige ich dazu möglichst viele Emotionswörter. Um sie zu sammeln, ist ein Lexikon gut geeignet. Zweitens benötige ich Daten aus Twitter, also Tweets, in denen die Emotionswörter vorkommen. Auch davon ist eine grössere Menge hilfreich. Nachfolgend beschreibe ich, wie ich diese Ressourcen aufgebaut habe und welche Herausforderungen sich dabei ergaben.

3.1 Aufbau des Emotionslexikons

In einem ersten Schritt habe ich ein Lexikon aus Emotionswörtern – nachfolgend Emotionslexikon genannt – aufgebaut. Darin wird jedes Emotionswort mit seiner Wortart (Nomen, Adjektiv, Verb) annotiert und in eine Emotionsklasse eingeordnet.

Wie in Abschnitt 2.3 beschrieben, existieren verschiedene Theorien über die menschlichen Basisemotionen und verschiedene Ansätze, um diese zu messen. Einige Studien aus der Sentimentanalyse-Forschung, die sich mit Emotionen in Tweets beschäftigt, stützen sich auf die Theorie der sechs Basisemotionen von Ekman [1992]. Dazu gehören z.B. die Arbeiten Roberts et al. [2012], Mohammad and Kiritchenko [2014], und Mohammad [2012]. Eher wenige Studien, z.B. die von Suttles and Ide [2013], stützen sich auf Plutchiks Theorie der acht Basisemotionen [Plutchik, 1980]. Andere verwendete Emotionstheorien in der Sentimentanalyse für Twitter sind z.B. das in der psychologischen Praxis häufig angewandte Profile of Mood States (POMS) von McNair et al. in der Studie von Bollen et al. [2011].

Für meine Arbeit verwende ich die Emotionstheorie von Plutchik [1980]. Der Grund dafür ist, dass sich komplexere Emotionen mithilfe Plutchiks dreidimensionalem Modell in einfacher Art auf die acht Basisemotionen zurückführen lassen. Entweder ist eine komplexe Emotion eine stärkere oder schwächere Variante einer der acht Basisemotionen, oder sie ist eine Mischung aus zwei Basisemotionen. *Langeweile* zum Beispiel ist eine schwächere Ausprägung der Basisemotion *Ekel*, während *Liebe* sich

aus den zwei Basisemotionen *Freude* und *Vertrauen* zusammensetzt. Dies ist für eine Annotation von Emotionswörtern insofern nützlich, da jedes Wort mit einer oder maximal zwei Basisemotionen annotiert werden kann. Die Auswahl an Kategorien wird beschränkt, gleichzeitig können komplexere Emotionen einfach eingeordnet werden. Somit sollte sich der zeitliche und „intellektuelle“ Aufwand für die Klassifizierung der Emotionswörter im Rahmen halten. Natürlich bringt diese Reduktion auf acht Emotionen auch Nachteile. Eine Einordnung in 33 Emotionskategorien, wie dies Hobbs and Gordon [2011] tun, würde dem menschlichen Spektrum der Emotionen wohl gerechter werden. Für den Rahmen dieser Arbeit sind Plutchiks Kategorien aber ausreichend detailliert. Um die Emotionswörter als Lexikon zu speichern, entschied ich mich für Excel. Mehrere Gründe sprechen für die Verwendung von Excel als Datenformat:

- Übersichtliche Darstellung der Daten
- Unkompliziertes Ergänzen und Annotieren der Daten
- Einfache Extraktion und Weiterverarbeitung der Daten
- Nützliche Funktionen wie Sortieren, Zählen, Markieren.

Für den Aufbau des Emotionslexikons konnte ich Teile des bereits bestehenden Polaritätslexikons von Klenner et al. [2014] nutzen. Dieses Polaritätslexikon enthält ca. 5600 deutsche Nomen und Adjektive, dazu ca. 300 Verben. Diese sind unter anderem nach ihrer Polarität (positiv, negativ, neutral) und nach ihrer Art der Polarität (faktual, moralisch, emotional) klassifiziert [Klenner et al., 2014, 2]. Daraus extrahierte ich die ca. 400 emotionalen Nomen und Adjektive als Basis für mein Emotionslexikon. Danach ergänzte ich diese Basis mit verschiedenen Quellen: mit der Liste der Emotionswörter von [Kemmler et al. 1991], mit durch Brainstorming und Wörterbuch-Lesen gefundenen Wörtern und mit zusätzlichen Wortlisten aus dem Internet (ABC [2014], Kluge Liste [2014], Medienwerkstatt Mühlacker [2009]). Zusätzlich fügte ich noch Emoticons hinzu.
Eine Unklarheit, die beim Hinzufügen von neuen Wörtern zum Emotionslexikon aufkam, ist der Umgang mit Partizipien. Die Problematik dabei ist folgende: „Einige Partizipien sind als Adjektive so selbstständig, dass ein zugrundeliegendes Verb gar nicht existiert oder eine vollständig andere Bedeutung hat“ [Wikipedia, 2014]. Deshalb habe ich folgende Entscheidungsmaxime formuliert:

- Wenn ein Partizip als Adjektiv verwendet wird, weil es keine Adjektivform zu diesem Verb gibt, wird es ins Lexikon aufgenommen.

Beispiele sind z.B. *eingeschnappt* und *nervenaufreibend*. Diese Wörter existieren

nur in der Partizipform, es gibt weder Verb noch Adjektiv dazu. Allerdings gibt es einige Wörter im Lexikon, die in beiden Formen – Adjektiv und Partizip – vorkommen, z.B. *erschüttert/erschütternd*. Die Frage, ob für alle Emotionswörter, wo dies möglich ist, beide Varianten ins Lexikon aufgenommen werden sollten, ist nicht abschliessend geklärt.

Nachdem eine erste Version des Lexikons stand, erweiterte ich es laufend, sodass es zum Schluss 951 Wörter enthält, davon 292 Nomen, 360 Adjektive, 169 Verben, und 130 Emoticons. In einem nächsten Schritt annotierte ich jedes Wort mit einer oder maximal zwei der acht Basisemotionen von Plutchik. Dazu benötigte ich die deutschen Übersetzungen der Emotionen *joy, sadness, anger, fear, trust, disgust, surprise*, und *anticipation* aus dem Englischen Original. Obwohl dies nur eine Formalität zu sein schien, ergaben sich einige Unklarheiten. Da keine deutschen Übersetzungen von Plutchiks Werken existieren, haben verschiedene Autoren die Begriffe unterschiedlich übersetzt. Besonders bei *fear, trust* und *disgust* scheint es Uneinigkeit zu geben, da die Übersetzungen oft zwischen „Angst“ oder „Furcht“, „Vertrauen“ oder „Akzeptanz“ und „Abneigung“ oder „Ekel“ variieren. Für *fear* habe ich das Wort „Furcht“ gewählt, weil mir „Angst“ bereits zu stark schein. *Trust* habe ich mit „Vertrauen“ übersetzt, weil laut Plutchik bereits die schwächere Emotion von *trust* „Akzeptanz“ (*acceptance*) heisst. Für *disgust* fand ich „Ekel“ die passendere Übersetzung, da mir „Ablehnung“ etwas zu schwach scheint. Da es keine offizielle Übersetzung gibt, wird jeder selbst für sich die passende Wahl begründen müssen. Grundsätzlich ist vor allem wichtig, dass alle Annotatoren einheitliches Vokabular verwenden.

Weitere Probleme beim Annotieren können sich ergeben, wenn Emotionswörter schwierig zu klassifizieren sind, und wenn dem Annotator/der Annotatorin die Wortbedeutung nicht klar ist. Eine Entscheidung für eine oder zwei Emotionen zu treffen fällt zum Beispiel bei Wörtern wie *berührt* schwer. Andere Wörter wie z.B. *indigniert* können Schwierigkeiten bereiten, weil sie aufgrund mangelnder Bekanntheit falsch klassifiziert werden. In solchen Fällen erwies sich die Konsultation des Duden [1] als sehr hilfreich, um mich für eine Emotion zu entscheiden. Die Idee, bei Unklarheiten zwei Emotionen zuzuordnen, hat Vor- und Nachteile. Einerseits hilft dies, schwierig zu klassifizierende Emotionswörter einzuordnen. Andererseits wird so auch der Spielraum für Interpretationen grösser, was die Sentimentanalyse schwieriger macht. Schaut man sich die Verteilung der Emotionswörter meines Lexikons in die Emotionsklassen in Abbildung 5 an, so fällt auf, dass Unterschiede in den Anzahlen bestehen (Emoticons wurden nicht gezählt).

[1] http://www.duden.de

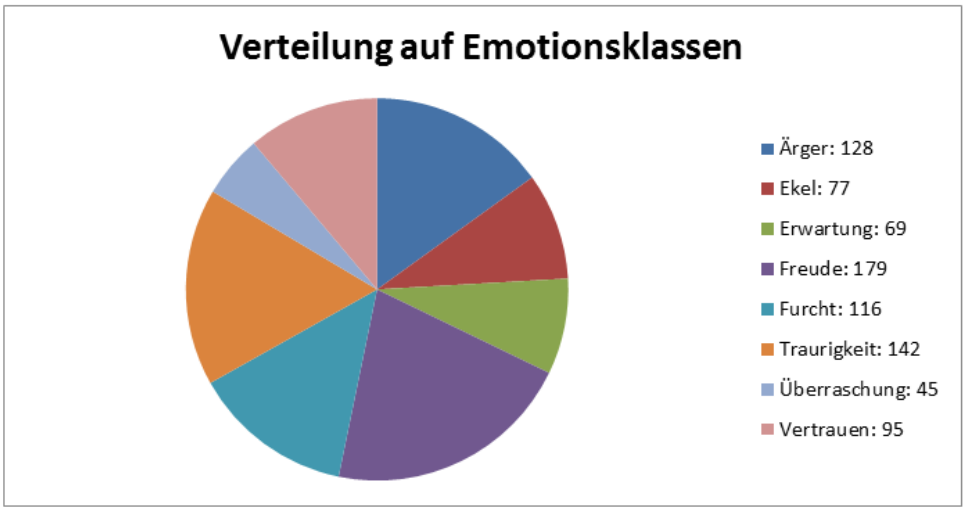

Abbildung 5: Verteilung der Emotionswörter auf Emotionsklassen

Um Emotionen wie *Freude, Traurigkeit*, und *Ärger* auszudrücken, scheint es also sehr viele Wörter zu geben. Im Gegensatz dazu gibt es für Emotionen wie *Überraschung* und *Erwartung* nur sehr wenige Wörter. Zählt man die „positiven“ Emotionen (*Erwartung, Freude, Überraschung, Vertrauen*) und die „negativen“ Emotionen (*Ärger, Ekel, Furcht, Traurigkeit*), so fällt auf, dass eine leichte Tendenz zu negativen Emotionswörtern besteht (463 negative vs. 388 positive).
Um die Verlässlichkeit der Annotation zu untersuchen, annotierte ich zuerst alle Emotionswörter selbst. Dann wurden die Nomen und Adjektive aus dem Emotionslexikon in drei ausgewogene Teile aufgeteilt, sodass drei Sets à je ca. 100 Nomen und 100 Adjektiven entstanden. Diese Sets wurden an 18 Studenten verteilt, die im Rahmen einer Übungsaufgabe für eine Vorlesung ihren Teil (also etwas weniger als 200 Wörter) annotierten [2]. Die Studenten erhielten dabei keine Instruktionen ausser der Anweisung, jedes Emotionswort mit einer oder maximal zwei von Plutchiks Basisemotionen zu annotieren. Mit diesen Daten konnte ich im Anschluss das Inter-Annotator-Agreement berechnen. Um gleiche Bedingungen zu haben, benutzte ich für jedes Set die Annotationen von fünf Studenten, dazu meine Annotation. Insgesamt berechne ich also die Übereinstimmung zwischen sechs Personen. Dabei unterscheide ich zwischen zwei Szenarien: im ersten Fall wird nur die erstgenannte Basisemotion pro Emotionswort gezählt; im zweiten Fall werden, falls zwei Annotationen pro Emotionswort vorhanden sind, beide gezählt. Die Wahrscheinlichkeit, dass die Annotatoren übereinstimmen, ist also für das zweite Szenario höher, weil jede Person maximal zwei Angaben machen kann. Eine Übereinstimmung ergibt sich,

[2] An dieser Stelle herzlichen Dank an die Studenten für ihre wertvolle Arbeit.

wenn alle Annotatoren einmal die gleiche Basisemotion annotiert haben, egal ob an erster oder an zweiter Stelle. In Tabelle 1 sind die Kappa-Werte (nach Cohen) für die einfache Annotation dargestellt, in Tabelle 2 die Kappa-Werte für die (maximal) zweifache Annotation.

	Nomen	**Adjektive**
Gruppe 1	0.31	0.41
Gruppe 2	0.20	0.26
Gruppe 3	0.24	0.18

Tabelle 1: Kappa-Koeffizient bei Zuordnung einer Basisemotion, sechs Annotatoren

	Nomen	**Adjektive**
Gruppe 1	0.47	0.63
Gruppe 2	0.31	0.28
Gruppe 3	0.36	0.29

Tabelle 2: Kappa-Koeffizient bei Zuordnung von ein bis zwei Basisemotionen, sechs Annotatoren

Wie man in den Tabellen 1 und 2 sehen kann, ist die Übereinstimmung zwischen den Annotatoren generell recht tief. Wenn zwei Basisemotionen zur Annotation verwendet werden dürfen, ist die Übereinstimmung schon etwas besser (Tabelle 2). Natürlich gibt es einige Einschränkungen in meiner Untersuchung, welche auch die tiefen Werte erklären können. Ein erster Punkt ist, dass keine Möglichkeit zur genauen Anleitung der Annotatoren bestand. Mit ein paar Instruktionen, wie z.B. im Zweifelsfalle vorzugehen ist, und der Möglichkeit, Tests zum Üben durchzuführen und zu besprechen, wäre bestimmt eine höhere Übereinstimmung möglich gewesen. Ein weiterer Punkt ist, dass ich die Motivation der Annotatoren nicht überprüfen konnte. In vereinzelten Datensätzen waren z.B. Schreibfehler zu finden, was auf eine eher tiefere Motivation oder Konzentration auf die Aufgabe schliessen lässt. Ein dritter Punkt ist, dass eine Übereinstimmung zwischen sechs Personen ein recht ambitioniertes Ziel ist. Hätte ich nur das Agreement zwischen zwei Personen gemessen, würden gewisse Paarungen um einiges höhere Kappa-Koeffizienten ergeben.
Das Messen des Kappa-Koeffizienten hat mir aber auch interessante Einsichten gegeben, wie andere Menschen Emotionswörter beurteilen. „Emotions are perceived differently by individuals, partly because of their life experiences and partly because of personality issues [...] and gender“, beschreiben [Thelwall et al., 2010, 2548] die Schwierigkeit treffend. Ein kleines Beispiel zur Veranschaulichung: Das Adjektiv „peinlich“ annotierten ich und eine weitere Person mit *Ärger*, drei Personen mit *Traurigkeit*, und eine Person mit *Furcht*. Für eine nächste Annotationsaufgabe würde

ich auf jeden Fall eine Tabelle vorbereiten, die nur eine Auswahl von vorgegebenen Worten zulässt. Damit können Fehler reduziert werden. Zudem würde ich, wenn möglich, eine Trainingsphase für die Annotatoren einplanen.
Auch interessant ist es, die Verteilung der Basisemotionen in meinem Emotionslexikon mit der Basisemotion-Annotation eines bestehenden Twitter-Korpus zu vergleichen. Wenn es unterschiedlich viele Wörter gibt, welche die Basisemotionen beschreiben, dann müsste es ja theoretisch auch eine entsprechende Verteilung der Tweets geben, die mit dieser Basisemotion annotiert sind. Dazu vergleiche ich meine Emotionsliste mit einem Twitter-Korpus von Suttles and Ide [2013], welches automatisch nach Plutchiks acht Basisemotionen annotiert wurde und ca. 3 Millionen Tweets enthält [127]. (Ob für eine Untersuchung wirklich drei Millionen Tweets notwendig sind, sei dahingestellt.) Der Vergleich zeigt grosse Ähnlichkeiten in der Verteilung auf Emotionsklassen. Die Tweets im Korpus von Suttles and Ide [2013] wurden am häufigsten mit *Freude* und am vierthäufigsten mit *Furcht* annotiert [128], dies stimmt mit meinem Emotionslexikon überein, in dem es am meisten Wörter für *Freude* und am viertmeisten Wörter für *Furcht* gibt. Auch *Traurigkeit* und *Ekel* waren bei beiden nur um einen Rang unterschiedlich platziert. Nur für *Ärger* und *Vertrauen* waren unsere Annotationen grundsätzlich verschieden, Suttles and Ide [2013] fanden viel mehr Tweets mit *Vertrauen*, während ich weit mehr Tweets mit *Ärger* fand. Aber so wie bei meinem Emotionslexikon am wenigsten Wörter mit *Erwartung* und *Überraschung* annotiert waren, waren auch bei Suttles and Ide [2013] am wenigsten Tweets mit *Erwartung* und *Überraschung* annotiert [128]. Es scheint also auch sprachübergreifend Ähnlichkeiten zu geben, wie viele Wörter jede Basisemotion beschreiben.

3.2 Gewinnung und Struktur der Twitter-Daten

Um für Forschungszwecke eine grosse Menge an Twitter-Daten zu gewinnen, gibt es verschiedene Möglichkeiten. Eine Möglichkeit ist die Verwendung der *Twitter API* [3]. Eine andere Möglichkeit ist die Verwendung von kostenpflichtigen Services wie z.B. *Tweet Archivist* [4]. Ich habe mich für das Tool *Pattern* [5] von Smedt and Daelemans [2012] entschieden. *Pattern* ist ein Package zur Verwendung in Python und enthält Funktionalitäten für Web Mining, für die Verarbeitung natürlicher Sprache, für Machine Learning und für die Analyse von Netzwerken [Smedt and Daelemans, 2012, 2063]. Um die Daten in der Form von Excel-Tabellen abzuspeichern, habe

[3] https://dev.twitter.com/streaming/overview
[4] http://www.tweetarchivist.com/
[5] http://www.clips.ua.ac.be/pattern

ich zusätzlich das Python-Modul *XlsxWriter* [6] verwendet, welches es erlaubt, mit Excel kompatible Tabellen zu erstellen. Mithilfe dieser zwei Tools habe ich ein kleines Programm geschrieben, welches eine gewünschte Anzahl aktueller Tweets, die eine bestimmte Zeichenfolge enthalten, von Twitter herunterlädt und in einer Excel-Tabelle speichert.
Um eine erste Datengrundlage für meine Untersuchungen zu schaffen, erstellte ich ein Korpus mit 800 Tweets. Dazu habe ich für jede von Plutchiks acht Basisemotionen (*Freude, Vertrauen, Furcht, Überraschung, Traurigkeit, Ekel, Ärger, Erwartung*) aus dem Emotionslexikon 20 Emotionswörter ausgewählt (je zehn Nomen und zehn Adjektive). Für jedes Emotionswort wurden dann fünf Tweets gesammelt. Da das Emotionslexikon natürlich mehr als 20 Wörter pro Emotion enthält, war eine Selektion notwendig. Um eine Auswahl von Emotionswörtern zu treffen und sie begründen zu können, habe ich folgende Selektionskriterien aufgestellt:

1. Jede der Basisemotionen von Plutchik soll als Nomen im Korpus enthalten sein (also z.B. *Traurigkeit* als Stichwort, um Tweets zu finden).

2. Wenn von einem Wort sowohl ein Nomen als auch ein Adjektiv existiert (z.B. *Zorn – zornig, Betroffenheit – betroffen*), so soll möglichst nur eine der beiden Formen für das Tweet-Mining verwendet werden. Dies hat zum Zweck, dass das Korpus möglichst viele unterschiedliche Wörter enthält.

3. Gibt es mehrere Emotionswörter, die sich nur gering in ihrer Bedeutung unterscheiden, so soll nur eines davon für die Datensammlung verwendet werden. Beispiele dazu sind: *ehrerbietig – ehrfürchtig, Missfallen – Missmut.*

4. Da viele Äusserungen auf Twitter umgangssprachlicher Natur sind, soll jeweils die „einfachere“ bzw. umgangssprachlichere Version von Wörtern in Tweets gesucht werden (z.B. *Begeisterung* anstatt *Euphorie*, *Skepsis* anstatt *Skeptizismus*).

5. Wörter, für die keine Tweets gefunden wurden, werden durch andere ersetzt.

Zu einigen Wörtern konnte ich beim Twitter-Mining keine Tweets finden. Beispiele dafür sind *Depressivität, Beklommenheit, zukunftsfroh*, und *ehrerbietig*. Schaut man sich den oft ungezwungenen und umgangssprachlichen Schreibstil von Twitterern an, ist dies keine Überraschung. Es gibt jedoch genügend andere Emotionswörter, um diese Lücken zu füllen. Nachdem ich das 800-Tweet-Korpus für meine Analysen verwendet hatte, sammelte ich weitere Daten, um Spezialfälle zu untersuchen. Hier verwendete ich zum Teil auch die normale Twitter-Suchfunktion [7], die einige

[6] https://xlsxwriter.readthedocs.org/
[7] https://twitter.com/search

grundlegende Operatoren bietet. Mittels AND und OR kann man bis zu 36 Wörter verknüpfen, welche in den Tweets gesucht werden. Zum einen sammelte ich deshalb Tweets, die zwei in ihrer Polarität entgegengesetzte Emotionswörter enthalten, wie z.B. *Vertrauen* (eher positiv) und *skeptisch* (eher negativ). Dafür unterteilte ich meine Liste von 100 ausgewählten Emotionswörtern in positive und negative Wörter. Zum anderen suchte ich nach Tweets, welche je mindestens ein Emoticon und ein Emotionswort enthalten, und nach Tweets, welche Hashtags gefolgt von Emotionswörtern (z.B. #Freude) enthalten.
Zum Schluss dieses Abschnitts möchte ich noch kurz auf ethische Überlegungen beim Datensammeln eingehen. Obwohl Tweets öffentlich sind, so stammen sie meist von Privatpersonen, denen nicht unbedingt bewusst ist, dass die von ihnen produzierten Daten für die Forschung oder für kommerzielle Zwecke verwendet werden könnten. Wie Neuhaus and Webmoor [2012, 45] beschreiben, gilt: „Informed consent and anonymizing data are the primary technical and procedural steps for minimizing risks to human subjects in non-medical research contexts.“ Wenn jedoch, wie bei der Twitter-Sentimentanalyse, unzählige “Studienteilnehmer” vorhanden sind, dann sind diese Anforderungen kaum erfüllbar. Problematisch bei Twitter ist zudem, dass nicht nur die Inhalte von Tweets gesammelt werden können, sondern auch Informationen über deren Produzenten, wie z.B. Username, Datum und Uhrzeit, Geschlecht, Sprache, und Ort [Neuhaus and Webmoor, 2012, 48]. Bei der Lektüre aktueller Studien fällt auf, dass in der Sentimentanalyse-Forschung diesem Thema keine Beachtung geschenkt wird. Dies ist z.B. in der linguistischen Forschung von Online-Kommunikation anders. Ein Grund für diesen Unterschied könnte sein, dass in computerlinguistischen Ansätzen der Fokus eher auf der Sammlung und Weiterverwendung der Daten liegt, und nicht auf deren detaillierten inhaltlichen Analyse. Beim Erstellen meines 800-Tweet-Korpus habe ich wenig Rücksicht auf die Privatsphäre der Nutzer genommen, da ich zu jedem Tweet auch den Usernamen und das Datum und die Uhrzeit extrahiert habe. Im Nachhinein stellte ich fest, dass diese zusätzlichen Informationen gar nicht notwendig gewesen wären. Um eine Art „verantwortungsvolle Forschung“ zu befolgen, werde ich im Rahmen dieser Arbeit deshalb nur die Tweets publizieren, nicht aber Details zu Produzenten und Produktionsumständen.

4 Analyse und Implikationen für die Sentimentanalyse

Nach dem Aufbau des Emotionslexikons und dem Sammeln von Twitter-Daten folgt nun der Hauptteil meiner Arbeit. Dabei untersuche ich die gesammelten Tweets auf die Verwendung von Emotionswörtern und befasse mich mit den Auswirkungen, die sich für die Sentimentanalyse ergeben. Als Resultat dieser Analyse mache ich Vorschläge, wie die Erkennung von Emotionen in der Sentimentanalyse verbessert werden kann. Ich konzentriere mich dabei auf folgende Szenarien: 1) Tweets, die ein Emotionswort enthalten, 2) Tweets, die mehrere Emotionswörter enthalten, 3) Emotionswörter, die zusammen mit Emoticons/Emojis verwendete werden, 4) Emotionswörter, die zusammen mit Hashtags verwendet werden.

4.1 Probleme mit Emotionswörtern für die Sentimentanalyse

Um einen ersten Überblick über die Verwendung von Emotionswörtern in Tweets zu erhalten, untersuchte ich das in Abschnitt 3.2 beschriebene Twitter-Korpus von 800 Tweets. Dabei stiess ich auf eine Vielfalt von Problemen, welche im Zusammenhang mit Emotionswörtern und Sentimentanalyse in Tweets auftauchen können. Nachfolgend stelle ich diese Schwierigkeiten vor und beschreibe, welche Lösungsansätze dafür geeignet sein könnten und ob es schon praktische Umsetzungen davon gibt.
Ein Problem, das im Zusammenhang mit Emotionswörtern in Tweets oft vorkommt, ist die falsche Lesart von Adjektiven. Dabei werden für Adjektive, die eine Emotion beschreiben, wie z.B. *aufgedreht*, fälschlicherweise Tweets zurückgeliefert, in denen das Emotionswort als nicht-emotionales Verb verwendet wird. Dies ist in (1) gut erkennbar:

(1) Die Lautstärke voll aufgedreht ist die Musik immer noch zu leise.

Die Problematik ist hier also, dass der Tweet kein Sentiment enthält/neutral ist,

obwohl er ein Emotionswort beinhaltet. Gerade Adjektive werden als sehr wichtig für die Sentimentanalyse angesehen, weil der Gebrauch von Adjektiven auf Subjektivität hindeutet [Shelke et al., 2012, 33]. Für solche Fälle braucht es deshalb eine Lösung um herauszufinden, ob es sich um den emotionalen oder sachlichen Gebrauch eines Wortes handelt.
Ein ähnliches Phänomen lässt sich in (2) beobachten. Wie man hier sehen kann, werden manche Emotionswörter auch als normale Eigennamen verwendet, wie z.B. *Kummer* oder *Angst*:

(2) #Lindenfels Große Mehrheit wählt Gerald Kummer zum Direktkandidaten - Echo-online: Echo-onlineGroße ... http://t.co/82UW0g74fy #Odenwald

Auch Tweet (2) würde kaum als Sentiment tragend beschrieben, obwohl er ein Emotionswort enthält. Eine mögliche Lösung für die in (1) und (2) gesehenen Schwierigkeiten könnte die Verarbeitung der Tweets mit Part-of-Speech-Tagging sein. Damit könnten etwa Eigennamen von normalen Nomen unterschieden werden. In der Sentimentanalyse für andere Domänen werden oft PoS-Tags verwendet, weshalb sie vielleicht auch für Twitter-Daten von Nutzen sind [Kouloumpis et al., 2011, 538].
Eine vielzitierte Untersuchung, die sich unter anderem mit der Verwendung von PoS-Tags für die Sentimentanalyse von Twitter beschäftigt, stammt von Pak and Paroubek [2010]. Darin analysieren die Autoren die Verteilung von PoS-Tags in subjektiven und objektiven Tweets und kommen zu einigen nützlichen Erkenntnissen. So werden etwa in subjektiven Texten (Tweets) mehr Adjektive im Superlativ, in objektiven Texten hingegen mehr Adjektive im Komparativ verwendet [1323]. Zudem sind Sentiment tragende Aussagen oft mit Verben in der ersten Person, oder mit Verben in Grundform zusammen mit einem Modalverb, und im Präteritum geschrieben [1323]. Deshalb kommen die Autoren zum Schluss: “authors use syntactic structures to describe emotions or state facts. Some POS-tags may be strong indicators of emotional text” [Pak and Paroubek, 2010, 1326]. PoS-Tags zur Verbesserung von Twitter-Sentimentanalyse scheinen also eine mögliche Lösung zu sein.
Allerdings gibt es auch Gegenstimmen: Kouloumpis et al. [2011] analysieren in ihrer Studie, wie sich verschiedene Features auf die Sentimentanalyse für Twitter auswirken. Untersuchte Features sind n-Gramme, ein Sentimentlexikon, PoS-Tags, und Abkürzungen/Emoticons/Verstärker [540]. Wie sich herausstellt, führt die Kombination aller Features, ausgenommen der PoS-Tags, zur besten Sentimentanalyse. Damit scheint das PoS-Tagging für die Twitter-Sentimentanalyse eher nicht von Nutzen zu sein. Allerdings könnte dies auch an der Genauigkeit des Taggers liegen, die aber nicht weiter untersucht wurde [Kouloumpis et al., 2011, 540].
Das Paradox, dass Tweets nicht emotional sind, obwohl sie ein Emotionswort enthalten, kommt auch oft in Tweets vor, die ich „unpersönlich“ nenne. Diese un-

persönlichen Tweets enthalten zwar ein Emotionswort, das auch tatsächlich ein Gefühl beschreibt. Allerdings bezieht sich die Emotion nicht auf den Autor des Tweets, sondern auf jemand anderen. Ein Beispiel für solche unpersönlichen Tweets ist (3):

(3) Neue Mona Lisa. Einfach fantastisch. Sie begeistert die Welt http://t.co/JEP Ad3NUEO

Zwar ist in (3) das emotionale Verb begeistern enthalten, aber es bezieht sich nicht auf den Urheber des Tweets. Diesen Tweet würde man also nicht als eine „begeisterte“ Äusserung einstufen. Um dieses Problem der unpersönlichen Aussagen zu lösen, muss man sicherstellen, dass ein Tweet auch wirklich eine emotionale Aussage einer Person über sich selbst enthält. Wie sich bei der Analyse gezeigt hat, scheint ein Emotionswort nur dann als persönlich zu gelten, wenn es zusammen mit einem Pronomen wie *ich, mein/e, mir, wir, uns* vorkommt. Die Regel könnte also lauten: Nur zusammen mit einem Personal-, Possessiv-, oder Reflexivpronomen in der 1. Person Singular oder 1. Person Plural bezieht sich ein Emotionswort auf den Autor des Tweets. Ein weiterer Hinweis ist zudem, wenn zusammen mit einem Emotionswort die 1. Person Singular/Plural des Verbs *sein* im Tweet vorkommt, also *bin* bzw. *sind*. Eine Beobachtung, welche diese aufgestellten „Regeln“ unterstützt, haben auch Pak and Paroubek [2010] in der oben erwähnten Studie gemacht: In ihrem Twitter-Korpus enthielten objektive Texte mehr Nomen und Eigennamen, während in subjektiven Tweets mehr Personalpronomen vorkamen [1322]. Um das Problem der „unpersönlichen Aussage trotz Emotionswort“ zu lösen, könnte PoS-Tagging deshalb nützlich sein. Für Pronomen könnte man zudem darauf hoffen, dass die Zuverlässigkeit des Taggers höher ist.
Auch in (4) zeigt sich, dass die Regel „Emotionswort plus (spezifische) Pronomen“ Sinn macht:

(4) Richtig cooles Video, hat mich echt zum Nachdenken angeregt... *@User* EIN BESSERES LEBEN — Julien Bam: http://t.co/JZzM8wYEcD

Viele Wörter wie z.B. *anregen* können auf zwei Arten verwendet werden, von denen nur eine emotional ist („Ich regte eine Diskussion an“ vs. „Die Diskussion regte mich an“). Um von diesen zwei Lesarten die für die Sentimentanalyse passende zu finden, reicht es wiederum als Bedingung, dass auch ein Reflexivpronomen im Tweet vorkommen muss.
Eine weitere Schwierigkeit für die Sentimentanalyse ist Ironie bzw. Sarkasmus, was in vielen Tweets vorkommt. Zur Differenzierung von Ironie und Sarkasmus: Ironie wird als „feiner, verdeckter Spott“ [Duden a, 2014], Sarkasmus als „beissender, verletzender Spott“ [Duden b, 2014] definiert. Mit Ironie wird die Technik bezeichnet,

dass durch Aussage des Gegenteils ausgedrückt wird, was tatsächlich gemeint ist. Im Sarkasmus kann diese Technik der Umkehrung verwendet werden, muss aber nicht [Sarkasmus, 2014]. In der (computerlinguistischen) Literatur werden diese Begriffe allerdings nicht differenziert und oft synonym verwendet.
Wie Liu [2012, 44] berichtet, wird Sarkasmus weniger in Produkt- oder Servicebewertungen verwendet, dafür umso mehr in politischen Diskussionen und Kommentaren. Benutzt ein Twitterer Ironie, wird die Bedeutung eines Emotionswortes genau ins Gegenteil gedreht, wie in (5) gut erkennbar ist:

(5) Montag. Und der Server ist hier irre lahm. Bin begeistert. #office

Die erste Lösung, die einem für das Ironie-Problem einfällt, ist die „Umkehrung“ der Annotation. Für das Beispiel (5), in dem das Emotionswort *begeistert* mit *Freude* annotiert wird, gälte also genau die gegenteilige Emotion, *Traurigkeit*. Dieses Vorgehen ist allerdings ziemlich problematisch. Klassifiziert ein menschlicher Annotator Tweet (5) nach den Basisemotionen, so würde er hier wohl eher *Ärger* auswählen. Ironie muss also nicht unbedingt die Umkehrung in die in Plutchiks Modell der Emotionen genau gegenüberliegende Basisemotion bedeuten. Ausserdem muss, bevor diese Umkehrung überhaupt stattfinden kann, zuerst ein Tweet als ironisch erkannt werden. Einen Algorithmus zu entwerfen, der erstens Ironie erkennt und zweitens mithilfe des Kontexts in jedem Fall neu entscheidet, welche Emotion am besten geeignet ist, scheint mir im Moment noch nicht möglich zu sein. Ein Versuch, dieses Problem zu lösen, wurde von González-Ibáñez et al. [2011] durchgeführt. Dabei erstellten sie ein Korpus aus positiven, negativen und sarkastischen Tweets, die sie mittels Hashtags wie „#sarcasm“ sammelten [581]. Um sarkastische Tweets automatisch klassifizieren zu können, verwendeten die Autoren lexikalische Features wie ein Lexikon, WordNetAffect, Interjektionen und Satzzeichen sowie pragmatische Features wie Emoticons und direkte Ansprachen anderer Nutzer mittels „@“ [582-83]. Die Ergebnisse zeigen allerdings, wie schwierig das Erkennen von Sarkasmus in Tweets ist. Nicht nur die automatische Erkennung war wenig erfolgreich, auch die manuelle Klassifizierung von sarkastischen Tweets durch menschliche Annotatoren ergab wenig Übereinstimmung. Der Grund dafür war laut Annotatoren der fehlende Kontext und die Kürze der Tweets [González-Ibáñez et al., 2011, 585]. Diese Untersuchung bestätigt deshalb meine Einschätzung der Schwierigkeit dieses Unterfangens.
Ein weiteres Problem, das sich bei der Analyse der Tweets gezeigt hat, sind „zukünftige“ Emotionen. Dabei kommt zwar ein Emotionswort im Tweet vor, das auch die Emotion des Tweet-Autors beschreibt. Allerdings bezieht sich dieses auf einen zukünftigen Gefühlszustand, der noch nicht stattgefunden hat. Dies ist z.B. in (6) gut erkennbar:

(6) @*User* Gib mir ein Lebenszeichen und ich bin dir ewig dankbar.😉💜☺

Im Zusammenhang mit diesem Phänomen kommen oft grammatikalische Fehler vor, z.B. in der Verwendung der Zeit. Mit konsequenter Anwendung der Zeitformen müsste (6) eigentlich heissen: „Gib mir ein Lebenszeichen und ich werde dir ewig dankbar sein“. Eine perfekte Rechtschreibung und Grammatik ist auf Twitter aber natürlich nicht zu erwarten. Die Schwierigkeit von Tweets wie (6) besteht darin, dass der Tweet zwar ein „echtes“ und auf den Autor bezogenes Emotionswort enthält, der Tweet insgesamt aber nicht als emotional eingeordnet werden soll. Schliesslich ist die Emotion an eine Bedingung geknüpft, die noch nicht erfüllt ist. Ähnlich wie diese „echten, aber zukünftigen“ Emotionen sind hypothetische Aussagen wie in (7) einzuordnen:

(7) *@_User_ @User* Ich wäre untröstlich. Für Jill natürlich.

Mit einem klassischen Sentimentanalyse-Algorithmus würde dieser Tweet aufgrund des Emotionswortes *untröstlich* mit der Basisemotion *Traurigkeit* annotiert. Dies ist für (7) natürlich falsch, denn der Autor des Tweets ist im Moment nicht untröstlich, er formuliert nur eine Hypothese.
Ein zusätzliches Problem bei Tweets wie (6) und (7) ist, dass nicht nur ein Emotionswort vorkommt, sondern auch ein Pronomen in der 1. Person Singular. Damit wird es umso schwieriger, diese zukünftigen oder hypothetischen Emotionen als nicht Sentiment tragend einzuordnen. Einen Hinweis auf die nicht reelle Existenz dieser Emotionen sind z.B. Wörter im Konjunktiv II wie „wäre“. Allerdings führt der informelle Schreibstil auf Twitter oft zum inkorrekten Gebrauch der Zeiten, wie etwa in (6). Damit fallen auch Hinweise wie der Gebrauch des Konjunktivs II weg.
Diese letzten zwei genannten Probleme (zukünftige und hypothetische Emotionen) sind Merkmale von „wenn, dann“- oder „konditionalen“ Sätzen. Solche konditionalen Sätze bestehen normalerweise aus zwei Teilen, die voneinander abhängen: einem Bedingungs- und einem Folgesatz [Liu, 2012, 43]. Die momentan einzige Studie, die sich mit Sentimentanalyse von Konditionalsätzen befasst, stammt von Narayanan et al. [2009]. Das in (6) und (7) erkennbare Problem sehen sie genauso: “A conditional sentence may contain many sentiment words or phrases, but express no opinion” [Narayanan et al., 2009, 181]. In ihrer Studie untersuchen die Autoren verschiedene Ansätze, um Sentimentanalyse für Konditionalsätze durchzuführen. Dabei gehen sie davon aus, dass zumindest ein Teil eines Konditionalsatzes eine Meinung über z.B. ein Produkt enthält, welche durch ihren Algorithmus als positiv, negativ, oder neutral bewertet werden soll. Ein Beispiel für solch einen Satz ist „If your Nokia phone is not good, buy this great Samsung phone” [Narayanan et al., 2009, 181]. In diesem Satz enthält der zweite Teil eine positive Meinung über ein Produkt. Ohne das „if“ zu Beginn würde auch der erste Teil eine Meinung über ein Produkt enthalten. Daran kann man gut erkennen, wie wichtig die detaillierte Analyse von Konditio-

nalsätzen ist. Ein Konditionalsatz lässt sich dabei nicht nur an einem „if“ erkennen, sondern auch an Verbindungswörtern wie „unless“, even if“, „until“, „only if“, etc. [181].
Ein Ansatz wie der von Narayanan et al. [2009] ist zwar von Nutzen, weil er auf die Eigenschaften von Konditionalsätzen eingeht und damit die Sentimentanalyse verbessern kann. Allerdings löst er immer noch nicht das Problem von Konditionalsätzen, die zwar ein Emotionswort enthalten, aber keine Emotion ausdrücken. Für die gesonderte Behandlung solcher Sätze scheint es leider noch keine Lösung zu geben.
Die Analyse der Twitter-Daten brachte eine weitere Quelle von Problemen für die Sentimentanalyse zutage, nämlich *Shifter* wie „Freiheit von“, „Mangel an“, „frei von“ in Kombination mit einem Emotionswort. Dabei geschieht ähnliches wie bei der Ironie: die Bedeutung des Emotionswortes wird umgekehrt, wie dies z.B. in (8) der Fall ist:

(8) Aus Mangel an Enthusiasmus an Kommunikation schreibe ich gerne mal vier Leuten die gleiche Nachricht.

Im Normalfall würde die Sentimentanalyse aufgrund des Emotionswortes *Enthusiasmus* den Tweet als positiv einordnen, was natürlich falsch ist. Eine mögliche Lösung wäre deshalb, den Tweet wie bei Fällen der Ironie mit der gemäss Plutchik entgegengesetzten Basisemotion zu klassifizieren. Allerdings herrscht hier auch wieder dieselbe Problematik, denn das Gegenteil von *Freude* (der Emotion vom Wort Enthusiasmus) ist *Traurigkeit*, die passende Annotation von (8) wäre allerdings eher *Ekel* (die vorherrschende Emotion im Tweet scheint *Langeweile* zu sein, wovon *Ekel* die Basisemotion ist). Das Aufstellen von Regeln, wie ein Shifter die Emotionen beeinflusst, hängt vom Kontext ab und ist deshalb schwierig. Für die allgemeine Sentimentanalyse wurden solche Regeln allerdings schon mit Erfolg umgesetzt. **?** entwickelten zum Beispiel das „@AM-System“, welches auf einer grossen Anzahl von Kompositionsregeln aufbaut. Dabei versuchen sie auch, die Umkehrung der Polarität von Emotionswörtern vorzunehmen, z.B. für Phrasen wie „reduced enthusiasm“ [808]. Auch Neviarouskaya et al. müssen feststellen: „it is not trivial to find pairs of opposite emotions in the case of a fine-grained classification, except for ‚joy‘ and ‚sadness‘“ [808]. Um dieses Problem zu lösen, ordnen die Autoren den positiven Emotionen *Interesse, Freude*, und *Überraschung* die gegensätzliche Emotion *Traurigkeit* zu und den sechs negativen Emotionen *Ärger, Ekel, Furcht, Schuld, Traurigkeit, Scham* die gegensätzliche Emotion *Freude* zu [Neviarouskaya et al., 2010, 808]. Diese Art der Einteilung könnte für die Sentimentanalyse in Twitter übernommen werden. Es würde dann also für alle positiven Basisemotionen von Plutchik (*Erwartung, Freude, Vertrauen, Überraschung*) als Gegenteil *Traurigkeit*

gelten, und für alle negativen Basisemotionen (*Ärger, Ekel, Traurigkeit, Furcht*) die Emotion *Freude*. Wie beschrieben hängt das Bestimmen der Emotion aber immer vom Kontext ab, weshalb dieses Vorgehen vor allem als erste grobe Einteilung verwendet werden könnte.
Dasselbe Problem wie gerade beschrieben gilt natürlich auch für Verneinungen wie etwa in (9):

(9) Tröstet mich jemand über meinen Alptraum hinweg? Ich konnte nicht mal angstfrei duschen...

Durch den Shifter „nicht“ verändert sich die Bedeutung des Tweets: während „angstfrei“ mit *Vertrauen* annotiert wird, ist für „nicht angstfrei“ die Basisemotion *Furcht* angemessen. Die Behandlung von Negationen sollte im Allgemeinen weniger problematisch sein, da man diese in einer (eher kurzen) Liste sammeln und dann automatisch erkennen kann. Allerdings bleibt auch hier das Problem, auf welche „umgekehrte“ Emotion ein Algorithmus zurückgreifen sollte, wenn eine Negation erkannt wird.
Ein zusätzliches Problem in der Sentimentanalyse, das als „Umkehrung“ der bisher vorgestellten Probleme angesehen werden kann, ist der Ausdruck einer Emotion ohne den Gebrauch von Emotionswörtern. Ein Beispiel dafür ist (10):

(10) Ich hab auch gerade wieder so ein ïch bin eh nur die zweite wahl”Gefühl

Ein menschlicher Annotator würde diesen Tweet wahrscheinlich mit der Basisemotion *Traurigkeit* oder *Ärger* beschreiben. Ein Sentimentanalyse-Programm steht bei solchen Tweets aber vor einer grösseren Herausforderung. Die Formulierung mit „ich bin“ könnte vielleicht als Hinweis für Emotion im Text verstanden werden, obwohl kein Emotionswort im Tweet vorhanden ist (allerdings würde dann ein Satz wie „Ich bin kurz in die Stadt gegangen“ auch als emotional bezeichnet werden, was natürlich falsch ist). Als einzige Möglichkeit bleibt deshalb wohl nur, auch den Kontext miteinzubeziehen – ein sehr schwieriges Unterfangen. Aus Platz- und Zeitgründen werde ich mich in der vorliegenden Arbeit deshalb auf Tweets beschränken, die Emotionswörter enthalten.
Die Analyse der Tweets hat auch zu einigen allgemeinen Erkenntnissen über deren Struktur und Inhalt geführt. Generell scheint es zum Beispiel der Fall zu sein, dass substantivierte Verben und damit eher „abstrakte“ Worte wie z.B. *Vereinsamung* vor allem in News-Tweets (also Schlagzeilen) vorkommen. Andere Wörter, wie z.B. *abgeneigt*, kommen ausschliesslich in der Verneinung vor, also *nicht abgeneigt*. Natürlich haben solche Einsichten keinen direkten Einfluss auf die Sentimentanalyse, aber sie sind in sprachlicher Hinsicht interessant.
Wenn man sich nun die vorgestellten Probleme mit Emotionswörtern in Tweets ge-

nauer anschaut, so erkennt man zwischen ihnen Zusammenhänge. Diese habe ich versucht in Tabelle 3 sichtbar zu machen:

Oberklasse	**Unterklasse**
Falsche Lesart von Emotionswörtern	Nicht-emotionale Verben werden als emotionale Adjektive gelesen („aufgedreht“)
	Nachnamen werden als emotionale Nomen gelesen („Kummer“)
Emotionswort ist unpersönlich	-
Umkehrung der Emotion	Ironie/Sarkasmus
	Shifter wie „Mangel an“, „Freiheit von“, und Verneinung
Konditionale Emotion	Zukünftige Emotion
	Hypothetische Emotion
Emotion wird nicht mit Emotionswörtern ausgedrückt	-

Tabelle 3: Probleme mit Emotionswörtern in Tweets

Zudem kann man die Oberklassen der „Twitter-Sentimentanalyse-Probleme“ noch in zwei grobe Richtungen unterteilen. Viele Probleme haben mit der falschen Interpretation von Wörtern zu tun. Im Gegensatz dazu entstehen andere Probleme, weil die Bedeutung von Wörtern durch den Kontext verändert wird.

4.2 Hinweise auf Emotion in Tweets

Die Analyse der Twitter-Daten hat viele Probleme zutage befördert, welche die korrekte Sentimentanalyse im Zusammenhang mit Emotionen erschweren oder verunmöglichen. Allerdings habe ich dabei auch Hinweise entdeckt, wann ein Emotionswort mit hoher Wahrscheinlichkeit in der Art verwendet wird, wie es für die Sentimentanalyse nützlich ist. Die folgenden Beispiele veranschaulichen, was ich unter diesen Hinweisen verstehe. Auffallend war zum Beispiel, dass bei allen Tweets in meinen Daten, die mit einem Binde- oder Gedankenstrich (- / –) gefolgt von einem Emotionswort enden, das Emotionswort in einer für die korrekte Sentimentanalyse nützlichen Art verwendet wird. Solche Vorkommen sieht man z.B. in (11) und (12).

(11) Vielen Dank für die Glückwünsche! – freudig
(12) 1 Tag in Freiheit und noch 5 Tage und der Rest von Heute, BARCELONA - glücklich

In diesen zwei Beispielen sind die Emotionswörter für die Sentimentanalyse insofern nützlich, da sie eine Art Eigenannotation des Autors der im Tweet beschriebenen

Emotion sind. Allerdings stellt sich hier die Frage, ob ein ironischer Gebrauch dieser Selbstannotation die Sentimentanalyse wieder erschweren könnte.
Ein weiterer guter Hinweis darauf, dass ein Emotionswort die Emotion des Tweet-Autors bezeichnet, ist das „Stern“-Zeichen (*). Steht ein Emotionswort zwischen zwei Sternen, so scheint es mit hoher Wahrscheinlichkeit auch emotional gemeint zu sein:

(13) Was guckt ihr denn da schon wieder für einen Mist, im TV? Und warum, zur Hölle, twittert ihr darüber? *groll*

Auch in (13) ist das Emotionswort wieder eine Eigenannotation des Autors, der damit klarstellt, wie er sich fühlt. Im Beispiel (14) sieht man zudem einen weiteren Hinweis, der auf die „korrekte“ Verwendung eines Emotionswortes hindeutet (korrekt in dem Sinne, dass das Emotionswort auch tatsächlich eine Emotion des Autors bezeichnet): die Dehnung des Wortes durch Einsetzen von zusätzlichen Vokalen (oder in anderen Beispielen auch Konsonanten).

(14) *@User* *Neeeeeeeeeeeeeeeeid* :D Viel Spaß mit deinem neuen Neid!

Es scheint, dass die Twitterer ihre Emotionen durch das Verlängern von Wörtern noch zusätzlich unterstreichen wollen. Um etwa das Beispiel (14) für die Sentimentanalyse verarbeiten zu können, ist deshalb ein Normalisierungsschritt notwendig. Bei der Normalisierung werden natürlich noch andere, Twitter-typische Phänomene wie Schreibfehler und Abkürzungen behandelt. Eine Methode, um „notoriously noisy“ [368] Tweets (und SMS) zu normalisieren, wurde von Han and Baldwin [2011] entwickelt. Dabei ist Normalisierung von Tweets nicht einfach eine Rechtschreibkorrektur, denn die Wohlgeformtheit in Tweets wird oft absichtlich verletzt – um Zeichen zu sparen, aufgrund der sozialen Identität des Nutzers, oder einfach aus Konvention [368]. Ein Problem bei der Verlängerung von Worten ist ausserdem, dass oft Ambiguität herrscht: so kann mit „Gooood“ entweder „Good“ oder „God“ gemeint sein. Die Methode von Han and Baldwin [2011] funktioniert deshalb in zwei Schritten: zuerst wird eine Liste von möglichen Kandidaten erstellt, danach wird anhand von Kontextfeatures die wahrscheinlichste Form ausgewählt [369]. Der Vorteil von diesem Vorgehen ist, dass ambige Wörter je nach Kontext unterschiedlich normalisiert werden [375]. Um die Sentimentanalyse zu verbessern, könnte ein Algorithmus wie der von Han and Baldwin [2011] eingesetzt werden. Falsch geschriebene, abgekürzte oder verlängerte Emotionswörter würden auf ihre ursprüngliche Form reduziert bzw. ausgebaut und könnten somit erfolgreich erkannt und im Emotionslexikon nachgeschlagen werden.
Ein weiterer nützlicher Hinweis für die Verarbeitung von Emotionswörtern für die Sentimentanalyse ist die Grossschreibung des gesamten Tweets wie in (15) und (16):

(15) @*User* OMG ICH BIN SO AUFGEREGT 😍😩😫

(16) BOAH ICH BIN SO WÜTEND UND DAS ALLES NUR WEGEN DIR!

Das Schreiben in Grossbuchstaben wird im Internet allgemein als „Schreien“ verstanden [Berk, 2011, 47]. Kommen in einem Tweet Emotionswörter verbunden mit Grossschreibung vor, deutet dies also auf eine starke Emotion hin. Man kann davon ausgehen, dass ein Emotionswort in diesem Kontext deshalb auch emotional (und nicht sachlich) verwendet wird.
Interessanterweise sind die genannten Hinweise auf Emotion – Binde- und Gedankenstrich, Sternchen, Verlängerung, Grossschreibung – alle „grafischer“ Natur. Solche grafischen Hinweise auf Emotion sind in der Sentimentanalyse natürlich um einiges einfacher zu verarbeiten als lexikalische oder semantische Hinweise.

4.3 Mehrere Emotionswörter in Tweets

Wie die Analyse von Tweets mit Emotionswörtern in Abschnitt 4.1 gezeigt hat, ergeben sich für die Sentimentanalyse einige Schwierigkeiten. In der Realität müssen Emotionswörter zudem keineswegs nur genau einmal pro Tweet vorkommen, wie dies in Abschnitt 4.1 der Fall war. Nachfolgend analysiere ich deshalb Tweets, die mindestens zwei Emotionswörter enthalten. Um es noch etwas interessanter zu machen, konzentriere ich mich ausserdem auf Tweets, die Emotionswörter unterschiedlicher Polarität enthalten, also z.B. ein eher positives und ein eher negatives Emotionswort.
Dabei ergibt sich schon ein erstes Problem für die Sentimentanalyse: enthält ein Tweet zwei Emotionswörter entgegengesetzter Polarität, stellt sich die Frage, welches für die Sentimentanalyse mehr Relevanz hat. Dies ist z.B. in (17) erkenntlich:

(17) Sehnsucht. Sehnsucht nach Schmetterlingen im Bauch. Sehnsucht nach diesem Glücksgefühl. Nach einem Lächeln, das nur für mich bestimmt ist...

Ein menschlicher Annotator erkennt wahrscheinlich, dass die in (16) vorherrschende Emotion *Erwartung* (Basisemotion von „Sehnsucht“) und nicht *Freude* (Basisemotion von „Glücksgefühl“) ist. Für einen Sentimentanalyse-Algorithmus ist dies jedoch schwierig zu erkennen. In diesem spezifischen Fall jedoch könnte die Lösung jedoch einfach sein, denn das Wort „Sehnsucht“ kommt drei Mal im Tweet vor. Eine Möglichkeit wäre deshalb, alle im Tweet verwendeten Emotionskategorien zu zählen und sich anschliessend für die Kategorie, die am meisten gezählt wurde, zu entscheiden.

Eine weitere Schwierigkeit mit mehreren Emotionswörtern pro Tweet ergibt sich, wenn die Wörter zwar wie gewünscht in einem emotionalen Sinn verwendet werden, aber trotzdem nicht die Emotion des gesamten Tweets ausdrücken. Die lässt sich z.B. in (18) beobachten:

(18) *@User* Ja das ist traurig, aber diese Art der Ignoranz geht einfach gar nicht! Du (und jeder andere) sollte froh sein, über jeden

In Tweet (18) wird „traurig“ mit *Traurigkeit* annotiert und „froh“ mit *Freude*. Trotzdem drückt der Tweet insgesamt weder Traurigkeit noch Freude aus, sondern eher *Ärger*. Dieses Beispiel ist sehr ähnlich zu dem in 4.1 erwähnten Problem, dass in einem emotionalen Tweet gar keine Emotionswörter vorkommen. Hier müssten also beide Emotionswörter ignoriert werden und mit dem restlichen Kontext die passende Emotion herausgefunden werden, was sehr schwierig umzusetzen ist.
Eine grundlegende Schwierigkeit, auf die man in Tweets mit mehreren Emotionswörtern trifft, ist zu entscheiden, welches Wort die momentane Emotion des Autors bezeichnet. (19) und (20) enthalten zwei bzw. drei Emotionswörter, die mit unterschiedlichen oder sogar entgegengesetzten Basisemotionen annotiert werden.

(19) *@User* och bitte nicht Trübsal blasen, wo ich doch gerade so glücklich bin! Sei du es auch! Bitte <3

(20) Angst machen mir beim Kinderturnen Mütter, die begeistert alle Lieder mitsingen und tanzen, während ihre Kinder gelangweilt daneben sitzen.

Soll also z.B. für (20) *Furcht* („Angst“), *Freude* („begeistert“), oder *Ekel* („gelangweilt“) gelten? Um in einer solchen Situation eine Entscheidung zu treffen, wäre eine wie in 4.1 schon beschriebene Möglichkeit, auf Pronomen zurückzugreifen. Steht ein Emotionswort in der Nähe eines Pronomens, ist die Wahrscheinlichkeit gross, dass diese zusammengehören. Den Tweet (20) würde man also insgesamt mit der Basisemotion *Furcht* annotieren, weil das Reflexivpronomen im Satz („mir“) am nächsten beim Emotionswort „Angst“ steht. Allerdings funktioniert dieses Vorgehen nicht immer: in (19) stehen sowohl zwischen „Trübsal“ und „ich“ als auch zwischen „ich“ und „glücklich“ drei Token. Die einfachste Lösung in so einem Fall ist natürlich, die Annotation mit Emotionswörtern wegzulassen bzw. „Uneinigkeit“ als Resultat zurückzugeben. Dies sollte allerdings im Interesse eines robusten Systems möglichst vermieden werden. Ein weiterer Lösungsansatz wäre deshalb, von der satzbasierten Sentimentanalyse auf die phrasenbasierte zu wechseln. Dabei könnte man als Resultat Personalpronomen mit der zugehörigen Emotion ausgeben. Als Beispiel dient Tweet (21):

(21) Ich finde es total traurig, dass ich immer nur ein halbes Schuljahr Philosophieunterricht habe. Ich liebe Philosophie so sehr.

Hier findet man zwar zwei Emotionswörter, aber auch zwei Personalpronomen mit jeweils gleichem Abstand von 2 Wörtern. Eine Sentimentanalyse auf Phrasenebene könnte also zwei Gefühle erkennen: „ich -¿ traurig“ und „ich -¿ liebe“. Damit verfehlt man allerdings wieder das Ziel, pro Tweet *eine* Einschätzung der Polarität abzugeben. Die verschiedenen Aspekte/Pronomen und ihre Basisemotion müssten mittels eines (noch zu entwickelnden) Algorithmus miteinander verrechnet werden, um zu einer abschliessenden Einschätzung der Polarität zu gelangen. Eine dritte, unkompliziertere Lösung könnte deshalb sein, nur das zuletzt erwähnte Emotionswort mit Plutchiks Basisemotionen zu annotieren, denn oft hat das letztgenannte Emotionswort in einem Tweet am meisten Gewicht. Twitter-Autoren scheinen, wenn sie mehrere Emotionswörter verwenden, die alle auf sie selbst bezogen sind, das letzte als eine Art „Zusammenfassung“ ihrer Grundstimmung zu verwenden. Dies ist z.B. in (21) der Fall, aber auch im folgenden (22):

(22) *@User* uhhh danke!!! Wir sind mega aufgeregt und haben etwas angst. Aber freuen uns arg!

Auch im Tweet (23) werden mehrere Emotionswörter verwendet:

(23) Bist du glücklich genießt du die Musik! Bist du traurig verstehst du den Text.

Allerdings deutet das Pronomen „du“ darauf hin, dass die Emotionswörter in diesem Tweet nicht auf den Autor bezogen sind, sondern auf eine andere Person. Der Tweet ist eher beschreibend als subjektiv und kann wie in 4.1 bereits eingeführt als „unpersönlich“ bezeichnet werden. Deshalb könnte für solche Fälle wiederum ein PoS-Tagging von Nutzen sein, um unpersönliche Aussagen auszusortieren.
In manchen Fällen kommen mehrere Emotionswörter zusammen mit verschiedenen Zeitformen vor, wie in (24):

(24) *@User* Ich bin so, so froh, zu lesen, dass alles halbwegs gut ist! Wir waren alle verwirrt und sehr besorgt!!! Alles Gute!! <33333333

Dabei bezieht sich das Emotionswort, das mit dem Präteritum verwendet wird, auf eine vergangene Emotion, während das mit dem Präsens verwendete Emotionswort die aktuelle Emotion beschreibt. Eine Möglichkeit wäre deshalb, nur die Emotionswörter, die zusammen mit einem Verb im Präsens vorkommen, mit einer Basisemotion annotieren. Für Tweet (24) würde also die Basisemotion *Freude* gelten, weil das „aktuelle“ Emotionswort „froh“ ist. Analog gilt natürlich dasselbe, wenn ein Emotionswort im Präsens und eines hypothetisch (z.B. im Konjunktiv)

verwendet werden.
Die Rolle der Verneinung wurde bereits in Abschnitt 4.1 erwähnt. Wenn ein Tweet nicht nur eines, sondern mehrere Emotionswörter enthält, können generell dieselben Vorgehensweisen befolgt werden. Ein Beispiel dafür ist Tweet (25):

(25) *@User* Gelangweilt habe ich mich auch nicht, aber begeistert bin ich auch nicht!

Wie man in (25) sieht, muss ein Shifter wie „nicht“ nicht unbedingt direkt vor einem Emotionswort platziert sein, sondern kann auch z.B. am Ende eines Teilsatzes vorkommen. Die Problematik des Vorgehens, aufgrund einer Verneinung die auf Plutchiks Darstellung der Emotionen die „gegenüberliegende“ Emotion auszuwählen, gilt auch hier: Die Umkehrungen der Basisemotionen von „gelangweilt“ (*Ekel*) und „begeistert“ (*Freude*) sind *Vertrauen* und *Traurigkeit*; aber beide Wörter beschreiben den Tweet nicht unbedingt passend.
Eine weitere interessante Situation ergibt sich, wenn in einem Tweet ein Emotionswort in Normalform zusammen mit einem verneinten Emotionswort vorkommt, wie in (26):

(26) *@User* hab' erstmal keine Angst u. bin froh, wenn SuS Dinge selbst in d. Hand nehmen. Trotzdem bin ich m. offenen Augen da! Cc *@User*

In solchen Fällen fällt die Entscheidung allerdings leichter, welches Emotionswort mehr über die Emotion des Tweets aussagt. Da eines der zwei Emotionswörter verneint ist, kann dieses ignoriert werden und der Tweet mit der Basisemotion des anderen Emotionsworts annotiert werden. Für Tweet (26) könnte ein Sentimentanalyse-Algorithmus also das verneinte Emotionswort („keine Angst“) ignorieren und den Tweet nur anhand des anderen Emotionswortes („froh“) mit einer Basisemotion annotieren (für „froh“ also z.B. *Freude*). Es ist allerdings wichtig zu betonen, dass ein solches Vorgehen nur funktioniert, wenn *zwei* Emotionswörter vorhanden sind.
Wenn ein Tweet mehrere Emotionswörter enthält, deutet dies oft auch darauf hin, dass der Tweet keine subjektive Aussage enthält. Es scheint viele sprachliche Wendungen zu geben, die mehrere Emotionswörter beinhalten. Diese sind in ihrem Aufbau Sprichwörtern ähnlich, man könnte sie vielleicht als Allgemeinplätze bezeichnen. Beispiele dafür sind (27) und (28):

(27) Wenn die Sehnsucht größer als die Angst ist, wird Mut geboren.
(28) Liebe und Hass sind selten kluge Ratgeber.

Um solche Allgemeinplätze auszusortieren, kann wiederum überprüft werden, ob Personalpronomen wie „ich“ oder „wir“ in einem Tweet vorkommen.
Zuletzt noch eine sprachliche Beobachtung: oft deuten Antwortpartikel wie „Danke“ oder „Merci“ darauf hin, dass ein Tweets eine Emotion beschreibt. Solche Wörter

könnte ich deshalb auch noch ins Emotionswörterlexikon aufnehmen.

4.4 Emotionswörter und Emoticons/Emojis

Als Nächstes untersuche ich nun das Zusammenspiel von Emotionswörtern und Emoticons/Emojis in Tweets, und ob dieses einen Einfluss auf die Sentimentanalyse haben könnte. Ein häufig beobachtetes Phänomen, welches ich schon im vorherigen Abschnitt genauer untersucht habe, ist das Vorkommen von mehreren, in ihrer Polarität gegensätzlichen Emotionswörtern in einem Tweet. Dabei besteht die Schwierigkeit darin, zu entscheiden, welches der Emotionswörter die aktuelle Emotion des Autors beschreibt. Kommen in einem Tweet zusätzlich Emoticons oder Emojis vor, kann dies diese Entscheidung erheblich erleichtern. Beispiele dafür sind Tweets (29) und (30):

(29) *@User* lieber jung naiv und verwirrt - Du schreibst sehr schöne Tweets - macht Freude sie zu lesen :)
(30) Hmm... Glücklich, aber doch traurig.. 😐

In (29) kommen mit „verwirrt“ und „Freude“ zwei Emotionswörter vor, die nach Plutchik verschiedene Basisemotionen ausdrücken. Während ein menschlicher Annotator schnell erkennt, dass der Autor des Tweets „Freude“ fühlt und sich das Wort „verwirrt“ auf eine andere Person bezieht, bereitet dies einem Sentimentanalyse-Algorithmus Schwierigkeiten. Um zu entscheiden, ob der Tweet eher mit *Überraschung* oder *Freude* annotiert werden soll, kann deshalb das Emoticon zu Hilfe gezogen werden. Das Smiley „:)“ wird auch mit *Freude* annotiert, somit erhält „Freude“ mehr Gewicht als „verwirrt“, und der Tweet insgesamt kann als Ausdruck von *Freude* klassifiziert werden. Dasselbe Vorgehen ist bei Tweet (30) nützlich: der Twitterer fühlt *Freude* und *Traurigkeit* zugleich. Auch ein menschlicher Annotator kann diesen Tweet nicht so einfach einordnen. Zieht man wiederum das Emoji zu Hilfe, kann die Entscheidung schnell getroffen werden. Emojipedia [2014] beschreibt dieses Emoji eloquent als „A pensive, remorseful face. Saddened by life. Quietly considering where things all went wrong“. Wenn dieses Emoji also *Traurigkeit* ausdrückt, dann erhält das Emotionswort „traurig“ mehr Gewicht als „glücklich“, und der Tweet kann auch insgesamt als *Traurigkeit* ausdrückend klassifiziert werden.
Generell sind Emojis und Emoticons ein sehr guter Hinweis darauf, dass ein Emotionswort nicht nur in einem emotionalen Sinn gemeint ist, sondern auch die Emotion des Tweet-Autors selbst beschreibt. Zudem scheinen die Emoticons nahezu immer die „wahre“ Emotion des Twitterers auszudrücken, was zum Beispiel die Erken-

nung von Ironie erleichtert. Die Vorteile, die Emoticons für die Sentimentanalyse möglicherweise bieten können, wurden auch von Liu et al. [2012] erkannt. In ihrer Studie entwickeln die Autoren ein System für die Sentimentanalyse, das sie „ESLAM (emoticon smoothed language model)“ nennen [Liu et al., 2012, 1679]. ESLAM wird mit manuell klassifizierten Tweets für die Sentimentanalyse trainiert, dann aber mit „noisy labeled data“ ergänzt, um die Performanz zu verbessern. Diese „noisy labeled data“ sind Tweets, die aufgrund der enthaltenen Emoticons automatisch als positiv oder negativ eingeordnet werden [1679]. Wie der Vergleich zeigt, funktioniert ein Sentimentanalyse-System konsistent etwas besser, wenn es nebst den manuell annotierten Daten auch noch die automatisch annotierten Emoticon-Daten verwendet [1681-1682].

Einen ähnlichen Ansatz wählten Davidov et al. [2010]. Um herauszufinden, welche die am häufigsten verwendeten Emoticons sind und mit welchen Emotionswörtern diese beschrieben werden, befragten die Autoren Menschen auf Amazon Mechanical Turk. Danach sammelten sie für die 15 meistgenannten Emoticons je 1000 Tweets, welche diese Emoticons enthalten [Davidov et al., 2010, 245]. Um zu testen, wie gut sich Emoticons als Kategorien für die Sentimentanalyse eignen, untersuchten die Autoren die Performanz verschiedener Algorithmen mit den Annotationen von menschlichen Annotatoren. Wie die Ergebnisse zeigen, stimmte die Zuordnung von Emoticons zwischen der automatischen und menschlichen Annotation für 84% der Tweets überein [247].

Diese zwei Studien unterstützen also meine These, dass Emoticons und Emojis zur Verbesserung der Sentimentanalyse beitragen können. Anstatt Emoticons nur als Labels zu verwenden, könnte man sie auch ins Lexikon aufnehmen und mit Plutchiks Basisemotionen annotieren. Es wäre dabei spannend zu untersuchen, ob ein automatisiertes Vorgehen gute Resultate liefert. Als Beispiel soll Tweet (31) dienen:

(31) Sehr geehrtes @*User*, ich bedanke mich, auch im Namen meiner Alter Egos, für Ihren großzügigen Sternenregen. Wir sind entzückt! □👍😍□👍□

Das Vorgehen wäre so, dass die in einem Tweet vorkommenden Emoticons und Emojis mit der Basisemotion des im Tweet erwähnten Emotionswortes annotiert werden. Hier würden also die Emojis „Daumen hoch“ und „lächelnder Smiley mit herzförmigen Augen“ beide mit dem Emotionswort „entzückt“ in Verbindung gesetzt. Danach ist es nur ein kleiner Schritt, im Emotionswörterlexikon die bereits annotierte Basisemotion nachzuschlagen.

4.5 Emotionswörter und Hashtags

Neben Emoticons bieten auch Hashtags im Zusammenhang mit Emotionswörtern einige interessante Möglichkeiten zur Untersuchung. Im Allgemeinen werden Hashtags benutzt, um Tweets nach Themen zu kategorisieren. Klickt man auf ein Hashtag, so werden weitere Tweets, die dieses Hashtag enthalten, angezeigt [Mollett et al., 2011, 3]. Oft werden Hashtags auch als „Hintergrund-Kommunikationsmittel“ verwendet, um z.B. an Konferenzen oder bei Fernsehprogrammen parallel zur Veranstaltung Meinungen auszutauschen [Mollett et al., 2011, 3]. Allerdings werden Hashtags nicht mehr nur verwendet, um Tweets mit Themen zu taggen. Viele User sind auch sehr kreativ im Umgang mit Hashtags und verwenden z.B. *#itwasntme, #sorrynotsorry*, oder *#bestdayofmylife*. Andere wiederum verwenden Hashtags mit „normalen“ Worten, die in den Tweet integriert sind, wie z.B. in (32):

(32) #Heute gibts außergewöhnliche #Shows und Spektakel in #Berlin: http://bit.ly/16l3e8O

Schaut man sich nun die Verwendung von Emotionswörtern mit Hashtags an, so fällt auf, dass diese oft am Ende von Tweets auftauchen. Anstelle des Themas des Tweets wird damit die Emotion des Users klassifiziert, wie in (33):

(33) Das Leben bietet so viel Schönheit und Überraschung. #glücklich #dankbar

Diese Emotionswort-Hashtags kommen damit einer „Selbstannotation“ des Tweets durch den Autor dar, da er seine Äusserung gleich selber mit einer Emotion klassifiziert. Damit sind solche Hashtags oft auch sehr hilfreich für die automatische Verarbeitung von Tweets in der Sentimentanalyse, denn vielfach ist nach dem Lesen des Tweets nicht klar, wie die Emotionslage des Users ist. Erst das Hashtag klärt auf, wie die Äusserung gemeint ist. Dies kann man zum Beispiel in (34) beobachten:

(34) Jaaaa,ich bin begeistert. Zug fällt doch aus... Nicht!!!Mal richtig früh bekannt gegeben..Nicht!!!! #sauer #hzl #Bahnstreik

Während das Emotionswort „begeistert“ zu einer Annotation mit *Freude* führen würde, lässt das „#sauer“ gegen Ende des Tweets eher auf *Ärger* schliessen. Wie meine Analyse gezeigt hat, scheinen die Hashtag-Emotionswörter mehr Gewicht zu haben, wenn ein Tweet mehrere Emotionswörter von unterschiedlicher Polarität enthält. Diese Erkenntnis könnte in einen Sentimentanalyse-Algorithmus eingebaut werden, indem z.B. Hashtag-Emotionswörter höher gewichtet werden als normale Emotionswörter.
Eine Studie, die sich mit Emotionswörter-Hashtags und ihrer Nützlichkeit für die Sentimentanalyse beschäftigt, stammt von Mohammad [2012]. Dabei untersucht der

Autor unter anderem, ob die „Selbstannotation“ von Tweets durch Hashtags konsistent ist [248]. Mohammad [2012] verwendet dazu die Theorie der Basisemotionen von Ekman und sammelt Tweets, die mit Hashtags mit Ekmans Basisemotionen enden (z.B. „#surprise“). Das resultierende Korpus enthält ca. 21‘000 Tweets von ungefähr 19‘000 Autoren [248-249]. Um nun herauszufinden, ob die Hashtag-„Annotationen“ der Twitterer übereinstimmen, entfernte Mohammad [2012] die Hashtags und trainierte einen Binary Classifier darauf, diese Hashtags selbst zu bestimmen. Über alle Labels hinweg erreicht der Classifier ein F-measure von ca. 44% [250]. Dies ist eine erstaunliche Übereinstimmung, wenn man bedenkt, dass nicht ein einzelner Annotator, sondern 19‘000 verschiedene die Tweets mit Hashtags „annotiert“ haben.
Eine weitere interessante Kombination, die ich bei meiner Tweet-Analyse entdecken konnte, ist das Vorkommen von Emoticons zusammen mit Emotionswörter-Hashtags in einem Tweet. In Beispiel (35) kann man anhand des negativen Emoticons „:-(“ sehen, dass der Autor des Tweets *Traurigkeit* empfindet. Allerdings schreibt der Autor selbst, dass er auch „#Neugier“ empfindet:

(35) Ja was verteilt denn der *@User* gerade im IZ NÖ Süd - leider war die Ampel grün :-(#Neugier

Im Gegensatz zu einem Computer kann ein menschlicher Annotator Feinheiten im Ausdruck herausspüren und etwa erkennen, dass die vorherrschende Basisemotion in (35) nicht *Traurigkeit*, sondern eher *Erwartung* ist.
Bereits in vorherigen Abschnitten konnte ich beobachten, dass besonders Emotionswörter am Ende eines Tweets und Emotionswörter in Grossbuchstaben grossen Einfluss auf die gesamte Emotion des Tweets haben. Auch im Zusammenhang mit Hashtags lässt sich dies bestätigen, z.B. anhand Tweet (36):

(36) *@User* ohhhhhh nein. der Laden macht doch nicht wirklich 200 m von meiner Haustür auf? Wer bezahlt mir das?! #angst #FREUDE

Annotiert man den Tweet (36) mit einer von Plutchiks Basisemotionen, so ist *Freude* die passende Wahl. Nur ein Mensch kann erkennen, dass „#angst“ vom User eher ironisch anstatt ernst gemeint ist. Die Tatsache, dass die „wahre“ Emotion allerdings an letzter Position im Tweet und zudem in Grossbuchstaben erscheint, könnte für einen Algorithmus genügend Hinweis sein, um die richtige Basisemotion automatisch zu erkennen.

5 Diskussion

Wie meine Untersuchung über die Verwendung von Emotionswörtern in Tweets in Kapitel 4 gezeigt hat, gibt es eine Vielzahl von Schwierigkeiten, die sich im Zusammenhang mit der Sentimentanalyse ergeben können. Zu den Hauptproblemen zählen die falsche Lesart von Emotionswörtern (z.B. wenn Nachnamen als emotionale Wörter gelesen werden); Emotionswörter, die sich nicht auf den Autor des Tweets beziehen („unpersönliche Emotionswörter“); die Umkehrung der Bedeutung von Emotionswörtern durch Ironie oder Shifter (z.B. Verneinung); Konditionalsätze, die dazu führen, dass Emotionswörter nur hypothetisch gemeint sind; und Fälle, wo eine Emotion ohne die Verwendung von Emotionswörtern ausgedrückt wird. Manche Probleme, wie z.B. die falsche Interpretation von Eigennamen, sollten durch die Verwendung von Part-of-Speech-Tagging relativ einfach zu lösen sein. „Unpersönliche“ Tweets können von der Sentimentanalyse ausgeschlossen werden, indem man die Bedingung aufstellt, dass mit Emotionswörtern auch gewisse Pronomen im Tweet vorkommen müssen. Verschiedene Autoren haben sich damit beschäftigt, Shifter wie „frei von“ oder die Negation in die Sentimentanalyse einzubeziehen. Während die Erkennung von Shiftern unproblematisch ist, stellt sich die Frage, welche „entgegengesetzte“ Emotion stattdessen ausgewählt werden soll, möchte man eine Annotation mit Basisemotionen beibehalten.

Im Gegensatz zu den bereits erwähnten Problemen gibt es auch Herausforderungen, die weniger einfach zu lösen sind. Dazu gehört etwa die Erkennung von Ironie. Sind Menschen ironisch, so geben sie stets durch ihre Stimmlage und Mimik Hinweise, dass ihre Aussage genau umgekehrt interpretiert werden soll. Die Frage ist deshalb, wie ein Computer Ironie erkennen soll, da in Texten solche Hinweise fehlen. Auch schwierig für die Sentimentanalyse sind konditionale Sätze. Diese enthalten oft Emotionswörter, ohne eine aktuelle Emotion zu beschreiben, und sollten daher nicht als Sentiment tragend betrachtet werden. Um solche Tweets auszusortieren, müssten Kontextfeatures miteinbezogen werden. Wie ironische oder konditionale Aussagen nur durch den Kontext erkannt werden können, bleibt deshalb weiterhin eine Herausforderung.

Es gibt allerdings auch klare Merkmale, die darauf hindeuten, dass in einem Tweet Emotionswörter in der Weise verwendet werden, dass sie für die Sentimentanaly-

se nützlich sind. Interessanterweise sind diese Hinweise oft grafischer Natur, also z.B. Sternchen, Bindestriche, Grossschreibung oder Dehnung von Wörtern durch Einfügen zusätzlicher Buchstaben. Kommen also in Tweets Emotionswörter kombiniert mit solchen Kennzeichen vor, kann man davon ausgehen, dass es sich um eine „wahre“ Emotion des Twitterers selbst handelt.
Wie die Analyse von Tweets mit mehreren Emotionswörtern gezeigt hat, finden sich dort ähnliche Herausforderungen für die Sentimentanalyse wie bei Tweets mit einzelnen Emotionswörtern. Die wichtigste Frage ist in solchen Kontexten, welches der Emotionswörter die „wahre“ Emotion des Twitter-Autors beschreibt. Manchmal werden auch verschiedenste Emotionswörter erwähnt, aber keines beschreibt die mit dem Tweet ausgedrückte Emotion. In solchen Fällen müsste wiederum mit dem Kontext eine Lösung gefunden werden, was sehr schwierig ist. In anderen Fällen können bereits erwähnte Lösungen mit Pronomen oder Verneinungen angewandt werden.
Die Analyse von Tweets, die Emoticons, Emojis und Hashtags sowie ein oder mehrere Emotionswörter enthalten, ergab interessante Resultate. Wenn zum Beispiel in einem Tweet mehrere Emotionswörter vorkommen, so können Emoticons Hinweise darauf geben, welches Emotionswort die aktuelle Emotion des Twitterers beschreibt. Bei Hashtags zeigt sich dasselbe Bild, denn oft klärt das Emotionswort, das zusammen mit einem Hashtag verwendet wird, über die Emotion im Tweet auf. Diese Hinweise scheinen ziemlich zuverlässig zu sein, weshalb ein solcher Ansatz für die Verbesserung der Sentimentanalyse für Twitter erfolgsversprechend sein könnte. Insbesondere das „unlösbare“ Problem der Ironie-Erkennung könnte damit vielleicht in einigen Fällen sogar gelöst werden.
Nebst meiner Analyse der Tweets hat mir besonders der Aufbau des Emotionswörterlexikons und der Tweet-Korpora zu vielen neuen Einsichten verholfen. Dabei ist mir bewusst geworden, dass der Aufbau solcher Ressourcen nicht einfach blosses Materialsammeln ist, sondern viele, manchmal schwierige Entscheidungen erfordert. Über solche Entscheidungen macht man sich, wenn man bestehende Ressourcen nutzt, meist nur wenig Gedanken. Einige Fragen, mit denen ich mich auseinandergesetzt habe, waren z.B.: Wie viele Daten braucht es, damit die Ressource nützlich ist? Wie viel Aufwand kann ich für das Sammeln der Daten betreiben? In welcher Form sollen die gesammelten Daten sein? Welche Inhalte sollen aufgenommen werden, und welche sollen ausgeschlossen werden? Welche Kategorien werden verwendet? Was passiert mit „unpassenden“ Daten? Bei der Auseinandersetzung mit diesen Fragen ist mir auch klar geworden, wie wichtig es ist, seine Entscheidungen zu begründen. Damit bietet man anderen die Grundlage, die getroffenen Entscheidungen nachzuvollziehen und evtl. auch zu bewerten.
Durch die Beschäftigung mit diesen Themen sind mir natürlich auch einige Ideen zur Weiterführung dieser Arbeit gekommen. Eine naheliegende Idee ist der Ausbau

des Emotionswörterlexikons. Um noch weitere Emotionswörter zu finden, wäre z.B. die Übersetzung von bestehenden anderssprachigen Ressourcen eine Möglichkeit, oder das Durchsuchen von Ressourcen wie *WordNet*. Zudem könnte das Lexikon auch noch mit Emojis ergänzt werden. Eine weitere Idee, die ich gerne noch verfolgt hätte, ist die Annotation des 800-Tweet-Korpus. Dazu würde ich die Tweets nach Plutchiks Basisemotionen annotieren und danach untersuchen, ob ein Zusammenhang zwischen der Annotation des gesamten Tweets und den Annotationen der Emotionswörter im Tweet besteht. Der logische nächste Schritt in meiner Arbeit wäre, die gewonnenen Erkenntnisse selbst in ein Sentimentanalyse-System einzubauen. Dies könnte zum Beispiel in der Form von Features geschehen, die man für einen Classifier bestimmt. Danach könnte man die Performanz des neu geschaffenen Systems mit bestehenden Sentimentanalyse-Algorithmen vergleichen. Natürlich wäre dann auch der Vergleich zwischen einem Machine Learning-Ansatz und einem rein lexikonbasierten Ansatz spannend.
Mit meiner Arbeit wollte ich eine qualitative Untersuchung von Schwierigkeiten der Sentimentanalyse für eine anspruchsvolle Textsorte wie Tweets durchführen. Mein Ziel ist es zu zeigen, dass auch qualitative Untersuchungen dazu beitragen können, Sprachtechnologien wie die Sentimentanalyse zu verbessern. Anstatt immer grössere automatisch erstellte Polaritätslexika zu verwenden, die womöglich viele Fehler enthalten, oder beliebige Features aneinanderzureihen, sollte man sich besser fragen, welche Probleme die meisten Fehler verursachen. Auch einige der präsentierten Studien stützen meine These, dass manchmal weniger, dafür mit Überlegung gewählte Features oder Trainingsmaterial zu einer grösseren Verbesserung führen kann.

6 Schlusswort

In der vorliegenden Arbeit habe ich mich mit dem Gebrauch von Emotionswörtern für die Sentimentanalyse von Twitter-Daten befasst. Dabei war das Ziel, das Vorkommen von Emotionswörtern in Tweets zu analysieren und mögliche Herausforderungen für die Sentimentanalyse zu beschreiben. Dazu erstellte ich in einem ersten Schritt ein Lexikon aus deutschen Emotionswörtern, die ich gemäss der Theorie der Basisemotionen von Robert Plutchik nach acht Basisemotionen annotierte. In einem zweiten Schritt sammelte ich Tweets, in denen die zusammengetragenen Emotionswörter in verschiedenen Kontexten – mit gegensätzlichen Emotionen, Emoticons, Hashtags – vorkamen. Wie die nachfolgende Analyse dieser Tweets gezeigt hat, gibt es eine Fülle an Schwierigkeiten, die sich für die Sentimentanalyse im Zusammenhang mit Emotionswörtern ergeben. Während Emotionswörter nützliche Hinweise für eine bessere Sentimentanalyse geben können, so müssen sie zuerst einmal korrekt erkannt und annotiert werden. Auch die Struktur der Twitter-Daten – ihre Kürze, der fehlende Kontext, die oft sehr kreative Verwendung der Sprache – stellen die Sentimentanalyse vor Herausforderungen. Während viele Sentimentanalyse-Ansätze versuchen, diese Hindernisse mit noch mehr Daten aus dem Weg zu räumen, gehört meiner Meinung nach auch eine detaillierte Analyse der Probleme zu einem erfolgreichen Lösungsansatz. Wenn man genau hinschaut, lassen sich einige Auffälligkeiten und Regelmässigkeiten in der Verwendung von Emotionswörtern in Tweets erkennen. Diese Muster gilt es nun in einem nächsten Schritt in Algorithmen zu verarbeiten. Ich bin davon überzeugt, dass dies ein weiterer „Baustein“ zu einer guten Sentimentanalyse ist. Wie Martínez-Cámara et al. [2014, 18] schreiben, erkennt die akademische Welt zunehmend das Interesse, das Informationen aus der Sentimentanalyse für Mikroblogging-Plattformen geschenkt wird. Damit dieses Wissen für die Endnutzer aber tatsächlich von Nutzen ist, muss es angemessen dargestellt und einfach zu verstehen sein [18]. Gelingt es Forschern, ihre Resultate auch für ein nicht-akademisches Publikum zugänglich zu machen, könnte dies zu mehr Ressourcen für dieses Forschungsgebiet führen, was hoffentlich wiederum in einer Verbesserung der Technologien resultiert. Für die Zukunft der Sentimentanalyse für Mikroblogging-Texte wäre es zudem hilfreich, wenn die Zusammenarbeit mit Wissenschaftlern ähnlicher Forschungsbereiche intensiviert würde. Davon könnten auch

andere Gebiete und Anwendungen der Verarbeitung natürlicher Sprache wie z.B. die automatische Zusammenfassung profitieren. Insbesondere für die deutsche Sprache gibt es noch genügend ungelöste Probleme, sodass die Sentimentanalyse-Forschung weiterhin spannend bleiben dürfte.

Literaturverzeichnis

ABC (2014). ABC der Gefühle und Adjektive. http://www.psychologie-psychotherapie.ch/chameleon/site/ABC.pdf. Letzter Zugriff: 01.09.2014.

Alexa (2014). Site Overview: twitter.com. http://www.alexa.com/siteinfo/twitter.com. Letzter Zugriff: 20.11.2014.

Berk, R. A. (2011). Top 12 Be-Attitudes of Netiquette for Academicians. *The Journal of Faculty Development*, 25(3):45–48.

Bollen, J., Mao, H., and Pepe, A. (2011). Modeling Public Mood and Emotion: Twitter Sentiment and Socio-economic Phenomena. In *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, pages 450–453.

Clematide, S., Gindl, S., Klenner, M., Petrakis, S., Remus, R., Ruppenhofer, J., Waltinger, U., and Wiegand, M. (2012). MLSA - A Multi-layered Reference Corpus for German Sentiment Analysis. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation*, pages 3551–3556.

Dahl, H. and Stengel, B. (1978). A Classification of Emotion Words: A Modification and Partial Test of De Rivera's Decision Theory of Emotions. *Psychoanalysis & Contemporary Thought*, (1):269–312.

Davidov, D., Tsur, O., and Rappoport, A. (2010). Enhanced Sentiment Learning Using Twitter Hashtags and Smileys. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 241–249.

Devitt, A. and Ahmad, K. (2013). Is There a Language of Sentiment? An Analysis of Lexical Resources for Sentiment Analysis. *Language Resources and Evaluation*, 47(2):475–511.

Donaldson, M. (2011). Plutchik's Model of Emotions. http://www.6seconds.org/2011/06/24/plutchiks-model-of-emotions/. Letzter Zugriff: 22.12.2014.

Duden a (2014). Ironie. http://www.duden.de/rechtschreibung/Ironie. Letzter Zugriff: 03.12.2014.

Duden b (2014). Sarkasmus. http://www.duden.de/rechtschreibung/Sarkasmus. Letzter Zugriff: 03.12.2014.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition & Emotion*, 6(3-4):169–200.

Emojipedia (2014). Pensive Face. http://emojipedia.org/pensive-face/. Letzter Zugriff: 26.12.2014.

Feldman, R. (2013). Techniques and Applications for Sentiment Analysis. *Communications of the ACM*, 56(4):82–89.

González-Ibáñez, R., Muresan, S., and Wacholder, N. (2011). Identifying Sarcasm in Twitter: A Closer Look. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 581–586.

Han, B. and Baldwin, T. (2011). Lexical Normalisation of Short Text Messages: Makn Sens a #twitter. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 368–378.

Hobbs, J. R. and Gordon, A. S. (2011). The Deep Lexical Semantics of Emotions. In Ahmad, K., editor, *Affective Computing and Sentiment Analysis*, pages 22–34, Dordrecht: Springer.

Klenner, M., Tron, S., Amsler, M., and Hollenstein, N. (2014). The Detection and Analysis of Bi-polar Phrases and Polarity Conflicts. In *Proceedings of the 11th International Workshop on Natural Language Processing and Cognitive Science*.

Kluge Liste (2014). Emotionen und Gefühle, Liste: Positive, Negative und Neutrale. http://www.kluge-liste.de/gefuehle_liste.html. Letzter Zugriff: 01.09.2014.

Kouloumpis, E., Wilson, T., and Moore, J. (2011). Twitter Sentiment Analysis: The Good the Bad and the OMG! In *Proceedings of the Fifth International AAAI Conference of Weblogs and Social Media*, pages 538–541.

Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. San Rafael: Morgan & Claypool.

Liu, K. L., Li, W. J., and Guo, M. (2012). Emoticon Smoothed Language Models for Twitter Sentiment Analysis. In *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, pages 1678–1683.

Martínez-Cámara, E., Martín-Valdivia, M. T., Ureña-López, L. A., and Montejo-Ráez, A. R. (2014). Sentiment Analysis in Twitter. *Natural Language Engineering*, 20(1):1–28.

Medhat, W., Hassan, A., and Korashy, H. (2014). Sentiment Analysis Algorithms and Applications: A Survey. *Ain Shams Engineering Journal*, 5(4):1093–1113.

Medienwerkstatt Mühlacker (2009). Gefühle mit Adjektiven beschreiben - Wortliste. https://medienwerkstatt-online.de/lws_wissen/vorlagen/showcard.php?id=19088&edit=0. Letzter Zugriff: 01.09.2014.

Mohammad, S. M. (2012). #Emotional Tweets. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics*, pages 246–255.

Mohammad, S. M. and Kiritchenko, S. (2014). Using Hashtags to Capture Fine Emotion Categories from Tweets. *Computational Intelligence*.

Mollett, A., Moran, D., and Dunleavy, P. (2011). *Using Twitter in University Research, Teaching and Impact Activities: A Guide for Academics and Researchers*. London School of Economics and Political Science: LSE Public Policy Group.

Nakov, P., Kozareva, Z., Ritter, A., Rosenthal, S., Stoyanov, V., and Wilson, T. (2013). Semeval-2013 Task 2: Sentiment Analysis in Twitter. In *Proceedings of the Second Joint Conference on Lexical and Computational Semantics*, pages 312–320.

Narayanan, R., Liu, B., and Choudhary, A. (2009). Sentiment Analysis of Conditional Sentences. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pages 180–189.

Neuhaus, F. and Webmoor, T. (2012). Agile Ethics for Massified Research and Visualization. *Information, Communication & Society*, 15(1):43–65.

Neviarouskaya, A., Prendinger, H., and Ishizuka, M. (2010). Recognition of Affect, Judgment, and Appreciation in Text. In *Proceedings of the 23rd International Conference on Computational Linguistics*, pages 806–814.

Nielsen, F. . (2011). A New ANEW: Evaluation of a Word List for Sentiment Analysis in Microblogs. In *Workshop on Making Sense of Microposts: Big Things Come in Small Packages*, pages 93–98.

Ortony, A., Clore, G. L., and Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press.

Pak, A. and Paroubek, P. (2010). Twitter as a Corpus for Sentiment Analysis and Opinion Mining. In *Proceedings of the International Conference on Language Resources and Evaluation*, pages 1320–1326.

Plutchik, R. (1980). *Emotion: A Psychoevolutionary Synthesis*. London: Harper & Row.

Plutchik, R. (1991). *The Emotions: Facts, Theories, and a New Model*. Lanham: University Press of America, rev. edition.

Remus, R., Quasthoff, U., and Heyer, G. (2010). SentiWS - A Publicly Available German-language Resource for Sentiment Analysis. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation*, pages 1168–1171.

Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., and Harabagiu, S. M. (2012). EmpaTweet: Annotating and Detecting Emotions on Twitter. In *Proceedings of the Eight International Conference on Language Resources and Evaluation*, pages 3806–3813.

Sarkasmus (2014). Was ist Sarkasmus? http://sarkasmus-ironie-zynismus.de/sarkasmus/. Letzter Zugriff: 03.12.2014.

Shelke, N. M., Deshpande, S., and Thakre, V. (2012). Survey of Techniques for Opinion Mining. *International Journal of Computer Applications*, 57(13):30–35.

Skinner, J., Ghiassi, M., and Zimbra, D. (2013). Twitter Specific Lexicon for Sentiment Analysis. In Bacciga, A. and Naliato, R., editors, *Recent Advances in Artificial Intelligence Research*, pages 1–30, New York: Nova Science.

Smedt, T. D. and Daelemans, W. (2012). Pattern for Python. *The Journal of Machine Learning Research*, 13(1):2063–2067.

Suttles, J. and Ide, N. (2013). Distant Supervision for Emotion Classification with Discrete Binary Values. In Gelbukh, A., editor, *Computational Linguistics and Intelligent Text Processing*, pages 121–136, Berlin: Springer.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A. (2010). Sentiment Strength Detection in Short Informal Text. *Journal of the American Society for Information Science and Technology*, 61(12):2544–2558.

Twitter (2014). Über Twitter, Inc. https://about.twitter.com/de/company. Letzter Zugriff: 20.11.2014.

Waltinger, U. (2010). GermanPolarityClues: A Lexical Resource for German Sentiment Analysis. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation*, pages 1638–1642.

Wikipedia (2014). Partizip. http://de.wikipedia.org/wiki/Partizip. Letzter Zugriff: 19.09.2014.