

BTU Cottbus - Senftenberg
Fakultät 3 - Maschinenbau, Elektrotechnik und
Wirtschaftsingenieurwesen
Lehrstuhl für Marketing und Innovationsmanagement
Vertretungsprofessor Dr. rer. pol. Henning Kreis



– Masterarbeit –

Sentiment Analyse von Online Consumer Reviews - Ein Vergleich zwischen Deutschland und China im Textileinzelhandel

Yicheng Wei
Wirtschaftsingenieurwesen

26. Januar 2016

Betreuer: Dr. phil. Alexandra Rese

Inhaltsverzeichnis

Abbildungsverzeichnis	v
Tabellenverzeichnis	vii
Abkürzungsverzeichnis	ix
1 Einführung	1
1.1 Online Textileinzelhandel in den beiden Ländern	2
1.2 Forschungsfragen	4
1.3 Die Struktur der Arbeit	4
2 Theoretische Grundlagen	9
2.1 Online Consumer Reviews in der Marketingforschung	9
2.1.1 Begriff	9
2.1.2 Forschungsstand zu Online Consumer Reviews	11
2.1.3 Motive für das Schreiben von Online Consumer Reviews	13
2.1.4 Wirkungen von Online Consumer Reviews	14
2.1.5 Online Consumer Reviews im Textileinzelhandel	19
2.2 Sentiment Analyse	21
2.2.1 Definition	21
2.2.2 Der Ablauf der Sentiment Analyse	22
2.2.3 Kernprozess	23
2.2.4 Der konkrete Algorithmus im Dokument-Level	25
2.3 Deutschland und China: ein kulturübergreifender Vergleich	27
2.3.1 Kulturelle Dimensionen	28
2.3.2 Machtdistanz	28
2.3.3 Individualismus gegenüber Kollektivismus	30
2.3.4 Maskulinität gegenüber Femininität	31
2.3.5 Unsicherheitsvermeidung	32
2.3.6 Langfristige Orientierung gegenüber kurzfristige Orientierung	33
2.3.7 Nachsicht gegenüber Beherrschtheit	34
2.3.8 Das sechsdimensionale Modell für Deutschland und China	34

2.4	Relevante Forschungsergebnisse zum Textileinzelhandel	36
2.5	Hypothesen	37
3	Datensammlung und Textverarbeitung	43
3.1	Datenquellen	43
3.1.1	Amazon weltweit und in Deutschland	43
3.1.2	Chinesische Plattform:“Tmall”	45
3.2	Sportkleidung als Untersuchungsobjekt	47
3.3	Technische Werkzeuge für die Datensammlung	48
3.3.1	Datensammlung aus Webseiten mit Python	48
3.3.2	XPath: Daten Lokalisierung in Webseiten	49
3.3.3	Reguläre Ausdrücke: Bearbeitung der Texte	50
3.4	Datensammelungsprozess	50
3.4.1	Sammeln der Daten aus Amazon.de	50
3.4.2	Sammeln der Daten aus Tmall.com	53
3.5	Textverarbeitung	54
3.6	Maschinelle Übersetzung für mehrsprachige Sentiment-Analyse	55
4	Empirische Ergebnisse	57
4.1	Die deskriptiven Ergebnisse zu den Online Consumer Reviews	57
4.2	Hypothesenüberprüfung	59
4.3	Hypothese 1	60
4.3.1	Hypothese 1 in unterschiedlichen Kategorien	61
4.3.2	Hypothese 1 bei unterschiedlichen Marken	62
4.3.3	Hypothese 1 im Allgemeinen	64
4.4	Hypothese 2	65
4.4.1	Hypothese 2 in unterschiedlichen Kategorien	65
4.4.2	Hypothese 2 bei unterschiedlichen Marken	67
4.4.3	Hypothese 2 in Allgemeinen	69
4.5	Hypothese 3	70
4.5.1	Hypothese 3 bei jedem Artikel	70
4.5.2	Hypothese 3 in Allgemeinen	73
4.6	Hypothese 4	74
4.6.1	Hypothese 4 bei jedem Artikel	75
4.6.2	Hypothese 4 in Allgemeinen	77
5	Diskussion	81
5.1	Qualitativ gegenüber Quantitativ	82

5.2	Positiv gegenüber Neutral	83
5.3	Dezentralisiert gegenüber Zentralisiert	85
5.4	Abhängig gegenüber Unabhängig	86
6	Zusammenfassung und Ausblick	89
6.1	Zusammenfassung	89
6.2	Limitationen	90
6.3	Zukünftige Forschung	92
	Literaturverzeichnis	xi

Abbildungsverzeichnis

1.1	42% Textilien werden in Deutschland online verkauft	3
1.2	Anteil der Online-Shopping in China	4
1.3	Online Umsatz für Kleidung in China	5
1.4	Die Struktur der Arbeit	6
2.1	Die Typologie von eWOM	10
2.2	Ein Beispiel von OCRs	10
2.3	Ein Beispiel von Online Rückkopplungsmechanismen	12
2.4	Die beiden beobachteten Forschungswege über OCRs	13
2.5	Auswirkungen der OCRs aus Konsumenten- und Unternehmenssicht	15
2.6	Die bemerkte Zeitachse von OCRs	16
2.7	Der Ablauf der Sentiment Analyse	22
2.8	Die Verteilung der absoluten Gewichte	27
2.9	Das sechsdimensionale Modell für Deutschland und China	35
3.1	Logo von Amazon Deutschland	43
3.2	Umsatzerlöse von Amazon.com	44
3.3	Amazon Deutschland Umsatz 2010-2014	45
3.4	Logo von Tmall.com	45
3.5	Tmall GMV	46
3.6	Kundenrezensionen von Amazon.de	51
3.7	Die Struktur der Webseite von Amazon.de	52
3.8	Kundenrezensionen von Tmall.com	53
4.1	die Wortwolken von chinesischen (Links) sowie deutschen (Rechts) OCRs .	57
4.2	die Top 10 Wörter von chinesischen (Links) sowie deutschen (Rechts) OCRs	58
4.3	Das Q-Q-Plot der quantitativen Valenzen für China und Deutschland	60
4.4	Das Q-Q-Plot der qualitativen Valenzen für China und Deutschland	61
4.5	Durchschnittliches Volumen von qualitativen OCRs über T-Shirt und Hose .	62
4.6	Durchschnittliches Volumen von qualitativen OCRs bei verschiedenen Marken	63
4.7	Durchschnittliches Volumen von qualitativen OCRs	64
4.8	Die Struktur von chinesischen und deutschen quantitativen OCRs über T-Shirt	65

4.9	Die Verteilung der Valenzen von chinesischen und deutschen qualitativen OCRs für T-Shirt	66
4.10	Die Struktur von chinesischen und deutschen quantitativen OCRs über Hose	67
4.11	Die Verteilung der Valenzen von chinesischen und deutschen qualitativen OCRs für Hose	68
4.12	Die Struktur von chinesischen quantitativen OCRs bei Nike und Puma . . .	69
4.13	Die Verteilung der chinesischen Valenzen von qualitativen OCRs bei zwei unterschiedlichen Marken	70
4.14	Die Struktur von deutschen quantitativen OCRs bei Nike und Puma	71
4.15	Die Verteilung der deutschen Valenzen von qualitativen OCRs bei zwei unterschiedlichen Marken	72
4.16	Die Struktur von den chinesischen und deutschen quantitativen OCRs in Allgemeinen	73
4.17	Die Verteilungen der chinesischen und deutschen Valenzen von den qualitativen OCRs in Allgemeinen	74
4.18	Die chinesischen und deutschen Valenzen von den quantitativen und qualitativen OCRs in Allgemeinen	79

Tabellenverzeichnis

2.1	Auswirkung der Valenz von OCRs	15
2.2	Ein Überblick über die Literature in quantitativer Aspekte	18
2.3	Überblick über den Inhalt des Lexikons	25
2.4	Die wichtige Unterschiede zwischen Gesellschaften mit geringer und großer Machtdistanz	29
2.5	Die relevanten kulturübergreifenden Forschungen	40
2.6	Die Attribute der quantitativen und qualitativen OCRs	41
3.1	Marktpositionen von Online Händler in China	46
3.2	Vergleich zwischen manueller und maschineller Übersetzung nach der Textverarbeitung	56
4.1	Der jeweilige Preis von jeder Kleidung in Deutschland und China	59
4.2	Die Ergebnisse durch Shapiro-Wilk-Test für China und Deutschland	60
4.3	Das Volumen der OCRs über T-Shirt und Hose von qualitativen und quantitativen Seiten	61
4.4	Das Volumen der OCRs bei verschiedenen Marken	63
4.5	Das Volumen der OCRs im Allgemeinen von qualitativen und quantitativen Seiten	64
4.6	Die Valenz von quantitativen und qualitativen OCRs über unterschiedlichen Kategorien in beiden Ländern	67
4.7	Die Valenzen von chinesischen und deutschen, quantitativen und qualitativen OCRs bei unterschiedlichen Marken	68
4.8	Die Valenzen der chinesischen und deutschen OCRs auf der quantitativen und qualitativen Seite	69
4.9	Die Ergebnisse des Levene-Tests beim jeden Artikel	71
4.10	Die Varianzen der quantitativen und qualitativen OCRs von jedem Artikel	72
4.11	Die Ergebnisse von Levene-Test in Allgemeinen	73
4.12	Die Varianzen von den chinesischen und deutschen OCRs	73
4.13	Die Ergebnisse von Pearsons Korrelationstest bei jedem Artikel	75
4.14	Die Pearsons Korrelationskoeffizienten ρ_p zwischen den chinesischen und deutschen Valenzen der quantitativen und qualitativen OCRs	76

4.15	Die Ergebnisse von Spearmans Rangkorrelationstest bei jedem Artikel . . .	76
4.16	Die Spearmans Rangkorrelationskoeffizienten ρ zwischen den chinesischen und deutschen Valenzen der quantitativen und qualitativen OCRs . .	77
4.17	Die Ergebnisse von Pearsons Korrelationstest zwischen quantitativen und qualitativen OCRs in China und Deutschland in Allgemeinen	78
4.18	Die Ergebnisse von Spearmans Rangkorrelationstest zwischen quantitativen und qualitativen OCRs in China und Deutschland in Allgemeinen	78
5.1	Die zusammenfassenden Ergebnisse von Hypothesen	81

Abkürzungsverzeichnis

B2C Business to Customer, (allgemein für Beziehungen zwischen Unternehmen und Konsumenten)

SEC Securities and Exchange Commission in USA

bzw. beziehungsweise

GMV Gross Merchandise Volume

XML Extensible Markup Language

XSLT Extensible Stylesheet Language: Transformations

CSV Comma-Separated Values

i.e. auf Englisch

usw und so weiter

PMI Pointwise Mutual Information

SO Semantische Orientierung

eWOM electronic Word-Of-Mouth

OCRs Online Consumer Reviews

WOM Word-Of-Mouth

1 Einführung

Auf der ganzen Welt brachten die vergangenen Jahre eine Fortführung des hohen Wachstums des Online Einzelhandels. Der Umsatz stieg im Jahr 2014 um über 20 Prozente weltweit auf fast 840 Milliarden US-Dollar, da Online-Händler weiter in neuen Regionen ausgebaut haben und die physischen Einzelhändler durch E-Commerce in neuen Märkten eingedrungen sind (Hana Ben-Shabat *et al.* , 2015). Das schnelle Wachstum des E-Commerce-Marktes hat umfangreiche Forschung angetrieben, um die Meinungen der Konsumenten zum Online Handel besser verstehen zu können (Tong, 2010).

Im Online Handel können die Meinungen der Konsumenten einfach als eine Form der Online Consumer Reviews (OCRs) erreicht werden. Mit diesen OCRs können eine große Menge an Proben ohne Beschränkung von Zeit und Ort erhalten werden. Anders als traditionelle Umfragen und Fragebogen, sind OCRs eine spontane Rückmeldung (Lu, 2015). Während in den meisten Studien eine Probe von Hunderten Menschen verwendet wird um die Meinungen der Konsumenten zu studieren, könnte ein beliebtes Produkt Tausende von OCRs haben. Wegen des Wachstums des E-Commerce werden eine enorme Menge an OCRs erzeugt.

Es gibt schon viele Forschungen über OCRs. Viele Studien haben die Wichtigkeit der OCRs diskutiert, und einige Studien haben schon versucht, die nützliche Information aus den OCRs zu extrahieren. In der Mehrheit dieser Studien werden die OCRs in zwei Teile unterteilt: der numerische Teil und der schriftliche Teil. Diese Studien fokussieren sich mehr auf den numerischen Teil während der schriftliche Teil von OCRs ignoriert wird, trotz der Tatsache, dass der schriftliche Teil mehrere Informationen enthält.

Sentiment Analyse ist ein Teilgebiet des Text Minings, welches den Menschen ermöglicht, die Meinungen der Kunden durch den Text der OCRs zu verstehen, ohne manuell den Text durchlesen und zusammenfassen zu müssen. Es ist wichtig, dass die Sentiment Analyse automatisch die Meinungen und die Gefühle der Verfasser der Texte identifizieren kann. Durch die Sentiment Analyse kann man mit den Meinungen der Kunden in dem Text der OCRs statistisch rechnen, damit man die Gesamtsituation der Meinungen bewerten kann.

Diese Gesamtsituation ist nützlich, besonders für die kulturübergreifende Forschung. Im Vergleich mit den Situationen aus verschiedenen Ländern kann man die kulturellen Einflüsse auf das Konsumverhalten besser erkennen. Trotz der Tatsache, dass die Unterschiede in der nationalen Kultur das Konsumverhalten beeinflussen könnten, haben die meisten Forschungen

über die OCRs die Wirkung der Kultur ignoriert (Gefen & Heart, 2006).

Diese Arbeit macht einen kulturübergreifenden Vergleich zwischen Deutschland und China. Gemeinsam haben die deutschen B2C-Mehrkanal-Online- und Versandhandel einen Umsatz von über 49 Milliarden Euro im Jahr 2014. E-Commerce erwirtschaftete mehr als 85 Prozent des gesamten Branchenumsatzes. In Deutschland repräsentierte der Sektor des elektronischen Handels neun Prozent der gesamten Einzelhandelbranche des Landes im Jahr 2014 mit der positiven Tendenz. (Späth, 2015)

Die chinesischen Online-Käufer sind anspruchsvolle Kunden mit einem großem Markenbewusstsein und Vertrauen in den größten Namen, einschließlich inländischer Marktführer wie Tmall und JD.com sowie internationaler wie Amazon und eBay. In China stellen die chinesischen Käufer im elektronischen Handel ein kulturelles Phänomen dar, vor allem am Ledigen Tag (11. November), der ähnlich wie der Cyber Monday in den USA ist. Alibaba (Plattformbetreiber von Tmall) hat 9,3 Milliarden Dollar Umsatz am Ledigen Tag 2014 gemeldet, was ungefähr sieben Prozent des Gesamtjahresumsatz des Landes entspricht. (Hana Ben-Shabat *et al.*, 2015)

1.1 Online Textileinzelhandel in den beiden Ländern

Dieser Abschnitt stellt die Situationen des Markts von Online Textileinzelhandel in Deutschland und China vor. Diese Vorstellung ist wichtig, man diese Situationen zu wissen, um die Arbeit besser zu verstehen.

Mit mehr als 51 Millionen (94 Prozent der Internetnutzer ab 14 Jahren) digitale Verbraucher im Jahr 2014, genießt Deutschland die größte E-Commerce-Kundenpotenzial in Europa - und ist damit der klare kontinentale Führer (Späth, 2015). In "The 2015 Global Retail E-Commerce Index" ist Deutschland am fünft Position (Hana Ben-Shabat *et al.*, 2015).

Kleidung, Elektronische Waren und Bücher gehören zu den beliebtesten Online-Kategorien. 88 Prozente der Befragter sagen, dass sie in den letzten drei Monaten Kleidungen online gekauft haben, während der Weltdurchschnitt 76% ist. (Hana Ben-Shabat *et al.*, 2015)

Und die Kategorie "Bekleidung, Textilien und Schuhe" hat eine der stärksten Wachstumsraten für den Zeitraum 2011-2014. Die höchsten Umsätze im E-Commerce werden in der Produktkategorien Bekleidungs (11,9 Mrd. EUR), Unterhaltungselektronik (5 Mrd. EUR) und Bücher (4,1 Mrd. EUR) generiert. Als die Abbildung 1.1 gezeigt, werden 42% Kleidungen in Deutschland online verkauft. (Späth, 2015)

China, das weltweit bevölkerungsreichsten Land (fast 1,4 Milliarden Menschen), ist online aktiv. Mehr als ein Drittel der Menschen, die online mindestens einmal pro Woche surfen, sind "ständig verbunden", und 58 Prozenten prüfen das Internet zwei bis vier Mal pro Tag.

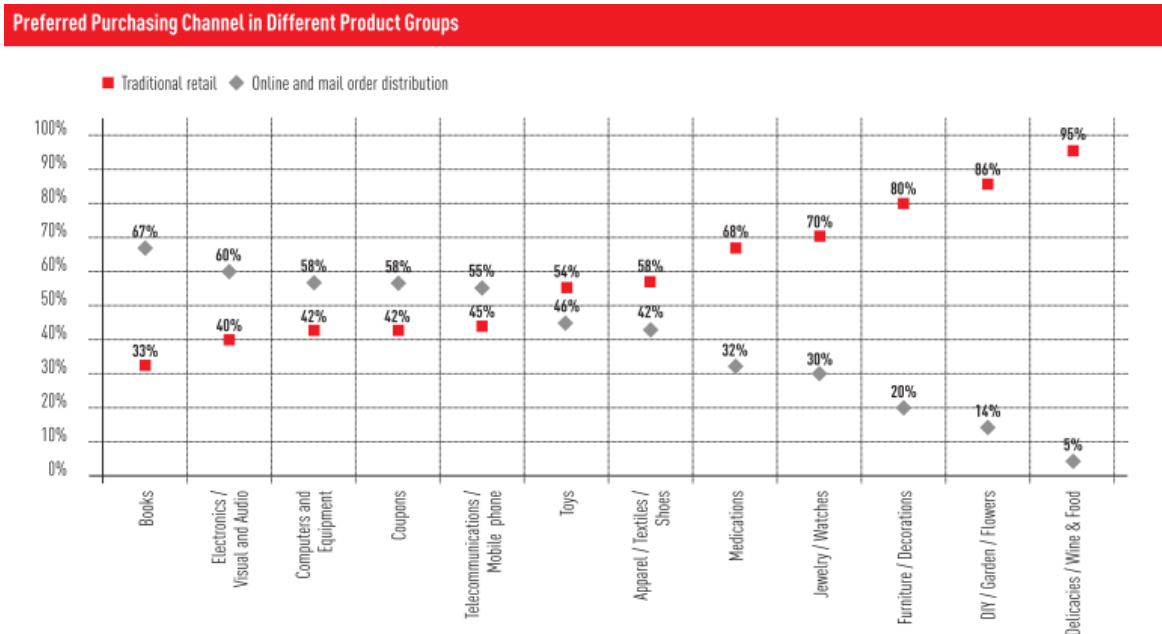


Abbildung 1.1: 42% Textilien werden in Deutschland online verkauft(Quelle:Späth, 2015)

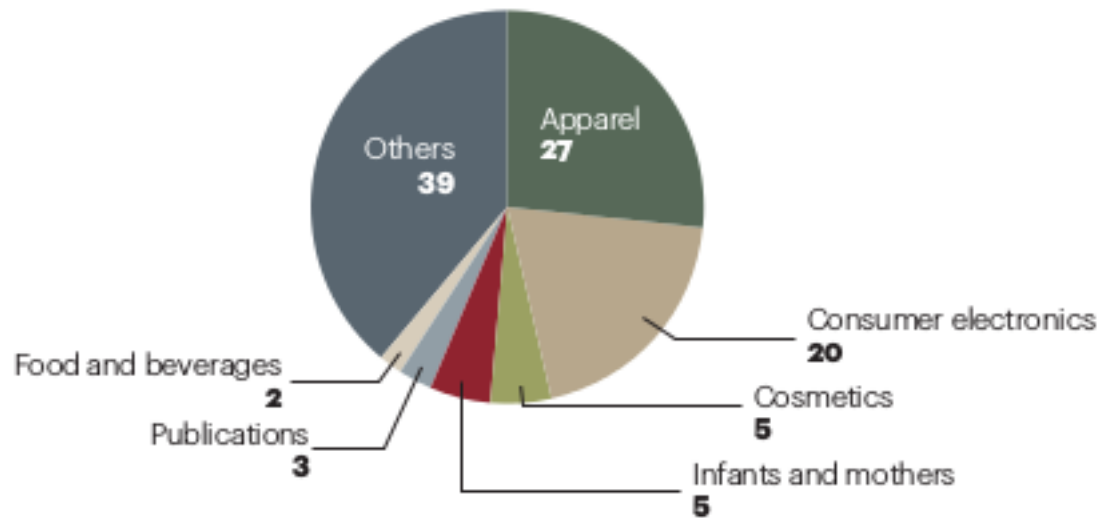
Verglichen mit Deutschland, ist China am zweiten Position in “The 2015 Global Retail E-Commerce Index”, nur nach den USA.(Hana Ben-Shabat *et al.* , 2015)

Online-Handel ist die am schnellsten wachsende Einzelhandels für Bekleidung in China (Fung & Fung, 2014). Abbildung 1.2 stellt fest, dass 27% des Umsatzes in 2013 durch Kleidung verkauft werden, größer als die Kategorie Unterhaltungselektronik. Wie die Abbildung 1.3 gezeigt, wuchs die gesamte Online-Bekleidungs-Transaktionswert in China deutlich um 42,6% gegenüber dem Vorjahr um 434,9 Milliarden Yuan (62,1 Milliarden Euro) im Jahr 2013 zu erreichen (Fung & Fung, 2014).

“Kleidung, Schuhe, Hüte, Taschen und Koffer, Outdoor-Produkte” sind die beliebtesten Kategorien in Online-Shopping im Jahr 2013 erworben, und diese Kategorie wird von 41,3% der Frauen und 28% der Männer inzwischen Top 10 Lieblings Waren in Online-Shopping gewählt (Fung & Fung, 2014). 97% der Befragter sagen, dass sie in den letzten drei Monaten Kleidungen online gekauft (Hana Ben-Shabat *et al.* , 2015).

Trotz der Tatsache, dass Online-Textileinzelhandel einer der größten und stark wachsenden Kategorien des elektronischen Einzelhandels ist, gibt es, insbesondere im Bereich Online Consumer Reviews, wenig kulturübergreifende Forschung auf diesem Gebiet. OCRs können die Kaufentscheidungen der Kunden stark beeinflussen. Zum Beispiel, wollen 40 Prozent der Online Käufer in China unmittelbar “kaufen oder nicht kaufen” Beratungen und Bewertungen (Hana Ben-Shabat *et al.* , 2015). Deutschland und China, haben große Unterschiede in den Kulturen, und die beiden Länder sind wichtig für die Wirtschaft der Welt, aber es gibt zur Zeit noch keine Forschungen, die über die Unterschiede von Kulturen und OCRs zwischen

Share of online shopping, by value (2013, %)



Note: Percentages may not resolve due to rounding.

Sources: iResearch; A.T. Kearney analysis

Abbildung 1.2: Anteil der Online-Shopping in China, nach Wert (2013, %)(Quelle:Mui-Fong & Chee Wee, 2014)

den beiden Ländern diskutiert.

1.2 Forschungsfragen

Diese Arbeit wird die folgende Forschungsfragen diskutieren:

1. Wie sollten die Online Consumer Reviews von China und Deutschland sein, basierend auf den theoretischen Grundlagen?
2. Welche Unterschiede haben die Online Consumer Reviews von China und Deutschland in der Tat durch Sentiment Analyse?
3. Welche Unterschiede gibt es zwischen der Praxis und der Theorie? Warum?
4. Welche Auswirkungen gibt es von diesen Unterschieden der Online Consumer Reviews in den beiden Ländern?

1.3 Die Struktur der Arbeit

Um diese Forschungsfragen zu beantworten, wird die Arbeit in der Struktur geschrieben, die in der Abbildung 1.4 gezeigt wird. Erst werden die deutschen und chinesischen OCRs

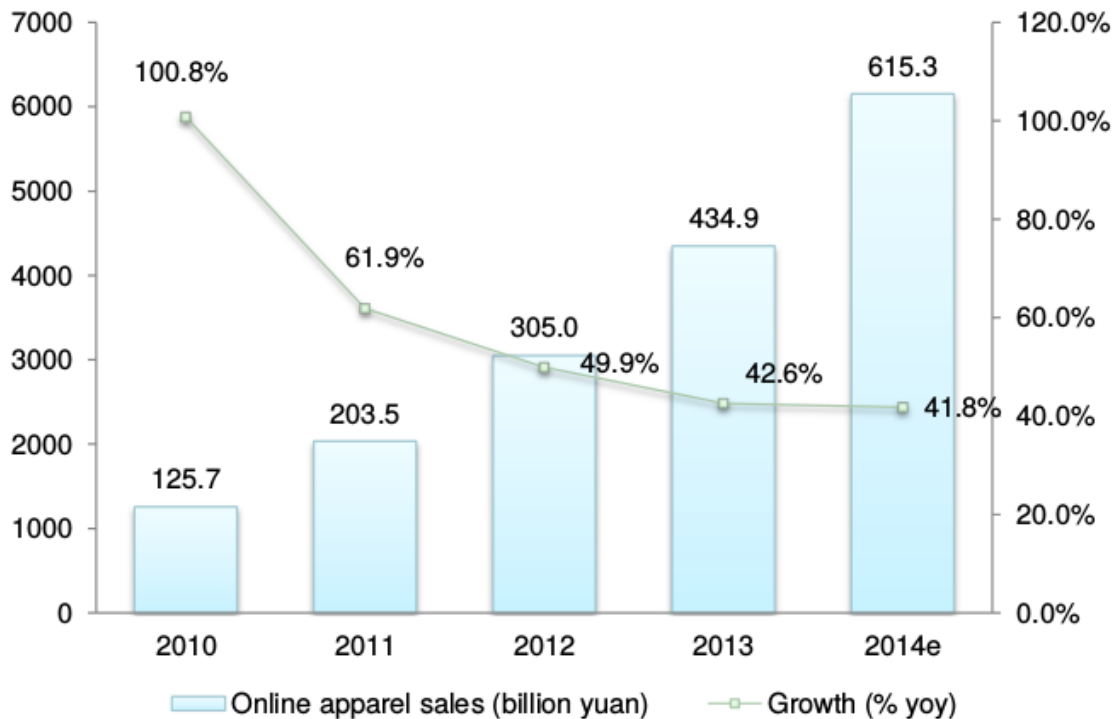


Abbildung 1.3: Online Umsatz für Kleidung in China, 2010 - 2014(Quelle:Fung & Fung, 2014)

gesammelt, dann werden sie durch die Sentiment Analyse analysiert und die Ergebnisse werden durch statistische Maßnahmen verglichen. Die Unterschiede werden mit den kulturellen Unterschieden diskutiert und begründet. Die Wirkungen dieser Unterschiede werden auch in verschiedenen Sichtweisen diskutiert.

Kapitel 2 In Kapitel 2 werden die theoretischen Grundlagen beschrieben. Hier werden erst die Online Consumer Reviews vorgestellt, einschließlich des Begriffs und Forschungsstands in der Marketingforschung, der Motive für das Schreiben der Online Consumer Reviews und der Wirkungen der Online Consumer Reviews im Textileinzelhandel. Danach werden die Definition, der Ablauf, der Kernprozess und der in dieser Arbeit verwendete Algorithmus der Sentiment Analyse eingeführt. Durch die Sentiment Analyse haben die Online Consumer Reviews einige statistische Attribute. Dann wird das sechsdimensionale Modell für Deutschland und China von Hofstede vorgestellt, und die Unterschiede zwischen den beiden Ländern werden herausgearbeitet. Die relevanten Forschungsergebnisse im kulturübergreifenden Bereich werden aufgeführt. Basierend auf diesen theoretischen Grundlagen werden vier Hypothesen für die Forschungsfragen aufgestellt.

Kapitel 3 Kapitel 3 stellt den Prozess der Datensammlung und Textverarbeitung vor. Die

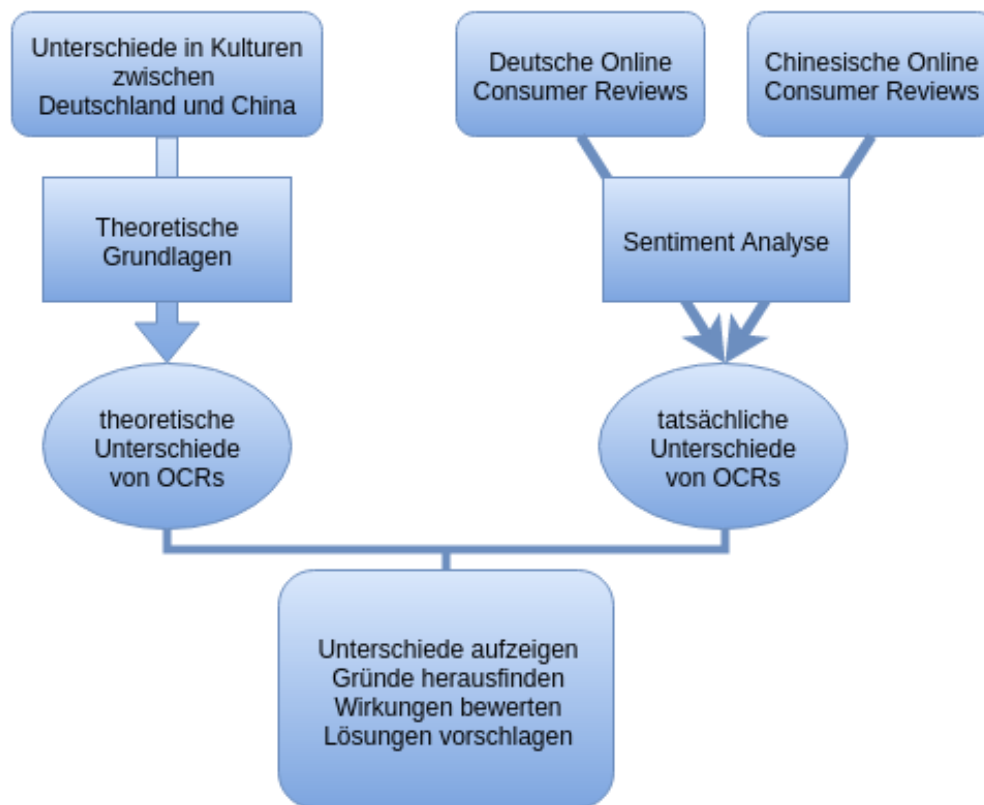


Abbildung 1.4: Die Struktur der Arbeit (Quelle: Eigene Darstellung)

zwei Business to Customer, (allgemein für Beziehungen zwischen Unternehmen und Konsumenten) (B2C) Plattformen Amazon Deutschland und Tmall (China) werden als Datenquellen dienen. Vier unterschiedliche Sportbekleidungen von Adidas, Nike und Puma werden als die Untersuchungsobjekte ausgewählt. Diese Untersuchungsobjekte müssen bestimmte Bedingungen erfüllen. Weiterhin werden die technischen Werkzeuge in der Datensammlung vorgestellt und der Datensammlungsprozess aus beiden Plattformen dokumentiert. Nach der Datensammlung werden die Rohdaten bestimmter Verarbeitungen unterzogen und die chinesischen Daten werden durch eine Übersetzungsmaschine auf Deutsch übersetzt, damit die Ergebnisse besser verglichen werden könnten. Die Einflüsse der Übersetzung werden auch in diesem Abschnitt diskutiert.

Kapitel 4 In diesem Kapitel werden die empirischen Ergebnisse deskriptiv erst gezeigt. Vor der Prüfungen der Hypothesen, muss man erst bestimmen, ob die Online Consumer Reviews statistisch normal verteilt sind. Dies ist die Basis dafür, welche Prüfverfahren man für die Hypothesen auswählen sollte. Je nach der Überprüfung der vier Hypothesen wird das Ergebnis gezeigt.

Kapitel 5 In Kapitel 5 wird erst der Zusammenhang der kulturellen Dimensionen und der Ergebnissen diskutiert, um die Gründe der Unterschiede zwischen China und

Deutschland zu entdecken. Und die Motive für das Schreiben der Online Consumer Reviews und die Wirkungen der Ergebnisse in Konsumenten-, Unternehmen- und Plattformenbetreiberssicht werden auch diskutiert.

Kapitel 6 In diesem Kapitel wird die Arbeit zusammengefasst, und einige Limitationen aus unterschiedlichen Gründen werden aufgestellt. Darüber hinaus wird der Ausblick auf die zukünftigen Forschung geführt.

2 Theoretische Grundlagen

2.1 Online Consumer Reviews in der Marketingforschung

2.1.1 Begriff

Online Consumer Reviews, oft auch als elektronische Mund-zu-Mund (auf Englisch (i.e.) electronic Word-Of-Mouth (eWOM)), Online-Bewertungen, Online-Empfehlungen, oder Online-Meinungen bezeichnet, hat an Bedeutung mit dem Aufkommen der neuen technologischen Werkzeugen gewonnen (Serra Cantallops & Salvi, 2014). Aber, OCRs haben noch kleine Unterschiede von eWOM Die Definition des eWOMs von Litvin *et al.* (2008) orientiert sich an der Definition von Westbrook (1987), und stellt nun klar: (i.e.)

“electronic Word-Of-Mouths are all informal communications directed at consumers through Internet-based technology related to the usage or characteristics of particular goods and services, or their sellers.”

Die Typologie von Litvin *et al.* ist zweidimensional: (a) Kommunikationsbereich: von eindeutig (E-Mails), mengenwertig (i.e. one to many) (Bewertungsseiten) oder Many-to-Many (virtuelle Gemeinschaften); (b) Level der Interaktivität: von Asynchron (E-Mails, Bewertungsseiten, Blogs) bis Synchron (Chatrooms, Newsgroups, instant messaging). Abbildung 2.1 zeigt die grafische Darstellung der Typologie.

Diese Typologie meint deutlich, dass die OCRs nur ein Unterbereich des eWOMs sind. In der Dimension des Kommunikationsbereich sind OCRs mengenwertig, und in Dimension der Interaktivität asynchron. Deshalb ist die Definition der OCRs in dieser Arbeit nach Hennig-Thurau *et al.* (2004):

“Online Consumer Review is any positive or negative statement made by potential, actual, or former customers about a product or company, which is made available to a multitude of people and institutions via the Internet”

Diese Definition unterscheidet Online Consumer Reviews deutlich von eWOM. Für diese Definition ist es wichtig, die Gefühle von den Reviews, die von der Kunden gemacht werden, zu identifizieren.

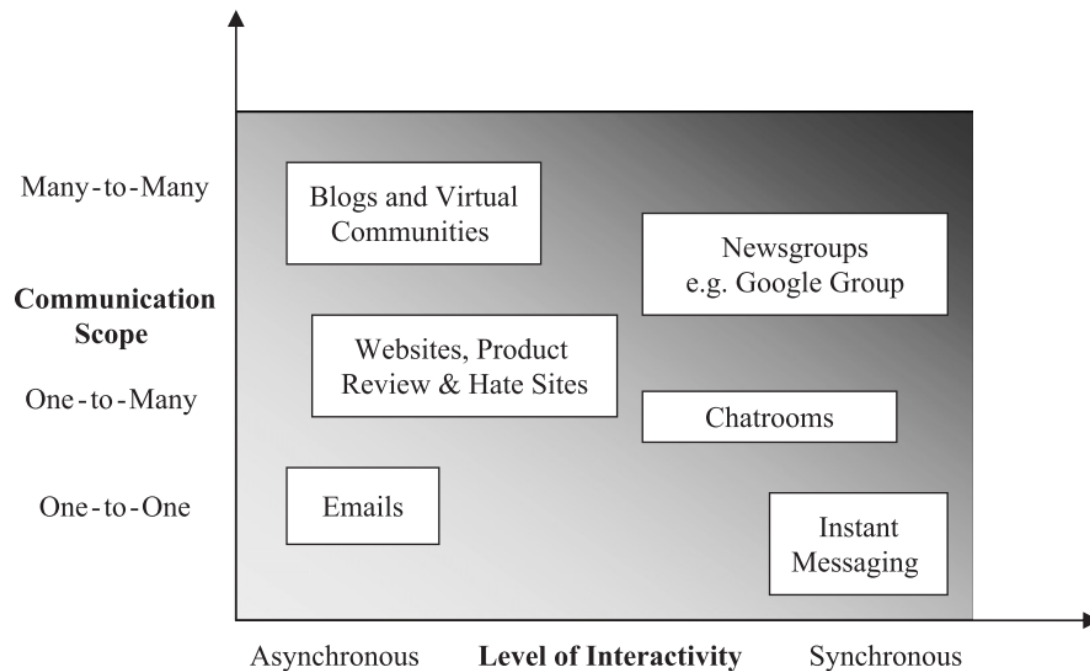


Abbildung 2.1: Die Typologie von eWOM (Quelle: Litvin *et al.*, 2008)

Nach Sridhar & Srinivasan (2012), werden die OCRs in zwei Gruppen unterteilt: qualitative und quantitative OCRs. Qualitative OCRs sind die schriftliche Beschreibungen der Kunden. Und die Quantitative OCRs sind die Werte der Online Ratings des Produkts. Im Falle einer quantitativen Online Consumer Review ist der Kunde gezwungen, seine oder ihre Auswertung in einer einzigen Bewertung oder Grad zusammenzufassen, und die einzelnen Ratings von Kunden werden in der Regel zusammen in eine Auswertungsstatistik gebündelt. (Kostyra *et al.*, 2015)

Abbildung 2.2 ist ein Beispiel der Online Consumer Review aus Amazon. Der quantitative Teil der Review im Beispiel ist diese "drei Sterne" in dem "Fünf-Sterne-Rating-System" in Amazon. Und die anderen schriftlichen Teile sind die qualitative Online Consumer Review.

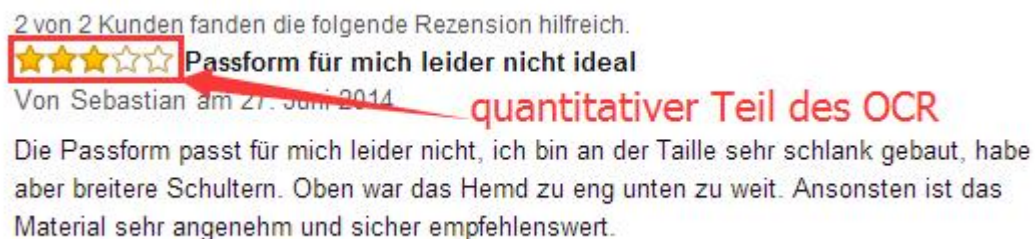


Abbildung 2.2: Ein Beispiel von OCRs, (Quelle: Amazon.de)

Gemäß Chintagunta *et al.* (2010) kann eine Online Consumer Review in die folgenden drei Attribute zerlegt werden:

1. **Volumen** ist die Gesamtzahl der Ratings von den Kunden.
2. **Valenz** ist die durchschnittliche Rating von den Kunden, und repräsentiert die durchschnittliche Kundenzufriedenheit.
3. **Varianz** ist die Variation der Ratings entlang der Rating-Skala und ist durch die Anzahl der Kundenbewertungen für jeden Valenzniveau beobachtbar. Varianz stellt den Grad der Meinungsverschiedenheiten oder Heterogenität der Auswertungen von den Kunden dar.

2.1.2 Forschungsstand zu Online Consumer Reviews

Es gibt zwei unterschiedliche Forschungswege OCRs zu studieren. Viele Wirtschaftswissenschaftler glauben, dass die OCRs eine Form von eWOM ist, und damit studieren sie die OCRs ähnlich wie eWOM. eWOM wird außerdem noch als ein Synonym von OCRs benutzt (Serra Cantallops & Salvi, 2014).

Der Begriff eWOM hat sich aus Word-Of-Mouth (WOM) entwickelt, die auf Deutsch Mund-zu-Mund heißt, ist ein Faktor, das in dem Prozess des Treffens der Entscheidungen von Kunden wichtig ist (Serra Cantallops & Salvi, 2014). Harrison-Walker (2001) definiert WOM als *“informal, person-to-person communication between a perceived noncommercial communicator and a receiver regarding a brand, a product, an organization, or a service.”* und Dick & Basu (1994) als *“a volitional post-purchase communication by consumers.”* Die meisten Studien analysieren WOM als Faktor, der zu einem größeren oder geringeren Grad den Verbraucher in der Auswahl von Produkten und Dienstleistungen beeinflusst (Serra Cantallops & Salvi, 2014).

Da sich die Kommunikationsumgebung geändert hat, wurde die Forschung über WOM aktualisiert (Vilpponen *et al.*, 2006). Zwar ist eWOM ähnlich wie die traditionelle Form des WOM, doch mit den einigen einzigartigen Eigenschaften. Häufiger ist eWOM zwischen Menschen zu beobachten, die wenig oder gar keine vorherige Beziehung zueinander haben (zum Beispiel Fremde) und kann anonym sein (Dellarocas, 2003, Goldsmith & Horowitz, 2006, Sen & Lerman, 2007). Diese Anonymität ermöglicht es den Verbrauchern, bequemer ihre Meinungen mitzuteilen, ohne Offenlegung ihrer Identität (Goldsmith & Horowitz, 2006). Die Bereitschaft von Einzelpersonen, ihre Meinung auszudrücken, kann durch die bestehenden öffentlichen Meinungen auf der Webseite beeinflusst werden, und zur gleichen Zeit wollten Personen auch, die öffentliche Meinung durch ihre Meinungen zu beeinflussen (Hong & Park, 2011). Dieses Phänomen ist viel häufiger in den Plattformen, die viele OCRs haben, weil die OCRs mengenwertig sind.

Diese einzigartigen Eigenschaften der OCRs ermutigen die Verbraucher, ihre Meinungen

mit mehreren anderen Verbrauchern zu teilen, so dass das Volumen der OCRs sich erhöht (Chatterjee, 2001). Dadurch gibt es eine größere Wahrscheinlichkeit, dass die Verbraucher andere Verbraucher mit Fachwissen über das Produkt in den Plattformen finden könnten (Duhan *et al.*, 1997).

Ein anderer Forschungsweg für die frühere Studien zum Verständnis der OCRs ist durch Online Rückkopplungsmechanismen (i.e. Online-Feedback-Mechanismus) (Dellarocas, 2003). Online Rückkopplungsmechanismen, die auch als Reputation-Systeme bekannt sind (Resnick *et al.*, 2000), sind unter Verwendung von bidirektionalen Kommunikationsfunktionen des Internet, künstlich die groß angelegte WOM-Netzwerke zu konstruieren, in denen die Menschen Erfahrungen und Meinungen über die vielfältigen Themen teilen können, inklusiv Unternehmen, Produkte, Dienstleistungen, und sogar Weltereignisse (Dellarocas, 2003).

Zum Beispiel, ist das Online-Feedback-System von eBay ein der am besten studierte Online Rückkopplungsmechanismen. Dieses System sammelt die Bewertungen von den Käufern und auch von den Verkäufern nach jedem Geschäft (Resnick & Zeckhauser, 2002). Abbildung 2.3 zeigt ein Beispiel von einem Käufer und Verkäufer. Resnick & Zeckhauser haben herausgefunden, dass diese Feedbacks fast immer positiv sind, und es gab eine hohe Korrelation der Feedbacks zwischen Käufern und Verkäufern.

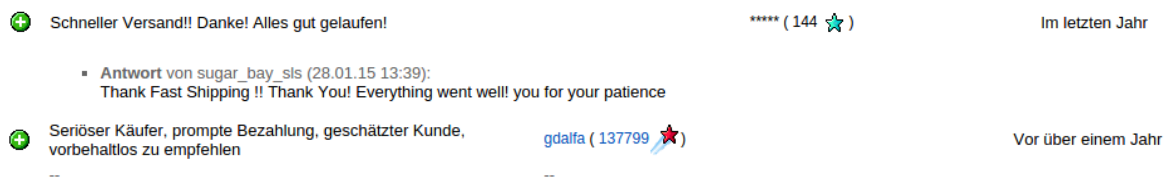


Abbildung 2.3: Ein Beispiel von Online Rückkopplungsmechanismen (Quelle: eBay.de)

Es gibt Anzeichen dafür, dass die Menschen zunehmend Meinungen in solchen Systemen schreiben wollen, um eine Vielzahl von Entscheidungen zu treffen, die sind, von welcher Film anzusehen, bis in welcher Aktien zu investieren (Guernsey, 2000). Erst vor fünf Jahren haben die gleichen Leute in erster Linie diese Entscheidungen getroffen, die auf Werbungen oder professionellen Beratungen basiert sind (Dellarocas, 2003).

Es ist interessant zu beobachten, dass die Forschungen durch diesen Weg die gleichen Ergebnisse wie Duhan *et al.* (1997), Hong & Park (2011) hervorgebracht haben. Die beiden Forschungswege werden in Abbildung 2.4 gezeigt.

Die bisherige Forschung über OCRs fokussiert sich im Allgemeinen auf zwei Forschungsrichtungen. Einerseits sind die Faktoren, die sich auf die Erzeugung von Kommentaren beziehen; und auf der anderen Seite ist die Wirkung der Kommentare (Serra Cantallops & Salvi, 2014). In Abschnitt 2.1.3 und 2.1.4 werden die beiden Forschungsrichtungen diskutiert.

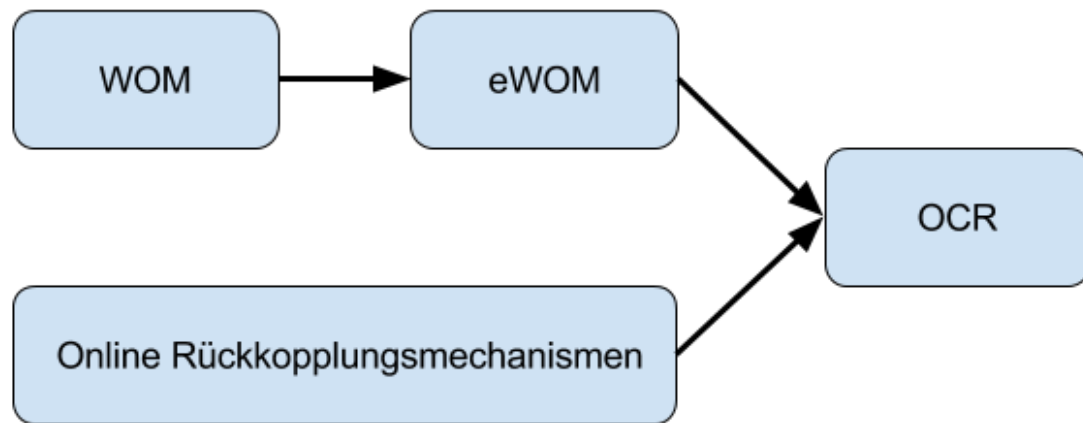


Abbildung 2.4: Die beide beobachtete Forschungswege über OCRs (Quelle: Eigene Darstellung)

2.1.3 Motive für das Schreiben von Online Consumer Reviews

In Bezug auf die Motive von OCRs, zeigen die meisten der untersuchten Studien Aspekte wie “Service-Qualität und Zufriedenheit”, “Fehler und Wiederherstellung”, “Unzufriedenheit der Kunden” und “Sinn für die Zugehörigkeit der Gemeinschaft” als Hauptmotive der Verbraucher für das Schreiben der Bewertungen (Kim *et al.* , 2009, Nusair *et al.* , 2011, Sánchez-García & Currás-Pérez, 2011, Sun & Qu, 2011, Swanson & Hsu, 2009). Diese Studien identifizieren eine direkte Beziehung zwischen Zufriedenheit oder Unzufriedenheit mit positiven oder negativen Bewertungen und ist ein ziemlich offensichtliches und vorhersehbares Konsumverhalten (Serra Cantallops & Salvi, 2014). Einige Autoren setzen die Themen bezüglich “Engagement”, “soziale Identität”, “Erwartungen vor dem Kaufen”, und “Kunden begeistert” als wichtige Aspekte bei den Motiven (Bronner & de Hoog, 2010, Casaló *et al.* , 2010, Crotts *et al.* , 2009).

Hennig-Thurau *et al.* (2004) machte eine Zusammenfassung von allen Motiven durch die Definition von WOM. Durch die Theorien von Dichter (1966), Engel *et al.* (1993), Sundaram *et al.* (1998) werden die Motive der Erstellung von OCRs in 11 Faktoren von Hennig-Thurau *et al.* unterteilt. Die sind: Sorge um andere Verbraucher; Unterstützung von Unternehmen; Erhalten von den sozialen Vorteilen; Ausübung von Macht; Postkaufberatung suchen; Selbstverbesserung; wirtschaftlichen Erfolg; Komfort bei der Suche nach Wiedergutmachung; Unterstützung der Plattformbetreiber; positive Emotionen auszudrücken; und negativen Gefühle zu erleichtern.

Diese 11 zusammenfassenden Faktoren wurden von rund 2000 Verbraucher, die aktiv auf den deutschen Web-basierten Meinungsplattformen sind, empirisch geprüft. Die Motivation, Sorge um andere Verbraucher, ist das Hauptmotiv. Danach folgen die wirtschaftliche Anreize und positive oder negative Gefühle auszudrücken. (Hennig-Thurau *et al.* , 2004)

2.1.4 Wirkungen von Online Consumer Reviews

Es gibt viele Aspekte über die Wirkungen von OCRs in Marketing. Dellarocas (2003) findet sich eine Auflistung von Forschungen, die Wirkungen der Valenz von OCRs aufzeigen, (das heißt, dass die OCRs positiv, negativ oder neutral sind), mit der Maßnahme Online Rückkopplungsmechanismen. Die Tabelle 2.1 zeigt diese Forschungen.

Quelle	Untersuchungs- objekte	Ergebnisse
Ba & Pavlou (2002)	Musik, Software	Positive Bewertungen erhöhen den geschätzten Preis, aber negativen haben keinen Effekt.
Bajari & Hortacsu (2003)	Münzen	Sowohl positive als auch negative Bewertungen beeinflussen die Wahrscheinlichkeit des Eintritts der Käufer in die Auktion, aber nur positive Bewertungen hatte einen signifikanten Einfluss auf den Endpreis.
Eaton <i>et al.</i> (2002)	Elektrische Gitarren	Negative Bewertung reduziert die Wahrscheinlichkeit auf den Verkauf, aber nicht auf den Preis der verkauften Produkte.
Houser & Wooders (2006)	Pentium-Chips	Positive Bewertung erhöht den Preis; negativer Kommentar reduziert Preis.
Kalyanam & McIntyre (2001)	Palm Pilot PDAs	Positive Bewertung erhöht den Preis; negativer Kommentar reduziert Preis.
Kauffman & Wood (2000)	Münzen	Keine besonderen Wirkungen, aber negativer Kommentar erhöht den Preis in der univariaten Analyse vielleicht.
Lee <i>et al.</i> (2000)	Computer-Monitore und Drucker	Negativer Kommentar reduziert den Preis, aber nur für die gebrauchte Ware.
Livingston (2005)	Golfklubs	Positive Bewertung erhöht sowohl die Wahrscheinlichkeit für den Verkauf als auch für den Preis; Effekt ausläuft, sobald ein Datensatz etabliert.
Lucking-Reiley <i>et al.</i> (2007)	Münzen	Keine Auswirkung von positiver Bewertung; negativer reduziert den Preis.

Melnik & Alm (2002)	Goldmünzen	Positive Bewertung erhöht den Preis; negativer Kommentar reduziert den Preis.
McDonald & Slawson (2002)	Puppen	Mehrere Bewertungen (positive oder negative) erhöht den Preis.
Resnick & Zeckhauser (2002)	MP3-Player, Beanie Babies	Beide Formen des Kommentars beeinflussen die Wahrscheinlichkeit für den Verkauf, aber es gibt keine Abhängigkeit auf den Preis.
Resnick <i>et al.</i> (2006)	Weinlesepostkarten	Geringe Wirkung von den kleinen Menge des negativen Kommentars.

Tabelle 2.1: Auswirkung der Valenz von OCRs, (Quelle: Dellarocas, 2003)

Die Auswirkungen der OCRs wurden in den letzten Jahren sowohl aus der Unternehmenssicht als auch der Konsumentensicht analysiert. Serra Cantallops & Salvi (2014) hat eine Zusammenfassung dafür im Bereich Hotelleitung gemacht, und fand, dass die Hauptauswirkungen von OCRs aus der Konsumentensicht: Entscheidungsprozess, Glaubwürdigkeit, Risikoreduktion, Produktakzeptanz, Loyalität, Markenbewusstsein, und so weiter (usw) sind. Aus der Unternehmenssicht sind die Auswirkungen häufig: Qualitätskontrolle und neue Produkte, Ertragsmanagement, Kundeninteraktion und Wiederherstellung, Kommunikation, die sie sich auf das Ziel konzentrieren, spezifische Marketing-Strategien, Online-Reputation, und Loyalitätserzeugung. Die Abbildung 2.5 stellt die Ergebnisse graphisch dar.

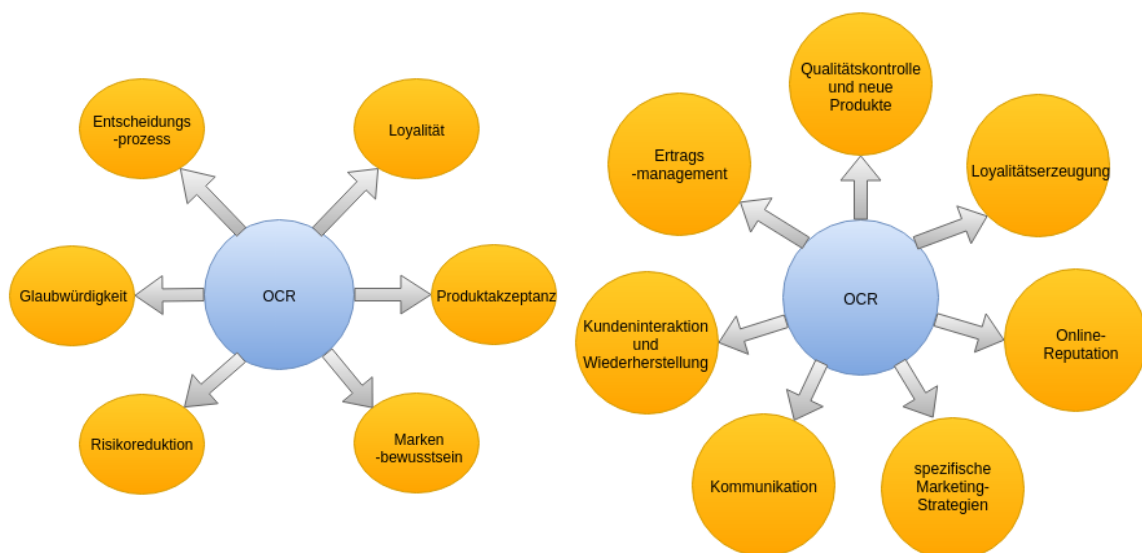


Abbildung 2.5: Auswirkungen der OCRs aus Konsumenten(Links)-und Unternehmenssicht(Rechts) (Quelle: Serra Cantallops & Salvi, 2014)

Cheung & Thadani (2012) beobachtete, dass es noch keine Forschung aus der Unternehmenssicht bis zum Jahr 2004 gab. Aber aus der Konsumentensicht gibt es schon ab 2001 Studien über OCRs. Im Jahr 2008 gab es insgesamt 15 Studien, 66% davon waren aus der Konsumentensicht, aber im Jahr 2010 war die Zahl nur 2, jedoch gab es 10 Studien aus der Unternehmenssicht. (Cheung & Thadani, 2012) Die Abbildung 2.6 zeigt die von Cheung & Thadani festgestellte Zeitachse von OCRs.

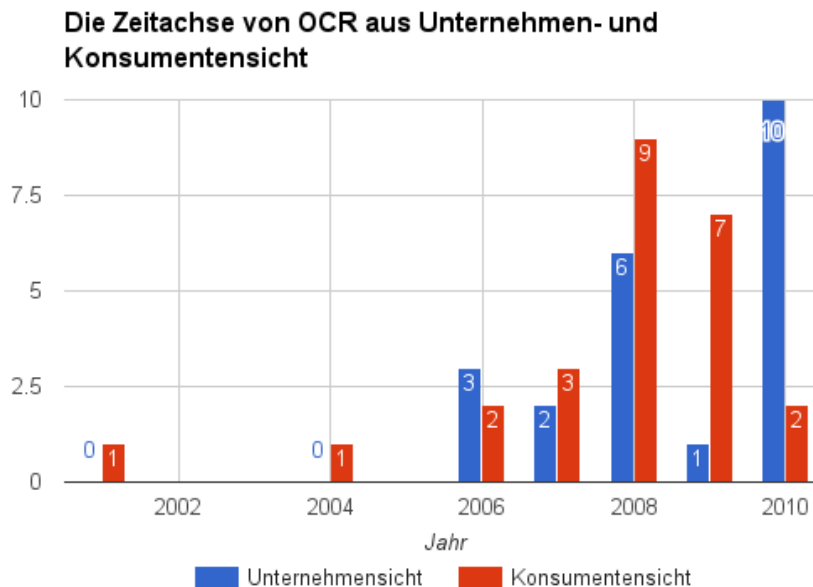


Abbildung 2.6: Die bemerkte Zeitachse von OCRs (Quelle: Cheung & Thadani, 2012)

Einige Studien, die die Auswirkungen der OCRs studieren, fokussieren sich auf die qualitativen und quantitativen Aspekten. In den qualitativen Bewertungen ist es dem Kunden völlig frei, wie Sie beschreiben, kritisieren und das Produkt bewerten (Jiménez & Mendoza, 2013). Im Falle einer quantitativen Bewertung ist der Kunde gezwungen, seine oder ihre Auswertung in einer einzigen Bewertung oder Grade zusammenzufassen, und die Einzelbewertungen von Kunden werden in der Regel in eine Auswertungsstatistik zusammengefasst (Kostyra *et al.*, 2015).

Diese Studien aber analysieren nur aus quantitativer Sicht. Sie haben schon viel über die Valenz (beispielsweise: 5 Sterne-Rating in Amazon), Volumen (Anzahl der OCRs), und Varianz (zeigt den Grad der Meinungsverschiedenheiten der Kunden) der OCRs in der quantitativen Aspekte studiert. Was der Kunde schreibt, ist für die Studien unwichtig. Tabelle 2.2 zeigt die Ergebnisse dieser Studien.

Quelle	Untersuchungs- objekte	Ergebnisse
--------	---------------------------	------------

Chen <i>et al.</i> (2011)	Kameras	Negative Bewertungen haben stärkere Auswirkung als positiven. OCRs zeigen eine abnehmende Auswirkung über die gesamte Lebensdauer.
Chevalier & Mayzlin (2006)	Bücher	Bessere Valenz von OCRs werden den Umsatz für Amazon verbessern, aber nicht für Barnes & Noble. Volumen verbessert den Umsatz für Amazon.
Chintagunta <i>et al.</i> (2010)	Filme	Valenz ist wichtig für die Vorhersage des Umsatzes, aber nicht Volumen oder Varianz.
Clemons <i>et al.</i> (2006)	Bier	Varianz ist signifikant mit dem Umsatzwachstum korreliert. Durchschnittliche Valenz ist signifikant.
Cui <i>et al.</i> (2012)	Videospiele und Elektronikläden	Valenz und Volumen sind signifikant. Valenz hat eine stärkere Auswirkung als Volumen, um Waren zu suchen. Aber wenn Kunde Erfahrung über die Ware hat, wechselt diese Beziehung.
Dellarocas <i>et al.</i> (2007)	Filme	Valenz und Volumen zeigen eine signifikante Beziehung mit dem zukünftigen Schalterverkauf.
Dhar & Chang (2009)	Musik	Valenz ist signifikant für eine Woche-voraus-Vorhersage in einem Modell mit festen Effekten. Volumen wird nicht implementiert.
Duan <i>et al.</i> (2008)	Filme	Valenz hat keine Effekte für den Umsatz, aber Volumen ist signifikant.
Ho-Dac <i>et al.</i> (2013)	DVD	OCRs haben eine signifikante Auswirkung auf die schwach aufgeladene Marken, aber keine signifikante Auswirkung auf die starken.
Moe & Trusov (2011)	Bath, Duft, und Schönheitspflege	Es gibt positive direkte Effekte der Valenz auf den Umsatz. Keine direkte Effekte von Varianz und Volumen auf den Umsatz.
Sun (2012)	Bücher	Auswirkung der Volumen ist positiv und signifikant für Amazon. Valenz ist bezeichnend für Amazon, aber nicht für Barnes & Noble. Interaktion mit der Valenz und Varianz ist signifikant.
Zhu & Zhang (2010)	Konsolenspiele	Valenz und Varianz sind signifikant für weniger bekannte Spiele mit einem Online-Modus. Volumen ist bezeichnend für Spiele mit dem Online-Modus.

Kostyra <i>et al.</i> (2015)	eBook-Reader	Valenz hat eine positive direkte Wirkung auf die Produktwahl. Volumen moderiert nur hohen Valenz. Varianz moderiert hohe und mittlere Valenz negativ.
------------------------------	--------------	---

Tabelle 2.2: Ein Überblick über die Literature in quantitativer Aspekte, in Anlehnung an Kostyra *et al.* (2015)

Es ist schwer eindeutige Zusammenhänge zu kommen, obwohl Wirtschaftler schon so viele Studien und Forschungen gemacht haben. Es wird einfach beobachtet, dass viele Ergebnisse in Tabelle 2.1 und 2.2 im Konflikt stehen, auch über das gleiche Untersuchungsobjekt oder in dem gleichen Bereich. Beispielsweise findet Chintagunta *et al.* (2010), dass Valenz wichtig für die Vorhersage des Umsatzes ist, aber Volumen oder Varianz nicht. Dagegen sagt Duan *et al.* (2008), dass Valenz keine Effekte für den Umsatz hat, aber Volumen ist signifikant. Kauffman & Wood (2000) findet dass negative Kommentare den Preis erhöhen, aber Lucking-Reiley *et al.* (2007) ist da ganz anderer Meinung.

Gründe für diese im Konflikt stehende Ergebnisse sind vielfältig. Es gibt meist die folgende drei verschiedene Gründe:

- Die Qualität und Glaubwürdigkeit zu bestimmen ist schwer für die potenzielle Kunden, weil die OCRs anonym sind (Chatterjee, 2001, Schindler & Bickart, 2005).
- Wegen der Anonymität haben einige Verkäufer versucht, OCRs zu verbessern durch die wirtschaftliche Anreize an die guten Bewertungen, und auch durch das Veröffentlichen ihrer eigenen Kommentaren (Chatterjee, 2001)
- Viele Konsumenten geben eine quantitative Bewertung, die trotz seiner oder ihrer Meinung (qualitative Bewertung) nicht passend ist. Das heißt, dass die quantitative Bewertungen nicht so glaubwürdig, wie viele Wirtschaftler geglaubt haben, sind.

Aus diesen Gründen lesen die Konsumenten häufig eine Vielzahl von OCRs durch, um die Qualität und Glaubwürdigkeit der Online-Information zu bestimmen (Greer, 2003). Eine Umfrage ergab, dass 90% der Verbraucher die OCRs lesen, während 83% von ihnen stimmt zu, dass OCRs ihre Kaufentscheidungen beeinflussen würden (Lu, 2015). Laut Zhu & Zhang (2010) hatten die OCRs einen stärkeren Einfluss auf weniger beliebte Produkten.

Und auf diese Gründe entsteht ein Bedarf für Unternehmen, Plattformbetreiber, Wirtschaftswissenschaftler, und auch die Konsumenten, die qualitative Kommentaren zu verstehen und zu analysieren. Das heißt, dass es eine theoretische Studie ist erforderlich, um das Wissen von Meinungen der Menschen gegenüber diesen Produkten, Services oder Attributen, zu extrahieren (Khan *et al.*, 2011).

2.1.5 Online Consumer Reviews im Textileinzelhandel

Ein Erfahrungsgut ist es, das nur erst nach dem Kauf des Produkts ausgewertet werden kann (Nelson, 1970). Einige Erfahrungsgüter wie Kleidung sind nun weitgehend, die über das Internet verkauft (Lee & Park, 2009). Der Einkauf der Kleidung erfordert "Touch-and-Feel" Auswertungen in einer Vor-Kaufphase. Da die Menschen die sensorischen Informationen unterschiedlich empfangen und verarbeiten, können sie unterschiedlichen Wahrnehmungen des gleichen Produkts haben (Armstrong *et al.*, 2014, p. 151). Frühere Studien haben mehrere Faktoren und Hinweise, die die Sicht der Verbraucher auf der Qualität der Kleidung beeinflussen, identifiziert.

Viele Studien kategorisieren die Attribute, die die Wahrnehmungen der Verbraucher bestimmen oder beeinflussen, in zwei allgemeine Typen - äußere und innere. Innere Attribute beziehen sich auf eigenen Eigenschaften des Produkts, die ohne das tatsächliches Produkt nicht geändert werden können, während die äußeren Attribute kein Teil des Produktes sind, sondern relevant (Olson *et al.*, 1972, p. 167).

Eckman *et al.* (1990) haben vor 1990 die Erkenntnisse zusammengefasst, dass der Preis und die Marke die am häufigsten genannte äußere Kriterien waren, während Produktzusammensetzung (Stil, Farbe/Design, Stoff, Aussehen) und Leistung (Pflege, Passform, Haltbarkeit und Komfort) die am häufigsten genannte innere Kriterien waren. Farbe/Muster, Stil, Stoff, Passform und Aussehen wurden als der wichtigsten Faktoren in dem In-Store-Kaufprozess der Kleidung identifiziert (Eckman *et al.*, 1990). Forsythe (1991) stellte fest, dass die Verbraucher die Auswertungen und Eindrücke von Kleidung gebildet waren, durch die Verwendung von äußeren Faktoren (Marke, Preis, usw.) und inneren Kriterien, Design, Stil, die Konstruktionsdetails, die hauptsächlichen Kriterien waren (Eckman *et al.*, 1990). Fiore & Damhorst (1992) konzentrierten ihre Studie über die innere Signale und haben gefunden, dass die ästhetische Attribute wie das Layout (Design) und Stoff am wichtigsten waren. Mit offenen Fragen und Content-Analyse, entdeckten Lennon & Fairhurst (1994), dass ästhetische Attribute die meisten identifiziert wurden, nach Leistung, äußeren Attributen und Nützlichkeit.

Abraham-Murali & Littrell (1995a) verwendeten Fokusgruppen-Interviews, um eine vollständige Liste der Attribute einschließlich physischer Erscheinung (Stoff, Farbe/ Muster/ Textur, Bau und Styling), die körperliche Leistungsfähigkeit (Stoff, Farbe, Pflege, Verarbeitung und Kleidungsstück), expressive Attribute ("Schaut auf mich", Angemessenheit, Kommentare von anderen) und äußere Attribute (Marke, Preis, Geschäft, Herkunftsland, und Service) zu erzeugen. Die physische Erscheinung wurde in meistens genannt (Abraham-Murali & Littrell, 1995a). Basierend auf dieser Liste, haben Abraham-Murali & Littrell (1995b) eine weitere Studie über die Sicht der Verbraucher in den Vor- und Nach-Kaufphase gemacht, und

festgestellt, dass Stoff und Bekleidungskonstruktion, Stil, Pflege, Verarbeitung und Kosten zu den wichtigsten Attributen im Allgemeinen gehörten.

Forsythe *et al.* (1996) untersucht dieses Thema durch Einkaufszentrum mit realen Produkten und realen Einkaufssituationen, und stellten fest, dass die Robustheit/ Haltbarkeit (inklusive Stoffkonstruktion im Lieferumfang) und Stil/ Ästhetik (Design, Styling, Aussehen) die Schlüsseldimensionen waren, die das Bewusstsein der Verbraucher am meisten beeinflussen. Haltbarkeit und Pflege waren wichtige Entscheidungsfaktoren, aber tragen nicht zu der Qualitätsbewertung bei. Dieses Ergebnis war im Widerspruch zu den meisten früheren Studien.

Swink & Hines (2006) haben herausgefunden, dass die Ästhetik der wichtigste Faktor war, gefolgt von Leistung (Haltbarkeit, Versorgung, usw.) und äußeren Kriterien. Aber keine innere Kriterien wurden identifiziert, aufgrund des Fehlens von momentan vorliegender Kleidung. Abraham-Murali & Littrell (1995a) und Swink & Hines (2006) stimmten zu, dass die Verbraucher immer multidimensionalen Kriterien verwenden, um die Kleidung zu bewerten.

Viele andere Studien berücksichtigten dieses Thema in verschiedenen Stufen. Rosenau & Wilson (2014) definierten die Qualitätsmessung der Verbraucher als "zweistufiges Verfahren" (p.263-264). Sie meinten, dass Ästhetik (Stil, Farbe, Stoff, Zierteile und passen) und Konstruktionsdetails wie Nähte, Stiche und Mustererkennung zuerst bewertet wurden, während die Haltbarkeit, Komfort, Pflege und das Aussehen in der zweiten Beurteilung waren. Viele Forscher haben festgestellt, dass die Verbraucher das Produkt mit dem Preis (Kosten) bewerten wurden (Kadolph, 2007, Keller *et al.*, 2011), aber die Ergebnisse für die Wichtigkeit der Kosten waren inkonsistent (Lu, 2015). Laut Goldsmith & Goldsmith (2002), standen die demographischen Variablen wie Alter und Geschlecht in keinem Zusammenhang mit dem Online-Bekleidungseinkauf.

Basierend auf früheren Studien, werden drei allgemeine Attributkategorien - Ästhetik, Performance, und äußere Attribute in dieser Studie verwendet:

1. **Ästhetik** inklusive Farbe, Muster, Stil (Silhouette, Modebewusstsein), Verarbeitung (Stiche, Nähte und Konstruktionsdetails), Vielseitigkeit (Anpassungsfähigkeit) und das Aussehen (wie das Kleidungsstück sieht für den Verbraucher).
2. **Performance** inklusive Passform (Größe), Komfort, Haltbarkeit und Pflege.
3. **Äußere Attribute** sind Marke, Preis und Service.

Diese drei Kategorien sind die Attribute, worüber die Kunden OCRs über die Kleidung schreiben wollen. Als Hu *et al.* (2008) erwähnt, werden die OCRs von Verbrauchern eine Hauptquelle für die Verbraucher als auch Unternehmen, um die Informationen über das Produkt zu zugreifen. Differenzen in den Kulturen haben ein unterschiedliches Konsumverhalten

in verschiedenen Ländern zu Folge (Keller *et al.* , 2011). Dieses Verhalten spiegelt sich auch im Schreiben der OCRs wider. Die kulturellen Differenzen werden im Abschnitt 2.4 diskutiert.

2.2 Sentiment Analyse

2.2.1 Definition

Sentiment Analyse (auch als Opinion Mining bekannt (Duenas-Fernández *et al.* , n.d., Tapia & Velásquez, 2014)) ist ein Teilgebiet des Text Minings, mit der Hauptaufgabe, die Meinungen von Web-Benutzer generierten Inhalten zu extrahieren (Balazs & Velásquez, 2016). Es existieren viele Definitionen, die verschiedene Bereiche und Ebenen der Granularität präsentieren i.e.:

Wilson *et al.* (2005) “the task of identifying positive and negative opinions, emotions and evaluations”.

Feldman (2013) “the task of finding the opinions of authors about specific entities”.

Vinodhini & Chandrasekaran (2012) “tracking the mood of the public about a particular product or topic”.

Cambria *et al.* (2013) “the task of polarity classification”.

In dieser Arbeit, wird die Meinung (i.e.:“Opinion”) laut der Definition von Liu (2010), als ein Tupel definiert:

$$(e_i, a_{ij}, s_{ijkl}, h_k, t_l)$$

, in dem die Bedeutungen sind:

e_i ist die Entität der i ten Meinung.

a_{ij} ist das j te Attribut von e_i .

h_k ist der k ten Autor der Meinung.

t_l ist das Zeitpunkt wann die Meinung abgegeben wird.

s_{ijkl} ist die Polarität der Meinung gegenüber das Attribut a_{ij} von der Entität e_i bei dem Autor h_k am Zeitpunkt t_l .

Deshalb ist die Hauptaufgabe von Sentiment Analyse: (i.e.)

“to find all the opinion tuples $(e_i, a_{ij}, s_{ijkl}, h_k, t_l)$ within a document, collection of documents (called corpus) or across many corpora.” (Liu, 2010)

Diese Polarität s_{ijkl} kann sowohl als positiv, negativ oder neutral, als auch numerisch dargestellt werden (Balazs & Velásquez, 2016). Wie zum Beispiel in dieser Arbeit wird 1 eine sehr negative Meinung sein, während 5 eine sehr positive Meinung ist. Auch wenn die Analyse nicht so viele Detailstufen erfordert, könnten die Attribute einer Entity entfallen und durch *GENERAL* statt a_{ij} bezeichnet werden (Balazs & Velásquez, 2016).

Es gibt noch andere Definitionen, wie die in Neviarouskaya *et al.* (2010) präsentiert wird, in dem die Autoren versuchen, emotionale Zustände, wie (i.e.) “anger”, “fear”, “joy”, oder “interest” nur in positiv oder negativ zu klassifizieren. In diesem Fall könnte das Modell von Liu (2010), durch Hinzufügen eines anderen Elements in das Tupel, die Informationen darstellen. (Balazs & Velásquez, 2016)

2.2.2 Der Ablauf der Sentiment Analyse

Der übliche Ablauf oder Prozess der Sentiment Analyse besteht aus einer Reihe von festgelegten Schritten (Arora & Srinivasa, 2014, Dey & Haque, 2009, Khan *et al.* , 2014), die Datensammlung (i.e. data acquisition), Texte-Vorverarbeitung (i.e. text preprocessing), Kernprozess (i.e. core process), Aggregation und Zusammenfassung der Ergebnisse und Visualisierung (i.e. aggregation and summarization of results and visualization) sind. Abbildung 2.7 zeigt den ganzen Ablauf.

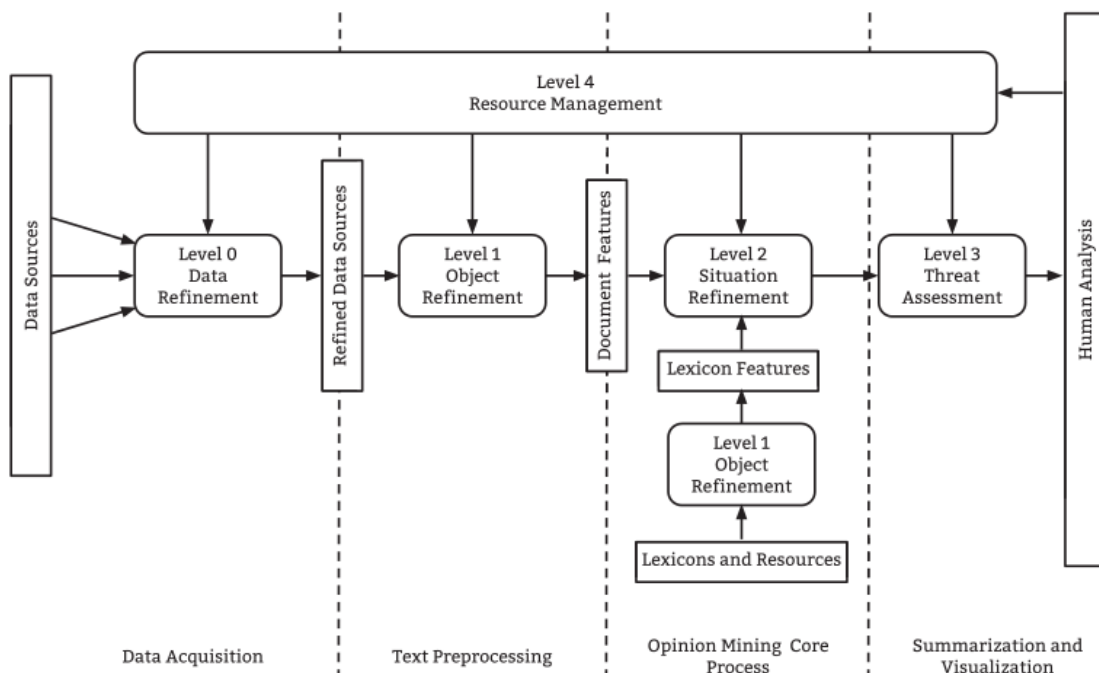


Abbildung 2.7: Der Ablauf der Sentiment Analyse (Quelle: Balazs & Velásquez, 2016)

In dieser Arbeit, werden die erste beide Schritten (beziehungsweise (bzw.) Datensammlung

und Texte-Vorverarbeitung) in Kapitel 3 diskutiert. Insbesondere wird der Kernprozess von Sentiment Analyse in Abschnitt 2.2.3 dargestellt. Die anderen Schritten, sowie Zusammenfassung und Visualisierung der Ergebnisse, werden in Kapitel 4 vorgestellt.

2.2.3 Kernprozess

Mit zustimmender Popularität der Sentiment Analyse, gibt es viele unterschiedliche Ebenen der Analyse, bzw. Dokument-Level, Satz-Level, Entität und Aspekt-Level usw (Balazs & Velásquez, 2016). In dieser Arbeit, wird die Analyse von dem Dokument-Level durchgeführt.

Dokument-Level: Sentiment Analyse auf dieser Ebene versucht, Dokumente in positiv oder negativ zu klassifizieren. Formal kann die Aufgabe von Sentiment Analyse in Dokument-Level als modifizierte Version, die in Abschnitt 2.2.1 vorgestellt, definiert werden und entspricht der Suche nach dem Tupeln:

$$(-, GENERAL, s_{GENERAL}, -, -)$$

, in dem die Entität e , der Autor, die Meinung h und die Zeit t , wann die Meinung geschrieben wird, bekannt oder ignoriert sein wird, und das Attribut a_j der Entität e als *GENERAL* entspricht. Das heißt, dass die Analyse nur die verallgemeinerte Polarität des Dokuments berücksichtigt. (Balazs & Velásquez, 2016)

In dieser Arbeit wird jede Review vom Kunden als ein Dokument analysiert. Gleichzeitig werden auch jede vom Kunden vergebene Sterne notiert. Das bedeutet, dass jede Review Sterne und Polarität von Sentiment haben. Es wird verglichen, welche Unterschiede es in den beiden und zwischen Deutschland und China gibt.

Satz-Level: Dieses Level ist analog zu dem vorherigen, da ein Satz als ein kurzes Dokument betrachtet werden kann. Allerdings wird ein zusätzlicher Vorverarbeitungsschritt erfordert, der aus Brechen des Dokumentes auf mehreren Sätzen besteht. Dieser Schritt stellt die ähnlichen Herausforderungen wie Tokenisierung, weil es Sprachen gibt, die nicht durch Punkte getrennt werden. (Balazs & Velásquez, 2016)

In dieser Arbeit ist das Level nicht von Bedienung, weil die chinesischen Reviews meistens nur ein Satz eingewiesen werden, außerdem werden sie auch als ein Dokument schon in dem Dokument-Level analysiert.

Entität und Aspekt-Level: Dieses Level stellt die detaillierteste Ebene dar, auf der die Sentiment Analyse durchgeführt wird. Hier ist die Aufgabe nicht nur um die Polarität der Meinung, sondern auch sein Ziel (Entität, Aspekt odr beides) zu finden, damit die in Abschnitt 2.2.1 beschriebene 5-Tupel-Definition vollständig gilt. Sowohl auf Dokument-Level als auch Satz-Level funktioniert die Analyse gut, wenn die geprüften

Texte nur eine Entität und Aspekt enthalten, aber sie sind ins Wanken geraten, wenn mehr vorhanden sind (Feldman, 2013). Die Aspekt-basierte Sentiment Analyse versucht, dieses Problem zu lösen, durch Erfassen jedes erwähnten Aspekt im Text und in Verknüpfung mit einer Meinung. (Balazs & Velásquez, 2016)

Die wichtigen Aspekte sind nach zwei Beobachtungen identifiziert: (Yu *et al.* , 2011)

1. Die wichtigen Aspekte eines Produktes werden üblicherweise durch eine große Zahl von Verbrauchern kommentiert.
2. Die Meinungen zu wichtigen Aspekten von Verbrauchern beeinflussen stark ihre allgemeine Meinung über das Produkt.

Wegen der Besonderheiten des elektronischen Einzelhandels werden die Autoren der Reviews sinnlos sein, weil die andere Kunden die Autoren der Reviews nicht kennen könnten.

Es gibt zwei gut etablierte Ansätze, zur Durchführung des Kernprozesses der Sentiment Analyse. Einen davon ist der unbeaufsichtigte Lexikon-basierten Approach (i.e. unsupervised lexicon-based approach), in dem sich der Prozess auf die, aus Sprachkenntnisse erhaltene, Regeln und Heuristiken stützt (Vilares *et al.* , 2015). Und der Andere ist der beaufsichtigte Maschine-Learning-Approach (i.e. machine learning approach), in dem die Algorithmen die zugrundeliegenden Informationen lernen, die aus vorher kommentierten Daten sind, damit sie die neuen, unmarkierten Daten klassifizieren können (Pang *et al.* , 2002).

Der unbeaufsichtigte Lexikon-basierte Ansatz: (i.e. “the unsupervised lexicon-based approach”) auch als Semantik basierter Approach bezeichnet, versuchen, die Polarität des Textes unter Verwendung einer Reihe von den aus Sprachkenntnisse erhaltenen Regeln und Heuristiken zu bestimmen (Balazs & Velásquez, 2016).

Die üblichen Schritte, um sie durchzuführen, sind: (1) jedes Wort und jeden Satz mit der entsprechenden Polarität mit Hilfe eines Lexikons zu klassifizieren. (2) die Analyse der Sentiment-Schiebern sowie deren Umfang (Verstärker und Verneinung) zu übernehmen.(3) Die adversative Klauseln (aber-Klauseln) durch das Verständnis, wie sie die Polarität beeinflusst zu behandeln, und dies in der letzten Sentiment-Score zu reflektieren (Liu, 2012).

Der beaufsichtigte Maschine Lerning Ansatz: (i.e. “the supervised machine learning approach”) auch als statistische Methoden zur Sentiment-Klassifikation bekannt. Die bestehen aus Algorithmen, die zugrundeliegenden Muster von Beispieldatei lernen (Rebolledo *et al.* , 2010), das heißt, die Daten, deren Klasse oder Label für jedes Exemplar bekannt sind, die später versuchen, neue, nicht-markierte Daten zu klassifizieren (Mitchell, 1997).

Die üblichen Schritte bestehen aus Projektierung der Merkmale, um das Objekt, dessen

Klasse prognostiziert wird, darzustellen, und dann mit ihrer Darstellung als Eingabe des Algorithmus (Balazs & Velásquez, 2016). Einige Merkmale, die in der Sentiment Analyse häufig verwendet werden, sind: Wortfrequenz, Wortart, die Stimmung der Wörter und Sätze, Regeln der Meinung, Sentiment-Schiebern und syntaktischen Abhängigkeit (Joshi & Penstein-Rosé, 2009, Liu, 2012).

In dieser Arbeit wird die unbeaufsichtigte Methode verwendet. Einer der Vorteile der Verwendung davon ist, nicht mehr auf große Datenmengen für die Ausbildung der Algorithmen zu beruhen. Dennoch ist es notwendig, ein Sentiment-Lexikon zu erhalten oder zu erstellen. Die unbeaufsichtigten Methoden sind auch weniger domainabhängig als die beaufsichtigten Verfahren. In der Tat haben die in einer Domain ausgebildeten Klassifikationen konsistent schlechtere Performance in anderen Bereichen. (Aue & Gamon, 2005, Blitzer *et al.* , 2007)

2.2.4 Der konkrete Algorithmus im Dokument-Level

Wie es oben gezeigt wird, ist es wichtig, ein gutes Sentiment-Lexikon zu erhalten. Die Genauigkeit der Polarität ist sehr abhängig von dem Sentiment-Lexikon für die unbeaufsichtigten Methoden. In dieser Arbeit wird das Lexikon von Remus *et al.* (2010) verwendet. Tabelle 2.3 zeigt einen Überblick über den Inhalt des Lexikons.

		Positiv	Negativ
Adjektive	Stammform	784	698
	Flexion	11.782	10.604
Adverb	Stammform	6	4
	Flexion	0	0
Nomen	Stammform	584	686
	Flexion	521	806
Verb	Stammform	312	430
	Flexion	2.453	3.100
Alle	Stammform	1.650	1.818
	Flexion	14.756	14.510
	Total	16.406	16.328

Tabelle 2.3: Überblick über den Inhalt des Lexikons. (Quelle: Remus *et al.* , 2010)

Bei der Messung der allgemeinen Einigung von Rater wird es mit Cohen *et al.* (1960) in einem Freirandvariante Brennan & Prediger (1981) durchgeführt. Die Interrater-Reliabilität ist $k_{free} = 0,76$ und damit gilt sie als zuverlässig. (Remus *et al.* , 2010)

Die Polaritätsgewichte wurden bei der von Church & Hanks (1990) vorgeschlagenen Anwendung eines Verfahrens abgerufen: die so genannte (i.e.) Pointwise Mutual Information (PMI). Dieser Ansatz wurde erfolgreich für die Sentiment Analyse wiederverwendet (Turney, 2002, Turney & Littman, 2003). Ihrer allgemeine Strategie ist es, Semantische Orientierung (SO)

von semantischen Assoziation zu folgern. Die SO eines gegebenen Wort w wird aus der Stärke ihrer Assoziation A mit einer manuell ausgewählten Menge der positiven Saatgut-Wörtern P minus der Stärke ihrer Verbindung mit einer Reihe von negativen Saatgut-Wörtern N berechnet (siehe Gleichung 2.1). (Remus *et al.* , 2010)

$$SO-A(w) = \sum_{p \in P} A(w, p) - \sum_{n \in N} A(w, n) \quad (2.1)$$

Das Wort w wird als eine positive SO bei $SO-A(w)$ positiv und negative SO bei $SO-A(w)$ als negative klassifiziert. Der absolute Wert des $SO-A(w)$ kann als die Stärke ihrer SO werden. (Remus *et al.* , 2010)

Parallel zu den Paradigmen von Turney & Littman (2003), benutzen Remus *et al.* (2010) die folgenden deutsche Saatgutsets P_{de} und N_{de} :

$$P_{de} = \left\{ \begin{array}{l} \text{gut, schön, richtig,} \\ \text{glücklich, erstklassig,} \\ \text{positiv, großartig, ausgezeichnet,} \\ \text{lieb, exzellent, phantastisch} \end{array} \right\} \quad (2.2)$$

$$N_{de} = \left\{ \begin{array}{l} \text{schlecht, unschön, falsch,} \\ \text{unglücklich, zweitklassig,} \\ \text{negativ, scheiße, minderwertig,} \\ \text{böse, armselig, mies} \end{array} \right\} \quad (2.3)$$

Die semantische Assoziationen $A(w, p)$ und $A(w, n)$ werden dann unter Verwendung des PMI berechnet. Der PMI zwischen zwei Wörtern w_1 und w_2 , welche nach Church & Hanks (1990) definiert ist, ist diese welche in Gleichung 2.4 gegeben wird.

$$PMI(w_1, w_2) = \log_2 \left(\frac{P(w_1 \& w_2)}{P(w_1) \cdot P(w_2)} \right) \quad (2.4)$$

$P(w)$ ist die Wahrscheinlichkeit, dass w auftritt, und $P(w_1 \& w_2)$ ist die Wahrscheinlichkeit, dass w_1 und w_2 gemeinsam auftreten. Diese Wahrscheinlichkeiten wurden unter Verwendung von Frequenzen und Kookkurrenz Statistiken über einen internen deutschsprachigen Korpus geschätzt, der aus rund 100 Millionen Sätze besteht. (Remus *et al.* , 2010)

Alle Gewichte sind in dem Intervall $[-1, 1]$ und auf 4 Dezimalstellen gerundet, mit $+1, 0$ absolut positiv und $-1, 0$ absolut negativ. Sehr positive Wörter sind beispielsweise Freude mit einem Gewicht von 0,6502 und perfekt mit einem Gewicht von 0,7299. Sehr negative Wörter sind, zum Beispiel betrügen mit einem Gewicht von -0,743 und schädlich mit einem Gewicht von -0,9269. Die Verteilung der absoluten Gewichte im Sentiment Wortschatz (siehe

Abbildung 2.8) folgt einer Zipf-artige Verteilung (Zipf, 1949): Sehr wenige Wortformen haben hohe Gewichte, einige Wortformen sind in der Mitte und eine große Menge von Wortformen haben ein wenig oder ein sehr wenig Gewichte. (Remus *et al.* , 2010)

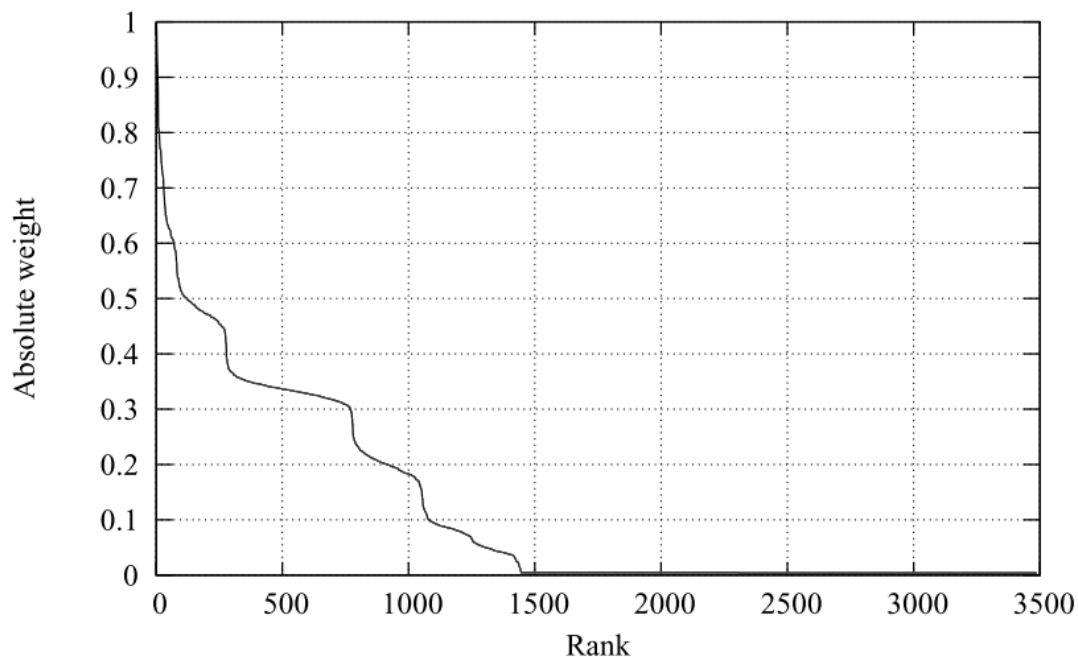


Abbildung 2.8: Die Verteilung der absoluten Gewichte. (Quelle: Remus *et al.* , 2010)

2.3 Deutschland und China: ein kulturübergreifender Vergleich

Deutschland und China unterscheiden sich deutlich im kulturellem Bereich. Das heißt, dass dieser Vergleich, der in dieser Arbeit gemacht wird, nicht nur länderübergreifend sondern auch kulturübergreifend ist.

Unter den vielen verschiedenen Möglichkeiten, wie die Kultur von Forschern klassifiziert wurde, sind die Theorie von Hofstede (2001) unter den Forscherinnen und Forschern am meisten verbreitet und zitiert (Fong & Burton, 2008). Hofstede (2001) definiert die Kultur als **“the collective programming of the mind, which distinguishes the members of one group from another”** (Hofstede, 2001, p. 9).

“Kultur” ist ein nationales Phänomen ist, die die unterschiedlichen Verhalten und Kenntnisse der Leute in den unterschiedlichen Ländern klassifizieren kann. Studien haben, dass Kultur einen großen Einfluss, auf die Kenntnisse und Verhalten der Konsumenten hat (McCort

& Malhotra, 1993, Triandis, 1972). In dieser Meinung machten schon viele Wirtschaftler erfolgreich. Zum Beispiel: Sia *et al.* (2009) hat herausgefunden, dass sich die Kultur von Online-Konsumenten die Wirksamkeit von verschiedener Vertrauen-Errichtung-Strategien auf ihren Überzeugung signifikant mäßigt. Und Mazaheri *et al.* (2011) hat auch festgestellt, dass die Auswirkung von dem Inhalt der Webseite und der Attitude des Services über die Absicht vom Kaufen durch die Orientierung der Kultur der Online-Konsumenten gemäßigt sein wird.

2.3.1 Kulturelle Dimensionen

Eine Dimension ist ein Aspekt der Kultur, der sich von anderen Kulturen unterscheidet. Nach Hofstede (1991) gibt es die folgendermaße benannten Diemensionen: *Machtdistanz (von gering bis groß), Kollektivismus gegenüber Individualismus, Femininität gegenüber Maskulinität, Unsicherheitsvermeidung (von schwach bis stark) und langfristige Orientierung gegenüber kurzfristiger Orientierung, Nachsicht gegenüber Beherrschtheit*. Zusammen bilden sie ein sechsdimensionales Modell von unterschiedlichen Kulturen. In diesem Modell hat jede Kultur eine Punktzahl in jeder Dimension, und die Kultur wird auch dadurch gekennzeichnet. (Hofstede, 1991, p. 29).

In den folgenden Abschnitte 2.3.2 bis 2.3.6 werden diese Dimensionen detaillierter. Darüber hinaus werden die wichtigen Unterschieden in den jeweiligen Dimensionen im Bereich der OCRs sowie Meinungen oder wirtschaftlichen Aspekten erfasst.

Die Logik von Gesellschaften entspricht aber nicht der Logik der die Gesellschaften betrachtenden Individuen. Vielleicht scheint keine logische Notwendigkeit für die Verknüpfung von Menschen zu bestehen, aber jede Dimension erfasst eine Reihe von Phänomenen in der Gesellschaft, die empirischen Untersuchungen zufolge in Kombination auftreten. Viele verschiedene Aspekte einer Phänomene werden zusammengefasst, diese treten in Kombination, und in trennbaren Verbindungen auf. In manchen Gesellschaften oder Kulturen gibt es einen Trend oder Aspekt, das gegen den allgemeinen Trend ist, der in den meisten anderen Gesellschaften gefunden wird.(Hofstede, 1991, p. 29)

2.3.2 Machtdistanz

Nach Hofstede (1991) kann man Machtdistanz auch definieren als *das Ausmaß, bis zu welchem die weniger mächtigen Mitglieder von Institutionen bzw. Organisationen eines Landes erwarten und akzeptieren, dass Macht ungleich verteilt ist*. Institutionen wie Familie, Schule und die Gemeinschaft bilden die Hauptelemente einer Gesellschaft; unter Organisation ist der Ort zu verstehen, wo die Leute arbeiten. (Hofstede, 1991, p. 42)

Geringe Machtdistanz	Große Machtdistanz
Ungleichheit unter den Menschen sollte so gering wie möglich sein.	Ungleichheit unter den Menschen wird erwartet und ist erwünscht.
Tendenz zu Dezentralisation.	Tendenz zu Zentralisation.
Mitarbeiter erwarten, in Entscheidungen miteinbezogen zu werden.	Mitarbeiter erwarten, Anweisung zu erhalten.
Der Einsatz von Macht muss legitimiert sein und wird danach beurteilt, was gut und was böse ist.	Macht geht vor Recht. Wer die Macht hat, ist legitimiert dazu und ist gut.
Alle haben die gleichen Rechte.	Die Mächtigen genießen Privilegien.
Hierarchische Struktur in einer Organisation bedeutet eine ungleiche Rollenverteilung aus praktischen Gründen.	Hierarchische Struktur in Organisationen sind ein Spiegelbild einer Ungleichheit von Natur aus zwischen oberer und unterer Schicht.
Die vorherrschenden Religionen und die philosophischen Systeme betonen die Gleichheit	Hierarchie und Einteilung der Gesellschaft in Klassen wird von Religion und philosophischer Gedankenwelt begünstigt.
Ausgeprägte Parteienlandschaft. Parteien der Mitte sind stark, extreme Links- und Rechtsparteien schwach.	Parteienspektrum schwach ausgeprägt. Schwaches Zentrum, starke Links- und Rechtsparteien.

Tabelle 2.4: Die wichtige Unterschiede zwischen Gesellschaften mit geringer und großer Machtdistanz (Quelle: in Anlehnung an Hofstede, 1991, p. 52, 57)

Tabelle 2.4 zeigt die wichtigen Unterschiede von Hofstede zwischen Gesellschaften mit geringer und großer Machtdistanz. Diffusionsforschung hat eine negative Beziehung zwischen der Bewertung von Machtdistanz und der Diffusionsrate des Produkts festgestellt (Van Everdingen & Waarts, 2003, Yenyurt & Townsend, 2003). Smith *et al.* (1995) zeigen, dass Kulturen mit geringer Individualismus und hohe Machtdistanz tendenziell eine geringere interne Kontrollüberzeugung haben. Diese Befunde legen nahe, dass Machtdistanz einen negativen Einfluss auf Produktdiffusion und -bewertung haben.

In Kulturen mit starker Machtdistanz, sind Ungleichheiten in der Gesellschaft erwartet. Diese Erwartung fördert die Vorstellung, dass Informationsaustausch auch ungleich ist. Die Mächtigen erwarten, dass mehr Informationen als Menschen mit weniger Macht zu halten. Diejenigen, die in der Machtdistanz eine geringe Punktzahl sind, geben ihre Ideen und Ansichten zum anderen eher, weil sie dazu neigen, jeden als gleich anzusehen. Diejenigen, die in hohen Machtdistanzpunkten sind, zeigen sich zurückhaltender in ihren Gruppeninteraktionen, vor allem im Umgang mit Menschen, die mächtiger sind. Ungleichheit der Macht fordert Mitglieder mit unterschiedlichen Machtsstufen zu verschiedenen Gruppen. In diesem Umfeld entsteht mehr Interaktionen (zum Beispiel: OCRs) mit denen, die zu den gleichen Machtsstufen gehören. (Lam *et al.* , 2009)

Es gibt noch einige Unterschiede dazwischen, aber sie sind nicht so wichtig in diesem

Bereich, welche diese Arbeit studiert. In Zusammenhang mit OCRs gibt es von Kunden und Verkäufern, und zwischen Kunden unterschiedliche Machtdistanz in verschiedenen Kulturen. Dazu beeinflusst diese Dimension die Inhalte und die Polarität der Gefühle in den OCRs groß.

2.3.3 Individualismus gegenüber Kollektivismus

Die als Individualismus gegenüber Kollektivismus bezeichnete Dimension ist folgendermaßen definiert: *Individualismus beschreibt Gesellschaften, in denen die Bedingung zwischen den Individuen locker sind: man erwartet von jedem, dass er für sich selbst und seine unmittelbare Familie sorgt. Sein Gegenstück, der Kollektivismus, beschreibt Gesellschaften, in denen der Mensch von Geburt an in starke, geschlossene Wir-Gruppen integriert ist, die ihn ein Leben lang schützen und dafür bedingungslose Loyalität verlangen.* (Hofstede, 1991, p. 67)

Nachfolgende Forschung von Greenfield (2000) betrachtet, dass diese Dimension die "Tiefenstruktur" unter diesen sechs Dimensionen ist, und einige weitere Studien (Sia *et al.*, 2009, Triandis, 2001) stellen vor, dass es die wichtigste Dimension ist, um die Unterschiede verschiedener Gesellschaften oder Nationen zu erklären.

Diese Dimension ist vielleicht die am häufigsten verwendete Dimension der kulturellen Variabilität der kulturübergreifenden Verbraucherforschung (Aaker & Maheswaran, 1997, Han & Shavitt, 1994). Mitglieder einer individualistischen Kultur, sowie der USA, neigen dazu, sich selbst als unabhängig von anderen, und sich auf Eigenständigkeit, interne Attribute, Getrenntheit und Entfernung von in-Gruppen zu betrachten (Singelis, 1994). Im Gegensatz dazu, betonen die Mitglieder in einer kollektivistischen Kultur wie China und Korea die Interdependenz und Wert auf ein harmonisches Arrangement, soziale Normen, Verbundenheit und in-Gruppenmitgliedschaften (Singelis, 1994).

Hofstede hat auch herausgefunden, dass diese Dimension, Kollektivismus gegenüber Individualismus, die Wahl und Entscheidungsfindung des Verbrauchers beeinflusst. Entscheidungen über das Konsumverhalten sind selten rein individuell. In kollektivistischen Kulturen, werden individuelle Entscheidungen im Konsens mit der Gruppe gemacht und es gibt keine rein individuellen Entscheidungen zu treffen. (Hofstede, 1998, p. 65)

In kollektivistischen Kulturen, sind die Menschen eher sich selbst in Bezug auf die Gruppenmitgliedschaft zu berücksichtigen und legen an erster Stelle das Wohl (Triandis, 1994). Im Gegensatz dazu sind die Menschen in individualistischen Kulturen eher selbst als autonome zu denken und machen die individuellen Interessen an erster Stelle (Shweder, 1990). In kollektivistischen Kulturen sind die nicht-lebensbedrohenden Verletzungen der sozialen Verantwortung eher moralisch betrachtet, während sie als Dinge der persönlichen Wahl in

individualistischen Ländern zu betrachten ist (Miller *et al.* , 1990). In individualistischen Kulturen haben die persönlichen Ziele in der Regel Vorrang vor Gruppenziele, aber in kollektivistischen Kulturen sind eher die Gruppenziele im Vordergrund (Triandis, 1994).

Im Bereich OCRs oder eWOM haben Studien bisher noch wenige Ergebnisse. Luo *et al.* (2014) haben herausgefunden, dass durch die stärkere Wirkung der zweiseitigen Informationen (im Vergleich zu einseitigen Informationen) die Wahrnehmung von Glaubwürdigkeit der Informationen der eWOM Leser gestärkt wird, wenn die Leser mit individualistischen Kulturen sind, verglichen mit denen, die mit kollektivistisch kulturellen Orientierung sind. Je höher kollektivistische Kultur die eWOM Leser bekennen, desto stärker wirken sich die Informationskonsistenz und die Informationsbewertung auf ihre Wahrnehmung von Informationsglaubwürdigkeit.

2.3.4 Maskulinität gegenüber Femininität

Nach Hofstede (1991), ist diese Dimension Femininität gegenüber Maskulinität zu nennen, da *diese Dimension die einzige ist, bei der die männlichen und die weiblichen Angestellten von IBM durchweg verschiedene Punktwerte erzielten*. Nur bei der Dimension wurde ein solcher Unterschied unter den Geschlechtern festgestellt. (Hofstede, 1991, p. 101)

“Legt man die Informationen über Unterschiede zwischen Gesellschaften bei dieser Dimension zugrunde, so kommt man zu folgender Definition: *Maskulinität* kennzeichnet eine Gesellschaft, in der die Rollen der Geschlechter klar gegeneinander abgegrenzt sind: Männer haben bestimmt, hart und materiell orientiert zu sein, Frauen müssen bescheidener sensibler sein und Wert auf Lebensqualität legen. *Femininität* kennzeichnet eine Gesellschaft, in der sich die Rollen der Geschlechter überschneiden: sowohl Frauen als auch Männer sollten bescheiden und feinfühlig sein und Wert auf Lebensqualität legen.” (Hofstede, 1991, p. 101)

Feminine Kulturen werden durch eine stärkere Beziehungsorientierung gekennzeichnet. Für sie sind die Lebensqualität und die Menschen wichtiger. Sie betonen, wer eine Person ist, und sie arbeiten, um nicht als die andersherum zu leben. (Schumann *et al.* , 2010)

Maskuline Kulturen werden durch einen stärkeren Ego-Orientierung gekennzeichnet, so dass die Menschen sich selbst und ihre Daseinsberechtigung nach ihrer Arbeit und Geld oder Gegenstände definieren (Schumann *et al.* , 2010). Wegen der materialistischen und Besitz orientierten Charakter der maskulinen Kulturen haben Forscher vorgeschlagen und fand Beweise für die höheren Ebenen der Informationsaustausch sowie Informationserfassungsaktivitäten (Dwyer *et al.* , 2005, Lam *et al.* , 2009, Liu *et al.* , 2001). Gegenstände sind in maskulinen Kulturen hoch geschätzt, weil sie Erfolg und Status widerspiegeln. Menschen definieren sich viel mehr durch ihren Besitz und deshalb legen sie mehr Nachdruck auf die Information mit diesen Beziehungen (Dwyer *et al.* , 2005).

Foscher haben deswegen vorgeschlagen, dass die Maskulinität Einflüsse auf die Erzeugung und Auswirkung der WOM haben, aber diese Hypothese wird nicht durch die Ergebnisse gestützt (Lam *et al.* , 2009, Schumann *et al.* , 2010). Das heißt, dass es noch keine direkte Beweise gibt, dass WOM bzw. OCRs sich bislang zwischen maskulinen und femininen Kulturen unterscheiden lassen.

2.3.5 Unsicherheitsvermeidung

“Unsicherheitsvermeidung lässt sich daher definieren als der *Grad, in dem die Mitglieder einer Kultur sich durch ungewisse oder unbekannte Situationen bedroht fühlen.*” Unsicherheitsvermeidung ist ein Gefühl, das unter anderem in nervösem Stress bringt, und dieses Gefühl fordert geschriebenen oder ungeschriebenen Regeln. (Hofstede, 1991, p. 133)

Kulturen mit starker Unsicherheitsvermeidung werden durch die Notwendigkeit gekennzeichnet, um Unklarheiten und Risiken zu reduzieren (Kale & Barnes, 1992). Verglichen mit Menschen in Kulturen mit niedriger Unsicherheitsvermeidung, nehmen die Mitglieder der hohen Unsicherheitsvermeidung, das Leben mehr als eine Bedrohung wahr, und erleben höhere Ebenen der Angst. Um diese Angst zu senken, werden sie stärker motiviert, um die wahrgenommene Zweideutigkeit und Unsicherheit des Lebens zu reduzieren (Doney *et al.* , 1998). Ein Weg um die Unklarheiten und Unsicherheiten zu reduzieren, ist die Beratung oder Zusicherung von vertrauenswürdigen anderen zu suchen. Im Einklang mit dieser Vorstellung wird eine hohe Unsicherheitsvermeidung mit einem höheren Niveau der Meinungs Austausch sowie Meinungssucht verbunden sein (Dawar *et al.* , 1996, Lam *et al.* , 2009, Liu *et al.* , 2001, Money *et al.* , 1998). Die Menschen in Kulturen mit hoher Unsicherungsvermeidung werden es auch versuchen, sich so in bereits vorhandenen geschäftlichen Beziehungen zu verhalten, weil sie ihre Meinung über die Verkäufer versichern wollen (Schumann *et al.* , 2010).

Personen in Kulturen mit niedriger Unsicherungsvermeidung haben größeren Glauben, dass sie ihr eigenes Leben und die Welt im Allgemeinen beeinflussen können (Hofstede, 2001). Daher sind sie für das Informationserfassungsverhalten weniger engagiert und sollten weniger anfällig gegen äußere Einflüsse auf ihr Konsumverhalten und Erkenntnisse sein (Dawar *et al.* , 1996, Money *et al.* , 1998).

Schumann *et al.* (2010) haben auch herausgefunden, dass die Auswirkung der WOM größer für die Kunden in Kulturen mit hoher Unsicherheitsvermeidung als die mit niedriger Unsicherheitsvermeidung ist. Liu *et al.* (2001) haben herausgefunden, dass mit schlechtem Service, Personen aus geringerer Unsicherheitsvermeidungskulturen häufiger sagten, dass sie wechseln wollen, negative Kommentaren geben, oder sich darüber beschwerten, als die Personen in Kulturen mit höherer Unsicherheitsvermeidung. Donthu & Yoo (1998)

zeigen, dass die Menschen in Kulturen mit hoher Unsicherheitsvermeidung im Allgemeinen positivere Kommntaren schreiben oder geben, als die mit niedriger Unsicherheitsvermeidung.

2.3.6 Langfristige Orientierung gegenüber kurzfristige Orientierung

Langfristige Orientierung gegenüber kurzfristiger Orientierung, als eine Dimension, machte man relativ neu von Unterschieden zwischen nationalen Kulturen aus (Hofstede, 1991, p. 29). Diese Dimension ist auch als die Konfuzianische Dynamik genannt (Hofstede, 1991, Lam *et al.* , 2009, Schumann *et al.* , 2010). Man kann folgende Werte am Pol, der 'langfristige Orientierung' bezeichnen: (Hofstede, 1991, p. 190)

- Ausdauer (Beharrlichkeit)
- Ordnung der Beziehungen nach dem Status
- Einhaltung dieser Ordnung
- Sparsamkeit
- Schamgefühl

Am entgegengesetzten Pol, der 'kurzfristigen Orientierung':

- Persönliche Standhaftigkeit und Festigkeit
- Wahrung des 'Gesichts'
- Respekt vor der Tradition
- Erwidern von Gruß, Gefälligkeiten und Geschenken

Die Wert des einen Pols sind eher auf die Zukunft hin ausgerichtet (insbesondere Beharrlichkeit und Sparsamkeit), und eher dynamisch. Dagegen sind mehr auf die Vergangenheit und Gegenwart hin ausgerichtet, und eher statisch. (Hofstede, 1991, p. 190)

Yoon (2009) hat herausgefunden, dass der Interaktionseffekt des Vertrauen \times langfristige Orientierung erhebliche Auswirkungen auf die Verwendungsabsicht hat. Aber die langfristige Orientierung hat keine direkte Wirkung auf die Verwendungsabsicht. Deshalb glaubte Yoon, dass langfristige Orientierung ein reiner Moderator des Vertrauens und der Verwendungsabsicht war. Das heißt: Je höher der Grad der langfristigen Orientierung ist, desto stärker ist die Wirkung des Vertrauens auf die Verwendungsabsicht. (Yoon, 2009)

In langfristig orientierten Gesellschaften suchen die Menschen die Tugend, während kurzfristig orientierte Gesellschaften daran glauben, die absolute Wahrheit festzustellen (Sohaib & Kang, 2014). Im Vergleich zu langfristig orientierten Konsumenten, erfahren die kurzfristig orientierten Konsumenten weniger wahrgenommene Kontrolle in unbefriedigenden Service-Begegnungen, und Entschädigung hätten einen stärkeren Effekt auf ihnen (Hui & Au, 2001,

Poon *et al.* , 2004).

2.3.7 Nachsicht gegenüber Beherrschtheit

Die sechste Dimension, Nachsicht gegenüber Beherrschtheit, wird auf Englisch “Indulgence versus Restraint” genannt. Diese Dimension ist ganz neu und wurde in dem Jahr 2010 in das Modell von Hofstede angelegt (Hofstede *et al.* , 2010, p. 285).

Die Definition nach Hofstede *et al.* (2010) ist folgende: “*Indulgence (Nachsicht) stands for a tendency to allow relatively free gratification of basic and natural human desires related to enjoying life and having fun. Restraint (Beherrschtheit) reflects a conviction that such gratification needs to be curbed and regulated by strict social norms.*”

Eine Umfrage in 2002-03 wurde durch Pew Research Center darüber gemacht, ob die ausländischen Filmen und Musik gut sind. Die Befragten, welche “sehr gut” geschrieben haben, sind positiv abhängig von der Nachsicht. Es wurde auch herausgefunden, dass nachsichtigere Gesellschaften mehrere Leute haben, die die Importierung von Entertainment bevorzugen. (Hofstede *et al.* , 2010, p. 292)

Kuppens *et al.* (2006) haben herausgefunden, dass Menschen, die in einer nachsichtigen Kultur sind, sich gerne an positiven Gefühlen erinnern. Schimmack *et al.* (2002) haben einen ähnlichen Ergebnis gefunden. In der Forschung berichten die Studenten mehr, die in nachsichtigerer Kultur sind, wenn sie positive Erfahrung haben. Durch den Vergleich der hoch und niedrig nachsichtigen Teiprobe, finden Zhou *et al.* (2015), dass Nachsicht die Wirkung des utilitaristischen Wert schwächt, und die Wirkung von hedonischen Wert auf affektivem Engagement stärkt.

Es gibt schon viele Forschungen im Bereich Online-Kaufen oder Kaufverhalten, die oft auf den Spontankauf fokussieren (Sharma & Sivakumaran, 2005). Im Bereich eWOM oder OCRs, gibt es noch wenige Ergebnisse.

Mangold *et al.* (1999) geben einen Einblick, um die Gründe, warum Menschen nachsichtig in eWOM-Kommunikationen sind, zu verstehen. Aber dieses Ergebnis betrifft keine Dimension von Hofstede.

2.3.8 Das sechsdimensionale Modell für Deutschland und China

Basierend auf diesem sechsdimensionalen Modell, kann man erkennen, dass Deutschland und China viele Unterschiede in der Kultur haben. Nach der Forschung von Cyr (2014), Hofstede (2001), Singh *et al.* (2006), unterschieden sich die Kulturen in Deutschland und

China deutlich. Abbildung 2.9 zeigt die Ungleichheiten zwischen Deutschland und China im sechsdimensionalen Modell.

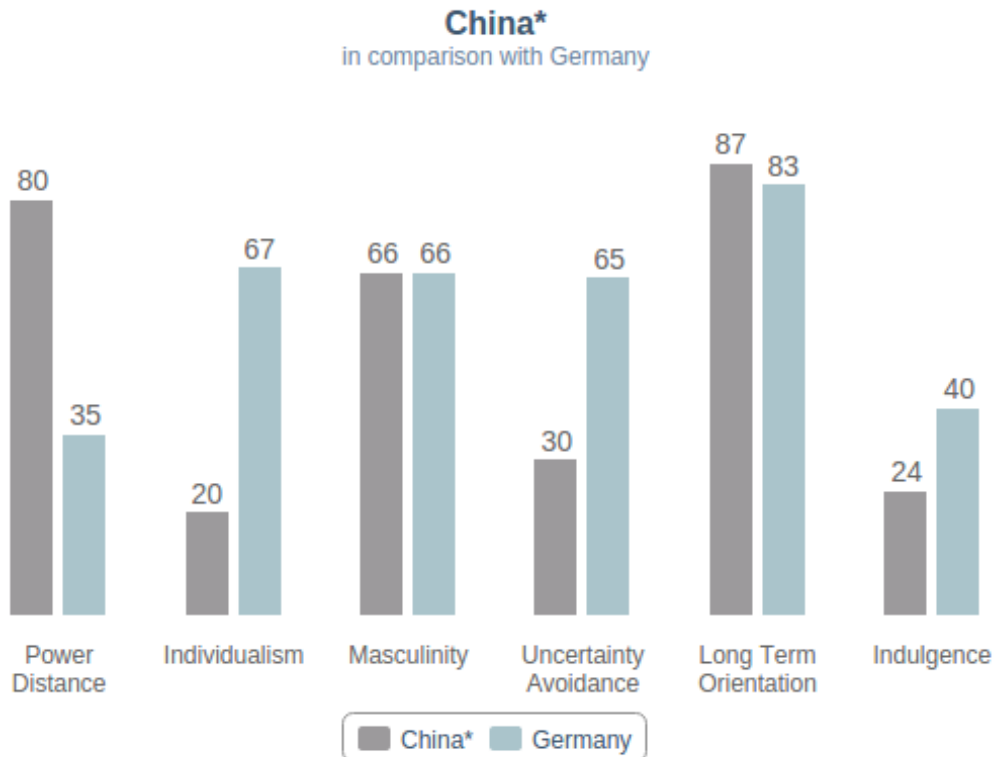


Abbildung 2.9: Das sechsdimensionale Modell für Deutschland und China (Quelle: Hofstede, cited November 2015)

Verglichen mit deutschen Menschen, haben die Chinesen größere Machtdistanz. Das heißt: die Ungleichheit ist überall, erwartet und erwünscht, nicht nur zwischen den Konsumenten und den Verkäufer im Internet, sogar auch in den Konsumenten untereinander. Die deutsche Kultur ist individualistisch aber die chinesische ist typisch kollektivistisch. Es wird vorgeschlagen, dass die deutschen Konsumenten sich unabhängiger von anderen verhalten, als die Chinesen. Die chinesischen Konsumenten verhalten sich mehr zentralisiert. In Dimensionen “Maskulinität gegenüber Femininität”, “Langfristige Orientierung gegenüber kurzfristige Orientierung” und “Nachsicht gegenüber Beherrschtheit” sind Deutschland und China auf gleicher Ebene. Aber bei der Unsicherheitsvermeidung ist ein großer Spalt. Nach Donthu & Yoo (1998), wird es vorgeschlagen, dass die Deutschen positivere Kommentaren schreiben wollen, als die Chinesen.

2.4 Relevante Forschungsergebnisse zum Textileinzelhandel

Im Bereich Textileinzelhandel oder Online Consumer Reviews gibt es schon viele Forschungen, aber die kulturübergreifenden Forschungen in diesem Bereich sind wenig (nur 8). Die meisten vergleichen die Kulturen zwischen Osten (meistens China) und Western. Fünf Forschungen haben das Modell von Hofstede (2001) genutzt. Tabelle 2.5 zeigt die relevanten kulturübergreifenden Forschungen im Bereich Online Consumer Reviews.

Von dieser Tabelle kann man sehen, dass es Unterschiede gibt, nicht nur zwischen Osten und Westen, sondern auch zwischen unterschiedlichen östlichen Ländern (India und China). Deshalb wird es vorgeschlagen: Aufgrund der Einflüssen von Kulturen, verhält man sich unterschiedlich, und man hat verschiedene Meinungen, wenn die Dimensionen der eigenen Kultur zueinander verschieden sind. Diese Unterschiede kann man auch in den OCRs finden.

Im Bereich Textileinzelhandel gibt es schon einige Ergebnisse über die OCRs, aber die meisten Ergebnisse sind über die chinesischen Verbrauchern. Rahman *et al.* (2010) haben gefunden, dass die chinesischen Befragten mehr befasst mit funktioneller Eigenschaft waren, während die Kanadier sich mehr auf ästhetische Attribute konzentrierten. Jin *et al.* (2010) erklärten, dass die Verarbeitung und Passform am wichtigsten für die US-Verbraucher sind, aber nicht für die chinesischen Verbraucher. Hsu & Burns (2002) machten einen Vergleich zwischen den Studenten in Taiwan und den USA, und stellten fest, dass die Flexibilität und Angemessenheit die wichtigen Bewertungskriterien von Kleidungen für die Studenten in Taiwan waren, nicht in den USA, aber Komfort und Passform waren für die Verbraucher aus beiden Ländern am wichtigsten. Zhang *et al.* (2002) haben entdeckt, dass Passform und Komfort die wichtigsten Faktoren für die chinesischen Verbraucher mit Stil, Farbe und Verarbeitung.

Viele Studien fanden auch, dass die äußeren Attribute wie Marken relativ mehr wichtig für die chinesischen Verbraucher waren. Zum Beispiel: Chun-Tung Lowe & Corkindale (1998) schlug vor, dass die chinesischen Verbraucher durch Markenloyalität gekennzeichnet waren und Zhou & Hui (2003) fanden auch, dass symbolische Attribute wichtiger als Performance zur Kaufabsicht der chinesischen Konsumenten waren. Die "Kollektivismus" Kultur macht die chinesischen Verbraucher mehr Sorgen um andere Meinungen und sozialen Selbstbewusstsein (Rahman *et al.* , 2010).

2.5 Hypothesen

Wie es in dem Abschnitt 2.2.3 beschrieben wird, wird jede Review von den Kunden als ein Dokument durch die Sentiment Analyse analysiert. Das heißt: jede Review hat eine eigene Polarität. Ähnlich wie die quantitativen OCRs, hat die qualitative Online Review auch deswegen folgende Attribute:

Volumen: Dieses Attribut für die quantitativen OCRs ist die Anzahl der Kommentare, die das Produkt hat (Sridhar & Srinivasan, 2012). Aber hier in dieser Arbeit, beschreibt das Volumen die Anzahl der Zeichen, die der Kunde in der qualitativen Online Review geschrieben hat.

Valenz: Dieses Attribut für die quantitativen OCRs ist die Anzahl Sterne, die der Kunde gegeben hat (Sridhar & Srinivasan, 2012). Aber in dieser Arbeit, beschreibt die Valenz die Polarität, welche die qualitative Online Review durch die Sentiment Analyse hat.

Deshalb hat eine Online Review (inklusive quantitativer und qualitativer Seite) die Attribute: Valenz der Sterne, Valenz der Polarität durch Sentiment Analyse, und Volumen von Zeichen. Für ein Produkt, gibt es viele Reviews, deswegen haben die qualitativen OCRs für ein Produkt oder mehrere Produkte noch folgende Attribute:

Varianz: Dieses Attribut beschreibt den Grad der Polaritätsverschiedenheiten von Kunden durch die Sentiment Analyse.

Korrelationskoeffizient: Dieser zeigt den Grad des Zusammenhangs zwischen quantitativen und qualitativen OCRs.

Weil die deutschen und chinesischen OCRs gleichzeitig quantitativen und qualitativen Inhalte haben, besitzen die Attribute Valenz, Volumen, und Varianz für die quantitativen und qualitativen Aspekte unterschiedliche Bedeutungen. Tabelle 2.6 zeigt dies Unterschiede.

Diese Attribute beschreiben die OCRs mathematisch und statistisch. Basierend auf diesen mathematischen und statistischen Maßen kann man einfach einige statistischen Analysen durchführen.

Aus der Abbildung 2.9 kann man deutlich sehen, dass Deutschland und China sich in Bezug auf Machtdistanz, Individualismus gegenüber Kollektivismus und Unsicherheitsvermeidung unterscheiden. Für die anderen drei Dimensionen sind Deutschland und China in gleicher Ebene, deshalb wird vorgeschlagen, dass die Dimensionen "Maskulinität gegenüber Femininität", "langfristige Orientierung gegenüber kurzfristiger Orientierung" und "Nachsicht gegenüber Beherrschtheit" wenige Einflüsse haben, um die deutschen und chinesischen OCRs zu unterscheiden.

Wenn es keine kulturellen Einflüsse auf das Schreiben der OCRs geben würde, wäre jede Erzeugung eines Online Reviews individuell und hätte keine Abhängigkeit mit anderen. Das

heißt: Das Gefühl von jedem Kunden für das Produkt ist unabhängig voneinander, und die Polaritäten durch Sentiment Analyse sollte die Normalverteilung angepasst werden. Deshalb wird es als die Nullhypothese H_0 vorgeschlagen:

Hypothese H_0 *Es gibt keine kulturelle Einflüsse auf die Online Consumer Reviews an sowohl der quantitativen Seite als auch der qualitativen Seite.*

Wie in dem Abschnitt 2.3.2 beschrieben, sind die Menschen in den Kulturen mit großer Machtdistanz zurückhaltend wenn es um den Informationsaustausch nach Lam *et al.* (2009) geht. Nach Dawar *et al.* (1996), Lam *et al.* (2009), Liu *et al.* (2001), Money *et al.* (1998) wird eine hohe Unsicherheitsvermeidung mit einem höheren Niveau des Meinungsaustausches sowie Meinungssucht verbunden sein. Das heißt:

Hypothese 1 *Das Volumen von chinesischen quantitativen und qualitativen Online Consumer Reviews ist kleiner als das von Deutschen.*

Nach Liu *et al.* (2001) und Donthu & Yoo (1998), geben die Menschen in Kulturen mit hoher Unsicherheitsvermeidung im Allgemeinen positivere Kommentare, als die mit niedriger Unsicherheitsvermeidung. Deshalb wird vorgeschlagen:

Hypothese 2 *Die deutschen qualitativen und quantitativen Online Consumer Reviews haben größere Valenzen als die chinesischen.*

Ungleichheit unter den Menschen in Kulturen mit großer Machtdistanz wird erwartet und ist erwünscht nach Hofstede (1991), und hierarchische Strukturen in Organisationen sind ein Spiegelbild einer Ungleichheit von Natur aus zwischen oberen und unteren Schichten. Diese sagen, dass die Varianz von Polaritäten durch Sentiment Analyse größer sein wird, wenn Menschen in Kulturen mit großer Machtdistanz leben.

Aber beim Kollektivismus wirken sich die Informationskonsistenz und die Informationsbewertung auf ihre Wahrnehmung von Informationsglaubwürdigkeit stärker aus (Luo *et al.*, 2014). Gemäß der Theorie der Hofstede (1998), machen die Menschen in den kollektivistischen Kulturen die Kaufentscheidungen im Konsens mit der Gruppe und kaum rein individuelle Entscheidungen im Gegensatz zu den Menschen in den individuellen Kulturen (Singelis, 1994). Auf Grund der größeren Machtdistanz, gibt es die zentralisierte Tendenz, aber auf der gegenüberliegenden Seite gibt es keine solche Tendenz (Hofstede, 1991). China hat eine sehr hohe Machtdistanz mit kollektivistischer Kultur, aber für Deutschland ist die Machtdistanz niedrig und die Menschen sind individuell. Es gibt unterschiedliche Einflüsse auf die chinesischen und deutschen OCRs. Das heißt: die OCRs werden konsistenter beim Kollektivismus sein, deswegen sollte die Varianz von Polaritäten kleiner als beim Individualismus sein.

Weil die vorherigen Forschungen vorgeschlagen haben, dass die Dimension “Individualismus gegenüber Kollektivismus” eine Tiefenstruktur und wichtigste Dimension ist (Greenfield, 2000, Sia *et al.* , 2009, Triandis, 2001), wird es vermutet, dass diese Dimension einen größeren Einfluss auf die OCRs hat, deshalb:

Hypothese 3 *Die chinesischen Valenzen von den qualitativen und quantitativen Online Consumer Reviews haben kleinere Varianzen als die Deutschen.*

Und Menschen in der kollektivistischer Kultur sind einfacher zu beeinflussen, eine quantitative Review zu geben, aber das Gefühl ist von jedem Kunden unterschiedlich und nicht einfach übereinzustimmen. Basierend auf die Hypothesen 2 und 3, wird die deutschen OCRs positiver und individueller als die chinesischen, aber der Zusammenhang zwischen quantitativen und qualitativen OCRs ist noch nicht in Sicht.

Resnick & Zeckhauser haben durch die Studien von Online Rückkopplungsmechanismen festgestellt, dass die Feedbacks (also die OCRs in eBay) eine hohe Korrelation zwischen den Käufern und Verkäufern haben. Und Stenberg & Sydow haben schon gefunden, dass die Attitüde und Absicht von Menschen in China und Schweden unterschiedlich sind. Die Aspekten, auf der man sich konzentriert, sind in den USA und China unterschiedlich (Lu, 2015). Das heißt: Die Kunden geben gleichen quantitativen OCRs aber haben vielleicht unterschiedliche Aspekte in den verschiedenen Kulturen. Anders gesagt, die kulturellen Einflüsse auf quantitative und qualitative OCRs sind unterschiedlich. Um diese Unterschiedlichkeit zu verdeutlichen, wird es vorgeschlagen:

Hypothese 4 *Der Zusammenhang zwischen den deutschen quantitativen und qualitativen Online Consumer Reviews ist stärker als der Zusammenhang von den chinesischen Online Consumer Reviews.*

Quelle	Verfahren von Kultur	Länder/Bereich	Ergebnisse
Adams (2005)	Dimension: Individualismus gegenüber Kollektivismus genutzt	Großbritannien, China/eBusiness	Unterschiede in der Benutzer-Praktiken und Umgebung weisen hin, dass die E-Commerce-Geschäftsmodellen für den Westen nicht ganz geeignet für den Osten sein können.
Chu & Choi (2011)	Dimension: Individualismus gegenüber Kollektivismus genutzt	USA, China/Soziale Netzwerke	Nationale Kultur spielt einen wesentlichen Faktor, der Engagement der Verbraucher in eWOM in sozialen Netzwerken, die beide Länder betrifft.
Fong & Burton (2008)	Dimension: Individualismus gegenüber Kollektivismus genutzt	USA, China/Ursprungsland-Effekte	China-basierte Diskussionsforen engagieren in höheren Ebenen der Informationssuche und niedrigere Informationshingabe als in den USA.
Jin <i>et al.</i> (2010)	keine Dimension von Hofstede genutzt	India, China/Kleidungen	Chinesische und indische Verbraucher bewerten Attribute von Kleidungen unterschiedlich.
Lu (2015)	keine Dimension von Hofstede genutzt	USA, China/Kleidungen	Chinesische und amerikanische Verbraucher bewerten Attribute von Kleidungen unterschiedlich.
Singh <i>et al.</i> (2006)	Alle Dimensionen von Hofstede genutzt	Deutschland, India, China/Wahrnehmung der Webseite-Anpassung von Verbrauchern	Kultur spielt eine große Rolle, um die Verbraucher der Webseite anzupassen. Und Kultur beeinflusst Glauben, Meinungen und Kaufabsicht der Verbraucher im Internet.
Sohaib & Kang (2014)	4 Dimensionen (außer langfristiger Orientierung und Nachsicht) von Hofstede getestet.	Pakistan und Australien/B2C-Webseite	B2C-Webseite spiegeln das kulturelle Umfeld, das die Online-Käufer umgibt. Es scheint dass die Kultur die Kaufabsicht von Online-Käufer beeinflusst.
Stenberg & Sydow (2014)	Unsicherheitsvermeidung und Individualismus werden verwendet	Schweden und China/Kleidung	Die Hauptunterschiede waren die wahrgenommene allgemeine Nützlichkeit von Online-Shopping und der Risikograd, den Verbraucher beim Online-Shopping wahrgenommen haben.

Tabelle 2.5: Die relevanten kulturübergreifenden Forschungen

	quantitative OCRs	qualitative OCRs
Volumen	Anzahl der OCRs	Anzahl der Kennzeichen
Valenz	Zahl der Sterne	Zahl der Polarität
Varianz	Grad der Meinungsverschiedenheiten	Grad der Polaritätenverschiedenheiten
Korrelationskoeffizient	der Grad des Zusammenhangs zwischen quantitativen und qualitativen OCRs	

Tabelle 2.6: Die Attribute der quantitativen und qualitativen OCRs (Quelle: eigene Darstellung)

3 Datensammlung und Textverarbeitung

Die Kommentaren auf deutschen und chinesischen Webseiten für die gleichen Produkte von sportlichen Kleidungen werden in dieser Arbeit gesammelt. Die chinesischen Kommentaren werden durch eine Übersetzungsmaschine (Google Translate) ins Deutsche übersetzt. Die Einflüsse von den Übersetzungen auf die Ergebnisse, besonders wegen der Ungenauigkeit von Übersetzungsmaschine, werden im Abschnitt 3.6 diskutiert.

3.1 Datenquellen

Die deutsche Daten werden aus Amazon.de gesammelt, und die chinesische aus Tmall.com. Die beiden sind die wichtigsten B2C Plattformen in den jeweiligen Ländern.

3.1.1 Amazon weltweit und in Deutschland

Amazon ist präsent auf den derzeit größten E-Commerce-Märkten weltweit und bietet einen schnellen und einfachen Zugang zu diesen. Amazon.de, die deutschsprachige Website, wird von der Amazon.com, Inc. betrieben. Die Amazon.com, Inc. ist ein amerikanischer Online-Versandhändler mit einer breit gefächerten Produktpalette. Vom Zeitpunkt des Debüts von Amazon.com als Webseite im Juli 1995 bis Anfang 1999 hat Amazon.com Inc. eine Marktkapitalisierung von \$6 Milliarden erreicht, das ist mehr als Barnes&Noble und Borders, die größten Rivalen on-und offline, zusammen. Im Letzten Quartal von 1998 beliefen sich die Internetverkäufe auf \$252,9 Millionen, das ist ein 283%iger Zuwachs über den Nettoumsatz von \$66 Millionen im letzten Quartal von 1997.(Saunders, 2001, p. 15)



Abbildung 3.1: Logo von Amazon Deutschland(Quelle:Amazon.de)

Nach dem Geschäftsbericht im Jahr 2014 von der amerikanischen Börsenaufsicht (Securities

and Exchange Commission in USA (SEC)), hat die Amazon.com, Inc. einen Nettoveräußerungserlös von \$88.988 Millionen(Bezos, 2014). Abbildung 3.2 zeigt das Wachstum der Umsatzerlöse von Amazon.com, Inc. von 2010 bis 2014.

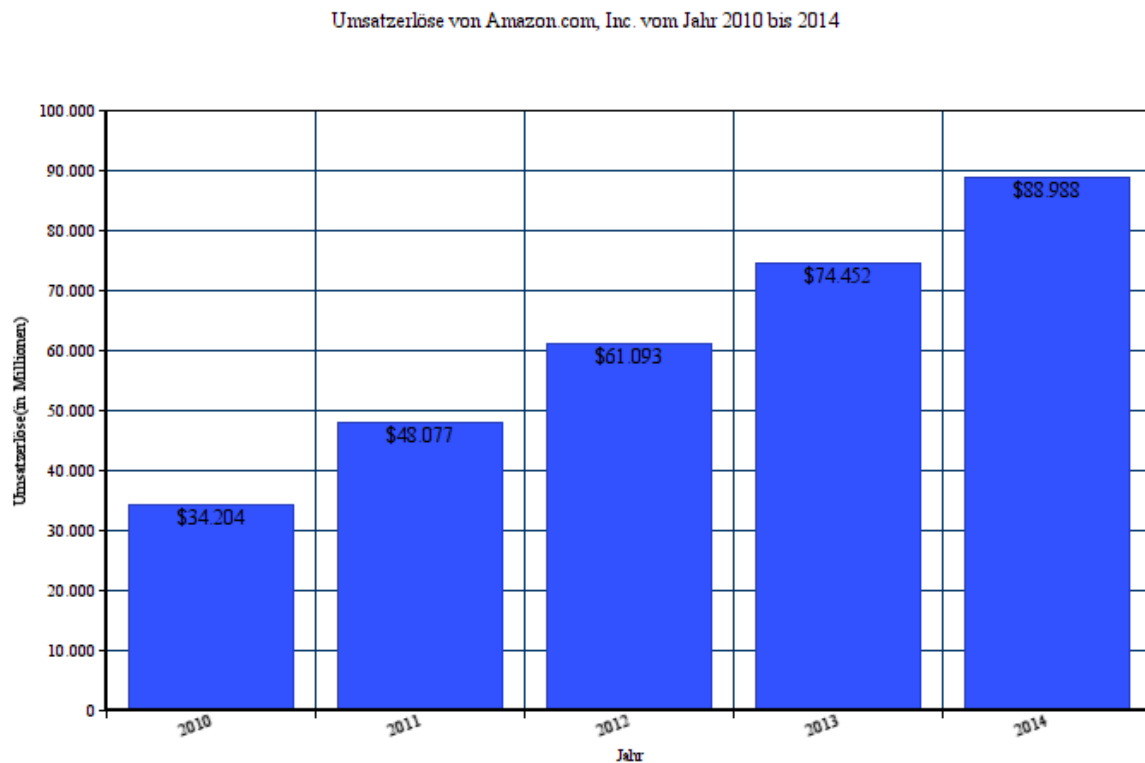


Abbildung 3.2: Umsatzerlöse (US Dollar in Millionen) von Amazon.com, Inc. 2010 - 2014 (Datenquelle: Bezos, 2013, 2014)

Für Amazon.com, Inc. ist Deutschland der umsatzstärkste Auslandsmarkt(Roland Lindner, 2013). Deutschland stand im Jahr 2014 für 13 Prozent des Gesamtumsatzes von Amazon, und im Jahr 2013 für 14 Prozent(Bezos, 2014). Der Bundesverband des Deutschen Versandhandels hat den Gesamtumsatz des deutschen Online-Handels im vergangenen Jahr auf 27,5 Milliarden Euro beziffert. Zahlen des Einzelhandelsverbands sprechen von einem Umsatz von 29,5 Milliarden Euro. Demnach kontrolliert Amazon ein gutes Fünftel oder sogar fast ein Viertel des gesamten deutschen Online-Versandhandels.(Roland Lindner, 2013) Abbildung 3.3 zeigt das starke Wachstum des Erlöses vom Deutschland-Geschäft ab 2010 bis 2014.

Mehrere internationale Textilhersteller sind über Amazon Deutschland aufgelistet, bzw. Adidas, Nike, Puma und so weiter. Deshalb ist Amazon Deutschland ein gutes Objekt im Bereich elektronischer Textileinzelhandel zu studieren.

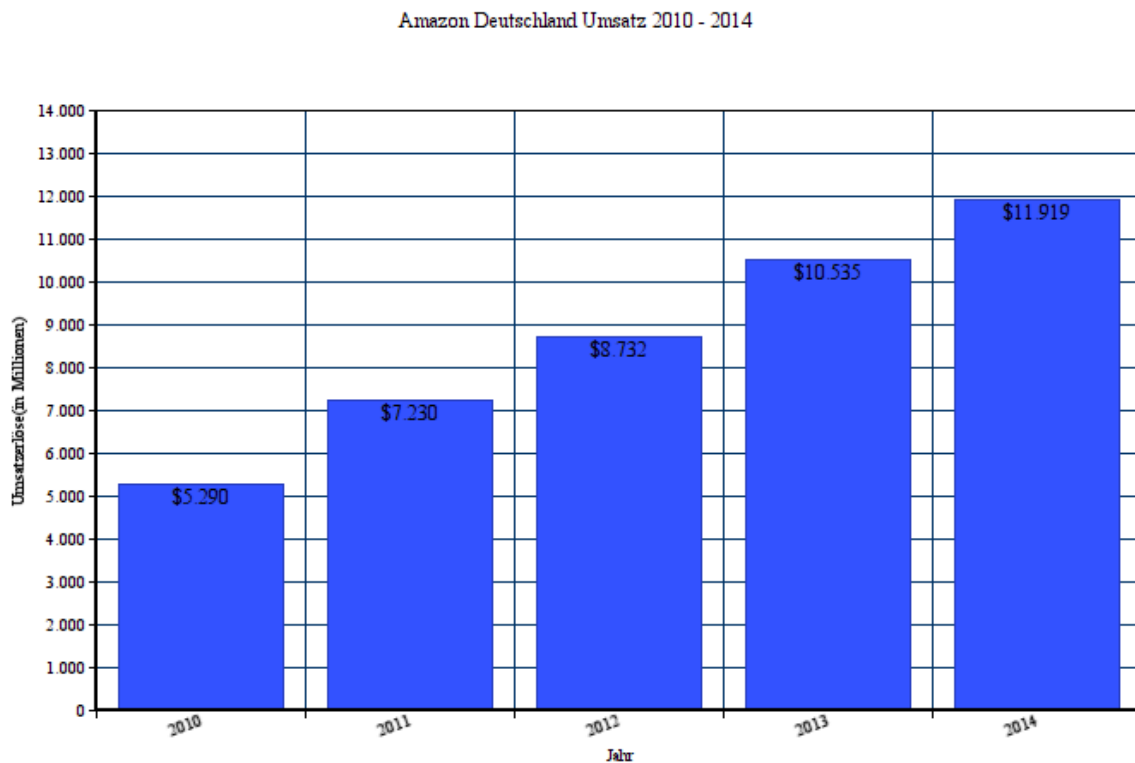


Abbildung 3.3: Amazon Deutschland Umsatz 2010 - 2014 (US Dollar in Millionen) (Datenquelle: Bezos, 2013, 2014)

3.1.2 Chinesische Plattform: "Tmall"

Tmall.com ist eine chinesischsprachige Webseite für B2C Online-Einzelhandel, die sich im Jahr 2008 von Taobao.com aufgespalten hatte. Es ist das größte Plattform für lokale chinesische und internationale Unternehmen, um Markenware für die Verbraucher in der Volks Republik China, Hongkong, Macao und Taiwan zu verkaufen. Im Jahr 2013 wurde



Abbildung 3.4: Logo von Tmall.com (Quelle: Tmall.com)

Tmall das größte Open-Portal für Online Händler und Endverbraucher, mit einem Anteil am chinesischen Online-Markt ca. 50%. (CECRC, 2014) Tabelle 3.1 zeigt fast alle Online Händler in der chinesischen Online-Marktposition.

Online Händler	Marktanteil(%)
Tmall	50,1
JD.com	22,4
Suning.com	4,9
Tecent B2C	3,1
Amazon China	2,7
Yihaodian	2,6
Vipshop	2,3
Dangdang	1,4
Gome	0,4

Tabelle 3.1: Marktpositionen von Online Händler in China.(Quelle: CECRC, 2014)

Mit mehr als 10 Millionen jährlich aktiven Anbieter und über 140.000 Marken in Tmall bis zum 31. März 2015 bieten die Marktplätzen von Tmall den Verbraucher wettbewerbsfähige Preisgestaltung in einer breit angelegten Skala von Kategorien. Wegen der Anwesenheit einer großen Anzahl von globalen Marken und den hohen Anforderungen an die auf Tmall betriebenen Händlern, hat sich eine Präsenz auf Tmall zu einer Validierung der Qualität, um die Händler die Markenbekanntheit zu bauen und erweitern. Gleichzeitig ist Tmall eine vertrauenswürdige Plattform für die chinesischsprachigen Verbraucher, um sowohl einheimische als auch internationale Markenprodukte zu kaufen, selbst wenn Produkte nicht verfügbar in den traditionellen Verkaufsstellen sind.(Ma, 2014)

Tmall ist zum Synonym für Online-Shopping in China geworden. Die Verbraucher kommen auf die Plattform mit starken kommerziellem Absicht, die zu hohen Umsätze für die Händlern führt. Abbildung 3.5 zeigt das Wachstum der Gross Merchandise Volume (GMV) von Tmall ab 30. Juni 2013 bis 31. März 2015.(Ma, 2014) Wegen dem großen Erfolg in China ist Tmall als eine gute Plattform für den chinesischen Online-Handel zu studieren.

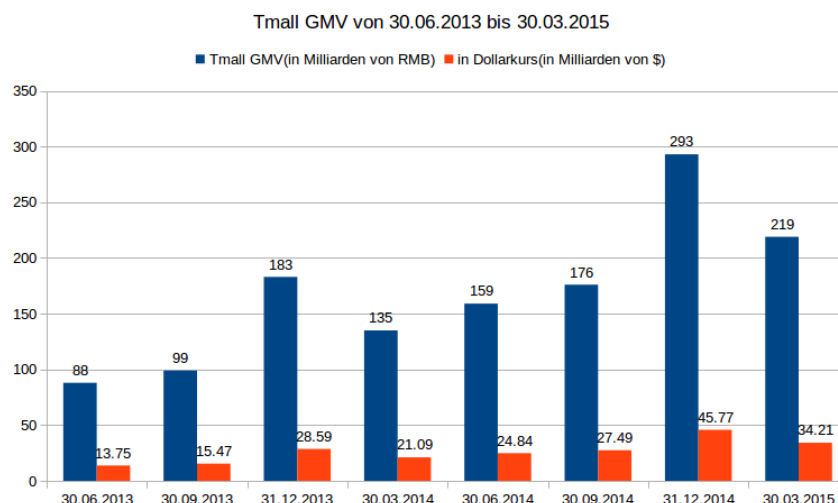


Abbildung 3.5: Tmall GMV von 30.06.2013 bis 30.03.2015 (Datenquelle:Ma, 2014)

3.2 Sportkleidung als Untersuchungsobjekt

Diese Arbeit macht einen Vergleich über Online-Kommentaren zwischen Deutschland und China im Bereich Textileinzelhandel. Um dieses Ziel zu erreichen, sollen die Studienobjekte folgende Bedingungen erfüllen:

1. Die Studienobjekte sollten Kleidung oder Textilien sein.
2. Die Studienobjekte sollten in Deutschland und China im Internet verkauft werden.
3. Die Studienobjekte sollten auf den gewählten Plattformen (Amazon.de und Tmall) viele Kommentaren haben.
4. Jedes ausgewählte Studienobjekt sollt in den beiden Ländern gleich sein.
5. Es sollte keine Fashion-Elemente geben.

Bedingung 1 beschränkt die Arbeit im textilen Bereich, und die Bedingung 2 und 3 werden es einem möglich machen die Online-Kommentaren zu analysieren. Mit der Erfüllung der vierten Bedingung macht der Vergleich akademischen und geschäftlichen Sinn, weil es große Änderungen in einem Zeitraum über Fashion gibt, und dazu Fashion ein ganz anderes Thema ist, wird diese Arbeit die Fashion-Elemente von den Kleidungen rausnehmen. Es gibt noch eine potenzielle Bedingung: Alle Studienobjekte sollten in den beiden Ländern sehr gut verkauft werden, damit sie viele Kommentaren haben.

Aufgrund der obigen Bedingungen werden drei internationalen Textilmarken (Adidas, Nike und Puma) aus unterschiedlichen Marken in dieser Arbeit ausgewählt. Alle drei Marken sind sportliche Artikel, um die Bedingung 5 zu erfüllen.

- **Nike, Inc.** ist ein amerikanischer multinationaler Konzern, die in der Konstruktion, Entwicklung, Produktion und weltweite Vermarktung und Verkauf von Schuhen, Bekleidung, Ausrüstung, Zubehör und Dienstleistungen tätig ist. Nike ist einer der weltweit größten Anbieter von Sportkleidung und Schuhe, und ein bedeutender Hersteller von Sportartikeln, mit einem Umsatz von mehr als 24,1 Milliarden US-Dollar im Geschäftsjahr 2012. Ab 2012 beschäftigt es mehr als 44.000 Mitarbeiter weltweit. Im Jahr 2014 wurde die Marke allein zu 19 Milliarden Dollar geschätzt und ist damit die wertvolle Marke unter den Sport Unternehmen. (Mahdi *et al.* , 2015)
- **Adidas AG** mit Sitz in Herzogenaurach ist ein deutscher Sportartikelhersteller mit den Marken Adidas, Reebok und TaylorMade. Seit dem 17. November 1995 ist der Konzern im Deutschen Aktienindex an der Frankfurter Wertpapierbörse notiert und gilt nach Nike als der zweitgrößte Sportartikelhersteller der Welt (Langenscheidt & Bauer, 2010). Es ist ein Haushalts Marke mit den drei Streifen Logo auf den Märkten in der ganzen Welt anerkannt. Das Produktportfolio des Unternehmens ist groß und reicht von Sportkleidung und Schuhe, und Accessoires wie Taschen, Uhren, Brillen und

andere sportbezogene Produkte und Ausrüstung. Adidas beschäftigt mehr als 46.000 Menschen Weltweit. Die Adidas Gruppe besteht aus rund 170 Tochtergesellschaften einschließlich Reebok, Taylormade-Adidas Golf, Rockport und CCM-Hockey. Der Hauptsitz der Gruppe ist in Herzogenaurach, Deutschland. Im dritten Quartal 2014 betrug der Umsatz der Gruppe 4,118 Milliarden Dollar. (Mahdi *et al.* , 2015)

- **Puma SE** mit Sitz in Herzogenaurach ist ein börsennotierter Sportartikelanbieter und Hersteller von Schuhen, Textilien und Accessoires (Peters, 2007). Puma ist mit rund drei Milliarden Euro Jahresumsatz, einem Konzerngewinn von 5,3 Millionen Euro und 10.982 Mitarbeitern im Jahr 2013 neben Adidas und Nike einer der weltweit größten Sportartikelhersteller (About Puma, 2014).

Die drei sind die weltweit größten sportlichen Textilhersteller, und in Deutschland und China verkaufen sie schon lange und gut. Deshalb ist es einfach, die Kleidungen zu finden, die die 5 Bedingungen erfüllen können. In dieser Arbeit werden insgesamt 4 solche Kleidungen als Studienobjekte genutzt. 2 Kleidungen davon sind Hosen und die anderen 2 sind T-Shirt. 3 Kleidungen sind für Herrn und eine für Damen. Die beiden Hosen werden von Adidas hergestellt, eine weibliche und eine männliche. Die beiden männlichen T-Shirts sind jeweils von Nike und Puma.

3.3 Technische Werkzeuge für die Datensammlung

3.3.1 Datensammlung aus Webseiten mit Python

Python hat von Data-Miner Gemeinschaft für Datensammlung entweder im Internet oder anderswo als eine perfekte Sprache gewonnen (Khwaldeh *et al.* , 2013). Tools und Methoden von Python können nach der ausgeführten Funktion oder der Klasse der verwendeten Anwendung klassifiziert werden, weil es viele positive Eigenschaften besitzt: (Segaran, 2008, p. X)

Kürzer

Code, der in dynamisch typisierten Sprachen wie Python geschrieben wird, ist meist kürzer als Code, der in anderen verbreiteten Sprachen entsteht. Das bedeutet, dass einfacher ist, den Algorithmus in Kopf unterzubringen und wirklich zu verstehen, was er tut.

Leicht zu lesen

Python wurde schon als “ausführbarer Pseudocode” beschrieben. Das ist natürlich übertrieben, aber es stimmt sehr wohl, dass die meisten erfahrenen Programmierer Python-Code lesen und verstehen können, was er tun soll.

Einfach erweiterbar

Python bringt standardmäßig vielen Bibliotheken mit, unter anderem mit mathematischen Funktionen, zum Parsen von Extensible Markup Language (XML) und zum Laden von Webseiten. Diese Bibliotheken sind frei erhältlich und lassen sich einfach herunterladen, installieren und nutzen.

Interaktive

Wenn man ein Code durcharbeitet, ist es hilfreich, die Funktionen schon auszuprobieren, während man sie schreibt, ohne ein anderes Programm nur zum Testen erstellen zu müssen. Python kann Programme direkt von der Befehlszeile aus ausführen und besitzt zudem einen interaktiven Prompt, an dem man Funktionen startet, Objekte erstellen und Pakete testen könnte.

Viele Paradigmen

Python unterstützt objektorientierte, prozedurale und funktionale Programmierung. Manchmal ist es sinnvoll, Funktionen als Parameter zu übergeben, während man ein anderes Mal den Zustand in einem Objekt ablegt. Python unterstützt beide Vorgehensweisen.

Viele Plattformen und freie Software

Python bietet eine Referenzimplementierung für alle wichtigen Plattformen und ist auf allen frei erhältlich. Der Code, der in dieser Arbeit beschrieben wird, funktioniert unter Windows, Linux und Mac.

Darüber hinaus kommt Python mit Modulen für den Zugriff auf das Internet wie *urllib* und *urllib2*, *Pattern*, *mechanisieren Paket* und *Scrapy*. In dieser Arbeit wird der letzte verwendet, weil Nutzer weniger Zeit und Skript als die anderen benötigt. (Khwaldeh *et al.* , 2013)

3.3.2 XPath: Daten Lokalisierung in Webseiten

XPath 1.0 gilt als eine Untersprache innerhalb eines Extensible Stylesheet Language: Transformations (XSLT) 1.0-Stylesheets und ist beim Transformationsprozess unerlässlich. XPath ist also auch unabhängig von XSLT nutzbar. XPath selbst verwendet keine XML-Syntax, sondern benutzt Funktionen und Ausdrücke, um XML-Dokumente bzw. deren Teile zu adressieren. XPath-Ausdrücke benutzen sogenannte Lokalisierungspfade (Location path), um die gewünschten Knoten im Dokument zu finden. (Sebestyen, 2010, p. 211-212)

Man kann sich solche Lokalisierungspfade wie einen Wegweiser vorstellen: “*Sie sind jetzt hier und jetzt müssen Sie die Treppe nach unten nehmen, dort nach links und dann die dritte Tür.*” Der XPath-Ausdruck “*/buch/autor/nachname*” selektiert das *<nachname>* Element im *<autor>* Element, das sich unterhalb des *<buch>* Element befindet. (Sebesty-

en, 2010, p. 212)

Mit XPath kann man einfach die gewünschten Daten in Webseiten finden und lokalisieren. Und in Python und Scrapy gibt es noch viele Bibliotheken, die XPath gut unterstützen. Deshalb ist XPath in dieser Arbeit eine wichtige Maßnahme, um die Kommentaren in den Webseiten zu bekommen.

3.3.3 Reguläre Ausdrücke: Bearbeitung der Texte

Reguläre Ausdrücke sind ein mächtiges, flexibles und effizientes Mittel, um Texte zu bearbeiten. Reguläre Ausdrücke im engeren Sinne sind eine generelle Notation zur Beschreibung von Textmustern, beinahe wie eine kleine Programmiersprache zum Prüfen und zum Manipulieren von Texten. Mit den zusätzlichen Mitteln eines bestimmten Programmierwerkzeugs können reguläre Ausdrücke benutzt werden, um alle Arten von Texten zu erweitern, zu reduzieren und in jeder Art zu misshandeln. Es kann sich um einfache Dinge handeln wie um die Suchfunktion eines Texteditors oder um komplexe wie eine ganze Textverarbeitungssprache. (Friedl, 2009, p. 1)

In dieser Arbeit, werden reguläre Ausdrücke benutzt, um die Kommentaren zu formulieren. Nur nach der Formulierung kann sich die Bearbeitung von der Kommentartexten bei einer Textverarbeitungssprache durchführen lassen. Also Python bittet eine sehr gute Unterstützung auf regulären Ausdrücken unter Einsatz einer externen Bibliothek.

3.4 Datensammelungsprozess

Weil die Analyseprogramme die Eingabe nur in CSV-Formen kennen, ist das Ziel der Datensammlung, alle Daten von Kundenkommentaren für die ausgewählten Studienobjekten in zwei Plattformen (Amazon.de und Tmall.com) in die CSV-Dokumente automatisch zu sammeln und zu speichern. Um dieses Ziel zu erreichen, muss man an die Unterschiede zwischen Amazon.de und Tmall denken, und zwei Datensammelungsprogramme schreiben.

3.4.1 Sammeln der Daten aus Amazon.de

Die Daten von Kundenkommentaren sind in der Seite Kundenrezensionen jedes Studienobjektes. Abbildung 3.6 zeigt das Bild der Seite. Aus dem Bild kann man einfach die Struktur der Webseite verstehen. Sie ist ganz einfach: Es steht das Logo von Amazon Deutschland und die Suchleiste oben. Danach sind die Situation der Sternen, die von jedem Kunde gemacht wurde. Rechts steht der Namen und links die Ware. Darunter ist eine Tabelle, in der es

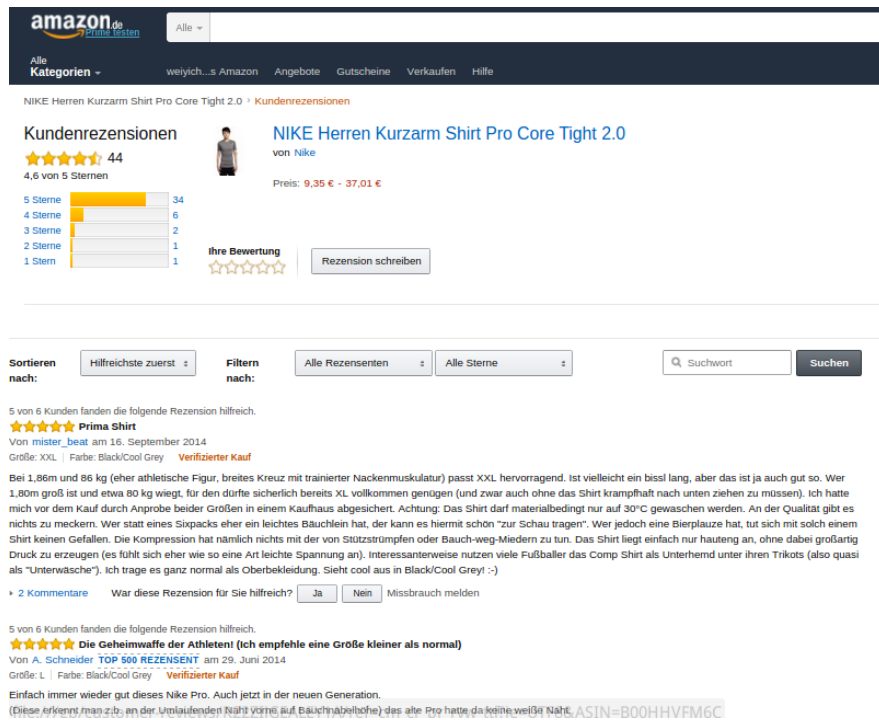


Abbildung 3.6: Kundenrezensionen von Amazon.de (Quelle: Amazon.de)

alle Kundenrezensionen gibt. Jede Kundenrezension ist inklusive: Title, Sterne, Autor der Rezension, Zeitpunkt, die Kommentartexte und so weiter. Abbildung 3.7 zeigt intuitiv die Struktur auf.

Bei der gezeigten Struktur ist der Daten-Scraping-Prozess einfach mit den obigen technischen Werkzeugen. Man kann die Tabelle von den Rezensionen lokalisieren mit folgendem Code in Python:

```
import scrapy

table = response.xpath('//*[@id="productReviews"]')

.xpath() ist eine Funktion, die von Bibliothek Scrapy geboten wird, die XPath-Ausdrücke
zu unterstützen. Dieser XPath-Ausdruck //*[@id="productReviews"] bedeutet,
dass das Element, das mit der Identität "productReviews" ist, selektiert wird.

divs = table.xpath('.//div[contains(@style, "margin-left:0.5em;")]')

Der obige Code selektiert die Elemente, die in der Tabelle von den Rezensionen sind. Jedes
selektierte Element ist ein Objekt, welches die Information von Titel, Autor, Zeitpunkt,
Kommentartexte und weiteren Attributen der Rezension inklusive ist. Für jedes Element
kann man mit folgendem Code die Informationen exportieren:
```

```
for div in divs:
    item = AmazonItem()
    item['titel']=div.xpath('.//span[contains(@style,
```

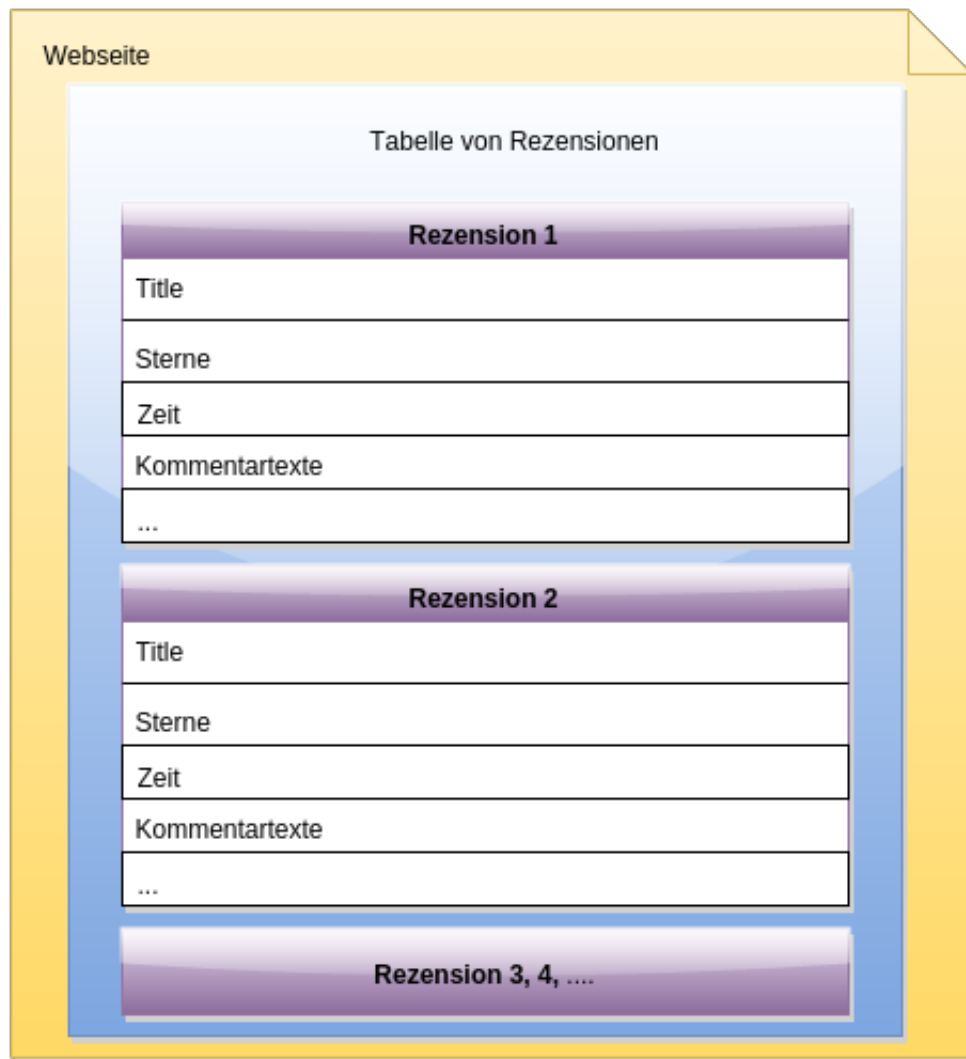


Abbildung 3.7: Die Struktur der Webseite von Amazon.de (Quelle: Eigene Darstellung)

```
"vertical-align:middle;")]/b/text()').extract()
item['kommentartexte']=div.xpath('.//div[@class="reviewText"]
/text()').extract()
yield item
```

Mit folgendem Bash-Code kann man alle Informationen in *.json*-Dokument exportieren. Ein *.json*-Dokument kann im Internet einfach ins CSV-Dokument transformiert werden. Dann ist das Ziel der Datensammlung von Amazon Deutschland erreicht.

```
scrapy crawl Amazon -o items.json
```


3.4.2 Sammeln der Daten aus Tmall.com

Man muss noch ein Programm in Python schreiben, um die Daten aus Tmall zu sammeln, weil Tmall und Amazon unterschiedliche Strukturen von Webseiten haben. Die folgende Abbildung 3.8 zeigt diese chinesischen Webseiten in Tmall.



Abbildung 3.8: Kundenrezensionen von Tmall.com (Quelle: Tmall.com)

Es ist einfach zu finden, da alle Daten der Kundenrezensionen in einem `.json`-File gespeichert sind. Deswegen braucht man nur dieses Dokument runterzuladen, und nach CSV-Datei zu transformieren. Mit folgendem Code kann man die Daten als `.json`-Datei speichern. Der reguläre Ausdruck `'\"rateList\":(\\[.*?\\])\\,\"tags\"'` macht das Dokument valid.

```
myjson = re.findall('\"rateList\":(\\[.*?\\])\\,\"tags\"',
,myweb.text)[0].encode('utf-8')
f = file('rate.json', 'a')
f.write(myjson)
```

Natürlich werden die gesammelten Daten auf chinesisch sein, weil fast alle Kunden in Tmall Chinesen sind. Um einen besseren Vergleich in dieser Arbeit zu machen, werden die chinesischen Daten ins Deutsche von Google Translator übersetzt. Dann ist das Ziel der Datensammlung aus Tmall erreicht.

3.5 Textverarbeitung

Vor dem Start der vollwertigen Analysen müssen die aus verschiedenen Quellen erfassten Rohdaten oft eine Vorverarbeitung machen (Ravi & Ravi, 2015). Einige in dieser Arbeit genutzten Vorverarbeitungsschritte sind:

Tokenisierung: Aufgabe zur Trennung des Volltextzeichenfolge in eine Liste von einzelnen Wörtern (Balazs & Velásquez, 2016). Es ist einfach dies in Leerzeichen getrennte Sprachen wie Englisch, Spanisch oder Deutsch durchzuführen, es wird aber wesentlich schwieriger in Sprachen in denen Wörter nicht durch Leerzeichen getrennte werden, wie Chinesisch oder Japanisch (Dale *et al.* , 2000, p. 9-30). Deshalb werden die Rohdaten von Tmall erst nach Deutsch übersetzt, und dann werden die Vorverarbeitungsschritte gemacht. Mit folgendem Code in R kann man diesen Schritt erledigen:

```
word.list = str_split(sentence, '\\s+')
```

Stemming: heuristische Verfahren zum Löschen von Wort Affixen und verlassen sie in einer invarianten kanonischen Form oder "i.e. stem" (Balazs & Velásquez, 2016). Zum Beispiel: Einkauf, Einkaufen, einkaufen werden als "Einkauf" gestemmt. Dieser Schritt wird in R so fertig gemacht:

```
docs <- tm_map(docs, stemDocument, language="german")
```

Stoppwörtern Entfernen: Aktivität zum Entfernen von Wörtern, die für die Strukturierung von Sprache verwendet werden, aber leisten keinen Beitrag zum Inhalt (Balazs & Velásquez, 2016). Einige dieser Wörter sind wie zum Beispiel: die, war, oder wird. In R wird es so gemacht:

```
docs <- tm_map(docs, removeWords, stopwords("german"))
```

Und man kann auch selbst einige Stoppwörter definieren. Zum Beispiel: beim, im, oder schon mit solchem Code:

```
docs <- tm_map(docs, removeWords, c("schon", "beim"))
```

Sonstiges Entfernen oder Änderungen: wie zum Beispiel: in Kleinbuchstaben umwandeln, die Zahlen und Satzzeichen entfernen. Mit folgendem Code werden sie beschäftigt:

```
docs <- tm_map(docs, removeNumbers)
```

```
docs <- tm_map(docs, content_transformer(tolower))
```

```
docs <- tm_map(docs, removePunctuation)
```

Nur mit diesen Vorverarbeitungsschritten können die Analysen in dieser Arbeit ausgeführt werden. Nach diesen Schritten werden die Vorbereitungen der Daten für die Analysen in dieser Arbeit geschaffen.

3.6 Maschinelle Übersetzung für mehrsprachige Sentiment-Analyse

Vor dem Beginn der Analyse, muss man erkennen, dass die in dieser Arbeit gemachte kulturrübergreifende Sentiment-Analyse eine mehrsprachige (Deutsch und Chinesisch) Analyse ist. Mit vielen Rohdaten, muss man in dieser Arbeit eine maschinelle Übersetzung machen. Deswegen werden die chinesischen Kommentaren ins Deutsche durch Google Translate übersetzt. In diesem Abschnitt wird diskutiert, ob die originalen Daten nach der maschinellen Übersetzung noch valid und zuverlässig sind.

Es gab schon in der Praxis einige Forschungen, die maschinelle Übersetzung für mehrsprachige Sentiment-Analyse benutzen, zum Beispiel: Demirtas (2013) machte eine mehrsprachige Sentiment-Analyse zwischen Türkisch und Englisch, bei der Google Translate für die türkischen OCRs von Filmen und verschiedenen Produkten benutzt wurde. Scholz *et al.* (2012) benutzten auch Google Translate für die fünfssprachige Sentiment-Analyse (Italienisch, Spanisch, Französisch, Deutsch und Englisch). Momtazi (2012) machte die Forschung von der deutschsprachigen Sentiment-Analyse auch mit Google Translate.

Andererseits machen die Wissenschaftler noch einige theoretische Grundlagen für die maschinelle Übersetzung. Balahur & Turchi (2014) haben herausgefunden, dass die Übersetzungsmaschine einen angemessenen Reifegrad erreichen muss, um ausreichend zuverlässige Daten für andere Sprachen zu produzieren. Die Lücke in der Klassifikationsleistung zwischen auf Englisch geschulten Systemen und übersetzten Daten ist minimal, mit einem Maximum von 11.8% zugunsten der Ausgangssprachedaten. Experimentelle Ergebnisse von Wan (2008) zeigen, dass die Analyse der Reviews auf Englisch, die durch die Übersetzungsdienst (zum Beispiel: Google Translate und Yahoo Babel Fish) übersetzt werden, an Leistung übertrifft, als die Analyse der ursprünglichen chinesischen Reviews.

In dieser Arbeit, werden die chinesischen Kommentaren auf Deutsch durch Google Translate übersetzt. Es wird wenige Einflüsse geben, nicht nur wegen den Ergebnissen von Praxis- und Theorieseiten, sondern auch wegen den folgenden Gründen:

1. Wie in Abschnitt 2.2.4 gesagt, wird das Lexikon von Remus *et al.* (2010) in dieser Arbeit verwendet. Dieses Lexikon nach Remus *et al.* (2010) wird auch teilweise durch Google Translate von Englisch auf Deutsch übersetzt.
2. Nach der Textverarbeitung, werden die Rohdaten in den einzelnen Wörtern getrennt, und in eine invariante kanonischen Form gestellt. Das heißt: die Reihenfolge des Satzes und Formen der Wörtern sind nicht wichtig. Zum Beispiel, es gibt solche Kommentare in Chinesisch:

东西很好, 昨天穿了刚刚好! 我180、74kg! 很不错, 就是客服比较傲, 经常

不回复。

Der Kunde meint:

Die Dinge sind gut, ich habe sie gestern angezogen, genau das Richtige! Ich bin 180cm groß, wiege 74kg. Sehr gut, aber der Mitarbeiter des Kundendienstes ist relativ stolz, weil er oft nicht auf mich reagiert.

Durch Google Translate:

Etwas Gutes gestern, tragen genau das Richtige! Ich 180,74kg! Sehr gut, ist es relativ stolz auf die Kunden-Service, die oft nicht reagieren.

Nach der Textverarbeitung werden nur die Wörtern behalten:

Chinesisch	好	刚刚好	不错	傲
Deutsch	gut	genau richtig	gut	stolz
Google Translate	gut	genau richtig	gut	stolz

Tabelle 3.2: Vergleich zwischen manueller und maschineller Übersetzung nach der Textverarbeitung (Quelle: eigene Darstellung)

Durch Tabelle 3.2 kann man deutlich sehen, dass die Übersetzungsergebnisse von Maschine oder vom Menschen gleich oder fast gleich sind. Verglichen mit der manuellen Übersetzung, macht sich die Maschine schon gut genug für die Forschung, und es ist schneller und kostet weniger. Aus diesen Gründen macht Google Translate die Übersetzung in dieser Arbeit.

4.1 Die deskriptiven Ergebnisse zu den Online Consumer Reviews

Eine Wortwolke ist eine Methode zur Informationsvisualisierung, bei der eine Liste aus Schlagworten, oft alphabetisch sortiert, flächig angezeigt wird, wobei einzelne unterschiedlich gewichtete Wörter größer oder auf andere Weise hervorgehoben dargestellt werden (Bateman *et al.* , 2008). Die Abbildung 4.1 zeigt die Wortwolken von chinesischen sowie deutschen OCRs.



Die beiden Wortwolken sind aus den Wörtern aufgebaut, bei denen die Menge des Auftretens

größer als ein Promille der Menge von chinesischen oder deutschen OCRs ist. Von Abbildung 4.1 kann man wissen, dass die Kleidungen von allen Marken in beiden Ländern als “gut” bewertet wurden. Worum die Kunden sich zuerst kümmern sind die Größe, Tragbarkeit, und Qualität. Die chinesischen Kunden ziehen mehr die logistische Dienstleistung in Erwägung und die Deutschen berücksichtigen mehr den Preis. Diese Gemeinsamkeiten und Unterschiede kann man auch durch die Top 10 Wörter in Abbildung 4.2 zusammenfassen.

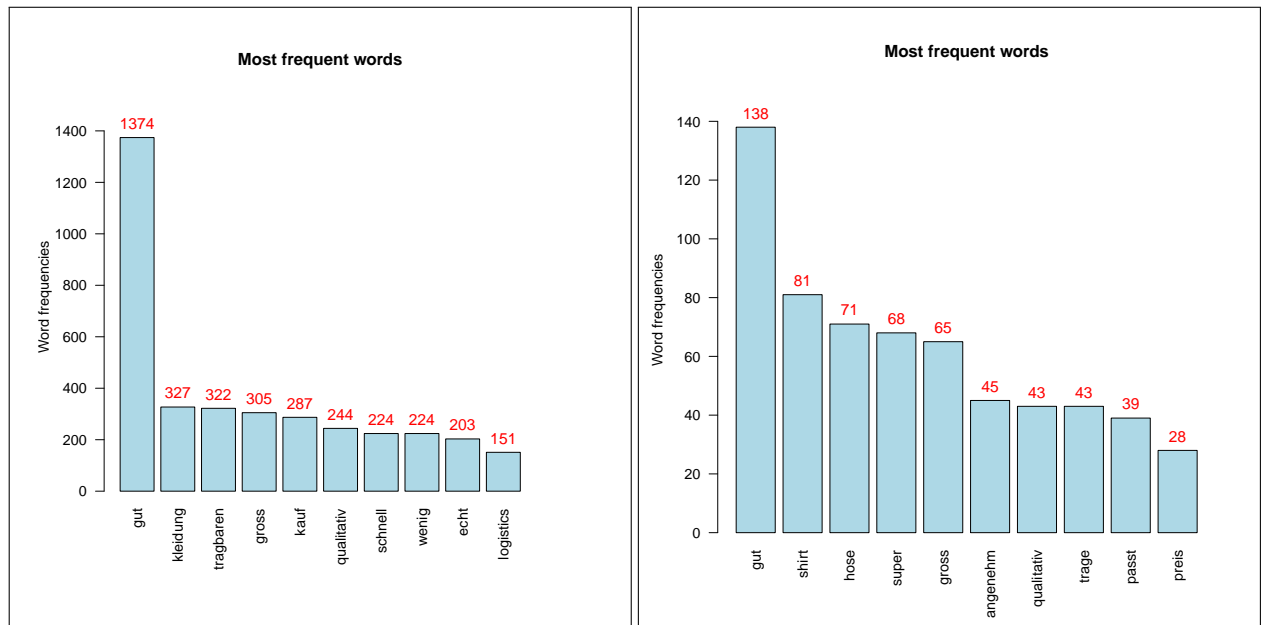


Abbildung 4.2: die Top 10 Wörter von chinesischen (Links) sowie deutschen (Rechts) OCRs (Quelle: Eigene Darstellung)

Von diesen chinesischen Top 10 Wörtern, sieht man auch, dass die chinesischen Kunden viel mehr das Wort “gut” verwenden wollen als die Deutschen. Das kann auch das Problem von maschineller Übersetzung sein. Außerdem beachten die Chinesen die Logistik (“logistics”, “schnell”) und ob die Waren zertifiziert sind (“echt”), während die Deutschen das Gefühl der Aussicht oder den Preis mehr berücksichtigen (“angenehm”, “passt”, “preis”).

Die deutschen Reviews werden in dem Zeitraum von 7. Januar 2012 bis 9. Mai 2015 geschrieben, und die Chinesen haben die Kommentare zwischen dem 10. Februar 2014 und dem 2. Juni 2015 geschrieben. Dieser Datensammlungsprozess ist genau am 2. Juni 2015 durchgeführt worden. Das heißt: die Chinesen wollen gern OCRs schreiben aber die deutschen Kunden vielleicht eher nicht.

In den Abschnitten 4.2 bis 4.6 wird jede Hypothese in unterschiedlichen Kategorien und Marken getestet. Danach wird sie in allgemeiner Hinsicht zusammengefasst. Eine geschlechts-spezifische Analyse wird aus folgenden Gründen nicht in dieser Arbeit erstellt:

1. Wie es in Abschnitt 2.1.5 geschrieben steht, waren die demographischen Variablen wie Alter und Geschlecht in keinem Zusammenhang mit Online-Bekleidungseinkauf

(Goldsmith & Goldsmith, 2002).

2. Die Kunden sind anonym im Internet. Durch die OCRs ist das Geschlecht nicht sichtbar. Es ist nicht so gewissenshaft zu sagen, dass die Frauen nur die weiblichen Produkte kaufen wollen.
3. Die Abbildung 2.9 zeigt, dass die Dimension: “Maskulinität gegenüber Femininität” von Deutschland und China die gleichen Punkte hat. Das heißt: Die geschlechtspezifische Auswirkung zwischen Deutschland und China in kulturellem Bereich ist gleich, deshalb braucht man diese nicht zu unterscheiden.

Die folgende Tabelle 4.1 zeigt den Preis des jeweiligen Produkts in beiden Ländern. Daran sieht man, dass der Preis des Produkts in China und Deutschland keinen großen Unterschied hat. Und laut Lu (2015) sind die Forschungsergebnisse für die Wichtigkeit des Preises inkonsistent. Deshalb wird der Faktor des Preises in dieser Arbeit nicht berücksichtigt.

Produkt	Deutschland	China
Adidas männlich	€ 29,99	CNY 199 (\approx € 29,70)
Adidas weiblich	€ 34,99	CNY 250 (\approx € 37,31)
Nike	€ 56	CNY 359 (\approx € 53,58)
Puma	€ 16,98	CNY 119 (\approx € 17,76)

Tabelle 4.1: Der jeweilige Preis von jeder Kleidung in Deutschland und China (Datenquelle von Amazon und Tmall, Euro gegenüber dem CNY-Wechselkurs von Bank of China: 6,70, am 2. Juni 2015)

4.2 Hypothesenüberprüfung

Die Nullhypothese behandelt die wesentliche Frage, ob es kulturelle Einflüsse auf den quantitativen und qualitativen OCRs gibt. Wenn es keine Einflüsse darauf gibt, sollten die quantitativen und qualitativen Valenzen beide normalverteilt sein.

In dieser Arbeit wird der Shapiro-Wilk-Test durchgeführt. Der Shapiro-Wilk-Test ist ein statistischer Signifikanztest, der die Hypothese überprüft, dass die zugrunde liegende Grundgesamtheit einer Stichprobe normalverteilt ist. Die Nullhypothese H_0 nimmt an, dass eine Normalverteilung der Grundgesamtheit vorliegt. Wird alternativ der p-Wert des Tests ermittelt, so wird die Nullhypothese in der Regel nicht abgelehnt, wenn der p-Wert größer ist als das festgelegte Signifikanzniveau α . Der Test kann zum Überprüfen von univariaten Stichproben mit 3 bis 5000 Beobachtungen eingesetzt werden. (Shapiro & Wilk, 1965)

Nach Royston (1995, p. 547) ist die Hypothese H_0 abgelehnt wenn der p-Wert kleiner als 0,1 ist. Tabelle 4.2 zeigt die Ergebnisse für die beiden Ländern.

p-Wert	quantitative OCRs	qualitative OCRs
China	$< 2.2 \times 10^{-16}$	$< 2.2 \times 10^{-16}$
Deutschland	$< 2.2 \times 10^{-16}$	0.003672

Tabelle 4.2: Die Ergebnisse durch Shapiro-Wilk-Test für China und Deutschland (Quelle: Eigene Darstellung)

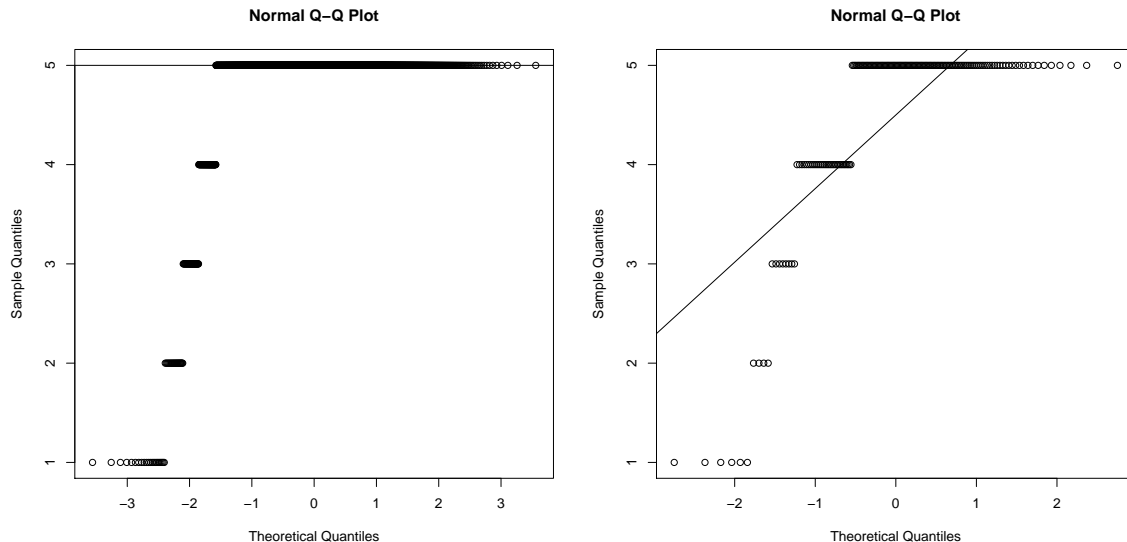


Abbildung 4.3: Das Q-Q-Plot der quantitativen Valenzen für China (Links) und Deutschland (Rechts) Quelle: Eigene Darstellung

Aus dieser Tabelle kann man sehen, dass alle vier p-Werte kleiner als 0,1 sind. Mit dem Quantile-Quantile-Plot (Q-Q-Plot) kann man die empirischen Häufigkeitsverteilung mit der hypothetischen Verteilung bzw. Normalverteilung vergleichen. Die Abbildungen 4.3 und 4.4 zeigen die Q-Q-Plot der quantitativen und qualitativen Valenzen graphisch auf. Besonders in dem Q-Q-Plot der deutschen qualitativen Valenz sind die Verteilungen sehr ähnlich wie die Normalverteilung, aber der p-Wert sieht anders aus. Durch diese Abbildungen und Tabelle 4.2 wird die Hypothese H_0 abgelehnt. Anders gesagt, es gibt tatsächlich kulturellen Einflüsse auf quantitativen und qualitativen Online Consumer Reviews.

Nur mit diesem Ergebnis kann man die weiteren Hypothesen testen. In Abschnitt 2.5 werden vier Attribute der OCRs an der quantitativen und qualitativen Seite erstellt. Die Hypothesen 1 bis 4 testen, ob es kulturelle Einflüsse auf die vier Attribute gibt. In den Abschnitten 4.3 bis 4.6 sind die empirischen Ergebnisse.

4.3 Hypothese 1

Hypothese 1 beschäftigt sich mit der Frage, ob es kulturelle Einflüsse auf die Menge und Länge der OCRs gibt. In diesem Abschnitt wird Hypothese 1 in Allgemeinen bzw. auch in

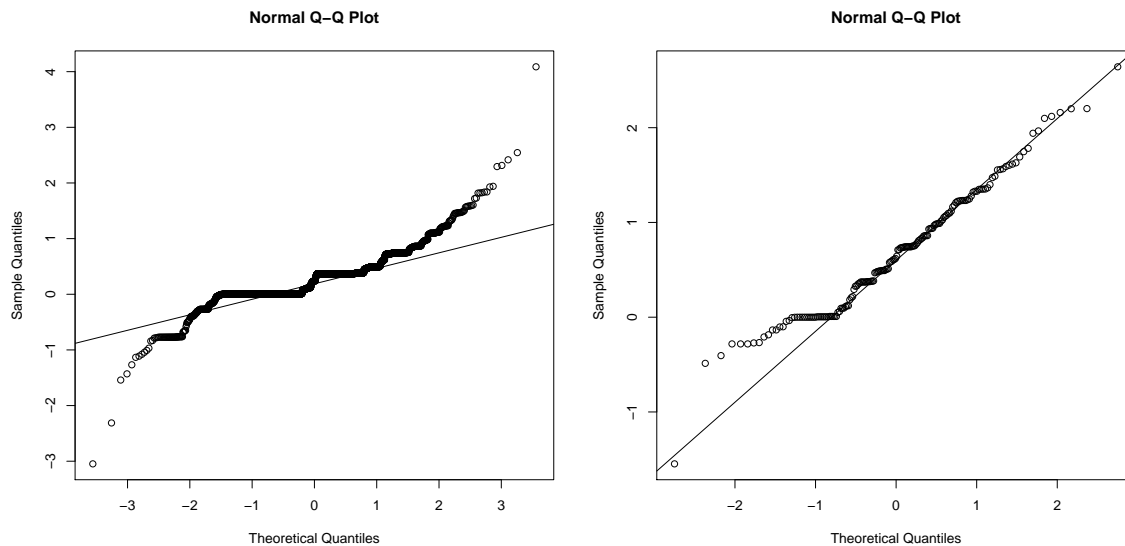


Abbildung 4.4: Das Q-Q-Plot der qualitativen Valenzen für China (Links) und Deutschland (Rechts)
Quelle: Eigene Darstellung

verschiedenen Kategorien (T-Shirt und Hose) und Marken (Adidas, Nike, Puma) getestet. Damit kann man nicht nur erkennen, ob Hypothese 1 in allgemeiner Hinsicht unterstützt wird, sondern auch in verschiedenen Situationen.

4.3.1 Hypothese 1 in unterschiedlichen Kategorien

Insgesamt wurden 2402 chinesische OCRs über T-Shirts gesammelt, während es bei den deutschen nur 98 sind. Über Hosen, wurden 266 chinesische OCRs gesammelt und 70 deutsche. Tabelle 4.3 zeigt die Ergebnisse.

T-Shirt	Wörter	Zeichen	Menge
China	22255	139258	2402
Deutschland	3821	24039	98
Hose	Wörter	Zeichen	Menge
China	2744	17021	266
Deutschland	2411	14658	70

Tabelle 4.3: Das Volumen der OCRs über T-Shirt und Hose von qualitativen und quantitativen Seiten
(Quelle: Eigene Darstellung)

Man kann einfach ausrechnen, dass die chinesischen Kunden weniger Wörter benutzt haben als die deutschen. Im Durchschnitt benutzt ein chinesischer Kunde nur 9 bis 10 Wörtern während der deutsche Kunden 34 bis 39 benutzt hat. Das heißt: die chinesischen OCRs haben meistens nur einen Satz und die Deutschen schreiben mehreren Sätzen. Durch diese Daten

kann man auch einfach das durchschnittliche Volumen ausrechnen. Dieses Volumen wird in Abbildung 4.5 verglichen.

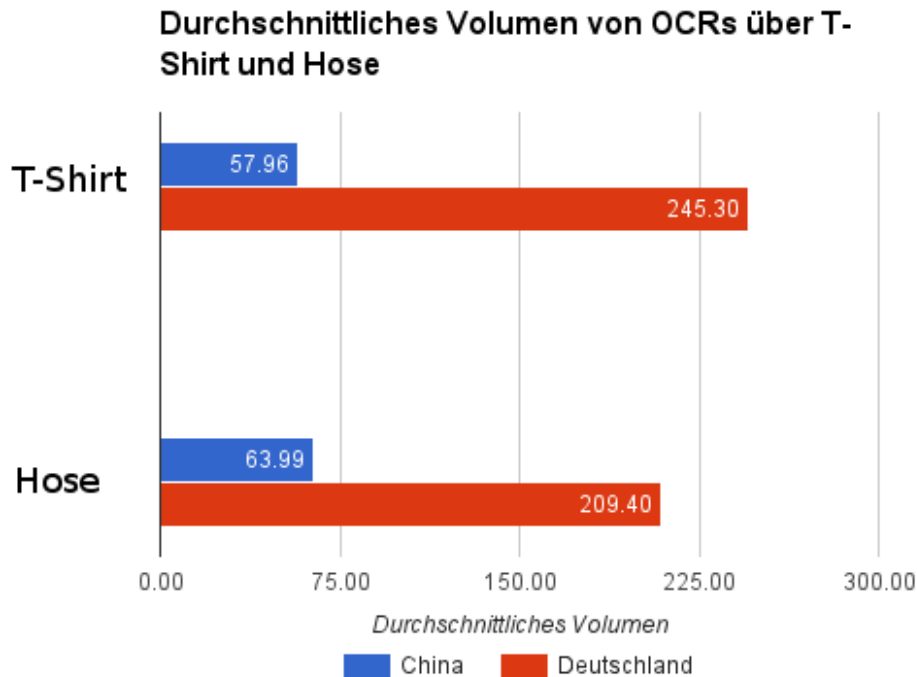


Abbildung 4.5: Durchschnittliches Volumen von qualitativen OCRs über T-Shirt und Hose (Quelle: Eigene Darstellung)

Aus dieser Abbildung kann man erfassen, dass das deutsche durchschnittliche Volumen drei bis vier mal so groß wie das chinesische ist. Das entspricht auch der Anzahl von Wörtern. Das heißt: Im Durchschnitt schreiben die deutschen Kunden drei bis vier Sätzen, während die Chinesen nur einen Satz schreiben, aber sie schreiben OCRs gerne nach dem Kauf. Deshalb wird Hypothese 1 im Bereich OCRs von T-Shirt und Hose nur in dem qualitativen Aspekt gestützt.

4.3.2 Hypothese 1 bei unterschiedlichen Marken

Wie vorher beschrieben wurde, wurden die Daten aus insgesamt drei unterschiedlichen Marken gesammelt: “Adidas”, “Nike” und “Puma”. Tabelle 4.4 zeigt das Volumen der qualitativen OCRs von den drei Marken.

Durch Tabelle 4.4 steht es fest, dass die Menge der chinesischen OCRs bei jeder Marke größer als die von den Deutschen ist. Bei der Marke “Nike” ist die Menge der chinesischen OCRs sogar fast das 50fache der deutschen. Bei der Marke “Adidas” ist die Menge der Wörtern von chinesischen und deutschen fast gleich, trotz der Tatsache dass die chinesischen OCRs drei mal mehr als die deutschen OCRs sind. Folgende Abbildung 4.6 zeigt den großen

Adidas	Wörter	Zeichen	Menge
China	2744	17021	266
Deutschland	2411	14658	70
Nike	Wörter	Zeichen	Menge
China	17541	109589	1923
Deutschland	1909	12023	40
Puma	Wörter	Zeichen	Menge
China	4714	29669	479
Deutschland	1912	12016	58

Tabelle 4.4: Das Volumen der OCRs bei verschiedenen Marken (Quelle: Eigene Darstellung)

Unterschied.

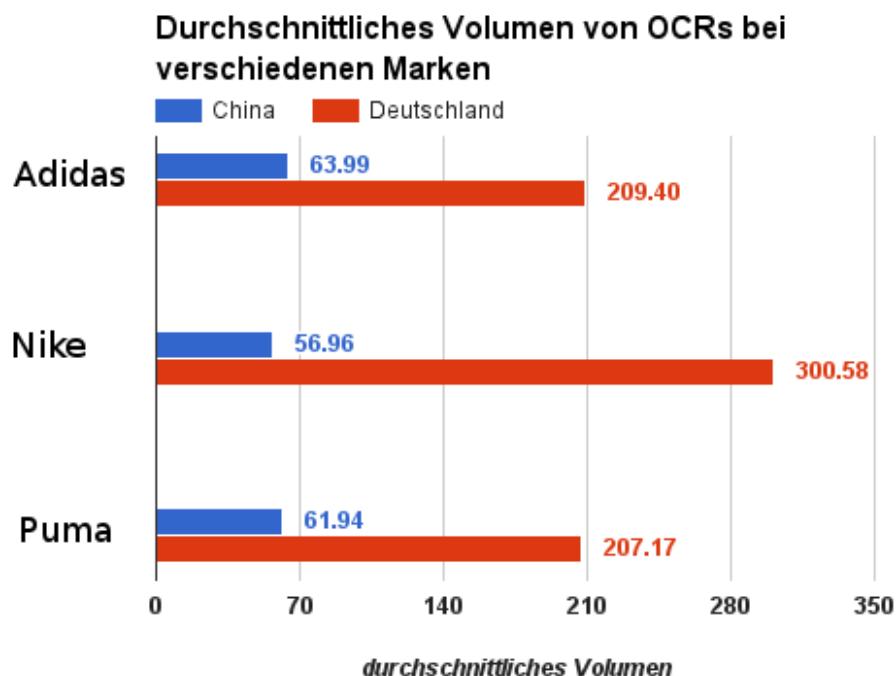


Abbildung 4.6: Durchschnittliches Volumen von qualitativen OCRs bei verschiedenen Marken (Quelle: Eigene Darstellung)

Daran sieht man dass die chinesischen Kunden beim Online-Einkauf bei diesen drei Marken nur ungefähr 60 Zeichen in den OCRs schreiben, aber die Deutschen schreiben mehr, von 200 bis 300 Zeichen. Insbesondere schreiben die deutschen Kunden von Nike mehr als die von den anderen zwei Marken. Dieses Ergebnis ist ähnlich wie das Ergebnis in Abbildung 4.5. Die deutschen Kunden schreiben längere aber weniger OCRs als die chinesischen Kunden. Das heißt, dass Hypothese 1 bei unterschiedlichen Marken auch nur in dem qualitativen Aspekt gestützt wird.

4.3.3 Hypothese 1 im Allgemeinen

Weil die Hypothese 1 in den beiden Kategorien und allen drei verschiedenen Marken im quantitativen Aspekt abgelehnt aber im qualitativen Aspekt gestützt wird, wird vorgeschlagen, dass die Hypothese 1 in der gleichen Situation in allgemeiner Weise ist. Tabelle 4.5 zeigt das allgemeine Ergebnis.

	Wörter	Zeichen	Menge
China	24999	156279	2668
Deutschland	6232	38697	168

Tabelle 4.5: Das Volumen der OCRs im Allgemeinen von qualitativen und quantitativen Seiten (Quelle: Eigene Darstellung)

Aus dieser Tabelle kann man erfassen, dass bei den chinesischen OCRs nur durchschnittlich 9,37 Wörtern (58,58 Zeichen) benutzt werden, aber die deutschen Kunden schreiben im Durchschnitt 37,10 Wörter (230,34 Zeichen). Abbildung 4.7 zeigt das Ergebnis graphisch.

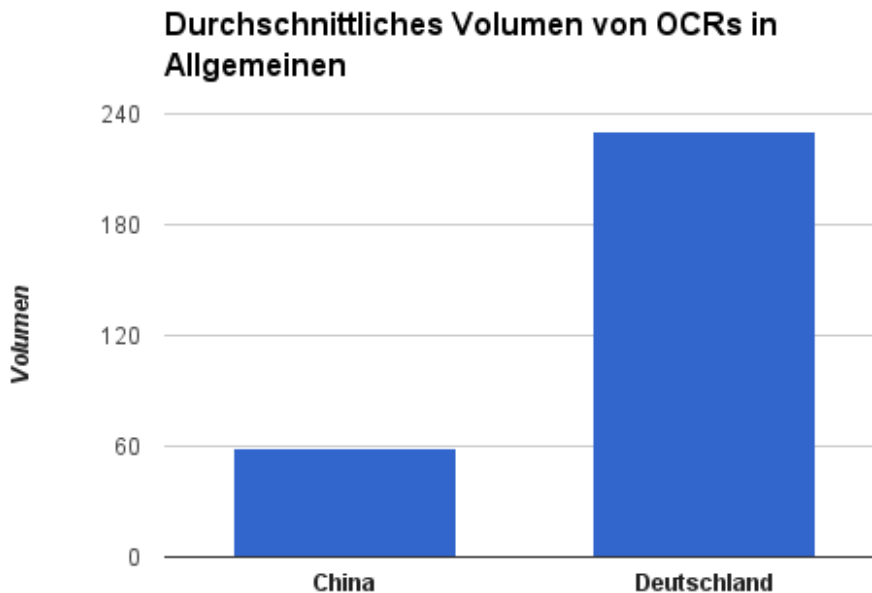


Abbildung 4.7: Durchschnittliches Volumen von qualitativen OCRs (Quelle: Eigene Darstellung)

Die meisten chinesische OCRs haben nur ein oder zwei Wörter, sowie: “gut”, “zufrieden”, “Lob”, “Okay”, usw. Aber die Deutschen schreiben mehr. Wer mehr schreibt, teilt mehr und gerne die Information. Deshalb sind die chinesischen Kunden zurückhaltender beim Informationsaustausch als die Deutschen.

Wegen der kollektivistischen Kultur bevorzugen die chinesischen Kunden, nacheinander OCRs zu schreiben, wenn es schon große Mengen von OCRs gibt. Deshalb sieht man in

Tabelle 4.4, dass es fast 2000 OCRs für die Marke “Nike” gibt während es nur 266 Reviews für “Adidas” sind.

Zusammengefasst sind die chinesischen Reviews deutlich kürzer, aber mehr als die deutschen Reviews. Das heißt: In Quantität sind die chinesischen OCRs viel mehr als die deutschen, aber in Qualität sind die deutschen OCRs viel besser. Die Hypothese 1 wird in dem quantitativen Aspekt abgelehnt aber in dem qualitativen Aspekt gestützt.

4.4 Hypothese 2

Hypothese 2 wird in dieser Arbeit vorgeschlagen, um die Valenz der quantitativen und qualitativen OCRs zu unterscheiden und zwischen den beiden Ländern zu vergleichen. Damit wird sie, ähnlich wie Hypothese 1, in unterschiedlichen Kategorien und Marken diskutiert.

4.4.1 Hypothese 2 in unterschiedlichen Kategorien

Wie in Abschnitt 4.3 beschrieben wird, haben die Daten in zwei Kategorien unterschiedliche Mengen. Deshalb ist es besser, die Analyse getrennt für T-Shirt und Hose durchzuführen.

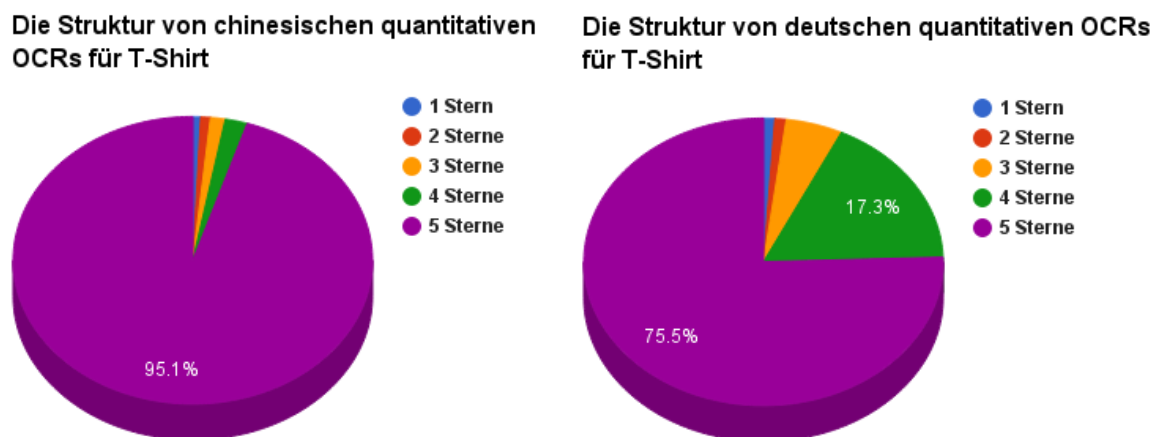


Abbildung 4.8: Die Struktur von chinesischen (Links) und deutschen (Rechts) quantitativen OCRs über T-Shirt (Quelle: Eigene Darstellung)

Abbildung 4.8 zeigt die Struktur der OCRs über T-Shirts aus quantitativer Sicht in den beiden Ländern. Dort kann man erkennen, dass die Mehrheit der OCRs aus den beiden Ländern fünf Sterne für das T-Shirt geben. Aber der Unterschied ist auch sichtbar. Die chinesischen Kunden geben viel lieber fünf Sterne als die Deutschen. 95% der chinesischen OCRs sind fünf Sterne, bei den deutschen dagegen nur 75%. Das heißt auch, dass die chinesischen quantitativen OCRs viel dichter als die deutschen sind.

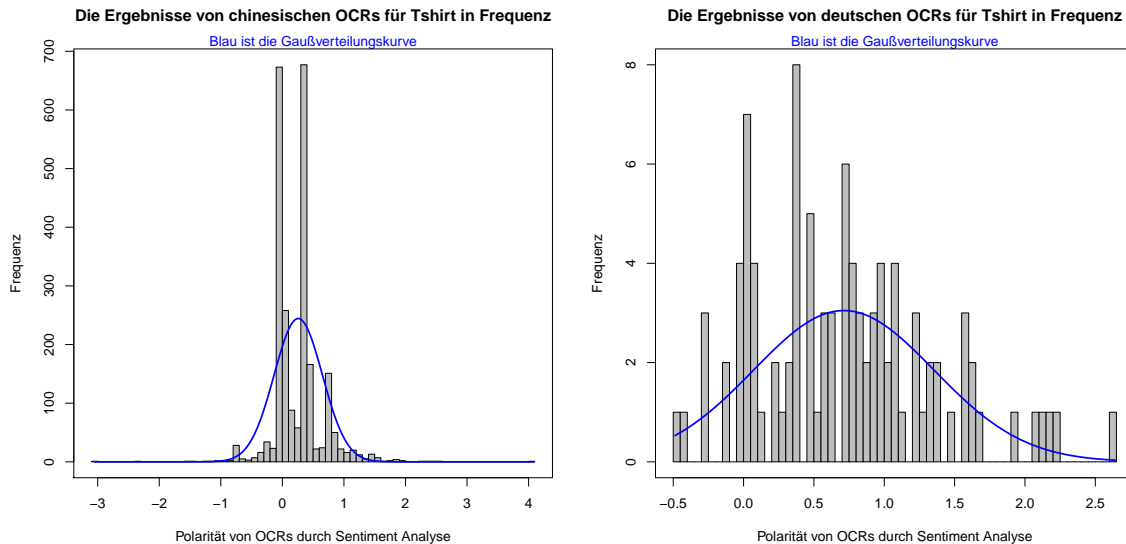


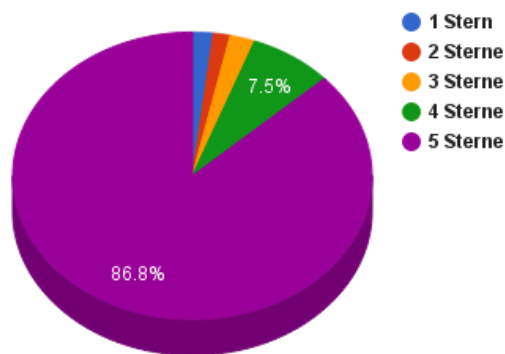
Abbildung 4.9: Die Verteilung der Valenzen von chinesischen und deutschen qualitativen OCRs für T-Shirt (Quelle: Eigene Darstellung)

Es lässt sich der Abbildung 4.9 entnehmen, dass die Valenzen von chinesischen qualitativen OCRs viel dichter und zentralisierter als die von Deutschen sind. Die chinesischen qualitativen OCRs haben zwei große Gruppen aber die deutschen Polaritäten sind vielfältiger und dezentralisierter. Nach der Einsicht in die Rohdaten ist es klar, dass die beiden Valenzen von den zwei großen Gruppen die Werte 0 und 0,3716 festsetzen. Durch das Lexikon weiß man, dass die Polarität des Wortes “gut” 0,3716 ist. Polarität 0 bei chinesischen OCRs tritt 589 mal auf, und 0,3716 tritt 502 mal auf. Das heißt auch, dass die chinesischen Kunden meistens kein Gefühl ausgedrückt oder nur “gut” geschrieben haben, aber die deutschen Kunden schreiben mehr von ihrem eigenen Gefühl über das T-Shirt. Die blauen Linien sind die entsprechenden Normalverteilungen. Aus der Abbildung geht hervor, dass die Valenzen nicht normal verteilt sind.

Für die Hosen ist das Ergebnis graphisch fast gleich. Abbildung 4.10 zeigt dass über 86% der chinesischen Kunden fünf Sterne für die Hosen geben aber die Zahl der Deutschen ist nur 64,3%. Aus der Abbildung 4.11 geht hervor, dass 24,81% der chinesischen Kunden nichts über ihr Gefühl gesagt haben und 14,29% der Chinesen nur “gut” geschrieben haben. Aber wie die deutschen Kunden in Abbildung 4.9 sind, bleiben hier die Deutschen in der Abbildung 4.11 dezentralisiert.

Weil die chinesischen Kunden wesentlich mehr “fünf Sterne” vergeben haben als die Deutschen (durch Abbildung 4.8 und 4.10), ist es natürlich, dass die durchschnittliche Valenz der deutschen quantitativen OCRs kleiner als die von den chinesischen ist. Das heißt: Auf der Seite der quantitativen OCRs, ist die Valenz der Chinesischen besser als die der Deutschen. Aber die Tabelle 4.6 zeigt Etwas anderes an der Seite der qualitativen OCRs. Die chinesischen

Die Struktur von chinesischen quantitativen OCRs für Hose



Die Struktur von deutschen quantitativen OCRs für Hose

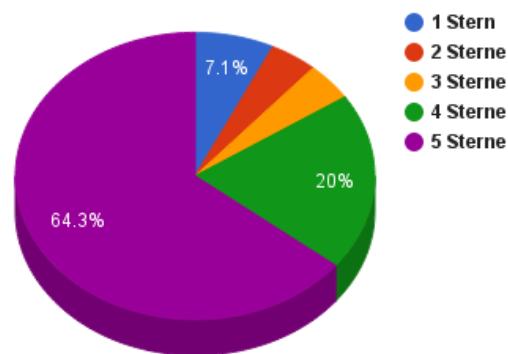


Abbildung 4.10: Die Struktur von chinesischen (Links) und deutschen (Rechts) quantitativen OCRs über Hose (Quelle: Eigene Darstellung)

qualitativen Valenzen sind in beiden Kategorien kleiner als die von den Deutschen. Das heißt: Basierend auf dem gleichen Produkt drücken die deutschen Kunden sich selbst positiver aus, als die Chinesen, obwohl sie schlechte Sterne vergeben haben. Die Tabelle bestätigt die Begründung 3 am Ende des Abschnitts 2.1.4. Zusammenfassend wird die Hypothese 2 in unterschiedlichen Kategorien bzw. T-Shirt und Hose in dieser Arbeit nur auf der qualitativen Seite gestützt.

	quantitative Valenz		qualitative Valenz	
	China	Deutschland	China	Deutschland
T-Shirt	4,898876	4,653061	0,2579923	0,7124276
Hose	4,759398	4,3	0,2023165	0,6110686

Tabelle 4.6: Die Valenz von quantitativen und qualitativen OCRs über unterschiedlichen Kategorien in beiden Ländern (Quelle: Eigene Darstellung)

4.4.2 Hypothese 2 bei unterschiedlichen Marken

In diesem Abschnitt wird die Hypothese 2 bei den drei ausgewählten Marken getestet. Weil die beiden Hosen von Adidas produziert werden, sind die Ergebnisse von Adidas gleich wie die Ergebnisse von den Hosen. Deshalb repräsentieren die Abbildungen 4.10 und 4.11 auch die Ergebnisse von der Marke "Adidas". Die Ergebnisse von den anderen zwei Marken werden nicht wie Abbildungen 4.8 oder 4.9 national verglichen, sondern bei unterschiedlichen Marken im einzelnen Land verglichen. Damit wird getestet, ob die Ergebnisse wirklich stabil oder nur im Durchschnitt stabil sind.

4 Empirische Ergebnisse

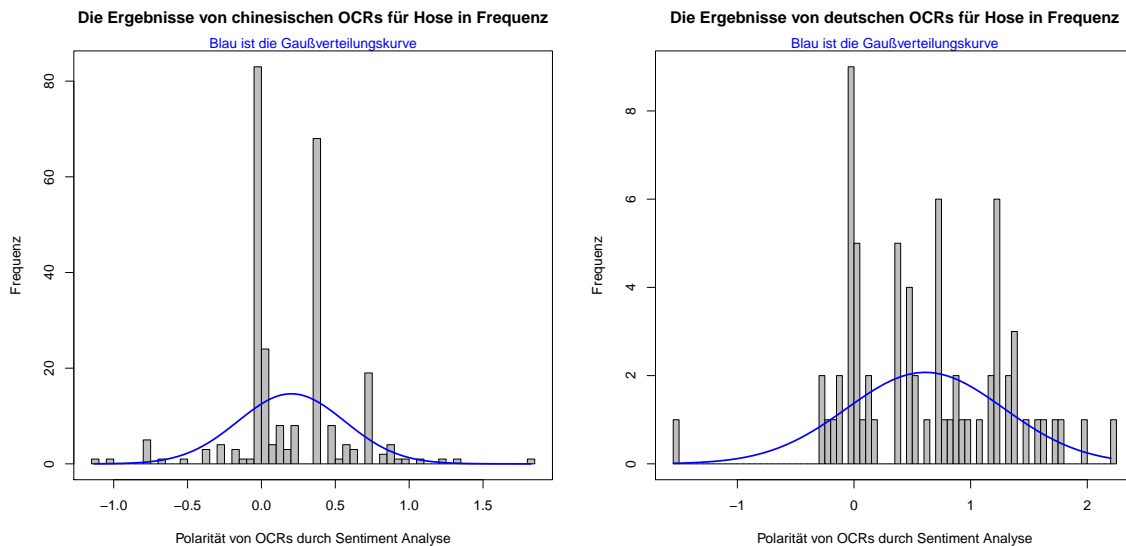


Abbildung 4.11: Die Verteilung der Valenzen von chinesischen und deutschen qualitativen OCRs für Hose (Die blauen Linien sind die entsprechende Normalverteilung.) (Quelle: Eigene Darstellung)

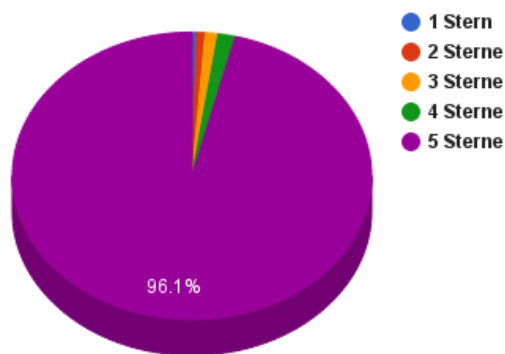
Durch die Abbildungen 4.12 und 4.13 kann man erfassen, dass über 90% der chinesischen quantitativen OCRs 5 Sterne haben, obwohl viele chinesische Kunden nichts über ihr Gefühl sagen oder einfach nur “Gut” schreiben. Diese Ergebnisse sind gleich den Ergebnissen über die chinesischen OCRs über T-Shirts und Hosen. Damit meint man: Die Ergebnisse für unterschiedliche Kategorien sind stabil für die Chinesen auch bei unterschiedlichen Marken. Die deutschen Ergebnisse von den zwei Marken werden durch Abbildungen 4.14 und 4.15 gezeigt. An den Abbildungen sieht man dass die deutschen Ergebnisse sich nicht ändern. Ungefähr 75% der deutschen Kunden vergeben “fünf Sterne”, und die Valenzen von deutschen qualitativen OCRs sind voneinander getrennt. Besonders bei Puma gibt es noch keine quantitative Bewertung, die einen oder zwei Sterne hat.

Deswegen wird vorgeschlagen, dass die Hypothese 2 bei unterschiedlichen Marken nur auf der qualitativen Seite gestützt ist. Die Tabelle 4.7 zeigt die Ergebnisse. Dadurch erfasst man dass die chinesischen quantitativen OCRs besser als die deutschen sind, aber die qualitativen OCRs nicht so gut sind.

Marke	quantitative Valenz		qualitative Valenz	
	China	Deutschland	China	Deutschland
Adidas	4,759398	4,3	0,2023165	0,6110686
Nike	4,921477	4,6	0,2688356	0,7057425
Puma	4,807933	4,689655	0,2149992	0,7170379

Tabelle 4.7: Die Valenzen von chinesischen und deutschen, quantitativen und qualitativen OCRs bei unterschiedlichen Marken (Quelle: Eigene Darstellung)

Die Struktur von chinesischen quantitativen OCRs bei Nike



Die Struktur von chinesischen quantitativen OCRs bei Puma

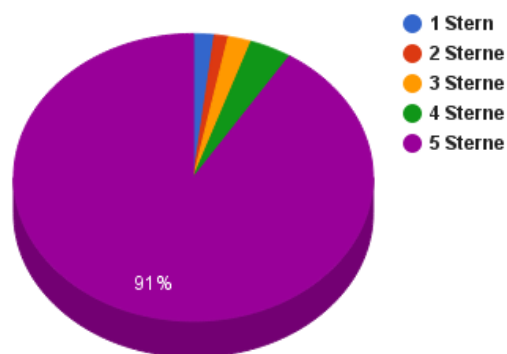


Abbildung 4.12: Die Struktur von chinesischen quantitativen OCRs bei Nike und Puma (Quelle: Eigene Darstellung)

4.4.3 Hypothese 2 in Allgemeinen

In diesem Abschnitt wird die Hypothese 2 in allgemeiner Hinsicht getestet. Obwohl es möglich ist, dass die Besonderheiten von Marken oder Kategorien im Durchschnitt ausgeglichen werden könnten, braucht man ein allgemeines Ergebnis. Außerdem ist es wichtig, eine allgemeine Regelung zu finden, damit sie meisten den Situationen entsprechen kann.

In den Abbildungen 4.16 und 4.17 kann man feststellen, dass die chinesischen Kunden viel lieber “fünf Sterne” als quantitativen Bewertungen für die Produkte vergeben, aber gleichzeitig drücken sie ihre Gefühle zentralisierter und dichter über die Produkte aus. Die Deutschen vergeben nicht so viele “fünf Sterne” Bewertungen, und sie berichten vielfältiger und individueller von ihren Gefühlen durch die qualitativen OCRs. Auch aus dieser Abbildung kann man die Nullhypothese in Abschnitt 4.2 ablehnen.

	Valenz von quantitativen OCRs	Valenz von qualitativen OCRs
China	4,884933	0,2525381
Deutschland	4,505952	0,6701946

Tabelle 4.8: Die Valenzen der chinesischen und deutschen OCRs auf der quantitativen und qualitativen Seite (Quelle: Eigene Darstellung)

In der Tabelle 4.8 sieht man schon, dass die deutschen qualitativen Bewertungen viel positiver als die Chinesischen sind, obwohl die Deutschen normalerweise weniger Sterne als die Chinesen vergeben haben. Durch diese Ergebnisse stimmt die Hypothese 2 nur auf der qualitativen Seite im Allgemeinen.

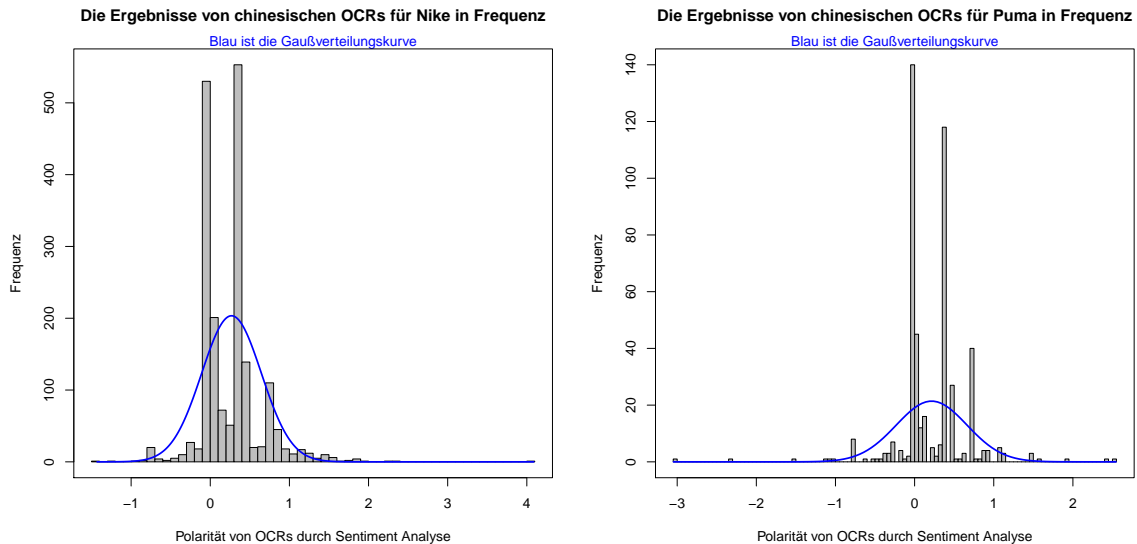


Abbildung 4.13: Die Verteilung der chinesischen Valenzen von qualitativen OCRs bei zwei unterschiedlichen Marken (Die blauen Linien sind die entsprechende Normalverteilung.) (Quelle: Eigene Darstellung)

4.5 Hypothese 3

Bei der Hypothese 3 gibt es einen theoretischen Widerspruch. Einerseits erzeugt große Machtdistanz Ungleichheit und damit große Varianz von Valenzen der OCRs nach Hofstede (1991). Andererseits meinen Luo *et al.* dass die OCRs konsistenter beim Kollektivismus sein werden. Deswegen sollte die Varianz von Valenzen der chinesischen OCRs kleiner sein. In diesem Abschnitt wird dieser Konflikt bei jedem Artikel und im Allgemeinen getestet, um eine dominierende Dimension herauszufinden.

4.5.1 Hypothese 3 bei jedem Artikel

Die Varianz beschreibt wie stark die Meinungen und Gefühle von Menschen in den OCRs schwanken. Verschiedene Artikeln, obwohl in gleicher Kategorie oder bei gleicher Marke, sind für die Menschen unterschiedlich, weil sie vielleicht unterschiedliche Qualität haben. Deshalb ist es besser, getrennt bei jeden Artikel zu analysieren, als in Kategorien oder Marken.

Wie es am Ende des Abschnitts 3.2 beschrieben wird, sind die zwei T-Shirts jeweils von Nike und Puma, und die beiden Hosen werden, jeweils eine männliche und eine weibliche, von Adidas produziert. Deshalb werden sie in dieser Arbeit nach Marken und Geschlecht genannt.

Um die Heterogenität der Varianzen besser zu bestimmen, wird hier der Levene-Test durch-

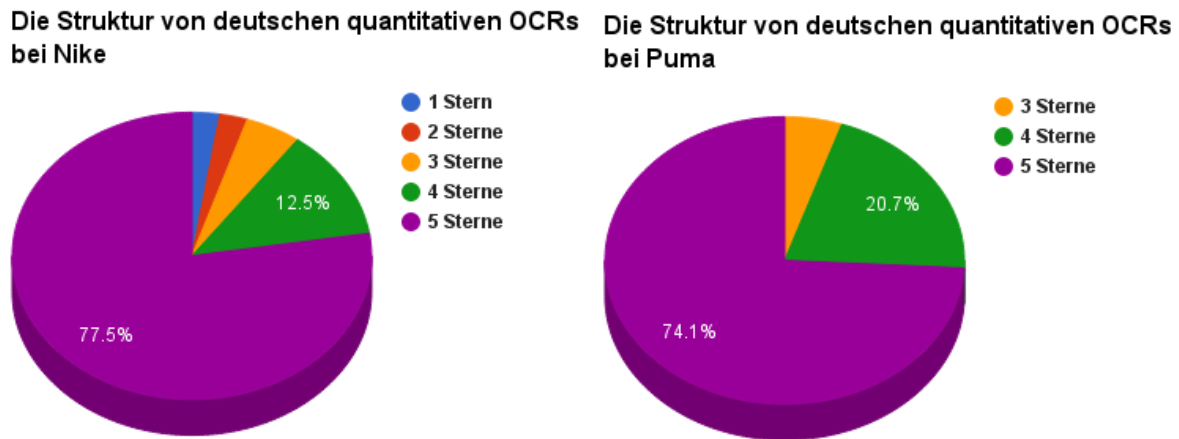


Abbildung 4.14: Die Struktur von deutschen quantitativen OCRs bei Nike und Puma (Quelle: Eigene Darstellung)

geführt. Der Levene-Test bezeichnet in der Statistik einen Signifikanztest, der auf Gleichheit der Varianzen (Homoskedastizität) von zwei oder mehr Grundgesamtheiten (Gruppen) prüft. Der Levene-Test prüft die Nullhypothese darauf, dass alle Gruppenvarianzen gleich sind. (Levene, 1960) Wenn sich der Signifikanzwert des Tests unter einem zuvor bestimmten Niveau befindet, sind die Heterogenität der Varianzen der Stichproben signifikant und die Nullhypothese der Varianzgleichheit kann abgelehnt werden. (Martens, 2003)

p-Wert	Varianzheterogenität in deutschen und chinesischen	
	quantitative OCRs	qualitative OCRs
Adidas männlich	0,04948	0,0006272
Adidas weiblich	0,03005	0,0001619
Nike	$9,003 \times 10^{-06}$	$2,511 \times 10^{-08}$
Puma	0,22	0,0002532

Tabelle 4.9: Die Ergebnisse des Levene-Tests beim jeden Artikel (Signifikanzniveau $\alpha = 0,05$) (Quelle: Eigene Darstellung)

Wie in der Tabelle ersichtlich, sind die p-Werte fast alle kleiner als das Signifikanzniveau α , außer den zwischen quantitativen OCRs von Puma. Der Wert 0,22 ist die einzige Ausnahme. Für diese Außerordentlichkeit kann man nicht die Nullhypothese des Levene-Tests ablehnen, aber für die Anderen zeigt der Levene-Test, dass die Varianzen ungleich sind. Das heißt, dass es die Varianzheterogenität in deutschen und chinesischen quantitativen und qualitativen OCRs bei der Mehrheit der Produkte gibt.

Tabelle 4.10 zeigt die Varianzen der quantitativen und qualitativen OCRs von jedem Artikel. Wie sich bereits aus der Tabelle ansehen lässt, ist jede Varianz von chinesischen qualitativen

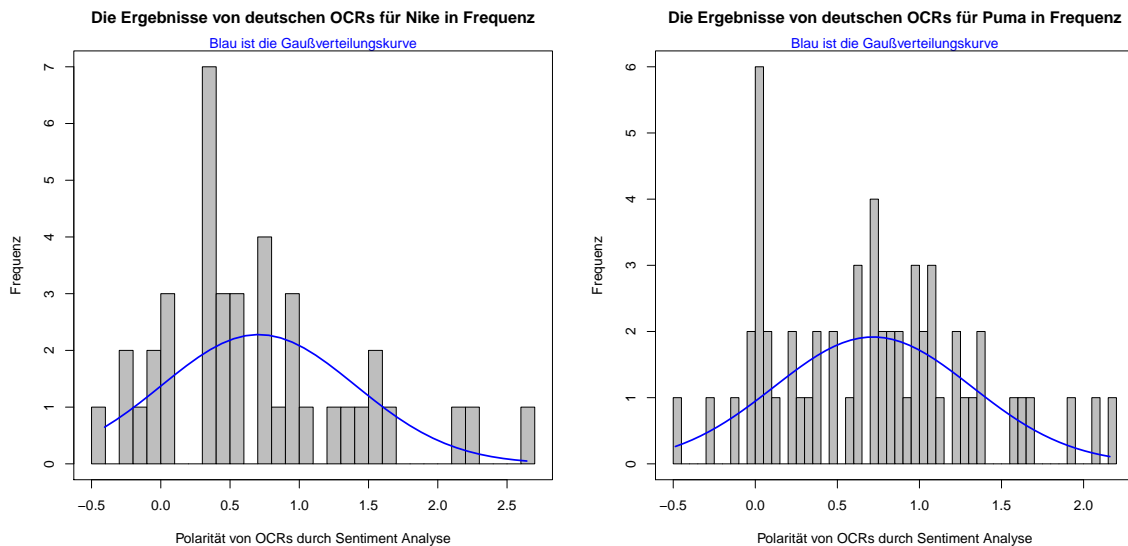


Abbildung 4.15: Die Verteilung der deutschen Valenzen von qualitativen OCRs bei zwei unterschiedlichen Marken (Die blauen Linien sind die entsprechende Normalverteilung.) (Quelle: Eigene Darstellung)

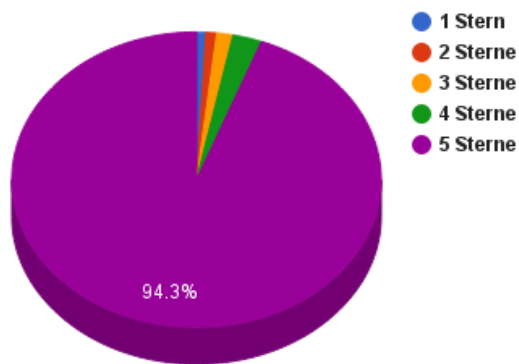
Artikel	Varianz von quantitativen OCRs		Varianz von qualitativen OCRs	
	China	Deutschland	China	Deutschland
Adidas männlich	0,6666667	1,552525	0,0950417	0,4937723
Adidas weiblich	0,5356501	1,25	0,134694	0,3634815
Nike	0,1920619	0,8102564	0,1421846	0,4902091
Puma	0,498598	0,323049	0,1991728	0,3641573

Tabelle 4.10: Die Varianzen der quantitativen und qualitativen OCRs von jedem Artikel (Quelle: Eigene Darstellung)

OCRs über das eigene Untersuchungsobjekt kleiner als die der deutschen, obwohl die Varianzen von deutschen quantitativen OCRs nicht alle größer als die von chinesischen sind. Das heißt: Die Grade der Verschiedenheiten von deutschen und chinesischen quantitativen OCRs sind unbedingt, aber die Valenzen von deutschen qualitativen OCRs schwanken stärker als die von chinesischen. Anders gesagt, es gibt mehr Ungleichheiten in den Gefühlen der deutschen als bei den chinesischen Kunden. Dies erklärt auch, dass die Gefühle der chinesischen Kunden über das Produkt viel dichter in der Nähe des Mittelwertes als die Gefühle von Deutschen sind.

Tabelle 4.10 kann man auch anders erfassen. Weil die beiden Produkte von Adidas Hosen sind, sind die Varianzen von deutschen quantitativen OCRs alle größer als die von den Chinesen. In der Kategorie T-Shirt ist die Situation anders. Die Artikeln von Nike und Puma sind T-Shirts, aber die Varianzen bleiben heterogen. Im Bereich “Varianz von quantitativen OCRs” sind die

Die Struktur von chinesischen quantitativen OCRs



Die Struktur von deutschen quantitativen OCRs

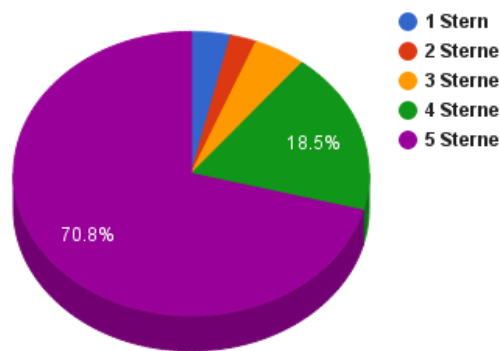


Abbildung 4.16: Die Struktur von den chinesischen und deutschen quantitativen OCRs in Allgemeinen (Quelle: Eigene Darstellung)

deutschen Valenzen von Produkten aus Adidas und Nike größer als die chinesischen, aber die Valenz von Puma ist eine Ausnahme, die schon in Tabelle 4.9 gekennzeichnet wird. Außer der Ausnahme, sind die Varianzen von chinesischen OCRs kleiner als die deutschen. Die Hypothese 3 wird bei jedem Artikel nur teilweise gestützt.

4.5.2 Hypothese 3 in Allgemeinen

In diesem Abschnitt werden die Varianzen in einer allgemeinen Hinsicht verglichen. Damit wird besser erfasst, welches Verhalten die deutschen und chinesischen Kunden im Allgemeinen haben. Die Unterschiede zwischen Marken oder Kategorien werden ausgeglichen.

allgemein	quantitative OCRs	qualitative OCRs
p-Wert	$< 2, 2 \times 10^{-16}$	$< 2, 2 \times 10^{-16}$

Tabelle 4.11: Die Ergebnisse von Levene-Test in Allgemeinen (Signifikanzwert $\alpha = 0,05$) (Quelle: Eigene Darstellung)

In dieser Tabelle kann man durch betrachten der p-Werte von Levene-Test schließen, dass sie alle kleiner als der Signifikanzwert α sind. Die Tabelle 4.11 stützt die Ergebnisse von Tabelle 4.12. Durch den Levene-Test sind die Varianzen zwischen den OCRs von China und Deutschland heterogen.

Varianz	China	Deutschland
von quantitativen OCRs	0,2855922	0,92212
von qualitativen OCRs	0,1518924	0,4284648

Tabelle 4.12: Die Varianzen von den chinesischen und deutschen OCRs (Quelle: Eigene Darstellung)

Tabelle 4.12 zeigt die Ergebnisse. Dieser Tabelle kann man auch entnehmen, dass die

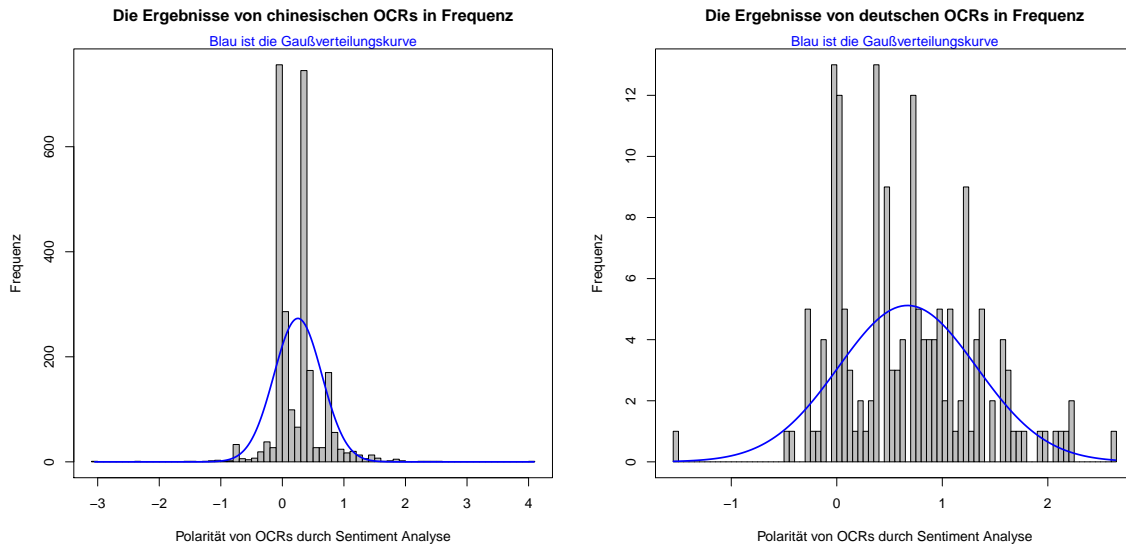


Abbildung 4.17: Die Verteilungen der chinesischen und deutschen Valenzen von den qualitativen OCRs in Allgemeinen (Die blauen Linien sind die entsprechende Normalverteilung.) (Quelle: Eigene Darstellung)

deutschen Varianzen größer als die chinesischen sind. Die Meinungen sowie die Gefühle der chinesischen Kunden liegen meistens in der Mitte während die Meinungen und Gefühle der deutschen Kunden hin und her schwanken. Die Varianzen von qualitativen OCRs sind kleiner als die von quantitativen OCRs für die Kunden aus beiden Ländern. Das bedeutet, dass die quantitativen OCRs viel zufallsbedingter als die qualitativen OCRs sind. Mit anderen Worten sind die quantitativen OCRs nicht so glaubwürdig, aber die Gefühle von Kunden werden nicht so viel von externen Faktoren beeinflusst. Durch diese Ergebnisse wird Hypothese 3 in allgemeiner Weise gestützt.

4.6 Hypothese 4

Obwohl es viele Forschungen über die Online Consumer Reviews schon gab, gibt es noch niemanden, der über den Zusammenhang zwischen quantitativen und qualitativen OCRs geforscht hat. Man dachte, dass es bestimmte Zusammenhänge zwischen ihnen gibt, aber ist es wirklich so? Wenn ja, ist der Zusammenhang linear oder unlinear? Gibt es kulturelle Einflüsse auf den Zusammenhang? Diese Fragen sind noch nie beantwortet geworden. In diesem Abschnitt werden die Antworten möglicherweise herausgefunden.

4.6.1 Hypothese 4 bei jedem Artikel

Wie bereits erwähnt weiß man noch nichts über den Zusammenhang zwischen den quantitativen und qualitativen OCRs. Die erste Aufgabe von Hypothese 4 ist, zu bestimmen, ob der Zusammenhang tatsächlich existiert. Deshalb wird erst der Test von jedem Artikel durchgeführt.

Als ersten Schritt ist es anzunehmen, dass der Zusammenhang linear ist. Von daher wird erst der Korrelationstest von Pearson zwischen den Valenzen von quantitativen und qualitativen OCRs durchgeführt. Der Test beginnt mit einer Nullhypothese, die setzt voraus, dass es keinen linearen Zusammenhang zwischen den Valenzen der quantitativen und qualitativen OCRs gibt. Ist der p-Wert kleiner als das Signifikanzniveau α , wird die Nullhypothese abgelehnt, also: Es gibt einen linearen Zusammenhang. (Pearson, 1895)

p-Wert	China	Deutschland
Adidas männlich	0,5251	0,001808
Adidas weiblich	0,0002521	0,07921
Nike	$2,911 \times 10^{-05}$	0,02091
Puma	$5,148 \times 10^{-07}$	0,9997

Tabelle 4.13: Die Ergebnisse von Pearsons Korrelationstest bei jedem Artikel (Signifikanzniveau $\alpha = 0,05$) (Quelle: Eigene Darstellung)

Die Ergebnistabelle 4.13 gibt die p-Werte von jedem Artikel an. Wie sich bereits aus der Tabelle ersehen lässt, sind die p-Werte nicht alle kleiner als das Signifikanzniveau α . Für die chinesischen OCRs, kann man nicht die Nullhypothese bei der männlichen Hose von Adidas ablehnen, aber bei den Anderen wird die Nullhypothese abgelehnt. Auf der deutschen Seite, ist die Situation anders. Bei der weiblichen Hose von Adidas und dem T-Shirt von Puma, wird die Nullhypothese nicht abgelehnt, während man bei den anderen zwei Artikeln die Nullhypothese ablehnen kann.

Dies bedeutet, dass es einen linearen Zusammenhang zwischen den chinesischen Valenzen von quantitativen und qualitativen OCRs gibt, außer der, deren OCRs für männlichen Hosen von Adidas sind. Für die Valenzen von den deutschen OCRs, existiert der lineare Zusammenhang bei der männlichen Hose von Adidas und dem T-Shirt von Nike.

Bei den Artikeln, deren p-Werte größer als 0,05 sind, kann man nicht direkt sagen, dass es keinen linearen Zusammenhang gibt. Ebenso, stellt ein kleiner p-Wert nicht einen großen Zusammenhang dar. Die tatsächlichen Ergebnisse von Pearson's Korrelationskoeffizient ρ_p werden durch Tabelle 4.14 dargestellt.

Aus der Tabelle ist ersichtlich, dass die Korrelationskoeffizienten ρ_p fast alle positiv sind, außer bei Puma Deutschland. Der Korrelationskoeffizient von den Valenzen der deutschen

ρ_p	China	Deutschland
Adidas männlich	0,1333677	0,4523909
Adidas weiblich	0,2337115	0,3576531
Nike	0,09517316	0,3640926
Puma	0,2269939	$-5,7147 \times 10^{-05}$

Tabelle 4.14: Die Pearsons Korrelationskoeffizienten ρ_p zwischen den chinesischen und deutschen Valenzen der quantitativen und qualitativen OCRs (Quelle: Eigene Darstellung)

OCRs ist eine Ausnahme, wie es im Abschnitt 4.5.1 schon einmal passiert ist. Diese Ausnahme zeigt auch, dass die deutschen quantitativen OCRs von Puma kaum einen Zusammenhang mit den qualitativen OCRs haben. Wenn man sich das ohne der Außerordentlichkeit ansieht, kann man erfassen, dass die anderen nicht nur positiv sind, die deutschen Korrelationskoeffizienten sind auch größer als die chinesischen.

Diese Ergebnisse bedeuten, dass der lineare Zusammenhang zwischen den meisten quantitativen und qualitativen OCRs existiert. Und sie sind positiv abhängig. In den meisten Situationen sind die linearen Zusammenhänge von chinesischen OCRs kleiner als die von deutschen OCRs. Auf diese Weise wird Hypothese 4 teilweise gestützt.

Aber vielleicht ist der Zusammenhang nicht linear. Um den tatsächlichen Zusammenhang zwischen den quantitativen und qualitativen OCRs herauszufinden, wird der Spearmans Rangkorrelation-Test durchgeführt.

p-Wert	China	Deutschland
Adidas männlich	0,9555	0,0002549
Adidas weiblich	0,001942	0,06097
Nike	$4,813 \times 10^{-07}$	0,002007
Puma	$1,173 \times 10^{-09}$	0,5814

Tabelle 4.15: Die Ergebnisse von Spearmans Rangkorrelationstest bei jedem Artikel (Signifikanzniveau $\alpha = 0,05$) (Quelle: Eigene Darstellung)

Der Spearmans Rangkorrelationskoeffizient ist ein parameterfreies Maß für Korrelationen. Das bedeutet, dass er misst, wie gut eine beliebige monotone Funktion den Zusammenhang zwischen zwei Variablen (In dieser Arbeit sind die quantitativen und qualitativen Valenzen) beschreiben kann, ohne irgendwelche der Variablen der Wahrscheinlichkeitsverteilung anzunehmen. (Meyer-Bahlburg, 1969)

Die Ergebnisse von Spearmans Rangkorrelationstest sind nicht anders als die Ergebnisse von Pearsons Korrelationstest. Der Tabelle 4.15 ist zu entnehmen, dass die drei p-Werte, die in gleichen Positionen in Tabelle 4.13 sind, größer als das Signifikanzniveau 0,05 sind. Und

Tabelle 4.16 zeigt die Spearmans Rangkorrelationskoeffizienten ρ bei jedem Artikel.

ρ	China	Deutschland
Adidas männlich	0,01176353	0,5195352
Adidas weiblich	0,1986592	0,3799971
Nike	0,1144833	0,473936
Puma	0,2733825	-0,07389644

Tabelle 4.16: Die Spearmans Rangkorrelationskoeffizienten ρ zwischen den chinesischen und deutschen Valenzen der quantitativen und qualitativen OCRs (Quelle: Eigene Darstellung)

Die Ergebnisse in Tabelle 4.16 sind sehr ähnlich wie die in Tabelle 4.14. Außer der Ausnahme von Puma Deutschland, sind die deutschen Spearmans Rangkorrelationskoeffizienten ρ größer als die chinesischen, und alle ρ sind größer als 0. Diese Ergebnisse zeigen, dass es einen positiven Zusammenhang zwischen den meisten quantitativen und qualitativen OCRs in beiden Ländern gibt. Und der Zusammenhang bei den deutschen Kunden ist mehrheitlich größer als der bei den Chinesen. Dadurch wird die Hypothese 4 in den meisten Fällen gestützt.

4.6.2 Hypothese 4 in Allgemeinen

Abschnitt 4.6.1 gibt schon die Übersicht über jeden Artikel an, und in diesem Abschnitt kann man eine allgemeine Einsicht in den Zusammenhang haben. Die Unterschiede von den jeweiligen Artikeln werden ausgeglichen, um besser das Kennzeichen von China oder Deutschland zu verstehen.

Abbildung 4.18 zeigt die Verteilung der quantitativen und qualitativen OCRs von China und Deutschland. Jeder Farbblock zeigt eine qualitative Valenz. Der dunkelrote Block repräsentiert Valenz 0, und der Pinke ist Valenz 0,3716, die in dem Lexion “gut” besteht. Der Abbildung ist zu entnehmen, dass viele chinesische Kunden nichts über ihre Attitüde oder nur ein “gut” geschrieben haben, obwohl sie schon fünf Sterne vergeben haben. Und die deutschen Kunden geben ihre qualitativen Bewertung viel individueller als die Chinesischen. Anders gesagt, die Korrelation zwischen den quantitativen und qualitativen OCRs sind in China und Deutschland unterschiedlich. Um diese Unterschiede zu erkennen, werden der Pearsons Korrelationstest und der Spearmans Rangkorrelationstest durchgeführt.

Tabelle 4.17 zeigt die Ergebnisse im Signifikanzniveau $\alpha = 0,05$ mit der Nullhypothese, die besagt, dass es keinen linearen Zusammenhang gibt. Der Tabelle ist zu entnehmen, dass alle p-Werte kleiner als das Signifikanzniveau sind. Das bedeutet, dass die Nullhypothese abgelehnt wird. Also, es gibt einen linearen Zusammenhang zwischen quantitativen und

Pearson	China	Deutschland
p-Wert	$2,665 \times 10^{-15}$	$1,629 \times 10^{-05}$
ρ_p	0,1522703	0,3258736

Tabelle 4.17: Die Ergebnisse von Pearsons Korrelationstest zwischen quantitativen und qualitativen OCRs in China und Deutschland in Allgemeinen (Signifikanzniveau $\alpha = 0,05$) (Quelle: Eigene Darstellung)

qualitativen OCRs in den beiden Ländern. ρ_p sind in beiden Ländern positiv. Das heißt, dass der Zusammenhang in Allgemeinen positiv ist. Aber die Korrelationskoeffizienten von den beiden Ländern sind klein, auch wenn der von Deutschland größer ist. Das sind die Ergebnisse über den linearen Zusammenhang. Die Tabelle 4.18 zeigt die Ergebnisse, wenn der Zusammenhang nicht linear ist.

Spearman	China	Deutschland
p-Wert	$< 2,2 \times 10^{-16}$	$5,967 \times 10^{-05}$
ρ_{rho}	0,1632395	0,3045776

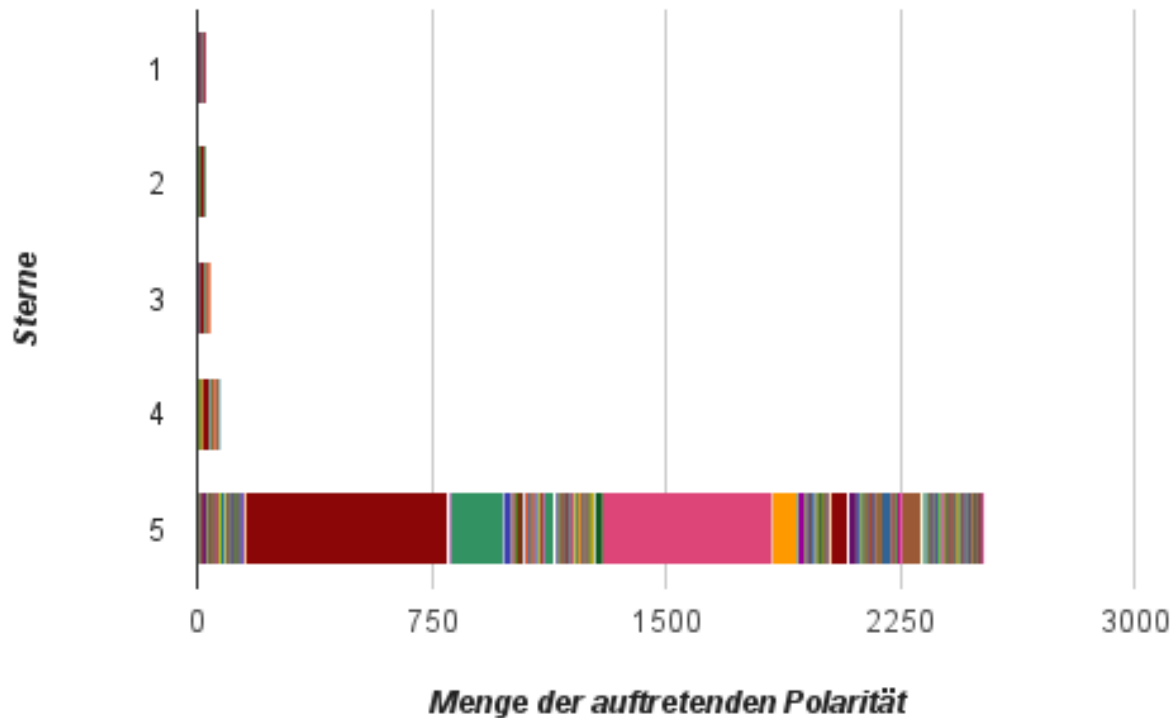
Tabelle 4.18: Die Ergebnisse von Spearmans Rangkorrelationstest zwischen quantitativen und qualitativen OCRs in China und Deutschland in Allgemeinen (Signifikanzniveau $\alpha = 0,05$) (Quelle: Eigene Darstellung)

Die Ergebnisse in Tabelle 4.18 sind nichts anderes als die in Tabelle 4.17. Dadurch wird bestimmt: obwohl die quantitativen und qualitativen OCRs einen positiven Zusammenhang haben, aber der Zusammenhang ist nicht stark. Anders gesagt, es ist nicht genug, wenn man nur die quantitativen OCRs studiert, weil der schriftlicher Teil mehr das wirkliche Gefühl der Kunden über das Produkt ausdrücken kann.

Diese Ähnlichkeit zeigt auch, dass dieser Zusammenhang linear sein könnte, obwohl er nicht so stark ist. Der Unterschied zwischen den Korrelationskoeffizienten oder Rangkorrelationskoeffizienten bedeutet auch, dass die kulturellen Einflüsse unterschiedlich auf den quantitativen und qualitativen OCRs sind.

In allgemeiner Hinsicht ist der Zusammenhang zwischen den deutschen quantitativen und qualitativen OCRs stärker. Besonders bei China, sind die meistens quantitativen Bewertung fünf Sterne (siehe Abbildung 4.16), aber der Korrelationskoeffizient zwischen quantitativen Bewertung und schriftlichen Meinung von Kunden ist kleiner als 0,2. Das zeigt dass die Sterne und die Polarität keinen großen Zusammenhang haben, obwohl sie von einander positiv abhängig sind. Anders gesagt, die quantitativen Bewertungen (besonders von chinesischen Kunden) sind nicht so glaubwürdig, wie viele Wirtschaftler geglaubt haben. Im Vergleich mit chinesischen Bewertungen, sind die deutschen quantitativen OCRs glaubwürdiger, weil der Rangkorrelationskoeffizient größer ist. Zusammenfassend wird die Hypothese 4 in allgemeiner Weise gestützt.

Unterschiedlichen Valenzen von chinesischen quantitativen und qualitativen OCRs



Unterschiedlichen Valenzen von deutschen quantitativen und qualitativen OCRs

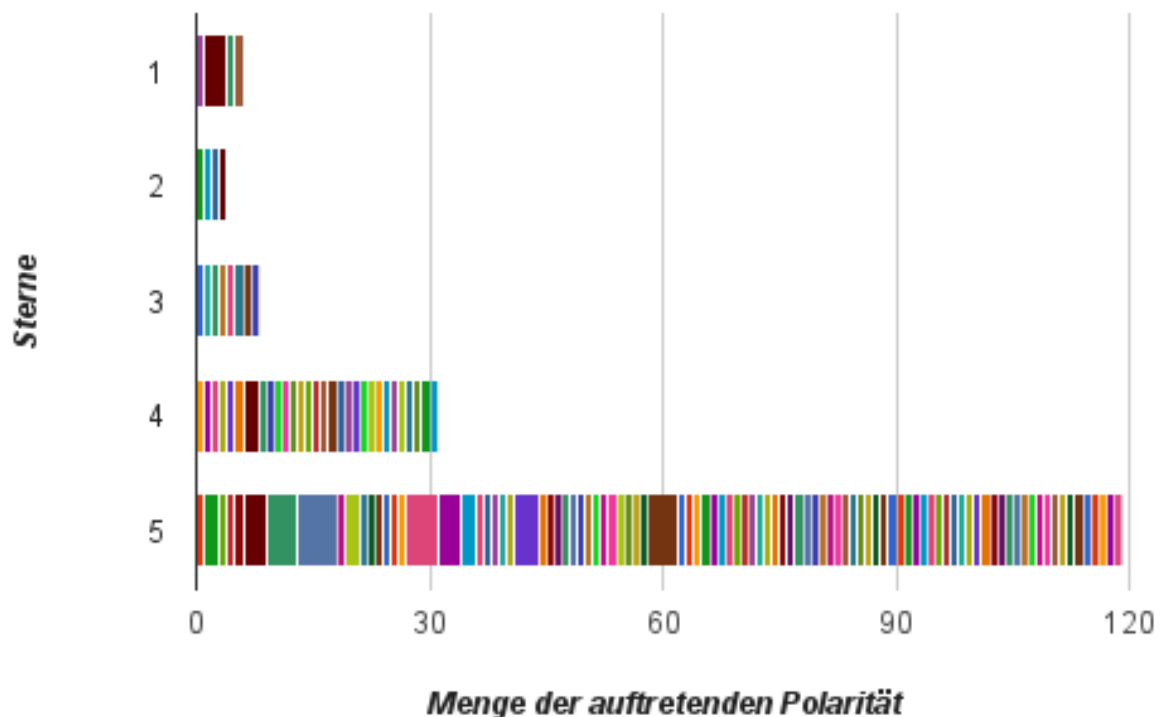


Abbildung 4.18: Die chinesischen und deutschen Valenzen von den quantitativen und qualitativen OCRs in Allgemeinen (Quelle: Eigene Darstellung)

5 Diskussion

Es gibt schon viele Forschungen über die Online Consumer Reviews, und der Begriff, die Motivationen und die Auswirkungen werden durch die Studien bereits reichlich studiert. Aber es gibt nicht viele Ergebnisse über die kulturellen Einflüsse. Gibt es Unterschied in den OCRs aufgrund der kulturellen Einflüsse? Wenn ja, welche Unterschiede, und warum? Diese Fragen sind noch zu beantworten. Besonders über die qualitativen OCRs, also dem schriftliche Teil, gibt es noch wenige Studien. In dieser Arbeit werden die OCRs nach Sridhar & Srinivasan (2012) in den quantitativen und qualitativen Teil getrennt und jedem Teil werden drei Attribute gegeben: Volumen, Valenz, und Varianz. Außerdem haben die beiden Teile ein gemeinsames Attribut: Der Korrelationskoeffizient. Durch diese digitalen Attribute kann man die statistischen Maßnahmen benutzen, um die allgemeinen Unterschiede hinter den Rohdaten herauszufinden.

Für diese Attribute werden vier Hypothesen gebildet, um die Unterschiede zwischen China und Deutschland zu entdecken. Hypothese 1, 2 und 3 beziehen sich auf das Volumen, die Valenz, und die Varianz der OCRs sowohl in dem quantitativen Aspekt als auch in dem qualitativen Aspekt. Hypothese 4 bezieht sich auf den Zusammenhang der quantitativen und qualitativen OCRs. Diese vier Hypothesen versuchen, durch die Attribute die statistischen Unterschiede herauszufinden. Tabelle 5.1 zeigt die zusammenfassenden Ergebnisse.

Hypothese	Ergebnisse	Quantitativer Aspekt	Qualitativer Aspekt
1	Volumen	China >Deutschland	China <Deutschland
2	Valenz	China >Deutschland	China <Deutschland
3	Varianz	China <Deutschland*	China <Deutschland
4	Korrelationskoeffizient	China <Deutschland*	

Tabelle 5.1: Die zusammenfassenden Ergebnisse von Hypothesen (*: Ausnahme ist Puma) (Quelle: Eigene Darstellung)

In den folgenden Teilen wird versucht, die kulturellen Gründen für diese Ergebnisse möglicherweise nach den theoretischen Grundlagen herauszufinden. Des weiteren wird diskutiert, welche Auswirkungen es auf den Kunden bzw. den Produzenten gibt.

5.1 Qualitativ gegenüber Quantitativ

Laut Lam *et al.* (2009) sind die Menschen zurückhaltend bei dem Informationsaustausch, wenn sie in einer Kultur mit großer Machtdistanz leben, und nach Dawar *et al.* (1996), Lam *et al.* (2009), Liu *et al.* (2001), Money *et al.* (1998) können die Menschen mit einer hohen Unsicherheitsvermeidung ein höheres Niveau des Meinungs austauschs sowie der Meinungssucht. In Abbildung 2.9 sieht man schon, dass es in China eine sehr große Machtdistanz gibt, und die deutsche Unsicherheitsvermeidung ist höher als in China. Deshalb wird in Hypothese 1 erwartet, dass das chinesische Volumen kleiner als das von den Deutschen ist, sowohl bei den quantitativen als auch bei den qualitativen OCRs. Diese Hypothese meint, dass die deutschen Kunden länger und mehr schreiben würden während die chinesischen Kunden kürzer und weniger schreiben würden.

Aber die Wirklichkeit sieht anders aus. Es stimmt, dass die Deutschen länger schreiben und die Chinesen kürzer schreiben. Aber die Chinesen schreiben viel lieber als die Deutschen. Die Menge der chinesischen OCRs ist das 15fache der deutschen OCRs bei den Untersuchungsobjekten in dieser Arbeit. In dieser Situation wird zusammengefasst, dass die Deutschen qualitativ schreiben wollen aber die chinesischen OCRs sind quantitativ.

Es ist klar, dass wer zurückhaltend beim Informationsaustausch ist, kürzer schreiben wird. Diese Zurückbehaltung begründet sich in der großen Machtdistanz in der chinesischen Kultur nach Lam *et al.* (2009). Die Deutschen, die mit kleiner Machtdistanz in ihrer Kultur leben, sind offener als die Chinesen und vermeiden die Unsicherheit lieber (Dawar *et al.* , 1996, Lam *et al.* , 2009, Liu *et al.* , 2001, Money *et al.* , 1998), daher wollen sie mehr Information austauschen, deswegen schreiben sie länger.

Aber die Frage, warum das chinesische quantitative Volumen so groß ist, steht noch in dem theoretischen Bereich offen. Es sollte mindestens eine extra Motivation der Chinesen für das Schreiben einer Review im Vergleich zu der Deutschen geben. Es wird vorgeschlagen, dass diese Motivation aus der kollektivistischen Kultur kommt. Bei der kollektivistischen Kultur ist die Einstellung “alle machen, dann mache ich auch” normal. Es entsteht die Motivation, eine Review für das Produkt zu schreiben, wenn das Produkt schon viele Reviews hat. Aber die Motivation ist nicht stark genug für eine lange Review. Die Chinesen möchten zeigen, dass sie noch in der Gruppe sind, in der alle Reviews schreiben. Diese in-Gruppenmitgliedschaften ist für sie wichtig. Aber die Deutschen, als die Mitglieder einer individualistischen Kultur, sind unabhängig von anderen, und betrachten die Getrenntheit und Entfernung von in-Gruppen. (Singelis, 1994)

Für die anderen chinesischen Konsumenten, sind die chinesischen OCRs schlechter als die Deutschen, weil die Kunden, die OCRs geschrieben haben, wenige Information gegeben haben, obwohl die Menge groß ist. Wenn man nur einfach “gut” oder “Okay” geschrieben

hat, wissen die anderen Leute nicht, worauf sich die Review bezieht. Deshalb müssen sie noch mehr lesen, um die richtigen und wichtigen Informationen zu finden. Die kurzen OCRs werden als “Noise Text” von den anderen Kunden oder potentiellen Kunden genannt. In dieser Hinsicht sollte der Kunde auch wissen, dass es keine Hilfe ist, wenn er oder sie nur eine kurze Review geschrieben hat. Der Kunde sollte vermeiden, eine kurze Review zu schreiben, sondern möglicherweise die Attribute des Produkts beschreiben, damit er oder sie den anderen helfen kann.

Andererseits gibt es diese Situation für die deutschen OCRs kaum. Aber bei den deutschen OCRs gibt es noch andere Probleme. Zu lange Reviews sind auch sinnlos, weil die anderen Menschen keine Zeit haben, um die lange Review durchzulesen. Deshalb könnte die lange Review nicht so viel helfen. Die Review sollte nicht zu kurz oder zu lang sein, aber die Attribute des Produkts sollten als Stichpunkte beschrieben werden. Zum Beispiel, ist es eine gute Review, wenn man über die Vorteile und Nachteile der Kleidung in den Attributen “Ästhetik”, “Performance” und “äußere Attribute” geschrieben hat.

Es war auch eine Herausforderung, die die Unternehmen oder die Betreiber der Plattformen meistern müssen, die “Noise Text” zu vermeiden oder wenigstens sich nicht an der ersten Stelle stehen zu lassen. Beim Vermeiden kann man einfach die minimale Anzahl der Wörter beschränken, aber wie gesagt, die Motivation ist nicht so stark, damit kann man einfach das Schreiben der Review aufgeben. Deshalb ordnen die Unternehmen gerne die OCRs nochmal, damit sie nicht in zeitlicher Reihenfolge, sondern in nützlicher Reihenfolge sind.

Die chinesische Plattform “Tmall” ordnet die OCRs so, dass man die längere OCRs mit Fotos zuerst lesen kann. Die deutsche Plattform “Amazon” ordnet die OCRs durch “Hilfreich”. Wenn die anderen Leute denken, dass dieser Kommentar für sie hilfreich ist, drücken sie “Ja” und die anderen sehen, dass dieser Kommentar einem Mensch hilft. Der Kommentar kann vorne stehen, damit mehrere Menschen ihn lesen können.

5.2 Positiv gegenüber Neutral

Nach Liu *et al.* (2001) geben die Personen aus geringerer Unsicherheitsvermeidung häufiger negative Kommentare, als die Personen in Kulturen mit höherer Unsicherheitsvermeidung, wenn sie schlecht bedient werden. Und laut Donthu & Yoo (1998), geben die Menschen in Kulturen mit hoher Unsicherheitsvermeidung im Allgemeinen positivere Kommentare, als die mit niedriger Unsicherheitsvermeidung. Und Wie aus Abbildung 2.9 ersichtlich ist, sind die Unsicherheitsvermeidungen der Deutschen relativ hoch im Vergleich zu den Chinesen. Deshalb wird in Hypothese 2 vorgeschlagen, dass die deutschen Valenzen größer als die chinesischen Valenzen der OCRs in dem quantitativen und auch in dem qualitativen Aspekt

sein sollten. Also, die deutschen OCRs sollten positiver als die chinesischen OCRs sein.

Aber es stimmt nur teilweise in der Praxis. Die deutschen qualitativen OCRs sind tatsächlich positiver als die chinesischen durch die Sentiment Analyse, aber in dem quantitativen Aspekt präsentieren sich die chinesischen OCRs positiver. Im Durchschnitt geben die Deutschen für die Untersuchungsobjekte nur 4,51 Sterne im Vergleich zu 4,88 Sterne von den Chinesen (Maximal 5 Sterne). Die durchschnittliche Valenz von den deutschen qualitativen OCRs ist 0,6702, schon relativ hoch (beim einzelnen Wort bis 1, und in dem Lexikon von Remus *et al.* (2010) ist 0,6702 zwischen “zuvorkommend, 0,6669” und “romantisch, 0,6965”). Die chinesische durchschnittliche Valenz ist 0,2525 und es ist im Lexikon von Remus *et al.* (2010) zwischen “fröhlich, 0,2501” und “gefallen, 0,2578”. Daher sind die deutschen OCRs positiv in dem qualitativen Aspekt und die chinesischen qualitativen OCRs sind neutral.

Die kulturelle Begründung der positiveren deutschen qualitativen OCRs ist nach den Forschungen von Donthu & Yoo (1998) klar, und laut Liu *et al.* (2001) sollen die chinesischen OCRs negativer als die deutschen sein, aber dieses Phänomen, dass die chinesischen Kunden viel lieber 5 Sterne geben (siehe in Abbildung 4.16), wird noch nicht begründet. Möglicherweise gab es einen wirtschaftliche Anreiz für die chinesischen Kunden, die 5 Sterne zu geben. Wenn man in den kulturellen Dimensionen denkt, wird vermutet, dass die kollektivistische Kultur der Chinesen das Phänomen erzeugt. Nach Hofstede (1998, p. 65), werden die individuellen Entscheidungen (zum Beispiel: wie viele Sterne gebe ich) in kollektivistischer Kultur im Konsens mit der Gruppe gemacht und es gibt keine rein individuellen Entscheidungen zu treffen. Aber im Gegensatz dazu denken die Menschen in individualistischen Kulturen eher selbst als autonom zu denken und stellen die individuellen Interessen an erster Stelle (Shweder, 1990).

Das heißt, dass die chinesische einzelne Valenz von den quantitativen Reviews (also Sterne) von den bereits existierten OCRs groß beeinflusst wird, während die einzelne deutsche Valenz individuell ist. Wenn es schon viele “5 Sterne” OCRs gab, könnte es sein, dass die chinesischen Kunden “5 Sterne” geben wollen, obwohl sie nicht zufrieden sind. Nach der Einsicht in die Rohdaten, wird das Beispiel gefunden: “老公说质量不怎么样。” Die Kundin meint, dass ihr Mann sagt, dass die Qualität des T-Shirts nicht so gut ist, trotzdem hat sie “5 Sterne” gegeben. Es ist nur ein Beispiel, aber es gibt noch einige ähnlichen Situationen in den chinesischen Daten.

Die chinesischen Kunden sollten vorsichtiger sein, als die Deutschen, weil die anderen chinesischen Kunden “5 Sterne” gegeben haben könnten, obwohl das Produkt nicht gut ist. Es ist nicht genug nur die Sterne zu lesen, man muss auch, besonders bei den chinesischen Kunden, den Text dazu lesen.

Für die Unternehmen, ist es wichtig zu wissen, dass es Probleme geben könnte, obwohl

die quantitativen Valenzen gut scheinen. Sie sollten auch die Texte durchlesen, und schnell reagieren, um die Probleme zu beseitigen.

Luo *et al.* (2014) haben herausgefunden, dass durch die stärkere Wirkung der zweiseitigen Informationen (im Vergleich zu einseitigen Informationen) die Wahrnehmung von Glaubwürdigkeit der Informationen der eWOM Leser gestärkt wird, wenn die Leser aus individualistischen Kulturen sind, verglichen mit denen, die aus kollektivistisch kulturellen Orientierung sind. Deshalb sollte die Reihenfolge der OCRs in Deutschland und China unterschiedlich sein. Die Betreiber von Amazon sollten nicht nur die positiven Bewertungen sondern auch einige negativen Bewertungen an die ersten Positionen stellen, damit für die deutschen Kunden die Glaubwürdigkeit der OCRs höher ist. Die Betreiber von Tmall brauchen das vielleicht nicht so zu machen.

5.3 Dezentralisiert gegenüber Zentralisiert

Laut Hofstede (1991) ist die Ungleichheit bei großer Machtdistanz erwünscht und wird erwartet, deswegen sollten die OCRs dezentralisiert sein. Aber nach der Forschung der Luo *et al.* sollten die OCRs konsistenter und dichter beim Kollektivismus sein. Die Chinesen leben in der kollektivistischen Kultur mit großer Machtdistanz. Die Deutschen leben in individueller Kultur mit geringer Machtdistanz. Weil die vorherigen Forscher die Dimension “Individualismus gegenüber Kollektivismus” als eine Tiefenstruktur denken (Greenfield, 2000, Sia *et al.*, 2009, Triandis, 2001), wird in Hypothese 3 vorgeschlagen, dass die chinesischen OCRs zentralisiert und die deutschen OCRs dezentralisiert sein sollten.

Die Tatsache stützt die Hypothese. In dem quantitativen und auch qualitativen Aspekt, sind die Varianzen von den chinesischen OCRs kleiner als die von den deutschen im Allgemeinen. Für das jeweilige Produkt, ist die Situation fast gleich, aber es gibt eine Ausnahme bei den quantitativen OCRs von Puma. Der Grund für diese Außerordentlichkeit ist noch nicht klar. Vielleicht ist es aufgrund der Machtdistanz nach Hofstede (1991), aber dafür ist nicht der Schatten eines Beweises zu erbringen. Durch diese Ergebnisse, ist es auch ersichtlich, dass die Dimension “Individualismus gegenüber Kollektivismus” eine Tiefenstruktur und wichtigste Dimension ist, wie die Forscher schon vorher gemacht haben.

Die chinesischen Kunden sollten vor dem Lesen der OCRs wissen, dass sie viel dichter sein werden. Damit können sich die Kunden nicht einfach von den guten OCRs bei der Kaufentscheidungen beeinflussen lassen. Man sollte mehr lesen oder andere Benutzererfahrungen suchen um die Schwächen und Stärken des Produkts herauszufinden und damit die Informationen eher umfassend sind. Die deutschen Kunden sollten auch wissen, dass die OCRs individuell und dezentralisiert sind. Das heißt, dass die OCRs nicht nur positiv sein

können. Individuelle Erfahrung oder Wahrnehmung könnten nicht selbst anpassend sein, wie zum Beispiel das ein Kunde schreibt “Die Farbe mag ich nicht” und deswegen nur einen Stern gibt.

Die Produzenten sollten auch diese Verteilung erkennen. In China sind die OCRs zentralisiert und dicht. Die Unternehmen versuchen, das Zentrum der OCRs möglicherweise zur positiven Seite zu entwickeln. Die wichtigen Schwächen werden durch die OCRs häufig gezeigt und diese Probleme sollten die Unternehmen zuerst lösen. In Deutschland sind die OCRs dezentralisiert, deswegen müssen die Unternehmen sie selbst zusammenfassen und die Stärken und Schwächen kennen, um bessere Produkte bieten zu können.

Wenn die Betreiber der Plattformen die Technologien anbieten könnten, den Kunden fokussierende Aspekte oder die Attitüden der Kunden aus den OCRs automatisch zusammenzufassen, wäre es besser. Damit die Kunden und Unternehmen nicht so viel lesen müssen und einfach erkennen worin die Stärken und Schwächen des Produkts liegen. Tmall aus China hat diese Zusammenfassung schon realisiert, aber bei Amazon gibt es diese noch nicht.

5.4 Abhängig gegenüber Unabhängig

Obwohl es schon viele Forschungen über die OCRs gibt, bleibt der Zusammenhang zwischen den quantitativen und qualitativen OCRs heutzutage nicht in der Sicht der Wissenschaftler. Gibt es bestimmten Zusammenhänge? Wenn ja, ist der Zusammenhang positiv oder negativ? Gibt es kulturelle Unterschiede in dem Zusammenhang? Diese Fragen werden noch nicht beantwortet. Hypothese 4 schlägt vor, dass es den Zusammenhang gibt, und der deutsche Zusammenhang zwischen quantitativen und qualitativen OCRs ist stärker als der von den chinesischen OCRs. Dieser Zusammenhang wird durch den Korrelationskoeffizienten und Rangkorrelationskoeffizienten gemessen.

Durch die Tests wird die Hypothese 4 in den meisten Fällen gestützt. Es gibt positive Zusammenhänge zwischen den quantitativen und qualitativen OCRs und der Zusammenhang von den deutschen OCRs ist stärker als der von den chinesischen OCRs. Diese Ergebnisse werden von den Ergebnissen der Hypothese 2 gestützt, weil die chinesische quantitative Valenz größer als die deutsche ist. Jedoch ist die deutsche Valenz in dem qualitativen Aspekt größer. Wie es in Abschnitt 5.2 durch das Beispiel schon gezeigt wird, sind die chinesischen quantitativen OCRs manchmal unabhängig von den qualitativen OCRs, aber die Deutschen geben die Sterne in Abhängigkeit davon, was sie schreiben wollen.

Es gibt in den Ergebnissen auch eine Ausnahme von Puma. Der deutsche Zusammenhang ist nicht mehr positiv sondern negativ. Gemäß des Korrelationskoeffizient ist dieser Zusammenhang fast Null. Der Grund der Außerordentlichkeit ist noch nicht bekannt.

Die Frage, warum die Chinesen ohne oder mit kleinerer Abhängigkeit Sterne geben und OCRs schreiben, ist noch offen. Vielleicht gibt es wirtschaftliche Anreize für die Kunden, die “5 Sterne” zu vergeben. In dem kulturellen Aspekt, werden die Sterne von chinesischen Kunden von den bereits existierten OCRs größer beeinflusst, als die Deutschen, wegen des Kollektivismus.

Die Kunden schreiben die qualitativen OCRs (Text), um die wirkliche Situation zu beschreiben und das eigene Gefühl auszudrücken. Je mehr man Text schreibt, desto mehr Informationen werden über die Attribute des Produkts beschrieben. Zum Beispiel: Die allgemeinen Attributkategorien für Kleidung sind Ästhetik, Performance, und äußere Attribute. Die qualitativen OCRs werden meistens über diese drei Attributkategorien geschrieben. Wie sich bereits aus Abbildung 4.2 ersehen lässt, schreiben die deutschen Kunden mehr über die Performance (“gross”, “angenehm”, “passt”, “trage” in Top 10 Wörtern) und die chinesischen Kunden schreiben mehr über die Performance (“tragbaren” und “gross”) und die äußere Attribute (“schnell”, “echt” und “logistics” in Top 10 Wörtern). Diese Informationen sind individuell und können nicht einfach von anderen Leuten beeinflusst werden. In diesem Sinne sind die qualitativen OCRs glaubwürdiger als die quantitativen OCRs.

Diese Ergebnisse sind wichtig für die Kunden und die Unternehmen. Die chinesischen Kunden sollten erkennen, dass die Sterne nicht so glaubwürdig sind, weil sie einfach beeinflusst werden könnten. Mehr zu lesen und die Attribute über die Kleidung zusammenzufassen ist besser als an die Sterne zu glauben, weil der Zusammenhang nicht stark ist. Für die Deutschen ist das Problem nicht so groß wegen des stärkeren Zusammenhang, aber für die Sicherheit ist es auch besser, die qualitativen Teile der OCRs durchzulesen.

Die Produzenten sollten auch nicht durch die Mehrheit der “5 Sterne” verwirrt sein. Sie sollten die qualitativen OCRs durchlesen und auf die Probleme schnell reagieren, damit sie bessere Produkte und Service anbieten können. Und wie vorher schon gesagt, wird es besser sein, dass die Betreiber der Plattform die Webseite so organisieren, dass die von Kunden fokussierenden Attribute automatisch zusammengefasst sind.

6 Zusammenfassung und Ausblick

6.1 Zusammenfassung

In dieser Arbeit, wird ein kultureller Vergleich zwischen Deutschland und China in Textileinzelhandel durch Sentiment Analyse von Online Consumer Reviews gemacht. Durch die Recherche der vorherigen Forschungen wird eine Online Consumer Review in zwei Teile geteilt: den quantitativen Teil und den qualitativen Teil. Der quantitative Teil ist die Anzahl der Sterne, die die Kunden gegeben haben, und der qualitative Teil ist die schriftliche Beschreibung des Kunden. Durch Sentiment Analyse wird der qualitative Teil auch durch eine Polarität in Zahlen ausgedrückt. Diese Polarität beschreibt, wie positiv oder negativ der Kunde das Produkt empfunden hat. Dadurch werden die OCRs in dem quantitativen und qualitativen Aspekt drei statistische Attribute haben: Volumen, Valenz, und Varianz. Es besteht auch ein Korrelationskoeffizient, der den Grad des Zusammenhangs zwischen den quantitativen und qualitativen OCRs beschreibt. Durch diese statistischen Attributen kann man die Gesamtsituation der OCRs besser erkennen.

Basierend auf dem sechsdimensionalen Modell von Hofstede kann man erfassen, dass Chinesen in der Kultur große Machtdistanz und geringe Unsicherheitsvermeidung haben, kollektivistisch sind, während die Deutschen geringe Machtdistanz und große Unsicherheitsvermeidung haben, individuell sind. Mit Hilfe der vorherigen Forschungen in dem kulturellen Bereich, wird vorgeschlagen, dass die chinesischen OCRs kürzer/weniger (kleineres Volumen), negativer (kleinere Valenz), und dichter (kleinere Varianz) in den beiden quantitativen und qualitativen Aspekten und unabhängiger (kleinerer Korrelationskoeffizient) zwischen den quantitativen und qualitativen Teilen als die deutschen OCRs sind.

Mit diesen Hypothesen werden die OCRs von den Untersuchungsobjekten aus den Plattformen des Online-Einzelhandels in den beiden Ländern gesammelt. In dieser Arbeit werden vier unterschiedliche Sportkleidungen als Untersuchungsobjekte ausgewählt. Die chinesischen OCRs werden durch Google Translate in Deutsch übersetzt. Nach der Textverarbeitung werden die gesammelten OCRs durch Sentiment Analyse durchgeführt, damit jede Review eine Polarität über den Grad des Gefühls des Kunden hat.

Nach der Prüfung der Hypothesen, wird das gefunden, was die unterschiedlichen kulturellen Einflüsse auf die OCRs in quantitativen und qualitativen Aspekten ausmacht. Im quantitativen

Aspekt, sind die chinesischen OCRs dichter als die deutschen, wie die Theorien gezeigt haben, aber die chinesischen OCRs sind mehr und positiver als die deutschen OCRs. In dem qualitativen Aspekt, sind die chinesischen OCRs kürzer, negativer, und dichter als die deutschen, genau wie die vorherigen Forschungen gezeigt haben. Die Zusammenhänge in den beiden Ländern sind positiv, und der chinesische Zusammenhang zwischen den quantitativen und qualitativen OCRs ist schwächer als der von den deutschen.

Die Gründe, warum es solche Unterschiede in der Theorie und der Praxis gibt, werden in Kapitel 5 diskutiert. Zusammenfassend werden die quantitativen OCRs stärker als die qualitativen durch die Kultur beeinflusst, besonders in China. Die kollektivistische Kultur der chinesischen Kunden motiviert sie, mehrere aber kürzere OCRs zu schreiben und beeinflusst groß die quantitativen OCRs, damit die chinesischen Kunden mehr “5 Sterne” Bewertungen geben, obwohl sie nicht zufrieden sind. Deshalb sind die chinesischen quantitativen OCRs viel positiver als die deutschen und gleichzeitig ist der Zusammenhang zwischen chinesischen quantitativen und qualitativen OCRs schwächer.

Mit dieser Kenntnisse sollten die Kunden beachten, dass die chinesischen Sterne (also die quantitativen OCRs) kein glaubwürdiges Kennzeichen in der Kaufentscheidung sind, während die deutschen Sterne ein Verweis sind. Die Kunden in den beiden Ländern sollten auch mehr die schriftlichen Teile der OCRs durchlesen, um die Glaubwürdigkeit zu erhöhen und mehrere Informationen über die Attribute des Produkts herauszufinden, damit sie die richtige Kaufentscheidung treffen können.

Die Hersteller eines Produkts sollten auch die schriftlichen Teile analysieren, um die Stärken und Schwächen durch die Rezensionen der Kunden kennenzulernen, und die auf die Probleme schnell zu reagieren. Durch diese Analyse könnten sie die durch den Kunden fokussierten Attribute erkennen und verbessern, und damit bessere Produkte oder Serviceleistungen anbieten.

Die Betreiber der Plattformen sollten diese Ergebnisse beachten. Sie sollten die OCRs möglicherweise gut in einer bestimmten Reihenfolge (aber nicht zeitlich) ordnen, damit die nutzlosen, beispielsweise die zu kurzen OCRs, nicht direkt an erster Stelle positioniert werden sondern die hilfreichen Kundenrezensionen, damit die anderen Kunden schnell die nützlichen Informationen finden könnten. Noch besser wäre es, wenn die Plattformen die vom Kunden fokussierten Aspekte oder die Attitüden aus den OCRs automatisch zusammenfassten.

6.2 Limitationen

Die in dieser Forschung ausgegebenen Ergebnisse sollten in Verbindung mit verschiedenen Limitationen berücksichtigt werden. Diese Limitationen betrachten aus einer Vielzahl von

Blickwinkeln:

1. Wirtschaftliche Limitationen:

- a) Diese Studie ist aus dem Bereich des Textileinzelhandels. Das bedeutet, dass sich die Ergebnisse dieser Arbeit nur auf den Textileinzelhandel beziehen. Es ist noch nachzuweisen, ob die Ergebnisse domainspezifisch sind oder nicht. Diese Untersuchungsobjekte sind die Sportkleidungen, die nicht alle Güter im Textileinzelhandel repräsentieren können.
- b) Die in dieser Arbeit genutzten Maßnahmen sind statistisch. Diese Maßnahmen müssen zahlreiche OCRs von gleichen Produkten in den beiden Ländern analysieren. Nur nach Kauf des Produkts können die Kunden OCRs geben, deshalb sind die Untersuchungsobjekte beschränkt, da sie in China und Deutschland sehr gut verkauft werden müssen. Deshalb sind die OCRs meistens positiv.

2. Limitationen von der Sentiment Analyse:

- a) Die in dieser Arbeit verwendete Sentiment Analyse ist nur im Dokument-Level. Es gibt auch den Aspekt-Level in Sentiment Analysen, in welchem die Analyse über die konkreten Aspekte der Kunden, zum Beispiel: wie gut ist die Kleidung in Ästhetik, Performance oder äußeren Attributen in den OCRs berücksichtigt werden. Diese Attitüde von Kunden über die Attribute des Produkts werden in dieser Arbeit nicht studiert.
- b) Das Lexikon sollte verbessert werden. Der Algorithmus betrachtet keine Phrase, keinen Komparativ und keinen Superlativ. Die Phrase “nicht gut” ist natürlich kein positiver Ausdruck, und die Wörter “besser” und “am besten” sollten auch nicht gleiche Polarität wie “gut” haben. Aber in diesem Algorithmus werden diese Situationen ignoriert.

3. Limitationen von Übersetzung:

- a) Der in dieser Arbeit gemachte Vergleich ist kulturübergreifend und mehrsprachig, deshalb ist es nötig eine Sprache in die andere zu übersetzen. Wegen der großen Menge der Rohdaten, wird hier eine Übersetzungsmaschine verwendet. Obwohl die Einflüsse klein sind, wie die Diskussion im Abschnitt 3.6 zeigt, gibt es sie.
- b) Weil China und Deutschland unterschiedliche Kulturen haben, haben die Wörter unterschiedliche Bedeutungen. Zum Beispiel: das Wort “Stolz” ist in Deutschland positiv aber in China negativ, obwohl es in beiden Ländern das gleiche Wort ist.

4. Limitationen von Statistik:

- a) In dieser Arbeit, werden einige Ergebnisse im Allgemeinen herausgefunden, aber es gibt auch eine Ausnahme in den deutschen Daten aus Puma. Der Grund für

diese Außerordentlichkeit ist noch nicht gefunden.

6.3 Zukünftige Forschung

In der Zukunft werden weiteren Forschungen mehrere Untersuchungsobjekte im Textileinzelhandel durchgeführt, um die Ergebnisse in dieser Arbeit zu beweisen. Wenn man mehrere Untersuchungsobjekte analysieren würde, könnte man auch erkennen, ob die obig genannte Ausnahme nur ein seltener Fall ist. Man sollte zwischen mehreren Ländern den kulturellen Vergleich durchführen, damit die kulturellen Einflüsse von den Dimensionen auf die Online Consumer Reviews besser erkannt werden könnten.

Für die weiteren Forschungen, ist es auch wichtig, die Limitationen von der Sentiment Analyse zu lösen. Man kann die Analyse im Aspekt-Level durchführen und herausfinden, welche Unterschiede es in den konkreten Aspekten der Kunden zwischen den Kulturen gibt, und welche Zusammenhänge es zwischen den Unterschieden und kulturellen Dimensionen von Hofstede gibt. Das Lexikon und der Algorithmus sollten auch in Zukunft verbessert werden, um bessere Ergebnisse zu bekommen.

Die kulturellen Gründe werden vorgeschlagen, aber noch nicht durch Studien bewiesen. Mit den Ergebnissen dieser Arbeit kann man weitere Forschungen über die Online Consumer Reviews betreiben, insbesondere eine Ursachenforschung durchführen, damit man die Zusammenhänge zwischen den kulturellen Dimensionen und den Ergebnissen bestimmen kann.

Die Motive und Wirkungen dieser Unterschiede sind auch in dieser Arbeit in der Konsumentensicht, Unternehmensicht und der Sicht der Plattformbetreiber diskutiert geworden. Aber es ist auch wichtig die obig dargestellten Motive und Wirkungen durch weitere Forschungen zu bestimmen.

In der Zusammenfassung, macht diese Arbeit einen kulturellen Vergleich von den Online Consumer Reviews im Textileinzelhandel zwischen Deutschland und China durch Sentiment Analyse, und einige Unterschiede zwischen den beiden Ländern werden in dieser Arbeit herausgefunden. Die deutschen Online Consumer Reviews sind qualitativer, die chinesischen quantitativer. In dem quantitativen Aspekt sind die chinesischen Online Consumer Reviews positiver aber in dem qualitativen Aspekt ist die Situation anders. Die deutschen Online Consumer Reviews sind dezentralisiert und die chinesischen Online Consumer Reviews sind dicht zentralisiert. Die deutschen Kunden geben die quantitativen Online Consumer Reviews abhängiger von dem schriftlichen Teil, als die chinesischen.

Es ist erwünscht, dass diese Ergebnisse und Diskussionen für die Konsumenten im Textileinzelhandel nützlich sind. Darüber hinaus ermöglichen es den Produzenten der Produkten

durch diese Analyse, effektiver auf ihre Angebote und Marketingaktivitäten zu zielen, um die Umsätze zu erhöhen, und eine effektive und kosteneffiziente Möglichkeit für die Betreiber der Plattformen, Online-Kunden in Deutschland und China zu erwerben und halten.

Eidesstattliche Erklärung

Der Verfasser erklärt an Eides statt, dass er die vorliegende Arbeit selbständig, ohne fremde Hilfe und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt hat. Die aus fremden Quellen (einschließlich elektronischer Quellen) direkt oder indirekt übernommenen Gedanken sind ausnahmslos als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden.

Ort, Datum

Unterschrift des Verfassers

Literaturverzeichnis

- Aaker, Jennifer L, & Maheswaran, Durairaj. 1997. The effect of cultural orientation on persuasion. *Journal of consumer research*, **24**(3), 315–328.
- Abraham-Murali, Liza, & Littrell, Mary Ann. 1995a. Consumers' conceptualization of apparel attributes. *Clothing and textiles research journal*, **13**(2), 65–74.
- Abraham-Murali, Liza, & Littrell, Mary Ann. 1995b. Consumers' perceptions of apparel quality over time: An exploratory study. *Clothing and textiles research journal*, **13**(3), 149–158.
- Adams, Carl. 2005. Will B2C E-commerce Developed in One Cultural Environment be Suitable for Another Culture : A Cross-Cultural Study between amazon . co . uk (UK) and. *Acm*, 236–243.
- Armstrong, Gary, Adam, Stewart, Denize, Sara, & Kotler, Philip. 2014. *Principles of marketing*. Pearson Australia.
- Arora, R., & Srinivasa, S. 2014 (Jan). A faceted characterization of the opinion mining landscape. *Pages 1–6 of: Communication systems and networks (comsnets), 2014 sixth international conference on*.
- Aue, Anthony, & Gamon, Michael. 2005. Customizing sentiment classifiers to new domains: A case study. *Pages 2–1 of: Proceedings of recent advances in natural language processing (ranlp)*, vol. 1.
- Ba, Sulin, & Pavlou, Paul A. 2002. Evidence of the effect of trust building technology in electronic markets: Price premiums and buyer behavior. *Mis quarterly*, 243–268.
- Bajari, Patrick, & Hortacsu, Ali. 2003. The winner's curse, reserve prices, and endogenous entry: Empirical insights from ebay auctions. *Rand journal of economics*, 329–355.
- Balahur, Alexandra, & Turchi, Marco. 2014. Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis. *Computer speech &*

language, **28**(1), 56–75.

- Balazs, Jorge A., & Velásquez, Juan D. 2016. Opinion Mining and Information Fusion: A survey. *Information fusion*, **27**(jan), 95–110.
- Bateman, Scott, Gutwin, Carl, & Nacenta, Miguel. 2008. Seeing things in the clouds: the effect of visual features on tag cloud selections. *Pages 193–202 of: Proceedings of the nineteenth acm conference on hypertext and hypermedia*. ACM.
- Bezos, Jeffrey P. 2013. *Annual report pursuant to section 13 or 15(d) of the securities exchange act of 1934 of amazon, inc.* Tech. rept. United States Securities and Exchange Commission.
- Bezos, Jeffrey P. 2014. *Annual report pursuant to section 13 or 15(d) of the securities exchange act of 1934 of amazon, inc.* Tech. rept. United States Securities and Exchange Commission.
- Blitzer, John, Dredze, Mark, Pereira, Fernando, *et al.* . 2007. Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. *Pages 440–447 of: Acl*, vol. 7.
- Brennan, Robert L., & Prediger, Dale J. 1981. Coefficient kappa: Some uses, misuses, and alternatives. *Educational and psychological measurement*, **41**(3), 687–699.
- Bronner, Fred, & de Hoog, Robert. 2010. Vacationers and ewom: Who posts, and why, where, and what? *Journal of travel research*.
- Cambria, E., Schuller, B., Xia, Yunqing, & Havasi, C. 2013. New avenues in opinion mining and sentiment analysis. *Intelligent systems, iee*, **28**(2), 15–21.
- Casaló, Luis V., Flavián, Carlos, & Guinalíu, Miguel. 2010. Determinants of the intention to participate in firm-hosted online travel communities and effects on consumer behavioral intentions. *Tourism management*, **31**(6), 898 – 911.
- CECRC. 2014. *2013 china online retail market data analysis report*. http://www.100ec.cn/zt/upload_data/down/20140304.pdf. (in Chinese).
- Chatterjee, Patrali. 2001. Online reviews: do consumers use them? *Advances in consumer research*, **28**, 129–133.

- Chen, Yubo, Wang, Qi, & Xie, Jinhong. 2011. Online social interactions: A natural experiment on word of mouth versus observational learning. *Journal of marketing research*, **48**(2), 238–254.
- Cheung, Christy M.K., & Thadani, Dimple R. 2012. The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision support systems*, **54**(1), 461–470.
- Chevalier, Judith A., & Mayzlin, Dina. 2006. The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, **43**(3), 345–354.
- Chintagunta, Pradeep K., Gopinath, Shyam, & Venkataraman, Sriram. 2010. The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing science*, **29**(5), 944–957.
- Chu, Shu-Chuan, & Choi, Sejung Marina. 2011. Electronic Word-of-Mouth in Social Networking Sites: A Cross-Cultural Study of the United States and China. *Journal of global marketing*, **24**(3), 263–281.
- Chun-Tung Lowe, Anthony, & Corkindale, David R. 1998. Differences in “cultural values” and their effects on responses to marketing stimuli: A cross-cultural study between australians and chinese from the people’s republic of china. *European journal of marketing*, **32**(9/10), 843–867.
- Church, Kenneth Ward, & Hanks, Patrick. 1990. Word association norms, mutual information, and lexicography. *Computational linguistics*, **16**(1), 22–29.
- Clemons, Eric K., Gao, Guodong Gordon, & Hitt, Lorin M. 2006. When online reviews meet hyperdifferentiation: A study of the craft beer industry. *Journal of management information systems*, **23**(2), 149–171.
- Cohen, Jacob, *et al.* . 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, **20**(1), 37–46.
- Crotts, John C, Mason, Peyton R, & Davis, Boyd. 2009. Measuring guest satisfaction and competitive position in the hospitality and tourism industry an application of stance-shift analysis to travel blog narratives. *Journal of travel research*, **48**(2), 139–151.
- Cui, Geng, Lui, Hon-Kwong, & Guo, Xiaoning. 2012. The effect of online consumer reviews

- on new product sales. *International journal of electronic commerce*, **17**(1), 39–58.
- Cyr, Dianne. 2014 (dec). *Modeling Web Site Design Across Cultures: Relationships to Trust, Satisfaction, and E-Loyalty*.
- Dale, Robert, Moisl, Hermann, & Somers, Harold. 2000. *Handbook of natural language processing*. CRC Press.
- Dawar, Niraj, Parker, Philip M, & Price, Lydia J. 1996. A cross-cultural study of interpersonal information exchange. *Journal of international business studies*, 497–516.
- Dellarocas, Chrysanthos. 2003. The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms. *Management science*, **49**(10), 1407–1424.
- Dellarocas, Chrysanthos, Zhang, Xiaoquan (Michael), & Awad, Neveen F. 2007. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of interactive marketing*, **21**(4), 23 – 45.
- Demirtas, Erkin. 2013. Cross-lingual sentiment analysis with machine translation.
- Dey, Lipika, & Haque, Sk.Mirajul. 2009. Opinion mining from noisy text data. *International journal on document analysis and recognition (ijdar)*, **12**(3), 205–226.
- Dhar, Vasant, & Chang, Elaine A. 2009. Does chatter matter? the impact of user-generated content on music sales. *Journal of interactive marketing*, **23**(4), 300 – 307.
- Dichter, Ernest. 1966. {How word-of-mouth advertising works}. *Harvard business review*, **44**(6), 147–160.
- Dick, Alan S, & Basu, Kunal. 1994. Customer loyalty: toward an integrated conceptual framework. *Journal of the academy of marketing science*, **22**(2), 99–113.
- Doney, Patricia M, Cannon, Joseph P, & Mullen, Michael R. 1998. Understanding the influence of national culture on the development of trust. *Academy of management review*, **23**(3), 601–620.
- Donthu, Naveen, & Yoo, Boonghee. 1998. Cultural influences on service quality expectations. *Journal of service research*, **1**(2), 178–186.
- Duan, Wenjing, Gu, Bin, & Whinston, Andrew B. 2008. Do online reviews matter? —

an empirical investigation of panel data. *Decision support systems*, **45**(4), 1007 – 1016.
Information Technology and Systems in the Internet-Era.

Duenas-Fernández, Rodrigo, L'Huillier, Gaston, & Velásquez, Juan D. Sentiment polarity of trends on the web using opinion mining and topic modeling.

Duhan, Dale F, Johnson, Scott D, Wilcox, James B, & Harrell, Gilbert D. 1997. Influences on consumer use of word-of-mouth recommendation sources. *Journal of the academy of marketing science*, **25**(4), 283–295.

Dwyer, Sean, Mesak, Hani, & Hsu, Maxwell. 2005. An exploratory examination of the influence of national culture on cross-national product diffusion. *Journal of international marketing*, **13**(2), 1–27.

Eaton, David H, *et al.* . 2002. Valuing information: Evidence from guitar auctions on ebay. *Murray, ky, murray state university*, **28**.

Eckman, Molly, Damhorst, Mary Lynn, & Kadolph, Sara J. 1990. Toward a model of the in-store purchase decision process: Consumer use of criteria for evaluating women's apparel. *Clothing and textiles research journal*, **8**(2), 13–22.

Engel, James F, Blackwell, Roger D, & Miniard, Paul W. 1993. *Consumer behavior: Harcourt broce joranovich collegepublishers*.

Feldman, Ronen. 2013. Techniques and applications for sentiment analysis. *Commun. acm*, **56**(4), 82–89.

Fiore, Ann Marie, & Damhorst, Mary Lynn. 1992. Intrinsic cues as predictors of perceived quality of apparel. *Journal of consumer satisfaction, dissatisfaction and complaining behavior*, **5**, 168.

Fong, John, & Burton, Suzan. 2008. A cross-cultural comparison of electronic word-of-mouth and country-of-origin effects. *Journal of business research*, **61**(3), 233–242.

Forsythe, Sandra, Presley, Ann Beth, & Caton, Karen Wilson. 1996. Dimensions of apparel quality influencing consumers' perceptions. *Perceptual and motor skills*, **83**(1), 299–305.

Forsythe, Sandra M. 1991. Effect of private, designer, and national brand names on shoppers' perception of apparel quality and price. *Clothing and textiles research journal*, **9**(2), 1–6.

- Friedl, Jeffrey EF. 2009. *Reguläre ausdrücke*. O'Reilly Germany.
- Fung, Victor K., & Fung, William K. 2014. *China's apparel market, 2014*. Tech. rept. Fung Business Intelligence Centre.
- Gefen, David, & Heart, Tsipi. 2006. On the need to include national culture as a central issue in e-commerce trust beliefs. *Journal of global information management*, **14**(4), 1.
- Goldsmith, Ronald E., & Goldsmith, Elizabeth B. 2002. Buying apparel over the Internet. *Journal of product & brand management*, **11**(2), 89–102.
- Goldsmith, Ronald E, & Horowitz, David. 2006. Measuring motivations for online opinion seeking. *Journal of interactive advertising*, **6**(2), 2–14.
- Greenfield, Patricia M. 2000. Three approaches to the psychology of culture: Where do they come from? where can they go? *Asian journal of social psychology*, **3**(3), 223–240.
- Greer, Jennifer D. 2003. Evaluating the credibility of online information: A test of source and advertising influence. *Mass communication and society*, **6**(1), 11–28.
- Guernsey, Lisa. 2000. Suddenly, everybody's an expert on everything. *The new york times*, **3**.
- Han, Sang-Pil, & Shavitt, Sharon. 1994. Persuasion and culture: Advertising appeals in individualistic and collectivistic societies. *Journal of experimental social psychology*, **30**(4), 326–350.
- Hana Ben-Shabat, Moriarty, Mike, Nilforoushan, Parvaneh, & Yuen, Christine. 2015. *Global Retail E-Commerce Keeps On Clicking*. Tech. rept. A.T. Kearney Global Consumer Institute.
- Harrison-Walker, L Jean. 2001. The measurement of word-of-mouth communication and an investigation of service quality and customer commitment as potential antecedents. *Journal of service research*, **4**(1), 60–75.
- Hennig-Thurau, Thorsten, Gwinner, Kevin P., Walsh, Gianfranco, & Gremler, Dwayne D. 2004. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of interactive marketing*, **18**(1), 38–52.

- Ho-Dac, Nga N., Carson, Stephen J., & Moore, William L. 2013. The effects of positive and negative online customer reviews: Do brand strength and category maturity matter? *Journal of marketing*, **77**(6), 37–53.
- Hofstede, G, Hofstede, GJ, & Minkov, M. 2010. *Cultures and organizations, software of the mind. intercultural cooperation and its importance for survival*. McGrawHill.
- Hofstede, Geert. 1991. *Interkulturelle zusammenarbeit: kulturen—organisationen—management*. Springer-Verlag.
- Hofstede, Geert. 1998. *Masculinity and femininity: The taboo dimension of national cultures*. Vol. 3. Sage.
- Hofstede, Geert H. 2001. *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage.
- Hofstede, Geert H. cited November 2015. *the hofstede centre*. <http://geert-hofstede.com/china.html>.
- Hong, Jaehyun, & Park, Hee Sun. 2011. User Ratings and Willingness to Express Opinions Online. *International journal of marketing studies*, **3**(2), 2–15.
- Houser, Daniel, & Wooders, John. 2006. Reputation in auctions: Theory, and evidence from ebay. *Journal of economics & management strategy*, **15**(2), 353–369.
- Hsu, Hsiu-Ju, & Burns, Leslie Davis. 2002. Clothing evaluative criteria: A cross-national comparison of taiwanese and united states consumers. *Clothing and textiles research journal*, **20**(4), 246–252.
- Hu, Nan, Liu, Ling, & Zhang, Jie Jennifer. 2008. Do online reviews affect product sales? the role of reviewer characteristics and temporal effects. *Information technology and management*, **9**(3), 201–214.
- Hui, Michael K., & Au, Kevin. 2001. Justice perceptions of complaint-handling: a cross-cultural comparison between {PRC} and canadian customers. *Journal of business research*, **52**(2), 161 – 173. Doing Business in China.
- Jiménez, Fernando R., & Mendoza, Norma A. 2013. Too popular to ignore: The influence of online reviews on purchase intentions of search and experience products. *Journal of*

interactive marketing, **27**(3), 226 – 235.

- Jin, Byoungcho, Yong Park, Jin, & Sang Ryu, Jay. 2010. Comparison of Chinese and Indian consumers' evaluative criteria when selecting denim jeans. *Journal of fashion marketing and management: An international journal*, **14**(1), 180–194.
- Joshi, Mahesh, & Penstein-Rosé, Carolyn. 2009. Generalizing dependency features for opinion mining. *Pages 313–316 of: Proceedings of the acl-ijcnlp 2009 conference short papers*. ACLShort '09. Stroudsburg, PA, USA: Association for Computational Linguistics.
- Kadolph, Sara J. 2007. *Quality assurance for textiles and apparel*. Fairchild Publications.
- Kale, Sudhir H, & Barnes, John W. 1992. Understanding the domain of cross-national buyer-seller interactions. *Journal of international business studies*, 101–132.
- Kalyanam, Kirthi, & McIntyre, Shelby. 2001. Returns to reputation in online auction markets. *Santa clara, ca, santa clara university. available on-line at http://business.scu.edu/faculty/research/working_papers/pdf/kalyanam_mcintyre_wp10.pdf*.
- Kauffman, Robert J, & Wood, Charles A. 2000. Running up the bid: Modeling seller opportunism in internet auctions. *Amcis 2000 proceedings*, 376.
- Keller, Kevin Lane, Parameswaran, MG, & Jacob, Isaac. 2011. *Strategic brand management: Building, measuring, and managing brand equity*. Pearson Education India.
- Khan, Aurangzeb, Baharudin, Baharum, & Khan, Khairullah. 2011. *Sentiment Classification Using Sentence-level Lexical Based Semantic Orientation of Online Reviews*.
- Khan, Farhan Hassan, Bashir, Saba, & Qamar, Usman. 2014. Tom: Twitter opinion mining framework using hybrid classification scheme. *Decision support systems*, **57**, 245 – 257.
- Khwaldeh, Ali, Tahat, Amani, Marti, Jordi, & Tahat, Mofleh. 2013. Atomic Data Mining Numerical Methods, Source Code SQLite with Python. *Procedia - social and behavioral sciences*, **73**(Feb.), 232–239.
- Kim, Taegoo (Terry), Kim, Woo Gon, & Kim, Hong-Bumm. 2009. The effects of perceived justice on recovery satisfaction, trust, word-of-mouth, and revisit intention in upscale hotels. *Tourism management*, **30**(1), 51 – 62.

- Kostyra, Daniel S., Reiner, Jochen, Natter, Martin, & Klapper, Daniel. 2015. Decomposing the effects of online customer reviews on brand, price, and product attributes. *International journal of research in marketing*, Feb.
- Kuppens, Peter, Ceulemans, Eva, Timmerman, Marieke E, Diener, Ed, & Kim-Prieto, Chu. 2006. Universal intracultural and intercultural dimensions of the recalled frequency of emotional experience. *Journal of cross-cultural psychology*, **37**(5), 491–515.
- Lam, Desmond, Lee, Alvin, & Mizerski, Richard. 2009. The Effects of Cultural Values in Word-of-Mouth Communication. *Journal of international marketing*, **17**(3), 55–70.
- Langenscheidt, Florian, & Bauer, Ulrike. 2010. *Lexikon der deutschen weltmarktführer:[die königsklasse deutscher unternehmen in wort und bild]*. Deutsche Standards Ed.
- Lee, Eun-Jung, & Park, Jung Kun. 2009. Online service personalization for apparel shopping. *Journal of retailing and consumer services*, **16**(2), 83–91.
- Lee, Zoonky, Im, Il, & Lee, Sang Jun. 2000. The effect of negative buyer feedback on prices in internet auction markets. *Pages 286–287 of: Proceedings of the twenty first international conference on information systems*. Association for Information Systems.
- Lennon, Sharron J, & Fairhurst, Ann E. 1994. Categorization of the quality concept. *Home economics research journal*, **22**(3), 267–285.
- Levene, Howard. 1960. Robust tests for equality of variances¹. *Contributions to probability and statistics: Essays in honor of harold hotelling*, **2**, 278–292.
- Litvin, Stephen W., Goldsmith, Ronald E., & Pan, Bing. 2008. Electronic word-of-mouth in hospitality and tourism management. *Tourism management*, **29**(3), 458–468.
- Liu, Ben Shaw-Ching, Furrer, Olivier, & Sudharshan, Devanathan. 2001. The relationships between culture and behavioral intentions toward services. *Journal of service research*, **4**(2), 118–129.
- Liu, Bing. 2010. Sentiment analysis and subjectivity. *Handbook of natural language processing*, **2**, 627–666.
- Liu, Bing. 2012. Sentiment Analysis and Opinion Mining. *Synthesis lectures on human language technologies*, **5**(1), 1–167.

- Livingston, Jeffrey A. 2005. How valuable is a good reputation? a sample selection model of internet auctions. *Review of economics and statistics*, **87**(3), 453–465.
- Lu, Yue. 2015. *Understanding Consumers' Perceptions of Apparel Quality in the US and China through Online Reviews*. Ph.D. thesis, North Carolina State University.
- Lucking-Reiley, David, Bryan, Doug, Prasad, Naghi, & Reeves, Daniel. 2007. Pennies from ebay: The determinants of price in online auctions*. *The journal of industrial economics*, **55**(2), 223–233.
- Luo, Chuan, Wu, Jing, Shi, Yani, & Xu, Yun. 2014. The effects of individualism–collectivism cultural orientation on eWOM information. *International journal of information management*, **34**(4), 446–456.
- Ma, Jack. 2014. *Annual report pursuant to section 13 or 15(d) of the securities exchange act of 1934 of alibaba group holding limited*. Tech. rept. United States Securities and Exchange Commission.
- Mahdi, Ali, Hussain, A, Abbas, Mohammed, Mazar, Taher Ilyas, & George, Shaju. 2015. A comparative analysis of strategies and business models of nike, inc. and adidas group with special reference to competitive advantage in the context of a dynamic and competitive environment. *International journal of business management & economic research*, **6**(3).
- Mangold, W. Glynn, Miller, Fred, & Brockway, Gary R. 1999. Word-of-mouth communication in the service marketplace. *Journal of services marketing*, **13**(1), 73–89.
- Martens, Jul. 2003. *Statistische datenanalyse mit spss für windows*. Oldenbourg Verlag.
- Mazaheri, Ebrahim, Richard, Marie-Odile, & Laroche, Michel. 2011. Online consumer behavior: Comparing canadian and chinese website visitors. *Journal of business research*, **64**(9), 958 – 965. Globalization, Culture, and Marketing Strategy.
- McCort, Daniel John, & Malhotra, Naresh K. 1993. Culture and consumer behavior: toward an understanding of cross-cultural consumer behavior in international marketing. *Journal of international consumer marketing*, **6**(2), 91–127.
- McDonald, Cynthia G, & Slawson, V Carlos. 2002. Reputation in an internet auction market. *Economic inquiry*, **40**(4), 633–650.

- Melnik, Mikhail I, & Alm, James. 2002. Does a seller's ecommerce reputation matter? evidence from ebay auctions. *Journal of industrial economics*, 337–349.
- Meyer-Bahlburg, Heino FL. 1969. Spearmans rho als punktbiserialer rangkorrelationskoeffizient. *Biometrische zeitschrift*, **11**(1), 60–66.
- Miller, Joan G, Bersoff, David M, & Harwood, Robin L. 1990. Perceptions of social responsibilities in india and in the united states: moral imperatives or personal decisions? *Journal of personality and social psychology*, **58**(1), 33.
- Mitchell, Tom M. 1997. *Machine learning*. wcb.
- Moe, Wendy W., & Trusov, Michael. 2011. The value of social dynamics in online product ratings forums. *Journal of marketing research*, **48**(3), 444–456.
- Momtazi, S. 2012. Fine-grained German Sentiment Analysis on Social Media. *Lrec*.
- Money, R Bruce, Gilly, Mary C, & Graham, John L. 1998. Explorations of national culture and word-of-mouth referral behavior in the purchase of industrial services in the united states and japan. *The journal of marketing*, 76–87.
- Mui-Fong, Goh, & Chee Wee, Gan. 2014. *China's E-Commerce Market in 2014: The Logistics Challenges*. Tech. rept. A.T. Kearney, Shanghai.
- Nelson, Phillip. 1970. Information and consumer behavior. *The journal of political economy*, 311–329.
- Neviarouskaya, Alena, Prendinger, Helmut, & Ishizuka, Mitsuru. 2010. Recognition of fine-grained emotions from text: An approach based on the compositionality principle. *Pages 179–207 of: Nishida, Toyoaki, Jain, LakhmiC., & Faucher, Colette (eds), Modeling machine emotions for realizing intelligence*. Smart Innovation, Systems and Technologies, vol. 1. Springer Berlin Heidelberg.
- Nusair, Khaldoon “Khal”, Parsa, H.G., & Cobanoglu, Cihan. 2011. Building a model of commitment for generation y: An empirical study on e-travel retailers. *Tourism management*, **32**(4), 833 – 843.
- Olson, Jerry C, Jacoby, Jacob, & Venkatesan, M. 1972. Proceedings of the third annual conference of the association for consumer research. Association for Consumer Research

Iowa City.

- Pang, Bo, Lee, Lillian, & Vaithyanathan, Shivakumar. 2002. Thumbs up?: Sentiment classification using machine learning techniques. *Pages 79–86 of: Proceedings of the acl-02 conference on empirical methods in natural language processing - volume 10*. EMNLP '02. Stroudsburg, PA, USA: Association for Computational Linguistics.
- Pearson, Karl. 1895. Note on regression and inheritance in the case of two parents. *Proceedings of the royal society of london*, 240–242.
- Peters, Rolf-Herbert. 2007. *Die puma-story*. Hanser.
- Poon, Patrick S., Hui, Michael K., & Au, Kevin. 2004. Attributions on dissatisfying service encounters: A cross cultural comparison between canadian and prc consumers. *European journal of marketing*, **38**(11/12), 1527–1540.
- Rahman, Osmud, Jiang, Yan, & Liu, Wing-sun. 2010. Evaluative criteria of denim jeans: A cross-national study of functional and aesthetic aspects. *The design journal*, **13**(3), 291–311.
- Ravi, Kumar, & Ravi, Vadlamani. 2015. A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-based systems*, jun.
- Rebolledo, VíctorL., L'Huillier, Gastón, & Velásquez, JuanD. 2010. Web pattern extraction and storage. *Pages 49–77 of: Velásquez, JuanD., & Jain, LakhmiC. (eds), Advanced techniques in web intelligence - i*. Studies in Computational Intelligence, vol. 311. Springer Berlin Heidelberg.
- Remus, R, Quasthoff, U, & Heyer, G. 2010. SentiWS-A Publicly Available German-language Resource for Sentiment Analysis. *Lrec*.
- Resnick, Paul, & Zeckhauser, Richard. 2002. Trust among strangers in internet transactions: Empirical analysis of ebay's reputation system. *The economics of the internet and e-commerce*, **11**(2), 23–25.
- Resnick, Paul, Kuwabara, Ko, Zeckhauser, Richard, & Friedman, Eric. 2000. Reputation systems. *Communications of the acm*, **43**(12), 45–48.
- Resnick, Paul, Zeckhauser, Richard, Swanson, John, & Lockwood, Kate. 2006. The value of

reputation on ebay: A controlled experiment. *Experimental economics*, **9**(2), 79–101.

Roland Lindner, Carsten Knop. 2013. *Online-handel: Amazon lüftet deutsches umsatzgeheimnis*. <http://www.faz.net/aktuell/wirtschaft/netzwirtschaft/online-handel-amazon-lueftet-deutsches-umsatzgeheimnis-12052201.html>. 05.02.2013.

Rosenau, Jeremy A, & Wilson, David L. 2014. *Apparel merchandising: The line starts here*. A&C Black.

Royston, Patrick. 1995. Remark as r94: A remark on algorithm as 181: The w-test for normality. *Journal of the royal statistical society. series c (applied statistics)*, **44**(4), 547–551.

Sánchez-García, Isabel, & Currás-Pérez, Rafael. 2011. Effects of dissatisfaction in tourist services: The role of anger and regret. *Tourism management*, **32**(6), 1397 – 1406.

Saunders, Rebecca. 2001. *Amazon. com: der schnelle weg gross zu werden*. FinanzBuch-Verlag.

Schimmack, Ulrich, Oishi, Shigehiro, & Diener, Ed. 2002. Cultural influences on the relation between pleasant emotions and unpleasant emotions: Asian dialectic philosophies or individualism-collectivism? *Cognition & emotion*, **16**(6), 705–719.

Schindler, Robert M, & Bickart, Barbara. 2005. Published word of mouth: Referable, consumer-generated information on the internet. *Online consumer psychology: Understanding and influencing consumer behavior in the virtual world*, 35–61.

Scholz, T, Conrad, S, & Hillekamps, L. 2012. Opinion mining on a german corpus of a media response analysis. *In: Text, speech and dialogue*.

Schumann, Jan H, v. Wangenheim, Florian, Stringfellow, Anne, Yang, Zhilin, Blazevic, Vera, Praxmarer, Sandra, Shainesh, G, Komor, Marcin, Shannon, Randall M, & Jiménez, Fernando R. 2010. Cross-Cultural Differences in the Effect of Received Word-of-Mouth Referral in Relational Service Exchange. *Journal of international marketing*, **18**(3), 62–80.

Sebestyen, Thomas. 2010. *Xml: Einstieg für anspruchsvolle*. Pearson Deutschland GmbH.

Segaran, Toby. 2008. *Kollektive intelligenz analysieren, programmieren und nutzen*. O'Reilly

Germany.

- Sen, Shahana, & Lerman, Dawn. 2007. Why are you telling me this? an examination into negative consumer reviews on the web. *Journal of interactive marketing*, **21**(4), 76–94.
- Serra Cantallops, Antoni, & Salvi, Fabiana. 2014. New consumer behavior: A review of research on eWOM and hotels. *International journal of hospitality management*, **36**(Jan.), 41–51.
- Shapiro, S. S., & Wilk, M. B. 1965. An analysis of variance test for normality (complete samples). *Biometrika*, **52**(3-4), 591–611.
- Sharma, Piyush, & Sivakumaran, Bharadhwaj. 2005. Self-indulgence or loss of self-control? or, is it a bit of both? investigating cross-cultural aspects of impulse buying behavior. *Adv consum res*, **6**(1), 151.
- Shweder, Richard A. 1990. In defense of moral realism: Reply to gabennesch. *Child development*, **61**(6), 2060–2067.
- Sia, Choon Ling, Lim, Kai H, Leung, Kwok, Lee, Matthew KO, Huang, Wayne Wei, & Benbasat, Izak. 2009. Web strategies to promote internet shopping: is cultural-customization needed? *Mis quarterly*, 491–512.
- Singelis, Theodore M. 1994. The measurement of independent and interdependent self-construals. *Personality and social psychology bulletin*, **20**(5), 580–591.
- Singh, Nitish, Fassott, Georg, Zhao, Hongxin, & Boughton, Paul D. 2006. A cross-cultural analysis of german, chinese and indian consumers' perception of web site adaptation. *Journal of consumer behaviour*, **5**(1), 56–68.
- Smith, Peter B, Trompenaars, Fons, & Dugan, Shaun. 1995. The rotter locus of control scale in 43 countries: A test of cultural relativity. *International journal of psychology*, **30**(3), 377–400.
- Sohaib, Osama, & Kang, Kyeong. 2014. Cultural aspects of business-to-consumer (B2C) E-commerce: A comparative analysis of Pakistan and Australia. *Electronic journal of information systems in developing countries*, **61**(1), 1–18.
- Späth, Nadine. 2015. *The e-commerce market in germany*. Tech. rept. Germany Trade &

Invest, Berlin.

Sridhar, Shrihari, & Srinivasan, Raji. 2012. Social influence effects in online product ratings. *Journal of marketing*, **76**(5), 70–88.

Stenberg, Martin, & Sydow, Leo. 2014. *Purchasing Apparel Online in China and in Sweden*. Ph.D. thesis.

Sun, Lucia Bongran, & Qu, Hailin. 2011. Is there any gender effect on the relationship between service quality and word-of-mouth? *Journal of travel & tourism marketing*, **28**(2), 210–224.

Sun, Monic. 2012. How does the variance of product ratings matter? *Management science*, **58**(4), 696–707.

Sundaram, Dinesh S, Mitra, Kaushik, & Webster, Cynthia. 1998. Word-of-mouth communications: A motivational analysis. *Advances in consumer research*, **25**(1), 527–531.

Swanson, Scott R., & Hsu, Maxwell K. 2009. Critical incidents in tourism: Failure, recovery, customer switching, and word of mouth behaviors. *Journal of travel & tourism marketing*, **26**(2), 180–194.

Swinker, Mary E, & Hines, Jean D. 2006. Understanding consumers' perception of clothing quality: A multidimensional approach. *International journal of consumer studies*, **30**(2), 218–223.

Tapia, Pablo A, & Velásquez, Juan D. 2014. Twitter sentiment polarity analysis: A novel approach for improving the automated labeling in a text corpora. *Pages 274–285 of: Active media technology*. Springer.

Tong, Xiao. 2010. A cross-national investigation of an extended technology acceptance model in the online shopping context. *International journal of retail & distribution management*, **38**(10), 742–759.

Triandis, Harry C. 1972. The analysis of subjective culture.

Triandis, Harry C. 1994. Theoretical and methodological approaches to the study of collectivism and individualism. *Cross cultural research and methodology series-sage-*, **18**, 41–41.

- Triandis, Harry C. 2001. Individualism-collectivism and personality. *Journal of personality*, **69**(6), 907–924.
- Turney, Peter D. 2002. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. *Pages 417–424 of: Proceedings of the 40th annual meeting on association for computational linguistics*. Association for Computational Linguistics.
- Turney, Peter D, & Littman, Michael L. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *Acm transactions on information systems (tois)*, **21**(4), 315–346.
- Van Everdingen, Yvonne M, & Waarts, Eric. 2003. The effect of national culture on the adoption of innovations. *Marketing letters*, **14**(3), 217–232.
- Vilares, David, Alonso, Miguel A., & Gómez-Rodríguez, Carlos. 2015. A syntactic approach for opinion mining on spanish reviews. *Natural language engineering*, **21**(1), 139–163.
- Vilpponen, Antti, Winter, Susanna, & Sundqvist, Sanna. 2006. Electronic Word-of-Mouth in Online Environments : Exploring Referral Network Structure and Adoption Behavior. *Journal of interactive advertising*, **6**(2), 63–77.
- Vinodhini, G, & Chandrasekaran, RM. 2012. Sentiment analysis and opinion mining: a survey. *International journal*, **2**(6).
- Wan, Xiaojun. 2008. Using bilingual knowledge and ensemble techniques for unsupervised Chinese sentiment analysis. oct, 553–561.
- Westbrook, Robert A. 1987. Product/consumption-based affective responses and postpurchase processes. *Journal of marketing research*, 258–270.
- Wilson, Theresa, Wiebe, Janyce, & Hoffmann, Paul. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. *Pages 347–354 of: Proceedings of the conference on human language technology and empirical methods in natural language processing*. HLT '05. Stroudsburg, PA, USA: Association for Computational Linguistics.
- Yeniyurt, Sengun, & Townsend, Janell D. 2003. Does culture explain acceptance of new products in a country? an empirical investigation. *International marketing review*, **20**(4), 377–396.

- Yoon, Cheolho. 2009. The effects of national culture values on consumer acceptance of e-commerce: Online shoppers in China. *Information & management*, **46**(5), 294–301.
- Yu, Jianxing, Zha, Zheng-Jun, Wang, Meng, & Chua, Tat-Seng. 2011. Aspect Ranking : Identifying Important Product Aspects from Online Consumer Reviews. *Computational linguistics*, 1496–1505.
- Zhang, Zhiming, Li, Yi, Gong, Chen, & Wu, Haidong. 2002. Casual wear product attributes: a chinese consumers' perspective. *Journal of fashion marketing and management: An international journal*, **6**(1), 53–62.
- Zhou, Lianxi, & Hui, Michael K. 2003. Symbolic value of foreign products in the people's republic of china. *Journal of international marketing*, **11**(2), 36–58.
- Zhou, Z., Jin, X., Fang, Y., & Vogel, D. 2015. Cultural values matter in svw continuance. *Eur j inf syst*, **24**(3), 247–261.
- Zhu, Feng, & Zhang, Xiaoquan (Michael). 2010. Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *Journal of marketing*, **74**(2), 133–148.
- Zipf, George Kingsley. 1949. Human behavior and the principle of least effort.